

Solutions to the Cost-of-Selection Dilemma

(substitutional load/gene substitution/evolutionary rate)

VERNE GRANT AND ROBERT H. FLAKE

Departments of Botany and Electrical Engineering, University of Texas, Austin, Texas 78712

Contributed by Verne Grant, July 22, 1974

ABSTRACT There are various biologically realistic ways around the cost-of-selection restriction on rapid multiple-gene substitution. Some of these ways depend upon particular forms of interaction or linkage in the genes undergoing substitution; other ways depend on particular conditions of population size and structure. The special genotypic and populational conditions required for rapid evolutionary change in genetically complex characters are not unusual in higher organisms.

The restrictive effect of the cost of selection, or of the substitutional genetic load, on the rate of evolution was pointed out by Haldane in his 1957 paper (1). The theoretical ceiling imposed by selection cost on evolutionary rate is consistent with empirical evidence in slow-rate and moderate-rate groups (1), but is apparently in disagreement with some other empirical evidence in high-rate lines. Some groups of organisms have undergone evolutionary changes in multifactorial characters and character combinations at rates apparently exceeding those imposed by a tolerable cost of selection (2).

This cost-of-selection dilemma has been extensively discussed in the literature and in informal conferences. In our experience, much of the discussion, both formal and informal, has displayed a marked single-mindedness, with different students of the problem preferring their own solutions and paying little attention to other workers' solutions. A broad approach to the problem indicates that there is no single solution to the cost-of-selection dilemma.

Haldane's two papers (1, 3) were short, condensed, and highly suggestive. Haldane set up the problem, briefly indicated some possible solutions, and foreshadowed others, but deliberately left most of the problem-solving to later workers. Since the problem is multi-faceted, it was possible and perhaps inevitable that different workers should seek solutions along different lines. In our attempt to study the problem broadly and systematically, we have found it necessary to amplify Haldane's original model, and to specify some assumptions which were only implicit in the original papers. When this is done, the various solutions in the literature fall into place as products of varying one assumption or another in the original model, and they will be presented in this way here.

Related problems

Although we are concerned with the substitutional load in this paper, the interest of many other workers has shifted to the related but different problem of the segregational load. Brief reference to the latter is desirable before proceeding with our main theme, since some of the solutions proposed for the segregational load problem may be applicable to the substitutional load problem, and vice versa.

It is well known that high levels of polymorphism for enzyme loci, as disclosed by electrophoretic methods, occur in

a wide variety of animals and plants. The observed high levels of polymorphism are difficult to reconcile with the cost-of-selection limitation if the polymorphic genes are in fact under the control of selection, whereas the problem disappears if the polymorphic genes are selectively neutral. The problem of the segregational load has been discussed in a number of studies (4-20).

The question as to whether many polymorphic genes in natural populations are selectively neutral or not is currently under active discussion, with selectionists and neutralists lined up on opposite sides of the question. We ourselves happen to be in the selectionist camp, and follow Allard, Wills, Clark, Ayala, and other similar authors (14, 15, 17, 18, 21, 22), but do not propose to discuss this controversy further here. The related problem of the rate of amino-acid substitutions in protein evolution has also been discussed recently from opposite viewpoints by Kimura and Ohta (23) and Clarke (22, 24), and is also controversial at present.

In view of the focus and emphasis in current discussions, it is worthwhile to note that the problem of occasional rapid evolution of complex adaptive characteristics remains with us no matter how the controversy over selective neutrality is resolved. Even if a large class of selectively neutral genes could be shown to exist in natural populations, we would still have to seek explanations for rapid evolutionary changes involving adaptively valuable genes.

Basic assumptions

Haldane's original model (1) contained a number of elements, which can be paraphrased as follows. (i) The population is very large or effectively infinite in size. (ii) The population size remains constant or essentially so from generation to generation. (iii) A new superior allele is present at a low initial frequency in the population. (iv) The selective value of the new allele remains constant or nearly so during the course of selection. Haldane later (3) dealt with complications of (iv) which arise when selective values change during the course of gene replacement.

Some other conditions were implicit but not developed in the original paper of Haldane (1). (v) The population has a fixed reproductive potential in excess of the bare replacement level which allows for a certain fraction of selective deaths. (vi) The reproductive potential allows for a regular quota of accidental deaths as well as selective deaths. (vii) The population lives under conditions of strong interspecific competition (compare ref. 1, p. 523).

In going from single-gene selection to selection for two or more genes Haldane made two further assumptions. The several genes undergoing selection are independent of one another (viii) in gene action and (ix) in inheritance.

We have recently shown (25) that postulates (ii) and (iv) are not necessary in deterministic models for obtaining the standard estimate of selection cost. Population size and selection intensity can fluctuate without affecting the calculated total cost.

It is obvious that deviations from the necessary conditions listed above can occur in natural situations, with effects on selection cost.

Avoidance of cost

One way to solve the cost-of-selection problem is to alter the assumption (iii) above that the allele or alleles progressing toward fixation have a consistent positive selective value. Kimura (26) and others have in fact suggested that some genes undergoing replacement by stochastic processes may be selectively neutral. The relevance of this suggestion hinges on the premise that substantial numbers of selectively neutral genes exist in natural populations, a premise which is far from being established. In any case, as we noted earlier, the proposed solution leaves untouched the central problem of the mode of fixation of adaptively valuable gene systems in a relatively short time span.

Van Valen (27) has argued that the cost of gene substitution may be avoided or nullified in those situations where an increase in population fitness accompanies the substitution. The underlying assumption here is that during the substitution process the carriers of a new favorable allele are increasing in numbers, with one or more on the average replacing each individual of the inferior type that is lost. Therefore, the mean population fitness is increasing during the rise in frequency of the slightly superior type. This results in a mild increase in the size of the population thus evading a rigid interpretation of condition (ii).

Brues (28, 29) has developed this line of thought further. The substitutional load is viewed as an artifact arising from the change in the base level of population fitness from the old optimum, when the new favorable allele was not in the picture, to a new optimum when it is present. In other words, a substitutional load accompanies a rise in population fitness, whereas the absence of a substitutional load may go hand in hand with extinction.

The real issue here, as regards both the Van Valen and Brues models, is not that the substitution could be successful in such cases, considering the substitution cost, but rather is the length of time required for the process. Haldane's cost concept still imposes a limit on how fast the substitution can take place.

Amelioration of cost

Wallace (30) has drawn a distinction between "hard selection" and "soft selection". In the case of hard selection, the difference between a selective death and a surviving adult is determined by rigid selective factors which operate independently of population size and normal environmental fluctuations. An example would be the difference between a lethal type and a viable genotype in a standard environment. In soft selection, by contrast, the proportion of genetic deaths is not fixed by rigid selective factors, but varies with changes in population size and environmental conditions. For example, zygotes of a given genotype might fail to survive in an unfavorable year when the whole population is reduced in numbers; but the same genotypic class of zygotes could survive and reach the adult breeding stage in a favorable year when the whole population is large.

With soft selection, in other words, the carriers of an unfavored allele may become selective deaths under stringent environmental conditions, but not under more permissive conditions. Accordingly, genes subject to soft selection would reduce the genetic load, as calculated on the basis of hard selection, and hence could reduce the cost of selection (10, 11, 30). This is in effect modifying the assumption (iv) of constant selective values.

Haldane (1) pointed out that the selection deaths would be less detrimental to the life of the population if they occurred in the embryonic or juvenile stages rather than in the adult stage.

Gene interaction, linkage, and cost

The solutions mentioned so far apply basically to the cost of single-gene substitutions in a population, and then, by simple extension, to substitutions of a larger number of genes. The solutions to be considered next differ from the foregoing in that they postulate particular relationships between two or more genes which lower the total cost of replacement. These solutions alter assumptions (viii) and (ix).

Several authors have pointed out that the cost of selection for a given number of separate genes could be reduced by certain modes of gene interaction. Mayr (31), and Mettler and Gregg (32) suggest that separate genes undergoing selection simultaneously may often tend to have correlated selective advantages or disadvantages. Different favored alleles will tend to occur together in the same genotype, and, conversely, different unfavored alleles will also tend to occur together in the same alternative genotype. Then, the genetic deaths for numerous separate genes will be pooled in a smaller number of individual deaths.

King (6, 7), Maynard Smith (33), and Crow (9) emphasize, and Felsenstein (34) discusses, a truncated model of selection, as found in artificial selection for quantitative characters, wherein the selective elimination falls at one end of a more or less normal frequency distribution of genotypes. The component genes determining the quantitative trait are supposed to have additional interactions so as to produce threshold effects in the phenotype and lead to the truncated mode of selection (3, 6, 7, 9). The cost of truncated selection for all component genes combined would be substantially less than that of the sum of the component genes taken individually. This model is similar to the one proposed by Mayr (31) and Mettler and Gregg (32).

Another possibility offering much hope for rapid evolutionary changes in complex characteristics entails the use of supergenes in place of sets of independent genes. If the various separate alleles comprising a new adaptive gene combination are brought into close linkage relations as a supergene, by means of the appropriate chromosomal rearrangements, the several genes involved can undergo selection together at the rate and cost of one gene. The substitutional load for a supergene is no greater than that for a single Mendelian gene. The population is thus able to acquire a new adaptive gene combination for the price of a single gene substitution.

A less extreme case would be that of strong linkage between the various genes comprising the valuable gene system. Kimura and Crow (35) touch on this point briefly.

Population size and cost

Haldane had disregarded the effects on gene frequencies of a random sampling of gametes in finite populations (assumption i). In finite populations there will of course be random

fluctuations in gene frequencies. Kimura and Maruyama (36) have treated the effects of the random element mathematically. They show that random sampling of gametes does have a significant effect on the substitutional load, and under some conditions reduces the load.

A suggestion put forward by Sved (37), postulates a probit fitness model in which population size is controlled by density-dependent factors. In this model, according to Sved, very high rates of gene substitution are possible with moderate selection intensities.

One could assume that the population, during some time period in its history, has a high excess fecundity which is not counterbalanced by a similarly high accidental mortality rate, so that it is expanding. This is in effect altering conditions *vi* and *vii*. Then the population can tolerate an unusually high number of selective deaths during the expansion period. Haldane (1) touched briefly on a possibility similar to this. This problem requires further study.

Population structure and cost

A colonial population system and a series of founder colonies seem to offer ways of escaping the cost-of-selection restriction that are not available in a large continuous population. These ways have been described in a previous paper (2) and will be mentioned only briefly here.

Although the overall cost of gene substitution is the same in a large continuous population and a subdivided population of the same total size, the accidental mortality rate will show colony-to-colony variation in the subdivided population; and the colonies with low accidental mortalities could tolerate the heavier burden of selective deaths that accompany rapid evolution.

Finally, consider a series of founder colonies derived from a large polymorphic ancestral population. A favorable allele which is rare in the parental population is likely to be established by random factors at middle or high frequencies in one or more of the daughter colonies. This nonselective process gets the colony over the critical stage when the frequency of the favored allele is low and the incremental cost of selection is greatest.

The main emphasis in cost-of-selection studies up to now has been on theoretical analysis of the substitution process. Various mechanisms for rapid substitution are now known in theory. However, there have been no empirical studies to date focused on cost of selection itself.

This study was supported in part by NSF Grant GB-16411 and in part by a special grant from the University Research Institute of the University of Texas; both grants are gratefully acknowledged.

- Haldane, J. B. S. (1957) "The cost of natural selection," *J. Genet.* 55, 511-524.
- Grant, V. & Flake, R. H. (1974) "Population structure in relation to cost of selection," *Proc. Nat. Acad. Sci. USA* 71, 1670-1671.
- Haldane, J. B. S. (1960) "More precise expressions for the cost of natural selection," *J. Genet.* 57, 351-360.
- Lewontin, R. C. & Hubby, J. L. (1966) "A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*," *Genetics* 54, 595-609.
- Lewontin, R. C. (1967) "Population genetics," *Annu. Rev. Genet.* 1, 37-70.
- King, J. L. (1966) "The gene interaction component of the genetic load," *Genetics* 53, 403-413.
- King, J. L. (1967) "Continuously distributed factors affecting fitness," *Genetics* 55, 483-492.
- Kojima, K. & Yarbrough, K. N. (1967) "Frequency dependent selection at the esterase-6 locus in *Drosophila melanogaster*," *Proc. Nat. Acad. Sci. USA* 57, 645-649.
- Crow, J. F. (1970) *Mathematical Topics in Population Genetics*, ed. Kojima, K. (Springer Verlag, New York).
- Wallace, B. (1970) *Genetic Load* (Prentice-Hall, Englewood Cliffs, N.J.).
- Dobzhansky, Th. (1970) *Genetics of the Evolutionary Process* (Columbia Univ. Press, New York).
- Kojima, K. (1971) "The distribution and comparison of 'genetic loads' under heterotic selection and simple frequency-dependent selection in finite populations," *Theoret. Pop. Biol.* 2, 159-173.
- Kimura, M. & Ohta, T. (1971) *Theoretical Aspects of Population Genetics* (Princeton Univ. Press, Princeton, N.J.).
- Clegg, M. T. & Allard, R. W. (1972) "Patterns of genetic differentiation in the slender wild oat species *Avena barbata*," *Proc. Nat. Acad. Sci. USA* 69, 1820-1824.
- Allard, R. W., Babbel, G. R., Clegg, M. T. & Kahler, A. L. (1972) "Evidence for coadaptation in *Avena barbata*," *Proc. Nat. Acad. Sci. USA* 69, 3043-3048.
- Stebbins, G. L. & Lewontin, R. C. (1972) *Darwinian, Neo-Darwinian, and Non-Darwinian Evolution*, eds. Le Cam, L. M., Neyman, J. & Scott, E. L. (Univ. California Press, Berkeley, Calif.).
- Ayala, F. J. & Anderson, W. W. (1973) "Evidence of natural selection in molecular evolution," *Nature* 241, 274-276.
- Ayala, F. J. & Tracey, M. L. (1974) "Genetic differentiation within and between species of the *Drosophila willistoni* group," *Proc. Nat. Acad. Sci. USA* 71, 999-1003.
- Bryant, E. H. (1974) "On the adaptive significance of enzyme polymorphisms in relation to environmental variability," *Amer. Natur.* 108, 1-19.
- Johnson, G. B. (1974) "Enzyme polymorphism and metabolism," *Science* 184, 28-37.
- Wills, C. (1973) "In defense of naive pan-selectionism," *Amer. Natur.* 107, 23-34.
- Clarke, B. (1973) "Neutralists vs. selectionists," *Science* 190, 600-601.
- Kimura, M. & Ohta, T. (1971) "On the rate of molecular evolution," *J. Mol. Evolu.* 1, 1-17.
- Clarke, B. (1969) "Darwinian evolution of proteins," *Science* 168, 1009-1011.
- Flake, R. H. & Grant, V. (1974) "An analysis of the cost-of-selection concept," *Proc. Nat. Acad. Sci. USA* 72, 3716-3720.
- Kimura, M. (1968) "Evolutionary rate at the molecular level," *Nature* 217, 624-626.
- Van Valen, L. (1963) "Haldane's dilemma, evolutionary rates and heterosis," *Amer. Natur.* 97, 185-190.
- Brues, A. M. (1964) "The cost of evolution vs. the cost of not evolving," *Evolution* 18, 379-383.
- Brues, A. M. (1969) "Genetic load and its varieties," *Science* 164, 1130-1136.
- Wallace, B. (1968) *Topics in Population Genetics* (Norton, New York).
- Mayr, E. (1963) *Animal Species and Evolution* (Harvard Univ. Press, Cambridge, Mass.).
- Mettler, L. E. & Gregg, T. G. (1969) *Population Genetics and Evolution* (Prentice-Hall, Englewood Cliffs, N.J.).
- Maynard Smith, J. (1968) "'Haldane's dilemma' and the rate of evolution," *Nature* 219, 1114-1116.
- Felsenstein, J. (1971) "On the biological significance of the cost of gene substitution," *Amer. Natur.* 105, 1-11.
- Kimura, M. & Crow, J. F. (1969) "Natural selection and gene substitution," *Genet. Res.* 13, 127-141.
- Kimura, M. & Maruyama, T. (1969) "The substitutional load in a finite population," *Heredity* 24, 101-114.
- Sved, J. A. (1968) "Possible rates of gene substitution in evolution," *Amer. Natur.* 102, 283-293.