A DIRECT ASSESSMENT OF THE ROLE OF GENETIC DRIFT IN DETERMINING ALLELE FREQUENCY VARIATION IN POPULATIONS OF *EUPHYDRYAS EDITHA*

LAURENCE D. MUELLER,* BRUCE A. WILCOX,** PAUL R. EHRLICH,**
DAVID G. HECKEL† AND DENNIS D. MURPHY**

**Department of Zoology and Program in Genetics, Washington State University, Pullman, Washington 99164; **Department of Biological Sciences, Stanford University, Stanford, California 94305; † Department of Biological Sciences, Clemson University, Clemson, South Carolina 29631*

Manuscript received March 9, 1984
Revised copy accepted March 2, 1985

ABSTRACT

Estimates of allele frequencies at six polymorphic loci were collected over eight generations in two populations of *Euphydryas editha*. We have estimated, in addition, the effective population size for each generation for both populations with results from mark-recapture and other field data. The variation in allele frequencies generated by random genetic drift was then studied using computer simulations and our direct estimates of effective population size. Substantial differences between observed values and computer-generated expected values assuming drift alone were found for three loci (*Got, Hk, Pgi*) in one population. These observations are consistent with natural selection in a variable environment.

FOR nearly 20 yr empirical population genetics has grappled with the problem posed by the large amounts of genetic variation found in natural populations. Evidence supporting the neutral theory of molecular evolution has usually come from comparisons of data to certain stationary properties of neutral models. Properties of natural populations such as the expected heterozygosity, variance in heterozygosity and frequency spectrum are seen to be in reasonable accord with predictions from the neutral model (FUERST, CHAKRABORTY and NEI 1977; CHAKRABORTY, FUERST and NEI 1980). Those workers who feel that most polymorphisms are maintained by natural selection have taken a different approach to the problem and have produced a number of independent investigations of particular polymorphic loci in several organisms (MCDONALD and AYALA 1978; PLACE and POWERS 1979; BURTON and FELDMAN 1983; WATT, 1983; WATT, CASSIN and SWAN 1983). Such studies aim to establish the biochemical and physiological differences between electrophoretic variants and then to relate these to pertinent ecological data to develop a coherent picture of the forces maintaining the particular polymorphism.

Recently, many of the statistical tests of the neutral theory have been questioned. ROTHMAN and TEMPLETON (1980) show that EWENS' (1972) sampling

theory and tests based on it (WATTERSON 1977) rely on a special assumption about the distribution of the number of copies of neutral alleles. When this assumption is relaxed the resulting neutral models can yield frequency spectra and homozygosities similar to those expected from heterosis. GILLESPIE (1979) has shown that the infinite allele model and his model of selection in a random environment have the same stationary distribution. Therefore, the agreement between empirical observations and the infinite allele model noted by FUERST, CHAKRABORTY and NEI (1977) can be used with equal strength to support Gillespie's model of natural selection. Another problem with these tests is the assumption that the study population has reached its stationary distribution. A population undergoing a variety of drift processes approaches its stationary distribution at a rate governed by the second largest eigenvalue which equals $1 - \text{const}/2N$, where N is the effective population size. If N is even moderately large this rate can be extremely slow. PERLOW (1979) has shown that the elegant test of the infinite allele model devised by WATTERSON (1977) is sensitive to the stationary assumption.

Several tests of the neutral theory have been devised which use observations of allele frequencies over several generations. The first such test was described by LEWONTIN and KRAKAUER (1973). Their test was based on calculating standardized variances of allele frequency changes over one or several generations in the same population. In their initial work, however, Lewontin and Krakauer did not properly account for the sampling process in allele frequency estimation; corrections for this have been noted since by PAMILO and VARVIO-AHO (1980), NEI and TAJIMA (1981) and POLLAK (1983). Yet, to use this method as a test of neutrality, the sampling distribution of the standardized variance, $\hat{F}$, must be known. NEI and TAJIMA (1981) have shown that $n\hat{F}/E(\hat{F})$, where n is the number of independent alleles (or loci) used to estimate F, has a χ^2 distribution with n degrees of freedom. However, we do not generally know $E(\hat{F})$ and it is not clear that, if this quantity is replaced with a small sample estimate, the χ^2 distribution will be preserved. Furthermore, even when $E(\hat{F})$ is known the χ^2 approximation becomes worse if $\hat{F}$ is based on observations separated by several generations. Finally, it is not at all clear what will happen to the distribution of $n\hat{F}/E(\hat{F})$ if more than one generation has passed and the effective population size does not remain constant over time.

Other tests of neutrality that utilize temporal data have been described by FISHER and FORD (1947), SCHAFFER, YARDLEY and ANDERSON (1977) and WATTERSON (1982). However, both of these tests require precise knowledge of the effective population size. Unfortunately, this is seldom known for natural populations. At best an estimate of total population size may be available from mark-recapture techniques.

In this paper we describe a test of the neutral theory that depends on the variation in allele frequencies utilizing genetic data and estimates of effective population size collected over many generations. The final test statistics are the result of extensive computer simulations and thus are useful only for our data sets. This loss of generality is more than compensated for by a test procedure which is relatively free of simplifying assumptions. This and other

advantages of computer intensive statistics have been discussed by DIACONIS and EFRON (1983).

We also investigate the ability of our test procedure to detect modes of natural selection that increase the variation in allele frequencies. We find that our methods are able to detect selection that increases allele frequency variation, such as selection in a random environment.

MATERIALS AND METHODS

Genetic and ecological data have been collected for the univoltine checkerspot butterfly *Euphydryas editha*, an endemic of serpentine grasslands in the San Francisco Bay Area. Two populations, known from areas H and C, were studied at Jasper Ridge Biological Preserve on the Stanford University campus, California, from 1973 and 1980 (EHRLICH *et al.* 1975; EHRLICH and MURPHY 1981). Adults fly and oviposit from approximately mid-March through late April. Thus, the life cycle is composed of truly discrete nonoverlapping generations.

Genetic data: Electrophoretic data were collected each year for six polymorphic loci: glutamate-oxaloacetate transaminase (*Got*), hexokinase (*Hk*), β-hydroxybutyric acid dehydrogenase (*Hbdh*), phosphoglucose-isomerase (*Pgi*), phosphoglucomutase (*Pgm*) and superoxide dismutase (*Sod*). Details of the electrophoretic methods are given by MCKECHNIE, EHRLICH and WHITE (1975).

Estimating adult numbers: Population size data were also collected each year using the mark-recapture methods of BRUSSARD, EHRLICH and SINGER (1974) and EHRLICH and DAVIDSON (1960). In brief, each year's study of a population consisted of several separate census visits to the site during which adult individuals were captured, individually marked if not already marked and released. These data were used to estimate the size of the population according to the Jolly-Seber model (SEBER 1973, chapter 5, correction for bias not employed). Two features of the Euphydryas census data make the Jolly-Seber method particularly appropriate here: (1) The population is open, *i.e.*, the adult population size is constantly changing in time as newly eclosed adults augment it and dying or emigrating adults diminish it. (2) Both weather (flight) conditions and sampling effort vary from time to time, and therefore, capture probabilities are potentially different for each separate census visit. The Jolly-Seber model is a widely used method that takes account of both of these features of the population. It provides estimates of N_i, the population size at census visit i, for all census visits except the first one. It also allows estimation of B_i, the number of new adults entering the population between samples i and $i + 1$. For our estimate of total population size B for a site at a given year, we summed all of the B_i corresponding to that year and site. The variance of the estimate of B was obtained from Jolly's (see SEBER 1973, chapter 5) formulas for the variances and covariances of the B_i and the general formula for the variance of a sum of random variables.

It became necessary to employ a slight modification of the Jolly-Seber method when m_i, the number of previously marked butterflies caught in the ith sample, equaled zero. Under these conditions, the estimate of N_i becomes undefined. Rather than to arbitrarily employ some bias-correcting alteration in the formula for N_i, we chose to pool the animals of sample i with those of sample $i - 1$ under these conditions. Also, the Jolly-Seber method does not give an estimate of N_1 (= B_0), the number of animals that joined the population before the first census visit and survived to that time. Usually, we have not attempted to estimate N_1; instead, we ignore it and confine our problem to estimating the total number of individuals that joined the population during the time spanned by the census visits. Because there are always *some* individuals present at the time of the first visit, we are thereby underestimating the total size of the population, but this is only a slight underestimation because the first census is placed early in the flight season and most of the butterflies eclose during the time spanned by the census. On two occasions (JRH 1974, 1978) our data pooling resulted in a large number of individuals being present at the first census. In these two cases we estimated the total number present on day 1 using the number captured and the probability of capture for the second census date. This number was added to the standard estimate of total population described previously.

Adult numbers and effective population size: Since males and females can be identified in the field,

separate estimates of the number of males B_m and females B_f are usually available. In some years the sex ratio was quite biased with females making up only 10–20% of the population. Unfortunately, the variance of the estimates of the separate male and female numbers is much larger than the estimate of total population size with males and females pooled. This is important in the decision to use separate B_m and B_f, or just pooled B assuming a 1:1 sex ratio, to estimate N_e. Thus, even though we expect an estimate of N_e that utilizes the separate sexes to be less biased, it is liable to have much larger variance. This, of course, will make it more difficult to detect deviations from the drift model. As a consequence we have examined four estimators N_e of the effective population size. In APPENDIX 1 we derive an inbreeding effective population size that takes into account unequal sex ratio, variance in female reproductive success and changing population size. Four versions of this effective population size are considered which make various assumptions about the sex ratio and population growth. As noted by EWENS (1982), the inbreeding effective size will not necessarily be the same as the variance or eigenvalue effective population size; however, its derivation appears to be much simpler in the present setting. The estimator with the smallest mean squared error was used as N_e in the computer simulations. This procedure does introduce a bias in our results which we consider in the discussion.

In a few years females were so rare that not enough were recaptured to allow a direct estimate from the Jolly-Seber procedure. In these years we estimated the frequency of females in the population, c, from their frequency in the captured pool of animals. Since males are almost twice as likely as females to be captured (EHRLICH, LAUNER and MURPHY 1984), c is probably biased downward. The effect of this will be estimates of N_e that are too small. We will discuss the effect of this bias on our test statistics later. The total number of females N_f was then estimated as $B \cdot c$, and the total number of males N_m as $B \cdot (1 - c)$. The variance of these estimates were obtained via the delta method (BISHOP, FIENBERG and HOLLAND 1975; pp. 486–488).

Variance in offspring number: Another factor that has a significant effect on the effective population size is the variance in offspring number and survivorship which is substantial between females. This variability can be partitioned into at least three components. First, there is variability in the number of eggs laid per egg mass and the number of egg masses. Females may lay between 40 and 200 eggs per egg mass, and in the laboratory they may lay up to seven egg masses or more, although four is perhaps more typical in the field. Second, there are marked differences in survival of egg masses laid in wet and dry environments. *E. editha* usually lays eggs on green *Plantago erecta* and *Orthocarpus densiflorus*. However, these plants may undergo rapid senescence, and the larvae may hatch into a dry environment in which mortality is virtually 100% (SINGER and EHRLICH 1979). Third, there is variability in larval survival even within wet environments.

Assuming that these three factors act independently of each other, we can derive formulas for the mean and the variance of female offspring number (see APPENDIX 2 for details). Let the random variable W_i be the number of surviving offspring from a female's ith egg mass. The mean number of surviving offspring from egg mass i is $E(W_i)$, and the formula for it is found to be

$$E(W_i) = \lambda(1 - \delta)u_{x_i} = \hat{W}_i,$$

where u_{x_i} is the mean number of eggs in egg mass i, δ is the probability that the egg mass hatches into a dry environment and λ is the individual probability of survival from larva to adult in a wet environment. The variance in number of surviving offspring from egg mass i is $\mathrm{Var}(W_i)$, which is

$$\mathrm{Var}(W_i) = \hat{W}_i[1 - \lambda - \hat{W}_i + \lambda\sigma^2_{x_i}/u_{x_i} + \lambda u_{x_i}]$$

where $\sigma^2_{x_i}$ is the variance of the number of eggs in egg mass i. If it is assumed that the separate egg masses of a female are independent, the total variance in female fecundity $\mathrm{Var}(F)$ is just the sum of the $\mathrm{Var}(W_i)$ over egg masses.

The actual calculations of $\mathrm{Var}(F)$ make several assumptions. Although females may lay up to seven egg masses, our experience indicates that four egg masses may be more typical in the field. The mean number of eggs per mass and the variance are estimated from the work of MURPHY, LAUNER and EHRLICH (1983) which looked at female fecundity as a function of diet. We assume the diet of sugar water and 0.004 M amino acids to be most similar to the diets available on Jasper Ridge. We also assume that the probability of hatching into a dry environment (δ) is 0.71. SINGER and EHRLICH (1979) cite a range for δ at 0.27–0.83 for the year 1971. We have chosen a value

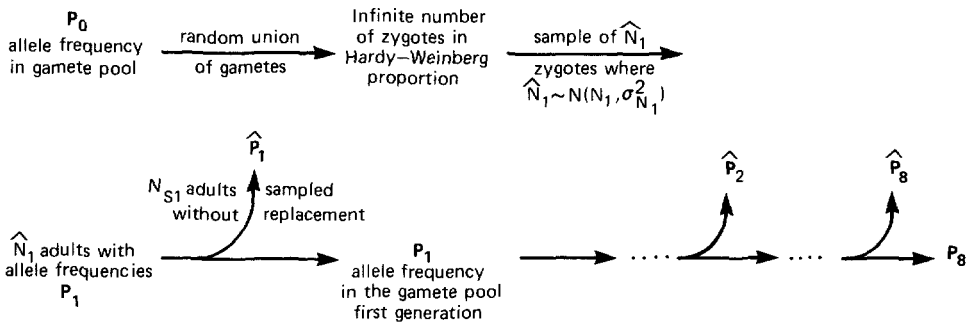

FIGURE 1.—The life cycle used in the computer simulation that incorporates random genetic drift and sampling of allele frequencies.

toward the higher end of this range (actually the value observed 1 wk into the 1971 flying season). The possible biases this choice might introduce is discussed later. Finally, λ is chosen such that each female at time t had, on average, $2N_{t+1}/N_t$ offspring in the next generation. For models I and III (see APPENDIX 1) it is assumed that $N_{t+1} = N_t$. Models II and IV allow for population increase and decrease and, therefore, use the actual estimates of total population size N_t and N_{t+1}. Carrying out the calculations produces an estimate for the variance in female fecundity of 5.49 when $N_{t+1} = N_t$.

Comparison with indirect estimates of effective population size: Recently, there have been several proposals to estimate effective population size indirectly from observations of allele frequency variation over time (NEI and TAJIMA 1981; POLLAK 1983). These methods assume that variation in allele frequencies over time is due solely to genetic drift and sampling for estimation purposes. Since we hope to use our estimates of effective population size to determine whether drift alone can account for the observed variation, it would seem inappropriate to use these methods to estimate effective population size. However, these methods consistently produce estimates of effective population size that are about one order of magnitude smaller than estimates obtained by several other methods (POLLAK 1983; BEGON, KRIMBAS and LOUKAS 1980). The same is true for our data. As we show in the next section Pollak's method of estimating effective population size is sensitive to the presence of variable natural selection. Thus, we feel justified in disregarding these estimates.

Computer simulations of Euphydryas population genetics: We now ouline the life cycle used in our computer simulations (Figure 1). The dynamics of a single locus with k alleles will be modeled. The frequency of the jth allele at time t is P_{jt}. The vector of all allele frequencies at time t is represented by $\mathbf{P}_t = (P_{1t}, \ldots, P_{k-1,t})^T$. At generation t drift occurs by sampling $\hat{N}_t$ zygotes from an infinite zygote pool. Since we only have estimates of the effective population size, $\hat{N}_t$ is a random variable chosen from a normal distribution with mean N_t and variance $\sigma^2_{N_t}$ which are given in Table 1. After the $\hat{N}_t$ adults have produced the gamete pool for the next generation, N_{st} adults are sampled without replacement to estimate the allele frequency vector $\mathbf{P}_t$. This estimate is denoted $\hat{\mathbf{P}}_t$. The N_{st} are known exactly and correspond to the actual sample sizes used to estimate allele frequencies in the various Euphydryas populations. After eight generations of drift the covariance matrix of the observations $\hat{\mathbf{P}}_1, \ldots, \hat{\mathbf{P}}_8$ was calculated. At this point, one iteration is complete.

A total of 1000 iterations were performed. Thus, there were 1000 covariance matrices for the drift-only process from which we made statistical inferences. All of the loci used in this survey were polymorphic, and there was no indication that any allele became fixed or lost in the eight generations of observations. Thus, if during the course of the computer simulation an allele was lost or fixed, that particular sequence of allele frequency vectors was ignored. Because of the small number of generations considered, mutation and migration can be ignored. The initial allele frequencies were set to the observed values in 1973.

For loci with just two alleles we simply followed the variance of the most common allele. The computer simulations provided information on the distribution of this variance. For loci with more

TABLE 1

Estimates of effective population size, the standard error, σN_e, and the method of estimation[a]

	Population					
	JRC			JRH		
Yr	N_e	σ_{N_e}	Method	N_e	σ_{N_e}	Method
1973	762.0	366.0	II	144.0	18.1	III
1974	164.0	23.9	II	27.5	2.6	II
1975	856.0	135.0	III	92.6	7.39	III
1976	4022.0	786.0	III	974.0	149.0	III
1977	374.0	6.2	I	179.0	18.1	I
1978	27.7	4.15	IV	16.8	1.54	IV
1979	182.0	20.7	IV	33.2	2.85	III
1980	195.0	32.1	III	90.2	13.2	I

[a] See APPENDIX 1.

than two alleles we have calculated the covariance matrix of **P**. Let the covariance matrix resulting from the ith iteration be $\mathbf{S}_i$. If we perform a total of m iterations, then we estimate $E(\mathbf{S}_i)$ as

$$E(\mathbf{S}_i) = \Sigma_0 = \sum_{i=1}^{m} \mathbf{S}_i/m.$$

For each sample covariance matrix we calculate the following statistic,

$$\lambda_i = (t-1)|\ln|\Sigma_0| - \ln|\mathbf{S}_i| + tr\mathbf{S}_i\Sigma_0^{-1} - (k-1)|,$$

where t is the number of generations (eight in our case) and k is the number of alleles at the particular locus. If the $\hat{\mathbf{P}}_i$ have a multivariate normal distribution then λ_i will be distributed as χ^2 $k(k-1)/2$ (MORRISON 1976, p. 248). There is no reason to assume that the $\hat{\mathbf{P}}_i$ has a multivariate normal distribution, so the purpose of the computer simulation is to describe empirically the distribution of the λ_i under the drift-only hypothesis.

RESULTS

The final estimates of effective population size with their standard errors are given in Table 1 for both populations JRC and JRH. In Table 2 we present the yearly allele frequencies and sample sizes for both JRC and JRH. Only the alleles used to calculate the covariance matrix are given in Table 2.

Following the scheme outlined in Figure 1 we have calculated the expected covariance matrices for each locus assuming the effective population sizes listed in Table 1. These expected covariance matrices and their observed values are given in Table 3. For loci with just two alleles we have computed the expected variance and the distribution of this variance. The cumulative distribution of this variance is given by $\phi(x)$ = prob. (the observed value of the variance is $\leq x$). Clearly, if this probability is very small, or very large, then our observation represents an unlikely event and we might question the underlying assumptions. In Table 3 we have listed the value of $\phi(x)$ for each two-allele locus (*Hk, Got, Sod*), where x = the observed variance of the most common allele.

For loci with more than two alleles our test statistic was λ_i. The value of λ_i gets larger as the observed covariance matrix gets either larger or smaller than

the expected value. In fact, when $\mathbf{S}_i = \Sigma_0$, $\lambda_i = 0$. From the computer simulations we have estimated the expected covariance matrix and the distribution of λ_i. For loci with more than two alleles (*Hbdh*, *Pgi*, *Pgm*) we have a cumulative distribution, $\phi(x)$ = prob. (observed value of λ_i is $\leq x$). Thus, only when $\phi(x)$ is very large (close to 1) will we question our underlying hypothesis. These values of $\phi(x)$ are also given in Table 3.

From Table 3 we see that there are three cases, out of 12, that show extreme test statistics, *e.g.*, $\phi(x) > 0.95$. These loci are *Got*, *Hk* and *Pgi* in population JRC. Each of these loci shows more variation than expected from drift alone. We expect that 12 independent applications of this test would produce 0.6 significant results by chance alone. In addition such chance deviations would be as likely to show less variation as more variation. Clearly, overdominance cannot account for these results. However, if selection varies over time, then we might expect more variation than drift alone would produce. To study this possibility we have examined the stochastic-additive-scale constant-fitness-function (SAS-CFF) model of GILLESPIE (1978). [We will review only a few important features of the symmetric SAS-CFF model here; GILLESPIE (1978) should be consulted for additional details.] We assume that the activity of the enzyme products coded for by genotype A_iA_i can be described by a random variable x_i. Heterozygotes are assumed to have an activity exactly intermediate between the two homozygotes; thus, the net activity of enzyme products of genotype A_iA_j is $(x_i + x_j)/2$. We assume further that $E(x_i) = 1$, $\text{Var}(x_i) = \sigma^2$ and $\text{cov}(x_ix_j) = \rho\sigma^2$. Finally, we assume there is some concave function that adequately describes the relationship between fitness and enzyme activity. For this study we have utilized the function $(1 + \alpha)x_i/(\alpha + x_i)$, where α is a constant. To conduct the appropriate computer simulations with the SAS-CFF model we must specify values for σ^2, ρ and α.

Once done, the previous simulation was carried out as before, but viability selection is assumed to act after drift and prior to the sampling of allele frequencies. After each round of eight generations of drift and selection, we compute the covariance matrix and compare it to the critical value which yields a 5% type I error rate. After approximately 1000 replications of this process, the frequency with which the drift plus selection process exceeds the critical value is computed. This is simply the power of the statistical procedure. The strength of selection in these computer simulations is summarized by computing the mean difference between the least and most fit genotype each generation. These results are given in Table 4.

As mentioned in the introduction, Pollak has described a method for estimating the effective population size from the variation in allele frequencies only. We have used this method in computer simulations with variable selection acting. The results are given in Table 5. It is evident that Pollak's method will be severely biased unless selection is quite weak. We should note that in the simulations that produced the results in Table 5 there was no variance due to allele frequency estimation N_s (sample size) = N_e; however, the standard deviation of these estimates was one to two orders of magnitude greater than the mean.

TABLE 2

Yearly fluctuations in allele frequencies at six loci for populations JRC and JRH

Locus	Population	Allele/sample size	Yr of sample							
			1973	1974	1975	1976	1977	1978	1979	1980
Hdbh	JRC	*n*	82	113	107	129	26	11	62	69
		1.00	0.77	0.81	0.76	0.77	0.79	1.00	0.88	0.80
		1.40	0.21	0.19	0.21	0.22	0.15	0.00	0.10	0.13
	JRH	*n*	50	51	83	85	48	56	60	35
		1.00	0.60	0.66	0.67	0.66	0.70	0.63	0.74	0.51
		1.40	0.36	0.33	0.33	0.32	0.29	0.38	0.26	0.49
Hk	JRC	*n*	79	104	104	119	32	11	62	62
		1.00	0.70	0.54	0.68	0.62	0.44	0.64	0.71	0.66
	JRH	*n*	50	50	78	85	38	56	60	35
		1.00	0.64	0.55	0.54	0.58	0.47	0.54	0.62	0.57
Got	JRC	*n*	83	118	107	129	51	11	62	69
		1.00	0.97	0.98	0.98	0.96	0.99	0.86	1.00	1.00
	JRH	*n*	52	52	83	86	50	56	58	35
		1.00	0.95	0.95	0.98	0.96	0.96	0.98	1.00	0.99
Pgi	JRC	*n*	80	114	108	128	27	11	62	68
		0.80	0.19	0.23	0.20	0.16	0.24	0.54	0.22	0.17
		1.00	0.47	0.51	0.51	0.47	0.35	0.23	0.53	0.55
		1.16	0.27	0.20	0.26	0.32	0.31	0.23	0.23	0.24
	JRH	*n*	51	52	77	86	43	56	59	35
		0.80	0.08	0.10	0.12	0.13	0.05	0.09	0.07	0.07
		1.00	0.50	0.49	0.41	0.39	0.47	0.49	0.59	0.64
		1.16	0.35	0.32	0.41	0.38	0.44	0.36	0.31	0.24

Pgm	JRC	*n*	80	115	108	123	47	11	61	68
		1.00	0.88	0.88	0.85	0.86	0.93	0.91	0.83	0.80
		1.06	0.11	0.10	0.11	0.10	0.07	0.09	0.15	0.17
	JRH	*n*	52	52	81	86	49	56	56	35
		1.00	0.79	0.84	0.81	0.76	0.83	0.79	0.78	0.84
		1.06	0.17	0.09	0.14	0.17	0.14	0.20	0.20	0.13
Sod	JRC	*n*	83	118	107	129	51	11	62	69
		1.00	0.93	0.92	0.94	0.93	0.97	0.86	0.89	0.93
	JRH	*n*	52	52	83	86	50	56	60	35
		1.00	0.98	0.94	0.96	0.97	0.97	0.94	0.96	0.96

n = number of individuals sampled.

TABLE 3

For each locus and population the expected and observed covariance matrix

Locus	Population	Covariance matrix		$\phi(x)^a$
		Expected	Observed	
Hbdh	JRC	0.0024, −0.0021 0.0022	0.0057, −0.0050 0.0049	0.44
	JRH	0.0038, −0.0032 0.0035	0.0042, −0.0041 0.0042	0.35
Hk	JRC	0.0028	0.0073	0.96
	JRH	0.0037	0.0024	0.40
Got	JRC	0.00043	0.0018	0.97
	JRH	0.00098	0.00031	0.25
Pgi	JRC	0.0021, −0.0012, −0.00068 0.0034, −0.0017 0.0026	0.013, −0.010, −0.0015 0.010, −0.00062 0.0015	0.96
	JRH	0.0011, −0.00066, −0.00046 0.0041, −0.0028 0.0037	0.00064, −0.0013, 0.00026 0.0061, −0.0038 0.0034	0.37
Pgm	JRC	0.0014, −0.0012 0.0012	0.0015, −0.0011 0.00092	0.16
	JRH	0.0037, −0.0019 0.0021	0.00078, −0.00079 0.0012	0.36
Sod	JRC	0.00087	0.00096	0.76
	JRH	0.00079	0.00018	0.23

[a] Cumulative distribution of the test statistic.

TABLE 4

The probability of detecting selection (power) with type I error set at 5%

σ^2	$\rho\sigma^2$	Average fitness difference	Power		
			Got	*Hk*	*Pgi*
0.36	0.32	0.198	0.40	0.48	
0.18	0.16	0.142	0.089	0.12	
0.036	0.032	0.062	0.054	0.077	
0.0036	0.0032	0.020	0.047	0.072	
0.36	−0.32	0.825	0.75	0.80	
0.18	−0.16	0.577	0.40	0.52	
0.036	−0.032	0.256	0.067	0.16	
0.036	−0.0032	0.081	0.062	0.074	
0.36	0.32	0.169			0.33
0.18	0.12	0.123			0.048
0.036	0.032	0.054			0.034
0.018	0.012	0.066			0.035
0.36	−0.10	0.574			0.482
0.18	−0.05	0.411			0.203
0.036	−0.01	0.182			0.048
0.018	−0.005	0.129			0.046

The parameters σ^2 and $\rho\sigma^2$ from the SAS-CFF model are listed: α was four in each case. The fitness difference is the average difference between the most and least fit genotypes.

TABLE 5

Pollak's estimates of effective population size, N_e, *with symmetric variable selection (SAS-CFF)*

σ^2	Average fitness diffrence	N_e	Var(N_e)
0.36	0.198	−64	4.6×10^6
0.18	0.142	23	3.4×10^7
0.036	0.062	168	1.2×10^7
0.0036	0.020	190	8.6×10^6
No selection	0	206	3.8×10^6

The average fitness difference between the two homozygotes is given along with the variance of the SAS-CFF model. Other model parameters are $\rho = 0.8$, $N_e = 200$, $N_s = 200$, $\alpha = 4$ and $p_0 = 0.5$.

DISCUSSION

We regard the estimates of effective population size in Table 1 as the best estimates. These values incorporate differences in sex ratios every generation, the variance in female reproductive success and differences in population growth. Three "extreme" covariance matrices are found in Table 3—*Got*, *Hk* and *Pgi* in the JRC population. The observed covariance matrices for *Got*, *Hk* and *Pgi* both show more variation than expected from drift and both have $\phi(x) > 0.95$. We would expect about one of the 12 tests at this level of "significance," even with complete neutrality of all loci. The observation of

three tests in this region ($\phi(x) > 0.95$) can be interpreted as an excess and, therefore, an indication of substantial deviations from some underlying assumption. As shown in Table 4, if the assumption of neutrality is removed and selection is allowed to vary over time, large covariance matrices can be detected with some reliability. However, even with variable selection, the power of the tests is low unless the magnitude of selection is quite strong. We should note here that, even if our inferences from these statistical tests are entirely correct, *i.e.*, variable selection has been responsible for the increased variation in allele frequencies, we cannot differentiate between selection on the *Got*, *Hk* and *Pgi* loci and selection on other genes linked to them. Only detailed studies of specific enzyme polymorphisms, such as those of WATT (1977, 1983) on *Pgi* in Colias butterflies, can directly address this latter issue. His demonstration of a "trade off" between the heat stability and kinetic efficiency of different *Pgi* genotypes suggests that variable selection may be acting directly on this polymorphism in natural populations.

It is of some interest to compare these results to those obtained in the first analysis of temporal data by FISHER and FORD (1947). They observed variation in the frequency of the *medionigra* allele over a 6-yr period and obtained estimates of the total number of *Panaxia dominula* moths in a natural population. They did not adjust the effective population size for ecological factors, considered in the present work; rather, they assumed that the population was no larger than their smallest observation of 1000 moths in 1943. Application of their statistical test indicated that allele frequencies showed much more variation over the 6-yr period than would be expected from drift and sampling alone.

Clearly, the results in Table 3 are dependent on accurate estimates of the effective population size. The weakest link in obtaining the final estimate of effective population size was the estimate of variance in female reproductive success. The final estimates of effective population size may be too high if (1) the probability of eggs hatching in a dry environment were considerably higher, (2) there are other significant sources of death that affect whole egg masses, (3) the average female has fewer than four egg masses per season, (4) there is large variance in male reproductive success, or (5) the sex ratio were highly skewed in the years we assumed it was 1:1.

Conditions 1–3 would each result in an increase in Var(F) which, we note from equation 1, would reduce N_e. We have chosen an estimate of δ (0.71) that was fairly high and thus do not consider condition 1 to be a likely source of bias. It certainly would be desirable, of course, to have reliable estimates of δ for each year. Condition 2 must occur to some degree; for instance, predation is probably patchy and affects whole families. The potential magnitude of predation on female variance in reproductive success remains to be studied. Our field observations indicate that condition 3 is not a significant source of bias. Experiments in the laboratory indicate differences in male reproductive success may exist; yet, quantitative values for these differences have not been obtained in either laboratory or natural populations. Since substantial differences in male reproductive success could reduce the estimates of effective population size produced here and methods exist for quantifying these differ-

ences in natural population (COBBS 1977; ØSTERGAARD and CHRISTIANSEN 1981), this remains an important area for future research. To assess the possible magnitude of condition 5 we have recomputed the JRC effective population sizes in Table 1 taking into account the sex ratios for every year (*i.e.*, using method IV). The harmonic mean of these new values is 133. The harmonic mean of the original JRC population sizes is 137. Therefore, condition 5 is unlikely to have had a serious effect on our final test statistics.

Our final estimates of N_e may be too small if (1) the probability of eggs hatching in a dry environment is closer to 0.2 rather than 0.7, (2) females lay five or more egg masses on average, or (3) differential capture rates of males and females lead to an underestimate of c.

We have already commented on conditions 1 and 2. Condition 3 could have substantial effects on N_e. For instance, if the frequency of females in JRC during 1978 were 0.5 instead of 0.34, the harmonic mean of all population sizes in JRC would increase from 137 to 170. We note that, if we have underestimated N_e, then the expected variance in allele frequencies due to drift should be even smaller than the predicted values given in Table 3. Thus, *Got*, *Hk* and *Pgi* would still show more variation than expected. We stress that to obtain direct estimates of effective population size for natural populations sources of mortality that affect whole families should be closely examined.

A qualitative conclusion from this work which is in accord with others (SCHAFFER, YARDLEY and ANDERSON 1977; WATTERSON 1982) is that selection has to be quite strong before patterns of allele frequency variation differ substantially from the neutral expectations. Other aspects of genetic differentiation of *E. editha* have been used to discount genetic drift. EHRLICH and WHITE (1980) have noted that Colorado populations of *E. editha* differ greatly from West Coast populations at the *Pgm* locus but are very similar at seven other loci. In the absence of significant migration they argue the lack of differentiation at most loci except *Pgm* must be due to some form of stabilizing selection or, if the populations only recently split, *Pgm* should not be so drastically different.

NEI and TATENO (1975) have shown that in some circumstances drift alone can produce patterns of differentiation in which identities (I_j) at some loci are high and at others low. Their simulation results, however, show approximately half of the loci in each category after 500 generations, with a trend toward increasing numbers of differentiated loci with generation time. Ehrlich and White's results, however, show a high to low identity ratio of 7:1 (binomial $P = 0.03$ if 4:4 expected). The approximate number of generations since the populations in their study were united is 7000, and the N_e was probably in the 50–500 range. Furthermore, all of the loci studied by Ehrlich and White are polymorphic in the vast majority of *E. editha* populations (MCKECHNIE, EHRLICH and WHITE 1975), and it is reasonable to assume that all eight loci were polymorphic when the populations separated. In contrast, Nei and Tateno started their simulation with about 60% of their loci monomorphic. Thus, one would expect fewer similar loci than found in the Ehrlich-White data, but the opposite is the case.

We have purposely tried to limit the number of assumptions necessary to

derive our final test statistics. The hope is that the final test will be more robust than previously described tests. The "cost" of making fewer assumptions is the need empirically to determine the distribution of test statistics via computer simulations. We point out that this procedure results in major differences in our final test criteria. For instance, the value of λ_i for population JRH and locus *Pgi* was 10.9. If we had assumed that the allele frequency vectors were samples from a trivariate normal distribution, then our test statistic would have a χ^2 distribution with 6 d.f. Consequently, for this value of λ_i, $\phi(x) \cong 0.91$, which indicates a nearly significant result. From our empirical distribution of λ_i we observe that the true value of $\phi(x)$ is 0.35. Note that the computer simulations provide our expected value of the covariance matrix in addition to determining the distribution of λ_i. To derive this quantity analytically would be difficult at best without a number of additional simplifying assumptions.

This work was supported by National Science Foundation grant DEB 82-06961 to P. R. EHRLICH, National Institutes of Health grant GM 28016 to M. W. FELDMAN, National Institutes of Health National Service Award GM 07310 from the National Institute of General Medical Sciences to L. D. MUELLER and a grant from the Koret Foundation of San Francisco. We thank MARCUS W. FELDMAN, JEFF POWELL and two anonymous referees for comments on a draft of this manuscript.

LITERATURE CITED

BEGON, M., C. B. KRIMBAS and M. LOUKAS, 1980 The genetics of *Drosophila subobscura*. XV. Effective size of a natural population estimated by three independent methods. Heredity **45:** 335–350.

BISHOP, Y. M., S. E. FIENBERG and P. W. HOLLAND, 1975 Discrete multivariate analysis: theory and practice. MIT Press, Cambridge, Massachusetts.

BRUSSARD, P., P. R. EHRLICH and M. C. SINGER, 1974 Adult movements and population structure in *Euphydryas editha*. Evolution **28:** 408–415.

BURTON, R. S. and M. W. FELDMAN, 1983 Physiological effects of an allozyme polymorphism: glutamate-pyruvate transaminase and response to hyperosmotic stress in the copepod *Tigriopus californicus*. Biochem. Genet. **21:** 239–251.

CHAKRABORTY, R., P. A. FUERST and M. NEI, 1980 Statistical studies on protein polymorphism in natural populations. III. Distribution of allele frequencies and the number of alleles per locus. Genetics **94:** 1039–1063.

COBBS, G., 1977 Multiple insemination and male sexual selection in natural populations of *Drosophila pseudoobscura*. Am. Nat. **111:** 641–656.

CROW, J. F. and N. E. MORTON, 1955 Measurement of gene frequency drift in small populations. Evolution **9:** 202–214.

DIACONIS, P. and B. EFRON, 1983 Computer-intensive methods in statistics. Sci. Am. **248:** 116–130.

EHRLICH, P. R. and S. E. DAVIDSON, 1960 Techniques for capture-recapture studies of Lepidoptera populations. J. Lepid. Soc. **14:** 227–229.

EHRLICH, P. R., A. E. LAUNER and D. D. MURPHY, 1984 Can sex ratio be defined or determined? The case of a population of checkerspot butterflies. Am. Nat. **124:** 527–539.

EHRLICH, P. R. and D. D. MURPHY, 1981 The population biology of checkerspot butterflies (Euphydryas). Biol. Zentralbl. **100:** 613–629.

EHRLICH, P. R. and R. R. WHITE, 1980 Colorado checkerspot butterflies: isolation, neutrality and the biospecies. Am. Nat. **115:** 328–341.

EHRLICH, P. R., R. R. WHITE, M. C. SINGER, S. W. MCKECHNIE and L. E. GILBERT, 1975 Checkerspot butterflies: a historical perspective. Science **188:** 221–228.

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

EWENS, W. J., 1979 *Mathematical Population Genetics.* Springer-Verlag, Berlin.

EWENS, W. J., 1982 On the concept of effective population size. Theor. Pop. Biol. **21:** 373–378.

FISHER, R. A. and E. B. FORD, 1947 The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula* L. Heredity **1:** 143–174.

FUERST, P. A., R. CHAKRABORTY and M. NEI, 1977 Statistical studies on protein polymorphism in natural populations. I. Distribution of single locus heterozygosity. Genetics **86:** 455–483.

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

GILLESPIE, J. H., 1979 Molecular evolution and polymorphism in a random environment. Genetics **93:** 737–754.

LEWONTIN, R. C. and J. KRAKAUER, 1973 Distribution of gene frequency as a test of the theory of selective neutrality of polymorphisms. Genetics **74:** 175–195.

MCDONALD, J. F. and F. J. AYALA, 1978 Gene regulation in adaptive evolution. Can. J. Genet. Cytol. **20:** 159–175.

MCKECHNIE, S. W., P. R. EHRLICH and R. R. WHITE, 1975 Population genetics of *Euphydryas* butterflies. I. Genetic variation and the neutrality hypothesis. Genetics **81:** 571–594.

MORRISON, D. F., 1976 *Multivariate Statistical Methods,* Ed. 2. McGraw Hill, New York.

MURPHY, D. D., A. E. LAUNER and P. R. EHRLICH, 1983 The role of adult feeding in egg production and population dynamics of the checkerspot butterfly *Euphydryas editha.* Oecologia **56:** 257–263.

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

NEI, M. and Y. TATENO, 1975 Interlocus variation of genetic distance and the neutral mutation theory. Proc. Natl. Acad. Sci. USA **72:** 2758–2760.

ØSTERGAARD, H. and F. B. CHRISTIANSEN, 1981 Selection component analysis of natural polymorphisms using population samples including mother-offspring combinations, II. Theor. Pop. Biol. **19:** 378–419.

PAMILO, P. and S. L. VARVIO-AHO, 1980 Letter to the editor on the estimation of population size from allele frequency changes. Genetics **95:** 1055–1057.

PERLOW, J., 1979 The transition density for multiple neutral alleles. Theor. Pop. Biol. **16:** 223–232.

PLACE, A. R. and D. A. POWERS, 1979 Genetic variation and relative catalytic efficiencies: lactate dehydrogenase B allozyme of *Fundulus heteroclitus.* Proc. Natl. Acad. Sci. USA **76:** 2354–2358.

POLLAK, E., 1983 A new method for estimating the effective population size from allele frequency changes. Genetics **104:** 531–548.

ROTHMAN, E. D. and A. R. TEMPLETON, 1980 A class of models of selectively neutral alleles. Theor. Pop. Biol. **18:** 135–150.

SCHAFFER, H. E., D. YARDLEY and W. W. ANDERSON, 1977 Drift or selection: a statistical test of gene frequency variation over generations. Genetics **87:** 371–379.

SEBER, G. A. F., 1973 *The Estimation of Animal Abundance and Related Parameters.* Griffin, London.

SINGER, M. C. and P. R. EHRLICH, 1979 Population dynamics of the checkerspot butterfly *Euphydryas editha*. Fortschr. Zool. **25:** 53–60.

WATT, W. B., 1977 Adaptation at specific loci. I. Natural selection on phosphoglucose isomerase of *Colias* butterflies: biochemical and population aspects. Genetics **87:** 177–194.

WATT, W. B. 1983 Adaptation at specific loci. II. Demographic and biochemical elements in the maintenance of the colias PGI polymorphism. Genetics **103:** 691–724.

WATT, W. B., R. C. CASSIN and M. S. SWAN, 1983 Adaptation at specific loci. III. Field behavior and survivorship differences among Colias PGI genotypes are predictable from in vitro biochemistry. Genetics **103:** 725–739.

WATTERSON, G. A., 1977 Heterosis or neutrality? Genetics **85:** 789–814.

WATTERSON, G. A., 1982 Testing selection at a single locus. Biometrics **38:** 323–331.

Communicating editor: J. R. POWELL

APPENDIX 1

In this appendix we derive an inbreeding effective population size, N_e, which takes into account the relevant ecological details of Euphydryas. Our derivation closely follows the methods outlined by EWENS (1979). The inbreeding effective population size is defined as the reciprocal of the probability (π) that two different alleles chosen at random in generation $t + 1$ are descended from the same individual in generation t. Let N_1 be the number of males and N_2 the number of females in generation t. The total population size in generation t, N_t, is clearly the sum of N_1 and N_2. For female i we denote the number of descendent alleles produced from each of i's two copies by m_i and m_{N_2+i}. For male i we denote similar quantities as p_i and p_{N_1+i}. Since each individual at time $t + 1$ has one allele from a female and one from a male the following relationships hold,

$$N_1\bar{p} = \sum_{i=1}^{N_1} (p_i + p_{N_1+i}) = N_{t+1},$$

$$N_2\bar{m} = \sum_{i=1}^{N_2} (m_i + m_{N_2+i}) = N_{t+1},$$

where N_{t+1} is the total number of individuals in generation $t + 1$. Following arguments similar to EWENS' (1979, p. 106) we see that

$$\pi = (N_{t+1} - 1)[2(2N_{t+1} - 1)]^{-1} \left\{ \sum_{i=1}^{N_1} (p_i + p_{N_1+i})(p_i + p_{N_1+i} - 1) \right.$$

$$\cdot[\Sigma(p_i + p_{N_1+i})]^{-1}[\Sigma(p_i + p_{N_1+i}) - 1]^{-1} +$$

$$\left. \sum_{i=1}^{N_2} (m_i + m_{N_2+i})(m_i + m_{N_2+i} - 1)[\Sigma(m_i + m_{N_2+i})]^{-1}[\Sigma(m_i + m_{N_2+i}) - 1]^{-1} \right\}.$$

If we let V_m be the variance of the m_i's and V_p the variance of the p_i's then,

$$\pi = [2(2\bar{m}N_2 - 1)]^{-1}[V_p/\bar{p} + V_m/\bar{m} + \bar{p} + \bar{m} - 2].$$

Thus,

$$N_e = 2(2\bar{m}N_2 - 1)(V_p/\bar{p} + V_m/\bar{m} + \bar{p} + \bar{m} - 2)^{-1}.$$

Since we have no information about the variance in male reproductive success, we assume $V_p/\bar{p} = 1$, yielding

$$N_e = 2(2\bar{m}N_2 - 1)[V_m/\bar{m} + \bar{m}(N_2/N_1 + 1) - 1]^{-1}. \quad (1A)$$

We consider four variations of (1A) that make various assumptions about the sex ratio and population growth.

Model I: Equal sex ratio, no population growth. These assumptions imply that $N_1 = N_2 = N_t/2$ and $\bar{m} = 2$; thus,

$$N_{el} = 2(2N_t - 1)(V_m/2 + 3)^{-1}. \tag{2A}$$

Model II: Equal sex ratio, population growth. With these assumptions we have $N_1 = N_2 = N_t/2$ and $\bar{m} = 2N_{t+1}/N_t$; thus,

$$N_{elI} = 4N_tN_{t+1}(2N_{t+1} - 1)(V_mN_t^2 + 8N_{t+1}^2 - 2N_tN_{t+1})^{-1}. \tag{3A}$$

Model III: Unequal sex ratio, no population growth. These assumptions imply $\bar{m} = N_t/N_2$; thus,

$$N_{elII} = 2N_tN_1N_2(2N_t - 1)[N_1N_2(N_2V_m - N_t) + N_t^3]^{-1}. \tag{4A}$$

Model IV: Unequal sex ratio, population growth. With this model $\bar{m} = N_{t+1}/N_2$; thus,

$$N_{elV} = 2N_1N_2N_{t+1}(2N_{t+1} - 1)[N_1N_2(N_2V_m - N_{t+1}) + N_{t+1}^2(N_1 + N_2)]^{-1}. \tag{5A}$$

APPENDIX 2

Let us define some terms. Let X_i be a random variable that represents the number of eggs laid by a female in her ith egg mass. A particular realization of this random variable will be denoted x_i. We also have $E(X_i) = u_{x_i}$ and $\mathrm{Var}(X_i) = \sigma_{x_i}^2$. If δ is the probability of an egg mass hatching into a dry environment, then let Y be an indicator random variable with

$$Y = \begin{cases} 0 \text{ with probability } \delta \\ 1 \text{ with probability } 1 - \delta \end{cases}$$

Z_j will be a Bernoulli random variable and, as before, z_j will be a realization of Z_j. If the larvae survive dessication, we assume they survive to become adults with probability λ. Thus, in a mass of x_i eggs that have survived dessication, the number that survive to adulthood, W_i, will be $\sum_{j=1}^{x_i} z_j$. The fate of egg j is survival to adulthood (if $z_j = 1$) or death (if $z_j = 0$). Finally, we let the total number of larvae from egg mass i that survive to adulthood to be W_i. Clearly, we now seek an expression for $E(W_i)$ and $\mathrm{Var}(W_i)$.

In both cases we find the expected values by conditioning successively on the values of X_i, Y and $\sum_{j=1}^{X_iY} Z_j$. Thus,

$$E(W_i) = \mathrm{E}\{E\{E\{E[W_i \mid x_i, y, \Sigma z_j]\}\}\}.$$

Solving this we get

$$E[W_i \mid x_i, y, \Sigma z_j] = \sum_{j=1}^{x_iy} z_j,$$

$$E\left[\sum_{j=1}^{x_iy} Z_j \mid x_i, y\right] = x_iy\lambda \qquad \text{since} \quad \Sigma Z_j \sim B(x_iy\lambda,\ \lambda(1 - \lambda)x_iy),$$

$$E[x_iY\lambda \mid x_i] = \lambda x_i(1 - \delta),$$

$E[\lambda(1 - \delta)X_i] = E(W_i) = \lambda(1 - \delta)u_{x_i} = \hat{W}_i$. We do the same sort of conditioning to get $\mathrm{Var}(W_i)$. After some algebra we see that

$$\mathrm{Var}(W_i) = \hat{W}_i[1 - \lambda - \hat{W}_i + \lambda\sigma_{x_i}^2/u_{x_i} + \lambda u_{x_i}].$$

Since the total fecundity, F, is equal to $\sum_{i=1}^{m} W_i$, where m is the last clutch, we note that $\mathrm{Var}(F) = \sum_{i=1}^{m} \mathrm{Var}(W_i)$. This derivation includes as special cases the "random survival" and "family survival" models considered by CROW and MORTON (1955).