THE EVOLUTION OF SELECTIVE ADVANTAGE IN A DELETERIOUS MUTATION

P. O'DONALD

Department of Zoology, University College of North Wales, Bangor

Received January 6, 1967

FISHER (1928) put forward a theory of dominance that is still the subject of controversy. If a deleterious mutation is continually occurring, it will appear mainly as a heterozygote. FISHER suggested that the fittest of these heterozygotes would be selected and ultimately become as fit as the wild type. The wild type would then be dominant to the deleterious mutation. The fitness of the heterozygote must therefore be subject to genetic variations which are selected for their effects as modifiers. This idea has been criticised because the selection of single genes modifying the expression of the mutant heterozygote is very slow (WRIGHT 1929a, b; CROSBY 1963; EWENS 1965, 1966). But as SHEPPARD and FORD (1966) point out, a very great deal of time may have been available for this selection. They suggest that several recessive mutations may have a very long history—for example that albinism may have a longer history than man and have been occurring for millions of generations. Even so we should really be thinking not of the selection of single modifiers, but of selection acting on a genetic variance in the fitness of the mutant heterozygote. This problem has been formulated by O'DONALD (1967) in a very simple way. Consider MENDEL's tall and short peas. Both phenotypes have very different means; but no doubt there is variation within each phenotype around its mean and if the population consisted of only one of the phenotypes we should then observe a simple quantitative character and would want to know its variance and heritability. Therefore, when we observe a polymorphism, we ought surely to consider the variance and heritability of each of the phenotypes. Thus we should consider the problem of the natural selection of the selective coefficient of a particular genotype whose fitness has a given variance.

Following O'DONALD (1967), let the fitness of an individual be $1 - s$, where fitness is to be regarded as the probability of survival. Let s have the relative frequency density $f(s)$ with mean $\bar{s}$ and variance V . An individual will therefore have the selective coefficient s with the probability $f(s)ds$. After selection the probability will be $(1 - s)f(s)ds$ and the total frequency will be $1 - \bar{s}$. Thus in the distribution of the selected individuals, the probability of an individual with the selective coefficient s is $(1 - s)f(s)ds/(1 - \bar{s})$. The mean of the selected distribution is therefore

$$\begin{aligned}
 \bar{s}' &= \frac{1}{(1-\bar{s})} \int_{-\infty}^{\infty} s(1-s)f(s)ds \\
 &= \frac{1}{(1-\bar{s})} \int_{-\infty}^{\infty} f(s) [\bar{s}^2 + s(1-2\bar{s}) - (s-\bar{s})^2] ds \\
 &= \frac{1}{(1-\bar{s})} [\bar{s}^2 + \bar{s}(1-2\bar{s}) - V] \\
 &= \bar{s} - \frac{V}{(1-\bar{s})}.
 \end{aligned}$$

The variance of the selected distribution is

$$V' = \frac{1}{(1-\bar{s})} \int_{-\infty}^{\infty} s^2(1-s)f(s)ds - \bar{s}'^2,$$

which gives

$$V' = V - \frac{V^2}{(1-\bar{s})^2} - \frac{\mu_3}{(1-\bar{s})},$$

where μ_3 is the third moment about the mean. If the distribution is symmetrical about the mean—if for example s is normally distributed—then $\mu_3 = 0$, and we have simply

$$V' = V - \frac{V^2}{(1-\bar{s})^2}.$$

Thus $\bar{s}'$ and V' are the mean and variance of the selective coefficients of the individuals that have survived. But in their offspring, only those components of the variance that are determined by additive genetic effects will be changed. We may consider s to have a phenotypic variance and a heritability that gives the proportion of the phenotypic variance determined by additive genetic causes. It is simpler to suppose that V is entirely additive genetic and to ignore the other components of the variance. This will be justified if there is no genotype-environment interaction, which is also a necessary assumption in the progress from one generation to the next. Suppose it is the selective coefficient of the heterozygote that is selected. If the frequency of the heterozygote is $2pq$, then the change in the selective coefficient will be

$$\Delta\bar{s} = -2pqV/(1-\bar{s}),$$

where V is the additive genetic variance. Hence the new mean will be

$$\bar{s}' = \bar{s} - 2pqV/(1-\bar{s}).$$

Similarly the new variance will be

$$V' = V - 2pqV^2/(1-\bar{s})^2.$$

These recurrence relations can be solved numerically with any model describing the changes in gene frequencies. Suppose two alleles at a locus give genotypes with frequencies and fitnesses as follows:

	AA	Aa	aa
Frequencies	p^2	$2pq$	q^2
Fitnesses	$(1-s_A)$	$(1-s)$	$1-s_a$

where s is the variable selective coefficient. Then

$$p' = (p - s_A p^2 - \bar{s} pq) / (1 - s_A p^2 - 2\bar{s} pq - s_a q^2).$$

We have three recurrence relations, therefore, giving the values of p , $\bar{s}$ and V in generation $n+1$ in terms of their values in generation n . They are very easy to solve numerically for particular starting values. We can also allow for a mutation rate λ of $A \rightarrow a$ by letting p take the value $p - \lambda p$ before selection.

As well as the numerical solutions of the recurrence equations, an approximate solution can be obtained algebraically for the case of a deleterious mutation maintained at low frequency by the mutation rate opposing selection. If V is small, changes in V , being of the second order, can be ignored. Thus if the mutation is at the equilibrium frequency of $q = \lambda/\bar{s}$ where λ is the mutation rate as before, then

$$\Delta\bar{s} = - (2\lambda/\bar{s})V/(1-\bar{s}).$$

Now we can put $\Delta\bar{s} = d\bar{s}/dt$ for changes in $\bar{s}$ are very small from one generation to the next. Thus integration gives the time required for the selective coefficient of the heterozygote to change from s to 0. We get for this change

$$t = \frac{1}{2\lambda V} \left[\frac{s^2}{2} - \frac{s^3}{3} \right].$$

To see how far the theory applies in nature it is necessary to get an idea of how large the variance in the expression of the mutant heterozygote may be. Experiments like CLARKE and SHEPPARD's (1963) on crosses in *Papilio dardanus* between an isolated polymorphic race and an isolated monomorphic race revealed very great variability in the expression of the F_2 hybrids. The characters observed, however, were the colours and patterns of particular mimics in the polymorphic race. There were very few individuals; but even if large numbers had been obtained, it would be very difficult to estimate the variances. FORD (1940) did an experiment to select for dominance in the moth *Abraxas grossulariata*. There is a variety *lutea* determined by a mutation that gives an intermediate heterozygote. FORD was able to select for the dominance of *lutea* in one line and its recessiveness in another. He divided the range of the phenotypes into eight classes of which classes 2 to 5 were the heterozygotes. Thus a rough estimate of the variance in the expression of the heterozygotes can be made. From the response to selection in the first generation, the heritability can also be estimated. In the initial cross of *semi-lutea* $\times$ wild type, the variance in the classes of the heterozygotes is about 0.6. A combined estimate of about 0.6 was obtained for the heritability. Therefore the additive genetic variance is about 0.36; and, considering additive genetic effects only, this gives a standard deviation of 0.6 class intervals. Now the initial difference between the wild type and the mean of the heterozygotes is 2.0 class intervals. Thus the additive genetic standard deviation is roughly

$$\sigma_A = 0.3 \times | \text{hom.} - \text{het.} |.$$

Suppose the homozygote has the fitness 1.0 and the heterozygote the fitness $1-s$. If $s = 0.1$, the corresponding additive standard deviation in fitness will be 0.03. The variance will therefore be 0.0009. This is a very small variance and the approximate solution we have obtained should give a fairly reasonable figure for

the number of generations needed for dominance to evolve. If $\lambda = 0.00001$ then applying the formula

$$t = \frac{1}{2\lambda V} \left[\frac{s^2}{2} - \frac{s^3}{3} \right]$$

we get $t = 259,000$ generations. If $\lambda = 0.0001$, then $t = 25,900$ generations. Dominance is therefore attained after a long period of time. But it must be remembered that certain deleterious mutations have been recurring for perhaps millions of generations, so these time-scales are not necessarily too long for dominance to evolve.

O'DONALD (1967) set up models for the selection of fitness of an advantageous gene as it spreads through a population. Semidominance, dominance or overdominance will evolve depending upon the genetic variance in the fitness of the heterozygote. Overdominance evolved and a balanced polymorphism was established if the standard deviation due to genetic causes in the fitness of the heterozygote was about 0.7 times the difference of the homozygote and heterozygote fitnesses at the start of selection. This is a standard deviation more than twice as large as the standard deviation of the expression of the heterozygote in *Abraxas grossulariata*. In general, variances may be larger than that in *Abraxas*. In CLARKE and SHEPPARD's experiments, the heterozygotes appeared to be extremely variable. It may therefore be interesting to see the results of selection for a deleterious mutation with different initial variances in the fitness of its heterozygote. The selective coefficients at the start of selection were for the genotype *AA*, 0.5; for *Aa*, 0.6; and for *aa*, 0.7. These selective coefficients were chosen to allow for overdominance of the heterozygote. It may be argued that the fitness of the wild type will also be subject to hereditary variations and be selected for. In practice, however, the wild type will have been selected for an optimum phenotype. It will therefore be subject to stabilising selection and further increases in its fitness will not be possible. A new mutation will be likely not to be at its optimum phenotypic value and improvements in its fitness will be possible. If the fitness of the wild type were subject to variation and were selected for, it would always be selected faster than the fitness of the mutant heterozygote, for the wild type is the most abundant genotype. Any model for the evolution of dominance and overdominance implies that the wild type is at or near an optimum phenotypic value above which its fitness cannot be improved. Thus given the selective coefficients 0.5, 0.6 and 0.7 for the three genotypes, the effects of selection were found for variances of s of 0.01, 0.008, 0.006 and 0.004. These were in fact phenotypic variances with an heritability of 0.5 giving additive genetic variances half these values.

Figure 1 shows the results of selection for s in populations with these variances when there is mutation of $A \rightarrow a$ at a rate 0.00001 per gamete per gene. When the variance is small more than 100,000 generations are needed for overdominance to evolve. Less than 50,000 generations are needed when the phenotypic variance is 0.01.

Table 1 shows the additive genetic variance at the point at which overdomi-

FIGURE 1.—Selection of the fitness of a deleterious heterozygote maintained by recurrent mutations. The initial selective coefficients were 0.5 for genotype AA , 0.6 for Aa and 0.7 for aa . The selective coefficients are plotted against the number of generations in thousands of generations. The results are given for a mutation rate of 0.00001 and for different initial variances shown in the figure. The additive genetic variances are half those shown.

nance had evolved. It shows there will still be variance for selection to act upon: indeed the fitness of the heterozygote continues to increase and the polymorphism which has been established becomes more and more stable. In practice this will not happen because the fitness of the heterozygote, like the wild type's, is not subject to direct variations but is determined by some character or characters that vary. Just as with the wild type, the fitnesses of the mutant genotypes can only be raised to their optimum values. Overdominance will evolve if the mutant heterozygote has an optimum fitness above the wild type's. If, however, the mutation only changes the value of the wild type, its fitness can only be raised to the wild type's fitness, and dominance but not overdominance will evolve. The mutation must therefore cause a qualitative change in the character having an optimum above the wild type's. As O'DONALD (1967) pointed out, this must have been so for the mutation causing the sickle-cell trait in Africa. The heterozygote certainly has an optimum above the wild type's in areas with malaria. But it is not necessary and most unlikely for the sickle-cell trait to have been at this optimum when it first arose: its fitness would have continually been raised by the process described in this paper until it reached its optimum.

When overdominance has evolved by this process, the mutant gene will start to increase in frequency and move towards the equilibrium point of the balanced polymorphism. The mutant homozygote will then start to appear in appreciable numbers. Sometimes it will be lethal as in the sickle-cell disease. No selection of its fitness could take place and a permanent balanced polymorphism will have been established. But if the mutant homozygote is not lethal, its fitness, too, will

TABLE 1

Initial additive genetic variance	Number of generations to produce overdominance	Additive genetic variance at point of overdominance
0.005	47,300	0.003999
0.004	59,100	0.003198
0.003	78,800	0.002394
0.002	118,100	0.001597

vary and be selected for. Thus the mutant homozygote may also be expected to reach its optimum. If its optimum is below the heterozygote's, a permanent balanced polymorphism will be established as before. If its optimum is equal to or above the heterozygote's, dominance or semidominance will finally evolve and at this point the mutation, though it was at first deleterious, will have become the new wild type. Thus in the long run, the fact that almost all mutations are deleterious when they first arise does not imply that evolutionary changes must necessarily be very slow or that useful adaptations will be very rare.

I am very grateful to W. F. BODMER and J. A. SVED who suggested a more general mathematical formulation of the problem than the one I used originally. I assumed that the selective coefficient would have a normal distribution, whereas they pointed out that my results would hold in general for any distribution $f(s)$. This leads to a more elegant formulation of the problem. I am also grateful for the valuable suggestions they made to improve this paper.

SUMMARY

A mathematical model is described of the effect of selection on a genotype with a variable selective coefficient. The selective coefficient s is assumed to have a frequency density $f(s)$ with mean $\bar{s}$ and variance V . Recurrence relations are derived giving the changes in $\bar{s}$ and V from one generation to the next. They are applied to the case of a deleterious mutation maintained at an equilibrium frequency by a given mutation rate. If the mutant heterozygote has an additive genetic variance in its fitness, selection slowly improves its fitness until dominance evolves. In certain conditions, overdominance, too, may evolve. Thus the initially deleterious mutation may evolve a selective advantage. The length of time for this process depends on the variance of the heterozygote. If the standard deviation due to additive genetic effects is equal to the initial difference between the wild type and the mutant heterozygote, about 48,000 generations are needed with a mutation rate of 0.00001. For smaller standard deviations and variances, longer periods of time are needed, but they may have been available, for certain mutations seem to have been recurring for millions of generations. It is suggested that the polymorphism in the sickle-cell trait may have evolved by this process.

LITERATURE CITED

- CLARKE, C. A., and P. M. SHEPPARD 1963 Interactions between major genes and polygenes in the determination of the mimetic patterns of *Papilio dardanus*. *Evolution* **17**: 404-413.
- CROSBY, J. L., 1963 The evolution and nature of dominance. *J. Theoret. Biol.* **5**: 35-51.
- EWENS, W. J., 1965 Further notes on the evolution of dominance. *Heredity* **20**: 443-450. — 1966 Linkage and the evolution of dominance. *Heredity* **21**: 363-370.
- FISHER, R. A., 1928 The possible modification of the response of the wild-type to recurrent mutations. *Am. Naturalist* **62**: 115-126. — 1929 The evolution of dominance; reply to Professor Sewall Wright. *Am. Naturalist* **63**: 553-556.
- FORD, E. B., 1940 Genetic research in the Lepidoptera. *Ann. Eugenics* **10**: 227-252.
- O'DONALD, P., 1967 On the evolution of dominance, over-dominance and balanced polymorphism. *Proc. Roy. Soc. London* (in press).
- SHEPPARD, P. M., and E. B. FORD, 1966 Natural selection and the evolution of dominance. *Heredity* **21**: 139-147.
- WRIGHT, S., 1929a Fisher's theory of dominance. *Am. Naturalist* **63**: 274-279. — 1929b The evolution of dominance. Comment on Dr. FISHER's reply. *Am. Naturalist* **63**: 556-561.