Aquaculture, 108 (1992) 215–232
Elsevier Science Publishers B.V., Amsterdam

AQUA 30024

Effective population numbers of shellfish broodstocks estimated from temporal variance in allelic frequencies

Dennis Hedgecock[a], Victor Chow[a] and Robin S. Waples[b]
[a]*University of California, Davis, Bodega Marine Laboratory, Bodega Bay, CA, USA*
[b]*Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA, USA*

(Accepted 30 May 1992)

ABSTRACT

Hedgecock, D., Chow, V. and Waples, R.S., 1992. Effective population numbers of shellfish broodstocks estimated from temporal variance in allelic frequencies. *Aquaculture*, 108: 215–232.

Effective population number (N_e) can be estimated from temporal changes in the frequencies of selectively neutral alleles in isolated populations. The utility and limitations of this temporal method for aquaculture broodstocks are examined using published allozyme data for 16 shellfish stocks and their wild progenitors. Estimates of N_e for these broodstocks are all less than 100, and 13 are less than 50. For eight stocks, estimated N_e agrees with records of breeding numbers (N_b), but in the remaining eight cases, N_b lies outside of the 95% confidence interval for estimated N_e.

Assumptions of the temporal method are evaluated. First, two independent tests, one based on the distribution of temporal variances at individual loci, the other on the expected loss of alleles from finite populations, suggest that allozymes behave as selectively neutral genetic markers suitable for estimating genetic drift and N_e. Second, migration into a stock can have diverse effects on estimates of N_e, depending on the genetic similarity of immigrants and the captive broodstock population. Exchange among diverged hatchery stocks of Kuruma shrimp appears to have inflated temporal change, reducing estimated N_e; on the other hand, additions of wild hard clams to a hatchery stock appear to have retarded temporal change, increasing estimated N_e. Four cases in which N_b is smaller than the lower confidence limit of estimated N_e are explained by contamination among stocks propagated simultaneously in the same hatchery. Finally, comparison of captive broodstocks with contemporary samples from natural sources confounds genetic drift that has occurred in both lineages; genetic drift is not negligible in some natural populations, leading to uncertainty in some N_e estimates.

For four cases in which N_b is much larger than estimated N_e, variance in reproductive success is proposed as the most likely explanation of loss of genetic diversity over time. The extremely high fecundities and variable fertilities of aquatic organisms must be taken into account in broodstock management, particularly if hatchery products are used to enhance natural populations. The temporal method, when correctly applied and interpreted, provides new insight into the genetics of aquaculture broodstock populations.

Correspondence to: Dr. D. Hedgecock, University of California, Davis, Bodega Marine Laboratory, P.O. Box 247, Bodega Bay, CA 94923, USA.

INTRODUCTION

The effective population number (N_e) holds a prominant place in population biology, allowing calculation of the expected rate of decay of genetic variability, the expected increase in inbreeding, and the strength of selection required to overcome random genetic drift in a finite population (Crow and Denniston, 1988). Yet, estimating N_e in real populations has proved to be a daunting problem. Instead of attempting to estimate rates of genetic change from independent knowledge of N_e, however, several authors have focused on estimating N_e indirectly from temporal genetic changes in finite populations (Krimbas and Tsakas, 1971; Nei and Tajima, 1981; Pollak, 1983). Recent simulation studies have shown this temporal method to be robust and particularly well suited to the study of small captive populations (Waples, 1989; Waples and Teel, 1990). Here we apply the temporal method to allozyme data from published studies that compared captive aquaculture broodstocks to the natural populations from which they were derived.

The effective population number of an aquaculture broodstock may be a critical parameter to measure. Refinements of aquatic hatchery technology and broodstock husbandry and management have permitted the establishment of *closed* aquaculture broodstocks, i.e. populations propagated solely by the cycling of mature animals from growout facilities back to the hatchery as broodstock, without input of individuals from wild populations or other captive broodstocks. Of particular concern is the extent to which random genetic drift and inbreeding occur in these closed aquaculture broodstocks. Because of the extremely high fecundities but variable fertilities of many cultured aquatic animals, hatchery seed may often come from just a few successfully spawning individuals (e.g. Lannan, 1980; Gharrett and Shirley, 1985; Withler, 1988; Gile and Ferguson, 1990; Gaffney et al., 1992). Under these conditions, effective population number may be only a fraction of the number of broodstock. Estimates of N_e that have thus far been made for aquaculture broodstocks (Taniguchi et al., 1983; Sbordoni et al., 1986, 1987; Hedgecock and Sly, 1990; Waples and Teel, 1990) indeed suggest that closed aquatic broodstocks are small populations in which random drift occurs at appreciable rates. Improvement of aquaculture broodstocks through domestication or artificial selection might be stymied by random drift in such effectively small populations.

Estimating N_e from published data on temporal genetic change in captive aquatic populations allows us to examine the assumptions and limitations of the method. One key assumption of the temporal method is that allozymes are selectively neutral; indeed, much of the past interest in studying temporal genetic changes has been directed at testing this assumption (Lewontin and Krakauer, 1973, 1975; Nei and Tajima, 1981; Mueller et al., 1985). We use two independent methods to evaluate the assumption of selective neutrality.

A second assumption is that the same closed population has been randomly sampled at two different points in time. In most cases, however, a broodstock that has been hatchery-propagated for a number of generations, t, is compared to a sample from the wild progenitor population, taken not at the founding of the broodstock but contemporaneously with the sampling of the t^{th} hatchery generation. Such a comparison confounds random genetic drift that has accumulated along the two independent lines of descent in the hatchery and natural populations. We use the temporal method to investigate the extent of genetic drift in several natural populations from which progenitors of hatchery stocks were selected. A third important assumption is that the broodstock population is indeed closed. The potentially diverse effects of immigration on the estimation of N_e are illustrated by two cases in which immigration is known to have occurred.

MATERIALS AND METHODS

Estimation of effective population number

The rationale for estimating N_e from its inverse relationship to temporal changes in allelic frequencies is reviewed by Waples (1989). In practice, allelic frequencies and an estimate of temporal variance in allelic frequency, F, which is standardized to compensate for differences among loci in initial allelic frequencies, are estimated from population samples. Several estimates of F have been proposed (Krimbas and Tsakas, 1971; Nei and Tajima, 1981; Pollak, 1983) and evaluated in simulations (Waples, 1989). Following Pollak's (1983, p. 543) suggestion, we use Nei and Tajima's (1981) estimator ($\hat{F}_c$) for loci with two alleles but Pollak's estimator ($\hat{F}_k$) for loci with three or more alleles. A mean of $\hat{F}$ from several loci, weighted by the number of alleles is taken as an estimate of the population mean $\bar{F}$.

For aquatic species, sampling to determine allelic frequencies is generally done before reproduction, without replacement, and independently of the subsequent selection of broodstock individuals (sampling plan II of Waples, 1989). Thus, the expected value of the drift variance, $E(F)$, is approximately equal to $t/(2N_e)+1/(2S_o)+1/(2S_t)$, where t is the number of generations in the time interval between the initial sample of S_o individuals and the second sample of S_t individuals. Rearrangement of this relationship yields an estimator of the harmonic mean effective population number over the time interval, $\hat{N}_k = t/(2[\bar{F} - 1/(2S_o) - 1/(2S_t)])$. If sample sizes differ among loci, the S-terms are replaced by harmonic means weighted by the number of independent alleles per locus. Ninety-five percent confidence limits (C.L.) on the estimated $\hat{N}_k$ are calculated according to Waples (1989).

This method of estimating effective population number assumes that selection is not operating to change allelic frequencies. If random genetic drift is the only cause of temporal change in allelic frequencies, then variance of $\hat{F}$

among loci should be stochastic; $n\hat{F}/E(\hat{F})$, where n is the total number of loci, should therefore be distributed approximately as chi-square with n degrees of freedom (Lewontin and Krakauer, 1973). Thus, one way to examine the assumption that allozymes are selectively neutral genetic markers is to compare the distribution of $n\hat{F}/\bar{F}$ with the χ^2, n d.f., either by inspection of a probability plot of observed vs. expected values (Gnanadesikan, 1977; Wilkinson, 1987) or by the Kolmogorov-Smirnov (KS) test. This general approach has been criticized by Mueller et al. (1985) because the sampling distribution of $\hat{F}$ is unknown and $E(\hat{F})$ must be estimated by the mean of $\hat{F}$ values in small samples. Simulations have nevertheless shown that the χ^2 is a good approximation unless t is large (Waples, 1989).

An independent, a posteriori test of whether allozyme markers behave as neutral alleles is to compare the numbers of alleles remaining in the population with the numbers expected to remain if $\hat{N}_k$ were the true N_e. These expectations are calculated as $n_t = \text{sum}\,[1 - P_o(p;t)]$, where the summation is over all alleles present in the founder of the broodstock population and $P_o(p;t)$ is the probability that a selectively neutral allele, initially at frequency p in the founder population at generation 0, is lost by generation t (Ewens, 1979; see Hedgecock and Sly, 1990). A χ^2 goodness-of-fit test with 1 d.f. is used to test whether the observed numbers of alleles lost and remaining differ from their expectations. [Alleles found in the sample at generation t but not in the sample representing generation 0 are assumed to have been missed by sampling error and are excluded from the calculation.]

Estimating effective population numbers of aquaculture stocks

We estimate F and N_e for 13 shellfish stocks for which allozyme data have been reported for both a closed broodstock population and a natural population representing that from which the cultured stock was originally derived: Kuruma shrimp (*Penaeus japonicus*) propagated for seven generations in Lesina Lagoon, Italy (Sbordoni et al., 1986, 1987); two stocks of hard clam (*Mercenaria mercenaria*) selected for large body size over four generations (Dillon and Manzi, 1987); four lines of pearl oysters (*Pinctada martensii*) selected for shell traits over six generations (Wada, 1986); one line of American oyster *Crassostrea virginica* selected for increased growth rate for eight generations (Paynter and DiMichele, 1990) and five lines of this same species selected for increased resistance to MSX disease for four–six generations (Vrijenhoek et al., 1990). For comparison, we tabulate estimates of F and N_e for three different stocks of the Pacific oyster *Crassostrea gigas* given by Hedgecock and Sly (1990). Additional data for the native progenitors of some oyster lines are from Buroker (1975, 1983, 1984).

The number of breeders used to propagate each line, N_b, was taken directly from each source reference or estimated from information provided. Sbordoni et al. (1986) gives breeding population size as number of mating pairs

per generation; harmonic means were used to estimate the per generation breeding number over various time intervals. Vrijenhoek et al. (1990) gives N_b's adjusted for differences in the numbers of males and females used to propagate the MSX-resistant oyster lines. Sex ratios were assumed to be equal in the remaining cases.

RESULTS

Penaeus japonicus

Harmonic mean effective population number for the Kuruma shrimp broodstock, over an interval of seven generations of hatchery propagation, is estimated to be only about 10 (Table 1). Breeding numbers, N_b, are within the 95% confidence limits (C.L.) of $\hat{N}_k$ for all time intervals beginning with the first generation, which comprised just two breeding pairs. The number of spawning pairs increased to 300 in the third generation (Sbordoni et al., 1986), so that, for example, N_b of 600 for the F5–F6 interval was well above the upper C.L. for N_k. A transient increase in $\bar{F}$ and a concomitant lowering of $\hat{N}_k$ to about 4 or 5 appears for time intervals spanning the fifth generation (e.g., the F5–F6 interval shown Table 1). N_e may have been higher in the F4–F5 and F6–F7 intervals ($\hat{N}_k$ nearly 50; estimates not shown), but these single-generation point estimates are imprecise and have infinite upper confidence limits. Distributions of $n\hat{F}/\bar{F}$ for comparisons of the first to the fourth, fifth,

TABLE 1

Estimates of temporal variance in allelic frequencies ($\bar{F}$) and variance effective population number ($\hat{N}_k$) in an Italian hatchery stock of *Penaeus japonicus*[a]. S_o and S_t are sample sizes for the progenitor and derived populations (mean over loci, weighted by the numbers of independent alleles); N_b is an estimate of the breeding number; n_p is the number of alleles in S_o, n_a the number in S_t, n_t the number expected to remain in a population of size $\hat{N}_k$ after t generations of random drift; goodness-of-fit of n_a vs. n_t is tested by chi-square with one degree of freedom

Interval	S_o, S_t	No. of loci	No. of gen.	$\bar{F}$	$\hat{N}_k$ (95% C.L.s)	N_b	n_p, n_a (n_t)	χ^2, 1 d.f. n_a vs. n_t
F1–F4	13.6, 21.7	8	3	0.220	9.8 (2.5–36.8)	11.3	17, 12 (14.6)	2.07
F1–F5	13.6, 25.2	7	4	0.232	11.7 (2.8–41.2)	15.0	16, 10 (13.4)	3.79
F1–F6	13.6, 19.2	8	5	0.296	10.9 (3.2–30.8)	18.6	17, 11 (13.8)	2.13
F5–F6	25.8, 19.2	4	1	0.169	4.1 (0.5–26.8)	600[b]	7, 7 (6.8)	—
F1–F7	13.6, 19.4	7	6	0.366	10.0 (2.8–26.3)	22.2	16, 9 (12.3)	2.72

[a]Data for the Lesina Lagoon population (Sbordoni et al., 1986, 1987).
[b]Includes an unknown number of broodstock obtained as postlarvae from a Japanese hatchery.

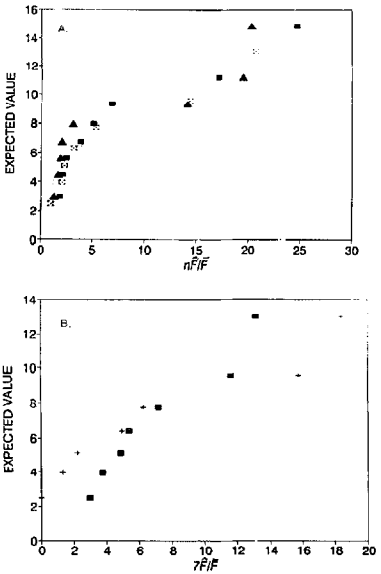

Fig. 1. Chi-square probability plots of $n\hat{F}/\bar{F}$ for n degrees of freedom. A. Temporal variances, F, calculated from allelic frequency data for the Lesina Lagoon population of Kuruma shrimp *Penaeus japonicus* sampled over four time intervals (from Sbordoni et al., 1986, 1987): filled squares=F1–F4, 8 d.f.; crossed squares=F1–F5, 7 d.f.; filled triangles=F1–F6, 8 d.f.; empty squares=F1–F7, 7 d.f. The distribution of $8\hat{F}/\bar{F}$ for the F1–F6 interval differs significantly from a chi-square distribution by the Kolmogorov-Smirnov test ($P=0.008$). B. Distributions of $7\hat{F}/\bar{F}$ for comparisons of the ARC (filled squares) and VIMS (crossed lines) hatchery stocks of hard clams *Mercenaria mercenaria* to their respective native progenitor populations (allelic frequency data from Dillon and Manzi, 1987).

sixth and seventh generations, respectively, are curvilinear, particularly for the F1–F6 and F1–F7 intervals. $\hat{F}$s below the mean are smaller than expected and those above the mean are larger than expected; yet only the F1–F6 comparison differs significantly from χ^2, 8 d.f. (Fig. 1A; $P=0.008$ for the KS test corresponding to this interval). Numbers of alleles remaining in this shrimp population over time are somewhat less than expected under random genetic drift, but all tests yield non-significant χ^2 values.

Mercenaria mercenaria

Estimates of N_e over the four generations separating the two stocks of hard clam (ARC, VIMS) from native progenitors are 81.5 and 54.7, respectively

TABLE 2

Estimates of temporal variance in allelic frequencies and variance effective population numbe, in two hatchery stocks of hard clam *Mercenaria mercenaria*[a]. Column headings are defined in Table 1

Stock[b]	S_o, S_t	No. of loci	No. of gen.	$\bar{F}$	$\hat{N}_k$ (95% C.L.s)	N_b	n_p, n_a (n_t)	χ^2, 1 d.f. n_a vs. n_t
ARC	76.2, 100	7	4	0.036	81.5 (32.1–205)	30–60	28, 26 (25.1)	0.064
VIMS	84.7, 74.1	7	4	0.050	54.7 (21.5–128)	30–60	26, 25 (22.7)	1.107

[a]Data from Dillon and Manzi (1987).
[b]Samples of native clams from Martha's Vineyard, MA, and Hog Island, VA, represent progenitors of the ARC and VIMS stocks, respectively.

(Table 2). N_b for both stocks reportedly ranged from 30 to 60, well within the 95% C.L.s for both $\hat{N}_k$ estimates; the $\hat{N}_k$ of 81.5 for the ARC stock, however, is substantially larger than maximum N_b. The two distributions of $7\hat{F}/\bar{F}$ are not significantly different from χ^2, 7 d.f. (Fig. 1B; $P=0.94$ and 0.22 for the KS tests), and the numbers of alleles remaining are close to those expected under the assumption of genetic drift in broodstocks having the estimated effective numbers (26 vs. 25.1 and 25 vs. 22.7, $\chi^2=0.064$ and 1.107, respectively).

Pinctada martensii

Estimates of N_e over six generations in four selected lines of pearl oyster are 30 to 40 in three of the lines but 84.3 in the SWSL-II line (Table 3). $\hat{N}_k$ is larger than N_b in all lines, but the breeding number lies within the 95% C.L. for all but the SWSL-II line, in which case N_b is below the lower limit. The four distributions of $4\hat{F}/\bar{F}$ are not significantly different from χ^2, 4 d.f. (smallest $P=0.561$ in the KS tests). The numbers of alleles remaining and lost are in good agreement with drift expectations for the three populations with $30 < \hat{N}_k < 40$ ($\chi^2=0.183, 3.178, 0.950$), but significantly fewer alleles than expected remain in the SWSL-II line ($\chi^2=5.643, P < 0.05$).

Crassostrea virginica

$\hat{N}_k$ for the American oyster line selected for fast growth over eight generations is 9.8 using Paynter and DiMichele's (1990) data for three loci in the Tolley Point, Chesapeake Bay, native population; using Buroker's (1984) Herring Bay sample to represent the progenitor, $\hat{N}_k$ of the Wilde line is estimated to be 26.2, based on four loci (Table 4). N_b lies within the 95% C.L. for both estimates of $\hat{N}_k$. In neither case are the numbers of alleles remaining and lost significantly different from those expected under random genetic drift in populations of the estimated sizes ($\chi^2=0.009$ and 2.385, respectively). A

TABLE 3

Estimates of temporal variance in allelic frequencies and variance effective population number in four selected lines of pearl oyster *Pinctada martensii*[a]. Column headings are defined in Table 1

Selection line[b]	S_o, S_t	No. of loci	No. of gen.	$\bar{F}$	$\hat{N}_k$ (95% C.L.s)	N_b	n_p, n_a (n_t)	χ^2, 1 d.f. n_a vs. n_t
SWSL-II	280, 25.2	4	6	0.057	84.3 (22.3–388)	15.2	16, 10 (13.8)	5.643*
SWSV-II	280, 63.8	4	6	0.086	39.3 (13.3–86.7)	19.7	16, 11 (12.2)	0.183
SWTL-III	280, 103	4	6	0.104	30.9 (10.9–64.3)	19.3	16, 8 (11.7)	3.178
SWTLS-II	280, 45.5	4	6	0.091	38.2 (12.7–87.9)	20.7	16, 10 (12.2)	0.950

[a]Data from Wada (1986).

[b]The progenitor population for all lines was represented by a pooled sample of lots of cultured pearl oysters from the Ehime Prefecture (see Wada, 1986).

comparison of the two samples of the native population sampled about two generations apart yields a surprisingly small $\hat{N}_k = 14.9$ (Table 5); significantly fewer alleles remain in the Tolley Point sample than expected (8 vs. 12.354, $\chi^2 = 4.399$, $P < 0.05$). Nevertheless, the distribution of $4\hat{F}/\bar{F}$ is not significantly different from χ^2, 4 d.f. ($P = 0.975$ in the KS test).

Estimates of $\hat{N}_k$ based on Vrijenhoek et al.'s (1990) data for five lines of American oyster that were selected over four–six generations for MSX resistance range from 12.5 in the D3-2 line to 35.4 for the NA line (Table 4). In three of four cases, estimates of N_e based on Buroker's (1983) earlier data for native populations are smaller, ranging from 8.3 to 19.0 (Table 4). N_b is less than or equal to the lower 95% C.L. of $\hat{N}_k$ for the DB-1, DB-2 and NA lines. Goodness-of-fit χ^2 values for numbers of alleles remaining and lost in these lines are significant only for DB-2 using Vrijenhoek et al.'s data; in this case 12 alleles remain but 17.5 are expected. Corresponding χ^2 values using Buroker's data are smaller and none is significant.

As for the upper Chesapeake Bay native population, temporal variances in allelic frequencies may be calculated over the two-generation interval separating samples of the Long Island Sound, Delaware Bay, and James River, Virginia, native populations studied by Vrijenhoek et al. (1990) and Buroker (1983). Estimates of N_e for these three localities are infinity, 30.5 and 27.6, respectively (Table 5). Tests for goodness-of-fit between observed and expected numbers of alleles are not significant for the latter two populations; no test is possible for the Long Island Sound population which is not expected to lose any alleles.

TABLE 4

Estimates of temporal variance in allelic frequencies and variance effective population number in hatchery stocks of oysters (*Crassostrea*). Column headings are defined in Table 1

Hatchery stock	S_c, S_t	No. of loci	No. of gen.	$\bar{F}$	$\hat{N}_k$ (95% C.L.s)	N_b	n_p, n_a (n_t)	χ^2, 1 d.f. n_a vs. n_t
C. virginica; the Wilde growth-selected line								
Wilde line[a]	92, 20	4	8	0.180	26.7 (9.2–62.3)	~20	17, 7 (10.6)	2.459
Wilde line[b]	20, 20	3	8	0.458	9.8 (1.1–35.0)	~20	7, 6 (5.6)	0.009
C. virginica; MSX-resistant lines								
DB-1[c]	104, 98.4	6	6	0.174	18.4 (8.9–32.2)	8.7	27, 16 (17.3)	0.275
DB-1[d]	79.4, 98.6	6	6	0.106	31.8 (15.0–57.8)	8.7	27, 16 (20.6)	3.476
DB-2[c]	104, 95.8	6	5	0.310	8.4 (4.1–14.3)	4.1	27, 12 (15.3)	1.156
DB-2[d]	79.4, 96	6	5	0.212	12.5 (6.1–21.8)	4.1	27, 12 (17.5)	4.005*
NA[d]	72.2, 48.6	6	5	0.088	5.4 (15.1–73.2)	6.5	26, 18 (21.3)	2.026
LI[c]	93.8, 95.7	6	6	0.166	19.3 (9.1–34.2)	16.2	26, 18 (19.0)	0.042
LI[d]	46.6, 95.8	6	6	0.186	17.6 (7.7–33.2)	16.2	23, 18 (18.1)	0.040
VA[c]	88, 48	6	4	0.190	11.5 (5.2–21.0)	9.6	26, 18 (17.5)	0.000
VA[d]	118, 47.9	6	4	0.113	20.4 (9.5–38.0)	9.6	28, 18 (19.9)	0.313
C. gigas; hatchery stocks[e]								
Humboldt	57.1, 56.6	14	3	0.201	8.2 (4.8–12.8)	>100	49, 30 (28.7)	0.048
Willapa	57.0, 57.8	14	3	0.055	39.6 (20.0–77.6)	>100	49, 41 (40.6)	0.001
Seasalter	80.9, 118	10	~3.5	0.091	22.2 (11.5–38.7)	75	38, 25 (29.6)	2.594

[a]Progenitor represented by Buroker's (1984) Herring Bay sample.
[b]Progenitor represented by Paynter and DiMichele's (1990))Tolley Pt. sample.
[c]Progenitors represented by native samples studied by Buroker (1983).
[d]Progenitors represented by native samples studied by Vrijenhoek et al. (1990).
[e]Reported by Hedgecock and Sly (1990).

Crassostrea gigas

Estimates of effective population numbers for three hatchery stocks of the Pacific oyster were given by Hedgecock and Sly (1990); these are presented for comparison in Table 4, together with the 95% C.L.s on $\hat{N}_k$, which were not given previously. N_b is much greater than the upper 95% C.L. of $\hat{N}_k$ for all

TABLE 5

Estimates of temporal variance in allelic frequencies and variance effective population number in natural populations of oysters (*Crassostrea*). Column headings are defined in Table 1

Species, locality	S_o, S_t	No. of loci	No. of gen.	$\bar{F}$	$\hat{N}_k$ (95% C.L.s)	n_p, n_a (n_t)	χ^2, 1 d.f. n_a vs. n_t
C. virginica							
Long island Sound[a]	93.9, 46.4	6	2	0.016	∞ (62.4–∞)	26, 22 —[d]	—
Delaware Bay, lower[a]	100, 80.0	6	2	0.042	33.8 (13.9–79.4)	27, 26 (23.8)	1.035
Chesapeake Bay, upper[b]	92.0, 20.0	4	2	0.097	14.9 (4.5–48.2)	17, 8 (12.4)	4.642*
James River[a]	88.8, 118	6	2	0.043	30.0 (13.5–60.8)	26, 24 (23.6)	0.009
C. gigas; Dabob Bay, WA[c]							
1971–85	129, 58.7	5	7	0.021	502 (93.0–∞)	14, 14 (13.9)	—
1972–85	80.5, 58.7	5	6	0.024	337 (68.0–∞)	15, 15 (14.8)	—
1971–72	111, 95.4	6	1	0.023	41.7 (13.4–219)	25, 25 (24.6)	—

[a]Comparable samples from Buroker (1983) and Vrijenhoek et al. (1990).
[b]Comparable samples from Buroker (1984) and Paynter and DiMichele (1990).
[c]Comparable samples from Buroker (1975) and Hedgecock and Sly (1990).
[d]n_t is zero when $\hat{N}_k$ is infinite; four alleles are lost, n_a=22.

three stocks. Data for oyster samples taken from Dabob Bay, Puget Sound, Washington, in 1971 and 1972 (Buroker, 1975) and in 1985 (Hedgecock and Sly, 1990) allow estimates of N_e for this natural population. $\hat{N}_k$ is 42 over the 1971–72 interval and several hundred over the two longer time intervals, although upper C.L.s are infinity in these last two cases.

DISCUSSION

Despite the widespread availability of electrophoretic techniques for obtaining genetic information from hatchery-propagated aquaculture populations, few estimates of temporal variance in allelic frequencies or effective population number have been made. Electrophoretic data have commonly been used only to estimate average heterozygosity, even though this static measure of genetic variability is relatively insensitive to allele frequency shifts in small populations (Allendorf, 1986; Fuerst and Maruyama, 1986; Hedgecock and Sly, 1990). We believe that methods for estimating temporal variances in allelic frequencies and effective population numbers provide impor-

tant new information concerning genetic change in closed broodstock populations.

Application of the temporal method to published data on 16 shellfish stocks reveal that recorded numbers of breeding individuals, N_b, lie within the 95% confidence limits of the corresponding estimates, $\hat{N}_k$, of effective population number in eight cases. In the remaining eight cases, N_b lies outside of the 95% C.L. of $\hat{N}_k$; in four cases, N_b is slightly less than the lower 95% C.L. of $\hat{N}_k$, and in another four cases, N_b is much greater than the upper 95% C.L. of $\hat{N}_k$. We evaluate these discrepancies between N_b and $\hat{N}_k$ in light of the assumptions of the temporal method.

The major assumptions in calculating temporal variance and estimating N_e are: (1) mutation is unimportant, (2) alleles are selectively neutral and not in gametic phase disequilibrium, (3) samples for genetic analysis are randomly drawn, and (4) there is no migration from other populations (Waples, 1989). Over the short time intervals involved in the derivation of hatchery-propagated broodstock from extant native populations, mutation cannot play an appreciable role in changing allelic frequencies. The remaining three assumptions, however, are potentially violated in our study and must be evaluated more carefully, although in the absence of single-genotype data we cannot assess gametic phase disequilibrium.

We evaluate the assumption that allozymes are selectively neutral genetic markers in two ways. First, we compare the distributions of $n\hat{F}/\bar{F}$ with χ^2 distributions having degrees of freedom, n, equal to the number of loci. Plots of $n\hat{F}/\bar{F}$ against expected χ^2 values corresponding to the cumulative probabilities of the ranked $n\hat{F}/\bar{F}$ data (Wilkinson, 1987) are close to linear, as expected under the null hypothesis, although in some cases values below the mean tend to be smaller than expected while those above the mean are larger than expected (Fig. 1). Only one of 25 Kolmogorov-Smirnov tests for goodness-of-fit of $n\hat{F}/\bar{F}$ to the appropriate χ^2 distribution is significant. Second, we have calculated the number of alleles that should have remained (and been lost) in each broodstock, assuming that random genetic drift were the only process removing alleles during descent from a progenitor over t generations with an average N_e equal to $\hat{N}_k$. This test does not presently account for error in detecting alleles in finite samples of sizes S_o and S_t. Nevertheless, of 24 possible tests, only two were significant at the $P=0.05$ level. One of these significant departures was non-significant when a different sample from the native population was used to represent the progenitor in the comparison. Thus, by both tests, allozymes appear to behave as selectively neutral genetic markers in these populations, most of which were under rigorous artificial selection.

The method of estimating temporal variance and effective population number also depends on random sampling of gametes and unbiased estimation of allelic frequencies in a closed population at different points in time.

We must assume that data from published sources come from truly random, temporal samples of the identified populations; violations of this assumption, for example local sampling of native populations that would miss variation actually present in the founders of a stock (Nei and Tajima, 1981), could easily account for some of the discrepancies between N_b and $\hat{N}_k$. We do note one important respect in which sampling in most of these studies deviates from the sampling theory of the temporal method. Although sampling is without replacement and before reproduction, conforming to sampling plan II for the temporal method (Waples, 1989), initial samples of the natural progenitor populations, S_o, are typically taken not at generation 0 but at generation t, contemporaneously with samples from the current generation of hatchery stocks. (The study of Sbordoni et al. (1986, 1987) is exceptional in this respect.) Extending the temporal method to cover this sampling scheme reveals the importance of tracking genetic change in native populations from which aquaculture stocks are derived.

Variance of allelic frequencies between contemporaneous samples from the native and hatchery populations represents an accumulation of changes along two lines of descent, one for the natural population, the other for the hatchery stock (Fig. 2). If t is not large, the expected value of this variance becomes $E(F) \sim t/(2N_{ne}) + t/(2N_{he}) + 1/(2S_n) + 1/(2S_h)$, where N_{ne}, N_{he}, S_n and S_h are the effective population numbers and sample sizes of the native and hatchery populations, respectively. This reduces to $E(F) \sim t/(N_e) + 1/(2S_n) + 1/(2S_h)$, where N_e is the harmonic mean of the effective population numbers *over both lines of descent*. Note that the estimator for N_e, $\hat{N}_k = t/[F - 1/(2S_n) - 1/(2S_h)]$, lacks the 2 present in the denominator of the estimator for the case in which the initial sample is from the parental gametic pool (see Methods). Thus, given the variances in allelic frequencies between native and hatchery stocks in Tables 2 to 4, the corresponding $\hat{N}_k$s should be multiplied by two and interpreted as harmonic mean effective population numbers for each pair of native and hatchery populations, over the time interval since founding of the hatchery stock. The $\hat{N}_k$ estimates given in these tables for hatchery stocks would be approached asymptotically as N_{ne} approached infinity. Thus, it is important to estimate genetic drift and N_e in the natural source populations. For those cases in which N_b is already below the 95% C.L. for $\hat{N}_k$, the discrepancy between breeding and effective population numbers may be even larger than indicated if allelic frequencies in the natural population undergo appreciable genetic drift.

Genetic drift has been measured in five natural populations of oysters, using allozyme data sets for samples collected at different times in the same localities. Estimates of N_e range from 15 to infinity, but most are several orders of magnitude less than estimated abundances (Table 5). The number of adult oysters harvested per year from Dabob Bay, for example, is estimated to be on the order of 10^8 (R. Birge, Washington State Department of Fisher-

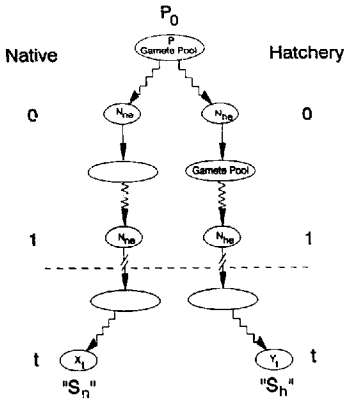

Fig. 2. Sampling plan for estimating temporal variance in a contemporaneous comparison of a hatchery stock to its native progenitor population. P_0 is the natural parental population which produces a gametic pool from which the generation 0, founding hatchery broodstock, are taken; x_t and y_t and S_n and S_h are allelic frequencies and numbers of individuals sampled for genetic analysis, respectively, in the native and hatchery populations at the t^{th} generation. N_{ne} and N_{he} are likewise the effective population numbers for these two populations. Wavy lines represent the binomial random sampling process whereby populations or samples are obtained from gametic pools.

ies, pers. commun.); the largest point estimate of N_e in this population is six orders of magnitude smaller. We discuss later an hypothesis to account for such discrepancies, but conclude here that genetic drift in natural populations cannot be assumed to be negligible.

The assumption that there has been no immigration into a closed broodstock population from either another hatchery stock or the native population is difficult to evaluate. Substantial temporal change in broodstocks, divergence of these stocks from native populations, and divergence among separate stocks propagated by the same hatchery (Wada, 1986; Hedgecock and Sly, 1990; Vrijenhoek et al., 1990) seemingly evidence genetic isolation, but admixture of just a few individuals might bias estimates of F and N_e without changing the qualitative evidence for divergence of a stock. Genetic studies of bivalves have often detected low level contamination of hatchery cultures despite careful husbandry (see references in Gaffney and Allen, 1993).

The effect of immigration on temporal genetic change within a broodstock population depends on how genetically different the immigrants are from both the founding population and the derived stock. On one hand, admixture of individuals from a large, genetically stable, native population from which the

stock was originally derived should pull allelic frequencies back toward their original values, retard genetic drift, reduce $\bar{F}$ and increase $\hat{N}_k$. On the other hand, admixture of individuals from a population distinct from either the native source population or the hatchery stock, e.g. a different geographic native population or another hatchery stock, could have diverse, locus-specific effects on $\bar{F}$ and $\hat{N}_k$. At some loci $\hat{F}$ might be reduced, while at others $\hat{F}$ might be increased relative to drift expectations; in general, mean $\hat{F}$ should increase with the degree of genetic divergence between immigrants and the broodstock. Known admixtures of broodstock in two cases (Dillon and Manzi, 1987; Sbordoni et al., 1987) illustrate the diverse effects that immigration can have on estimates of temporal variance.

Dillon and Manzi (1987) report that a small and unknown number of native hard clams were added to the ARC broodstock over the course of four generations. Mean temporal variance in the ARC stock of hard clams, 0.036, is less than in the VIMS stock, 0.050, even though both stocks were supposedly propagated by 30–60 broodstock in each generation. Although these breeding numbers fall well within the 95% confidence limits of the $\hat{N}_k$ estimates for both clam stocks, it is interesting that the point estimate of $\hat{N}_k$ for the ARC stock, 81.5, is well above 60, the maximum N_b. The greater estimated effective population number for the ARC stock may be attributable to continual immigration of small numbers of native broodstock. Allelic frequency changes in the ARC stock are nevertheless indistinguishable from random genetic drift.

Potential effects of immigration in the Italian population of Kuruma shrimp are much more complex. An unknown number of broodstock, which were reared from a single batch of postlarvae imported from a Japanese hatchery, were added to the Italian shrimp stock at the fifth generation. These immigrants had substantially different allelic frequencies and noticeably reduced genetic variability compared to the first generation broodstock (see Sbordoni et al., 1986 for details). Their introduction appears to have caused unusual shifts in allelic frequencies at the sixth generation, so that mean $\hat{F}$ is 0.169 for the F5–F6 interval, compared to estimates of 0.062 and 0.067 for the F4–F5 and F6–F7 intervals, respectively. Interestingly, the probability plot of $8\hat{F}/\bar{F}$ for the F1–F6 interval has the most exaggerated inflection (Fig. 1A) and is the only one to deviate significantly from the corresponding χ^2 distribution. Immigration of broodstock that were themselves hatchery-propagated appears to have inflated $\hat{F}$ for the *Ca-1* locus but reduced it at the *Est-2*, *Ldh-1* and *Phi* loci, relative to their drift expectations. The probability plot for the F1–F7 interval retains some of the inflection signature left by the immigration event. It would be useful to determine whether such inflection of the $n\hat{F}/\bar{F}$ probability plot might be diagnostic of admixture.

Immigration, particularly admixture among hatchery stocks, might also explain four cases in which breeding numbers, N_b, are below the lower 95% con-

fidence limits of effective numbers, $\hat{N}_k$. The first of these is the SWSL-II line of pearl oysters, which was propagated alongside three other lines (Wada, 1986) and for which $N_b = 15.2$ and the lower C.L. of $\hat{N}_k = 22.3$ (Table 3); N_bs for the other three lines are smaller than the $\hat{N}_k$s but still within the confidence limits. Of the five MSX-resistant strains of American oysters developed at Rutgers University by Dr. Haskin (see Vrijenhoek et al., 1990), the DB-1, DB-2, and NA lines had breeding numbers smaller than the lower C.L. of $\hat{N}_k$ (Table 4). Interestingly, these three lines were propagated side by side each year, whereas the VA and LI lines, which show a better agreement of N_b and $\hat{N}_k$, were spawned each at separate times (Dr. Stan Allen, Rutgers University, Haskin Shellfish Research Laboratory, pers. communication). The SWSL-II and DB-2 stocks also had significantly fewer alleles remaining than expected on the basis of random genetic drift, the only significant results obtained in a total of 24 such tests. This observation is consistent with admixture among hatchery stocks driving frequencies of common alleles back towards initial conditions — thus reducing apparent drift variance and increasing apparent N_e — but failing to restore less common alleles lost in early generations (Fuerst and Maruyama, 1986). Other potential explanations of why N_e would be greater than N_b — linked genes, selection for heterozygosity (Vrijenhoek et al., 1990), faulty data, or less-than-binomial variance in reproductive success among individuals (see below) — cannot account simultaneously for this loss of alleles.

In four broodstocks (the Italian Kuruma shrimp, and three hatchery stocks of Pacific oysters) and three natural populations of oysters, breeding numbers appear to be substantially greater than the upper 95% confidence limit of $\hat{N}_k$. This result is consistent with previous observations that actual populations appear to be very much larger than their effective population numbers for a broad spectrum of animal species (Hedgecock et al., 1982; Nei, 1983; Nei and Graur, 1984). Such large discrepancies can be explained either by widespread population bottlenecks (Nei and Graur, 1984) or, for highly fecund organisms, by large variances in reproductive success, as measured by numbers of offspring contributed per parent to the breeding population of the next generation (Hedgecock et al., 1982).

For many aquatic animals, large variances in reproductive success are made possible by a combination of great fecundity with sweepstakes survival of broods and their subsequent recruitment to the breeding population. In aquaculture situations, hatchery seed may often come from just a few successfully spawning individuals (e.g. Lannan, 1980; Gharrett and Shirley, 1985; Withler, 1988; Gile and Ferguson, 1990; Gaffney et al., 1992). Variance in offspring contribution, V_k, is related to the effective population number by $N_e = (4N_b - 4)/(V_k + 2)$, assuming population size is constant and males and females have equal progeny distributions (see Crow and Denniston, 1988). V_k can be several orders of magnitude larger than the mean number of off-

spring per parent in a population of highly fecund aquatic animals, reducing effective population numbers to small fractions of breeding numbers. We return below to the counter-intuitive implication that if V_k is artificially held below the binomial or Poisson variance allowed by genetics theory, random genetic drift is retarded relative to that in the theoretically ideal population, and the effective number becomes larger than the number of broodstock.

Small N_e in aquatic broodstocks may result simply from use of insufficient numbers of breeders or inequality of sex ratios; these problems are easily corrected. However, variance in reproductive success may be a more potent and insidious factor reducing N_e; the extremely high fecundities and variable fertilities of aquatic organisms must be managed properly to avoid loss of genetic diversity from this source. Genetic drift can be controlled by the development and use of pedigreed broodstocks, in which variance in reproductive success and inbreeding can be monitored directly. Indeed, as noted above, there is opportunity to increase N_e by artificially reducing or eliminating V_k; for example, when $V_k=0$, $N_e=2(N_b - 1)$. Achieving such control over reproductive contributions in closed aquaculture broodstocks will necessitate changes in common management practices. Separation of pedigreed broodstock from production broodstock, for example, may be necessary.

The development of large-scale hatchery technology and closed broodstock populations have been necessary first steps towards founding of aquatic livestocks. By themselves, however, these steps are not sufficient to guarantee genetic gains. Estimates of N_e for 16 shellfish broodstocks, all less than 100, 13 less than 50, are smaller than recommended for long-term maintenance of genetic diversity (see Hedrick and Miller, 1992). Although conservation of genetic diversity is not the primary goal of genetic improvement programs, rapid loss of genetic diversity by drift in the early stages of domestication could preclude subsequent genetic gains. On the other hand, conservation of genetic diversity in hatchery-propagated aquatic stocks is a primary concern if hatchery products are used to enhance natural populations. In either case, the temporal method provides a powerful means for monitoring genetic drift and estimating the effective population numbers of closed aquatic broodstocks from readily obtained allozyme data.

REFERENCES

Allendorf, F.W., 1986. Genetic drift and the loss of alleles versus heterozygosity. Zoo Biol., 5: 181–190.

Buroker, N.E., 1975. A survey of protein variation in populations of the Pacific oyster, *Crassostrea gigas*. MS dissertation, Univ. Washington, Seattle.

Buroker, N.E., 1983. Population genetics of the American oyster *Crassostrea virginica* along the Atlantic coast and the Gulf of Mexico. Mar. Biol., 75: 99–112.

Buroker, N.E., 1984. Genetic differentiation and population structure of the American oyster *Crassostrea virginica* (Gmelin) in Chesapeake Bay. J. Shellfish. Res., 3: 153–167.

Crow, J.F. and Denniston, C., 1988. Inbreeding and variance effective population numbers. Evolution, 42: 482–495.

Dillon, R.T. and Manzi, J.J., 1987. Hard clam, *Mercenaria mercenaria*, broodstocks: genetic drift and loss of rare alleles without reduction in heterozygosity. Aquaculture, 60: 99–105.

Ewens, J.J., 1979. Mathematical Population Genetics. Springer-Verlag, Berlin, 325 pp.

Fuerst, P.A. and Maruyama, T., 1986. Considerations on the conservations of alleles and of genic heterozygosity in small managed populations. Zoo Biol., 5: 171–179

Gaffney, P.M. and Allen, S.K., 1993. Hybridization among *Crassostrea* species: a critical review. Aquaculture, under review.

Gaffney, P.M., Davis, C.V. and Hawes, R.O., 1992. Assessment of drift and selection in hatchery populations of oysters (*Crassostrea virginica*). Aquaculture, 105: 1–20.

Gharrett, A.J. and Shirley, S.M., 1985. A genetic examination of spawning methodology in a salmon hatchery. Aquaculture, 47: 245–256.

Gile, S.R. and Ferguson, M.M., 1990. Crossing methodology and genotypic diversity in a hatchery strain of rainbow trout (*Oncorhynchus mykiss*). Can. J. Fish. Aquat. Sci., 47: 719–724.

Gnanadesikan, R., 1977. Methods for Statistical Data Analysis of Multivariate Observations. J. Wiley and Sons, New York.

Hedgecock, D. and Sly, F.L., 1990. Genetic drift and effective population sizes of hatchery-propagated stocks of the Pacific oyster *Crassostrea gigas*. Aquaculture, 88: 21–38.

Hedgecock, D., Tracey, M.L. and Nelson, K., 1982. Genetics. In: L. G. Abele (Editor), The Biology of Crustacea, Vol. 2. Academic Press, New York, pp. 297–403.

Hedrick, P.W. and Miller, P.S., 1992. Conservation genetics: techniques and fundamentals. Ecol. Appl., 2: 30–46.

Krimbas, C.B. and Tsakas, S., 1971. The genetics of *Dacus oleae*. V. Changes of esterase polymorphism in a natural population following insecticide control. Selection or drift? Evolution, 25: 454–462.

Lannan, J.E., 1980. Broodstock management of *Crassostrea gigas*. I. Genetic variation in survival in the larval rearing system. Aquaculture, 21: 323–336.

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

Lewontin, R.C. and Krakauer, J., 1975. Testing the heterogeneity of *F* values. Genetics, 80: 397–398.

Mueller, L.D., Wilcox, B.A., Ehrlich, P.R., Heckel, D.G. and Murphy, D.D., 1985. A direct assessment of the role of genetic drift in determining allele frequency variation in populations of *Euphydryas editha*. Genetics, 110: 495–511.

Nei, M., 1983. Genetic polymorphism and the role of mutation in evolution. In: M. Nei and R. Koehn (Editors), Evolution of Genes and Proteins. Sinauer Associates, Sunderland, MA, pp. 165–190.

Nei, M. and Graur, D., 1984. Extent of protein polymorphism and the neutral mutation theory. Evol. Biol., 17: 73–118.

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

Paynter, K.T. and DiMichele, L., 1990. Growth of tray-cultured oysters (*Crassostrea virginica* Gmelin) in Chesapeake Bay. Aquaculture, 87: 289–297.

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

Sbordoni, V., De Matthaeis, E., Cobolli-Sbordoni, M., La Rosa, G. and Mattoccia, M., 1986. Bottleneck effects and the depression of genetic variability in hatchery stocks of *Penaeus japonicus* (Crustacea, Decapoda). Aquaculture, 57: 239–251.

Sbordoni, V., La Rosa, G., Mattoccia, M., Cobolli-Sbordoni, M. and De Matthaeis, E., 1987. Genetic changes in seven generations of hatchery stocks of the Kuruma prawn, *Penaeus ja-*

ponicus (Crustacea Decapoda). In: K. Tiews (Editor), Selection, Hybridization and Genetic Engineering in Aquaculture. Heenemann Verlag, Berlin, pp. 143–155.

Taniguchi, N., Sumantadinata, K. and Iyama, S., 1983. Genetic change in the first and second generations of hatchery stock of black seabream. Aquaculture, 35: 309–320.

Vrijenhoek, R.C., Ford, S.E. and Haskin, H.H., 1990. Maintenance of heterozygosity during selective breeding of oysters for resistance to MSX disease. J. Hered., 81: 418–423.

Wada. K.T., 1986. Genetic variability at four polymorphic loci in Japanese pearl oysters, *Pinctada fucata martensii*, selected for six generations. Aquaculture, 59: 139–146.

Wilkinson, L., 1987. SYSTAT: The System for Statistics. SYSTAT Inc., Evanston, IL, 417 pp.

Waples, R. S., 1989. A generalized approach for estimating effective population size from temporal changes in allele frequency. Genetics, 121: 379–391.

Waples, R.S. and Teel, D.J., 1990. Conservation genetics of Pacific salmon. I. Temporal changes in allele frequency. Conserv. Biol., 4: 144–156.

Withler, R.E., 1988. Genetic consequences of fertilizing chinook salmon (*Oncorhynchus tshawytscha*) eggs with pooled milt. Aquaculture, 68: 15–25.