

Tiny estimates of the N_e/N ratio in marine fishes: Are they real?*

R. S. WAPLES

NOAA Fisheries, Northwest Fisheries Science Center, 2725 Montlake Blvd East, Seattle, WA,
98112, U.S.A.

Theory and empirical estimates agree that the ratio of effective size (N_e) to census size (N) falls roughly in the range 0.1–0.5 for most populations. In a number of marine species, however, genetic estimates of contemporary N_e/N are as much as 5–6 orders of magnitude lower. Although some mechanisms that could produce such tiny N_e/N ratios have been proposed, the subject remains controversial. This issue is important to resolve: if N_e/N can be 10^{-3} or smaller, marine fish populations that are quite large could be at genetic risk. Based on a recently-improved understanding of factors that influence N_e and N_e/N in species with overlapping generations, this paper evaluates conditions necessary to produce tiny N_e/N ratios in actual populations. These analyses show that although increased longevity, fecundity and variance in reproductive success that increase with age, and increased egg quality with age [the big old fat fecund female fish (BOFFFF) hypothesis] all reduce N_e/N , extreme scenarios are required to reduce N_e/N below about 0.01. Therefore, tiny N_e/N ratios require some version of Hedgecock's 'sweepstakes' hypothesis, whereby only a few families reproduce successfully. Simulations using common genetically-based estimators show that, when true N_e is very large ($\geq 10^6$), a substantial fraction of point estimates of N_e/N can be 10^{-3} or smaller. These results mean that tiny, genetically-based point estimates of N_e/N in large marine populations are expected to be quite common, even when the true N_e/N ratio is 'normal' (~ 0.1 or higher). Very large samples of individuals can reduce, but not eliminate, this problem. The simulation results also emphasize the importance of considering deviations from model assumptions (e.g. non-random sampling; weak selection or migration) that may be relatively small (and hence can generally be ignored when the signal is strong) but can lead to substantial biases when the drift signal is weak, as is likely for large marine populations. Empirical studies of this topic need to be able to distinguish between episodes of sweepstakes reproductive success that are ephemeral and lead to chaotic genetic patchiness, and those that are consistent enough across space and time to produce persistent evolutionary consequences.

Published 2016. This article is a U.S. Government work and is in the public domain in the USA.

Key words: age structure; effective population size; fecundity; overlapping generations; sweepstakes reproductive success.

INTRODUCTION

Marine species have long captured our imaginations, and this has been true of scientific investigations as well as the popular media. Even before the first large allozyme studies revealed dramatically higher-than-expected levels of genetic diversity in

Tel.: +1 206 860 3254; email robin.waples@noaa.gov

*Twenty-fifth J. W. Jones Lecture.

^aThis paper was presented at the FSBI Symposium, Bangor, in July 2016. Its content may not follow the usual style and format of the *Journal of Fish Biology*.

humans (Harris, 1966) and *Drosophila* (Hubby & Lewontin, 1966), genetic methods were being used to gain insights into ecological and evolutionary processes in marine fish populations (Frydenberg *et al.*, 1965; Waples *et al.*, 2008). Genetic studies of marine species have explored population structure (Cross & Payne, 1978; Burton & Feldman, 1981), larval dispersal and gene flow (Strathmann, 1978; Johnson & Black, 1984), natural selection (Tracey *et al.*, 1975; Koehn *et al.*, 1976; DiMichele & Powers, 1982; Gaffney, 1994), sampling and kinship (Hansen *et al.*, 1997; Buston *et al.*, 2009), fishery management (Ryman & Utter, 1987), unusual life histories (Koehn & Williams, 1978; Aarestrup *et al.*, 2009), mating systems (Bierne *et al.*, 1998; Bekkevold *et al.*, 2002), fecundity and maternal age (Berkeley *et al.*, 2004; Hixon *et al.*, 2014), phylogeography (Grant & Bowen, 1998), speciation (Palumbi, 1994) and even needle-in-a-haystack parentage assignments (Christie *et al.*, 2010). These studies have led to many surprises and, collectively, greatly enriched the understanding of how natural populations function in the real world.

One important topic that remains controversial is whether marine species with high fecundity can have effective population sizes (N_e) that are many orders of magnitude smaller than the census size (N). Conventional evolutionary theory holds that the N_e/N ratio should not deviate too far from 0.5 and rather special circumstances are required to produce N_e/N as small as 0.1 (Nunney, 1993). Hedgecock (1994), however, proposed that, through a variation of Hjordt's (1914) larval mismatch hypothesis, N_e/N in highly fecund marine species could be very small if a typical year class of surviving offspring is not derived randomly from the huge number of adults, but instead from only a few families that by chance happen to produce eggs and larvae that end up at the right place and time to allow them to survive. This idea has been referred to as the sweepstakes reproductive success (SRS) hypothesis. Subsequently, a number of empirical studies using indirect genetic methods obtained estimates of N_e/N in marine species ranging from 10^{-3} to 10^{-6} or even smaller. This topic was reviewed by Hauser & Carvalho (2008) and Hedgecock & Pudovkin (2011), and Hedrick (2005) used some simple theoretical models to identify scenarios that could potentially produce very small N_e/N .

This study extends these previous analyses in two ways. First, the analytical models considered by Hedrick (2005) are extended to account for age structure and overlapping generations, with the goal of identifying life history traits that can and cannot be expected to produce low N_e/N . Second, the conditions under which commonly used genetic methods can be expected to produce tiny estimates of the N_e/N ratio, even when N_e is large and N_e/N is close to 1, are evaluated. Finally, experimental procedures that can evaluate hypotheses regarding estimates of N_e/N in marine fishes are discussed.

DEMOGRAPHIC FACTORS THAT CAN PRODUCE LOW N_e/N

DISCRETE GENERATION MODELS

Consider a marine species with a large population size ($N \geq 10^6$). The focus is on scenarios where N_e/N is very small. For purposes of this study, a tiny estimate of N_e/N is defined as one that is $\leq 10^{-3}$. The choice of which individuals to include in N can strongly affect the estimated N_e/N . The analyses below use the definition that is most widely accepted in the literature: N = the number of mature adults (Nunney & Elam, 1994). In species with fixed age at maturity, this can be calculated as the number in all age classes from age at maturity to the maximum age, ω .

Hedgcock (1994) proposed that N_e/N could be arbitrarily small if only a small fraction of all N adults successfully reproduced (in this context, successful means production of at least one offspring that survive to be an adult). Let N_p be the number of these successful parents, so the focus is on scenarios in which $N_p \ll N$. Hedrick (2005) quantitatively evaluated some simple scenarios of this type and here this idea is expanded using the parentage analysis without parents (PWOP) approach of Waples & Waples (2011). In the PWOP formulation, the standard discrete-generation formula for inbreeding effective size is recast as:

$$N_e = \frac{\sum k_i - 1}{\frac{\sum(k_i^2)}{\sum k_i} - 1}, \quad (1)$$

where k_i is the number of offspring produced by the i^{th} parent. Because parents who produce no adult offspring ($k_i = 0$) do not contribute to either $\sum k_i$ or $\sum(k_i^2)$, this formulation shows that N_e does not depend on the total adult population size or the number of unsuccessful parents. Nevertheless, to facilitate evaluation of N_e/N , the connection to N will be retained. The analyses below consider how overall N_e and N_e/N depend on the mean and variance in offspring number of these N_p successful parents.

If the population is stable, then the N total adults in generation 1 produce N adults in generation 2 and so on. For diploid species, each of the adults must on average contribute half the genes to each of two offspring, so overall $\bar{k} = 2$. Considering only the successful parents, they still contribute $\sum k_i = 2N$ genes to the N offspring, so the mean reproductive contribution of the successful parents is $\bar{k}_p = 2N/N_p$. Evaluation of $V_{kp} = \text{variance}(k_i)$ for the successful parents is simplified by taking advantage of the property that a variance is the mean of the squares minus the square of the mean. In current notation, $V_{kp} = \sum(k_i^2)/N_p - (\bar{k}_p)^2$. Rearrangement produces $\sum(k_i^2) = N_p [V_{kp} + (\bar{k}_p)^2]$. Consider first the scenario in which the N_p successful parents have random reproductive success, in which case $V_{kp} \approx \bar{k}_p$, so $\sum(k_i^2) \approx N_p \bar{k}_p (1 + \bar{k}_p) = N_p (2N/N_p) (1 + 2N/N_p) = 2N (1 + 2N/N_p)$. Substituting into equation 1 and dropping the -1 in the numerator, since it will be trivial compared with N , produces:

$$N_e \approx \frac{2N}{\frac{2N(1+2N/N_p)}{2N} - 1} = \frac{2N}{1 - 1 + 2N/N_p} = N_p, \text{ so } \frac{N_e}{N} \approx \frac{N_p}{N} \quad (2)$$

(assumes random reproductive success of the N_p successful parents). This produces the intuitive result that, if the N_p successful parents behave like a mini Wright-Fisher ideal population, the overall effective size of the entire population is N_p and the N_e/N ratio is just N_p/N .

This analysis can be generalized by allowing V_{kp} to be any multiple (α) of the mean reproductive success: $V_{kp} = \alpha \bar{k}_p$. In that case, $\Sigma(k_i^2) = 2N(\alpha + 2N/N_p)$ and

$$N_e \approx \frac{2N}{\frac{2N(\alpha + 2N/N_p)}{2N} - 1} = \frac{2N}{\alpha - 1 + 2N/N_p}. \quad (3)$$

Situations in which N_p/N is very small are of primary interest, in which case the $2N/N_p$ term in the denominator of equation 3 will be very large, so the -1 term in the denominator also can be ignored, producing:

$$N_e \approx \frac{2N}{\alpha + 2N/N_p}, \text{ so } \frac{N_e}{N} \approx \frac{2}{\alpha + 2N/N_p}. \quad (4)$$

Unless α is very large, it also will be dwarfed by the other term in the denominator, which again will produce the result that $N_e/N \approx N_p/N$. For example, even if α is as large as \bar{k}_p (*i.e.* the variance in reproductive success is \bar{k}_p times the mean, where \bar{k}_p is a large number), N_e is only reduced by 50%, so $N_e/N \approx 0.5N_p/N$.

Hedrick (2005) considered a variation of this scenario in which each of the N_p successful parents produced exactly the same (large) number of progeny. In this case, $\alpha = 0$, which again produces $N_e/N \approx N_p/N$ from equation (4). Hedrick (2005) also showed that considering a third class of parents (those who produce exactly two offspring each) leads to the following result:

$$\frac{N_e}{N} \approx \frac{N_p}{N(1-y)^2} \quad (5)$$

where y is the fraction of all N adults that produce exactly two offspring. Using some numerical examples, Hedrick (2005) showed that N_e/N remains quite low in the presence of this third class of parents unless they make up a large fraction of the population.

To summarise, under the discrete generation model, when a small number (N_p) of parents dominate reproduction, the N_e/N ratio will be close to N_p/N regardless how reproductive success is partitioned among the successful parents and regardless whether some other parents manage to contribute small numbers of offspring.

OVERLAPPING GENERATIONS

An important limitation of the above analyses is that they implicitly assume discrete generations and fail to consider age structure, whereas most species (and all of those for which very small N_e/N ratios have been reported), have overlapping generations. Overlapping generations and reproduction in >1 year or season (iteroparity) have some important consequences for small N_e/N ratios that have not been quantitatively evaluated before. After considering the general model developed by Hill (1972), which assumes a constant population size, stable age structure and independence of birth and death rates over time, the consequences of some violations of these assumptions are evaluated.

Hill's model

To account for overlapping generations, the discrete-generation formula for N_e can be modified as follows (Hill, 1972):

$$N_e = \frac{4N_1T}{V_{k\bullet} + 2} \quad (6)$$

where N_1 is the number of offspring produced each time period, T is generation length and $V_{k\bullet}$ is lifetime variance in reproductive success of the N_1 individuals in a cohort. Any age up to the age of first reproduction can be used to enumerate the individuals in a cohort, provided that reproductive success and $V_{k\bullet}$ are also based on production of offspring of that same age. To see the effect of age at maturity on N_e/N , consider two hypothetical species, species A with age at maturity = 1 year and species B with age at maturity = 1 + z years. Further, assume that both species have adult lifespan = L_A years, the same annual adult survival and the same pattern of age-specific fecundity. Because species B delays onset of reproduction, generation length for species B will be higher by z years, while $V_{k\bullet}$ will not be affected. As a consequence N_e/N is higher in species B than in species A. In fact, N_e/N can be >1 in species that delay maturity for many years or reproductive cycles (Waples *et al.*, 2013). Because interest here is on factors that can produce tiny N_e/N ratios, it is assumed that age at maturity = 1 and that N_1 is the number of individuals in a cohort that survive to age 1 year. This does not mean that species with delayed maturity cannot have tiny N_e/N , just that it is a little harder than it is for species that mature at age 1 year.

Apart from age at maturity, the other key demographic traits that affect the N_e/N ratio are: adult lifespan or longevity, which is determined by the annual survival rate; age-specific patterns of change in survival (s_x) and especially fecundity (b_x), where x indicates age, and age-specific patterns of change in $\Phi_x = V_x/b_x$ = the ratio of the variance to the mean reproductive success in one time period for individuals of the same age and sex. These will be examined sequentially using the AgeNe model (Waples *et al.*, 2011), which calculates lifetime $V_{k\bullet}$ and N_e (using equation 6) by grouping individuals by age at death.

In age-structured species, some individuals live longer than others and hence have more opportunities to reproduce. This increases lifetime variance in reproductive success; because $V_{k\bullet}$ appears in the denominator of equation (6), this (all else being equal) reduces N_e and N_e/N . Increasing the adult lifespan, however, also increases generation length (which appears in the numerator of equation 6) as well as the number of adults in the population (which appears in the denominator of N_e/N), so all these factors must be considered jointly to assess overall effects on the effective size–census size ratio. Effects of longevity on N_e/N are isolated by considering a population that has constant fecundity with age and constant $\Phi_x = 1$, which means that (for example) all age 7 year males have random reproductive success among themselves. This hypothetical population also has a constant adult survival rate that produces total life spans (and maximum ages, ω) of 10, 20, 30, 40 and 50 years. The latter was achieved by defining cumulative survival (l_x) to be 1.0 at age 1 year and choosing constant s_x values that produce $l_\omega = 0.01$ (1% survival through the maximum age) for each value of ω . Corresponding s_x values for $\omega = 10$ –50 were 0.600, 0.785, 0.853, 0.891 and 0.910. Results (Fig. 1) show that, under these conditions of constant vital rates, increasing longevity reduces

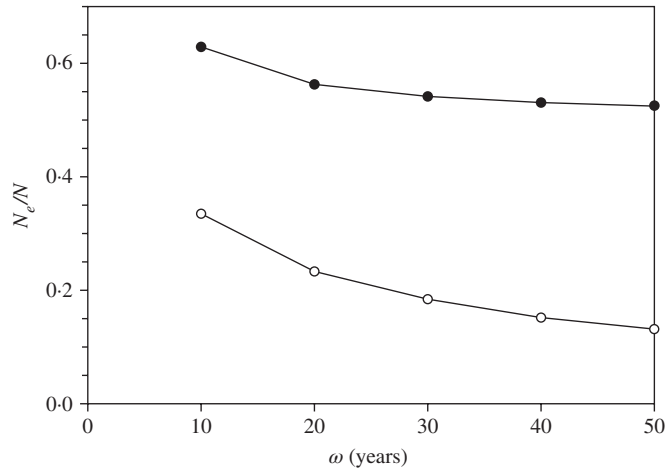


FIG. 1. The ratio N_e/N as a function of longevity (maximum age, ω years). In each case, adult survival was held constant at a value that produced cumulative survival (l_x) of $l_x = 0.01$ at age ω years. —●—, scenarios in which fecundity (b_x) and Φ_x do not vary with age; —○—, scenarios in which b_x and Φ_x are proportional to age (x).

N_e/N , but the reduction is rather modest (from 0.629 with $\omega = 10$ to 0.525 with $\omega = 50$, a decline of only 17%). Therefore, by itself increased longevity cannot be expected to produce tiny N_e/N ratios. Unless otherwise specified, the analyses below assume that maximum longevity is 40 years.

Although many species (such as many birds and mammals) have vital rates that are approximately constant across their adult lifespan, the same is not true for most marine ectotherms with indeterminate growth. In these species, older individuals are larger and generally have higher fecundity; this increases the reproductive payoff for individuals that survive to reproduce many times and further increases $V_{k\bullet}$. In addition, older females not only have more eggs, they may produce better eggs that have a higher chance of producing a viable offspring—the ‘big old fat fecund female fish’ (BOFFFF) hypothesis (Berkeley *et al.*, 2004; Hixon *et al.*, 2014). The same could potentially be true of males. To the extent that such effects occur, they would further enhance the reproductive disparities associated with increased longevity and fecundity that increases with age. Because only offspring that survive to age 1 year are considered, both increased number of eggs and increased survival of eggs for females of older ages can be accommodated by appropriate scaling of effective age-specific fecundity.

One simple way to do this is to assume that effective fecundity (in terms of production of offspring that survive to age 1 year) increases linearly with age, which is not uncommon in long-lived marine fishes (Fig. 2). In the first evaluation of this general scenario, relative fecundity was assumed to be proportional to age, while the other vital rates were constant at $s_x = 0.891$ and $\Phi_x = 1$, with $\omega = 40$. This pattern of increasing fecundity with age reduced N_e/N to 0.42, a decline of 22% from the value (0.53) obtained when all vital rates are constant with age (Fig. 3). In a more extreme scenario of this type, effective fecundity increases exponentially with age. If it is assumed that relative fecundity doubles every year after age 1 year, N_e/N drops sharply to 0.059 (Fig. 3). This particular scenario, however, is not very plausible biologically, as it would

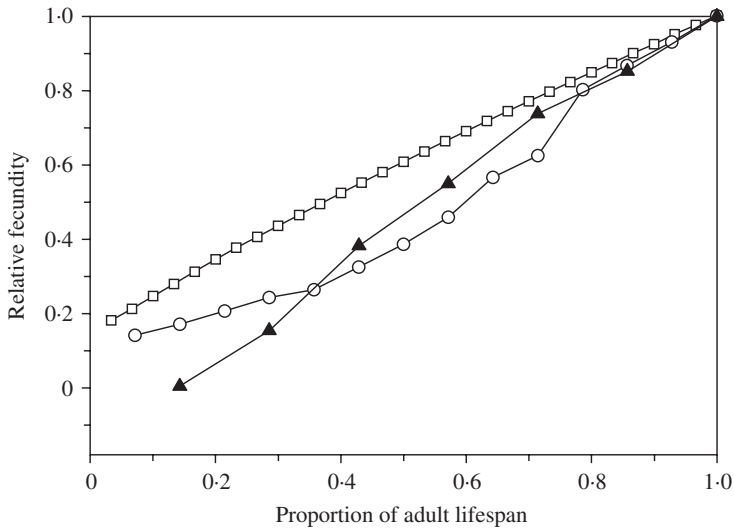


FIG. 2. Pattern of change in fecundity with age for three marine species: —○—, Atlantic cod, *Gadus morhua*; —□—, red drum, *Sciaenops ocellatus*; —▲—, orange cup coral *Balanophyllia elegans* (data are from Waples *et al.*, 2013).

require that a 40 year old individual produce 5×10^{11} times as many offspring as an age 1 year adult. Furthermore, although the resulting N_e/N value for this scenario is a bit below 0.1, it is nowhere as small as the tiny values that have been reported for some marine species in the literature. Therefore, these tiny N_e/N estimates cannot be explained entirely by increasing fecundity with age, although this is probably a contributing factor. The remaining analyses will use the more plausible scenario in which effective fecundity is proportional to age, as in Fig. 2.

If increasing fecundity with age creates larger disparities in lifetime reproductive success between those individuals that do and do not survive to old age, why does not this have a larger effect on the N_e/N ratio? The answer is straightforward: shifting more and more reproduction to older age classes also increases generation length. Because T appears in the numerator of equation 6 and $V_{k\bullet}$ appears in the denominator, the effects on T and $V_{k\bullet}$ largely cancel each other, leading to only modest net changes in N_e/N .

This brings us to the final major factor that determines N_e and N_e/N : Φ_x . Φ_x has no effect on generation length because the mean age of parents is unaffected; it only affects the variance around the mean reproductive success. But Φ_x can have a large influence on $V_{k\bullet}$ and hence N_e/N . In the next scenario, therefore, both b_x and Φ_x were proportional to age, as has been estimated for Atlantic cod *Gadus morhua* L. 1758 (Kuparinen *et al.*, 2016). This produced $N_e/N = 0.15$, 63% lower than obtained when only fecundity was proportional to age. Still, this is far above what can be considered a tiny ratio of N_e/N .

The final series of analyses evaluated how large Φ_x (assumed to be fixed) must be to produce very low N_e/N ratios. If Φ is constant at 10 for every age, N_e/N is actually a little higher than with additive Φ_x , but with $\Phi = 1000$ (variance, $V = 1000$ times the mean), N_e/N drops to 3.4×10^{-3} , which finally is in the upper end of the range of tiny

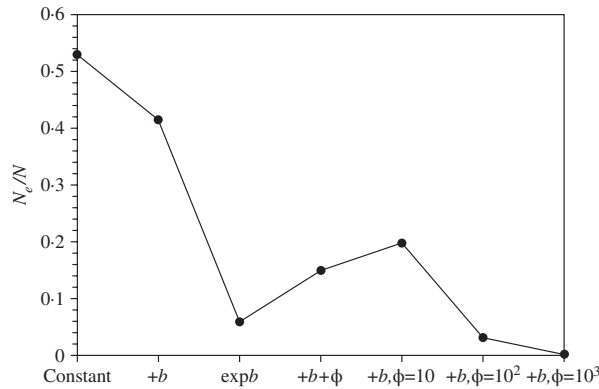


FIG. 3. The ratio N_e/N for hypothetical scenarios involving age-specific changes in vital rates. N_e/N was calculated from age-specific vital rates using AgeNe (Waples *et al.*, 2011). In all scenarios, individuals matured at age 1 year and lived to a maximum age $\omega = 40$ years with constant adult survival of $s_x = 0.891$. In the first three scenarios, $\Phi_x = 1$ for every age. Constant, fecundity is constant with age; +b, fecundity is proportional to age; expb, relative fecundity doubles each year; +b+ Φ , fecundity is proportional to age, and $\Phi_x = x$; +b, $\Phi = 10$, fecundity is proportional to age, and $\Phi_x = 10$ for every age; +b, $\Phi = 10^2$, fecundity is proportional to age, and $\Phi_x = 100$ for every age; +b, $\Phi = 10^3$, fecundity is proportional to age, and $\Phi_x = 1000$ for every age ($N_e/N = 3.4 \times 10^3$).

N_e/N that have been proposed for marine species. Effects of large Φ can be evaluated more generally by taking advantage of the property (Waples *et al.*, 2011) that, when Φ is constant with age in a stable population, an increase in Φ by the amount Δ_Φ units causes $V_{k\bullet}$ to increase by $2\Delta_\Phi$ units. A modified form of equation 6 that reflects this is:

$$N_e = \frac{4N_1T}{V_{k1\bullet} + 2\Delta_\Phi + 2} \quad (7)$$

(assumes Φ is constant with age), where $V_{k1\bullet}$ is the lifetime variance in reproductive success when Φ is fixed at 1. Situations where Φ and hence Δ_Φ are very large are of primary interest, in which case $\Phi \approx \Delta_\Phi$ and, to a good approximation:

$$N_e \approx \frac{4N_1T}{2\Phi} = \frac{2N_1T}{\Phi} \quad (8)$$

(assumes Φ is large and constant with age). That is, when Φ is large and constant with age, N_e (and hence N_e/N) are inversely proportional to Φ .

This effect can be illustrated using the above example with maximum age = 40, fecundity proportional to age and Φ fixed at 1. For this scenario, if N_1 is set at production of 10 000 age 1 year recruits year⁻¹, adult $N = 90\,835$, $T = 15.5$ years and $V_{k\bullet} = 14.4$, leading to $N_e = 37\,760$ and $N_e/N = 0.416$. Substituting into equation 8 to evaluate effects of large Φ produces this:

$$\frac{N_e}{N} \approx \frac{2N_1T}{N\Phi} = \frac{2 \times 10\,000 \times 15.5}{90\,835\Phi} \approx \frac{3.4}{\Phi}. \quad (9)$$

With $\Phi = 1000$, this approximation yields $N_e \approx 3.4 \times 10^{-3}$, in good agreement with the exact value shown in Fig. 3 calculated using equation (6). It is easy to see that N_e/N in species with overlapping generations can be arbitrarily small if Φ is assumed to be arbitrarily large. A good order-of-magnitude approximation is the following:

$$N_e/N \approx 1/\Phi \quad (10)$$

It is worth noting here that Φ represents the ratio of the variance to the mean reproductive success for a single age class of one sex. Φ can only be very large if the number in the age–sex class (N_x) is very large. For long-lived iteroparous species, Φ generally will be constrained to be $\leq N_x$, although this will not always hold for older age classes if fecundity increases steeply with age.

A simple example illustrates the biological meaning of large values of Φ . Consider a large marine fish population that each year produces $N_1 = 10^6$ recruits that survive to age at maturity. If annual mortality is constant at $d = 0.15$, the total number of adults will be approximately $N = N_1/d = 6.67 \times 10^6$, so the mean genetic contribution in each time period by all adults will be $\bar{k} = 2N_1/N = 2d = 0.3$. Considering a single age class and assuming that the number of adults in the age class is 5% of the total (which would be typical for a relatively long-lived species), then $N_x = N/20 = 3.33 \times 10^5$. Assuming these parents have the same mean reproductive success as the overall population, they would produce 10^5 offspring that survive to age at maturity. Now further assume that of the N_x potential parents of age x years, only N_p successfully reproduce, with random variation in offspring number. For $N_p = 3333, 333, 33$ or 3 (which represent a fraction $10^{-2}, 10^{-3}, 10^{-4}$ or 10^{-5} of all N_x potential parents), the resulting Φ_x values are 30.7, 301, 3031 and 33 334, respectively (Table I). Therefore, to produce $\Phi_x \approx 1000$ (as in the example above) requires that only about one in 1000 potential parents reproduces successfully. For realistic scenarios involving long-lived iteroparous species, Φ_x is the same order of magnitude as the inverse of the fraction of potential parents that successfully reproduce ($\Phi_x \approx N_x/N_p$). Using the $N_e/N \approx 1/\Phi$ approximation noted above, this means that with overlapping generations, constant vital rates and sweepstakes reproduction, the effective size–census size ratio is approximately equal to the fraction of potential parents that successfully reproduce in a given time period ($N_e/N \approx N_p/N_x$). This is similar to the result derived above for discrete generations, except that the N_p/N ratio applies to reproduction by each year class in 1 year or season rather than a full generation.

Variations to Hill's model

Variable recruitment: although Hill's model assumes that population size is constant and age structure is stable, the method is robust to random demographic stochasticity (Waples *et al.*, 2011, 2014). Furthermore, Felsenstein (1971) showed that his related model is still accurate if a population increases or declines at a steady rate. Some long-lived marine species, however, have highly variable recruitment, with little or no successful reproduction in many years and occasional large pulses of strong recruitment. If successful recruitment occurs less frequently than once per generation, the population is not likely to be viable in the long term. Taken to a plausible extreme, therefore, this type of scenario can be evaluated by assuming that the population has a strong

TABLE I. Relationship between age-specific values of Φ_x and the number of successful parents (N_p) producing a fixed number of progeny in one time period. All scenarios consider a typical age class of $N_x = 3.33 \times 10^5$ adults that produce 10^5 progeny each time period, for a mean reproductive success of $\bar{k} = 0.30$. In each column, a different number of successful parents is assumed to produce all the offspring, with random variation in reproductive success. $\bar{k}_{Np} = 2 \times 10^5 / N_p$ is the mean genetic contribution of the N_p successful parents, V_k is overall variance in reproductive success among all N_x potential parents and $\Phi_x = V_k / \bar{k}$ is the ratio of the variance to the mean reproductive success for individuals of age x years

\bar{k}	$N_x (\times 10^5)$	Progeny	N_p	\bar{k}_{Np}	V_k	Φ_x	$N_p N_x$
0.30	3.33	10^5	3333	30.0	9.2	30.7	10^{-2}
0.30	3.33	10^5	333	300.3	90.3	301.0	10^{-3}
0.30	3.33	10^5	33	3030.3	909.3	3031.0	10^{-4}
0.30	3.33	10^5	3	33333.3	10000.2	33334.0	10^{-5}

recruitment once per generation, with zero successful reproduction in the intervening years. But this is just a discrete-generation model, with consequences as discussed above. Therefore, variable recruitment by itself is not likely to lead to tiny N_e/N ratios, although it could if the N_p individuals responsible for the successful recruitment are a tiny fraction of all adults.

Persistent individual differences: the assumption in Hill's model that survival and reproduction are independent across time is necessary to make the analysis tractable but unrealistic for many populations. Intuitively, N_e and N_e/N should be reduced if the same individuals are consistently good or bad at reproducing across multiple time periods, and Lee *et al.* (2011) showed that this indeed is the case, although the effect was rather modest: in the most extreme scenario they considered, N_e/N was reduced by less than one order of magnitude (from about 0.5 to 0.1–0.2). Is this issue a weak link in the SRS hypothesis; *i.e.* is it necessary to assume that the same very few individuals are sweepstakes winners year after year after year? This seems highly implausible, given Hedgcock's (1994) concept of the sweepstakes winners being the parents that (by luck) just happened to deliver their families of eggs or larvae to one of the few places in the ocean where they could survive and grow. This is not a serious limitation for the SRS hypothesis, however, for the following reason.

Consider two scenarios for a long-lived species in which only a small number of parents successfully reproduce each year: first, the same N_p parents are successful every year across a generation; second, each year the N_p successful parents are randomly chosen from the population as a whole. In scenario 1 the effective size per generation will be approximately N_p , in which case the consequences for N_e/N are the same as they are in the discrete generation model. In scenario 2, assuming the total population of adults is very large, the successful parents that are randomly chosen each year are expected to be nearly or completely non-overlapping (*i.e.* the chances that any individual will win the sweepstakes more than once is very small). In that case, the number of successful parents over a period of a generation (and hence the approximate effective size per generation) will be approximately TN_p . Given the typical range of generation lengths for long-lived marine species (10–20 years or so), assuming a complete turnover of successful parents each year would only increase N_e/N by roughly one order of

magnitude. This would not preclude tiny N_e/N values, provided that the fraction of successful spawners each year is sufficiently small.

Intermittent breeding: the converse of positive correlations between reproductive success of individuals across time is inverse correlations caused by intermittent or skip breeding, which occurs when energetic costs of reproduction (including any associated migrations) reduce the chances that an individual that reproduces in 1 year will reproduce in a subsequent year (Shaw & Levin, 2013). It is common for females of large mammal species to skip breeding for one or more cycles after giving birth, and the same can be true for many other species and, occasionally, for males as well. Although skip breeding reduces the number of adults available to breed in any given year [and hence can sharply reduce the effective number of breeders per year (N_b) in some species], this phenomenon serves to reduce lifetime V_k and hence raise N_e/N slightly (Waples & Antao, 2014). Because the effect on N_e is small and expected to be positive, whether a species employs skip breeding is not likely to have an appreciable effect on opportunities for tiny N_e/N ratios.

To summarise, all else being equal, increasing longevity reduces N_e/N , but by itself the effect is rather modest. Stronger reductions can occur if a long adult lifespan is coupled with fecundity and Φ_x that increase with age, but plausible scenarios of this type, even those that incorporate the BOFFFF hypothesis, are still unlikely to produce N_e/N smaller than about 10^{-1} . Similarly, some life history variants such as variable recruitment or persistent individual differences in reproductive success can reduce N_e/N , but again their effects are expected to be relatively modest. In species with overlapping generations, tiny N_e/N ratios are only possible if the variance in reproductive success of same-age, same-sex individuals is orders of magnitude higher than the mean. This, in turn, is possible only if just a relatively few individuals are responsible for most of the successful reproduction. Therefore, although age-structure and iteroparity can influence effective size in predictable ways, these life-history traits cannot by themselves produce tiny N_e/N ratios; that still requires some variation of Hedgcock's SRS hypothesis. With sweepstakes reproduction and overlapping generations, if the same adults consistently produce most of the offspring, N_e/N will approximately equal the fraction of potential adults that successfully reproduce in a given year, just as in the discrete-generation model. If (mostly) different parents are sweepstakes winners each year, N_e/N will be increased by a factor equal to the generation length, which will be roughly one order of magnitude for many marine species. Because very little is known about variance in reproductive success of individuals of the same age and sex, collecting empirical data of this type would considerably improve the understanding of SRS.

ESTIMATING N_e AND N_e/N

ESTIMATING N

In a computer model, it is easy to count the number of individuals in the population, but that is not the case in the real world, particularly for large marine populations that can include millions or even billions of individuals that are (at best) difficult to observe directly. Also, calculation of adult N ideally would account for the fraction in each age class that are sexually mature. Again, this is easy in a computer model but much more challenging in real-world populations and this contributes uncertainty to estimates of

N . Most of what follows focuses on estimates of N_e , but it is important to remember that estimating N_e/N involves separate estimation of two parameters, each of which presents major challenges for marine species.

ESTIMATING N_e

Practical considerations

All of the tiny estimates of contemporary N_e/N for marine species are from indirect genetic methods that use a genetic index that is expected to be a function of $1/N_e$. Most of the estimates are from the temporal method (which requires two or more samples spaced in time) or the linkage disequilibrium (LD) method, which uses single samples. The respective genetic indices and their expected values are as follows:

$$E(F) \approx \frac{t}{2N_e} + \frac{1}{S} \quad (11)$$

$$E(r^2) \approx \frac{1}{3N_e} + \frac{1}{S}. \quad (12)$$

In the above, S is the number of individuals in a sample, F is the standardized variance in allele frequency between two temporal samples, t is the number of generations between temporal samples and r^2 is the squared correlation of alleles at different gene loci. The basic approach of these moment-based methods is straightforward: (1) develop theoretical expectations for contributions of drift and sampling error to the genetic index of interest (as in the equations above); (2) compute the overall index; (3) subtract from that the expected contribution from sampling error; (4) use the result to estimate N_e using a simple rearrangement of equation 11 (temporal method; Waples, 1989) or equation 12 (LD method; Hill, 1981).

Inspection of equations (11 and 12) makes it clear why estimating effective size in large populations is very challenging with indirect genetic methods. For a sample size that is common for marine species ($S = 50$), the contribution of random sampling error to the genetic index will dwarf the signal from drift unless true N_e is very low (Fig. 4). Vastly increasing sample size to $S = 5000$ can substantially improve performance if true N_e is no larger than about 10^4 , but even such large samples are ineffective in reducing the signal-to-noise ratio problem if true N_e is as large as 10^6 (Fig. 4). When $N_e \gg S$, the crucial step in the estimation procedure is (3), because a small error in correcting for sampling error can have a huge effect on \hat{N}_e . Even if the theoretical sampling terms in equations (10 and 11) are exactly correct and all model assumptions are met, the problem remains that both genetic drift and sampling error are random stochastic processes and the mean values only approach the theoretical expectations with a great deal of replication. For example, to get within one order of magnitude of the expected drift signal from LD when true $N_e = 10^6$, after accounting for sampling error a user must find a mean value of r^2 between 0.00000003 and 0.00000333. For estimating N_e in very large populations, this means that the genetic index must be based on very large amounts of data (samples of individuals or gene loci) to provide any hope of high precision.

These challenges can be illustrated using simulated data for populations with a wide range of true N_e values and applying the estimators and amounts of data that were

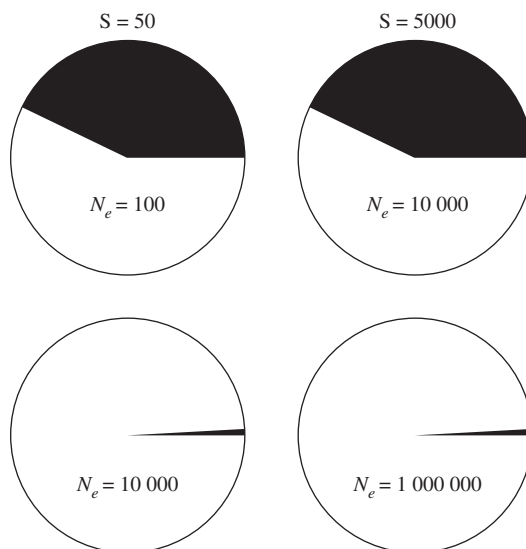


FIG. 4. Relative contributions of genetic drift (■) and random sampling error (◻) to the standardized variance in allele frequencies (F) between samples taken three generations apart. Results are based on equation 11 for the standard temporal method with plan II sampling, assuming sample sizes of $S = 50$ or $S = 5000$ individuals and N_e as shown.

commonly used to generate published tiny estimates of N_e/N . One hundred diallelic, single nucleotide polymorphism (SNP) loci were simulated in ideal populations of constant size $N_e = N = 100$ to 1 000 000 for a burn-in period of five generations (enough to reach equilibrium for LD; Waples, 2005). An initial sample of $S = 50$ individuals was taken using Plan II sampling (Nei & Tajima, 1981; Waples, 1989) and a final temporal sample was taken after simulating an additional three generations of drift. The second temporal sample was also used for the LD estimates. For some scenarios, 1000 loci were simulated, or samples sizes of $S = 200$, 1000 or 5000 were used. Simulations using $N_e = 10^7$ were also conducted for the temporal method only (memory limitations precluded generating such large matrices for LD analyses). Finally, infinite N_e was modelled by randomly choosing genotypes for LD samples (equivalent to having an infinite number of parents) and by taking two replicate samples from the same population for the temporal estimates. [Simulations were conducted in R (www.r-project.org) using code that is available from the author on request.]

Diallelic loci are easier to simulate than the microsatellites used in most of the published studies, but the 100 SNP loci modelled here provide approximately the same amount of information as 10 moderately variable microsatellites [the 11 microsatellite studies reviewed by Hauser & Carvalho (2008) used a mean of eight loci]. Samples were used to estimate N_e using both the standard temporal method (using Nei and Tajima's estimator of F) and the bias-corrected LD method (Waples, 2006); in addition, a combined estimate was computed as the harmonic mean of the two estimates.

For small populations (true $N_e = 100$), both methods are largely unbiased and have good precision, as has been reported elsewhere. Almost no estimates were below 50 and very few were above 200 [Fig. 5(a)]. The situation changes dramatically, however, if true N_e is 100 times as large ($N_e = 10\,000$): about half the estimates from each method

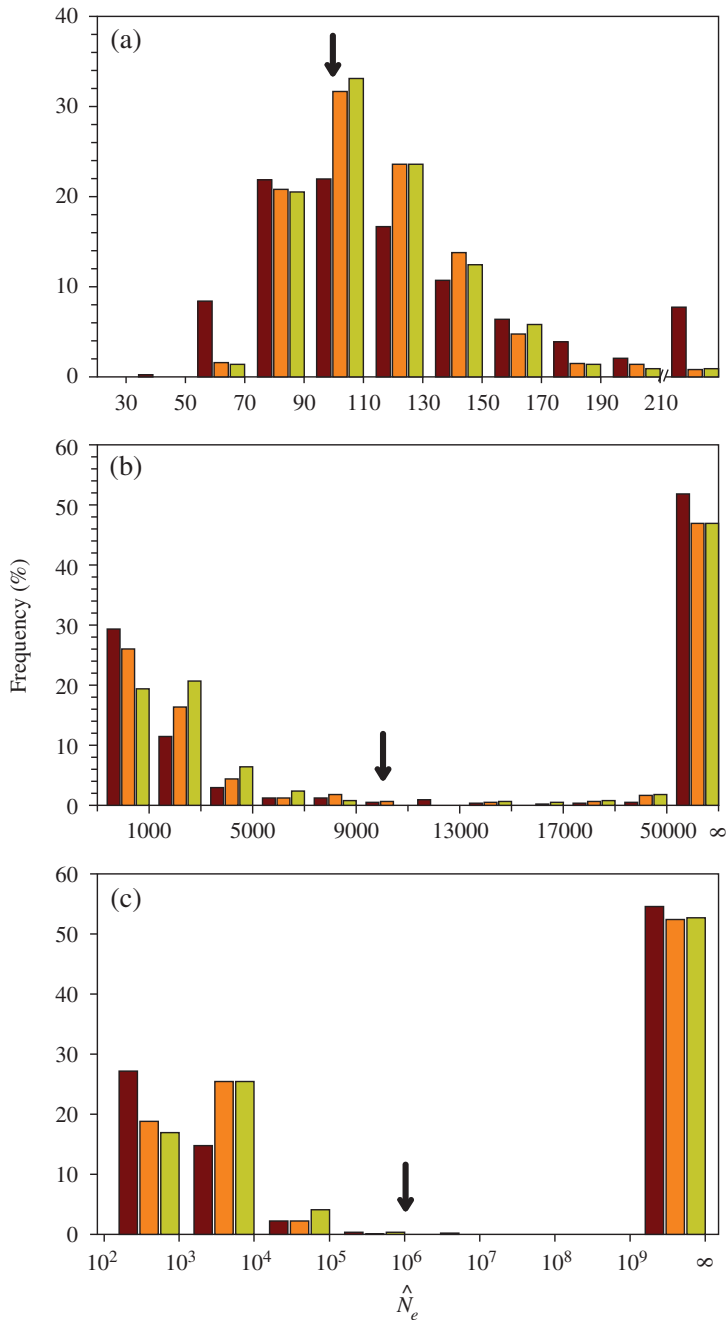


FIG. 5. Frequency distribution of estimates of effective population size (\hat{N}_e) for simulated data where (a) true N_e (\downarrow) = 100 (using 1000 replicates), (b) true $N_e = 10^4$ (500 replicates), (c) true $N_e = 10^6$ (500 replicates) using: ■, the standard temporal method, computed for samples taken three generations apart; ■, the linkage disequilibrium method; ■, the harmonic mean of the two estimates. All results are for samples of $S = 50$ individuals genotyped for 100 diallelic (single nucleotide polymorphism, SNP) loci.

are infinitely large and most of the remainder are <25% of the true value [Fig. 5(b)]. This bimodal pattern to \hat{N}_e is even more dramatic when true N_e is 10^6 : over half the estimates are infinite, while most of the rest are in the hundreds or low thousands, downwardly biased by three to four orders of magnitude [Fig. 5(c)]. Almost no estimates are within an order of magnitude of the true N_e . It might be expected that combining estimates from the two methods would lead to greater precision, but there was little difference in performance of the combined and the LD estimates in this example.

Note that as true N_e gets larger, the lower estimates stay in about the same range (hundreds to low thousands with $S = 50$; Table II and Fig. S1, Supporting Information). This means that a large fraction of point estimates of N_e/N can be expected to be increasingly small as true N_e becomes increasingly large. These results do not reflect a large systematic bias in the estimates; instead, they reflect the increasingly bimodal pattern of the estimates as effective size increases, such that virtually all estimates are either infinity or very small compared to true N_e .

It is interesting to note that this bimodal pattern also occurs with small-sample estimates based on the true pedigree. The simulations kept track of the parents of offspring that appeared in the samples of 50 individuals each and N_e was estimated for each sample using equation (1). With true $N_e = 10^6$, one does not expect to find any siblings in a sample of only 50 offspring, and that was the case for 499 of the 500 replicates. When each offspring has two unique parents, $\Sigma k_i^2 = \Sigma k_i$ and inbreeding effective size is infinitely large (*cf.* equation 1). The largest non-infinite estimate of N_e from parentage analysis that is possible with a sample of 50 offspring occurs when a single pair of half-sibs is found; this produces an estimate of $\hat{N}_e = 4950$, which was the value obtained in the 500th replicate. That is, with a sample of 50 offspring, even if parentage can be determined with 100% accuracy, it is impossible to obtain an estimate of N_e that is larger than 4950 and smaller than infinity. Put another way, with parentage analysis and N_e estimation based on standard demographic equations (*e.g.* equation 1), if true effective size is very large, it is impossible to obtain an estimate of N_e that is within several orders of magnitude of the true value unless the sample of individuals is also very large.

Fortunately, two approaches can help mitigate this rather gloomy prospect for obtaining reliable estimates of effective size and N_e/N when true N_e is large. First, all indirect methods for estimating N_e have ways of placing confidence intervals (C.I.s) around the point estimates; for the two moment-based methods used here, the C.I.s are based on the parametric χ^2 -distribution, as detailed by Waples (1989, 2006). As would be expected for methods that are largely unbiased, most of the small point estimates for $N_e = 10^4$ and 10^6 depicted in Fig. 5 had upper bounds that included the true N_e . This illustrates the importance of considering not only the point estimates but also the upper bounds of the C.I.s. This is tricky, however, because these C.I.s never incorporate all of the uncertainty associated with the estimates. For example, they do not account for uncertainty or biases associated with violations of simplistic model assumptions (*e.g.* discrete generations, closed populations and selective neutrality), nor do they account for biases associated with the nearly impossible task of achieving a completely random sample from a large wild population. Although model misspecifications like these were not explicitly evaluated (collectively, they encompass an enormous parameter space), the likely consequences are easy to predict. Given that, when true N_e is large and all model assumptions are met, a large fraction of point estimates will be orders of magnitude smaller, even modest downward biases in the estimates could easily cause the upper

TABLE II. Percentage of genetic estimates of N_e falling in specified ranges, for simulations with different values for N_e , sample size (S) and number of diallelic loci (L)

N_e	S	L	Metric	\hat{N}_e											Modal \hat{N}_e/N_e	
				10-100	100-1000	10^3-10^4	10^4-10^5	10^5-10^6	10^6-10^7	10^7-10^8	10^8-10^9	∞				
100	50	100	Temp	41.3	58.4	0.3	-	-	-	-	-	-	-	-	-	~1.0
			LD	39.5	60.5	-	-	-	-	-	-	-	-	-	-	-
1000	50	100	Temp	-	41.2	19.0	2.4	-	-	-	-	-	-	-	-	37.4
			LD	-	47.0	27.0	2.4	-	-	-	-	-	-	-	-	-
10^4	50	100	Temp	-	26.0	17.0	2.0	-	-	-	-	-	-	-	-	55.0
			LD	-	19.5	26.0	3.0	0.5	-	-	-	-	-	-	-	-
10^5	50	100	Temp	-	22.4	20.6	2.6	0.2	-	-	-	-	-	-	-	54.2
			LD	-	19.4	28.8	2.8	-	-	-	-	-	-	-	-	-
10^5	50	1000	Temp	-	3.0	28.0	5.0	0.4	0.2	0.2	0.2	-	-	-	-	63.4
			LD	-	-	27.0	28.8	2.4	0.4	0.4	0.4	-	-	-	-	-
10^6	50	100	Temp	-	27.4	15.0	2.4	0.4	-	-	-	-	-	-	-	54.8
			LD	-	19.0	25.6	2.4	0.2	0.2	0.2	-	-	-	-	-	-
10^6	200	100	Temp	-	2.2	35.2	6.0	0.4	-	-	-	-	-	-	-	56.2
			LD	-	0.2	35.8	10.6	1.2	0.2	0.2	-	-	-	-	-	-
10^6	1000	100	Temp	-	-	16.4	31.0	4.2	0.6	0.6	-	-	-	-	-	47.8
			LD	-	-	5.4	37.4	4.6	0.2	0.2	-	-	-	-	-	-

TABLE II. Continued

N_e	S	L	Metric	\hat{N}_e											Modal \hat{N}_e/N_e
				10-100	100-1000	10^3-10^4	10^4-10^5	10^5-10^6	10^6-10^7	10^7-10^8	10^8-10^9	∞			
10^6	5000	100	Temp	-	-	-	30.0	19.2	1.4	-	-	-	49.4	0.01-0.1	
			LD	-	-	24.4	31.2	2.2	0.2	42.0	0.1-1				
10^7	50	100	Temp	-	23.6	17.6	2.0	-	-	-	-	-	57.0	$10^{-5}-10^{-4}$	
			Temp	-	25.0	16.0	2.6	-	-	-	-	-	56.4	∞^{-1}	
∞	5000	100	Temp	-	19.2	27.2	2.8	0.6	-	-	-	-	50.2	∞^{-1}	
			LD	-	-	-	24.8	17.8	2.2	0.2	-	-	55.0	∞^{-1}	
∞	5000	100	Temp	-	-	-	19.8	28.2	3.0	-	-	-	49.0	∞^{-1}	
			LD	-	-	-	-	-	-	-	-	-	-	-	

Temp, estimate using the temporal method with samples separated by three generations.
 LD, estimates for the linkage disequilibrium method.
 ∞ , all estimates for which sampling error explained all of the observed variation.
 \hat{N}_e/N_e indicates how far off the mode of the finite \hat{N}_e estimates are from the true N_e .

c.i.s to fall below the true N_e . Finally, c.i.s around \hat{N}_e cannot be directly converted to c.i.s around N_e/N without also explicitly considering uncertainty in the estimate of N and the covariance of N_e and N , but that is seldom done.

To the extent that it is feasible, one can increase precision by obtaining more data (larger samples of individuals and genetic markers). Although the number of individuals and loci included in the simulations were comparable with those actually used in many of the studies that have reported tiny N_e/N (summarized by Hauser & Carvalho, 2008), it is now relatively easy to obtain many thousands of SNPs even for non-model species, and considerably larger samples of individuals are possible for some species. With true $N_e = 10^6$, increasing sample size from $S = 50$ to 200 improved performance of the estimators somewhat (fewer estimates were in the hundreds), but the distribution of \hat{N}_e was still strongly bimodal; over half the estimates were infinite and most of the finite estimates were $<10^4$ and would produce estimates of N_e/N in the range 10^{-2} – 10^{-3} [Table II and Fig. S2(b), Supporting Information]. In some cases, very large samples (up to 5000 individuals or so) can be collected from marine species (MacBeth *et al.*, 2013). With true $N_e = 10^6$, increasing sample size by two orders of magnitude (from $S = 50$ to 5000) also shifted the bulk of the finite estimates by two orders of magnitude (from 10^2 – 10^4 to 10^4 – 10^6), with most of the remainder still being infinite (Table II). Under these conditions, no estimates for either method were as much as two orders of magnitude lower than the true effective size. When modelled N_e was infinitely large, however, the distribution of estimates using $S = 5000$ remained unchanged (Table II). This means that, even with very large sample sizes, tiny estimates of N_e/N can occur if true N_e is large enough.

Memory limitations precluded simulating >100 SNP loci with $N_e = 10^6$, but results for 1000 SNPs and $N_e = 10^5$ also produced fewer estimates in the hundreds and slightly fewer infinite estimates [Fig. S1(a), Supporting Information]. Still, about half of the finite estimates were in the low thousands, which would produce estimates of N_e/N in the range 10^{-1} – 10^{-2} .

It is important to note here that all of the simulations evaluating performance of genetic estimates of N_e used discrete generations and ideal populations in which $N_e = N$. Even so, all scenarios with large true N_e produced a large fraction of estimates of N_e/N that were orders of magnitude too small. But N_e will generally be $<N$ in real populations, even without invoking extreme variance in reproductive success (Frankham, 1995; Palstra & Fraser, 2012; Fig. 3). If, for example, $N_e = 10^6$ and $N = 10^7$ (so that true $N_e/N = 0.1$), then the $\sim 50\%$ of estimates of N_e in Fig. 5(c) that fall between 10^2 and 10^4 would produce estimates of N_e/N in the range 10^{-3} – 10^{-5} rather than 10^{-2} – 10^{-4} .

In summary, genetic methods for estimating contemporary N_e are sensitive to a signal proportional to $1/N_e$, which is very small for populations with large true effective sizes. As a consequence, when true N_e is large and only moderate amounts of data are available, the distribution of \hat{N}_e becomes strongly bimodal, with most estimates either infinitely large or in the hundreds to low thousands. This distribution of finite \hat{N}_e estimates does not appreciably change regardless how large N_e gets. This means that when true N_e is $\geq 10^6$ and true N_e/N is ~ 0.1 , a large fraction (perhaps up to 50%) of the point estimates can be at least three to five orders of magnitude too small. The larger the population, the greater the scope for downward bias in the subset of finite \hat{N}_e values, even if the overall method is not biased. Using much larger samples (thousands

of individuals) and paying close attention to the upper bounds of confidence intervals can help to mitigate (but not eliminate) this problem.

DISCUSSION

Accounting for iteroparity and overlapping generations shows that plausible patterns of change in age-specific fecundity cannot by themselves produce tiny N_e/N ratios ($N_e/N \leq 10^{-3}$); these are only possible if individuals of the same age and sex have greatly over-dispersed variance in reproductive success ($\Phi_x \approx 10^3$ or higher). This in turn requires that only a small fraction ($N_p/N_x \approx 1/\Phi_x$) of potential parents within an age group is successful at producing offspring in any given year. This is similar to the conclusion reached by others who have evaluated the SRS hypothesis using simpler discrete-generation models, although opportunities for tiny N_e/N are reduced by about one order of magnitude if different randomly chosen parents are sweepstakes winners each year.

To date, little published information has been available regarding performance of estimators of contemporary N_e when effective size is very large. The largest effective sizes evaluated in some key papers were 100 (Nei & Tajima, 1981), 200 (Wang, 2009), 500 (Waples, 1989) and 1000 (Wang & Whitlock, 2003). Waples & Do (2010) and Gilbert & Whitlock (2015) evaluated N_e up to 5000 and Ovenden *et al.* (2007) and MacBeth *et al.* (2013) simulated some scenarios with $N_e = 10^4$, but this is still orders of magnitude smaller than effective sizes that might characterize large marine populations. This range was expanded to arbitrarily large populations in the present study. Simulated data using samples of individuals and genetic markers comparable to those used in most published estimates of tiny N_e/N demonstrate that, when true N_e is large ($10^5 - 10^6$ or higher), the distribution of \hat{N}_e is strongly bimodal, with roughly half of the estimates being infinitely large and most of the remainder being several orders of magnitude smaller than true N_e . If true N_e is large, genetic estimators have a characteristic sweet spot where almost all of the finite estimates land. For $S = 50$ and 100 diallelic loci, the range of this sweet spot is the hundreds to low thousands, and this range does not change appreciably no matter how large true N_e is (Table II and Fig. S1, Supporting Information). Thus, this sweet spot also represents a blind spot with respect to the true N_e . Huge increases in sample size shift the sweet spot to higher \hat{N}_e values (up to $10^4 - 10^6$ for $S = 5000$), but again this distribution of \hat{N}_e remains static regardless how large true N_e is (Table II).

These results mean that tiny, genetically based point estimates of N_e/N in large marine populations are expected to be quite common, even when the true N_e/N ratio is normal (c. 0.1 or higher). Notably, this pattern of spuriously low estimates of N_e/N agrees almost exactly with the three characteristics Hauser & Carvalho (2008) identified as typical of empirical estimates for marine species: first, most point estimates are in the hundreds or low thousands; second, resulting estimates of N_e are two to five orders of magnitude lower than estimates of N ; third, the estimated N_e/N ratio decreases as population size increases (Fig. S1, Supporting Information). Of course, this does not mean that any particular published estimate of a tiny N_e/N is wrong. The fact that false positives for tiny N_e/N are expected to be quite common when true N_e is large, however, argues for considerable caution in interpreting genetically based estimates for large populations.

Results presented here have several practical implications for evaluation of estimates of N_e/N that are unusually small. First, this scenario is likely to produce publication bias toward small estimates, as noted by Hauser & Carvalho (2008), but the problem is not simply that small estimates of N_e/N are interesting and provocative, while normal estimates with $\hat{N}_e \approx \hat{N}$ are boring. Because \hat{N}_e is strongly bimodal for large populations, almost no estimates will produce a value of \hat{N}_e close to the estimated census size; they will nearly all be much too small or infinite. Often, researchers are reluctant to report infinite point estimates because all real populations must be finite. These infinite point estimates, however, contain useful information; they are best interpreted as very large, and they will be rare if true N_e is small and even moderate amounts of data are available [Fig. 5(a)].

Because a large fraction of false-positives for low N_e/N can be expected when using genetic methods to estimate effective size in large marine populations, careful attention must be paid to the upper bounds of C.I.s and the full range of N_e/N that are consistent with the point estimate. As noted above, however, the C.I.s around genetically based estimates of effective size will seldom include all the uncertainty associated with the estimates. In particular, because the genetic signal is proportional to $1/N_e$, C.I.s are non-linear and asymmetrical, being much larger on the high side of the point estimate. One consequence of this is that any downward bias in the point estimate will have a disproportionately large effect on the upper bound of the C.I., which could (falsely) rule out the possibility that true N_e is very large. Gradually, some of the potential biases associated with violations of simplistic assumptions of genetic methods for estimating contemporary N_e are being critically evaluated (Waples & Yokota, 2007; Waples & England, 2011; Neel *et al.*, 2013; Waples *et al.*, 2014; Gilbert & Whitlock, 2015; Wang, 2016) and it will be important to carefully consider these results in interpreting low estimates of N_e/N .

Analysis of the simulated data used the moment-based temporal and LD methods, which are easy to calculate and which have been among the most commonly used methods to generate tiny estimates of N_e/N (Hauser & Carvalho, 2008; Hedgecock & Pudovkin, 2011). Likelihood-based or approximate-Bayesian-computation (ABC) methods (Wang & Whitlock, 2003; Tallmon *et al.*, 2008) have the potential to reduce biases and increase precision, but they have not been rigorously evaluated with very large populations; furthermore, most require one to specify an upper limit to N_e , which can be problematical for very large populations. Although the combined estimates computed here as a simple, unweighted harmonic mean of \hat{N}_e for the temporal and LD methods did not perform much better than the LD method alone (Fig. 5 and Fig. S1, Supporting Information), this idea merits more detailed study, as the two estimators are largely uncorrelated for small to moderate N_e (Waples, in press) and hence should provide largely independent information about effective size. The most robust results can be expected when multiple methods produce comparable estimates (Hauser *et al.*, 2002). Unfortunately, results for parentage analyses based on the known pedigree show that, even if parents of each offspring can be assigned with complete certainty, relatively small samples from very large populations will not provide much useful information about N_e unless effective size is very small compared with N . This means that the single-sample sibship method (Wang, 2009) is not likely to be useful for evaluating populations with large N_e , although it could help confirm events of SRS in which the individuals sampled could include many siblings.

Increasing the number of genetic markers can help improve precision, but only to a certain degree (Table II and Fig. S2, Supporting Information). An important caveat applies to evaluating the benefits of using many thousands of genetic markers, as it is now possible to obtain even for non-model species. In a computer it is easy to simulate arbitrarily large numbers of loci that assort independently and hence provide non-redundant information (up to 1000 such loci were modelled here). In the real world, however, loci must be packaged into a relatively small number of chromosomes, and physical linkage creates dependencies among the markers that reduce the overall information content. This means that increasing the number of loci by a factor of 10 [as in Fig. S2(a), Supporting Information] will not increase precision by the same proportion. This is an important topic that merits more rigorous study, but preliminary results (Jones *et al.*, 2016; Waples *et al.*, 2016) indicate that, for the LD method, the effective number of loci (in terms of information content) can be much lower than the actual number. Therefore, although using large numbers of SNP markers will increase precision (Hoffman *et al.*, 2014), this by itself is unlikely to solve all problems associated with estimating N_e in large populations. A rigorous evaluation of this issue will require conducting simulations using linked markers in populations with very large N_e .

Empirical studies to test the hypothesis of tiny N_e/N are tricky, for several reasons apart from statistical behaviour of the genetic estimators. Results presented above demonstrate that, after accounting for life-history traits typical of many marine species (iteroparity and long adult lifespan), it remains that case that tiny N_e/N require some type of SRS, whereby only a small fraction of adults successfully reproduce in any given season. The converse is not necessarily true. Not all types of sweepstakes reproduction produce tiny N_e/N , it depends on the spatial and temporal scale on which SRS occurs. In some cases, SRS can produce chaotic genetic patchiness (Johnson & Black, 1982; Broquet *et al.*, 2013) without permanent population structure or small overall N_e . Selkoe *et al.* (2006) and Buston *et al.* (2009) suggested that any effects of SRS are likely to be ephemeral and disappear when individual cohorts are integrated into the population as a whole. Whether this is true, however, depends on the nature of the SRS. As demonstrated above, if entire cohorts of a long-lived species are consistently produced by SRS (*e.g.* if Φ_x is consistently very high for all ages of adults), then N_e/N can be very small ($N_e/N \approx 1/\Phi$; equation 10).

The concept of N_e applies most directly to a full generation in a single, completely isolated population. Life histories of many marine species pose a major challenge in this regard. Many marine fishes (and some marine invertebrates) are highly vagile as adults and many have long larval stages that provide opportunities for dispersal. As a consequence, marine populations are often ill-defined spatially and in at least some cases better fit one or two-dimensional isolation-by-distance models with continuous distributions than they do models with semi-discrete subpopulations. If one wants to draw inferences about N_e and N_e/N for a metapopulation rather than a single isolated population, then spatially varying productivity could affect the result (as proposed by Turner *et al.*, 2002), but that scenario is beyond the scope of this paper to consider in detail.

Hedgcock (1994) proposed some tests of the sweepstakes-reproduction hypothesis, and these have been discussed by many subsequent authors (Selkoe *et al.*, 2006; Hauser & Carvalho, 2008; Taris *et al.*, 2009). One of the predictions is that genetic diversity within larval cohorts should be reduced (and LD increased) compared with adults. As subsequently noted by Hedgcock & Pudovkin (2011), this qualitative prediction is

tricky to evaluate. Robust tests must await a quantitative treatment that fully accounts for all the sampling issues involved and the different expectations for different measures of genetic diversity (*e.g.* number of alleles *v.* heterozygosity). Another of Hedgecock's (1994) predictions, that genetic differences among cohorts should be large compared with samples from the adult population, has more direct relevance to evaluation of N_e/N ratios. In iteroparous species, if effective size actually is relatively small, then the magnitude of allele frequency differences among cohorts can be used to estimate N_e , using the method of Jorde & Ryman (1995) and extended by Jorde (2012).

The most robust way to quantitatively evaluate the hypothesis that SRS leads to tiny N_e/N ratios is to implement a sampling programme that combines both spatial and temporal replication (akin to the Lagrangian and Eulerian frames of reference discussed by Hedgecock & Pudovkin, 2011). The spatial scale should be broad enough to identify population boundaries (if they exist) and to account for effects of immigration. Providing the appropriate temporal dimension to the data collected is likely to be more challenging, because at least two temporal components must be considered. First, many marine species spawn over extended periods of time (and tropical species can spawn throughout the year), so it is important to have a way to integrate all reproduction events across each season. Second, the strongest evidence for tiny N_e/N attributable to sweepstakes reproduction will be to demonstrate temporal stability of geographic patterns across multiple generations. If these patterns are not dynamically stable, the signal could be one that reflects ephemeral patterns of reproduction of local groups of parents (chaotic genetic patchiness) rather than small effective size of the entire population across a generation. Evaluating this will be difficult in long-lived species unless historic samples (*e.g.* archived scales) are available.

Finally, point estimates of N_e and N_e/N for marine species can be evaluated in the context of the species' life history and other genetic analyses, as suggested by Hedgecock & Pudovkin (2011). N_e/N cannot be tiny unless individuals are capable of producing at least thousands of offspring that survive to age at maturity, so tiny estimates for species with low fecundity would be unlikely to be valid. If N_e of a large marine species really is in the hundreds to low thousands and this pattern has persisted through time, it should be reflected in low overall genetic diversity, shallow coalescent structure and a star phylogeny of DNA sequences.

I appreciate the invitation to prepare this manuscript for the symposium edition of the Journal. I thank S. Sogard for useful information about BOFFFFs and M. Hare, L. Hauser and two anonymous reviewers for useful comments on an earlier draft.

Supporting Information

Supporting Information may be found in the online version of this paper:

FIG. S1. Changes in the frequency distribution of estimates of effective population size (\hat{N}_e) for simulated data as true effective population size (N_e) increases: (a) true $N_e = 10^4$ (†), (b) true $N_e = 10^5$ and (c) true $N_e = 10^6$. ■■■, The standard temporal method, computed for samples taken three generations apart; ■■■, the linkage disequilibrium method; ■■■, the harmonic mean of the two estimates; 'inf' = infinity. Each set of 500 replicate populations used 100 diallelic (single nucleotide polymorphism, SNP) loci and $S = 50$ individuals sampled. Note that the distribution of finite estimates remains essentially unchanged as true N_e increases.

FIG. S2. Frequency distribution of estimates of effective population size (\hat{N}_e) for simulated data using the linkage disequilibrium method. (a) Results for 100 (■) and 1000 (▨) diallelic (single nucleotide polymorphism, SNP) loci, $S = 50$ individuals sampled and true $N_e = 10^5$ (†) and (b) results for $S = 50$ (■) and 200 (▨), with 100 SNP loci and true $N_e = 10^6$. ‘inf’ = infinity.

References

- Aarestrup, K., Økland, F., Hansen, M. M., Righton, D., Gargan, P., Castonguay, M., Bernatchez, L., Howey, P., Sparholt, H., Pedersen, M. I. & McKinley, R. S. (2009). Oceanic spawning migration of the European eel (*Anguilla anguilla*). *Science* **325**, 1660.
- Bekkevold, D., Hansen, M. M. & Loeschcke, V. (2002). Male reproductive competition in spawning aggregations of cod (*Gadus morhua* L.). *Molecular Ecology* **11**, 91–102.
- Berkeley, S. A., Chapman, C. & Sogard, S. M. (2004). Maternal age as a determinant of larval growth and survival in a marine fish, *Sebastes melanops*. *Ecology* **85**, 1258–1264.
- Bierne, N., Launey, S., Naciri-Graven, Y. & Bonhomme, F. (1998). Early effect of inbreeding as revealed by microsatellite analyses on *Ostrea edulis* larvae. *Genetics* **148**, 1893–1906.
- Broquet, T., Viard, F. & Yearsley, J. M. (2013). Genetic drift and collective dispersal can result in chaotic genetic patchiness. *Evolution* **67**, 1660–1675.
- Burton, R. S. & Feldman, M. W. (1981). Population genetics of *Tigriopus californicus*. II. Differentiation among neighboring populations. *Evolution* **35**, 1192–1205.
- Buston, P. M., Fauvelot, C., Wong, M. Y. L. & Planes, S. (2009). Genetic relatedness in groups of the humbug damselfish *Dascyllus aruanus*: small, similar-sized individuals may be close kin. *Molecular Ecology* **18**, 4707–4715.
- Christie, M. R., Johnson, D. W., Stallings, C. D. & Hixon, M. A. (2010). Self-recruitment and sweepstakes reproduction amid extensive gene flow in a coral-reef fish. *Molecular Ecology* **19**, 1042–1057.
- Cross, T. F. & Payne, R. H. (1978). Geographic variation in Atlantic cod, *Gadus morhua*, off eastern North America: a biochemical systematics approach. *Journal of the Fisheries Research Board of Canada* **35**, 117–123.
- DiMichele, L. & Powers, D. A. (1982). Physiological basis for swimming endurance differences between LDH-B genotypes of *Fundulus heteroclitus*. *Science* **216**, 1014–1016.
- Felsenstein, J. (1971). Inbreeding and variance effective numbers in populations with overlapping generations. *Genetics* **68**, 581–597.
- Frankham, R. (1995). Effective population size/adult population size in wildlife: a review. *Genetical Research* **66**, 95–107.
- Frydenberg, O. V. E., Møller, D. A. G., Nævdal, G. & Sick, K. (1965). Haemoglobin polymorphism in Norwegian cod populations. *Hereditas* **53**, 257–271.
- Gaffney, P. M. (1994). Heterosis and heterozygote deficiencies in marine bivalves: more light? In *Genetics and Evolution of Aquatic Organisms* (Beaumont, A., ed), pp. 146–153. London: Chapman & Hall.
- Gilbert, K. J. & Whitlock, M. C. (2015). Evaluating methods for estimating local effective population size with and without migration. *Evolution* **69**, 2154–2166.
- Grant, W. S. & Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity* **89**, 415–426.
- Hansen, M. M., Nielsen, E. E. & Mensberg, K.-L. D. (1997). The problem of sampling families rather than populations: relatedness among individuals in samples of juvenile brown trout *Salmo trutta* L. *Molecular Ecology* **6**, 469–474.
- Harris, H. (1966). Enzyme polymorphisms in man. *Proceedings of the Royal Society B* **164**, 298–310.
- Hauser, L. & Carvalho, G. R. (2008). Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries* **9**, 333–362.
- Hauser, L., Adcock, G. J., Smith, P. J., Ramirez, J. H. B. & Carvalho, G. R. (2002). Loss of microsatellite diversity and low effective population size in an overexploited population

- of New Zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences, U.S.A.* **99**, 11742–11747.
- Hedgcock, D. (1994). Does variance in reproductive success limit effective population size of marine organisms? In *Genetics and Evolution of Aquatic Organisms* (Beaumont, A., ed), pp. 122–134. London: Chapman & Hall.
- Hedgcock, D. & Pudovkin, A. H. (2011). Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. *Bulletin of Marine Science* **87**, 971–1002.
- Hedrick, P. (2005). Large variance in reproductive success and the N_e/N ratio. *Evolution* **59**, 1596–1599.
- Hill, W. G. (1972). Effective size of population with overlapping generations. *Theoretical Population Biology* **3**, 278–289.
- Hill, W. G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetical Research* **38**, 209–216.
- Hixon, M. A., Johnson, D. W. & Sogard, S. M. (2014). BOFFFFs: on the importance of conserving old-growth age structure in fishery populations. *ICES Journal of Marine Science* **71**, 2171–2185.
- Hjort, J. (1914). Fluctuations in the great fisheries of northern Europe viewed in the light of biological research. *Rapport et Procès-Verbaux des Réunions du Conseil international pour l'Exploration de la Mer* **20**, 1–228.
- Hoffman, J. I., Simpson, F., David, P., Rijks, J. M., Kuiken, T., Thorne, M. A., Lacy, R. C. & Dasmahapatra, K. K. (2014). High-throughput sequencing reveals inbreeding depression in a natural population. *Proceedings of the National Academy of Sciences* **111**, 3775–3780.
- Hubby, J. L. & Lewontin, R. C. (1966). A molecular approach to the study of genetic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. *Genetics* **54**, 577–594.
- Johnson, M. S. & Black, R. (1982). Chaotic genetic patchiness in an intertidal limpet, *Siphonaria* sp. *Marine Biology* **70**, 157–164.
- Johnson, M. S. & Black, R. (1984). Pattern beneath the chaos: the effect of recruitment on genetic patchiness in an intertidal limpet. *Evolution* **38**, 1371–1383.
- Jones, A. T., Ovenden, J. R. & Wang, Y.-G. (2016). Improved confidence intervals for the linkage disequilibrium method for estimating effective population size. *Heredity* **117**, 217–223.
- Jorde, P. E. (2012). Allele frequency covariance among cohorts and its use in estimating effective size of age-structured populations. *Molecular Ecology Resources* **12**, 476–480.
- Jorde, P. E. & Ryman, N. (1995). Temporal allele frequency change and estimation of effective size in populations with overlapping generations. *Genetics* **139**, 1077–1090.
- Koehn, R. K. & Williams, G. C. (1978). Genetic differentiation without isolation in the American eel, *Anguilla rostrata*. II. Temporal stability of geographic patterns. *Evolution* **27**, 192–204.
- Koehn, R. K., Milkman, R. & Mitton, J. B. (1976). Population genetics of marine pelecypods. IV. Selection, migration and genetic differentiation in the blue mussel *Mytilus edulis*. *Evolution* **30**, 2–32.
- Kuparinen, A., Hutchings, J. & Waples, R. S. (2016). Harvest-induced evolution and effective population size. *Evolutionary Applications* **9**, 658–672.
- Lee, A. M., Engen, S. & Saether, B.-E. (2011). The influence of persistent individual differences and age at maturity on effective population size. *Proceedings of the Royal Society B* **278**, 3303–3312.
- Macbeth, G. M., Broderick, D., Buckworth, R. C. & Ovenden, J. R. (2013). Linkage disequilibrium estimation of effective population size with immigrants from divergent populations: a case study on Spanish mackerel (*Scomberomorus commerson*). *G3: Genes | Genomes | Genetics* **3**, 709–717.
- Neel, M. C., McKelvey, K. S., Ryman, N., Lloyd, M. W., Short Bull, R., Allendorf, F. W., Schwartz, M. K. & Waples, R. S. (2013). Estimation of effective population size in continuously distributed populations: there goes the neighborhood. *Heredity* **111**, 189–199.
- Nei, M. & Tajima, F. (1981). Genetic drift and estimation of effective population size. *Genetics* **98**, 625–640.

- Nunney, L. (1993). The influence of mating system and overlapping generations on effective population size. *Evolution* **47**, 1329–1341.
- Nunney, L. & Elam, D. R. (1994). Estimating the effective population size of conserved populations. *Conservation Biology* **8**, 175–184.
- Ovenden, J., Peel, D., Street, R., Courtney, A. J., Hoyle, S. D., Peel, S. L. & Podlich, H. (2007). The genetic effective and adult census size of an Australian population of tiger prawns (*Penaeus esculentus*). *Molecular Ecology* **16**, 127–138.
- Palstra, F. P. & Fraser, D. J. (2012). Effective/census population size ratio estimation: a compendium and appraisal. *Ecology and Evolution* **2**, 2357–2365.
- Palumbi, S. R. (1994). Genetic divergence, reproductive isolation and marine speciation. *Annual Review of Ecology and Systematics* **25**, 547–572.
- Ryman, N. & Utter, F. (1987). *Population Genetics and Fishery Management*. Seattle, WA: University of Washington Press.
- Selkoe, K. A., Gaines, S. D., Caselle, J. E. & Warner, R. R. (2006). Current shifts and kin aggregation explain genetic patchiness in fish recruits. *Ecology* **87**, 3082–3094.
- Shaw, A. K. & Levin, S. A. (2013). The evolution of intermittent breeding. *Journal of Mathematical Biology* **66**, 685–703.
- Strathmann, R. (1978). Length of pelagic period in echinoderms with feeding larvae from the Northeast Pacific. *Journal of Experimental Marine Biology and Ecology* **34**, 23–27.
- Tallmon, D. A., Koyuk, A., Luikart, G. & Beaumont, M. A. (2008). OneSamp: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources* **8**, 299–301.
- Taris, N., Boudry, P., Bonhomme, F., Camara, M. D. & Lapegue, S. (2009). Mitochondrial and nuclear DNA analysis of genetic heterogeneity among recruitment cohorts of the European flat oyster *Ostrea edulis*. *Biological Bulletin* **217**, 233–241.
- Tracey, M. L., Bellet, N. F. & Gravem, C. D. (1975). Excess allozyme homozygosity and breeding population structure in mussel *Mytilus californianus*. *Marine Biology* **32**, 303–311.
- Turner, T. F., Wares, J. P. & Gold, J. R. (2002). Genetic effective size is three orders of magnitude smaller than adult census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). *Genetics* **162**, 1329–1339.
- Wang, J. (2009). A new method for estimating effective population size from a single sample of multilocus genotypes. *Molecular Ecology* **18**, 2148–2164.
- Wang, J. (2016). A comparison of single-sample estimators of effective population sizes from genetic marker data. *Molecular Ecology*. doi: 10.1111/mec.13725
- Wang, J. & Whitlock, M. C. (2003). Estimating effective population size and migration rates from genetic samples over space and time. *Genetics* **163**, 429–446.
- Waples, R. K., Larson, W. A. & Waples, R. S. (2016). Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. *Heredity* **117**, 233–240.
- 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. (2005). Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Molecular Ecology* **14**, 3335–3352.
- Waples, R. S. (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics* **7**, 167–184.
- Waples, R. S. & Antao, T. (2014). Intermittent breeding and constraints on litter size: consequences for effective population size per generation (N_e) and per reproductive cycle (N_b). *Evolution* **68**, 1722–1734.
- Waples, R. S. & Do, C. (2010). Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications* **3**, 244–262.
- Waples, R. S. & England, P. R. (2011). Estimating contemporary effective population size based on linkage disequilibrium in the face of migration. *Genetics* **189**, 633–644.
- Waples, R. S. & Waples, R. K. (2011). Inbreeding effective population size and parentage analysis without parents. *Molecular Ecology Resources* **11**(Suppl. 1), 162–171.

- Waples, R. S. & Yokota, M. (2007). Temporal estimates of effective population size in species with overlapping generations. *Genetics* **175**, 219–233.
- Waples, R. S., Dickhoff, W. W., Hauser, L. & Ryman, N. (2008). Six decades of fishery genetics: taking stock. *Fisheries* **33**, 76–79.
- Waples, R. S., Do, C. & Choquet, J. (2011). Calculating N_e and N_e/N in age-structured populations: a hybrid Felsenstein-Hill approach. *Ecology* **92**, 1513–1522.
- Waples, R. S., Luikart, G., Faulkner, J. R. & Tallmon, D. A. (2013). Simple life history traits explain key effective population size across diverse taxa. *Proceedings of the Royal Society B* **280**, 20131339. doi: 10.1098/rspb.2013.1339
- Waples, R. S., Antao, T. & Luikart, G. (2014). Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics* **197**, 769–780.