SIMULATION STUDIES ON ELECTROPHORETICALLY DETECTABLE GENETIC VARIABILITY IN A FINITE POPULATION[1]

TOMOKO OHTA AND MOTOO KIMURA

National Institute of Genetics, Mishima, Japan

Manuscript received May 30, 1973
Revised copy received October 17, 1973

ABSTRACT

Using a new model of isoalleles, extensive Monte Carlo experiments were performed to examine the pattern of allelic distribution in a finite population. In this model it was assumed that the set of allelic states is represented by discrete points on a one-dimensional lattice and that change of state by mutation occurs in such a way that an allele moves either one step in the positive direction or one step in the negative direction on the lattice. Such a model was considered to be appropriate for estimating theoretically the number of electrophoretically detectable alleles within a population. The evenness of allelic distribution was measured by the ratio of the effective to the actual number of alleles (n_e/n_a). The results of the Monte Carlo experiments have shown that this ratio is generally larger under the new model of isoalleles than under the conventional KIMURA-CROW model of neutral isoalleles. In other words, the distribution of allelic frequencies within a population is expected to be more uniform in the new model. By comparing the Monte Carlo results with actual observations, it was concluded that the observed deviation from what is predicted under the new model with selective neutrality is not in the direction of conforming to the overdominance hypothesis but is, in fact, in the opposite direction.

IT is now well known that abundant genetic variability exists in natural populations at the enzyme level. In analyzing theoretically the underlying mechanism for the maintenance of such variability, the "infinite allele model" proposed by KIMURA and CROW (1964) has been used extensively. In this model, every mutant allele that arises at a locus represents a new allelic state not preexisting in the population. Recently, EWENS (1972) and JOHNSON (1972), for example, tried to test selective neutrality of protein polymorphisms by investigating if observed allelic distributions deviate significantly from what is expected from this model; if the observed distribution of allelic frequencies within a population tends to be more uniform than expected, "balancing selection" is suggested. In such an application, the model may be appropriate if the variability can be detected at the level of amino acid site or nucleotide site, since there are so many ways a gene (cistron) can mutate to produce a new allele.

[1] Contribution No. 968 from the National Institute of Genetics, Mishima, Shizuoka-ken, 411 Japan.

However, in the ordinary studies of variability using electrophoretic methods, the basic assumption of the model may not be met completely, because these methods can only detect a difference in the electric charges of molecules as discrete bands on the gels. If an amino acid substitution occurs in the molecule which causes a charge difference the band usually moves one "step" in either the positive or negative direction. For this reason we have recently proposed a new model of allelic mutation in which mutational change is represented as a stepwise movement on a one-dimensional lattice (OHTA and KIMURA 1973). Simple as it might seem, this new model turned out to be mathematically much less tractable than the previous KIMURA-CROW model. However, we have been able to obtain a simple formula for the "effective" number of alleles maintained in a finite population at equilibrium. From this it has been shown that the effective number of alleles for this model is much smaller than expected from the previous model when in both cases the product of mutation rate and the population number is large.

In the present paper we present some results of Monte Carlo experiments showing that the pattern of allelic distribution is also different under the new isoallelic model. Then we shall discuss the bearing of these results on actual observations of the genetic variability at the enzyme level.

The Model of stepwise production of alleles

It may be convenient to express the entire set of allelic states by integers. We assume that if an allele changes its state by mutation, the change occurs in such a way that it moves either one step in the positive direction or one step in the negative direction in the state space. To facilitate both the mathematical treatment and the Monte Carlo experiments, we regard the infinite number of allelic states corresponding to the integers as a limit attained by making the number (n) of possible allelic states indefinitely large in the finite model where allelic states are arranged on a circle (Figure 1). In our previous report we have shown that when n is sufficiently large, the effective number of neutral alleles maintained in a finite population, or the reciprocal of the sum of squares of allelic frequencies, is

$$n_e = \sqrt{1 + 8N_e v}, \tag{1}$$

where N_e is the effective population size and v is the mutation rate per gamete for neutral alleles (OHTA and KIMURA 1973). By comparing this with the corresponding formula, i. e.

$$n_e = 1 + 4N_e v \tag{2}$$

for the model of KIMURA and CROW (1964), we can see that the present model gives a much smaller number of alleles when $N_e v$ is large, although both formulae give similar results when $N_e v$ is small. Also, in one of our previous reports (KIMURA and OHTA 1973) the essential feature of the new model was briefly explained in graphical form in our discussion of the sequential arrangement of rare and common electrophoretic alleles on the gels. In this model, it was assumed that one positive and one negative change in charge cancel each other, leading the allele back to the original state. An actual example of such changes is given by the work of HENNING and YANOFSKY (1963) on the electrophoretic mobility of mutants of the A protein of tryptophan synthetase of *E. coli*: mutant protein A11 moves in the negative direction; mutant A46 moves in the positive direction; but the mobility of the double mutant is identical to that of the wild-type protein.

Recently, BULMER (1971) pointed out that the most common allele at a locus within a species almost always occurs in the middle of the sequence when the multiple alleles are arranged in

order of the electrophoretic mobility. He claims that this observation is against the neutral hypothesis of KIMURA (1968). Both MAYNARD SMITH (1972) and KIMURA and OHTA (1973), however, pointed out that BULMER's observation is compatible with the neutral theory if we consider a stepwise production of alleles. BULMER's observation suggests that this type of mutational scheme should be taken into account in the theoretical analyses of electrophoretically detectable variability.

More recently, KING (1973) considered the possibility that the number of electrophoretically detectable alleles in a large population is controlled by the number of finite variable sites in the molecule rather than by the finite population size. He thinks that the observed pattern of polymorphisms agrees with this hypothesis.

In the present paper we shall investigate the effect of finite population size on the pattern of allelic distribution using the model of stepwise production of alleles.

Monte Carlo Experiments

Since the mathematical form of the frequency distribution of alleles has not been obtained, we resorted to extensive Monte Carlo experiments to find out several statistics relating to the distribution. In particular, the ratio of the effective to the actual (or average) number of alleles is investigated in detail, since this ratio serves as an index to represent evenness of allele distribution.

Each generation of the experiment consisted of mutation and sampling. Mutational changes were carried out deterministically. Namely, if x_i be the frequency of the i-th allele, then the change in x_i by mutation is given by

$$\Delta x_i = \frac{v}{2}(x_{i+1} + x_{i-1} - 2x_i).$$

Since the allelic states are arranged on a circle (Fig. 1), the first and the last states are neighboring states and are connected by mutational change. The total number of allelic states assumed was either 50 or 80. These are sufficient to simulate the model, since the number of segregating alleles in each generation in our experiments turned out to be a small fraction of the total number of states assumed.

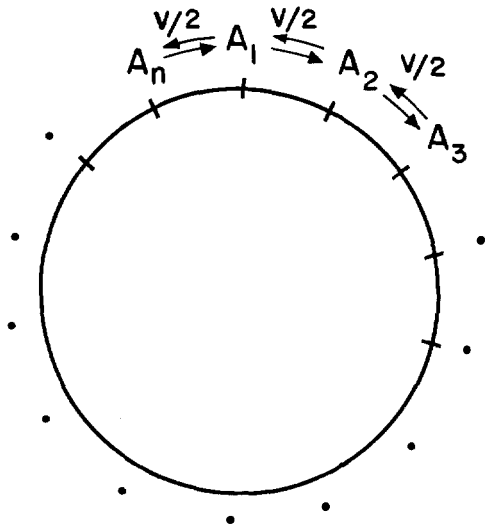

FIGURE 1.—The model of stepwise production of alleles with a finite number of allelic states arranged on a circle.

The sampling of gametes was performed following the scheme used by HILL and ROBERTSON (1966), i. e., one gamete with the i-th allele is sampled if the uniform random number falls in the interval, $\sum_{j=1}^{i} x_j - \sum_{j=1}^{i-1} x_j$. The sampling was repeated $2N_e$ times to produce the gene pool of the next generation. This is essentially a haploid model, but since no selection was practiced, pairing of gametes to produce zygotes was not required in our simulation of the distribution of allelic frequencies within a population.

We made three sets of experiments. In the first set, the average values of parameters such as the effective number (n_e) and the actual number (n_a) of alleles and their ratio (n_e/n_a) were obtained. Each experiment started from a homogeneous population, and the allele numbers were counted from the 101st generation onward until the 2100th generation. Thus, each experimental outcome is the average of 2000 genereations.

Figure 2 illustrates the outcome of this set of experiments and it shows the relationship between the average of the actual number of alleles ($\bar{n}_a$) and the average of the ratio of the effective to the actual number of alleles ($\overline{n_e/n_a}$). As mentioned already, this ratio serves as an indicator for the evenness of the allelic distribution: if this ratio is small, uneven distribution due to the presence of many rare alleles is suggested, whereas if the ratio is large, more uniform distribution is suggested. Balancing selection such as overdominance will tend to make the allelic distribution more uniform. Each point in the figure represents the result of one experimental run. Three levels of population size (50, 100 and 200) were employed, and at each population size, the mutation rate (v) was varied to make 19 values of $N_e v$ (0.1, 0.2, . . . 0.9, 1.0, 2.0, 3.0, . . . , 10.0). So, there are $19 \times 3 = 57$ output points in all. Of these, circles are for $N_e = 50$, squares are for $N_e = 100$ and triangles for $N_e = 200$.

From the figure, it can be seen that the average ratio, $\overline{n_e/n_a}$ mostly stays around 0.6 and is relatively little influenced by the mean actual number of alleles, $\bar{n}_a$, although there is a slight decrease at the neighborhood of $\bar{n}_a = 5.0$. This is in contrast to the conventional isoallele model of KIMURA and CROW, in which $\overline{n_e/n_a}$ is expected to decrease steadily with increasing $\bar{n}_a$ (cf.

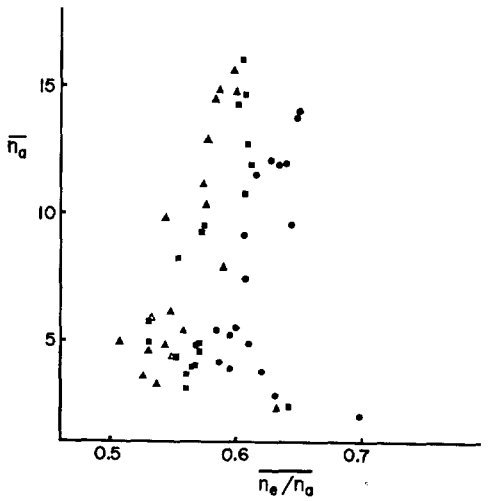

FIGURE 2.—Results of Monte Carlo experiments, showing the relationship between the average of the actual number of alleles ($\bar{n}_a$) and the average of the ratio of the effective to the actual number of alleles ($\overline{n_e/n_a}$). Each point is the average of 2000 generations. Circles represent the results for $N_e = 50$, squares for $N_e = 100$ and triangles for $N_e = 200$. The values of $N_e v$ range from 0.1 to 10.0 for each value of N_e.

KIMURA and CROW 1964; JOHNSON 1972). It may also be noticed that this ratio is larger in the present model than in the model of KIMURA and CROW. In the latter, the average ratio is less than 0.5 if $\overline{n_a}$ is larger than 2.

In the second set of experiments, allelic frequencies are counted starting from the generation 500 at an interval of 500 generations and the results at each counting, rather than the average values, are examined. A larger population size, i. e., $N_e = 10^3$, was assumed and at each counting a small sample of 100 was extracted. The sample statistics were calculated and compared with the population parameters. Four levels of mutation rate ($v = 10^{-3}$, 5×10^{-4}, 10^{-4} and 5×10^{-5}) were used.

Each experiment, starting from a homogeneous population, was continued until the generation 10,000. Then, with the same mutation rate, the experiment was repeated, giving the total of 40 outputs.

Figure 3 shows the results of this set of experiments, illustrating directly the relationship between n_e/n_a and n_a. Each point represents an individual outcome, circles for $v = 10^{-3}$, squares for $v = 5 \times 10^{-4}$, triangles for $v = 10^{-4}$ and crosses for $v = 5 \times 10^{-5}$. In the figure, solid (black) circles etc. represent those for the population, with open (white) circles etc. for the sample. For

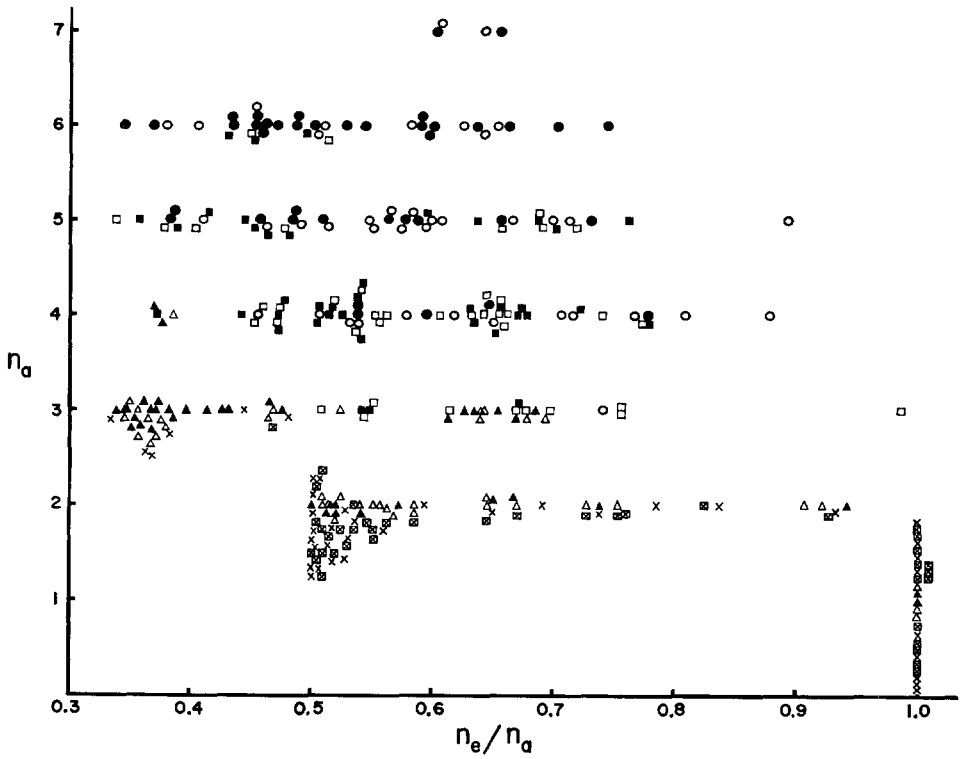

FIGURE 3.—Results of Monte Carlo experiments, showing the relationship between individual observations of n_a and n_e/n_a. The population size was assumed to be 10^3, from which a sample of the size 10^2 was extracted at the interval of 500 generations. Four levels of mutation rate ($v = 10^{-3}$, 5×10^{-4}, 10^{-4}, and 5×10^{-5}) were used. Black (solid) circles represent results for the population and white (open) circles are those for the sample both for $v = 10^{-3}$, black squares for the population and white squares for the sample both for $v = 5 \times 10^{-4}$, black triangles for the population and white triangles for the sample both for $v = 10^{-4}$, and crosses for the population and crossed squares for the sample both for $v = 5 \times 10^{-5}$.

TABLE 1

*Summary of the results of the Monte Carlo experiment shown in Figure 3**

Population:				
v	10^{-3}	5×10^{-4}	10^{-4}	5×10^{-5}
n_e	2.999 ± 0.637	2.394 ± 0.480	1.301 ± 0.326	1.129 ± 0.210
Theoretical n_e	3.000	2.236	1.342	1.183
n_a	5.525 ± 0.784	4.375 ± 0.740	2.675 ± 0.648	1.925 ± 0.608
n_e/n_a	0.545 ± 0.106	0.552 ± 0.112	0.516 ± 0.177	0.647 ± 0.223
$\bar{n}_e/\bar{n}_a$	0.543	0.547	0.486	0.586
Sample:				
n_e	2.954 ± 0.608	2.380 ± 0.474	1.307 ± 0.330	1.128 ± 0.207
n_a	4.975 ± 0.920	4.050 ± 0.815	2.400 ± 0.663	1.700 ± 0.510
n_e/n_a	0.601 ± 0.120	0.598 ± 0.127	0.581 ± 0.185	0.720 ± 0.215
$\bar{n}_e/\bar{n}_a$	0.594	0.588	0.545	0.664

* The averages and the standard errors of n_e, n_a and n_e/n_a are shown.

the case of $v = 5 \times 10^{-5}$, crosses represent population values and crossed squares the sample values. As expected, the ratio n_e/n_a is larger in the sample than in the population, since the population may contain rare variant alleles missed by a sample. Also there is negative correlation between n_a and n_e/n_a for a given mutation rate. In other words, if the mutation rate is kept constant, the ratio is negatively correlated with n_a. However, as seen from Figure 2, if the outputs from a wide range of mutation rates are plotted together (simulating the cases when observations from different gene loci are mixed) and the average values compared, little correlation appears between these two.

Tables 1 and 2 present the results embodied in Figure 3 in more detail. The average and the standard error of n_e, n_a and n_e/n_a are listed in Table 1. All the figures are the averages of 40 observations from simulation experiments. Theoretical values of n_e (same for the population and the sample) are given for comparison and they agree quite well with experimental results. The average ratio, $\overline{n_e/n_a}$ is about 0.6 in the sample, whereas it is about 0.55 in the population. In Table 2 the observed ratio, n_e/n_a (given by the mean and the standard error), is listed for each value of n_a, ranging from 1 to 7. The number in parenthesis is the number of observations. From the table, it can be seen that the average ratio decreases with increasing n_a for each set of parameters (except for $n_a = 7$, for which we have only four observations and therefore subject to large error). In other words, the ratio gets smaller when more alleles happen to be segregating merely by chance. For example, when $n_a = 3$, the average ratio $(\overline{n_e/n_a})$ in the sample is 0.473 for $N_e v = 0.1$ but it is 0.677 for $N_e v = 0.5$. Note here that the actual allele number of three ($n_a = 3$) represents the number above average for $N_e v = 0.1$, but below average for $N_e v = 0.5$. The range of n_a for a given mutation rate should also be noticed: we have, as n_a values for the population, $4 \sim 7$, $3 \sim 6$, $1 \sim 4$ and $1 \sim 3$ corresponding to $N_e v = 1.0$, 0.5, 0.1 and 0.05, respectively. These ranges are narrower than those observed under the KIMURA-CROW model (cf. KIMURA 1968).

The third set of experiments was designed to examine the effects of population size while keeping the product $N_e v$ constant: Namely, assuming $N_e v = 0.5$, two levels of N_e (200 and 2000) were chosen to find out the effect of change in population size. Population parameters such as n_e, n_a or n_e/n_a were examined at the interval of 500 generations for the case of $N_e = 2000$ and of 50 generations for $N_e = 200$. Simultaneously, a sample of size 100 was extracted to calculate sample statistics. The experiments were run for 20,000 generations with $N_e = 2{,}000$, and for 2,000 generations with $N_e = 200$. Table 3 shows the averages (with standard errors) of n_e, n_a and n_e/n_a of these experiments. In both cases, the average ratio $(\overline{n_e/n_a})$ for the sample turned out to be about 0.56 and no significant effect of population size was detected. This ratio should also

TABLE 2

*The mean and the standard error of n_e/n_a for each class of n_a ranging from 1 to 7 for the experimental results shown in Figure 3**

n_a	$v = 10^{-3}$		$v = 5 \times 10^{-4}$		$v = 10^{-4}$		$v = 5 \times 10^{-5}$	
	Population	Sample	Population	Sample	Population	Sample	Population	Sample
1	....	....	....	....	1.0 (2)	1.0 (3)	1.0 (9)	1.0 (13)
2	....	....	....	....	0.608 ± 0.129 (11)	0.622 ± 0.122 (19)	0.580 ± 0.120 (25)	0.590 ± 0.114 (26)
3	....	0.743 (1)	0.588 ± 0.073 (3)	0.677 ± 0.140 (10)	0.449 ± 0.118 (25)	0.473 ± 0.130 (17)	0.397 ± 0.051 (6)	0.470 (1)
4	0.621 ± 0.100 (5)	0.648 ± 0.131 (12)	0.570 ± 0.103 (22)	0.591 ± 0.092 (20)	0.374 ± 0.004 (2)	0.387 (1)	....	....
5	0.530 ± 0.108 (11)	0.593 ± 0.114 (16)	0.533 ± 0.138 (12)	0.545 ± 0.161 (8)	....	....	....	....
6	0.527 ± 0.105 (22)	0.530 ± 0.103 (9)	0.460 ± 0.033 (3)	0.482 ± 0.045 (2)	....	....	....	....
7	0.632 ± 0.037 (2)	0.628 ± 0.025 (2)	....	....	....	....	....	....

* The number in parenthesis is the number of observations.

TABLE 3

*Results of experiments showing the effect of population size**

	$N_e = 2000$		$N_e = 200$	
	Population	Sample	Population	Sample
n_e	1.943±0.320	1.930±0.301	2.432±0.644	2.424±0.673
n_a	3.950±0.669	3.500±0.592	4.400±0.970	4.325±0.877
n_e/n_a	0.496±0.065	0.560±0.092	0.560±0.119	0.563±0.120

* The averages and the standard errors are based on 40 outputs made every $N_e/4$ generations continued total of $10N_e$ generations.

TABLE 4

*Pattern of allelic distribution in the third sets of experiments**

Allelic state	Generation $2N_e$	$4N_e$	$6N_e$	$8N_e$	$10N_e$
$N_e = 200$					
18	..	..	..	..	..
19	..	..	..	..	..
20	127	41	23	..	..
21	134	77	..	4	..
22	135	94	13	109	23
23	4	187	331	269	207
24	..	1	33	18	123
25	..	..	..	..	40
26	..	..	..	..	7
$N_e = 2000$					
21	108	4	..	..	..
22	2814	1469	237	116	..
23	980	2254	2775	1447	97
24	98	273	988	2398	3235
25	..	..	..	39	668

* The experiments started from a single genotype (21st allele).

be compared with the results of previous sets of experiments. In all cases the average ratio falls in the range $0.5 \sim 0.6$ for $N_e v = 0.2 \sim 1.0$, with the exception of the case $N_e = 50$.

Table 4 gives some examples from the third set of experiments to show the typical pattern of arrangement of the alleles. In almost all cases, the most common alleles are clustered in the middle of the sequence flanked by less common alleles, and this pattern is identical with that of the observed protein polymorphisms (BULMER 1971).

DISCUSSION

One important conclusion drawn from the present study is that the ratio n_e/n_a is larger under the new model of stepwise production of alleles than under the conventional KIMURA-CROW model when many alleles are segregating simultaneously within a population. In other words, the allelic frequencies tend to be more evenly distributed in the new model. Also, the typical pattern of distribution on the electrophoretic gels is such that the common alleles are clustered in

the middle of the sequence. The observed values of $\overline{n_e/n_a}$ lie somewhere around $0.5 \sim 0.6$. This pattern is identical with that predicted by KING (1973), who assumes that the number of alleles is controlled by the number of finite acceptable sites within a cistron. He lists the equilibrium frequency distribution with various numbers of acceptable sites (Table 1 of KING 1973). The ratio n_e/n_a, assuming the sample size of 50, turned out to be mostly about 0.55 from his table, which is very close to our result. Thus, a very similar pattern of allelic distribution within a population is expected, whether the restriction is due to finiteness of variable sites in a cistron, or is due to finite population size with stepwise production of alleles with infinite possible allelic states.

Compared with the present (step allele) model, the conventional model of KIMURA and CROW differs in three respects: first, in the K-C model, common alleles are not necessarily in the middle of the sequence on electrophoretic gels. This is against BULMER's observation. Secondly, the ratio n_e/n_a is smaller, because a new mutant is produced from every member of allelic series currently existing in the population, resulting in many rare alleles. Thus, the average ratio is less than 0.5 if $n_a \geq 3$ (KIMURA and CROW 1964; JOHNSON 1972).

Finally, the variation of the allele number (n_e or n_a) for a given mutation rate is less in the present model than in the KIMURA-CROW model. For example, n_a varied from 3 to 6 for $N_e v = 0.5$ in our experiments, whereas it varied from 2 to 10 for $N_e v = 0.25$ in KIMURA's experiment, which used the KIMURA-CROW model (KIMURA 1970). Generally, in the present model the standard deviation of n_a is likely to be less than 1 as shown in Table 1, but it is not so in the KIMURA-CROW model. In fact, it is noted from Table 2 that n_a seldom becomes more than 3 when $N_e v \leq 0.1$. Thus, under the step-allele model, if many alleles (e.g., $n_a \geq 4$) are observed in a sample of about 100, it is unlikely that $N_e v$ is small (e.g., $N_e v \leq 0.1$).

Do the observed data agree with the new model? Unexpectedly, as shown by KIRBY and HALLIDAY (1973) and by YAMAZAKI and MARUYAMA (1973), as far as the ratio (n_e/n_a) is concerned, the observed values from enzyme polymorphisms seem to agree well with the conventional KIMURA-CROW model. These authors have made an extensive compilation and statistical analysis of published data on enzyme polymorphisms. They have shown that the average ratio $(\overline{n_e/n_a})$ stays around $0.4 \sim 0.5$ if $n_a \geq 3$. In other words, there exist more rare alleles than would be expected from our new model, at least for the cases of $n_a \geq 4$. It is possible that none of the above models is satisfactory to describe the real situation. Another possibility is that the observed polymorphisms are not equilibrium. In this event some of the rare alleles may be on their way to an increase in their frequency. Still another possibility is that some of the rare alleles are not selectively neutral, but are very slightly disadvantageous and therefore no increase is possible. In fact O'BRIEN, WALLACE and MACINTYRE (1972) reported that the "null" allele at the α-glycerophosphate dehydrogenase-1 locus of *Drosophila melanogaster* is deleterious.

We would like to emphasize that the observed deviation from the expectation based on selective neutrality, if it has a real meaning, is to the direction of

decreasing the ratio, n_e/n_a. Evidently very slightly deleterious mutations whose selection coefficient is comparable with the mutation rate will decrease this ratio. In contrast with this type of selection, overdominance and other balancing selections will tend to increase the ratio. This is because these alleles spend longer time in the middle range of their frequencies rather than in the extreme frequencies, unless the equilibrium point itself lies in the extremities. Thus, unless one assumes overdominance with extreme equilibrium frequencies, as a general phenomenon, the observed deviation is in the opposite direction to what is expected from the overdominance hypothesis. We would like to add that this inference applies to the average behavior of many polymorphisms; the presence of overdominance or other balancing selection is not excluded in particular cases. Also, "associative overdominance" due to linked selected loci is apparently not effective in increasing the ratio. As pointed out by OHTA (1973), it probably works only as an inertia against sudden local perturbation of gene frequencies, although it can not create any directed pressure.

LITERATURE CITED

BULMER, M. G., 1971 Protein polymorphism. Nature **234**: 410–411.

EWENS, W. J., 1972 The sampling theory of selectively neutral alleles. Theoret. Pop. Biol. **3**: 87–112.

HENNING, U. and C. YANOFSKY, 1963 An electrophoretic study of mutationally altered A proteins of the tryptophan synthetase of *Escherichia coli*. J. Mol. Biol. **6**: 16–21.

HILL, W. G. and A. ROBERTSON, 1966 The effect of linkage on limits to artificial selection. Genet. Res. **8**: 269–294.

JOHNSON, G. B., 1972 Evidence that enzyme polymorphisms are not selectively neutral. Nature New Biology **237**: 170–171.

KIMURA, M., 1968 Evolutionary rate at the molecular level. Nature **217**: 624–626. ——, 1970 Stochastic processes in population genetics, with special reference to distribution of gene frequencies and probability of gene fixation. pp. 178–209. In: *Mathematical Topics in Population Genetics*. Edited by K. KOJIMA. Springer-Verlag, Berlin.

KIMURA, M. and J. F. CROW, 1964 The number of alleles that can be maintained in a finite population. Genetics **49**: 725–738.

KIMURA, M. and T. OHTA, 1973 Mutation and evolution at the molecular level. Genetics **73**: s19–35.

KING, J. L., 1973 Isoallele frequencies in very large populations. Genetics (In press.)

KIRBY, G. C. and R. B. HALLIDAY, 1973 Another view of neutral alleles in natural populations. Nature **241**: 463–464.

MAYNARD SMITH, J., 1972 Protein polymorphism. Nature New Biology **237**: 31.

O'BRIEN, S. J., B. WALLACE and R. J. MACINTYRE, 1972 The α-glycerophosphate cycle in *Drosophila melanogaster*. III. Relative viability of "null" mutants at the α-glycerophosphate dehydrogenase-1 locus. Am. Naturalist **106**: 767–771.

OHTA, T., 1973 Effect of linkage on behavior of mutant genes in finite populations. Theoret. Pop. Biol. **4**: 145–162.

OHTA, T. and M. KIMURA, 1973 A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genet. Res. **22**: 201–204.

YAMAZAKI, T. and T. MARUYAMA, 1973 Evidence that enzyme polymorphisms are selectively neutral. Nature New Biology **245**: 140–141.

Corresponding editor: R. ALLARD