GENETIC LOAD: THREE VIEWS[1]

2. The Concept of Genetic Load: A Reply

JAMES F. CROW

Department of Medical Genetics, School of Medicine, University of Wisconsin, Madison, Wisconsin

I AM GLAD TO ACCEPT the invitation of the Editor to reply to Dr. Sanghvi's article which appears in this issue of the *American Journal of Human Genetics*. I should like also to use this opportunity to answer an article by C. C. Li (1963a) in which some of the same points arise. Dr. Li's arguments are repeated in another paper (1963b). Since all three articles involve interpretation of formulae and the choice of definitions, I would have preferred to deal with these in private correspondence with the two authors. On the other hand, if some previously obscure points are clarified, this public exchange is perhaps useful.

The basic idea of the genetic load was first discussed by Haldane (1937) in a paper entitled "The effect of variation on fitness." In this paper Haldane showed that the decrease of fitness in a species as a consequence of recurrent mutation is approximately equal to the total mutation rate multiplied by a factor that ranges from 1 to 2 depending on dominance, inbreeding, and possible sex-linkage. On the basis of prevailing data on mutation rates, Haldane estimated that the average fitness of the population is decreased by about 5 per cent from this cause. The principle is remarkable in that the decrease in fitness at equilibrium does not depend on the degree of harmfulness of the mutant, provided that its average selective disadvantage in the various genotypes that comprise the population is large relative to the mutation rate and that it is deleterious or at best neutral in the heterozygous state. Essentially the same principle was discovered independently a few years later by Muller and was discussed fully in his 1950 paper, "Our load of mutations." Haldane also pointed out in his 1937 paper that for a locus in which the heterozygote is superior in fitness to either homozygote the average population fitness will be less than that of the heterozygote (although greater than either homozygote). He noted further that the decrease in population fitness, relative to the best genotype, is usually considerably larger for a locus whose variability is maintained by heterozygote superiority than for one maintained by recurrent mutation.

The decreases in fitness from these two causes have been designated as the *mutation load* and the *segregation load* (Morton, Crow and Muller, 1956; Crow, 1958). One can also consider the load that is ordinarily hidden by heterozygosis, but which would be revealed if the population were somehow made homozygous. This was called the inbred load and can be estimated by extrapolation from data on the progeny of parents of a specified degree of relationship, which are homozygous for a known fraction of otherwise heterozygous loci.

Received May 31, 1963.

[1]See Editor's Note, page 298.

The expected ratio of the inbred load to the randomly mating load can be expressed in terms of meaningful genetic parameters. With a locus where two or more alleles are maintained by superior fitness of heterozygous combinations I have shown (1958, 1961) that the ratio is equal to or less than the number of alleles maintained in the population (or 2, if there are only two alleles). For a locus where the mutant alleles are maintained by mutation the ratio is approximately equal to the reciprocal of twice the average dominance, provided the mutants have sufficient dominance that most mutant gene eliminations occur through heterozygous effects. Data on dominance of Drosophila lethals suggest a ratio of about 20. For a completely recessive mutant the ratio is usually still greater, being $\sqrt{s/u}$, where s is the selective disadvantage of the mutant homozygote and u is the mutation rate. Thus, a high ratio is an argument that the majority of loci are mutational, or if segregational that a large number of mutually heterotic alleles are maintained in the population. These alleles must be sufficiently more advantageous in heterozygous combinations than in homozygotes that a large number are maintained in a stable polymorphism, despite the loss of alleles by random drift which increases roughly as the square of the number of alleles (Kimura, 1955, 1956), a circumstance that Morton, Crow and Muller (1956) regarded as unlikely for more than a minority of loci.

In the special case of the segregation load where there are only two alleles (the case treated by Li and Sanghvi), the ratio of the inbred to the random mating load is two. Li says that this is an artifact of my particular system of notation. However, the ratio as I have defined it is the same in any notation. The genetic load was defined in the paper Li refers to (Crow, 1958) as "the proportion by which the population fitness (or whatever other trait is being considered) is decreased in comparison with an optimum genotype." If w_m is the fitness of the genotype with the highest average fitness, the genetic load of a randomly mating and an inbred population are defined in Li's terminology as:

$$L_0 = \frac{w_m - \bar{w}_0}{w_m}, \quad L_1 = \frac{w_m - \bar{w}_1}{w_m}, \quad \frac{L_1}{L_0} = \frac{w_m - \bar{w}_1}{w_m - \bar{w}_0}$$

where the subscript 0 means random mating ($F = 0$) and 1 means completely homozygous ($F = 1$).

Using Li's (1963a) "System I" notation, which he says gives a different result, $w_m = 1 + Hs$, $\bar{w}_0 = 1 + Hsq$, $\bar{w}_1 = 1 - sq$, and the equilibrium value of $q = H/(1 + 2H)$. Substituting these into the above equations leads to $L_1/L_0 = 2$, as expected. The quantity L_1/L_0, as I have defined it, is just as "invariant" with respect to changes in notation as the ratio, $\bar{w}_1/\bar{w}_0$, that Li and Sanghvi advocate.

Since, as emphasized in my paper, the load involves only the relative fitnesses of the three genotypes and not their absolute values, I arbitrarily assigned the value 1 to w_m. This was done solely for algebraic convenience, and I hope no reader was led to think that this value has some absolute meaning.

Sanghvi makes the point on page 303 that if two homozygotes are of unequal fitness, most of the segregation load is attributable to the less harmful homozygote. I agree, and have already emphasized the same point (Crow, 1961; Hiraizumi and Crow, 1960). Sanghvi goes on to say that for this and other reasons the segregation load as I have defined it is an "algebraic artifact" and argues that

the reference point should not be the best genotype, but rather the best *homozygous* genotype. This is, of course, one possible definition; but it should be no surprise that a different definition leads to different results.

I counted myself fortunate in having found a simple relationship, L_1/L_o, which has a strikingly different value for two kinds of gene loci (a locus where harmful alleles are maintained by mutation as contrasted to a small number of alleles maintained by heterozygote superiority), and which therefore offers the possibility of distinguishing between them experimentally. Li and Sanghvi say that the ratio $\bar{w}_1/\bar{w}_o$ is more appropriate. It may be for some purposes, but not for this, for its value usually differs only slightly for the two gene models and doesn't have (for me, at least) any such transparent interpretation in terms of allele numbers and average dominance, as does L_1/L_o. Therefore it provides no basis for distinguishing between these two contrasting models of inbreeding effect.

Li says that I wrongly conclude that "population fitness decreases on inbreeding to a much greater extent for a mutational equilibrium population than for a heterotic one." I reached no such conclusion. My conclusion was rather the one that I have just discussed, which is that the *ratio* of the inbred to the random *load* is greater for a mutational locus than for a segregational one. (My 1958 paper has a carelessly worded statement at the bottom of page 10 that when taken out of context may be misleading. I said: "Unless the segregation load is based on loci with a large number of alleles, maintained in balanced polymorphism, it is not changed much by inbreeding. In general, I would conclude that if the fitness is greatly decreased by inbreeding it is to that extent largely attributable to the mutation load rather than the segregational." I should perhaps have said the fitness decrease *relative to the random mating load,* but I thought this would be clear from the overall context, and in particular from the preceding sentence. If this sentence is the source of Dr. Li's misunderstanding, I apologize to him and to any other reader who was confused by it.)

Sanghvi implies that I was in error in not taking into consideration the change of gene frequencies by selection during inbreeding. On the contrary, I stated explicitly in the 1958 paper and in subsequent papers that the L_1/L_o ratio can be interpreted in terms of allele numbers and dominance only if the inbreeding occurs without a change in gene frequencies in a population that has reached equilibrium under the previous mating system. For this reason, my colleagues and I (Morton, Crow and Muller, 1956; Morton, 1960; Greenberg and Crow, 1960) have been careful to apply the theory only to populations where all the inbreeding occurred in a single generation, thus permitting no gene frequency changes. The methods are not applicable (without considerable modification) for populations where there is a history of inbreeding over several generations, as in most domestic livestock.

Both Li and Sanghvi believe that the standard of reference for the segregation load should not be the best genotype, as I have chosen, but rather the best homozygous genotype. However, the relationships that I have just pointed out depend on the proper definition. Other definitions are of course possible, but I would urge that the words "genetic load" be restricted to my definition and that other words be substituted when the definition is changed. To change definitions

at this stage leads only to confusion. I should also emphasize that the relationships that I have discussed are applicable only to data that are measured in a way appropriate to the definition. For example, since the measurement is from an optimum genotype the applicability is to data where there is a clear optimum—for example, premature death rate where the optimum is obviously zero (Morton, Crow and Muller, 1956).

To me it has always seemed natural on other grounds as well to measure the segregation load as a deviation from the best genotype, not from the best *homozygous* genotype. With a large number of alleles, the choosing of an inferior genotype simply because it is homozygous seems more arbitrary. Also, I have difficulty in defining the load for a balanced lethal using a homozygote as the reference point. In a later paper (1961) I suggested another definition: the genetic load is the proportion by which the average fitness (or other trait) is decreased in comparison with what it would be if the factor under consideration were absent. This was not (as Sanghvi suggests) an attempt to rectify any defect in the earlier definition, but rather to view the question from another perspective. With this definition the choice of the best genotype (in this case a heterozygote) as the reference point for the segregation load is clear; in a population without segregation (*i.e.*, reproducing asexually) the best genotype, heterozygote or homozygote, would replace the others.

However, the value of a definition is mainly determined by its ability to aid in the discovery of new relationships. Besides the L_1/L_0 test already referred to, our definition leads to a method of estimating the segregation load for any number of alleles from information on the fitness and frequency of any one of them (Morton, 1960; Crow, 1961). It has also been useful in more difficult problems such as arise with maternal-foetal incompatibility (Crow and Morton, 1960), selection for an intermediate metrical phenotype, and the effects of finite populations and varying environmental conditions (Kimura and Crow, unpublished).

On the other hand, I do not want to make any exaggerated claims for the method. It may turn out that the assumptions are too restrictive to be applicable to many real situations. The L_1/L_0 ratio suffers from the fact that while a high value suggests mutational loci as the major cause of inbreeding decline a low value offers no evidence for the contrary hypothesis, since it may simply be due to errors of measurement, to a mixture of dominant and recessive mutants, to inflation of the denominator by a large non-genetic component, or to other irrelevant factors. The theory applies only to traits that are highly correlated with fitness (strictly, only to fitness itself). The $L/_1L_0$ criterion does not apply when selection of different genotypes is on different traits or on different parts of the life cycle. For example, in a two allele segregational locus if one homozygote has a high death rate while the other has a low fertility, a study of death rates alone would be misleading. For this reason, Morton (1960) argued for mutational rather than segregational maintenance of alleles for mental deficiency primarily because of the large value of L_0 rather than because of a high L_1/L_0 ratio. Another difficulty is that a population in a heterogeneous environment and with division of labor may not have an optimum genotype with respect to some loci. In other situations, for example in Wright's theory of evolution in

a subdivided population, the average fitness and genetic load may not be the most important considerations from a long time view. Finally, there are many difficulties and deficiencies at the data level, and as emphasized before, the data used must be consistent with the assumptions of the model. Despite such difficulties, I believe the methods when used correctly are helpful for a number of problems in population genetics.

My own views of population structure are misinterpreted by Sanghvi on page 301. Morton, Muller and I (1956) did not regard our analysis as "evidence that segregational loci do not make any substantial contribution to our hereditary burden." If Sanghvi means, as I think he does, the *expressed* hereditary burden, our study offered very little evidence on this point; what we did suggest from the high L_1/L_0 ratio was that such loci make a relatively small contribution to the *hidden* load revealed by inbreeding. I have argued long ago (1948, 1952) for the importance of overdominant loci and other selectively balanced systems in population fitness. But, because of their tendency to accumulate homozygous-detrimental alleles in the population at high frequencies, such loci exert an influence on population fitness out of all proportion to their number, and therefore a large effect could be caused by a relatively small number of loci.

The effect of mutation on the normal incidence of disability and on fitness in general is very difficult to assess from ordinary data, because of the confounding effects of such things as loci maintained by selective balance and non-genetic effects. Except for specific traits that are amenable to segregation analysis it is almost impossible to assess the mutational component. On the other hand, the inbred load is influenced relatively more by the kind of mutant whose frequency depends on the mutation rate than is the random load, so the study of consanguineous matings is a way of making an assessment of mutational effects that is less contaminated from segregational loci than if it were based on the study of progeny of unrelated parents.

There are a number of other points where I believe that Li and Sanghvi are in error, or have misunderstood our papers, but these are of less general interest and are best discussed in private correspondence.

REFERENCES

CROW, J. F. 1948. Alternative hypotheses of hybrid vigor. *Genetics* 33: 477-487.

CROW, J. F. 1952. Dominance and overdominance. In *Heterosis*, J. W. Gowen, ed., pp. 282-297. Ames: Iowa State College Press.

CROW, J. F. 1958. Some possibilities for measuring selection intensities in man. *Hum. Biol.* 30: 1-13.

CROW, J. F. 1961. Population genetics. *Amer. J. Hum. Genet.* 13: 137-150.

CROW, J. F., AND MORTON, N. E. 1960. The genetic load due to mother-child incompatibility. *Amer. Natur.* 94: 413-419.

GREENBERG, R., AND CROW, J. F. 1960. A comparison of the effect of lethal and detrimental chromosomes from Drosophila populations. *Genetics* 45: 1153-1168.

HALDANE, J. B. S. 1937. The effect of variation on fitness. *Amer. Natur.* 71: 337-349.

HIRAIZUMI, Y., AND CROW, J. F. 1960. Heterozygous effects on viability, fertility, rate of development, and longevity of Drosophila chromosomes that are lethal when homozygous. *Genetics* 45: 1071-1083.

KIMURA, M. 1955. Random genetic drift in a multi-allelic locus. *Evolution* 9: 419-435.

KIMURA, M. 1956. Rules for testing stability of a selective polymorphism. *Proc. Nat. Acad. Sci.* 42: 336-340.

KIMURA, M. 1961. Some calculations on the mutational load. *Jap. J. Genet.* 36 (suppl): 179-190.

LI, C. C. 1963a. Decrease of population fitness upon inbreeding. *Proc. Nat. Acad. Sci.* 49: 439-445.

LI, C. C. 1963b. Genetic aspects of consanguinity. *Amer. J. Med.* 34: 702-714.

MORTON, N. E. 1960. The mutational load due to detrimental genes in man. *Amer. J. Hum. Genet.* 12: 348-364.

MORTON, N. E., CROW, J. F., AND MULLER, H. J. 1956. An estimate of the mutational damage in man from data on consanguineous marriages. *Proc. Nat. Acad. Sci.* 42: 855-863.

MULLER, H. J. 1950. Our load of mutations. *Amer. J. Hum. Genet.* 2: 111-176.