2006

Seed Banks, Salmon, and Sleeping Genes: Effective Population Size in Semelparous, Age-Structured Species with Fluctuating Abundance

Robin Waples
National Marine Fisheries Service, robin.waples@noaa.gov

Ⓔ

Seed Banks, Salmon, and Sleeping Genes: Effective Population Size in Semelparous, Age-Structured Species with Fluctuating Abundance

Robin S. Waples[*]

National Marine Fisheries Service, Northwest Fisheries Science Center, Seattle, Washington 98112; and Museum of Vertebrate Zoology, University of California, Berkeley, California 94720

ABSTRACT: Previous studies reached contrasting conclusions regarding how fluctuations in abundance affect N_e in semelparous species with variable age at maturity: that N_e is determined by the arithmetic mean N among the T years within a generation ($N_e \approx T\overline{N}_t$; monocarpic plants with seed banks) or the harmonic mean ($N_e \approx T\tilde{N}_t$; Pacific salmon). I show that these conclusions arise from different model assumptions rather than inherent differences between the species. Sequentially applying standard, discrete-generation formulas for inbreeding N_e to a series of nominal generations accurately predicts the multigenerational rate of increase in inbreeding. Variability in mean realized reproductive success across years ($\overline{k}_t$) is the most important factor determining N_e and N_e/N. When abundance is driven by random variation in $\overline{k}_t$, $N_e \leq T\tilde{N}_t < T\overline{N}_t$. With random variation in N_t and constant per capita seed production (C), variation in $\overline{k}_t$ is low and $N_e \sim T\tilde{N}_t$; however, if C varies among years, N_e can be closer to $T\tilde{N}_t$. Because population regulation affects the genetic contribution of entire cohorts of monocarpic perennials, N_e for these species may be more closely approximated by $T\tilde{N}_t$ than by $T\overline{N}_t$. With density-dependent compensation, $\mathrm{Cov}(k_t, N_t) < 0$, and N_e is further reduced because relatively few breeders make a disproportionate contribution to the next generation.

Keywords: effective population size, fluctuating population size, semelparity, monocarpy, overlapping generations, density dependence.

* Address for correspondence: National Marine Fisheries Service, Northwest Fisheries Science Center, 2725 Montlake Boulevard East, Seattle, Washington 98112; e-mail: robin.waples@noaa.gov.

It might appear that Pacific salmon (*Oncorhynchus* spp.) and monocarpic plants with seed banks have little in common. After all, they are in different kingdoms; plants are autotrophs and are stationary as adults, whereas salmon are heterotrophs and can migrate thousands of kilometers in fresh water and the ocean. However, both groups share an unusual combination of life-history traits (semelparity with variable age at maturity) that have important evolutionary consequences. Seed banks serve as a genetic reservoir of "sleeping genes" that can awaken and contribute their genetic material—a phenomenon that can extend the generation time, increase effective population size (N_e), and slow the rate of genetic change (Templeton and Levin 1979; Hairston and De Stasio 1988; Kaj et al. 2001). In monocarpic perennials, vegetative rosettes provide a genetic storage effect equivalent to that of a seed bank (Vitalis et al. 2004). A somewhat similar role in Pacific salmon is played by cohorts of juveniles and subadults, which can be thought of as carrying sleeping genes that awaken when the fish return to their natal stream to spawn.

Recently, two articles considered how annual variation in abundance affects N_e in semelparous species and came to contrasting conclusions. Nunney (2002) concluded that in annual plants with seed banks, N_e per generation is primarily determined by the arithmetic mean of the population sizes in the years making up a generation; for Pacific salmon, I found that N_e is more strongly influenced by the harmonic mean of the annual sizes (Waples 2002).

The purpose of this article is twofold. First, I show that the apparent discrepancy in the conclusions of the two articles is due more to different model assumptions than to inherent differences between salmon and monocarpic plants. Second, I consider life-history and demographic factors that influence genetic variability and show that, in these species, N_e and the ratio N_e/N are determined primarily by the method of population regulation—specifi-

cally, the degree to which random and deterministic processes lead to variation in mean reproductive success among breeders in different years. These results provide a basis for evaluating the genetic consequences of demographic and life-history characteristics of salmon, monocarpic plants, and other semelparous species with variable age at maturity, such as crustaceans with diapausing eggs.

Description of the Models

Model description involves three steps. First, I summarize common features of the two models and review the notation used here (table 1). Second, I describe general life-history differences between salmon and plants with seed banks. Finally, I review specific assumptions made by previous authors in implementing the two models.
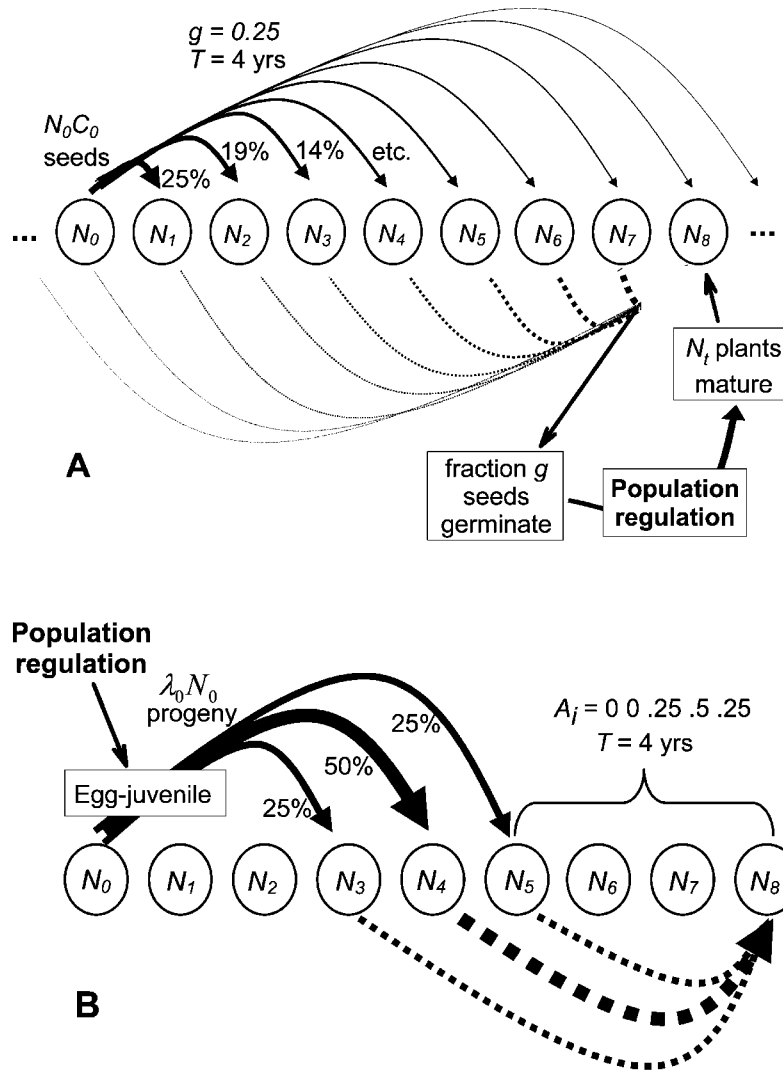
Common Features

In both models, in year t a certain number of individuals (N_t) mature and reproduce (fig. 1). Their progeny mature in future years according to a fixed maturity schedule A_i, where A_i represents the probability of maturing in year $t + i$. If population size is constant, the age distribution of adult breeders is also given by A_i, and generation length = average age of parents is given by $T = \sum_{i=1,M} iA_i$, where M is the maximum age at reproduction. If population size fluctuates, age structure will as well, and $\sum iA_i$ will only approximate the generation length.

The N_t breeders in year t produce a total of R_t offspring that eventually mature as adults in subsequent years. Demographic contributions to the next generation are quantified by the ratio $\lambda_t = R_t/N_t$, which is the replacement rate (or population growth rate) for adults reproducing in that year. The magnitude of variation in λ_t is measured

Table 1: Notation

Parameters	Definition
Demographic:	
N_t	Number of adults reproducing in year t
M	Maximum age at reproduction
A_i	Probability that a progeny produced in year t matures in year $t + i$ ($\sum_{i=1,M} A_i = 1$)
T	Generation length (mean age of parents) in years ($T = \sum_{i=1,M} iA_i$)
N_T	Total number of adults in a generation ($N_T = \sum_{t=1,T} N_t$)
$\overline{N}_t$	Arithmetic mean of the N_t within a generation
$\tilde{N}_t$	Harmonic mean of the N_t within a generation
J_N	Index of population fluctuation based on log-transformed values of N_t; $J_N = 10^{2\sigma L}$, where σL is the standard deviation of log (N_t)
C_t	Per capita production of seeds by plants flowering in year t
J_C	Analogue to J_N that quantifies yearly variation in C_t
B_t	Total number of seeds in the seed bank in year t
$r_{t,i}$	Number of offspring maturing in year $t + i$ produced by adults reproducing in year t
R_t	Total number of adult offspring in the next generation produced by adults reproducing in year t ($R_t = \sum_{i=1,M} r_{t,i}$)
λ_t	Mean number of adult offspring in the next generation produced by each adult reproducing in year t ($\lambda_t = R_t/N_t$)
σ_λ	Standard deviation of λ across years
R_T	Total adults in the next generation produced by all the adults in the current generation ($R_T = \sum_{t=1,T} R_t$)
X_t	Proportional contribution of adults reproducing in year t to the next generation ($X_t = R_t/R_T$)
Genetic:	
k	Number of gametes contributed to the next generation by an individual
$\overline{k}_t$	Mean k for individuals breeding in year t ($\overline{k}_t = 2\lambda_t$)
$\overline{K}$	Mean k for all individuals breeding within a generation ($\overline{K} = 2R_T/N_T$)
$\overline{k}_*$	Mean of the $\overline{k}_t$ within a generation
$\sigma^2_{k(t)}$	Variance in k among individuals within year t
V_k	Variance in k among all individuals within a generation
$N_{b(t)}$	Effective number of breeders in year t
N_e	Effective population size for a generation
$N_{e(n)}$	Effective population size over a period of n generations ($N_{e(n)} \approx$ harmonic mean of the N_e in the separate generations)

Figure 1: Schematic diagram of life-history features of the two types of species. Common features of both models: *circles*, N_t adults maturing in year *t*; *curved solid arrows*, reproductive contribution of 1 year of adults to future years, with size of the arrows proportional to the fraction maturing at each age (A_i). Relative contribution of different parental years to a future year of adults (*curved dotted arrows*) depends not only on A_i but also on annual population size and productivity. Life stage typically subject to population regulation is indicated. *A*, Seed-bank model. N_t adults flowering in year *t* produce a total of C_tN_t seeds that enter the seed bank. Probability of seed germination is constant at *g*/year, which, assuming no mortality in the seed bank, leads to a generation length of $T = 1/g$ years. Of the C_tN_t seeds contributed to the seed bank in year *t*, the largest fraction (*g*) mature in year $t + 1$, the next largest fraction ($g[1 - g]$) mature in year $t + 2$, etc. In this example, $g = 0.25$, so $T = 4$ (only 8 years of seed germination are shown here). In a given year (e.g., year 8), the fraction *g* of all the seeds in the seed bank germinate, but due to variable carrying capacity, only N_t are allowed to mature. *B*, Salmon model. N_t adults spawning in year *t* produce a total of $\lambda_tN_t = (\bar{k}_t/2)N_t$ progeny that survive to reproduce in future years. In this example, maturity occurs at ages 3, 4, or 5, so a nominal generation includes $T = \Sigma iA_i = 4$ years of spawners, and the N_t adults spawning in a given year (e.g., year 8) are derived from parents spawning 3–5 years previously.

by σ_λ. Over the *T* years making up a generation, the total number of breeders ($N_T = \Sigma N_t$) contributes a total of $R_T = \Sigma R_t$ offspring that mature in subsequent years, so R_T/N_T is a measure of population growth rate per generation.

Yearly genetic contributions are quantified by $\bar{k}_t$ and $\sigma^2_{k(t)}$, the mean and variance (among individuals within a single year) of the number of gametes contributed to the next generation, and $\bar{K}$ and V_k are the corresponding values computed across all individuals within a generation. Since we are dealing with diploid species, $\bar{k}_t = 2\lambda_t$ and $\bar{K} = 2R_T/N_T$.

Species-Specific Details

Seed Bank. In the seed-bank model (fig. 1A), in year t, N_t monoecious plants mature, flower, and die. Each adult produces an average of C_t seeds, and C is generally high to ensure that the seed bank is large compared with the number of adults and hence does not limit future population size. Each seed in the seed bank has a fixed annual probability g of germinating; as a consequence, the proportion of a cohort of seeds that germinate in each successive year (the A_i values) is a monotonically (and exponentially) decreasing function of time (fig. 1A). If a seed germinates, it either becomes a flowering adult that year or dies before maturing. If no mortality occurs in the seed bank, the average age of germinating seedlings is $1/g$; if population size is constant, $1/g$ also represents the average age of parents and hence the generation time, T.

Salmon. In the Pacific salmon model (fig. 1B), each female in year t contributes several thousand eggs, but high mortality in juvenile and early ocean phases largely determines the growth rate (λ_t) for that cohort. Variation in male reproductive success could be modeled separately but is generally considered to be comparable to that of the female. The $\lambda_t N_t$ progeny from adults in year t that survive to maturity return to spawn at a variety of ages i, with i typically ranging from ≥ 2 to ≤ 6 (leading to $T \approx 3$ to 5 years). Age at maturity within a cohort can be evenly spread over two or more years or, more commonly, dominated by a single age.

Previous Implementations of the Two Models

Seed Bank. In Nunney's (2002) implementation of the seed-bank model, each plant in every year produced exactly C seeds. Survival in the seed bank was random, and genotypes in the N_t successful adults in year t were randomly drawn from seeds germinating that year. Within each year, therefore, the maturing adults approximated an ideal Wright-Fisher population in which the variance in realized reproductive success was Poisson ($\sigma^2_{k(t)} \approx \overline{k}_t$) and the effective number of breeders ($N_{b(t)}$) was the same as the census number (N_t).

Variation in population size was accomplished by randomly drawing a number N_t from a lognormal distribution with a specified geometric mean (generally 100; Nunney also considered a scenario involving random fluctuation between "good" and "bad" years, but that variation will not be considered here). An index of variability in N_t was calculated as $J_N =$ the ratio of (log-transformed) population size 1 SD above the mean to population size 1 SD below the mean. A typical time series of N_t values using this model is shown in figure 2 (*top*).

Under the assumption that mating is random and selfing does not occur, Nunney (2002, eq. [28]) obtained the following approximation for effective size per generation:

$$N_e \approx \tilde{N}_t + (T - 1)\overline{N}_t, \tag{1}$$

where $\overline{N}_t$ and $\tilde{N}_t$ are the arithmetic and harmonic means of the N_t, respectively. As T increases, the $\overline{N}_t$ term in equation (1) quickly dominates. Consequently, Nunney (2002) concluded that in most situations, N_e in annual plants with seed banks is determined primarily by the arithmetic mean of the number of adult plants per year.

Vitalis et al. (2004) also evaluated a seed-bank model that allowed consideration of monocarpic perennials. Their model allowed a maximum longevity to be imposed on seeds in the seed bank as well as an annual cost (risk of mortality), both of which may be more realistic than the assumptions of Nunney's model. However, Vitalis et al. only considered asymptotic behavior in populations of constant size, so most of their results are not directly relevant to the issues discussed here. Nevertheless, their treatment of density dependence is illuminating and is considered below in the "Discussion."

Salmon. Like Nunney, my primary interest has been the effects of variable population size on N_e per generation rather than variation among individuals within years, and hence for simplicity I have assumed (Waples 2002) that the population is ideal each year so that $N_{b(t)} = N_t$. Previously (Waples 1990a), I showed that if population size is constant in the salmon model, effective size per generation (N_e) is approximately equal to N_t times the generation length: $N_e \approx TN_t$. Based on limited evaluations of variable population size, I also concluded (Waples 1990b) that effective size per generation is a function of the harmonic mean, not arithmetic mean, of the N_t across the years making up a generation:

$$N_e \approx T\tilde{N}_t. \tag{2}$$

Subsequently, however, it became evident that the relationship between N_e and N_b in salmon depends on the population dynamic model and that different demographic scenarios lead to contrasting genetic results (Waples 2002).

Scenario 1. Each year's breeding population contributes an equal number of progeny to the next generation regardless of the number of breeders. This is similar to the model considered by Waples (1990b). With R_t fixed within a generation at $R_t = R_T/T$, if N_t varies, the mean replacement rate ($\lambda_t = R_t/N_t = \overline{k}_t/2$) also varies across years. Under these conditions, N_e per generation is a function of the harmonic mean of N_t (eq. [2]).
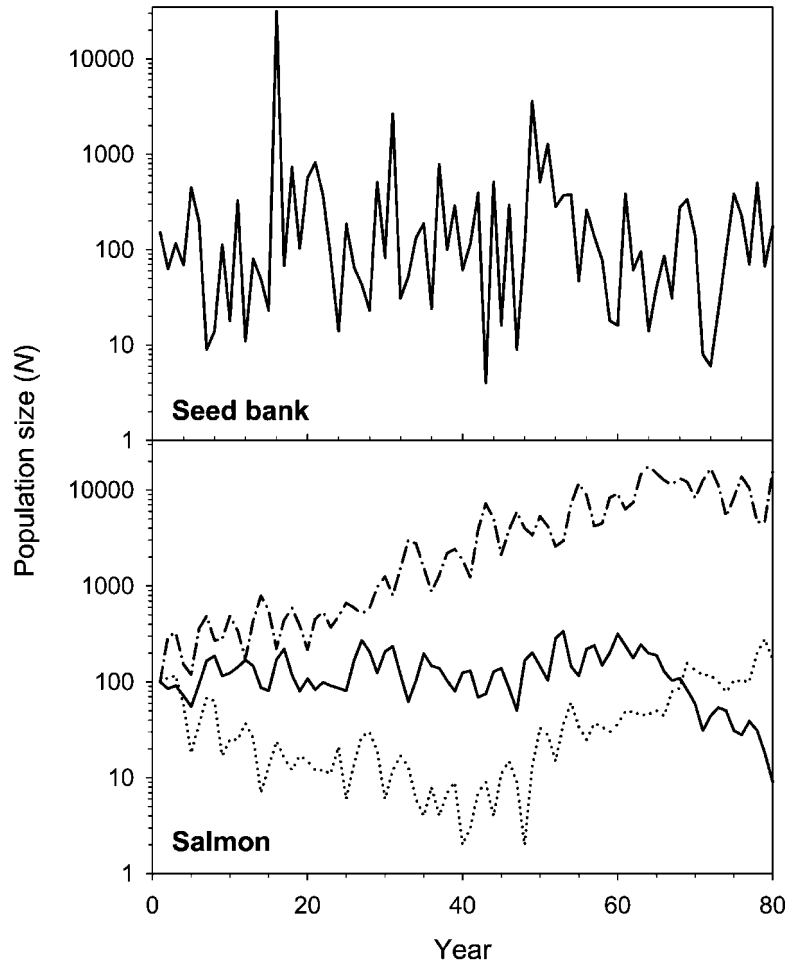
Figure 2: Contrasting patterns of variation in population size in the two models. In the seed-bank model (*top*), N_t is a random, lognormally distributed variable. Data shown are one time series with J_N = index of variability in $N = 25 \times$. In the salmon model (*bottom*), variation in N_t is generated by random, lognormal variation in λ, resulting in a higher-order Markov process. Shown are three replicate time series with $\sigma_\lambda = 1.0$ and $T = 4$. In both models, initial $N_0 = 100$, and population size was constrained by $2 \leq N_t \leq 30{,}000$.

Scenario 2. Each year's breeding population contributes to the next generation in direct proportion to the number of breeders. In this scenario, λ_t and $\overline{k}_t$ are constant across years within a generation, and N_e is a function of the arithmetic mean N_t:

$$N_e = \sum_{t=1,T} N_t = T\overline{N}_t = N_T. \tag{3}$$

Scenario 2 is also referred to as the additive model because $N_e = N_T$ is simply the sum of all the mature adults across a generation.

Neither scenario is completely realistic. If λ_t is constant, population size quickly stabilizes at a constant value, apart from small random fluctuations in age at maturity. The same result occurs if R_t is fixed over time. Thus, although

either of these two scenarios could apply to any particular generation, neither is a plausible general description of demographic processes in fluctuating salmon populations. Accordingly, I used computer simulations to model changes of allele frequency in salmon populations that fluctuated in size due to random variation in λ across years (Waples 2002). Three typical time series of N_t values using this type of model are shown in figure 2 (*bottom*). Simulation results (Waples 2002) showed that for most parameter sets considered, variance $N_e \leq T\tilde{N}_t < T\overline{N}_t$. Analytical evaluation of a time series of data for one salmon population led to a similar result: in most generations, inbreeding N_e calculated from demographic data was less than $T\tilde{N}_t$ and much less than $T\overline{N}_t$. Based on these collective results, Waples (2002) concluded that in most Pacific

salmon populations, equation (2) (harmonic mean method) is the better predictor of N_e.

The salmon and seed-bank models thus have a number of similarities: the populations are closed (Nunney also presented results for an island model with migration); the breeders each year represent an "ideal" population; the genes considered are assumed to be neutral, and gene frequency change was tracked for a single, diallelic locus; and each replicate simulation ran for 80–100 years. However, the two models have important differences in age structure, magnitude of variation in N_p, and mechanisms of population regulation. To evaluate these factors quantitatively, I used a combination of analytical and numerical methods that focus on inbreeding effective size.

Methods

Analytical Methods for Computing N_e

Inbreeding N_e quantifies the per generation rate of increase in inbreeding; it is proportional to the inverse of the probability f that two homologous alleles in a randomly chosen individual are identical by descent (IBD; derived from the same gene copy in the same individual in a prior generation): $N_e = 1/(2f)$. If random selfing occurs, the inbreeding coefficient in generation $t(f_t)$ is determined by N_e the previous generation: $E(f_t) = 1/(2N_{e(t-1)})$. If selfing is not allowed, alleles that are identical by descent cannot unite in individuals until the grand progeny generation, so f is a function of N_e two generations before: $E(f_t) = 1/(2N_{e(t-2)})$.

I used an analytical method modified from Ryman and Laikre (1991) to quantify the effects of population dynamic processes on f and inbreeding effective population size over single- and multiple-generation intervals. Ryman and Laikre (1991) considered the effective size of a population that is a composite of two or more groups of individuals with different effective sizes and mean reproductive rates. The inbreeding effective size of the composite population is

$$N_e = \frac{1}{\Sigma \left(\frac{X_t^2}{N_t} \right)}, \tag{4}$$

where N_t is the effective size of segment t and $X_t = R_t/R_T$ is the proportional contribution of breeders in segment t to the progeny generation. Ryman and Laikre used this approach to analyze the effective size of a hatchery-wild system in which one population segment is taken into captivity and enjoys enhanced reproductive success. For species that reproduce only once, this same approach can

be applied to the years comprising a generation (Waples 2002).

This problem can also be formulated in a more traditional way using the familiar equation for inbreeding effective size (Crow and Kimura 1970; Crow and Denniston 1988):

$$N_e = \frac{\overline{K}N_T - 2}{\overline{K} - 1 + \frac{V_k}{\overline{K}}}, \tag{5}$$

where $N_T = \Sigma N_t$ is the census population size for a generation and $\overline{K}$ and V_k are calculated across all individuals within the generation. If selfing does not occur, equation (5) applies both to monoecious species and to species with separate sexes (Crow and Denniston 1988). It can be shown that equation (5) is mathematically equivalent to equation (4) (see app. A in the online edition of the *American Naturalist*). Equation (5) is more convenient and explanatory of the N_e/N ratio when data are expressed in terms of generational $\overline{K}$ and V_k, whereas equation (4) is better suited for use with data expressed in terms of mean progeny production per year.

These equations also have a straightforward interpretation for variable population size in age-structured, semelparous species. Because yearly reproductive output also represents lifetime reproduction in these species, calculation of $\overline{K}$ and V_k across all individuals in a nominal generation provides a way to quantify the increase in inbreeding contributed by breeders in that generation. Two differences compared with the discrete generation model can be noted. First, in age-structured populations, it can take some time for IBD alleles to percolate across years and unite in individuals. Therefore, the increase in f arising from matings in generation t will not necessarily all occur by generation $t + 2$. However, equation (5) describes the increase in IBD alleles in the population as a whole in generation $t + 1$, and it quantifies the eventual contribution of generation t breeders to f in subsequent generations. The second difference is that in age-structured populations, although it is possible to define an average generation length T, within any particular time series the choice of the start and end of a generation is rather arbitrary. Using empirical data, Waples (2002) showed that N_e calculated using equation (4) can vary depending on which specific years are included in a nominal generation. It is thus important to ask the following question: if equation (4) or (5) were applied sequentially to a time series of demographic data partitioned into nominal generations of length T years, would the result accurately predict the true multigenerational effective size of the population?

Individual-Based Genetic Model

To evaluate this question quantitatively, I used an individual-based computer model that followed the population dynamics of the salmon model but also considered age structures more typical of plants with seed banks. Three basic steps (described in more detail below) were involved: (1) A time series of variable population sizes was generated, and demographic data were used to compute multigenerational N_e using three methods: the harmonic mean method ($N_e = T\tilde{N}_t$; eq. [2]), the arithmetic mean method ($N_e = T\overline{N}_t = N_T$; eq. [3]), and the discrete-generation model (eq. [4]); (2) The rate of increase in f for individuals following this population size trajectory was monitored; (3) The multigenerational N_e's were used to compute expectations of f, which were compared with observed f from the simulated populations to evaluate performance of the three methods.

Population Size Trajectories. Each year t ($t = 1$ to 80), a value for λ_t was drawn randomly from a lognormal distribution with specified mean (generally $\bar{\lambda} \approx 1$) and standard deviation σ_λ. The size of the first M years is fixed at $N_t = N_0$ individuals. The N_t adults maturing in year t produced a total of $R_t = N_t\lambda_t$ progeny that matured in future years. Age at maturity for each individual was chosen by a random number with probability of maturing at year $t + i$ equal to A_i. This model leads to a higher-order Markov chain with N_t depending on population sizes and growth rates in several years prior to year t (fig. 2, *bottom*). This presents a difficulty for averaging f and N_e across replicates because the different population size trajectories are not independent and identically distributed (IID). To get around this difficulty, for each population size trajectory I replicated the process of genetic drift 1,000 times and averaged f values across replicates; this is equivalent to averaging across 1,000 independent gene loci in the same individuals. The mean observed f was then compared with expected f for the three methods, and the entire process was replicated for 100 random population size trajectories to generate a distribution of Obs(f)/E(f) values for each of the three methods.

Genetic Drift. In each replicate of the process of genetic drift, genotypes in the N_0 adults for years 1 to M were drawn by sampling randomly from $2TN_0$ distinct alleles with the proviso that the two alleles for a new individual could not come from the same parent. This modeled sampling from a single generation (generation 0) of an ideal monoecious population of size TN_0 in which selfing was not allowed and $f = 0$ (i.e., $N_{e(-1)} = \infty$). Beginning in year 1, progeny were produced by mating adults in a Wright-Fisher process without selfing (hence, $N_{b(t)} \approx N_t + 0.5$; Ca-

ballero 1994; Balloux 2004). Genotypes of progeny produced that matured in years 2 to M replaced those of individuals chosen as described above but did not increase population size. Each year, f was measured directly in all maturing adults. Table 2 shows a sample of demographic and genetic data from a single time series for one parameter set.

Multigenerational N_e *and* $E(f)$. To compute multigenerational N_e, the N_t and λ_t values were partitioned into a consecutive series of nominal generations of length T years. For each of the three methods, N_e was computed each generation, and an n-generation effective size ($\tilde{N}_{e(n)}$) was computed as the harmonic mean of the single-generation values. Although the A_i were constant across years within a parameter set, with variation in population size, the mean age of parents of adult breeders (and hence the true generation length) also varies over time. Therefore, for each population trajectory, the true number of elapsed generations, n (based on the running mean age of parents each year), was computed for each time interval up to 80 years.

The $\tilde{N}_{e(n)}$ values were used to compute conditional expectations for f as follows. With respect to a reference generation in which $f = 0$, the expected increase in f in a single generation is $E(\Delta f) = 1/(2N_e)$. The expected cumulative increase in f due to n generations of drift can be calculated from the panmictic index $(1 - f)$ as

$$E(f_n) = 1 - [1 - 1/(2\tilde{N}_{e(n)})]^n, \qquad (6)$$

where $\tilde{N}_{e(n)}$ is the harmonic mean N_e across n generations. Since selfing was not allowed, $\tilde{N}_{e(n)}$ and $E(f_n)$ at generation n were computed from single-generation effective-size estimates from generations -1 to $n - 2$. As noted above, $N_{e(-1)} = \infty$ and $N_{e(0)} = TN_0$. Although the common practice of representing the long-term N_e by the harmonic mean of single-generation values is not exact, the approximation has been shown to be robust in most practical circumstances (Motro and Thompson 1982; Hühn and Piepho 2004; Jagers and Sagitov 2004). Use of the exact expression for $E(f_n)$ as $1 - \prod_{i=1,n}[1 - 1/(2N_{e(i)})]$ did not materially change the results (data not shown).

Demographic Models

Two demographic simulation models were also used. The first model used the same population dynamics as the individual-based salmon model but did not track individual genotypes; this allowed consideration of the relationship between true N_e, $\overline{N}_t$, and $\tilde{N}_t$ under a wider range of parameter values.

Table 2: Simulated demographic and genetic data for selected generations of a semelparous species with salmon age structure

Generation and year	N_t	λ_t	R_t	X_t	N_e	Observed f 1	2	1,000	$\bar{f}$
5:									
17	68	.43	29	.036	$T\bar{N}_t = 834$	.0000	.0000	.0049	.0068
18	216	.20	42	.052	$T\tilde{N}_t = 594$	.0093	.0093	.0077	.0069
19	334	1.00	334	.415	Eq. (4) = 592	.0150	.0120	.0087	.0070
20	216	1.85	400	.497		.0000	.0000	.0041	.0072
6:									
21	69	.82	56	.066	$T\bar{N}_t = 755$	.0145	.0145	.0064	.0069
22	112	4.83	541	.639	$T\tilde{N}_t = 526$	.0268	.0134	.0060	.0089
23	277	.36	101	.119	Eq. (4) = 259	.0072	.0090	.0084	.0087
24	297	.50	149	.176		.0067	.0101	.0090	.0084
20:									
77	28	.60	17	.258	$T\bar{N}_t = 89$	.0000	.0500	.0778	.0632
78	18	.77	14	.212	$T\tilde{N}_t = 87$	.1000	.1000	.0667	.0627
79	20	.85	17	.258	Eq. (4) = 88	.0000	.0000	.0667	.0681
80	23	.77	18	.273		.0909	.0455	.0202	.0693

Note: $A_i = 0\ 0\ 0.25\ 0.5\ 0.25$; $T = 4$ years. Typical data are shown for one randomly generated time series with initial $N_0 = 100$ and $\sigma_\lambda = 1$. For each generation, N_e was calculated three ways from the demographic data, using the arithmetic mean method (eq. [3]), the harmonic mean method (eq. [2]), and equation (4). $R_t = \lambda_t N_t$ and $X_t = R_t / \Sigma R_t$. Observed f values are shown for replicates 1, 2, and 1,000; $\bar{f}$ = mean f over all 1,000 replicates.

The second demographic approach evaluated how variation in N_t in the seed-bank model affected generational V_k and hence N_e. Random series of N_t values were generated that approximated the range of scenarios considered by Nunney (2002): the N_t were lognormally distributed about a geometric mean of 100, with variance such that $J_N = 4\times$ to $64\times$. The seed bank was initiated with $N_0 CT$ seeds (N_0 and C were both set at 100 to correspond with Nunney's values, although C_t was allowed to vary across years in some parameter sets), and the system was allowed to run for 100 years to reach a dynamic equilibrium with N_t and B_t (the total number of seeds in the seed bank) both varying over time. After the burn-in period, each simulation was run for another 1,000 years, mean lifetime reproductive success was calculated for each cohort of N_t parents, and for each span of T consecutive years, generational values of V_k and $\bar{K}$ were computed. Geometric means of the single-generation $V_k/\bar{K}$ ratios were used to compute the N_e/N_T ratio using equation (5). Although seeds were effectively immortal in Nunney's model, I imposed a maximum longevity of $M = 100$ to facilitate compilation of data on lifetime reproductive success.

Results

Individual-Based Model

First, I verified that the individual-based inbreeding model correctly predicts the rate of increase in f when applied to simulated data for a species with discrete generations and variable N_t. For a variety of values of N_0 and σ_λ, Obs(f)/E(f) ratios for different time series clustered tightly around 1.0 (data not shown). As expected from theory, the variance among time series asymptotically approached 0 as the number of replicates of the gene-sampling process increased (comparable to averaging across an increasing number of independent gene loci).

Age Structure and Generation Length. Figure 3*A* shows the rate of increase in f in a constant population with age structure typical of Pacific salmon. Nominal generations of length $T = 4$ years are shown by dashed vertical lines; $N_t = 100$ each year, so effective size computed by all three methods is the same: $N_e = T\bar{N}_t = T\tilde{N}_t = N_T = 400$. As expected given that no selfing is allowed, $f = 0$ for generation 1 (years 1–4) and rises to close to the expected value (E(f) = $1 - \{1 - 1/[2(400)]\} = 0.00125$) for years 5–7 in generation 2. In year 8, f is slightly higher because some individuals are progeny of adults maturing in the first year within the generation (year 5) and so have experienced an extra generation of inbreeding. This phenomenon is repeated in the following generation but is more muted as age structure effects begin to even out. By about year 16 (end of generation 4), the pattern has largely stabilized; in subsequent years, f increases at the expected rate, but the absolute value of f is slightly elevated compared with the expectation. This latter effect is due primarily to the fact (mentioned above) that within each
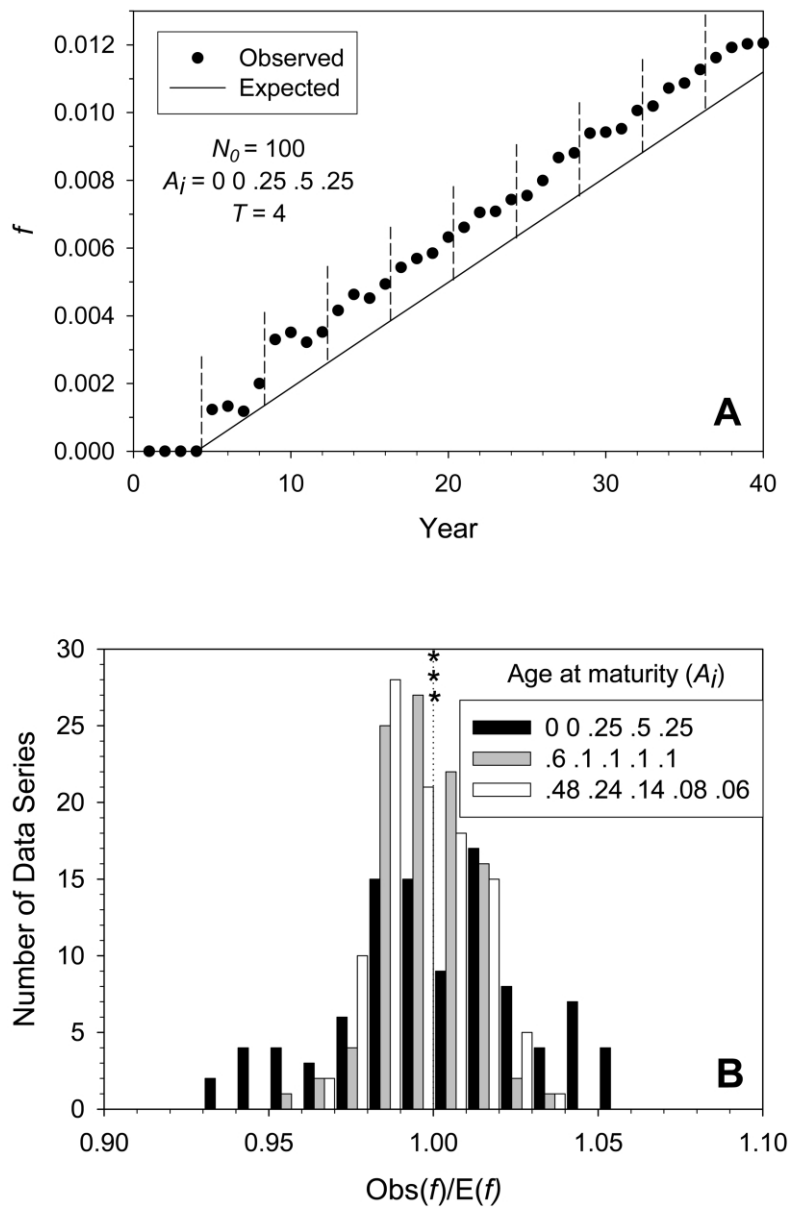
Figure 3: Comparison of observed f (Obs[f]) and expected f (E[f]) for the individual-based salmon model with constant population size $N_t = 100$. Obs(f) was averaged across 1,000 replicates; E(f) was calculated using equation (6), and N_e each generation $= TN_t = N_T$. *A*, Salmon age structure; maturity at ages 3, 4, or 5 years. Vertical dashed lines identify the nominal generations. *B*, Distribution of the ratios Obs(f)/E(f) for 100 replicate time series using three different age structures. Obs(f) and E(f) were calculated as the difference between values at generation 4 and the final generation. Asterisks indicate geometric mean ratios.

generation, at least one year class is composed in part of progeny of adults reproducing earlier in the generation. Waples (1990*a*) observed a similar phenomenon in evaluating the rate of allele frequency change in Pacific salmon populations: the rate of change in allele frequency was as predicted from $N_e = TN_b$, but for any given pair of years, the absolute magnitude of allele frequency change was higher than the expectation because one year's spawners represent only part of a generation. In the remaining analyses (results shown in figs. 3*B*, 4 and table 3), I computed Obs(f)/E(f) using differences between values for generation 4 (after initial age structure effects had stabilized) and the end of the replicate (80 years $= 80/T$ generations). With N constant, E(f) is the same for all three models,
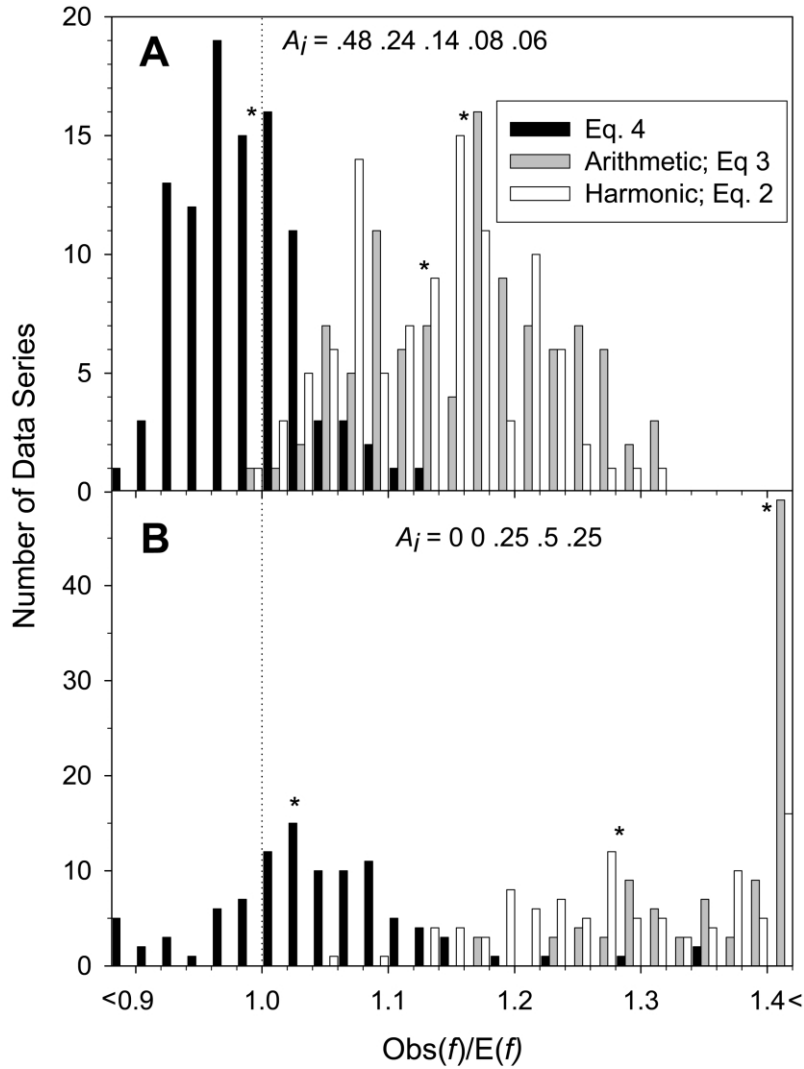
Figure 4: Comparison of observed f (Obs[f]) and expected f (E[f]) for the individual-based salmon model but with variable population size ($N_0 = 100$; $\sigma_\lambda = 1.0$). For each time series, E(f) was computed from equation (6) based on three different ways of calculating N_e; Obs (f) is the arithmetic mean f across 1,000 replicates of the gene sampling process. *A*, Age structure similar to that of annual plants with seed banks. *B*, Age structure typical of Pacific salmon. Asterisks show geometric mean ratios.

and in this case, multigeneration Obs(f)/E(f) ratios clustered tightly around 1.0 for age structures typical of both salmon and plants with seed banks (fig. 3*B*).

When N_t varies, E(f) differs for the three methods of computing N_e. Figure 4 shows typical results for species with age structure similar to that of plants with seed banks (fig. 4*A*) and Pacific salmon (fig. 4*B*). In both cases, the distribution of Obs(f)/E(f) calculated from equation (4) is centered around 1.0 with a geometric mean close to 1. In contrast, Obs(f) exceeds E(f) for both the harmonic mean and the arithmetic mean methods, indicating that those approximations overestimate true N_e. Results for a

more extensive range of simulations (table 3) were similar: equation (4) consistently estimates Obs(f) accurately, whereas the other two methods consistently underestimate the rate of increase in f and thus overestimate N_e. This indicates that true N_e was even less than predicted by the harmonic mean method, although the difference was somewhat less with an age structure typical of plants with seed banks (cf. fig. 4*A*, 4*B*).

Magnitude of Variability in N. Nunney (2002) noted that the magnitude of variability in N_t considered in the salmon model was small compared with the variability in the seed-

Table 3: Ratio of observed f to expected f (Obs[f]/E[f]) for various age structures (A_i), generation lengths (T), and variabilities in population growth rate (σ_λ)

A_i and T[a]	σ_λ[b]	Obs(f)/E(f)		
			Additive	Harmonic
		Eq. (4)	($N_e = T\overline{N}_t$)	($N_e = T\tilde{N}_t$)
.6 .1 .1 .1 .1:				
2	1.0	.99	1.17	1.13
.4 .3 .2 .1:				
2	1.0	.98	1.17	1.14
.48 .24 .14 .08 .06:				
2	.5	.99	1.07	1.06
2	1.0	.99	1.16	1.13
2	1.5	.98	1.19	1.16
2	2.0	.99	1.20	1.17
.33 .33 .33:				
2	1.0	.98	1.14	1.11
3	1.0	.99	1.28	1.23
4	1.0	.99	1.34	1.25
.25 .5 .25:				
2	1.0	1.01	1.16	1.14
3	1.0	1.03	1.39	1.27
4	.5	1.01	1.16	1.12
4	1.0	1.03	1.41	1.29
4	1.5	1.05	1.54	1.36
4	2.0	1.05	1.65	1.42

Note: E(f) was calculated from equation (6) based on three different ways of calculating N_e. Values shown are geometric means of Obs(f)/E(f) for 100 randomly generated time series of N_t and λ_t; in each parameter set, $N_0 = 100$.

[a] Generation length was controlled by varying age at first maturity.

[b] Constant population size ($\sigma_\lambda = 0$) was also considered for all age structures, in which case all three methods yield the same N_e. With constant N, geometric mean Obs(f)/E(f) was in the range 0.993–0.997 for all age structures.

bank model, and he predicted that "given larger annual fluctuations in the number of breeding fish, the error introduced by using the harmonic mean alone would become apparent" (p. 202). This hypothesis, however, is not supported by results shown in table 3. The range of variability in σ_λ (0.5–2.0) considered in those simulations corresponds to variability in N_t that spans the range ($J_N = 4\times$ to $64\times$) considered by Nunney (2002). For both seed-bank and salmon age structures, increasing variability in N_t did not lead to better agreement with the arithmetic mean method; in fact, the opposite was true— as variability increased, upward bias of both the arithmetic mean and harmonic mean methods increased. Even with high levels of variability in N_t and λ_t, however, equation (4) still provided a very good estimate of true multigeneration N_e.

Salmon Demographic Model

Nunney (2002) noted that the generation time of Pacific salmon is relatively short and predicted that N_e would converge to that predicted by the arithmetic mean method if T were larger. To evaluate this hypothesis, I conducted additional demographic simulations using salmon population dynamics and longer generation length (controlled by varying the minimum age of maturity). Results (fig. 5) show that regardless of generation length ($T = 2$ to 9 years) and whether maturity occurred at three or five different years, N_e computed from demographic data using equation (4) was always less than or equal to that predicted by the harmonic mean method. Results in figure 5 used a modal age structure; qualitatively similar results were obtained when maturity was evenly spread across ages.

I also conducted simulations with $A_i = 0.25, 0.19, 0.14, 0.105, 0.08, 0.06, 0.045, 0.034, 0.026, 0.07$ for ages $i = 1$ to 10; this approximated age structure in a seed-bank model having a fixed germination probability of 0.25/year and $T \approx 4$, but with A_{10} representing the sum of all probabilities for $i \geq 10$ (see Templeton and Levin 1979 for a similar example). Results (data not shown) indicate that spreading maturity across many years does tend to dampen annual fluctuations in N_t. Nevertheless, true N_e (eq. [4]) was still less than estimated using the harmonic mean method.

Seed-Bank Demographic Model

Variation in Population Size and Seed Production. Several features characteristic of the seed-bank model involving constant C (as assumed by Nunney 2002) can be noted (fig. 6, *top*). First, with a long generation length ($T = 10$) and modest ($J_N = 4\times$) variation in population size, V_k was close to the Poisson variance ($V_k = 1.03\overline{K}$), so true N_e was in close agreement with predictions of the arithmetic mean method ($N_e/N_T = 0.99$; eq. [5]). With more dramatic variation ($J_N = 50\times$), $V_k/\overline{K}$ was somewhat higher (1.34), and hence effective size was lower ($N_e/N_T = 0.85$). Second, for both levels of variability in N_t, $V_k/\overline{K}$ increased and N_e/N_T decreased as generation length was reduced from $T = 10$ to $T = 2$. Third, as expected, with increasing variability in population size, the difference between the harmonic mean and the arithmetic mean N_t increased. Fourth, with constant C, relatively good agreement was found between N_e calculated using equation (5) and that predicted by Nunney's formula (eq. [1]).

Results are quite different when per capita seed production is allowed to vary (fig. 6, *bottom*). With J_N set at $25\times$, N_e calculated from the demographic data depends heavily on variation in C_t, being much lower for $J_C = 50\times$ than for $J_C = 4\times$. Furthermore, when C_t varies, N_e/N_T decreases as generation length decreases—the opposite of the pattern found when C_t is constant. Finally, with yearly variation in C_t, under most circumstances, equation
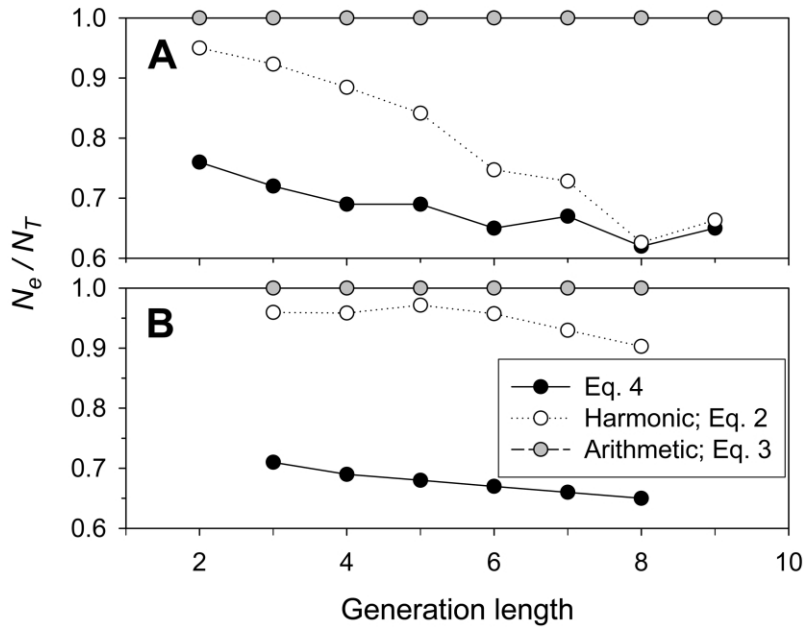
Figure 5: N_e/N_T as a function of age structure and generation length for simulated demographic data in the salmon model. In each simulation, geometric mean $N_t \approx 200$ and $\sigma_\lambda = 1.0$. The time series of N_t and $\overline{k}_t$ values was used to calculate N_e three ways, with results for equation (4) (*black circles*) and the harmonic mean method ($N_e = T\tilde{N}_t$; *open circles*) expressed as a fraction of the values for the arithmetic mean method ($N_e = T\overline{N}_t = N_T$; *gray circles*). Plotted values are geometric mean ratios of multigeneration N_e values. *A*, Maturity at 3 consecutive years, with probabilities 0.25, 0.5, and 0.25. *B*, Maturity at 5 consecutive years, with probabilities 0.12, 0.22, 0.32, 0.22, and 0.12. Variation in generation length was achieved by varying age at first maturity.

(1) is a poor predictor of true N_e, and performance declines as J_C and T increase. With $J_C \geq 25\times$, $T\tilde{N}_t$ predicted true N_e better than did equation (1).

Density Dependence. In the salmon model considered by Waples (2002) and here, λ_t was chosen independently of N_t, so density dependence was absent except as a result of boundary constraints on population size. In the seed-bank model, only N_t plants are allowed to mature each year from among the germinating seeds, with N_t chosen randomly. This is equivalent to random fluctuation in carrying capacity, with density-dependent survival at the stage of germinating seeds (fig. 1A). Density dependence thus occurs after seeds of various ages (and derived from parents in many previous years) have germinated, so it does little to create variation in $\overline{k}_t$ across years and hence has little effect on N_e. However, the constant C model leads to an indirect form of density-dependent compensation. Although seed production is constant for all plants in all years, the fraction of their seeds that eventually mature depends on N_t in subsequent years (see app. B in the online edition of the *American Naturalist*). On average, seeds that reproduce in years with low N_t will be followed by years with relatively higher N_t (and vice versa). This will (on

average) enhance the realized reproductive success for years of low abundance and reduce reproductive output for years of high abundance, leading to a negative correlation of N_t and $\overline{k}_t$. The effects of variation in N_t will be most pronounced when generation length is short, in which case lifetime reproductive success is strongly determined by N_t in the one or two years immediately following seed production. This explains the result seen in figure 6, *top*: with constant C, variation in N_t depresses N_e more for $T = 2$ than for longer generation lengths.

Joint Consideration of Variation and Covariation of Population Size and Growth Rate

The large differences in model assumptions regarding population regulation indicate that it would be useful to analytically consider how N_e is affected by variation and covariation of population size and growth rate. Although a completely general formula for N_e in terms of $\text{Cov}(N_t, \overline{k}_t)$ and the means and variances of N_t and $\overline{k}_t$ does not seem possible, analytical results can be derived for some specific cases (see app. C in the online edition of the *American Naturalist* for details). If mean reproductive success ($\overline{k}_t$) is constant across years within a generation,
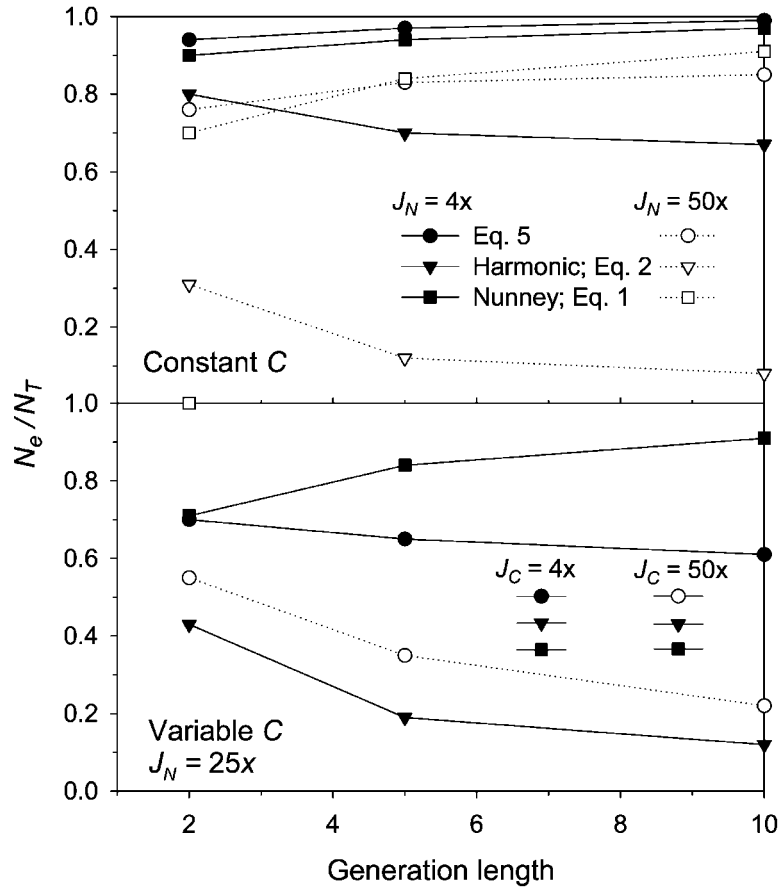
Figure 6: N_e/N_T as a function of generation length and variability in N_t and C_t in the seed-bank model. For each parameter set, N_e/N_T was calculated using equation (5) based on realized mean and variance in reproductive success over a generation ($\overline{K}, V_k$), and the result was compared with estimates based on the harmonic mean method ($N_e \approx T\tilde{N}_t$; eq. [2]) and Nunney's model ($N_e \approx \tilde{N}_t + [T-1]\overline{N}_t$; eq. [1]). *Top*, seed production is constant at $C = 100$; $J_N = 4\times$ and $50\times$. *Bottom*, $J_N = 25\times$ and C_t varies across years, with $J_C = 4\times$ or $50\times$.

$N_e = \sum N_t = N_T$ regardless of the pattern of variation in N_t. If population size is constant within a generation, then the effective size : census size ratio is a simple function of $CV(\overline{k}_t)$, the coefficient of variation of $\overline{k}_t$ (eq. [C2] in the online edition of the *American Naturalist*):

$$\frac{N_e}{N_T} = \frac{1}{1 + CV^2(\overline{k}_t)}. \tag{7}$$

The expression $CV^2(\overline{k}_t)$ is similar in form to I, the standardized variance in reproductive success ($I = V_k/\overline{K}^2$; Clutton-Brock 1988; Nunney and Elam 1994; Nunney 2002), but with an important difference: whereas I measures variation among individuals within a generation, $CV^2(\overline{k}_t)$ measures variation in mean reproductive success across years.

In more realistic scenarios, both N_t and $\overline{k}_t$ vary across years. In this case, it is apparent that N_e and N_e/N should

also be affected by $Cov(N_t, \overline{k}_t)$ because (for example) inbreeding effective size should be reduced when $\overline{k}_t$ and N are negatively correlated (i.e., relatively few breeders produce disproportionate numbers of offspring, and relatively large numbers of breeders produce few offspring). Trial-and-error numerical evaluation indicates that the following formula, which includes a term for the standardized covariance of N_t and $\overline{k}_t$ ($\overline{k}_*$ is the mean of the $\overline{k}_t$ among years within a generation), provides a reasonable approximation of the true N_e/N:

$$\frac{N_e}{N_T} \approx \frac{1}{1 + CV^2(\overline{k}_t)} + \frac{Cov(\overline{k}_t, N_t)}{2.5\overline{k}_*\overline{N}_t}. \tag{8}$$

The usefulness of this approximation can be seen in figure 7, which displays results for 200 replicate genera-
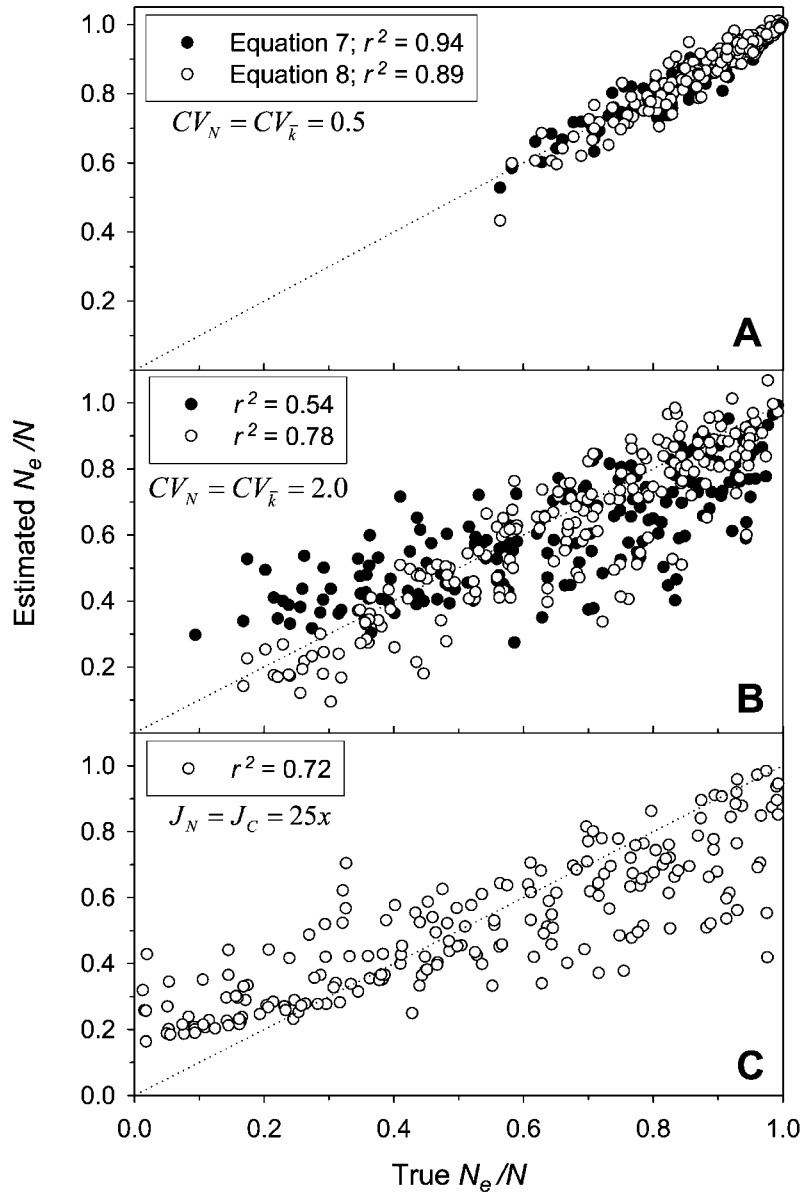
Figure 7: Relationship between the true N_e/N_T ratio (eq. [4]) and that predicted by equations (7) (*filled circles*) and (8) (*open circles*). Each data point represents N_e/N_T for a single generation of $T = 4$ years. In *A* and *B*, N_t and $\overline{k}_t$ values were chosen randomly and independently from lognormal distributions with $\overline{N}_t = 200$, $\overline{k}_* =$ mean $\overline{k}_t = 2$, and $CV(N_t) = CV(\overline{k}_t) = 0.5$ (*A*) or 2.0 (*B*). In *C*, data for 200 consecutive generations in the seed-bank model with $J_N = J_C = 25 \times$ and initial $N_0 = 100$.

tions of 4 years of simulated data. In figure 7*A* and 7*B*, N_t and $\overline{k}_t$ were randomly and independently chosen from lognormal distributions and N_e/N was computed using equation (4). Under these conditions, although $E(\text{Cov}\,(N_t, \overline{k}_t)) = 0$, random variation within generations leads to a diversity of actual values for $\text{Cov}\,(N_t, \overline{k}_t)$ and means and variances of $\overline{k}_t$ and N_t. With a moderate level of parametric variation ($CV(N_t) = CV(\overline{k}_t) = 0.5$; fig. 7*A*),

equation (7) explains 94% of the variance in N_e/N, and equation (8) provides no improvement. When variability is high ($CV(N_t) = CV(\overline{k}_t) = 2.0$; fig. 7*B*), equation (7) is not quite as effective and has a slight upward bias for low N_e/N; in this case, including the covariance term in equation (8) improves the fit considerably. Figure 7*C* shows data for 200 consecutive generations in the seed-bank model with $J_N = J_C = 25 \times$. Equation (8) does not per-

form quite as well as in the random model, but it still explains over two-thirds of the variance in N_e/N. I also evaluated scenarios using conditions as in figure 7B and 7C but with generation length ranging from $T = 2$ to 10. Results (data not shown) indicate the following: in both models, performance of equation (8) is better for short generation length and declines somewhat as T increases; in the random model, the difference between predictions of equations (7) and (8) is largest when T is small, and adding the covariance term does not improve performance with $T = 10$; in the seed-bank model, equation (8) was significantly better at predicting N_e for all values of T considered. Although equation (8) obviously does not capture all of the factors that affect N_e in these species, under even the least favorable conditions evaluated it still explains over 50% of the variance in N_e/N_T among replicates. This expression, therefore, can be useful as a heuristic tool for evaluating at least the relative importance of factors that affect N_e.

Discussion

An important result of the above analyses (table 3; fig. 4) is that discrete generation formulas for inbreeding N_e (eqq. [4], [5]) provide a robust means of calculating effective size over single and multiple generations in age-structured, semelparous species. Previously, Waples (2002) used this approach with an empirical data set and found qualitative agreement with results based on monitoring changes in allele frequency in simulated salmon populations. However, a direct comparison of N_e calculated from demographic data and actual levels of inbreeding calculated from genetic data for the same individuals has not previously been made. Results presented here used integer values of T; if T is not an integer, it can be rounded to the nearest integer (T^*) to produce a nominal generation length. In that case, N_e computed from equation (4) or (5) should be scaled by the factor T/T^* to produce an unbiased estimate of N_e per generation.

Collectively, results show that the contrasting conclusions reached by Nunney (2002) and Waples (2002) are due primarily to different model assumptions. The assumption that C is constant across years in the seed-bank model severely constrains variance in realized reproductive success, with the result that generational $V_k/\overline{K}$ is not substantially >1 even with large variation in N_t. As a consequence, N_e is close to N_T, in rough agreement with the prediction of Nunney's model (eq. [1]). As variation in N_t increases, the ratio $\tilde{N}_t/\overline{N}_t$ decreases, and the harmonic mean method increasingly underestimates true N_e.

When C is allowed to vary in the seed-bank model, however, under most circumstances equation (1) no longer provides an adequate approximation to N_e. With moderate variation in C and relatively long generation times, true N_e can be much closer to $T\tilde{N}_t$ than to the prediction from equation (1). Still, under all seed-bank scenarios evaluated, N_e computed using equation (5) was intermediate to values predicted by the arithmetic mean and harmonic mean methods.

Results for the salmon model are quite different: true N_e was never intermediate to $T\tilde{N}_t$ and $T\overline{N}_t$, as predicted by equation (1); rather, in all scenarios considered it was even less than $T\tilde{N}_t$. It is useful to explore possible explanations for this fundamental difference. Nunney (2002) predicted that agreement of N_e in the salmon model with equation (1) would improve for larger T and J_N. As shown in table 3, however, that is not the case. Instead, upward bias from using the arithmetic mean method increased with increasing variability in N_t. Another possible explanation is that the seed-bank model assumes a monoecious diploid with the possibility of selfing, whereas Pacific salmon have separate sexes. However, Nunney's model is general enough to allow consideration of random mating without selfing, in which case the two models are equivalent with respect to the expected rate of genetic change (cf. eqq. [2] and [2″] in Crow and Denniston 1988). Therefore, neither different assumptions about T and J_N nor the different mating systems of plants and salmon can explain the different results for the two models.

A more fundamental difference in the two models is the method of population regulation. Two factors are noteworthy in this regard. First, in the seed-bank model, variance in N within a generation is nearly as large as the variance for the entire time series (fig. 2, *top*), so even moderate variation in N leads to strong divergence between $\overline{N}_t$ and $\tilde{N}_t$ within a generation. In contrast, N changes more gradually in the salmon model because of strong dependence on population size in two or more parental years. Thus, although variance in N over the 80-year time series shown in figure 2 is approximately equal for the two models, the variance within a generation is much less in the salmon model and, as a consequence, $\overline{N}_t$ and $\tilde{N}_t$ are not as divergent. Because of these different relationships between $\overline{N}_t$ and $\tilde{N}_t$ within a generation, it is perhaps not surprising that $N_e < T\tilde{N}_t$ in the salmon model but $T\tilde{N}_t < N_e < T\overline{N}_t$ in the seed-bank model.

Second, the different methods of population regulation result in different relationships between N_t and $\overline{k}_t$. If $\overline{k}_t$ is constant and $\sigma^2_{k(t)}$ is binomial within each year, then the entire generation acts as an ideal population with $\overline{K} = \overline{k}_t \approx V_k$ and $N_e \approx N_T = T\overline{N}_t$. In the seed-bank model, C is constant and N_t is a random variable and only indirectly leads to modest variation in $\overline{k}_t$, so agreement with the arithmetic mean model is good. In contrast, in the salmon model, variation in N_t arises directly from random variation in $\overline{k}_t$. By definition, variation in $\overline{k}_t$ among years

within a generation increases the generational variance in reproductive success (V_k) and hence reduces N_e. The magnitude of this reduction is determined primarily by $CV(\overline{k}_t)$, but $Cov(N_t, \overline{k}_t)$ is also a factor (eq. [8]; fig. 7). Therefore, it is important to consider the factors primarily responsible for variation and covariation in N_t and $\overline{k}_t$ across years: variation in fecundity and cohort-specific survival and mechanisms of density dependence. Relevance of these factors for salmon and monocarpic plants is discussed below.

Pacific Salmon

Mean fecundity varies among years in Pacific salmon (Quinn and Bloomberg 1992), but the magnitude of variation is usually dwarfed by large (often density independent) annual fluctuations in juvenile survival that are common in both the freshwater and marine environments. For example, λ_t (and hence $\overline{k}_t$) can vary across several orders of magnitude over a short period of time (Cass and Riddell 1999; Waples 2002). Simulation results for salmon reported previously (Waples 2002) and in this article assumed random variation in $\overline{k}_t$ that was uncorrelated with variation in population size. However, it is a fundamental tenet of fisheries management that per capita production increases at low population size and is depressed or asymptotes at high density (Hilborn and Walters 1992), both of which lead to $Cov(N_t, \overline{k}_t) < 0$. Equation (8) and figure 7 indicate that, to the extent that density-dependent compensation occurs in salmon, reductions in N_e should be even larger than indicated by a model assuming random, uncorrelated variation in N_t and λ_t.

Monocarpic Plants

Perennials. A recent review of the evolutionary demography of monocarpic perennials (Metcalf et al. 2003) indicates that the following traits appear to be common in these species: plants of the same size can vary substantially in growth rate, survival and growth rates vary substantially across years, seed production (and perhaps offspring quality) increases with parental size, reproductive success among plants of the same size maturing in the same year can vary due to timing of flowering, and the transition from vegetative to reproductive state is affected by multiple genetic and environmental cues. Collectively, these factors indicate that in monocarpic perennials, variation in $\overline{k}_t$ across years is likely to be the rule rather than the exception and that true N_e will be lower than predicted using the arithmetic mean model. In addition, most population models for these species assume that density dependence occurs at the seedling stage, due to a limited number of "safe sites" for germination (Metcalf et al. 2003). As

pointed out by Vitalis et al. (2004), this leads to an important difference compared with the seed-bank model, because in this case the limitation applies directly to an entire cohort. If the number of safe sites varies over time, then $\overline{k}_t$ will as well, with variation imposed on mean reproductive success of plants flowering the previous year. The number of safe sites might vary over time due to variation in the number of empty sites or variation in environmental conditions that allow more or less of the empty sites to be suitable for germination in any given year. A major contributor to the first factor is habitat disturbance, which almost certainly varies over space and time. It is also not hard to imagine that particularly favorable (or unfavorable) environmental conditions might increase (or decrease) the number of empty sites suitable for germination. Therefore, it seems likely that in many perennial monocarps, variation in N_t caused by random variation in carrying capacity might act in such a way to increase variation in $\overline{k}_t$. As a consequence, N_e in these species might be better described by results of the salmon model than the seed-bank model.

Is it likely that density dependence will take the form of compensation, leading to a negative $Cov(N_t, \overline{k}_t)$ and further reducing N_e? Compensation would occur if, in years in which relatively few plants flower, they either have unusually high seed production or produce seeds with higher probabilities of maturing and flowering. The extent to which this occurs probably varies considerably across species and, perhaps, across space and time within species.

Seed Banks. The assumption (Nunney 2002) that density dependence acts at the germination stage (at which point it applies to seeds randomly drawn from the seed bank rather than to only a specific cohort) is widely used in seed-bank models and seems reasonable for these species. Therefore, it does not seem likely that density dependence has a strong effect on N_e in plants with seed banks. However, the assumption that fecundity is constant over time may not be realistic for many annual species, given the evidence for large annual variation in seed production in other monocarps. Similarly, given that seed banks are most commonly found in populations that live in unpredictable and/or extreme environments, it seems likely that the probability of germination varies depending on environmental conditions within a particular year (Levine and Rees 2004). Finally, allowing for a maximum seed longevity and an annual cost (chance of mortality) for seeds in the seed bank (Vitalis et al. 2004) also would seem to be more realistic. Mortality rates in the seed bank could also vary from year to year depending on environmental conditions. To the extent that they occur, all of these factors would tend to increase variation in $\overline{k}_t$ and reduce N_e compared with predictions of the arithmetic mean model.

Effective : Census Size Ratios per Year

In order to focus on effects of fluctuating population size, Nunney (2002) and Waples (2002) both assumed that $\beta_t = N_{b(t)}/N_t = 1$ each year. However, $N_{b(t)}/N_t < 1$ in most natural populations, so results presented above should be regarded as maximum estimates of N_e per generation. Yearly reductions in $N_{b(t)}$ compared with N_t can be accounted for in a general way as follows: seed banks, constant C: $N_e \approx \tilde{N}_{b(t)} + (T-1)\overline{N}_{b(t)}$; seed banks, variable C: $T\tilde{N}_{b(t)} < N_e < T\overline{N}_{b(t)}$; salmon, monocarpic perennials: $N_e < T\tilde{N}_{b(t)} \ll T\overline{N}_{b(t)}$.

Other Species

A variety of other semelparous species have variable age at maturity, most notably crustaceans with diapausing eggs. Results presented here should provide a framework for evaluating how fluctuations in population size and population demography interact to influence effective size in these species.

A number of authors (e.g., Felsenstein 1971; Hill 1972, 1979; Orive 1993; Pollak 2000) have considered effective size of iteroparous, age-structured species, but all assumed constant population size and stable age distributions, so results are not directly applicable to the topics considered here (but see Engen at al. 2005). However, it can be shown that if N_t is constant, the discrete-generation formulas used here to estimate N_e from demographic data are equivalent to the expression derived by Hill (1972, 1979) for N_e in species with overlapping generations and a stable age distribution.

The method used here takes advantage of the fact that semelparous species reproduce only once, in which case grouping T years of reproducing adults into nominal generations provides a good approximation to multigenerational increases in inbreeding. It might be possible to extend this approach to iteroparous species, but the latter are more complicated because of the uncertain relationship between reproductive success in year t and lifetime reproductive output.

Acknowledgments

Literature Cited

Balloux, F. 2004. Heterozygote excess in small populations and the heterozygote-excess effective population size. Evolution 58:1891–1900.

Caballero, A. 1994. Developments in the prediction of effective population size. Heredity 73:657–679.

Cass, A., and B. Riddell. 1999. A life history model for assessing alternative management policies for depressed chinook salmon. ICES (International Council for the Exploration of the Sea) Journal of Marine Science 56:414–421.

Clutton-Brock, T. H., ed. 1988. Reproductive success: studies of individual variation in contrasting breeding systems. University of Chicago Press, Chicago.

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

Crow, J. F., and M. Kimura. 1970. An introduction to population genetics theory. Harper & Row, New York.

Engen, S., R. Lande, and B.-E. Saether. 2005. Effective size of a fluctuating age-structured population. Genetics 170:941–954.

Felsenstein, J. 1971. Inbreeding and variance effective numbers in populations with overlapping generations. Genetics 68:581–597.

Hairston, N. G., Jr., and B. T. De Stasio Jr. 1988. Rate of evolution slowed by a dormant propagule pool. Nature 336:239–242.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman & Hall, New York.

Hill, W. G. 1972. Effective size of populations with overlapping generations. Theoretical Population Biology 3:278–289.

———. 1979. A note on effective population size with overlapping generations. Genetics 92:317–322.

Hühn, M., and H.-S. Piepho. 2004. Inbreeding coefficients for stochastically varying small population sizes: bias of calculation based on effective numbers. Journal of Theoretical Biology 226:467–475.

Jagers, P., and S. Sagitov. 2004. Convergence to the coalescent in populations of substantially varying size. Journal of Applied Probability 41:368–378.

Kaj, I., S. M. Krone, and M. Lascoux. 2001. Coalescent theory for seed bank models. Journal of Applied Probability 38:285–300.

Levine, J. M., and M. Rees. 2004. Effects of temporal variability on rare plant persistence in annual systems. American Naturalist 164:350–363.

Metcalf, J. C., K. E. Rose, and M. Rees. 2003. Evolutionary demography of monocarpic perennials. Trends in Ecology & Evolution 18:471–480.

Motro, U., and G. Thompson. 1982. On heterozygosity and the effective size of populations subject to size changes. Evolution 36:1059–1066.

Nunney, L. 2002. The effective size of annual plant populations: the interaction of a seed bank with fluctuating population size in maintaining genetic variation. American Naturalist 160:195–204.

Nunney, L., and D. R. Elam. 1994. Estimating the effective population size of conserved populations. Conservation Biology 8:175–184.

Orive, M. E. 1993. Effective population size in organisms with complex life-histories. Theoretical Population Biology 44:316–340.

Pollak, E. 2000. The effective population size of some age-structured populations. Mathematical Biosciences 168:39–56.

Quinn, T. P., and S. Bloomberg. 1992. Fecundity of chinook salmon (*Oncorhynchus tshawytscha*) from the Waitaki and Rakaia River,

New Zealand. New Zealand Journal of Marine and Freshwater Research 26:429–434.

Ryman, N., and L. Laikre. 1991. Effects of supportive breeding on the genetically effective population size. Conservation Biology 5: 325–329.

Ryman, N., P. E. Jorde, and L. Laikre. 1995. Supportive breeding and variance effective population size. Conservation Biology 9:1619–1628.

Templeton, A. R., and D. A. Levin. 1979. Evolutionary consequences of seed pools. American Naturalist 114:232–249.

Vitalis, R., S. Glémin, and I. Olivieri. 2004. When genes go to sleep: the population genetic consequences of seed dormancy and monocarpic perenniality. American Naturalist 163:295–311.

Waples, R. S. 1990a. Conservation genetics of Pacific salmon. II. Effective population size and the rate of loss of genetic variability. Journal of Heredity 81:267–276.

———. 1990b. Conservation genetics of Pacific salmon. III. Estimating effective population size. Journal of Heredity 81:277–289.

———. 2002. The effective size of fluctuating salmon populations. Genetics 161:783–791.