3

Extent of Protein Polymorphism and the Neutral Mutation Theory

MASATOSHI NEI AND DAN GRAUR

Center for Demographic and Population Genetics
University of Texas at Houston
Houston, Texas 77225

INTRODUCTION

In recent years Nei and his colleagues (Nei *et al.*, 1976*b*, 1978; Fuerst *et al.*, 1977; Chakraborty *et al.*, 1978, 1980; Nei, 1980*a*; Maruyama and Nei, 1981) have conducted a variety of statistical tests concerning the applicability of the neutral mutation hypothesis (Kimura, 1968*a*) to data on protein polymorphism. In these tests the relationships among such quantities as heterozygosity, allele frequency distribution, genetic distance, and subunit molecular weight of protein were evaluated for various groups of organisms and the agreements between data and the predictions from the neutral mutation hypothesis (often abbreviated as the "neutral theory") were examined. So far they have not been able to reject the "null" hypothesis of neutral mutations.

The statistical tests employed by them may be called indirect tests, since they were conducted without knowing the effective population size N_e and mutation rate v, which are the key parameters determining the level of protein polymorphism (heterozygosity or gene diversity). If we know N_e and v, however, we can make a direct test of the agreement between the observed and expected levels of protein polymorphism.

Under the "null" hypothesis of neutral mutations v can be estimated from the rate of gene substitution α, since α is equal to v for neutral mutations (Kimura and Ohta, 1971). Estimation of N_e is generally more difficult than that of v. Particularly when one wants to compute the expected gene diversity, one has to know the long-term effective size for the last thousand or million years (Nei *et al.*, 1975). This difficulty has kept investigators from using this direct test despite its obvious advantage that the consistency of the level of protein polymorphism and the rate of gene substitution in long-term evolution can be tested by this method.

In the past 15 years, however, the average gene diversity for protein loci has been studied for more than 400 species [for reviews see Powell (1975), Fuerst *et al.* (1977), Nevo (1978), and Hamrick *et al.* (1979)], and in several of these species information on current population size is available. Furthermore, if we make certain assumptions, it is possible to obtain rough estimates of population size for many other species. Of course, the current population size can be drastically different from the long-term effective size, but, as will be explained later, the long-term effective size is generally smaller than the current population size. Therefore, we can still test the null hypothesis of neutral mutations by examining whether the observed gene diversity is lower than the expected value or not. We note that if overdominant selection or other similar types of diversity-enhancing selection are important, the observed level of gene diversity can be much higher than the neutral expectation (Kimura and Crow, 1964; Nei, 1980*a*; Maruyama and Nei, 1981). Furthermore, if the level of gene diversity is determined primarily by ecological conditions rather than by the mutation rate and population size, there may be no correlation between gene diversity and population size. With this in mind we have studied the relationship between gene diversity for protein loci and population size for various groups of organisms.

The purpose of this chapter is twofold. First, we report the results of the above study and examine their compatibility with the neutral theory. Second, we discuss the applicability of various alternative hypotheses to protein polymorphism data. We believe this type of discussion is appropriate because no comprehensive reviews on this subject have recently been published. In this discussion we shall take into account the results obtained by previous authors as well as ours (both theoretical and experimental) and present our view on the maintenance of protein polymorphism. In this chapter emphasis will be given to the consistent explanation of data on protein polymorphism and those on amino acid substitution in long-term evolution. If a theory cannot explain these two aspects of molecular evolution, it is not a viable theory.

GENE DIVERSITY AND POPULATION SIZE

Hypotheses To Be Tested

Following Nei (1975), we define the gene diversity for a locus by $h = 1 - \sum x_i^2$, where x_i is the frequency of the ith allele and $\sum$ stands for the summation over all alleles at the locus. Average gene diversity H is simply the average of h over all loci examined. Gene diversity as defined above is a measure of genetic variation and equal to the heterozygosity in a randomly mating diploid population. In inbreeding populations or haploid organisms, however, it does not measure the frequency of heterozygotes. In this chapter we use H, since this has many favorable properties from the theoretical point of view (Nei, 1975).

Kimura and Crow (1964) studied the so-called infinite-allele model of neutral mutations and derived a formula for the expected value of H. It is given by

$$H = 4N_e v/(1 + 4N_e v) \tag{1}$$

Ohta and Kimura (1973) argued that the infinite-allele model is not appropriate for electrophoretic data, since there is a possibility of back mutations for electrophoretically detectable alleles, and presented an alternative model, the stepwise mutation model. In this model the expected gene diversity is given by

$$H = 1 - (1 + 8N_e v)^{-1/2} \tag{2}$$

Recent experimental data (Ramshaw *et al.*, 1979; Fuerst and Ferrell, 1980; McCommas, 1983), however, suggest that formula (1) is more appropriate than (2) even for electrophoretic data, so that we shall use (1) in this study. Of course, when H is small, as is generally the case, there is not much difference between the values given by (1) and (2).

In the neutral theory the mutation rate is assumed to be constant per year rather than per generation (Kimura and Ohta, 1971; Nei, 1975). (The rate of *deleterious* mutations is known to be approximately constant per generation. For the possible reason for the difference between neutral mutations and deleterious mutations, see Nei (1975, pp. 31–34).] Since we are testing the "null" hypothesis of neutral mutations, we assume that this is the case. The mutation rate per generation is then give by

$$v = kv, \tag{3}$$

where g is the generation time measured in years and v_s is the mutation rate per year. Considering the rate of amino acid substitution, the molecular size of proteins that are used for electrophoresis, and the detectability of protein variation by electrophoresis, Kimura and Ohta (1971) and Nei (1975) have estimated that the rate of neutral mutations that are detectable by electrophoresis is approximately 10^{-7} per locus per year. This estimate is, of course, very crude, but since no other estimate exists, we shall use this as a first approximation. In this connection, we note that estimates of mutation rate for protein loci obtained by Mukai and Cockerham (1977), Nei (1977), Neel and Rothman (1978), Bhatia *et al.* (1979), and Voelker *et al.* (1980) refer to the total rate of mutations, including deleterious ones. Since deleterious mutations are disregarded in neutral theory, these estimates cannot be used for our purpose.

In the present chapter we are interested in testing the following two "null" hypotheses:

1. Mean gene diversity increases as $N_e g$ increases.
2. The observed values of mean gene diversity are equal to or smaller than the expected values given by formula (1).

The rationale for the first hypothesis is obvious from Eq. (1). In most species so far studied the observed mean gene diversity for electromorphs is equal to or smaller than 0.3 when 20 or more loci are examined (Fuerst *et al.*, 1977) (see also Fig. 4), and in the range of $H = 0$–0.3, H is roughly linearly related to $N_e v \equiv N_e g v_s$. Since v_s is constant, H is expected to increase as $N_e g$ increases. This test does not require knowledge of v_s. In practice, of course, v_s may not be the same for all data sets even if the neutral theory is correct. This is because in different organisms different sets of protein loci have often been studied and the mutation rate (or the rate of amino acid substitution) is known to vary from locus to locus (Nei *et al.*, 1976*a*; Zouros, 1979). This would reduce the correlation between H and $N_e g$ to some extent. As mentioned earlier, it is difficult to know N_e in natural populations, so we shall use the actual population size N instead of N_e, assuming that there is a high correlation between N and N_e. Rejection of this "null" hypothesis would argue in favor of some ecological theories mentioned earlier (also the theory of advantageous mutation with genic selection, as will be discussed later), but would not distinguish among many population genetics theories, because in these theories average gene diversity is expected to increase with increasing N.

Our second hypothesis is based on the expectation that the long-term effective population size is generally equal to or lower than the current population size. The reasons are threefold. First, when population size fluctuates over evolutionary time, the long-term effective population size

is close to the smallest size (Wright, 1938), and the current size may not be the smallest. Second, in the last 1 million years there have been several glaciations in many different parts of the world, and the last one (Würm–Wisconsin period) ended only about 10,000 years ago. In these glaciation periods the population sizes of many different species apparently declined drastically compared with their current sizes. It is believed that in the Würm–Wisconsin period a large proportion of mammalian species became extinct (Martin and Wright, 1967), whereas many new species have appeared during the last 1 million years (Rensch, 1959; Flint, 1976). Since the effect of bottlenecks on genetic variability is expected to last for more than 1 million years (Nei *et al.*, 1975), many extant species of animals and plants would have a lower gene diversity than the value expected from current population size. In other words, the long-term effective size for these species is expected to be generally smaller than their current population size. Third, the effective population size is generally smaller than the actual size because of overlapping generations and non-Poisson distributions of progeny size, the ratio of the former to the latter varying considerably from species to species. In man this ratio is probably 0.3–0.4, whereas in insects it could be about 0.1 or even less (Malpica and Briscoe, 1981; Nei and Tajima, 1981). At any rate, for the above reasons the observed gene diversity is expected to be usually lower than the value given by $H = 4N_e v/(1 + 4N_e v)$, if the neutral mutation hypothesis is correct. On the other hand, if diversity-enhancing selection such as overdominance is operating, gene diversity is expected to be much higher than the neutral level (Kimura and Crow, 1964; Nei, 1980*a*; Maruyama and Nei, 1981). Therefore, if we can reject our "null" hypotheses, the neutral theory will have a serious problem.

In this connection it should be noted that the effective population size that is important for the neutral theory is not the size for a local population but that for the entire species, unless the species is divided into subpopulations between which virtually no gene migration occurs (Kimura and Maruyama, 1971). This is because even a small rate of gene migration is sufficient to prevent the genetic differentiation of subpopulations for neutral alleles. In a study similar to ours, Soulé (1976) apparently did not pay much attention to this distinction. For example, he assumed that the effective population sizes of *Mus musculus*, *Drosophila melanogaster*, and *Uta stansburiana* are 10^3–10^4, 10^7, and 10^2, respectively (M. Soulé, personal communication). We believe that these conjectures represent local effective sizes, and thus they are not appropriate. Soulé also used a mutation rate of 10^{-5}–10^{-6} per generation irrespective of the generation time of the organisms used. These two factors caused a downward bias of the expected gene diversity in some organisms and

led him to conclude that the observed gene diversity is higher than the expected value in a substantial number of organisms. Clearly, we need more careful work.

Data Used for Statistical Study

Gene Diversity (Heterozygosity)

In this study we used only those species in which gene frequencies obtained from 10 or more genomes were available for 20 or more protein loci. We used this criterion because of the large interlocus variation of gene diversity (Nei and Roychoudhury, 1974). The mean gene diversity (heterozygosity) was estimated by the formula described earlier unless the number of individuals examined was less than 15. When the number was less than 15, we used Nei's (1978) formula for small sample size. We did not use the observed proportion of heterozygous loci per individual, because this quantity is affected by such factors as mating structure and inbreeding and is thus not very useful for comparing the genetic variabilities of different species (Nei, 1975).

Population Size and Generation Time

Estimates of population sizes were obtained for 77 species for which the estimates of heterozygosity based on 20 or more loci were available (Table I). In some species estimates of population size were already available in the literature and in this case we used them directly. The species for which the population size was available were *Homo sapiens*, *Pan troglodytes* (chimpanzee; Reynolds, 1967), *Gorilla gorilla* (Dixon, 1981), *Mirounga angustirostris* (elephant seal; Boonell and Selander, 1974), *Cervus canadensis* (Canadian elk; Cameron and Vyse, 1978; D. G. Cameron, personal communication), *Alligator mississippiensis* (American alligator; Chartside *et al.*, 1977; McNease and Joanen, 1978), *Pagophilus groenlandicus* (harp seal; Barchard, 1978; Lavigne *et al.*, 1982; Ronald and Dougan, 1982), *Oryctolagus cuniculus* (wild rabbit; B. J. Richardson, personal communication), *Oncorhynchus nerka* and *Salmo salar* (salmons; MacDonald and Smith, 1980; Johansson, 1981), two species of *Macaca* (macaque; Nozawa *et al.*, 1975, and personal communication), *Sphincterochila aharonii* (land snail; E. Nevo, personal communication), four species of *Spalax* (mole rats; Nevo *et al.*, 1982), *Actinonyx jubatus* (South African cheetah; O'Brien *et al.*, 1983), and *Pinus torreyana* (Torey pine; Ledig and Conkle, 1983). In some of these species the population size was estimated by direct counting, whereas in the others it was estimated by multiplying the average population density by the geographical distribution. In the case of *Macaca fascicularis* the population size was estimated by multiplying the average effective population size of a troop by the total number of troops. Elephant seals are known to have experienced a severe bottleneck (N = 20) around 1890, so that we used this size rather than the current size (N = 30,000).

In many species, however, direct estimates of population sizes were not available, so that we had to estimate the sizes by using various methods. In the case of game species we used hunting records. For example, in the Scandinavian moose (*Alces alces*) we have reliable hunting records for the last several decades (Ryman *et al.*, 1977, and reference therein) and it is known that the proportion of hunted animals never exceeds about one-third of the entire population. Consequently, we can obtain a rough minimum estimate of the population size for this species. The same method was applied to the red fox (*Vulpes vulpes*), the stoat (*Mustela erminzeal*, the pole cat (*Mustela putorius*), the badger (*Meles meles*), and the beech marten (*Martes foina*) (Simonsen, 1982) as well as to *Macaca fascicularis* (Matsubayashi and Sajuthi, 1981). In many of the above cases the estimates of population size could also be obtained from information on population densities and geographic distributions (Serafinski, 1969; Southwick and Cadigan, 1972; Morow, 1975). The estimates obtained in this way were roughly in agreement with those obtained from hunting records.

In *Peromyscus* (deer mice or white-footed mice) the population density has been estimated to be 60–8000 per km^2 (Lawlor, 1965; Sheppe, 1966; Douglas, 1969; J. A. King, 1968; Avise *et al.*, 1979). Because we did not want to make it difficult to reject our second "null" hypothesis, we used the minimum estimate of 60 individuals per km^2 except in *P. maniculatus* and *P. polionotus*, where the minimum density is about 600 individuals per km^2 (Davenport, 1964; Redfield *et al.*, 1977). We used the same density (60 per km^2) for *Sigmodon* and *Eutamias*, since these organisms have a habitat similar to that of *Peromyscus*. The areas of the geographic distributions of these rodent species were obtained from Hall and Kelson's (1959) book and other sources [for references see Avise *et al.* (1974*a*,*b*; 1979)]. Using this information, we estimated the population sizes. The population size of *Ochotona princeps* (pikas) was obtained by multiplying the density [500–2500 per km^2 (Broadbooks, 1965)] by the distribution area (Hall and Kelson, 1959). This species inhabits "terrestrial islands," and the population size is directly related to the size of these islands (Glover *et al.*, 1977). The population sizes of the species of *Geomys* and *Thomomys* were estimated by using information on the pop-

ulation densities provided by Nevo (1979, and personal communication) and their geographic distributions (Hall and Kelson, 1959; Russell, 1968; Selander *et al.*, 1974). Similarly, the population size of *Odocoileus virginianus* (white-tailed deer) was estimated by multiplying the average population density [8 per km^2 (Dapson *et al.*, 1979)] by the range of geographic distribution (Hall and Kelson, 1959). The generation times for mammalian species were estimated from information on longevity and reproductive biology (Hall and Kelson, 1959; Altman and Dittmer, 1962; Wood, 1982).

The species of *Anolis* lizards used in this study inhabit the Caribbean Islands, mostly the Lesser Antilles. The population densities of *A. carolinensis*, *A. sagrei*, *A. cristatellus*, *A. distichus*, and *A. oculatus* have been estimated to be 1.8×10^4, 7.7×10^4, 1.45×10^4, 9×10^3, and 4.7×10^2 per km^2, respectively (Andrews, 1979). There are no direct estimates of population densities for the other *Anolis* species used here, and we used 6×10^3 per km^2 for them. This density is for *A. limifrons* and is the smallest density known among the *Anolis* species (Fleming, cited by Andrews, 1979). The geographic distributions and generation times of the species used were obtained from Ruibal and Philobosian (1970), Lazell (1972), Soule and Yang (1973), Yang *et al.* (1974), Gorman and Harwood (1977), and Bennett and Gorman (1979). The population density for *Taricha rivularis* (red-bellied newt) has been estimated to be 20,000 per km stream (Hedgecock, 1978, and personal communication). The areas of the geographic distribution for this species were obtained from Twitty (1955, 1966), Twitty *et al.* (1967), and Stebbins (1962).

The geographic distribution for the two species of darters used in this study (*Etheostoma spectabile* and *E. caeruleum*) were taken from Pflieger (1975). We used density estimates (10^4 per km stream) provided by L. M. Page (personal communication).

Estimation of population size is much more difficult in *Drosophila* than in either mammalian or reptilian species, and only in a limited number of species could we obtain a rough estimate. According to Steiner (1979*a*,*b*), *D. engyochracea* and *D. mimica* are largely confined to two locations of the island of Hawaii (10,433 km^2), i.e., Kipuka Ki (575,000 m^2) and Kipuka Puaulu (297,000 m^2), though a small number of specimens of these species have been observed at distantly removed Hawaii sites (R. H. Richardson, 1974; H. Carson, personal communication). Using the capture–recapture method, Fontdevila and Carson (1978) estimated the population size of *D. engyochracea* in a small area (7432 m^2) of Kipuka Ki to be about 40,000 (5.4 flies per m^2). This area is only 1.3% of the total area of Kipuka Ki, but it is densely populated by this species. There is at least one more densely populated area in Kipuka Ki (R. H. Richardson, personal communication), and the total population size in this location could be about 100,000. The population of *D. engyochracea* in Kipuka Puaulu seems to be about four times larger than that in Kipuka Ki (W. W. M. Steiner, personal communication). Therefore, the total population size of this species is estimated to be about half a million. *Drosophila mimica* is more abundant than *D. engyochracea*, the population density of the former being at least five times higher than that of the latter (W. W. M. Steiner and R. H. Richardson, personal communications). Therefore, the population size of this species seems to be about 3 million. Carson (1976) has stated that Hawaiian "picture-winged" drosophilids are large, slow-breeding flies, and there may be no more than two generations a year. However, Steiner informs us that *D. engyochracea* and *D. mimica* are relatively small and there may be about five generations a year.

In one continental species of *Drosophila*, i.e., *D. nigrospiracula*, the population size has been estimated to be 10^{11}. This species feeds on necrotic Saguaro cacti in the southwestern United States desert. The population size was estimated from information on the number of cacti in the area (J. S. Johnson and Heed, 1976). Unfortunately, there are no gene diversity data for this species. J. S. Johnson and Heed (1976) speculate that the population size of *D. pseudoobscura* is 100–1000 times larger than that of *D. nigrospiracula*. We also have rough estimates of relative species sizes for the following species groups: 1: 10^6 for *D. mojavensis*: *D. pseudoobscura*: *D. willistoni* (Stone *et al.*, 1960); 1:2:2:10 for *D. tropicalis*: *D. equinoxialis*: *D. paulistorum*: *D. willistoni* (Spassky *et al.*, 1971); and 1:1 for *D. nebulosa*: *D. willistoni* (Ayala *et al.*, 1974). Therefore, we can get crude estimates of population sizes N of these species. If we assume that *D. pseudoobscura* is 100 times more abundant than *D. nigrospiracula*, we obtain an estimate of $N = 10^{13}$ for *D. willistoni* and *D. nebulosa*. This estimate is, however, apparently too high. *Drosophila willistoni* inhabits a large territory of South America and Central America from northern Argentina to Mexico, the Caribbean Islands, and Florida, the total area of which is about 2×10^7 km^2. If our estimate of N is correct, the population density would be 5000 per m^2. This density is absurdly high, since a large proportion of this territory is not inhabitable by flies. The average population density for the entire territory would probably be at most 5 per m^2. We therefore used $N = 10^{12}$ for *D. willistoni* and *D. nebulosa*. The population sizes for the other *Drosophila* species were obtained by using the ratios given above relative to this number.

Escherichia coli occur normally in the intestines of mammalian species. There are about 4000 living species of mammals, and the average population size of one mammalian species would be of the order of 10^8. Thus, we estimate that the total number of mammalian individuals is of the order of 10^{11}. The average number of *E. coli* bacteria per mammalian

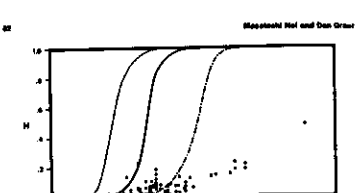

FIG. 1. Relationship between average gene diversity H and Ns. Dots represent observed values. The number of dots is less than 77, because some observed values overlap. Solid line: expected relationship for neutral alleles with a mutation rate of $v = 10^{-7}$. Dotted line: expected relationship for overdominant alleles with $v = 10^{-7}$ and $s = 0.001$. Dot-dashed line: expected relationship for slightly deleterious alleles with $v = 10^{-7}$ and $s = 0.002$. This relationship was obtained by using Kimura's (1979) formulas (23) and (24) with $\beta = 0.5$.

individual would be at least of the order of 10^9. We therefore estimate that the total number of *E. coli* individuals in the world is of the order of 10^{20}. This estimate is, of course, very, very crude, and could be several orders of magnitude off, but it indicates that the number is enormous. Furthermore, unlike higher organisms, the effective population size of *E. coli* would be drastically different from the actual size, as will be discussed later. *Escherichia coli* is a haploid organism, but this is a very minor factor affecting the effective size. Milkman (1976) has estimated that there would be about 1000 generations a year in natural populations of this organism.

Relationship between Gene Diversity and Population Size

Table I gives the observed value of mean gene diversity H_o, estimates of N and s, and the expected gene diversity H_E obtained from $H_E = 4Ns\mu/(4Ns\mu + 1)$ for each of the 77 species examined, and Fig. 1 shows the relationship between H_o and Ns. It is clear from Fig. 1 that there is a high correlation between H_o and Ns. The correlation coefficient is 0.651 and is significant at the 0.1% level. Thus, we cannot reject our first "null" hypothesis. Of course, our correlation coefficient is not close to 1, but

TABLE I. Average Gene Diversity H, Estimates of Population Size N, and Generation time g for Species in Which 20 or More Protein Loci Have Been Investigated

Group	Species	Number of loci	Gene diversity H		N	g	Source*	$Ns\mu$*
			Observed	Expected				
Primates	*Homo sapiens*	121	0.143	1.000	4×10^4	30	1	10^3
	Pongo pygmaeus	22	0.046	0.242	2×10^4	15	2	8×10^2
	Pan troglodytes	43	0.011	0.044	10^4	15	3	2×10^2
	Macaca cyclopis	29	0.041	0.074	2×10^4	10	4	10^3
	M. fascicularis	32	0.041	0.074	2×10^4	10	5	6×10^2
	Papio papio	29	0.053	0.167	5×10^4	10	6	10^3
Seals	*Phoca vitulina*	29	0.096	0.167	10^4	5	6	10^3
	Pagophilus groenlandicus	29	0.021	0.667	10^4	5	6	10^3
	Mirounga angustirostris	24	0.006	0.048	5×10^3	5	9	10^2
Ungulates	*Alces alces*	38	0.000	0.143	10^4	5	10	6×10^2
	Odocoileus virginianus	34	0.012	0.545	6×10^4	5	11	5×10^3
	Ovis dalli dalli	26	0.059	0.999	5×10^3	0.5	12	5×10^2
Lagomorphs	*Oryctolagus cuniculus*	25	0.011	0.667	2×10^4	0.5	13	8×10^2
	Ochotona princeps	25	0.011	0.833	2×10^4	0.5	14	7×10^2
Rodents	*Peromyscus gambia*	25	0.000	0.023	2×10^4	0.5	14	0
	P. aztecus	25	0.000	0.033	2×10^4	0.5	14	8×10^2
	P. merriami	25	0.016	0.833	2×10^4	0.5	14	0
	P. stephani	25	0.000	0.143	2×10^4	0.5	15	2×10^2
	P. floridanus	25	0.063	0.833	6×10^4	0.5	15	3×10^2
	P. leucopus	44	0.011	0.833	2×10^4	0.5	15	5×10^2
	P. polionotus	31–32	0.065	0.923	9×10^4	0.5	16	10^3
	P. pectoralis	29	0.122	0.999	6×10^4	0.5	15	7×10^3
	P. maniculatus	30	0.033	0.999	2×10^5	0.5	15	10^4
	P. californicus	23	0.021	0.960	2×10^4	0.5	16	10^3
	P. eremicus	25	0.060	0.706	2×10^4	0.5	8	2×10^3
	P. difficilis	25	0.040	0.909	2×10^4	0.5	8	2×10^3
	P. truei	25	0.060	0.645	6×10^4	0.5	15	3×10^3
	Eutamias palmeriae	36	0.003	0.455	2×10^4	0.5	17	10^2

(continued)

TABLE I. (Continued)

Group	Species	Number of loci	Gene diversity H		N	g	Source[a]	N_{min}[b]
			Observed	Expected				
	Sigmodon arizonae	26	0.033	0.941	6×10^7	0.5	15	2×10^7
	S. hispidus	26	0.020	0.992	5×10^8	0.5	15	10^7
	Spalax ehrenbergi (52)[c]	25	0.066	0.074	10^3	2	8	9×10^3
	S. ehrenbergi (54)[c]	25	0.038	0.138	2×10^4	2	8	5×10^4
	S. ehrenbergi (58)[c]	25	0.016	0.194	3×10^5	2	8	2×10^5
	S. ehrenbergi (60)[c]	29	0.035	0.444	10^6	2	8	5×10^6
	Thomomys umbrinus	27	0.031	0.998	5×10^8	2	8	4×10^7
	T. bottae	27	0.093	0.998	5×10^8	2	15	10^7
	Geomys personatus	24	0.027	0.998	10^8	2	15	3×10^7
	G. tropicalis	24	0.000	0.002	3×10^5	2	8	0
	G. bursarius[d]	24	0.063	0.998	5×10^8	2	15	8×10^6
Carnivores	*Vulpes vulpes*	21	0.000	0.016	2×10^4	2	17	0
	Mustela erminea	21	0.000	0.008	8×10^3	2	17	0
	M. putorius	21	0.000	0.008	8×10^3	2	17	0
	Martes fiana	21	0.000	0.008	8×10^3	2	17	0
	Meles meles	21	0.000	0.004	5×10^3	2	17	0
	Acinonyx jubatus[e]	47	0.000	0.002	2×10^3	3	18	0
Lizards	*Anolis trinitatis*	22	0.061	0.960	3×10^7	2	15	8×10^6
	A. carolinensis[e]	25	0.073	0.889	9×10^6	2	15	10^6
	A. grtaeus	22	0.020	0.941	2×10^7	2	15	3×10^6
	A. segrei[b]	24	0.010	0.970	4×10^7	2	15	10^6
	A. lucius	26	0.089	0.800	6×10^6	2	15	10^6
	A. cristellus	20	0.120	0.999	9×10^8	2	15	2×10^7
	A. wattsi	22	0.046	0.800	6×10^6	2	15	6×10^6
	A. distichus	23	0.051	0.800	5×10^6	2	15	7×10^6
	A. bimaculatus	22	0.053	0.194	3×10^5	2	15	7×10^6
	A. oculatus	22	0.050	0.997	4×10^8	2	15	7×10^6
	A. sabanus	22	0.044	0.194	3×10^5	2	15	6×10^6
	A. gingivinus	22	0.100	0.444	10^6	2	15	10^7
	A. grahami	24	0.078	0.976	6×10^7	2	15	10^8
	A. marmoratus	22	0.051	0.889	2×10^7	2	15	7×10^6
	A. roquet	25	0.074	0.800	6×10^6	2	15	10^7
	A. lividus	22	0.053	0.444	10^6	2	15	7×10^6
Alligators	*Alligator mississippiensis*	49	0.021	0.286	10^5	10	19	3×10^6
Newts	*Taricha rivularis*	40	0.068	0.737	7×10^6	10	20	2×10^6
Bony fish	*Ethestoma spectabile*	26	0.069	0.988	2×10^8	1	21	2×10^7
	E. coeruleum	26	0.066	0.992	3×10^8	1	21	2×10^7
	Oncorhynchus nerka	23	0.018	0.941	4×10^7	1	8	5×10^6
	Salmo salar	37	0.035	0.444	2×10^6	1	22	8×10^5
Fruitflies	*Drosophila nebulosa*	30	0.218	1.000	10^{16}	0.1	15	7×10^6
	D. tropicalis	30	0.153	1.000	10^{16}	0.1	15	5×10^6
	D. paulistorum	32	0.228	1.000	2×10^{13}	0.1	15	7×10^6
	D. willistoni	31	0.183	1.000	10^{14}	0.1	15	6×10^6
	D. equinoxialis	30	0.185	1.000	2×10^{13}	0.1	15	6×10^6
	D. pseudoobscura	46	0.136	1.000	5×10^{11}	0.1	15	4×10^6
	D. miranda	21	0.222	0.194	3×10^6	0.2	15	4×10^6
	D. engyochracea	20	0.127	0.040	5×10^7	0.2	15	2×10^6
Landsnails	*Sphincterochila aharoni*	29	0.067	0.074	10^3	1	23	2×10^3
Plants	*Pinus torreyana*	39	0.000	0.563	9×10^3	10^2	24	0
Bacteria	*Escherichia coli*	20	0.472	1.000	10^{20}	10^{-3}	25	2×10^9

[a] 1, Nei and Roychoudhury (1982); 2, Bruce and Ayala (1979); 3, M.-C. King and Wilson (1975); 4, Nozawa *et al.* (1977); 5, Nozawa *et al.* (1982); 6, Kawamoto *et al.* (1981); 7, Lavigne *et al.* (1982); 8, Nevo (1978), and references therein; 9, Ryman *et al.* (1977); 10, Manlove *et al.* (1977), and references therein; 16, Avise *et al.* (1974a); 11, Glover *et al.* (1977); 14, Avise *et al.* (1974b); 15, Fuerst and Hedgecock (1978); 21, McLean *et al.* (1980); 22, Ståhl (1982); 23, Nevo (1982); 24, Ledig and Conkle (1983); 25, Selander and Levin (1980).

[b] N_{min} represents the effective population size required for explaining the observed gene diversity in terms of the neutral theory.

[c] Average for *albifrons, lucubrans, polionotus, subgriseus,* and *pharaa* subspecies.

[d] Average for *lacrianus* and *collinus* subspecies.

[e] The number in parentheses is the diploid number of chromosomes.

[f] Average for central and peripheral subpopulations.

[g] *Jubatus* subspecies.

[h] Bimini populations.

since our estimates of population size are very crude, the correlation is not expected to be high anyway. As will be discussed later in this section, the ratio of effective size N_e to actual size N is expected to be smaller when N is large than when N is small. It is therefore interesting to see the correlation between H_O and $\log Ng$. In the present case it becomes 0.775. Thus, this supports our expectation, though the difference between the two correlations is small.

The positive correlation between H_o and Ng eliminates ecological theories in which the effect of genetic drift is discounted, but it is consistent with a number of genetic theories in which the drift effect is considered. Thus, we must examine our second "null" hypothesis.

The solid line in Fig. 1 shows the expected gene diversity for neutral alleles with the assumption of $v_n = 10^{-7}$. It is clear that all observed mean gene diversities are approximately equal to or lower than the expected values except in one species, *Drosophila engyochracea*. The observed value for *D. engyochracea* is 0.127, whereas the expected value is 0.038. The expected standard error of the observed value is 0.040 [formula (6.18) in Nei (1975)], so that the difference is statistically significant. A close examination of Table 1 shows that the observed heterozygosity of *Drosophila mimica* is also slightly higher than the expected value. This difference is not statistically significant, but we must note that in the computation of H_E we used N rather than N_e. If the N_e of this species is substantially smaller than N, then the difference between H_E and H_O would become significant. We also note that in *Macaca cyclopis*, *M. fascicularis*, *Peromyscus guardia*, *Spalax ehrenbergi* (diploid number of chromosomes: 52), and *Sphincterochilla aharoni* the expected heterozygosity is less than two times the observed heterozygosity. If N_e is about one-third or one-fourth of N in these species, then H_O would be higher than H_E. However, in these species we used a minimum estimate of population density to compute N, so that our estimate of N may not be much larger than N_e. In all other species the observed heterozygosity is substantially lower than the expected heterozygosity.

There are two possible explanations for the discrepancy between the observed and expected gene diversities in *D. engyochracea* and *D. mimica*. The first is the possibility that there is overdominant selection or a similar balancing selection operating in these species. If this is the case, the distributions of allele frequencies and single-locus gene diversities are expected to be different from those for neutral alleles (Li, 1978; Maruyama and Nei, 1981). In the present case the distribution of single-locus gene diversities is not very helpful, because only 20 loci have been studied. The observed distribution of allele frequencies for *D. engyochracea* is given in Fig. 2 together with the theoretical distributions for neutral and

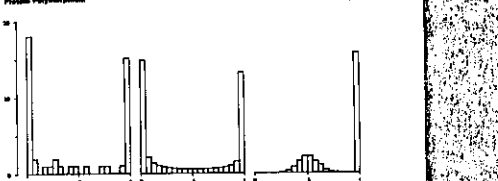

Fig. 2. Observed and expected distributions of allele frequencies. (a) Observed distribution for *D. engyochracea* ($H = 0.127$). (b) Expected distribution for neutral alleles with $H = 0.127$. (c) Expected distribution for overdominant alleles with $H = 0.127$.

overdominant alleles. The theoretical distributions were obtained under the condition that the expected gene diversity is equal to the observed value, i.e., 0.127 (Chakraborty *et al.*, 1980). In the case of neutral alleles this corresponds to $4Nv = 0.145$, whereas in the case of overdominant alleles $Ns \approx 30$ and $4Nv \approx 2.3 \times 10^{-7}$ were used, where s is the selective advantage for heterozygotes (Li, 1978). We note that even a small degree of overdominant selection is so powerful in maintaining genetic variability that an extremely low rate of mutation is sufficient to attain the level of $H = 0.127$ (Li, 1978; Nei, 1980a). Figure 2 indicates that the observed distribution of allele frequencies is very close to the neutral expectation but considerably different from the overdominant pattern. This suggests that the major factor for the discrepancy between the observed and expected gene diversity in *D. engyochracea* is not overdominance or similar balancing selection. Essentially the same conclusion was obtained for *D. mimica*.

The second possible explanation is that our estimates of N are not correct and the actual value is several times higher than our estimates. In our view this is quite possible, since no one has directly estimated the total population sizes of these species and our indirect estimates are very crude. We note that if the population size is several times higher than our estimates, the discrepancy between the expected and observed gene diversities disappears. Our population estimates are heavily dependent on Fontdevila and Carson's (1978) population survey in *D. engyochracea* by

using the capture–recapture method in one spot of Kipuka Ki. The estimate obtained by the capture–recapture method is known to be subject to a large sampling error (Cormack, 1973), so that our estimates may be unduly low. Furthermore, the possibility that these two species live in other locations in addition to the two kipukas cannot be excluded, since Richardson and Carson collected a few specimens of these species in far removed locations of the island of Hawaii as mentioned earlier. At any rate, at the present time no definite conclusion can be made about the cause of the discrepancy between the expected and observed gene diversities in this species. It would be rewarding to study the population size and structure of this species in further detail.

While we could not reject our second "null" hypothesis except in two possible species, the discrepancy between the expected and observed mean gene diversities increases as population size increases (Fig. 1). There are three possible explanations for this relationship. The first is variation in the mutation rate among different loci. Nei *et al.* (1976*a*) have shown that when the mutation rate varies from locus to locus the increase in average gene diversity with increasing population size is slower than when the mutation rate is the same. However, their Fig. 4 shows that the effect of this factor is relatively minor unless the extent of variation is extremely large.

The second explanation is that the ratio of effective population size to actual size N_e/N is generally much smaller in organisms with large N. For example, the actual number of *E. coli* cells in the world is apparently enormous, as mentioned earlier, but the effective size must be a tiny fraction of the actual size (Nei, 1976*b*; Maruyama and Kimura, 1980; B. R. Levin, 1981). This is because under certain circumstances *E. coli* strains rapidly grow but under other circumstances they easily become extinct, and in this case the effective population size is much smaller than the actual size. It is well known that after the Second World War drug-resistant strains of *E. coli* rapidly spread in many countries. It should also be noted that the population sizes of host species (mammalian species) must have been quite small at the times of glaciations. A large fluctuation of population size apparently occurs also in continental *Drosophila*. Crumpacker and Williams (1973) and Jones *et al.* (1981) have shown that the *D. pseudoobscura* populations in Colorado and California rapidly colonize new territories in some seasons or years, but quickly disappear in others, so that the seasonal or yearly fluctuation of population size seems to be enormous. In *D. melanogaster* and *D. simulans* on even larger scale of colonization and population replacement seems to be occurring. *Drosophila melanogaster* has a worldwide distribution at the present time, but it seems that the species was confined primarily to West Africa until

Europeans started worldwide navigation in the 16th century (Lemeunier and Ashburner, 1976; Ishii and Charlesworth, 1977). Colonization of *D. simulans* is also often very rapid. This species was virtually nonexistent in Japan until recently. Starting around 1972, however, it spread through the entire Japanese islands from south to north, and in 1974 these flies were already collected in the northern island of Hokkaido (Watanabe and Kawanishi, 1976). In general, fluctuation of population size occurs more often in organisms with high reproductive rate than in organisms with low reproductive rate. Since the former organisms tend to have larger population sizes, we would expect the discrepancy between observed and expected gene diversities to be larger in these species.

The third possible factor is the long-term effect of bottlenecks of population size. Earlier we discussed the possibility that in a large proportion of extant organisms the genetic variability is still under the influence of the bottlenecks that occurred during the ice ages of the Pleistocene. This seems to be true even in such tropical fruitflies as *D. willistoni* and *D. nebulosa*. In northern South America and Central America, where these flies live, there were no glaciations during the Pleistocene, but in the ice ages this area was apparently cool and dry, so that the number of flies must have been much smaller than the present size. It is also possible that they emerged as new species through bottlenecks in the last 1 or 2 million years. In man we have a fairly good estimate of population growth in the last 10,000 years (Cavalli-Sforza and Bodmer, 1971; Boaz, 1979), and the relatively low level of gene diversity in the present human population is apparently due to the small population size at the time of the Würm–Wisconsin glaciation. After the glaciation, population size has apparently increased in many organisms, but the increase in gene diversity is slow, and in large populations a longer period of time is required for the equilibrium value to be attained than in small populations (see Fig. 3). Therefore, this factor also contributes to a larger discrepancy between the observed and expected gene diversities in large populations.

NEUTRAL THEORY AND ITS ALTERNATIVES

We have seen that our data on the relationship between gene diversity and population size can be accommodated with the neutral theory, but this itself does not prove the validity of this theory, since some other hypotheses may explain the data equally well. We shall therefore examine the compatibility of our data with other hypotheses in which selection is invoked. In this chapter we consider only those hypotheses or theories

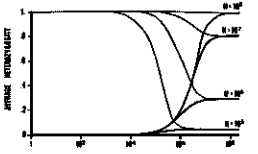

FIG. 3. Changes of expected gene diversity (heterozygosity) for neutral alleles with evolutionary time. Solid lines represent cases where the initial gene diversity is zero, and dotted lines represent cases where the initial gene diversity is one. The mutation rate v assumed is 10^{-6} per generation. The expected gene diversity at generation t is given by $H_t = H_\infty + (H_0 - H_\infty)\exp[(2v + 1/2N_e)t]$, where H_0 is the initial value and $H_\infty = 4N_e v/(1 + 4N_e v)$ (Nei, 1975). $N = N_e$.

in which a quantitative or qualitative prediction can be made about the level of protein polymorphism. Theoretically, one can imagine a very flexible theory in which selection regimes vary with time, location, locus, and allele set, so that it can explain almost every set of data. However, such a theory is unfalsifiable and thus unscientific (Popper, 1959). In this section we shall consider other sets of data obtained by previous authors as well as ours and attempt to derive a general conclusion about the maintenance of protein polymorphism. Before going into the detail of the discussion, however, we shall first discuss the general property of the neutral theory and the criticisms raised against it, since this theory has been misinterpreted by many authors.

Neutral Theory and Bottleneck Effects

The neutral theory of molecular evolution may be characterized by the following statements.

1. It refers only to molecular variation and evolution and does not apply to morphological evolution.

2. It is concerned with the behavior of a "majority" of genes that are incorporated into the population during evolution and allows for the existence of a small proportion of advantageous or overdominant mutations.

3. It is assumed that the majority of fresh mutations are deleterious but they are quickly eliminated from the population and thus contribute little to the genetic variation or gene substitution in a population. This large fraction of deleterious mutations occurs because most new mutations disturb the function of the protein encoded (Kimura, 1983).

4. Neutral genes are not functionless genes but are generally of vital importance to the organism. At any locus a pair of alleles are called neutral if they are functionally equivalent and thus equally important to the survival of the organism. In population genetics, however, the definition of neutrality of a gene depends on whether the behavior of the gene in a population is dictated by genetic drift or not. Suppose that there are two alleles at a locus, and let s be the selective advantage of one allele over the other. Then, if $N_e s \leq 1$, the pair of alleles are called neutral (Kimura, 1968*b*; Li, 1978). Therefore, a mutant gene that is advantageous in a large population may become neutral in small populations. For more detailed properties of the neutral theory, see Nei (1975) and Kimura (1983).

A number of authors (e.g., Selander 1976; Valentine, 1976) have questioned the validity of neutral theory on the ground that Hawaiian species of *Drosophila* are on the average as polymorphic as continental species are, although many of them are confined to single islands and have population sizes several orders of magnitude smaller than those of continental species. Ayala (1975) and Steiner (cited in Nevo, 1978) have reported that Hawaiian *Drosophila* species other than those used here have an average gene diversity of 0.025–0.24. As we discussed earlier, it is possible that the gene diversities of *D. engyochracea* and *D. mimica* are higher than the neutral expectations, but this seems to be an exception. Actually, the population sizes of most Hawaiian species of *Drosophila* are much larger than those of *D. engyochracea* and *D. mimica*, since they have a much wider geographic distribution (Steiner, 1978*a*,*b*); their population size is probably well over 10^6. Thus, the gene diversities of these species are not really higher but probably lower than the neutral expectations.

It is known that the island of Hawaii was formed only about 500,000 years ago. Furthermore, using chromosomal markers, Carson (1970) has shown that most species of Hawaiian *Drosophila* were apparently established rather recently through repeated bottlenecks. It then appears that the primary factor of determining the level of protein polymorphism is the time since the last bottleneck, as in the case of continental species. If this time is nearly the same for the Hawaiian and continental species, the level of polymorphism is expected to be virtually the same for them. This can be seen from Fig. 3, where the pattern of increase of gene diversity is given for various population sizes. It is clear that in the early

generations gene diversity increases virtually at the same rate for all population sizes, and the differences in gene diversity among populations of different sizes becomes apparent only when the evolutionary time is very long. Another factor that would have contributed to the same level of gene diversity for Hawaiian and continental species is the large seasonal or annual fluctuation of population size in continental species. Therefore, the N_e/N ratio is expected to be higher in Hawaiian *Drosophila* than in continental *Drosophila*.

The reduction of effective population size due to fluctuation of population size seems to be an important factor in reducing genetic diversity for a wide variety of organisms. For example, *Sigmodon hispidus* and *Peromyscus maniculatus* are known to have similar average population sizes, but the former experiences a more drastic fluctuation in population size than the latter (Smith *et al.*, 1975). Thus, the former is expected to have a lower gene diversity. This is indeed the case (see Table I). The wild rabbit *Oryctolagus cuniculus* is also known to undergo periodic bottlenecks in population size due to myxomatosis outbreaks and droughts (Richardson *et al.*, 1980). The relatively low level of gene diversity in this and some other organisms, such as the polar bears (Allendorf *et al.*, 1979), could be a result of fluctuation in population size (see Hecht (1974) for the population size of the polar bear).

Every population geneticist is aware of the importance of random genetic drift in small populations. It is well established theoretically (Wright, 1931) and experimentally (e.g., Dobzhansky and Spassky, 1947; Rich *et al.*, 1979). A number of authors have noted that the level of protein polymorphism is indeed lower in small populations than in large populations in the same species. For example, Selander *et al.* (1971) showed that the Santa Rosa Island (off the Gulf Coast of the Florida panhandle) population of *Peromyscus polionatus*, of which the size is known to be of the order of 12,000, has a gene diversity of 0.018, compared with the gene diversity of 0.086 in the Florida population. The cave populations (200–500 individuals) of the characid fish *Astyanax mexicanus* also have a very low gene diversity compared with the nearby surface population (Avise and Selander, 1972). Similar observations have been made in mammals (Price and Kennedy, 1980), lizards (Webster *et al.*, 1972; Gorman *et al.*, 1975), insects (Prakash *et al.*, 1969; Saura *et al.*, 1973) and plants (Hamrick *et al.*, 1979; Levin *et al.*, 1979). Nevertheless, these authors did not accept the importance of genetic drift except in small populations.

In our view it is illogical not to recognize the effect of genetic drift in large populations while accepting it in small populations. If the effect is seen in small populations, it must be operating in large populations as well, though the extent of the effect is inversely proportional to the population size. The present study supports this view. Studying 12 protein loci, Schmidtke and Engel (1980) also reported that the mean gene diversities in two tunicate species "are clearly related to population size as predicted by the neutral mutation hypothesis." Varvio-Aho (1981) conducted a similar study on eight species of Finnish water-striders and concluded that "the genetic variation does not correlate with local population sizes but correlates clearly with (species) effective population sizes, which can qualitatively be estimated on the basis of dispersion efficiencies, abundances, and habitat stabilities of the species." Selander (1975) studied the genetic structure of populations of the landsnail *Helix aspersa* and concluded that "to a surprising degree . . . the complex structure of the California population of *Helix aspersa* can be accounted for by differentiation of subpopulations as a result of random genetic drift" (see also Selander and Whittam, 1983).

Nelson and Hedgecock (1980) denied the importance of genetic drift for protein polymorphism on the grounds that different loci do not show the same degree of polymorphism in the same species. Actually this observation is perfectly compatible with the neutral expectation. There are two reasons for this. First, the mutation rate apparently varies greatly from locus to locus (Nei *et al.*, 1976*a*; Koehn and Eanes, 1978; Zouros, 1979), so that the gene diversity is expected to vary with locus. Second, and more importantly, genetic drift alone is known to produce a large variation of single-locus gene diversity (Ewens and Gillespie, 1974; Nei *et al.*, 1976*b*). Therefore, different loci in the same species are expected generally to have different gene diversities.

Balancing Selection

In the last 10 years many different selective mechanisms have been proposed for explaining the observed level of protein polymorphism (e.g., Ayala, 1976). Most of these mechanisms invoke some sort of balancing selection, and in a large population, gene frequency equilibrium is assumed to be reached. The classic example is overdominance, but there are other types of balancing selection, such as frequency-dependent selection (Wright, 1969) and habitat selection (Powell and Taylor, 1979). Nei (1980*b*) has called this type of selection diversity-enhancing (or diversity-retaining) selection, since this type of selection increases genetic variability compared with the case of neutral genes. There are other types of selection that tend to reduce genetic variability. One such example is purifying selection, in which mutant genotypes have a lowered fitness. Nei (1980*a*) has called this type of selection diversity-reducing selection.

Many selectionists (e.g., Ayala, 1972; Wills, 1973; Milkman, 1976; G. B. Johnson, 1976) believe that a large part of protein polymorphism is maintained by diversity-enhancing selection. This view was probably generated by a misconception about Harris' (1966) and Lewontin and Hubby's (1966) brilliant discovery that natural populations are highly polymorphic at the protein level. This discovery has somehow misled many biologists to believe that the polymorphism is so extensive that some kind of balancing selection must be operating (e.g., Sved *et al.*, 1967; J. L. King, 1967; Milkman, 1967). The present study indicates that the level of protein polymorphism is actually much lower than the neutral expectation and that if the bottleneck effect is not sufficient for explaining the observed level, the type of selection to be considered is not diversity-enhancing selection but diversity-reducing selection.

Figure 1 includes the expected relationship between gene diversity and population size when v is assumed to be 10^{-7} and all heterozygotes have a selective advantage of $s = 0.001$ compared with homozygotes. This expected relationship was obtained by the methods of Li (1978) and Maruyama and Nei (1981). It is clear that even if a small magnitude of overdominance is considered, the amount of expected gene diversity increases tremendously compared with the case of neutral mutations and is far above the observed gene diversities. The curve in Fig. 1 refers to the case of constant s for all heterozygotes, but the result is nearly the same even if s varies with allele combination as long as the mean of s remains the same (Maruyama and Nei, 1981).

The unimportance of balancing selection is also indicated by the patterns of distributions of single-locus gene diversities and of allele frequencies. Fuerst *et al.* (1977) compared the observed distribution of gene diversities with the neutral expectation in 68 animal species, but in none of these species was the former significantly different from the latter. Earlier we showed that the observed distribution of allele frequencies for *Drosophila engyochracea* is close to the neutral expectation rather than to the overdominant expectation. A more extensive study of this kind was conducted by Chakraborty *et al.* (1980), who examined allele frequency data from 138 populations (species or subspecies). In all of these populations the distribution was U-shaped, as expected for neutral alleles, and there was no indication of overdominant selection. We can, therefore, conclude that overdominant selection is generally unimportant for explaining protein polymorphism. We also note that *E. coli* has the highest gene diversity so far observed despite the fact that this organism is haploid, and thus there is no way for this organism to be subjected to overdominant selection. A relatively high degree of gene diversity in haploid populations was also observed in *Neurospora intermedia* (Spieth, 1975), *Histoplasma*

capsulatum (Gaur *et al.*, 1981), and two haploid moss species (Krzakowa and Szweykowski, 1979; Yamazaki, 1981).

Recently, Powell and Taylor (1979) tried to explain protein polymorphism in terms of habitat selection. In this model each individual is supposed to choose the habitat most suited to its genotype. However, this is a strong form of balancing selection (Taylor, 1976), and if this type of selection is prevalent, we would expect a much higher level of gene diversity than we observe now. It should also be noted that natural populations are polymorphic at many loci (10–60% at the protein level) and, thus, the number of possible genotypes in a population is enormously large. In this case it is not clear how habitat selection occurs and how effective it is as a diversity-enhancing selection. A similar criticism is raised against Kojima and Yarbrough's (1967) hypothesis of frequency-dependent selection [see Lewontin (1974) and Nei (1975) for additional comments].

In the present discussion we have not considered the relationship between polymorphism and gene substitution for balancing selection. This relationship is somewhat complicated, but roughly speaking, balancing selection is so powerful for maintaining polymorphism that it increases the level of gene diversity substantially compared with the case of neutral genes but decreases the rate of gene substitution when the mutation rate is fixed (Nei, 1980a; Maruyama and Nei, 1981). Therefore, if we consider polymorphism and gene substitution simultaneously, the balancing selection hypothesis faces a serious problem.

Adaptive Strategy and Heterogeneous Environments

After seeing no clearcut evidence for overdominant selection in the maintenance of protein polymorphism, an increasing number of authors have turned to the hypothesis of adaptive strategy to explain this polymorphism. This hypothesis is primarily due to Levins (1968), who stated that heterogeneous and changing (coarse-grained) environments maintain a larger amount of genetic polymorphism than constant (fine-grained) environments. He used the so-called fitness set theory to support his statement. However, this theory depends upon a number of overly simplified assumptions about the fitness relationship among different environments. Actually, his arguments are no more than the usual intuitive arguments that are often made by biologists [see, however, Strobeck (1975) for a special case]. He has not really considered the Mendelian segregation of genes, which is the basis of population genetics. A more solid mathematical study on genetic polymorphism in heterogeneous niches was con-

ducted earlier by Levene (1953), who showed that in heterogeneous niches even genic or dominant selection may produce stable polymorphism. His mathematical model has recently been extended to various cases by a number of authors, but most of the models studied depend on the assumption that in each niche population size is regulated [for reviews see Hedrick *et al.* (1976) and Felsenstein (1976)]. If there is no regulation of population size (Wright, 1969; Christiansen, 1974) or if population size regulation occurs in the entire population rather than in each niche (Nei, 1971), then no stable polymorphism will be established without overdominance. In practice, we do not know whether or not regulation of population size occurs in every generation, and, if it does, how. Indeed, we do not even know whether or not most electromorphs produce fitness differences large enough to create stable polymorphism, and, if they do, whether fitness variation occurs in such a way that the stability condition is guaranteed in every generation. In Levene's model the marginal fitness of heterozygotes must be higher than that of homozygotes for the case of two alleles. In the presence of multiple alleles the condition is more severe. In our view it is quite unlikely that the delicate assumptions underlying Levene's model and its modifications (e.g., Hedrick *et al.*, 1976) can be satisfied at all in every capricious environment and every organism in nature. We note that weather alone almost never shows exactly the same pattern in different years. Finite population size also reduces the stability of polymorphism substantially (Hedrick, 1974; Nei and Yokoyama, 1976; Takahata, 1981). Gillespie (1977) claimed that the condition for polymorphism in heterogeneous environments is less stringent than that for overdominant selection in the presence of multiple alleles, but his model of heterogeneous environmental selection is very specific and there are no experimental data to support it.

Theoretically, even in a single niche nonoverdominant stable equilibrium may be generated if selection coefficients vary from generation to generation in a specific manner in large populations (Dempster, 1955; Haldane and Jayakar, 1963; Jensen and Poltak, 1969; Karlin and Levikson, 1974; Gillespie, 1978). However, this also depends on how selection operates and how population size is regulated (Nei, 1976*a*; Nei and Yokoyama, 1976), and again there are no experimental data to support any specific model so far developed.

In population cage experiments with *Drosophila* Powell (1971) and McDonald and Ayala (1974) have reported that the gene diversity of the populations maintained in heterogeneous environments is generally higher than that of the populations maintained in constant environments. A number of authors have taken this as evidence for supporting the view that varying selection intensity results in balancing selection. However, as

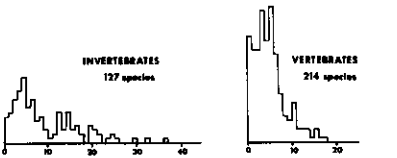

FIG. 4. Distributions of average gene diversity (heterozygosity) for species of invertebrates and vertebrates. Only species in which 20 or more loci were examined are included.

noted by Nei (1980*a*), their data show that in both constant and varying environments the gene diversity declined at a rate higher than expected under pure random genetic drift. In other words, heterogeneous environments did not retard but actually accelerated the reduction of gene diversity.

Selander and Kaufman (1973*b*) observed that the mean gene diversity is generally higher in invertebrates than in vertebrates. They explained this difference in terms of Levins' (1968) theory. Later investigations, however, showed that there are many invertebrate species that show very low gene diversity (e.g., Selander and Kaufman, 1973*a*; Tracey *et al.*, 1975), particularly in the Hymenoptera (Snyder, 1974; Metcalf *et al.*, 1975; Pamilo *et al.*, 1975, 1978*a*,*b*; Lester and Selander, 1979; Halliday, 1981; Varvio-Aho, 1981). On the other hand, there are examples of vertebrates that show high gene diversity (e.g., Avise and Smith, 1974; Milton and Kochn, 1975; Matthews, 1975; Nei and Roychoudhury, 1982). In collaboration with Paul A. Fuerst, we compiled mean gene diversity data for various groups of organisms. Figure 4 shows the distribution of mean gene diversity for vertebrates and invertebrates. It is clear that although invertebrates have a higher mean gene diversity than vertebrates, there are many species showing a very low value in invertebrates. Therefore, this set of data does not support Levins' hypothesis, but rather refutes it if there is any difference in environmental grain between vertebrates and invertebrates. We also note that Selander and Kaufman's (1973*a*) data on the landsnail *Rumina decollata* are incompatible with Levins' theory. This organism is sessile, so that its environmental grain must be coarse ac-

cording to Levins' definition. After being introduced to North America from Europe about 150 years ago, this self-fertilizing species has invaded a large territory of the southern United States, Mexico, and the Caribbean Islands. Despite their remarkable success in colonization of the North American continent and the polymorphic status in the original European populations (Selander and Hudson, 1976), the American populations of this species are entirely monomorphic for the 25 protein loci examined. Later observations that many organisms living in the deep sea have high degrees of gene diversity despite their stable environments (Gooch and Schopf, 1972; Valentine and Ayala, 1975; Ayala *et al.*, 1975; Ayala and Valentine, 1979) while offshore organisms living in unstable habitats have low degrees of heterozygosity (Tracey *et al.*, 1975; Nevo *et al.*, 1980; D. Wool and T. Perl, personal communication) are also incompatible with Levins' hypothesis.

After falsification of Levins' hypothesis, Valentine and Ayala (1975) and Ayala and Valentine (1979) proposed another type of adaptive strategy hypothesis. Ayala and Valentine (1979, p. 28) stated:

> functionally flexible alleles are assumed to be favored in species that inhabit temporally variable environments; . . . where individual organisms range widely through a variety of environments; in the more extreme cases the species tend to be monomorphic for the most flexible alleles. In stable environments a variety of alleles, with narrower functional ranges, are found in the gene pool of species whose individuals have restricted environmental ranges. The multiple alleles are maintained by balancing selection.

This hypothesis was presented as an *ad hoc* explanation for their observations and is just the opposite of Levins' theory. Actually, there is no theoretical basis for Valentine and Ayala's hypothesis; indeed, it is very difficult to conduct a mathematical study. Furthermore, there are empirical data that are incompatible with their hypothesis. For example, *E. coli* inhabits the intestines of many different mammalian species with different diet types, so that the environment must be quite heterogeneous. (Of course, this judgment is as subjective as Ayala and Valentine's argument!) Yet it has the highest gene diversity observed so far. On the other hand, cave animals like *Astyanax mexicanus* apparently live in highly stable environments, but they have low degrees of genetic polymorphism. Nevo's (1978) specialist–generalist theory is equally subjective; no one knows how to quantify the degree of specialization without using vague definitions that raise conceptual and methodological problems (Cody, 1974; Maynard Smith, 1974; Futuyma, 1976). Furthermore, specialization is often a local phenomenon rather than a species property. Fox and Morrow (1981) have shown that many herbivorous insects have generalized diets over the species' entire geographic ranges but they function as specialists with restricted diets in local communities. It should also be noted that some "generalist" species, such as the landsnail *Rumina decollata* (Selander and Kaufman, 1973*a*) and the parthenogenetic earthworm *Ortolasion tyrtaeum* (Jaenike *et al.*, 1980), have little genetic variability, unlike the prediction from Nevo's theory.

A number of authors (e.g., Bryant 1974*a*,*b*; Nevo, 1978; Hamrick *et al.*, 1979; Nelson and Hedgecock, 1980) have studied the correlations between gene diversity (or heterozygosity) and various environmental variables such as temperature or humidity to find the determinants of genetic variability. Critical review of these works, however, suggests that many of the correlations identified by these authors are spurious and do not indicate real causal relationships (see Schnell and Selander (1981) for an excellent review). Furthermore, in these studies little attention has been paid to the underlying genetic mechanism (or model), such as overdominant selection and selection in heterogeneous environments, so that no one knows what is the expected correlation for any given environmental condition. In other words, no particular genetic hypothesis is really tested in these studies. It should also be noted that the gene diversity or heterozygosity of a population is a product of long-term evolution (sometimes over 1 million years) whether there is selection or not, and it is difficult to believe that each of the populations has had the same environmental condition for such a long time. Similar comments can be made about the studies on the correlations between allele frequencies and environmental variables.

It is noted that all adaptive strategy hypotheses assume some kind of balancing selection. As mentioned earlier, however, what is really necessary for explaining the observed levels of gene diversity is not diversity-enhancing selection but diversity-reducing selection. From this point alone, these hypotheses do not appear to be important. Proponents of adaptive strategy hypotheses have not realized this point, because they have never tried to explain polymorphism and gene substitution simultaneously.

Advantageous Mutations

In neo-Darwinism gene substitution is assumed to occur when a new advantageous mutation is introduced or when an environmental change causes a previously disadvantageous mutation to become advantageous. If this process continues for a large number of loci, transient polymorphism is generated in a certain proportion of loci. The allele frequency distribution for this type of mutation (sequentially advantageous mutations)

is similar to that of neutral mutations (Chakraborty *et al.*, 1977). However, when the rate of gene substitution is fixed, the level of gene diversity is extremely low compared with the observed level in many natural populations.

This can be seen by using Kimura's (1969) infinite-site model of mutation. In this model the expected heterozygosities for neutral mutations H_n and advantageous mutations with genic selection H_s are given by

$$H_n = 1 - e^{-4N_e v} \tag{4}$$

$$H_s = 1 - e^{-4N_e v} \tag{5}$$

approximately. We note that (4) gives essentially the same value as that of (1) for $4N_e v \leq 0.4$ (or $H_n \leq 0.3$). The rate of gene substitution α is equal to v for neutral mutations, whereas it is $4N_e sv$ for advantageous mutations with genic selection, where s is the selective advantage for a mutant gene and $4N_e s \gg 1$ (Kimura and Ohta, 1971). Therefore, $v = \alpha$ for the former and $v = \alpha/(4N_e s)$ for the latter. Thus, for a given rate of gene substitution α, the expected heterozygosities are given by

$$H_n = 1 - e^{-4N_e \alpha} \tag{6}$$

$$H_s = 1 - e^{-\theta\alpha/s} \tag{7}$$

Expression (6) gives a value similar to that of (1) as long as $4N_e \alpha \leq 0.4$. Thus, when $\alpha = 10^{-7}$, (1) and (6) give 0.0196 and 0.0198, respectively, for $N_e = 5 \times 10^6$, whereas for $N_e = 10^8$, they give 0.29 and 0.33, respectively. On the other hand, H_s is determined by α and s and is independent of N_e as long as $4N_e s \gg 1$. Thus, when $\alpha = 10^{-7}$, H_s is 0.0002 for $s = 0.001$ and 0.00002 for $s = 0.01$. Therefore, unless the selective advantage of a new mutant allele is extremely small, average gene diversity is very small. It should also be noted that H_s is independent of N_e, unlike the actual relationship between H and N we have seen.

The reason for this is that advantageous mutations are quickly fixed in or lost from the population compared with neutral mutations, so that they do not generate extensive polymorphism. Theoretically, the average time required for an advantageous mutation to be fixed in the population is $(2/s)\ \ln(2N_e)$ generations approximately, whereas the fixation time for a neutral mutation is $4N_e$ generations. Thus, if $N_e = 10^6$, it takes 4 million generations for a neutral mutation to be fixed. However, the fixation time for an advantageous gene with a selective advantage of $s = 0.02$ will be 1450 generations. Therefore, the fixation time for neutral mutations is about 3000 times longer than that for advantageous mutations (see Fig. 5 for a comparison of the population dynamics of neutral and advantageous mutations). Furthermore, there are many neutral alleles that are not fixed in the population but become polymorphic and contribute to polymorphism.

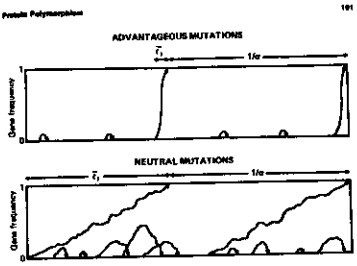

FIG. 5. Schematic patterns of gene substitutions for advantageous and neutral mutations. Advantageous mutations are either quickly lost from or quickly fixed in the population, so that their contribution to polymorphism for a given rate of gene substitution α is small. On the other hand, the frequency changes of neutral alleles are very slow in large populations and generate a great deal of transient polymorphism. The expected time for a mutation to be fixed is approximately $(2/s)\ \ln(2N_e)$ generations for advantageous mutations and $4N_e$ generations for neutral mutations, where N_e and s are the effective population size and selection coefficient. The expected time interval between two consecutive gene substitutions is $1/\alpha$, where $\alpha = 4N_e sv$ for advantageous mutations and $\alpha = 1/4v$ for neutral mutations. v is the mutation rate per generation.

Of course, if we increase α relative to s in (7), we can achieve a higher gene diversity similar to that observed in many organisms. In this case, however, the rate of gene substitution becomes too high compared with the observed rate. The population dynamics of dominant or recessive advantageous mutations is slightly different from that of mutations with no dominance (genic selection), but the essential feature is similar. Therefore, the hypothesis that protein polymorphism is generated largely by advantageous mutations is not satisfactory.

At this point, one might argue that both polymorphism and gene substitution can be explained if we consider a mixture of advantageous

mutations and overdominant mutations. This explanation seems to be attractive, but in this hypothesis the set of alleles showing overdominance is expected to persist in the population for a very long time—longer than species life (Maruyama and Nei, 1981). In practice, however, most polymorphisms appear to be transient, and the observed proportion of shared polymorphisms in related species can be explained by the neutral theory (Li and Nei, 1977). Of course, one can argue that because of environmental change balancing selection is relatively short-lived and gene substitution is facilitated. This argument is difficult to test since we do not know what kind of environmental change disturbs balancing selection and how often it occurs. Furthermore, if balancing selection is short-lived and turnover of genes occurs as often as in the case of neutral mutations there will be no difference in the evolutionary consequences between the two hypotheses.

It should also be noted that when long-term evolution encompassing many different phyla is considered, the rate of gene substitution is unlikely to be constant under selection, since in this case the substitution rate depends on the mutation rate, selection coefficient, generation time, and population size (Kimura and Ohta, 1971). This is not consistent with the actual observation that the rate of amino acid substitution is approximately the same for all branches of life (Wilson *et al.*, 1977). Van Valen (1973, 1974) claimed that his Red Queen hypothesis can explain the constant rate of molecular evolution. Actually, his claim is not justified, since the same comment as the above applies to his hypothesis. Furthermore, as indicated by Maynard Smith (1976) and Castrodeza (1979), the Red Queen hypothesis is expected to generate nonconstant rates in the presence of the species interaction Van Valen has postulated even if population size and generation time remain constant.

Slightly Deleterious Selection

Mendelian geneticists have established that almost all genes are subject to deleterious mutations. These mutations are quickly eliminated from the population and almost never become polymorphic. Therefore, this class of mutations has been ignored in the neutral mutation hypothesis (Kimura and Ohta, 1973). However, slightly deleterious mutations with very small selection coefficients may become polymorphic with an appreciable probability. Ohta (1974, 1976) has argued that in small populations these slightly deleterious mutations would behave just like neutral alleles, but in large populations there would be a mutation–selection balance between the type allele and other deleterious alleles and this balance

gives an upper limit for the gene diversity in a population. This hypothesis certainly explains the observed gene diversity lower than the neutral expectation, but it creates a new problem. Namely, since the mutation–selection balance is supposed to be obtained between a type allele (the best allele) and other slightly deleterious alleles in large populations, gene substitution is expected to occur very slowly or even stop in these populations. As mentioned above, however, gene substitution seems to have occurred roughly at a constant rate in all branches of life, whether population size is large or small. Furthermore, if this hypothesis is correct, the distribution of allele frequencies in large populations is expected to be bell-shaped. In all species examined, however, the actual distribution is U-shaped and is in accordance with the neutral expectation (Chakraborty *et al.*, 1980).

The mathematical theory of slightly deleterious mutations has recently been extended to the case where the selection coefficients of newly arisen mutations are continuously distributed from zero to one (Ohta, 1977; Li, 1978; Kimura, 1979). This extension mitigates the deficiency of Ohta's original hypothesis, but even this modified form does not seem to be able to explain observed data without recourse to the bottleneck effect. Figure 1 gives the expected relationship between H and N_E for the slightly deleterious mutation hypothesis (Kimura, 1979). The curve for this relationship is certainly lower than that for neutral genes, but the agreement between this curve and observed data is no better than that for neutral genes. Furthermore, there are many species in which the observed gene diversity is higher than the expected value. Kimura (1979) also attempted to explain the constant rate of substitution for this type of allele by assuming that generation time is inversely proportional to $\sqrt{N_e}$ and the effective neutral mutation rate is constant per generation. However, the effective neutral mutation rate thus defined depends on population size, and we are not sure whether the delicate relationship required among mutation rate, selection coefficient, generation time, and population size is really satisfied in nature or not. In fact, considering our set of data (Table 1), we find no significant correlation of generation time with N ($r = 0.031$), $\sqrt{N}$ ($r = 0.045$) or $\log N$ ($r = -0.196$).

Fluctuating Selection as a Diversity-Reducing Factor

In the past 10 years many authors have conducted mathematical studies of fluctuating selection as a cause of diversity-enhancing selection (e.g., Jensen and Pollak, 1969; Gillespie, 1973, 1978; Karlin and Levikson, 1974; Takahata *et al.*, 1975). However, as Nei (1976a) and Nei and Yo-

koyama (1976) indicated, the conclusions derived from these studies are very sensitive to the minor details of the model used, and if selection occurs through competition among genotypes, no balancing effect arises. Furthermore, in finite populations fluctuating selection generally tends to reduce genetic variability (Nei and Yokoyama, 1976; Takahata and Kimura, 1979). Recently, Takahata (1981) studied the expected gene diversity in finite populations using Gillespie's (1978) model. He showed that unless population size is very large, fluctuating selection reduces the expected gene diversity substantially. In other words, even Gillespie's model, which was originally thought to be diversity-enhancing, is actually diversity-reducing except in very large populations. These studies support Nei's (1975) view that fluctuating selection could be one factor that has made the observed gene diversity lower than the neutral expectation.

In our opinion, however, fluctuating selection becomes important only when the selection coefficient s is very small. In the hypothesis of slightly deleterious mutations, a selection intensity of the order of the mutation rate is considered. Namely, s is of the order of 10^{-5}. If the selection coefficient is as small as this, it would be very difficult to maintain a constant value of s in all generations. In this case, we believe, the selection coefficient fluctuates from generation to generation to a considerable extent, and thus the effect of variation of s must be considered. In Nei and Yokoyama's formulation these types of alleles are expected to behave just like neutral alleles when the mean of s is equal to zero, and the rate of gene substitution is not affected by population size. The only effect of this type of fluctuating selection is to increase the amount of genetic drift per generation and thus to reduce genetic variability. At any rate, if we combine this effect with the bottleneck effect, most of the discrepancies between the observed gene diversity and the neutral expectation can be explained.

CONCLUSION

In this and previous studies (Nei *et al.*, 1976*b*, 1978; Fuerst *et al.*, 1977; Chakraborty *et al.*, 1978, 1980) we have tested various "null" hypotheses concerning the neutral theory, but none of them could be rejected unconditionally. Rather, as discussed in this chapter, available data on protein polymorphism can be explained more easily by this hypothesis than by other, competing hypotheses. Of course, our statistical tests are not powerful enough to detect a very small extent of natural selection. Furthermore, there are some data that strongly suggest selection for electrophoretically detectable alleles (e.g., Koehn, 1969; Koehn *et al.*, 1983). However, these alleles seem to be rare, and our studies indicate that stochastic factors play a much more important role than deterministic factors in protein evolution. At the present time the neutral theory seems to be the simplest model to explain protein polymorphism and gene substitution adequately. Gillespie (1979) claimed that his SAS-CFF model (Gillespie, 1978) can explain them, but we doubt it (Nei, 1980*b*; Takahata, 1981).

In the past many authors have attempted to relate the level of gene diversity to various environmental factors as well as to biochemical parameters. Many initial claims were not supported by subsequent studies (Selander, 1976; Schnell and Selander, 1981). However, there are three factors that are clearly related to gene diversity. They are (1) quarternary structure of proteins (Zouros, 1976; Harris *et al.*, 1977; Koehn and Eanes, 1978), (2) subunit molecular weight of proteins (Koehn and Eanes, 1977; Ward, 1977; Nei *et al.*, 1978; Turner *et al.*, 1979), and (3) population size (this chapter; Nei, 1975; Soulé, 1976). We note that the relationships of gene diversity with these factors are all consistent with the neutral theory. In the future many other factors may be investigated to see their relationship with gene diversity. In these studies it is important not to confound the effects of the above three factors with those of the other factors to be studied.

Recent data on polymorphism and evolution at the nucleotide level also support the neutral theory. A number of authors (e.g., Kimura, 1977, 1983; Jukes, 1980; Perler *et al.*, 1980) have shown that the rate of nucleotide substitution is higher in the regions of genes that have less functional constraint than in those that have stronger functional constraint. Indeed, the highest rate has been observed in pseudogenes, which apparently do not have any function (Miyata and Yasunaga, 1981; Li *et al.*, 1981). Data on DNA sequence and restriction-site polymorphism suggest that the extent of polymorphism at the DNA level is much higher than that at the protein level (e.g., W. M. Brown, 1980, 1981; A. J. L. Brown and Ish-Horowicz, 1981; Nei, 1983). Apparently there is a large amount of polymorphism due to deletion or insertion in the noncoding region of DNA. If these observations are true for all genes, the contribution of these polymorphisms to fitness differences among individuals must be very small on the average.

It now seems clear that at the DNA level a large number of new mutations (nucleotide sequences) occur in every generation in the genome of higher organisms, and the majority of the mutations are eliminated quickly from the population because of their deleteriousness. A large proportion of the polymorphic mutations or the mutations that become

fixed in the population appear to be neutral or nearly neutral and apparently do not affect the function of the gene or the protein encoded appreciably. Of course, some proportion of nondeleterious mutations must be advantageous; otherwise no adaptive evolution can occur. However, if a few percent of nondeleterious mutations are advantageous as postulated by the neutral mutation theory, the adaptive change of organisms in evolution can easily be explained (Nei, 1975).

SUMMARY

The empirical relationship between mean gene diversity or heterozygosity H for protein loci and the product of population size N and generation time g is examined by using data from 77 different species. Although our estimates of population sizes are very crude, there is a highly significant correlation between H and Ng. When the mutation rate is estimated from data on amino acid substitution in proteins, the observed gene diversity is lower than the expected value under the neutral mutation hypothesis in all but two species. The tendency that the observed value is lower than the expected value apparently reflects the fact that the long-term effective population size is much smaller than the present size. It may also be partly due to fluctuating selection as conceived by Nei and Yokoyama (1976). The compatibility of various competing hypotheses with our present and previous results is also examined. Any hypothesis assuming balancing selection does not seem to be able to explain the observed level of protein polymorphism, since this makes the expected gene diversity much higher than the neutral expectation, which in turn is generally much higher than the observed value. The patterns of the distribution of single-locus gene diversity and allele frequencies are also close to the neutral pattern rather than to the overdominant pattern in all species examined. Most adaptive strategy hypotheses so far proposed are naively constructed and do not appear to be valid, since they are based on some type of balancing selection and there are many contradictory cases. Similarly, the hypotheses of sequentially advantageous mutations or slightly deleterious mutations are not appealing, because it is difficult to explain protein polymorphism and gene substitution simultaneously for a wide variety of organisms. It is concluded that the available data on protein polymorphism are most easily explained by a modified form of the neutral mutation hypothesis in which the effects of bottlenecks and fluctuating selection are taken into account.

ACKNOWLEDGMENTS

We thank the following investigators for helping us in gathering the data used in this study: Drs. Charles F. Aquadro, John C. Avise, Francisco J. Ayala, W. Wayne Barchard, D. G. Cameron, Hampton L. Carson, Ranajit Chakraborty, John Damuth, Paul A. Fuerst, George C. Gorman, Dennis Hedgecock, Ted Joanen, Richard K. Koehn, David M. Lavigne, James N. Layne, Wen-Hsiung Li, Ronald Lindahl, Michael N. Manlove, Michael J. McLeod, Eviatar Nevo, Ken Nozawa, Lawrence M. Page, Tova Pert, William Pfleger, B. J. Richardson, Richard H. Richardson, Nils Ryman, Valerio Shodroui, Jörg Schmidtke, Robert K. Selander, Shiva M. Singh, Michael H. Smith, Michael Soulé, William W. M. Steiner, Martin L. Tracey, David Wool, and David L. Wynes. This study was supported by research grants from the National Science Foundation and National Institute of Health.

REFERENCES

Allendorf, F. W., Christiansen, F. B., Dobson, T., Eanes, W. F., and Frydenberg, O., 1979, Electrophoretic variation in large-mammals. I. The polar bear, *Thalarctos maritimus*. *Hereditas* **91**:19–22.

Altman, P. L., and Dittmer, D. S., 1962, *Growth*, Federation of American Societies for Experimental Biology, Washington, D.C.

Andrews, R. M., 1979, Evolution of life histories: A comparison of *Anolis* lizards from matched island and mainland habitats, *Breviora* **454**:1–51.

Avise, J. C., and Selander, R. K., 1972, Evolutionary genetics of cave-dwelling fishes of the genus *Astyanax*, *Evolution* **26**:1–19.

Avise, J. C., and Smith, M. H., 1974, Biochemical genetics of sunfish. II. Genetic similarity between hybridizing species, *Am. Nat.* **108**:458–472.

Avise, J. C., Smith, M. H., and Selander, R. K., 1974a, Biochemical polymorphism and systematics in the genus *Peromyscus*. VI. The *boylii* species group, *J. Mammal.* **55**:751–763.

Avise, J. C., Smith, M. H., Selander, R. K., Lawlor, T. E., and Ramsay, P. R., 1974b, Biochemical polymorphism and systematics in the genus *Peromyscus*. V. Insular and mainland species of the subgenus *Haplomylomys*, *Syst. Zool.* **23**:226–238.

Avise, J. C., Smith, M. H., and Selander, R. K., 1975, Biochemical polymorphism and systematics in the genus *Peromyscus*. VII. Geographic differentiation in members of the *truei* and *maniculatus* species group, *J. Mammal.* **56**:177–192.

Ayala, F. J., 1972, Darwinian versus non-Darwinian evolution in natural populations of *Drosophila*, in: *Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 5: *Darwinian, Neo-Darwinian and Non-Darwinian Evolution* (L. M. LeCam, J. Neyman, and E. L. Scott, eds.), pp. 211–236, University of California Press, Berkeley.

Ayala, F. J., 1975, Genetic differentiation during the speciation process, *Evol. Biol.* **8**:1–78.

Ayala, F. J. (ed.), 1976, *Molecular Evolution*, Sinauer, Sunderland, Massachusetts.
Ayala, F. J., and Valentine, J. W., 1979, Genetic variability in the pelagic environment: A paradox?, *Ecology* 60:24–29.
Ayala, F. J., Tracey, M. L., Barr, L. G., McDonald, J. P., and Pérez-Salas, S., 1974, Genetic variation in natural populations of five *Drosophila* species and the hypothesis of the selective neutrality of protein polymorphism, *Genetics* 77:343–384.
Ayala, F. J., Valentine, J. W., Hedgecock, D., and Barr, L. G., 1975, Deep-sea asteroids: High genetic variability in a stable environment, *Evolution* 29:203–212.
Barchard, W. W., 1978, Estimates of Historical Stock Size for Northwest Atlantic Harp Seals, *Pagophilus groenlandicus*, M. S. Thesis, University of Guelph, Guelph, Ontario, Canada.
Bennett, A. F., and Gorman, G. C., 1979, Population density and energetics of lizards of a tropical island, *Oecologia* 42:339–358.
Bhasin, K. K., Blake, N. M., and Kirk, R. L., 1979, The frequency of private electrophoretic variants in Australian aborigines and indirect estimates of mutation rate, *Am. J. Hum. Genet.* 31:731–740.
Boaz, N. T., 1979, Early hominid population densities: New estimates, *Science* 206:592–595.
Bonnell, M. L., and Selander, R. K., 1974, Elephant seals: Genetic variation and near extinction, *Science* 184:908–909.
Brookbanks, H. E., 1965, Ecology and distribution of the pikas of Washington and Alaska, *Am. Midl. Nat.* 73:299–355.
Brown, A. J. L., and Ieh-Hermoncz, D., 1981, Evolution of the 87A and 87C heat-shock loci in *Drosophila*, *Nature* 290:677–682.
Brown, W. M., 1980, Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis, *Proc. Natl. Acad. Sci. USA* 77:3605–3609.
Brown, W. M., 1981, Mechanisms of evolution in animal mitochondrial DNA, *Ann. N. Y. Acad. Sci.* 361:119–134.
Bruce, E. J., and Ayala, F. J., 1979, Phylogenetic relationship between man and the apes: Electrophoretic evidence, *Evolution* 33:1040–1056.
Bryant, E. H., 1974a, On the adaptive significance of enzyme polymorphism in relation to environmental variability, *Am. Nat.* 108:1–19.
Bryant, E. H., 1974b, An addendum on the statistical relationship between enzyme polymorphisms and environmental variability, *Am. Nat.* 108:698–701.
Cameron, D. G., and Vyse, E. R., 1978, Heterozygosity in Yellowstone Park elk, *Cervus canadensis*, *Biochem. Genet.* 16:651–657.
Carson, H. L., 1970, Chromosome tracers of the origin of species, *Science* 168:1414–1418.
Carson, H. L., 1976, Inference of the time of origin of some *Drosophila* species, *Nature* 259:395–396.
Castrodeza, C., 1979, Non-progressive evolution, the Red Queen hypothesis, and the balance of nature, *Acta Biotheor.* 28:11–18.
Caveli-Sforza, L. L., and Bodmer, W. F., 1971, *The Genetics of Human Populations*, Freeman, San Francisco.
Chakraborty, R., Fuerst, P. A., and Nei, M., 1977, A comparative study of genetic variation within and between populations under the neutral mutation hypothesis and the model of occasionally advantageous mutations, *Genetics* 86:s98–s111.
Chakraborty, R., Fuerst, P. A., and Nei, M., 1978, Statistical studies on protein polymorphism in natural populations. II. Gene differentiation between populations, *Genetics* 88:367–390.

Chakraborty, R., Fuerst, P. A., and Nei, M., 1980, Statistical studies on protein polymorphism in natural populations. III. Distribution of allele frequencies and the number of alleles per locus, *Genetics* 94:1079–1093.
Christiansen, F. B., 1974, Sufficient conditions for protected polymorphism in a subdivided population, *Am. Nat.* 108:157–166.
Cody, M. L., 1974, *Competition and the Structure of Bird Communities*, Princeton University Press, Princeton.
Cormack, R. M., 1973, Common sense estimates from capture-recapture studies, in: *The Mathematical Theory of the Dynamics of Biological Populations* (M. S. Bartlett and R. W. Horns, eds.), pp. 225–234, Academic Press, London.
Crumpacker, D. W., and Williams, J. S., 1973, Density, dispersion, and population structure in *Drosophila pseudoobscura*, *Evol. Monogr.* 43:699–538.
Dagnot, R. W., Ramsey, F. R., Smith, M. H., and Urbston, D. F., 1979, Demographic differences in contiguous populations of white-tailed deer, *J. Wildlife Management* 43:889–898.
Davenport, L. B., Jr., 1966, Structure of two *Peromyscus polionotus* populations in old-field ecosystems at the AEC Savannah River Plant, *J. Mammal.* 45:95–113.
Dempster, E. R., 1955, Maintenance of genetic heterogeneity, *Cold Spring Harbor Symp. Quant. Biol.* 20:25–32.
Dixon, A. F., 1981, *The Natural History of the Gorilla*, Columbia University Press, New York.
Dobzhansky, Th., and Spassky, B., 1947, Evolutionary changes in laboratory cultures of *Drosophila pseudoobscura*, *Evolution* 1:191–216.
Douglas, C. L., 1969, Comparative ecology of Peromyscus mice and deer mice in Mesa Verde National Park, Colorado, *Univ. Kansas Publ. Mus. Nat. Hist.* 18:421–504.
Ewens, W. J., and Gillespie, J. H., 1974, Some simulation results for the neutral allele model, with interpretations, *Theor. Popul. Biol.* 6:35–57.
Feinestein, J., 1976, The theoretical population genetics of variable selection and migration, *Annu. Rev. Genet.* 10:253–280.
Flint, R. F., 1976, Pleistocene epoch, in: *Encyclopedia Britannica*, Vol. 14, pp. 558–569, University of Chicago Press, Chicago.
Fondevilla, A., and Carson, H. L., 1978, Spatial distribution and dispersal in a population of *Drosophila*, *Am. Nat.* 112:365–380.
Fox, L. R., and Morrow, P. A., 1981, Specialization: Species property or local phenomenon?, *Science* 211:887–893.
Fuerst, P. A., and Ferrell, R. E., 1980, The stepwise mutation model: An experimental evaluation utilizing hemoglobin variants, *Genetics* 94:185–201.
Fuerst, P. A., Chakraborty, R., and Nei, M., 1977, Statistical studies on protein polymorphism in natural populations. I. Distribution of single locus heterozygosity, *Genetics* 86:455–483.
Fukuyma, D. J., 1976, Food plant specialization and environmental predictability in Lepidoptera, *Am. Nat.* 110:285–292.
Garlside, D. F., Dessauer, H. C., and Jonnes, T., 1977, Genetic homozygosity in an ancient reptile (*Alligator mississippiensis*), *Biochem. Genet.* 15:655–663.
Gaur, P. K., Lichtwardt, R. W., and Huurnick, J. L., 1981, Isozyme variation among soil isolates of *Histoplasma capsulatum*, *Exp. Mycol.* 5:69–77.
Gillespie, J. H., 1973, Natural selection with varying selection coefficients—A haploid model, *Genet. Res.* 21:115–120.
Gillespie, J. H., 1977, A general model to account for enzyme variation in natural populations. II. Multiple alleles, *Evolution* 31:85–90.

Lazell, J. D., 1972, The anoles (Sauria, Iguanidae) of the Lesser Antilles, *Bull. Mus. Comp. Zool.* 143:1–115.

Ledig, F. T., and Conkle, M. T., 1983, Gene diversity and genetic structure in a narrow endemic, Torrey pine (*Pinus torreyana* Parry ex Carr.), *Evolution* 37:79–85.

Lemaunier, F., and Ashburner, M., 1976, Relationships within the *melanogaster* species subgroup of the genus *Drosophila* (Sophophora). II. Phylogenetic relationship between six species based upon polytene chromosome banding sequences, *Proc. R. Soc. London Ser. B* 193:275–294.

Lester, L. J., and Selander, R. K., 1979, Population genetics of haplodiploid insects, *Genetics* 92:1329–1345.

Levene, H., 1953, Genetic equilibrium when more than one ecological niche is available, *Am. Nat.* 87:331–333.

Levin, B. R., 1982, Periodic selection, infectious gene exchange and the genetic structure of *E. coli* populations, *Genetics* 99:1–23.

Levin, D. A., Ritter, K., and Ellstrand, N. C., 1979, Protein polymorphism in the narrow endemic *Ornithera organensis*, *Evolution* 33:534–542.

Levins, R., 1968, *Evolution in Changing Environments*, Princeton University Press, Princeton.

Lewontin, R. C., 1974, *The Genetic Basis of Evolutionary Change*, Columbia University Press, New York.

Lewontin, R. C., and Hubby, J. L., 1966, A molecular approach to the study of genetic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*, *Genetics* 54:595–609.

Li, W.-H., 1978, Maintenance of genetic variability under the joint effect of mutation, selection, and random genetic drift, *Genetics* 90:349–382.

Li, W.-H., and Nei, M., 1977, Persistence of common alleles in two related populations or species, *Genetics* 86:901–914.

Li, W.-H., Gojobori, T., and Nei, M., 1981, Pseudogenes as a paradigm of neutral evolution, *Nature* 292:237–239.

McCommas, S. A., 1983, A taxonomic approach to evaluation of the charge state model using twelve species of sea anemone, *Genetics* 103:741–752.

McDonald, J. F., and Ayala, F. J., 1974, Genetic response to environmental heterogeneity, *Nature* 250:572–574.

MacDonald, P. D. M., and Smith, H. D., 1980, Mark-recapture estimation of salmon smolt runs, *Biometrics* 36:401–417.

McLeod, M. J., Wynes, D. L., and Guttman, S. I., 1980, Lack of biochemical evidence for hybridization between two species of darters, *Comp. Biochem. Physiol.* 67B:321–325.

McNease, L., and Joanen, T., 1978, Distribution and relative abundance of the alligator in Louisiana coastal marshes, *Proc. Annu. Conf. S.E. Assoc. Fish Wildlife Agencies* 32:182–186.

Malpica, J. M., and Briscoe, D. A., 1981, Effective population number estimates of laboratory populations of *Drosophila melanogaster*, *Experientia* 37:947–948.

Manlove, M. N., Smith, M. H., Hildebrand, H. O., Fuller, S. E., Johns, P. E., and Straney, D. O., 1978, Genetic subdivision in a herd of white-tailed deer as demonstrated by spatial shifts in gene frequencies, *Proc. Annu. Conf. S.E. Assoc. Game Fish Comm.* 30:487–492.

Martin, P. S., and Wright, H. E. (eds.), 1967, *Pleistocene Extinctions: The Search for a Cause*, Yale University Press, New Haven.

Maruyama, T., and Kimura, M., 1980, Genetic variability and effective population size when local extinction and recolonization of subpopulations are frequent, *Proc. Natl. Acad. Sci. USA* 77:6710–6714.

Maruyama, T., and Nei, M., 1981, Genetic variability maintained by mutation and overdominant selection in finite populations, *Genetics* 98:441–459.

Matsubayashi, K., and Sagishi, H., 1961, Microbiological and clinical examinations of Cynomolgus monkeys in Indonesia, *Kyoto Univ. Overseas Rep. Stud. Indonesian Macaque* 1:47–56.

Matthews, T. C., 1975, Biochemical polymorphism in populations of the Argentine toad *Bufo arenarum*, *Copeia* 1975:454–465.

Maynard Smith, J., 1974, *Models in Ecology*, Cambridge University Press, London.

Maynard Smith, J., 1976, A comment on the Red Queen, *Am. Nat.* 110:325–330.

Metcalf, R. A., Martin, J. C., and Whitt, G. S., 1975, Low level of genetic heterozygosity in Hymenoptera, *Nature* 257:792–794.

Milkman, R. D., 1967, Heterosis as a major cause of heterozygosity in nature, *Genetics* 55:493–495.

Milkman, R. D., 1976, Selection is the major determinant, *Trends Biochem. Sci.* 1:N152–N154.

Milton, J. B., and Koehn, R. K., 1975, Genetic organization and adaptive response of allozymes to ecological variables in *Fundulus heteroclitus*, *Genetics* 79:97–111.

Miyata, T., and Yasunaga, T., 1981, Rapidly evolving mouse α-globin-related pseudogene and its evolutionary history, *Proc. Natl. Acad. Sci. USA* 78:450–453.

Morrow, K., 1975, Moose population characteristics and range use in the Augustow Forest, *Glad. Polskie* 33:693–706.

Mukai, T., and Cockerham, C. C., 1977, Spontaneous mutation rates at enzyme loci in *Drosophila melanogaster*, *Proc. Natl. Acad. Sci. USA* 74:2514–2517.

Neel, J. V., and Robertson, E. D., 1978, Indirect estimates of mutation rates in tribal Amerindians, *Proc. Natl. Acad. Sci. USA* 75:5585–5588.

Nei, M., 1971, Fertility excess necessary for gene substitution in regulated populations, *Genetics* 68:169–184.

Nei, M., 1975, *Molecular Population Genetics and Evolution*, *North-Holland, Amsterdam.*

Nei, M., 1976a, Comments on "Stochastic selection and the maintenance of genetic variation" by D. L. Hartl and R. D. Cook, in: *Population Genetics and Ecology* (S. Karlin and E. Nevo, eds.), pp. 612–613, Academic Press, New York.

Nei, M., 1976b, The case of natural selection and the extent of enzyme polymorphism, *Trends Biochem. Sci.* 1:N147–N148.

Nei, M., 1977, Estimation of mutation rate from rare protein variants, *Am. J. Hum. Genet.* 29:225–232.

Nei, M., 1978, Estimation of average heterozygosity and genetic distance from a small number of individuals, *Genetics* 89:583–590.

Nei, M., 1980a, Stochastic theory of population genetics and evolution, in: *Vito Volterra Symposium on Mathematical Models in Biology* (C. Barigozzi, ed.), pp. 17–47, Springer-Verlag, Berlin.

Nei, M., 1980b, Protein polymorphism and the SAS-CFF model, *Genetics* 94:1085–1087.

Nei, M., 1983, Genetic polymorphism and the rate of mutation in evolution, in: *Evolution of Genes and Proteins* (M. Nei and R. K. Koehn, eds.), pp. 165–190, Sinauer, Sunderland, Massachusetts.

Nei, M., and Roychoudhury, A. K., 1974, Sampling variances of heterozygosity and genetic distance, *Genetics* 76:379–390.

Nei, M., and Roychoudhury, A. K., 1982, Genetic relationship and evolution of human races, *Evol. Biol.* 14:1–59.

Nei, M., and Tajima, F., 1981, Genetic drift and estimation of effective population size, *Genetics* 98:625–640.

Nei, M., and Yokoyama, S., 1976, Effects of random fluctuations of selection intensity on genetic variability in a finite population. *Jpn. J. Genet.* 51:355–368.

Nei, M., Maruyama, T., and Chakraborty, R., 1975, The bottleneck effect and genetic variability in populations. *Evolution* 29:1–10.

Nei, M., Chakraborty, R., and Fuerst, P. A., 1976a, Infinite allele model with varying mutation rate, *Proc. Natl. Acad. Sci. USA* 73:4164–4168.

Nei, M., Fuerst, P. A., and Chakraborty, R., 1976b, Testing the neutral mutation hypothesis by distribution of single locus heterozygosity. *Nature* 263:491–493.

Nei, M., Fuerst, P. A., and Chakraborty, R., 1978, Subunit molecular weight and genetic variability of proteins in natural populations. *Proc. Natl. Acad. Sci. USA* 75:3359–3362.

Nelson, K., and Hedgecock, D., 1980, Enzyme polymorphism and adaptive strategy in the decapod Crustacea. *Am. Nat.* 116:238–280.

Nevo, E., 1978, Genetic variation in natural populations: Patterns and theory. *Theor. Popul. Biol.* 13:121–177.

Nevo, E., 1979, Adaptive convergence and divergence in subterranean mammals. *Annu. Rev. Ecol. Syst.* 10:269–308.

Nevo, E., Perl, T., Beiles, A., Wool, D., and Zoller, U., 1980, Genetic structure as a potential monitor of marine pollution, *J. Etud. Pollutions* 3:61–68.

Nevo, E., Huth, O., and Beiles, A., 1982, Population structure and evolution in subterranean mole rats, *Evolution* 36:1281–1289.

Nozawa, K., Shotake, T., Ohkura, Y., Kitajima, M., and Tanabe, Y., 1975, Genetic variations within and between troops of *Macaca fuscata fuscata*, in: *Contemporary Primatology* (S. Kondo, M. Kawai, and A. Ehara, eds.), pp. 75–89, S. Karger, Basel.

Nozawa, K., Shotake, T., Ohkura, Y., and Tanabe, Y., 1977, Genetic variations within and between species of Asian macaques, *Jpn. J. Genet.* 52:15–30.

Nozawa, K., Shotake, T., Kawamoto, Y., and Tanabe, Y., 1982, Population genetics of Japanese monkeys. II. Blood protein polymorphisms and population structure, *Primates* 23:252–271.

O'Brien, S. J., Wildt, D. E., Goldman, D., Merril, C. R., and Bush, M., 1983, The cheetah is depauperate in genetic variation, *Science* 221:460–462.

Ohta, T., 1974, Mutational pressure as the main cause of molecular evolution and polymorphism, *Nature* 252:351–354.

Ohta, T., 1976, Role of very slightly deleterious mutations in molecular evolution and polymorphism. *Theor. Popul. Biol.* 10:254–275.

Ohta, T., 1977, Extension to the neutral mutation random drift hypothesis, in: *Molecular Evolution and Polymorphism* (M. Kimura, ed.), pp. 148–167, National Institute of Genetics, Mishima, Japan.

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

Pamilo, P., Vepsalainen, K., and Rosengren, R. 1975. Low allozymic variability in *Formica* ants. *Hereditas* 80:293–296.

Pamilo, P., Rosengren, R., Vepsalainen, K., Varvio-Aho, S.-L., and Pisarski, B., 1978a, Population genetics of *Formica* ants. I. Patterns of enzyme gene variation, *Hereditas* 89:233–248.

Pamilo, P., Varvio-Aho, S.-L., and Pekkarinen, A., 1978b, Low enzyme gene variability in Hymenoptera as a consequence of haplodiploidy, *Hereditas* 88:93–99.

Pecker, F., Blumenfeld, A., Lamedico, P., Gilbert, W., Kolnaiser, R., and Dodgson, J., 1980, Evolution of genes: The chicken preproinsulin gene, *Cell* 20:553–566.

Pflieger, W. L., 1975, *The Fishes of Missouri*, Missouri Department of Conservation, Jackson, Missouri.

Popper, K. R., 1959, *The Logic of Scientific Discovery*, Hutchinson, London.

Powell, J. R., 1971, Genetic polymorphisms in varied environments. *Science* 174:1035–1036.

Powell, J. R., 1975, Protein variation in natural populations of animals, *Evol. Biol.* 8:79–119.

Powell, J. R., and Taylor, C. E., 1979, Genetic variation in ecologically diverse environments, *Am. Sci.* 67:590–596.

Prakash, S., Lewontin, R. C., and Hubby, J. L., 1969, A molecular approach to the study of genetic heterozygosity in natural populations IV. Patterns of genetic variation in central, marginal and isolated populations of *Drosophila pseudoobscura*. *Genetics* 61:841–858.

Price, P. K., and Kennedy, M. L., 1980, Genetic relationships in the white-footed mouse, *Peromyscus leucopus*, and the cotton mouse, *Peromyscus gossypinus*, Am. Midl. Nat. 103:73–84.

Ramshaw, J. A. M., Coyne, L. A., and Lewontin, R. C., 1979, The sensitivity of gel electrophoresis as a detector of genetic variation, *Genetics* 93:1019–1037.

Redfield, J. A., Krebs, C. J., and Taitt, M. J., 1977, Competition between *Peromyscus maniculatus* and *Microtus townsendii* in grasslands of coastal British Columbia, *J. Anim. Ecol.* 46:607–616.

Rensch, B., 1959, *Evolution above the Species Level*, Columbia University Press, New York.

Reynolds, V., 1967, *The Apes*, Dutton, New York.

Rich, S. S., Bell, A. E., and Wilson, S. P., 1979, Genetic drift in small populations of *Trichoism*, *Evolution* 33:579–584.

Richardson, B. J., Rogers, P. M., and Hewitt, G. M., 1980, Ecological genetics of the wild rabbit in Australia. II. Protein variation in British, French and Australian rabbits and the geographical distribution of the variation in Australia, *Austr. J. Biol. Sci.* 33:371–383.

Richardson, R. H., 1974, Effects of dispersal, habitat selection and competition on a speciation pattern of *Drosophila* endemic to Hawaii, in: *Genetic Mechanism of Speciation in Insects* (M. J. D. White, ed.), pp. 140–164, Reidel, Sydney, Australia.

Ronald, K., and Dougan, J. L., 1982, The ice lover: Biology of the harp seal (*Phoca groenlandica*), *Science* 215:928–933.

Rubial, R., and Philbrotian, R., 1970, Eathermy and niche expansion in lizards, *Copeia* 4:643–655.

Russell, R. J., 1968, Evolution and classification of the pocket gophers of the subfamily Geomisae, *Univ. Kansas Publ. Mus. Nat. Hist.* 16:473–579.

Ryman, N., Beckman, G., Rhoun Petersen, G., and Reuterwall, C., 1977, Variability of red cell enzymes and genetic implications of management policies in Scandinavian moose (*Alces alces*), *Hereditas* 85:157–162.

Saura, A., Halkka, O., and Lokki, J., 1972, Enzyme gene heterozygosity in small island populations of *Philaenus spumarius* (L.) (Homoptera), *Genetica* 46:459–472.

Schmidtke, J., and Engel, W., 1980, Gene diversity in tunicate populations. *Biochem. Genet.* 18:501–508.

Schnell, G. D., and Selander, R. K., 1981, Environmental and morphological correlates of genetic variation in mammals, in: *Mammalian Population Genetics* (M. H. Smith and T. Toule, eds.), pp. 68–99, University of Georgia Press, Athens, Georgia.

Selander, R. K., 1975, Stochastic processes in genetic population structure, in: *Proceedings 8th International Conference in Numerical Taxonomy* (G. F. Estabrook, ed.), pp. 284–332, Freeman, San Francisco.

Selander, R. K., 1976, Genic variation in natural populations, in: *Molecular Evolution* (F. J. Ayala, ed.), pp. 21–45, Sinauer, Sunderland, Massachusetts.

Selander, R. K., and Hudson, R. O., 1976, Animal population structure under close inbreeding: The land snail *Rumina* in southern France, *Am. Nat.* 110:695–718.

Selander, R. K., and Kaufman, D. W., 1973a, Self-fertilization and genic population structure in a colonizing land snail, *Proc. Natl. Acad. Sci. USA* 70:1186–1190.

Selander, R. K., and Kaufman, D. W., 1973b, Genic variability and strategies of adaptation in animals, *Proc. Natl. Acad. Sci. USA* 70:1875–1877.

Selander, R. K., and Levin, B. R., 1980, Genetic diversity and structure in *Escherichia coli* populations, *Science* 210:545–547.

Selander, R. K., and Whittam, T. S., 1983, Protein polymorphism and the genetic structure of populations, in: *Evolution of Genes and Proteins* (M. Nei and R. K. Koehn, eds.), pp. 89–114, Sinauer, Sunderland, Massachusetts.

Selander, R. K., Smith, M. H., Yang, S. Y., Johnson, W. E., and Gentry, J. B., 1971, Biochemical polymorphism and systematics in the genus *Peromyscus*. I. Variation in the old-field mouse (*Peromyscus polionotus*), *Univ. Texas Publ. Stud. Genet.* 6:49–90.

Selander, R. K., Kaufman, D. W., Baker, R. J., and Williams, S. L., 1974, Genic and chromosomal differentiation in pocket gophers of the *Geomys bursarius* group, *Evolution* 28:557–564.

Serafinski, W., 1969, Reproduction and dynamics of enone (*Alces alces* L.) population in the Kampinos National Park, *Ekol. Polska* 17:709–748.

Sheppe, W., 1966, Determinants of home range in the deer mouse, *Peromyscus leucopus*, *Proc. Calif. Acad. Sci.* 34:377–418.

Simonsen, V., 1982, Electrophoretic variation in large mammals. II. The red fox, *Vulpes vulpes*, the stoat, *Mustela erminea*, the weasel, *Mustela nivalis*, the pole cat, *Mustela putorius*, the pine marten, *Martes martes*, the beech marten, *Martes foina*, and the badger, *Meles meles*, *Hereditas* 96:299–303.

Smith, M. H., Gerlen, C. T., Jr., and Ramsey, P. R., 1975, Genic heterozygosity and population dynamics in small mammals, in: *Isozymes IV. Genetics and Evolution* (C. L. Markert, ed.), pp. 85–102, Academic Press, New York.

Snyder, T. P., 1974, Lack of allozymic variability in three bee species, *Evolution* 28:687–689.

Soulé, M., 1976, Allozyme variation: Its determinants in space and time, in: *Molecular Evolution* (F. J. Ayala, ed.), pp. 60–77, Sinauer, Sunderland, Massachusetts.

Soulé, M., and Yang, S. Y., 1973, Genetic variation in side-blotched lizards in the gulf of California, *Evolution* 27:593–600.

Southwick, C. H., and Cośgun, F. C., Jr., 1972, Population studies of Malaysian primates, *Primates* 13:1–18.

Spassky, B., Richmond, R. C., Pérez-Salas, S., Pavlovsky, O., Mourão, C. A., Hunter, A. S., Heusinger, H., Dobzhansky, T., and Ayala, F. J., 1971, Geographic variation of the sibling species related to *Drosophila willistoni*, and of the semispecies of the *Drosophila paulistorum* complex, *Evolution* 25:129–143.

Spieth, P. T., 1975, Population genetics of allozyme variation in *Neurospora intermedia*, *Genetics* 80:785–805.

Stahl, G., 1983, Differences in the amount and distribution of genetic variation between natural populations and hatchery stocks of Atlantic salmon, *Aquaculture* 33:23–32.

Stebbins, R. C., 1962, *Amphibians and Reptiles of Western North America*, McGraw-Hill, New York.

Steiner, W. W. M., 1979a, Genetic variation in Hawaiian *Drosophila*. VII. Seasonally dependent gene changes in *Drosophila silvica*, *Evolution* 33:543–562.

Steiner, W. W. M., 1979b, Genetic variation in Hawaiian *Drosophila*. VIII. Heterozygosity and genetic changes in isolated populations of *D. ruptrichaeta*, *Biochem. Genet.* 17:645–664.

Stone, W. S., Guest, W. C., and Wilson, F. D., 1960, The evolutionary implications of the cytological polymorphism and phylogeny of the *virilis* group of *Drosophila*, *Proc. Natl. Acad. Sci. USA* 46:390–361.

Strobeck, C., 1975, Selection in a fine-grained environment, *Am. Nat.* 109:419–425.

Sved, J. A., Reed, T. E., and Bodmer, W. F., 1967, The number of balanced polymorphisms that can be maintained in a natural population, *Genetics* 55:469–481.

Takahata, N., 1981, Genetic variability and rate of gene substitution in a finite population under mutation and fluctuating selection, *Genetics* 98:427–440.

Takahata, N., and Kimura, M., 1979, Genetic variability maintained in a finite population under mutation and autocorrelated random fluctuation of selection intensity, *Proc. Natl. Acad. Sci. USA* 76:5812–5817.

Takahata, N., Ishii, K., and Matsuda, H., 1975, Effect of temporal fluctuation of selection coefficients on gene frequency in a population, *Proc. Natl. Acad. Sci. USA* 72:4541–4545.

Taylor, C. E., 1976, Genetic variation in heterogeneous environments, *Genetics* 83:807–894.

Tracey, M. L., Nelson, K., Hedgecock, D., Shleser, R. A., and Pressick, M. L., 1975, Biochemical genetics of lobsters: Genetic variation and the structure of American lobster (*Homarus americanus*) populations, *J. Fish. Res. Board Can.* 32:2091–2101.

Turner, J. R. G., Johnson, M. S., and Eanes, W. F., 1979, Contrasted modes of evolution in the same genome: Allozymes and adaptive change in *Heliconius*, *Proc. Natl. Acad. Sci. USA* 76:1924–1928.

Twitty, V. C., 1955, Field experiments on the biology and genetic relationships of the California species of *Triturus*, *J. Exp. Zool.* 129:129–148.

Twitty, V. C., 1966, *Of Scientists and Salamanders*, Freeman, San Francisco.

Twitty, V. C., Grant, D., and Anderson, O., 1967, Home range in relation to homing in the newt *Taricha rivularis* (Amphibia: Caudata), *Copeia* 1967:649–653.

Valentine, J. W., 1976, Genetic strategies of adaptation, in: *Molecular Evolution* (F. J. Ayala, ed.), pp. 78–94, Sinauer, Sunderland, Massachusetts.

Valentine, J. W., and Ayala, F. J., 1975, Genetic variation in *Fridelia hulli*, a deep-sea brachiopod, *Deep-Sea Res.* 22:37–44.

Van Valen, L., 1973, A new evolutionary law, *Evol. Theory* 1:1–30.

Van Valen, L., 1974, Two modes of evolution, *Nature* 252:298–300.

Varvio-Aho, S.-L., 1981, The effects of ecological differences on the amount of enzyme variation in Finnish water-strider (*Gerris*) species, *Hereditas* 94:35–39.

Voelker, R. A., Schaffer, H. E., and Mukai, T., 1980, Spontaneous allozyme mutations in *Drosophila melanogaster*: Rate of occurrence and nature of the mutants, *Genetics* 94:961–966.

Ward, R. D., 1977, Subunit size of enzymes and genetic heterozygosity in vertebrates, *Biochem. Genet.* 15:123–133.

Watanabe, T. K., and Kawanishi, M., 1976, Colonization of *Drosophila simulans* in Japan, *Proc. Jpn. Acad.* 52:191–194.

Webster, T. P., Soulé, M., and Yang, S. Y., 1972, Genetic variability and similarity in the *Anolis* lizards of Bimini, *Evolution* 26:513–535.

Wilk, C., 1973, In defense of naive pan-selectionism, *Am. Nat.* 107:23–34.

Wilson, A. C., Carlson, S. S., and White, T. J., 1977, Biochemical evolution, *Annu. Rev. Biochem.* 46:573–639.

Wood, J. W., 1982, Models of human evolution, *Science* 217:297–298.

Wright, S., 1931, Evolution in Mendelian populations, *Genetics* 16:97–159.

Wright, S., 1938, Size of population and breeding structure in relation to evolution, *Science* 178:56–58.

Wright, S., 1969, *Evolution and the Genetics of Populations*, University of Chicago Press, Chicago.

Yamazaki, T., 1981, Genetic variabilities in natural populations of haploid plant, *Cornuphalium conicum*. I. The amount of heterozygosity, *Jpn. J. Genet.* 56:373–383.

Yang, S. Y., Soulé, M., and Gorman, G. C., 1974, *Anolis* lizards of the eastern Caribbean: A case study in evolution. I. Genetic relationship, phylogeny and colonization sequence of the *roquet* group, *Syst. Zool.* 23:387–399.

Zouros, E., 1976, Hybrid molecules and the superiority of the heterozygote, *Nature* 269:227–229.

Zouros, E., 1979, Mutation rates, population sizes and amounts of electrophoretic variation of enzyme loci in natural populations, *Genetics* 93:623–646.

Gillespie, J. H., 1978. A general model to account for enzyme variation in natural populations. V. The SAS-CFF model. *Theor. Popul. Biol.* 14:1–45.

Gillespie, J. H., 1979. Molecular evolution and polymorphism in a random environment. *Genetics* 92:757–754.

Glover, D. G., Smith, M. H., Ames, L., Bude, J., and Dubach, J. M., 1977. Genetic variation in pika populations. *Can. J. Zool.* 55:1841–1845.

Gooch, J. L., and Schopf, T. J. M., 1972. Genetic variability in the deep sea: Relation to environmental variability. *Evolution* 26:545–552.

Gorman, G. C., and Harwood, R., 1977. Notes on population density, vagility, and activity patterns of the Puerto Rican grass lizard *Anolis pulchellus* (Reptilia, Lacertilia, Iguanidae). *J. Herpetol.* 11:361–365.

Gorman, G. C., Soulé, M., Yang, S. Y., and Nevo, E., 1975. Evolutionary genetics of insular Adriatic lizards. *Evolution* 29:52–71.

Haldane, J. B. S., and Jayakar, S. D., 1963. Polymorphism due to selection of varying direction. *J. Genet.* 58:237–242.

Hall, E. R., and K. R. Kelson, 1959. *The Mammals of North America*, Ronald Press, New York.

Halliday, R. B., 1981. Heterozygosity and genetic distance in sibling species of meat ants (*Iridomyrmex purpureus* group). *Evolution* 35:254–342.

Hamrick, J. L., Linhart, Y. B., and Mitton, J. B., 1979. Relationship between life history characteristics and electrophoretically detectable genetic variation in plants. *Annu. Rev. Ecol. Syst.* 10:173–200.

Harris, H., 1966. Enzyme polymorphism in man. *Proc. R. Soc. London Ser. B* 164:298–310.

Harris, H., Hopkinson, D. A., and Edwards, Y. H., 1977. Polymorphism and the subunit structure of enzymes: A contribution to the neutralist–selectionist controversy. *Proc. Natl. Acad. Sci. USA* 74:698–701.

Hecht, K., 1976. Morphological transformation, the fossil record, and the mechanisms of evolution: A debate. I. The statement and the critique. *Evol. Biol.* 7:295–305.

Hedgecock, D., 1978. Population subdivision and genetic divergence in the red-bellied newt, *Taricha rivularis*. *Evolution* 32:271–286.

Hedrick, P. W., 1974. Genetic variation in a heterogeneous environment. I. Temporal heterogeneity and the absolute dominance model. *Genetics* 78:757–770.

Hedrick, P. W., Ginevan, M. E., and Ewing, E. P., 1976. Genetic polymorphism in heterogeneous environments. *Annu. Rev. Ecol. Syst.* 7:1–32.

Ishii, K., and Charlesworth, B., 1977. Associations between allozyme loci and gene arrangements due to hitch-hiking effects of new inversions. *Genet. Res.* 30:93–100.

Jaenike, J., Parker, E. D., and Selander, R. K., 1980. Clonal niche structure in the parthenogenetic earthworm *Octolasion tyrtaeum*. *Am. Nat.* 116:196–203.

Jensen, L., and Pollak, E., 1969. Random selective advantages of a gene in a finite population. *J. Appl. Prob.* 6:19–37.

Johansson, N., 1981. General problems in Atlantic salmon rearing in Sweden, in: *Fish Gene Pools* (N. Ryman, ed.). *Ecol. Bull (Stockholm)* 34:75–83.

Johnson, G. B., 1976a. Genetic polymorphism and enzyme function, in: *Molecular Evolution* (F. J. Ayala, ed.). pp. 46–59. *Sinauer, Sunderland*, Massachusetts.

Johnson, J. S., and Hood, W. B., 1976. Dispersal of desert-adapted *Drosophila*: The Saguaro-breeding *D. nigrospiracula*. *Am. Nat.* 110:629–651.

Jones, J. S., Bryant, S. H., Lewontin, R. C., Moore, J. A., and Prout, T., 1981. Gene flow and the geographical distribution of a molecular polymorphism in *Drosophila pseudoobscura*. *Genetics* 98:157–178.

Jukes, T. H., 1980. Silent nucleotide substitutions during evolution. *Naturwissenschaften* 67:534–539.

Karlin, S., and Levikson, B., 1974. Temporal fluctuations in selection intensities: Case of small population size. *Theor. Popul. Biol.* 6:383–412.

Kawamoto, Y., Nozawa, K., and Ischak, T. M., 1981. Genetic variability and differentiation of local populations in the Indonesian crab-eating macaque (*Macaca fascicularis*). *Kyoto Univ. Overseas Rep. Stud. Indonesian Macaque* 1:15–39.

Kimura, M., 1968a. Evolutionary rate at the molecular level. *Nature* 217:624–626.

Kimura, M., 1968b. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. *Genet. Res.* 11:247–269.

Kimura, M., 1969. The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics* 61:893–903.

Kimura, M., 1977. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature* 267:275–276.

Kimura, M., 1979. Model of effectively neutral mutations in which selective constraint is incorporated. *Proc. Natl. Acad. Sci. USA* 76:3440–3444.

Kimura, M., 1981. The neutral theory of molecular evolution, in: *Evolution of Genes and Proteins* (M. Nei and R. K. Koehn, eds.). pp. 208–233. Sinauer, Sunderland, Massachusetts.

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

Kimura, M., and Maruyama, T., 1971. Pattern of neutral polymorphism in a geographically structured population. *Genet. Res.* 18:125–131.

Kimura, M., and Ohta, T., 1971. Protein polymorphism as a phase of molecular evolution. *Nature* 229:467–469.

Kimura, M., and Ohta, T., 1973. Mutation and evolution at the molecular level. *Genetics* 73(suppl.):19–35.

King, J. A., 1968. *Biology of Peromyscus*, American Society Mammalogists Special Publication, Vol. 2.

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

King, M.-C., and Wilson, A. C., 1975. Evolution at two levels in humans and chimpanzees. *Science* 188:107–116.

Koehn, R. K., 1969. Esterase heterogeneity: Dynamics of a polymorphism. *Science* 163:943–944.

Koehn, R. K., and Eanes, W. F., 1977. Subunit size and genetic variation of enzymes in natural populations of *Drosophila*. *Theor. Popul. Biol.* 11:330–341.

Koehn, R. K., and Eanes, W. F., 1978. Molecular structure and protein variation within and among populations. *Evol. Biol.* 11:79–100.

Koehn, R. K., Zera, A. J., and Hall, J. G., 1983. Enzyme polymorphism and natural selection, in: *Evolution of Genes and Proteins* (M. Nei and R. K. Koehn, eds.). pp. 115–136. Sinauer, Sunderland, Massachusetts.

Kojima, K., and Yarbrough, K. M., 1967. Frequency dependent selection at the esterase 6 locus in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 57:645–649.

Krzakowa, M., and Szweykowski, J., 1979. Isozyme polymorphism in natural populations of a liverwort, *Plagiochila asplenioides*. *Genetica* 93:711–719.

Lavigne, D. M., Bogart, J. P., Duvener, R. G. H., Dawzmann, R. G., Barchard, W. W., and Earle, M., 1982. Genetic variability in northwest Atlantic harp seals, *Pagophilus groenlandicus*, Canadian Atlantic Fisheries Scientific Advisory Committee Working Paper 78/45.

Lawlor, T. E., 1965. The Yucatan deer mouse, *Peromyscus yucatanicus*, *Univ. Kansas Publ. Mus. Nat. Hist.* 16:421–458.