MAINTENANCE OF GENETIC VARIATION WITH A FREQUENCY-DEPENDENT SELECTION MODEL AS COMPARED TO THE OVERDOMINANT MODEL

PHILIP W. HEDRICK

Department of Systematics and Ecology, University of Kansas, Lawrence, Kansas 66044

Manuscript received July 10, 1972
Revised copy received August 3, 1972

ABSTRACT

A frequency-dependent selection model proposed by HUANG, SINGH and KOJIMA (1971) was found to be more effective at maintaining genetic variation in a finite population than the overdominant model. The fourth moment parameter of the distribution of unfixed states showed that there was a more platykurtic distribution for the frequency-dependent model. This agreed well with the expected gene frequency change found for an infinite population.

UNTIL the late sixties there was no strong evidence whether only a few genes were segregating for different alleles or a high percentage of loci had naturally occurring variants. In a period of several years, surveys of genes controlling enzyme and protein production by LEWONTIN and HUBBY (1966), SELANDER, HUNT and YANG (1969) and others thoroughly demonstrated that a large number of loci in many different organisms are segregating for variant forms. KOJIMA and his co-workers (KOJIMA and YARBROUGH 1967; HUANG, SINGH and KOJIMA 1971) believe that frequency-dependent selection may be particularly important in maintaining genetic variants at loci exhibiting electrophoretic variation.

For any balanced selection model, the major factor in most situations which causes elimination of genetic variation is that of sampling due to finite population size. In small populations only extremely strong selection can maintain segregation for a long period of time. As a result a comparison of the ability of different balanced selection models to maintain genetic variation in a finite population would give a measure of the effectiveness of these models in maintaining genetic variation in natural populations.

The purpose of this study is (1) to compare the expected gene frequency change in an infinite population using the data and model of HUANG, SINGH and KOJIMA (1971) to that of an overdominant model with the same area under the curve giving the expected gene frequency change and (2) to compare the effectiveness of these two models in maintaining genetic variation in a finite population.

METHODS

Infinite population size: The frequency-dependent model used by HUANG, SINGH and KOJIMA (1971) for a small chromosome segment marked by the esterase-6 locus can be called a marginal average model. In this model the viability of a particular genotype is a function of the frequencies of the three possible genotypes, *FF*, *FS* and *SS*, and nine specific viability estimates. For example, the relative viability of the *i*th genotype is

$$\bar{x}_{ik} = p_k^2 x_{i2} + 2p_k q_k x_{i1} + q_k^2 x_{i0} \tag{1}$$

where $\bar{x}_{ik}$ is the marginal average viability for the *i*th genotype when the frequency of the two alleles, *F* and *S*, are p_k and q_k and the relative viability of the *i*th genotype on media conditioned by genotypes *FF*, *FS* and *SS* is x_{i2}, x_{i1} and x_{i0}, respectively. The expected change in gene frequency for this model is

$$\Delta q_k = \frac{\bar{x}_{1k} p_k q_k + \bar{x}_{0k} q_k^2}{\bar{x}_{2k} p_k^2 + \bar{x}_{1k} 2p_k q_k + \bar{x}_{0k} q_k^2} - q_k \tag{2}$$

Using the relative viability values given in Table 1 of HUANG, SINGH and KOJIMA (1971) the equilibrium frequency of the *F* allele was found to be 0.474.

The expected change in gene frequency for an overdominant model is

$$\Delta q_k = \frac{v_1 p_k q_k + v_0 q_k^2}{v_2 p_k^2 + v_1 2p_k q_k + v_0 q_k^2} - q_k \tag{3}$$

where v_2, v_1 and v_0 are the relative viabilities of genotypes *FF*, *FS* and *SS*. By iteration, values of v_2, v_1 and v_0 were found that gave the same gene frequency equilibrium and had the same area between the Δq curve and the horizontal axis as found for the model and data of HUANG, SINGH and KOJIMA.

Finite population size: The methods used to determine the relative effectiveness of the two models in maintaining genetic variation in a finite population were similar to that of HEDRICK (1970). In this approach a transition matrix where the elements are equal to the probability of transition from *m S* alleles in generation *x* to *i S* alleles in generation $x + 1$ is generated. The gene frequency of the *S* allele after selection, q_k', is equal to the first term on the right side of expressions (2) and (3) for the frequency-dependent and overdominant models, respectively. In order to have a symmetrical distribution the viability values were slightly changed so that the equilibrium gene frequency was 0.5 instead of 0.474. The initial gene frequency was also set at 0.5.

RESULTS

Infinite population size: The gene frequency change for the frequency-dependent model using the viability values of HUANG, SINGH and KOJIMA (1971) and the expected change for the overdominant model with values of 0.88438, 1.0 and 0.89584 for v_2, v_1 and v_0 are given in Figure 1. The gene frequency change for the frequency-dependent model is greater than the overdominant model at extreme gene frequencies. At intermediate gene frequencies the change is greater for the overdominant model.

The slope of the curve at equilibrium is a measure of the rate of return to equilibrium after a slight displacement away from it. Near equilibrium, Δq is nearly linear with respect to q and can be given as

$$\Delta q = b\,(q - q_e) \tag{4}$$

where b is the slope near the equilibrium gene frequency, q_e (CAVALLI-SFORZA

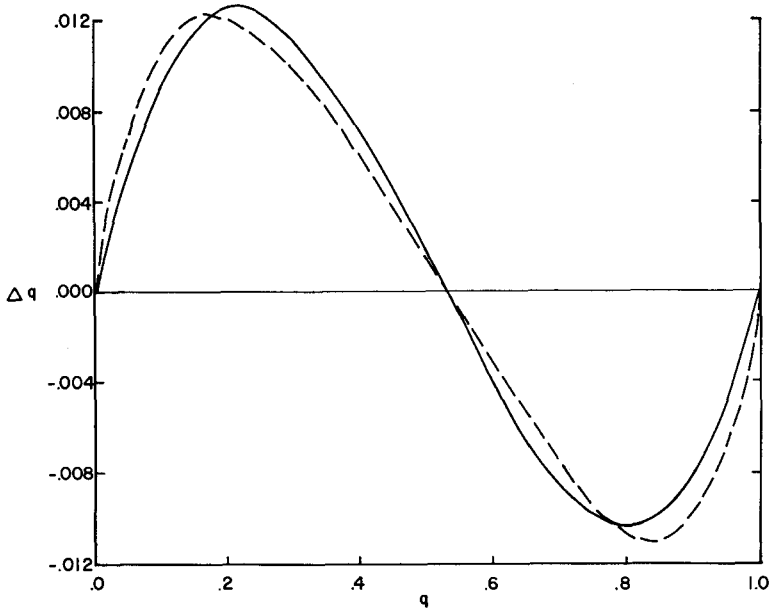

FIGURE 1.—The expected gene frequency change for the frequency-dependent model (broken line) and the overdominant model (solid line).

and BODMER 1971). For the overdominant case the value of b was -0.0577 while for the frequency-dependent model b was -0.0469, about 80% of the value for the overdominant model. In other words, the rate of return to equilibrium with a small displacement is faster for the over-dominant model. But when the displacement is large enough so that the Δq for the frequency-dependent model becomes greater, the initial rate of return for this model also becomes greater.

Finite population size: The effectiveness of the two selection models in maintaining genetic variation can be measured by examining the number of generations it takes until 95% of the populations of a particular size would be fixed for either the F or S allele. The number of generations (interpolated to the nearest tenth and divided by N) for 95% fixation for the two selection models and for no selection are given in Table 1. The time to fixation is very similar for the two

TABLE 1

The number of generations (divided by N*) until 95% of the populations are fixed and the retardation factor (r.f.)*

Population size (N)	No selection		Frequency-dependent selection		Overdominant selection	
	95%	r.f.	95%	r.f.	95%	r.f.
5	6.24	1.0	7.56	1.22	7.32	1.22
10	6.50	1.0	10.15	1.61	9.82	1.58
20	6.64	1.0	18.57	3.00	17.30	2.78
40	6.71	1.0	82.52	13.67	66.16	10.94

selection models at the smallest population sizes. At the largest population size (N=40), the time to 95% fixation is 24.7% longer for the frequency-dependent selection.

In a finite population after a period of time (the length of which is dependent upon the population size, the initial distribution and the selection model), the gene frequency distribution reaches a point of steady decay. After this point is reached all the unfixed gene frequency states decrease each generation by a specific amount, known as the rate of steady decay, that is characteristic of the population size and the selection model. A function of the rate of steady decay called the retardation factor (the reciprocal of $2N$ times the rate of steady decay) was introduced by ROBERTSON (1962) to illustrate the effect of selection models on the rate of fixation.

The retardation factor for the models is also given in Table 1. As expected from the time to fixation data the retardation factor for the frequency-dependent model is larger than for the overdominant model, particularly at the larger population sizes. At the largest population size (N=40), the retardation factor for frequency-dependent selection is 25.0% larger than for the overdominant model, a value quite close to the difference in the time to 95% fixation for this population size. This relationship does not hold with smaller population sizes because with smaller N many populations become fixed before steady decay is reached. At the larger population sizes though, the time to fixation as well as the retardation factor becomes a function of the rate of steady decay.

When the gene frequency distribution reaches a state of steady decay, the moments of the unfixed states of the gene frequency distribution also becomes constant. The variance and the fourth moment parameter (g_2) were calculated and are given in Table 2. The variance of the unfixed states for the two selection models is quite similar although the variance of the frequency-dependent model is somewhat larger for all population sizes. The value of g_2 for the frequency-dependent model is smaller than the overdominant model for all population sizes and the difference becomes more pronounced as the population size becomes larger. This indicates the unfixed distribution for the frequency-dependent model is more platykurtic, i.e., it has a greater proportion of its distribution in the extreme gene frequency states than does the overdominant model.

TABLE 2

The variance of gene frequency (σ^2_q) and the fourth moment parameter (g_2) of the gene frequency distribution for the unfixed states with steady decay

Population size (N)	No selection		Frequency-dependent selection		Overdominant selection	
	σ^2_q	g_2	σ^2_q	g_2	σ^2_q	g_2
5	.0628	—1.177	.0592	—1.116	.0589	—1.109
10	.0715	—1.175	.0615	—1.025	.0609	—1.004
20	.0766	—1.183	.0536	— .833	.0525	— .780
40	.0795	—1.190	.0355	— .445	.0341	— .308

DISCUSSION

The frequency-dependent selection model suggested by HUANG, SINGH and KOJIMA (1971) is of particular biological appeal since it is dependent only on the frequencies of the three genotypes and the estimated viabilities of the genotypes under different conditions. As shown above it also has a capacity to maintain genetic variation that is much greater than overdominant selection. It is also apparent from these results that the retardation factor or other measures of the rate of fixation are more dependent upon the magnitude of gene frequency change at extreme gene frequencies than the change at intermediate gene frequencies.

The frequency-dependent model would be most effective relative to the overdominant model at maintaining variation when there was a large displacement away from the equilibrium gene frequency. This could of course happen due to sampling in a founder population. In the latter case the genetic constitution in later generations would be highly dependent on the ability of the population to quickly return toward the equilibrium from an extreme gene frequency. An example with these conditions ($2N=40$, $q=0.025$) verified this and showed that the frequency-dependent model delayed the time to 95% fixation by 17.3% when compared to the overdominant model.

I would like to express my appreciation to Drs. R. NASSAR and G. SCHLAGER for reading the manuscript. The research was supported by funds from the University of Kansas Research Committee and Computation Center.

LITERATURE CITED

CAVALLI-SFORZA, L. L. and W. F. BODMER, 1971 *The Genetics of Human Populations*. W. H. Freeman and Co., San Francisco.

HEDRICK, P. W., 1970 Selection in finite populations. I. The probability of fixation and rate of response using transition matrix iteration. Genetics **65**: 157–173.

HUANG, S. L., M. SINGH and K. KOJIMA, 1971 A study of frequency-dependent selection observed in the esterase-6 locus of *Drosophila melanogaster* using a conditioned media method. Genetics **68**: 97–104.

KOJIMA, K. and K. M. YARBROUGH, 1967 Frequency-dependent selection at the esterase-6 locus in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.* **57**: 645–649.

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.

SELANDER, R. K., W. G. HUNT and S. Y. YANG, 1969 Protein polymorphism and genic heterozygosity in two European subspecies of the house mouse. Evolution **23**: 379–391.

ROBERTSON, A., 1962 Selection for heterozygotes in small populations. Genetics **47**: 1291–1300.