GENETIC LOAD: THREE VIEWS[1]

3. The Way the Load Ratio Works

C. C. LI

Department of Biostatistics, University of Pittsburgh, Pittsburgh, Pennsylvania

THE CONCEPT OF GENETIC load and the method of load ratio have been an interesting and popular subject in population and human genetics for the last few years as judged by the large amount of literature citing the concept and the method. The open discussions of Sanghvi and Crow, presented in the foregoing pages, are most timely, and they have undoubtedly contributed a great deal toward clarification of the method of load ratio. In every frontier of science there are disagreements and controversial subjects, and genetics, being in a healthy condition, is no exception. I believe that it is through these discussions that we make progress in formulating our concepts and in devising our methodology. At this stage, it is not at all a matter of right or wrong. The history of science shows that scientists often start with "wrong" ideas and as knowledge accumulates they gradually modify original concepts and reach better ones. From the historical viewpoint, an original "wrong" paper may contribute just as much as a final "correct" one. In the present case, it happens that I have also discussed the problem of genetic load elsewhere (Li, 1963 a, b). The only reason that I write this rejoinder is my belief that open discussions will benefit, not harm, everyone concerned, in the long run if not immediately. In the following I shall confine myself to the main features and implications of the load ratio. As brief as it is, I shall try to make this communication more or less self-contained so that it may be read independently without constantly referring to other papers.

THE BASIC PROBLEM

Let w_1, w_2, w_3 be the relative fitness of genotypes *AA*, *Aa*, *aa*, respectively, and let us consider the case in which w_3 is much smaller than w_1 and w_2; that is, the genotype *aa* leads to the development of a serious disease or condition that impairs reproductive ability (not necessarily involving deaths). We then say that selection is against genotype *aa*. In an equilibrium population, there must be some mechanism of compensation for the loss of gene *a* through selection. Indeed, there could be many different kinds of compensation in nature, about which, unfortunately, we know too little. However, if $w_1 > w_2 > w_3$, the simplest compensation mechanism is recurrent mutation from allele *A* to allele *a*. We shall call this a "mutational" equilibrium for brevity. On the other hand, if $w_1 < w_2 > w_3$, the higher reproductive ability of the heterozygote (*Aa*) will maintain the gene *a* in the population without the help of new mutations in

Received June 5, 1963.

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

every generation. We may call this a "heterotic" equilibrium for brevity. Since we deal with the case in which w_3 is much smaller than w_1 and w_2, the type of equilibrium is essentially determined by $w_1 > w_2$ or $w_1 < w_2$. The two types of equilibria mentioned above are by no means the only possible ones in nature but we shall confine our discussion to these two types.

Now, consider a random mating population in which the frequency of gene *A* is p = .98 and that of gene *a* is q = .02. The population consists of, phenotypically,

$$.9604 + .0392 = .9996\ (AA + Aa) \text{ and } .0004\ (aa).$$

Genotypes *AA* and *Aa* are all apparently normal, and even if they are different to some slight degree in their fitness values, they are certainly indistinguishable without detailed direct studies. The basic problem is: what type of equilibrium is this? Mutational or heterotic? There is no way to tell by looking at the genetic composition of the population. Since genotypes *AA* and *Aa* are hardly distinguishable, the relative magnitudes of w_1 and w_2 are totally unknown. If we render the population completely homozygous (actually by partial inbreeding and extrapolation) without changing the gene frequencies, the population would consist of

$$p = .9800\ (AA) \text{ and } q = .0200\ (aa).$$

The most conspicuous result of complete inbreeding is that the proportion of *aa* has increased from .0004 to .0200; or $q/q^2 = 1/q = 1/.02 = 50$ times. This will be so whether the equilibrium is mutational or heterotic as the genes remember no history (the Markov property). All that sudden and complete inbreeding does is to transform the genotypic distribution (p^2, 2pq, q^2) to (p, 0, q), whatever the type of the equilibrium. This is one of the arguments advanced by Li (details in 1963b) who concluded that inbreeding cannot help us to distinguish the two types of equilibria; in other words, inbreeding is no tool to diagnose why deleterious genes are in the population or how they are maintained there. Crow's (1958, 1963) contention is that inbreeding results can distinguish mutational equilibrium from an heterotic one, and this is accomplished by using a quantity known as "genetic load." Let us now examine how it works.

THE LOAD RATIO METHOD

Adopting the notation of Crow (1963), we let $w_m = \max(w_1, w_2, w_3)$ be the largest of the three fitness values. The genetic load at the random mating state is $L_o = (w_m - \bar{w}_o)/w_m$ and that at the completely inbred state is $L_1 = (w_m - \bar{w}_1)/w_m$, so that the ratio of these two loads is

$$L_1/L_o = (w_m - \bar{w}_1)/(w_m - \bar{w}_o).$$

Crow (1958, 1963) showed that the value of this ratio is always equal to 2 for a population in heterotic equilibrium and the ratio is large for a population in mutational equilibrium. Thus, through inbreeding and calculating the value of L-ratio, one would be able to tell whether the equilibrium is maintained by

heterosis or by recurrent mutation. Table 1 gives a numerical illustration of the method using the population with $p = .98$ and $q = .02$ mentioned previously. There are an infinite number of possible heterotic equilibria that can maintain the frequency of the deleterious gene at the $q = .02$ level, but for the particular

TABLE 1. TWO MODELS OF EQUILIBRIUM (HETEROTIC, H; AND MUTATIONAL, M) WITH GENE FREQUENCIES $p = .98$ AND $q = .02$

Genotype	Frequency (random mating)	Frequency (completely inbred)		Population H fitness w	Population M fitness w
AA	.9604	.9800		980	1000
Aa	.0392	0		1000	980
aa	.0004	.0200		20	20
random mating		mean	$\bar{w}_0$	980.400	998.824
		load	L_0	19.600	1.176
completely inbred		mean	$\bar{w}_1$	960.800	980.400
		load	L_1	39.200	19.600
ratio of means			$\bar{w}_1/\bar{w}_0$	0.9800	0.9816
load ratio			L_1/L_0	2.00	16.7

numerical example chosen $w_1 : w_2 : w_3 = 98 : 100 : 2$. Likewise, there are an infinite number of possible mutational equilibria that can also maintain the level of the deleterious gene at $q = .02$, and I have chosen the situation $w_1 : w_2 : w_3 = 100 : 98 : 2$ for an example. It is important to note that $w_m = w_2$ in the former case, and $w_m = w_1$ in the latter. In calculating the genetic load (lower portion of Table 1), I have merely used the deviation $w_m - \bar{w}$, omitting the denominator w_m for simplicity, as this does not affect the value of the ratio L_1/L_0.

From the numerical calculations it is clear that the mean fitness $\bar{w}_0$ of the random mating population is always close to w_1 whether the equilibrium is due to heterosis or due to mutation, simply because the great majority of the individuals in the population are of genotype *AA*. After eliminating the heterozygotes from the population, the new mean fitness $\bar{w}_1$ is about 2 per cent lower than $\bar{w}_0$ in both populations. Hence Sanghvi and Li do not think that the two types of populations react to inbreeding very differently. However, if one looks at the load ratio, he will find that it is exactly 2 for population H and 16.7 for population M. Based on this difference, Crow thinks that he can distinguish the two types of equilibria through inbreeding. The artificial nature of this comparison has been discussed by Sanghvi and will not be repeated here. Instead, I will raise an operational question, *viz.*, how could we ever calculate the value of a genetic load, not mentioning the ratio of two loads, without first knowing whether $w_m = w_1$ or $w_m = w_2$?

We should remember that the original problem was to determine the type of equilibrium or, equivalently, to detect whether $w_1 > w_2$ or $w_2 > w_1$ in the population .9996 (*AA* + *Aa*) and .0004 (*aa*) through inbreeding. Now, as exemplified in Table 1, in order to calculate a quantity that is different for the

two models, we must know *a priori* that $w_m = w_2$ in one case and $w_m = w_1$ in another. How could we know this when facing an unknown population like .9996 (*AA* + *Aa*) and .0004 (*aa*) ? There are 3.92 per cent heterozygotes in the population. How could we first isolate them? If we can separate *AA* from *Aa*, their relative fitness could then be studied directly and why we should resort to inbreeding at all? If we know $w_1 > w_2$ or $w_2 > w_1$, then there is no problem to begin with. In one word, *the method of load ratio requires the prior knowledge of the answer itself.* It cannot be applied as a tool of diagnosis to an unknown population. However, Crow (1958) reported that for a certain body of data on consanguineous matings, the load ratio is found to be 17, implying that it is a mutational equilibrium for the locus under consideration. The raw data and the step-by-step arithmetic procedure of arriving at this ratio have not been given and we have no way to tell precisely how values of w_m, $\bar{w}_0$ and $\bar{w}_1$ have been obtained.

The discussions of Li and of Sanghvi were written before Crow supplied an explicit expression for the load ratio. The essential "misunderstanding" on the part of Sanghvi and Li is that they employed w_1 of *AA* in place of w_m of an unknown genotype to measure the genetic load. It appears that this measurement is more realistic, at least in describing the population situation (see last section).

THE MORE SIMILAR, THE MORE DIFFERENT

In addition to the operational difficulty of the load ratio method, there is another property that is very curious, *viz.*, the more similar the two population models are, the greater the difference in their load ratios! To illustrate, let us consider the four pairs of populations in table 2. In the first pair (H_1 and M_1),

TABLE 2. THE RELATIVE FITNESS VALUE (w) AND GENE FREQUENCY (q) OF FOUR PAIRS OF POPULATIONS; H = HETEROTIC EQUILIBRIUM, M = MUTATIONAL EQUILIBRIUM

	q = .100		q = .020		q = .004		q = .001	
	H_1	M_1	H_2	M_2	H_3	M_3	H_4	M_4
w_1	900	1000	980	1000	996	1000	999	1000
w_2	1000	900	1000	980	1000	996	1000	999
w_3	100	100	20	20	4	4	1	1
$\bar{w}_0$	910	973	980.4	998.8	996.016	999.952	999.001	999.997
$\bar{w}_1$	820	910	960.8	980.4	992.032	996.016	998.002	999.001
$\bar{w}_1/\bar{w}_0$	.901	.935	.980	.982	.99600	.99606	.9990	.9990
L_1/L_0*	2	3.33	2	16.7	2	83.3	2	333

*ratio = $(w_2 - \bar{w}_1)/(w_2 - \bar{w}_0)$ for H populations.
ratio = $(w_1 - \bar{w}_1)/(w_1 - \bar{w}_0)$ for M populations.

the genotypes *AA* and *Aa* have a 10 per cent difference in their fitness values. A difference of such magnitude may be detected by direct studies as in the case of sickle cell trait versus homozygous normals. Yet, the load ratio for M_1 is so low (3.33) that it is not too different from the ratio value 2 for H_1. The second pair is the same as that in table 1; the genotypes *AA* and *Aa* have only a 2 per

cent difference in their fitness values but the load ratio of M_2 is 16.7, much higher than the ratio 2 of H_2. In the third pair, the fitness values of *AA* and *Aa* differ by less than 1/2 per cent and are probably indistinguishable by any actual studies, but the load ratios of H_3 and M_3 differ greatly. In the fourth pair, there is very little difference in fitness values between *AA* and *Aa*, but the load ratios for H_4 and M_4 are 2 and 333, respectively. If the method of load ratio is a valid tool for distinguishing the two types of equilibria, one must conclude that it is much easier to distinguish M_4 from H_4 than to distinguish M_3 from H_3, which in turn is much easier than to distinguish M_2 from H_2, and so on. In other words, the smaller the difference between w_1 and w_2 in the two types of populations, the greater the difference between them as measured by load ratios, and the greater the ease with which they can be distinguished through inbreeding. I thought the opposite should be true by any valid criterion of classification. I see very little difference between M_4 and H_4 and I would venture to say that it would be extremely difficult to tell which is in heterotic and which in mutational equilibrium by any method that may be devised.

THE MORE BENEFICIAL, THE MORE HARMFUL

Finally, we may examine the descriptive consequence of defining genetic load in terms of the maximum fitness value (w_m). Consider, as an example, a population of 1000 *AA* individuals; one of them becomes *Aa* through a new mutation. Suppose that the mutation is a favorable one, conferring directly a 2 per cent advantage in fitness on the heterozygote over the original homozygotes in example I and a 25 per cent advantage in example II:

Genotype	Frequency f	Fitness (I) w	Fitness (II) w
AA	.999	100	100
Aa	.001	102	125
aa	0	—	—
mean fitness	$\bar{w}$	100.002	100.025
load	$(w_m - \bar{w})/w_m$	.01959	.19980
gain	$\bar{w} - w_1$	.002	.025

If we use the fitness of genotype *AA* as the standard of comparison, we would say that in case I, the average fitness of the population has *gained* an amount .002 while in case II, the gain is .025 on account of the beneficial mutation. I think this is a reasonable description of the situation. However, if we calculate the genetic load in terms of the highest fitness value, the population will never have any gain no matter how beneficial the mutation is. In fact, the more beneficial the mutation, the greater the genetic load, implying that the population is suffering from a greater amount of genetic elimination and is worse off from an "optimum" genotype. Then it seems that the shortest way to alleviate the situation is not to have more favorable mutants but to kill off the existing one, so that there will be no selection and therefore no genetic load. This and other considerations has led both Sanghvi and Li to doubt if the expression $(w_m - \bar{w})/w_m$ has biological significance or usefulness as a descriptive index.

When considerations are extended beyond one locus, the concept of an "optimum genotype" will be even more unattainable and the average population fitness will play an increasingly important role, as stressed by Dobzhansky and Spassky (1963). The method of Sanghvi and of Li, allowing for both gain and loss and thus achieving a more realistic description of the population changes, is conceptually the same as that allowing for "genetic load" as well as for "genetic elite," as suggested by Dobzhansky and Spassky.

It may be argued (Crow, personal communication) that genetic load is designed to describe an equilibrium population, not to describe the effect of mutants. "That you can contrive a biologically nonsensical result by applying the definition in a circumstance for which it was never intended is no valid criticism of the theory." But there is no essential difference between describing the effects of mutants in a population and describing an equilibrium population. For instance, in a random mating population with $p = .9995$ and $q = .0005$, the genetic composition would be approximately

AA, .999
Aa, .001
aa, negligible

and this is in practice the same as having one mutant in every thousand individuals, and the same conclusion applies.

SUMMARY

The discussions presented in this communication may be summarized thus: Inbreeding results cannot distinguish one type of equilibrium from another. The method of load ratio is operationally backwards, as it requires prior knowledge of the type of equilibrium before calculating the ratio. The load ratio, if calculated, has the very unusual property that the more alike the two types of populations, the greater the difference in their load ratio and hence the greater the ease with which they can be distinguished. I doubt this very much. The definition of genetic load is such that the more beneficial the new mutant, the more harmful it is to the population, and this is hardly an accurate description of the effects of beneficial mutations. In my opinion, we need more, not less, discussion on the entire concept of genetic load and its possible applications so that we (pro and con) may eventually reach a better solution.

REFERENCES

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

Crow, J. F. 1963. The concept of genetic load: a reply. *Amer. J. Hum. Genet.* 15: 310-315.

Dobzhansky, T., and Spassky, B. 1963. Genetics of natural populations XXXIV. Adaptive norm, genetic load, and genetic elite in Drosophila pseudoobscura. *Genetics*, in press.

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.

Sanghvi, L. D. 1963. The concept of genetic load: a critique. *Amer. J. Hum. Genet.* 15: 298-309.