HETEROSIS AS A MAJOR CAUSE OF HETEROZYGOSITY
IN NATURE[1]

ROGER D. MILKMAN

Department of Zoology, Syracuse University, Syracuse, New York and Biological Laboratories, Harvard University, Cambridge, Massachusetts[2] 02138

H UBBY and LEWONTIN (1966; LEWONTIN and HUBBY 1966) have developed and begun to use a method for the basic description of natural genetic variation, in which they employ extremely reliable electrophoresis and enzyme-specific staining to determine the number and frequencies of alleles at particular loci in several *Drosophila pseudoobscura* populations. This work is certain to be a landmark, and it is likely to have a quick and constructive effect on the field of population genetics.

They find a high proportion of allelic polymorphism. In their first reports, 9 of 21 loci have more than one allele. Their discussion of the causality of this genetic variation, however, appears to contain a fundamental and far-reaching error. In view of the importance of this paper, and because of the great attention currently being given the concept of genetic load, I wish to reply to a single conclusion and, I believe, correct it.

LEWONTIN and HUBBY (1966) argue that allelic variation at many loci cannot be due to heterosis (selection against homozygotes) on the following grounds. To remove 10% of the homozygotes at one locus in each generation, one reduces the relative population fitness by 5% (assuming two alleles with equal frequencies, and therefore an initial heterozygote frequency of 0.50) to 0.95. For 2000 loci, this fitness would be $(0.95)^{2000} = 10^{-46}$, an intolerable level. But this is wrong.

The error here is to start with one locus and to consider selection at each locus independently, with individual fitnesses calculated and combined as a product. The locus, however, is not the unit object of selection; the individual is. And artificial selection has shown us nothing, if not the cumulative effects of genes at many loci.

To calculate the relation between selection pressure on individuals and selection pressure on loci, one may use a model in which the population is divided into classes on the basis of the number of heterozygous loci. The size of each class is given by the binomial distribution. Apart from the effects at many loci being cumulative, no assumptions are made that are not implicit in the method of LEWONTIN and HUBBY. All alleles are assumed to have a frequency of 0.5, and linkage is ignored.

Selecting the top 15.87% on the basis of the number of heterozygous loci, one

obtains a sample whose members contain heterozygous loci in numbers exceeding the average for the unselected population by *at least* one standard deviation. The sample average is considerably greater. For example, consider the case of 256 loci (the rest being monomorphic). With two alleles at equal frequency at each locus, 50% of all loci will be heterozygous, and an individual will thus average 128 heterozygous loci. In a binomial distribution of $(p + q)^n$ where $p + q = 1$, $\sigma = (npq)^{1/2}$. So in the top 15.87%, all individuals would have at least $128 + 8$ heterozygous loci, the sample average being greater, of course.

To match the selection pressure at each locus used by LEWONTIN and HUBBY, a 10% reduction in homozygotes would be achieved in the present specific case by selecting the top 13.5%, a practical procedure in Drosophila, whose females produce scores of eggs. By their calculations, however, the population fitness would drop to $(0.95)^{256} = 0.0000023$. This calculation is based on an erroneous concept of the independence of loci in selection; and its highly erroneous result leads to the improper exclusion of heterosis as a major mechanism in the maintenance of heterozygosity in nature.

For a general appraisal of the relationship between number of cumulative factors, selection differential, and selection pressure, LERNER's text (1958) is a good place to begin, and PEARSON (1924) has the necessary table. We wish to know what proportion of a population must be selected (selection pressure) in order to cause the proportion of heterozygotes at each of a number of loci (cumulative factors) to exceed by a certain percentage (selection differential) the proportion of heterozygotes in the unselected population. The selection differential turns out to be equivalent to the average selection pressure at individual loci: in a sample where each individual averages 10% more heterozygous loci, each locus will be heterozygous 10% more frequently.

Selection differential is often given in units of standard deviation. It must be converted to variate units and then can be expressed as a proportion of the unselected mean. Because σ increases as $n^{1/2}$, it is necessary to apply increasingly great selection pressure to the population as n increases, in order to apply a given selection pressure at each locus.

SVED, REED and BODMER (1967) point out that any model relating heterozygosity to fitness must be consistent with observed changes in fitness after inbreeding. These changes are often not great. The present model, a strictly competitive one, involves the survival of the individuals with the greater number of heterozygous loci; but it does not, for example, entail the extinction even of complete homozygotes, if there are no heterozygotes present.

In these two superb papers, the error I cite is a small one; but the subject is important in the genetics of populations.

SUMMARY

According to a simple model, heterosis is acceptable as a major cause of heterozygosity in nature; heterosis at hundreds of loci is consistent with observed reproductive potential and with the rather moderate loss of fitness often observed after inbreeding.—It is generally incorrect to calculate fitness as the product of fractional values assigned to each locus; the individual is the unit object of selection, and loci can contribute cumulatively.

LITERATURE CITED

HUBBY, J. L,. and R. C. LEWONTIN, 1966 A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. Genetics **54**: 577–594.

KING, J. L., 1967 Continuously distributed factors affecting fitness. Genetics **55**: 483–492.

LERNER, I. M., 1958 *The Genetic Basis of Selection*. Wiley, New York. pp. 112–114.

LEWONTIN, R. C., and J. L. HUBBY, 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.

PEARSON, K., 1924 *Tables for Statisticians and Biometricians*. Table II. Biometric Laboratory, University College, London, England.

SVED, J. A., T. E. REED, and W. F. BODMER, 1967 The number of balanced polymorphisms which can be maintained in a natural population. Genetics **55**: 469–481.