Genetic Methods for Estimating the Effective Size of Cetacean Populations

Robin S. Waples

Northwest Fisheries Center, National Marine Fisheries Service, 2725 Montlake Boulevard East, Seattle, Washington 98112, USA

ABSTRACT

Some indirect (genetic) methods for estimating effective population size (N_e) are evaluated for their suitability in studying cetacean populations. The methods can be grouped into those that (1) estimate current N_e, (2) estimate long-term N_e and (3) provide information about recent genetic bottlenecks. The methods that estimate current effective size are best suited for the analysis of small populations, and nonrandom sampling and population subdivision are probably the most serious sources of potential bias. Methods that estimate long-term N_e are best suited to the analysis of large populations or entire species, may be more strongly influenced by natural selection and depend on accurate estimates of mutation or DNA base substitution rates. Precision of the estimates of N_e is likely to be a limiting factor in many applications of the indirect methods.

Keywords: genetics; assessment; cetaceans – general; evolution.

INTRODUCTION

Population size is one of the most important factors that determine the rate of various evolutionary processes, and it appears as a parameter in many of the fundamental equations of population genetics. However, knowledge merely of the total number of individuals (N) in a population is not sufficient for an accurate description of these evolutionary processes. Because of the influence of demographic parameters, two populations of the same total size may experience very different rates of genetic change. Wright (1931; 1938) developed the concept of effective population size (N_e) as a way of summarising relevant demographic information so that one can predict the evolutionary consequences of finite population size (see Fig. 1). For those interested in biological conservation, N_e is important chiefly because it determines the rate of loss of genetic variability and the rate of increase in inbreeding in a population.

N_e is defined as the size of an ideal population that experiences genetic change at the same rate as the population under consideration. In an ideal population, the sex ratio is equal and the lifetime variance in the number of offspring (V_k) is binomial; if population size is constant, this variance is equal to the mean number of offspring produced per individual (i.e. $V_k = \bar{k} = 2$). Most natural populations depart from the ideal in that $V_k > 2$, and in many cases the sex ratio of breeders is uneven as well. Both factors cause the effective size to be smaller than the census number of a population.

In some cases, calculation of the effective population number is straightforward given the necessary demographic information (see Crow and Denniston, 1988). The problem, of course, is that the relevant demographic parameters are notoriously difficult to measure in natural populations. Although sex ratios can be observed for many organisms at some life-history stage, the sex ratio of those that successfully breed is often not so easy to determine. The factor with potentially the greatest influence on N_e is V_k, and even in well-studied populations estimating this parameter is difficult. To compute V_k, it is not

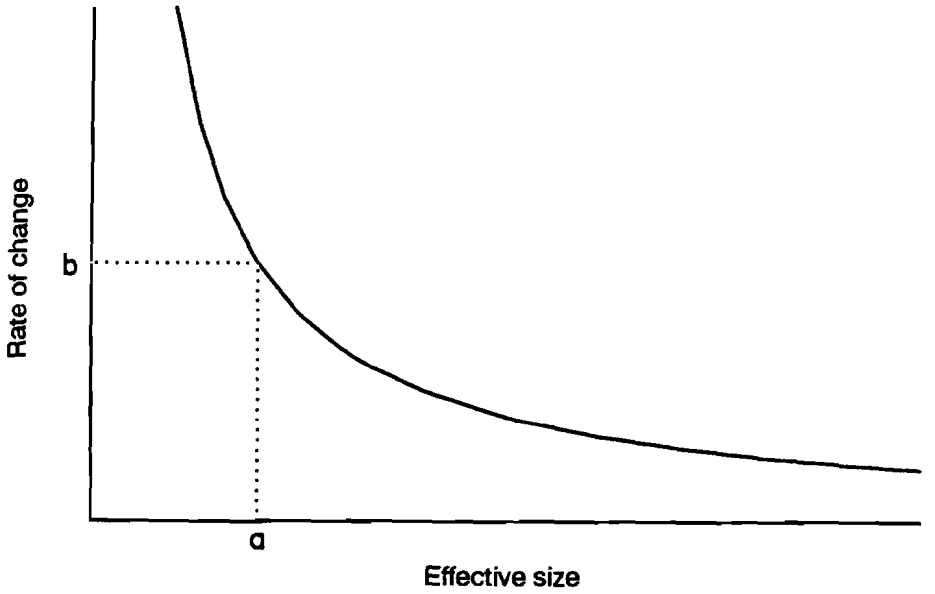

Fig. 1. Schematic representation of the relationship between effective size (N_e) and the rate of genetic change in a population. Many genetic processes caused by genetic drift (e.g. the rate of allele frequency change or the rate of loss of genetic variability) are inversely proportional to N_e. If one knows the effective size of a population (size a in the example shown), it is possible to predict the expected rate of change (b). Conversely, if one can measure the rate of change and assume that it is caused by genetic drift, then the theoretical relationship provides a means of estimating N_e. Other indirect methods for estimating N_e focus on different genetic parameters, but the approach is similar. Note that methods (such as the temporal method) that measure processes dominated by terms in $1/N_e$ are best suited to the study of small populations. If population size is large (flat part of the curve), substantial changes in N_e have little effect on the rate of change.

sufficient merely to know the lifetime variance in number of offspring produced; the necessary data are the lifetime variance in the number of offspring that survive to reproduce in the next generation. Such information is difficult to obtain even for highly-visible terrestrial organisms; it may be impossible to measure accurately in cetaceans, that spend most of their time underwater and whose populations may extend across entire oceans (Hammond, 1987).

The difficulties inherent in obtaining reliable data for the direct computation of N_e has fostered a variety of attempts to estimate N_e by indirect methods. The logic behind these approaches is simple, as illustrated schematically in Fig. 1: if effective population size is a key factor driving various evolutionary processes, quantifying the magnitude or rate of these processes should make it possible to estimate N_e. These indirect methods may also be called genetic methods because they measure the genetic consequences of finite population size.

The purpose of this paper is to examine various indirect methods that have been proposed and evaluate their usefulness for the analysis of cetacean populations. The indirect methods differ in the type of genetic and demographic data used, the assumptions on which they are based and their sensitivity. They can be grouped into methods that (1) estimate current N_e, (2) estimate long-term N_e and (3) provide information about recent

genetic bottlenecks. The following descriptions are necessarily brief and the reader is referred to the original sources for more details about the methods.

METHODS THAT ESTIMATE CURRENT N_e

Temporal changes in allele frequency

Approach

Perhaps the first application of an indirect method was the attempt of Krimbas and Tsakas (1971) to estimate N_e in the olive fruit fly *(Dacus oleae)* from the magnitude of allele frequency change over time. The rationale for the temporal method is that a measure of allele frequency change (F) is a function of N_e and elapsed time in generations (t). For small t ($t<<2N_e$) and assuming that the changes are due to drift, the following approximation is useful:

$$E(F) \approx t/(2N_e) \quad (1)$$

F has been called the standardised variance of allele frequency change because its numerator (the squared difference in allele frequency in two temporally-spaced samples) is like a variance and the effects of initial allele frequency are compensated for in the denominator. As seen in (1), F has obvious potential as a means of estimating N_e. However, F as defined above is a population parameter, and in practice sample (rather than population) allele frequencies are generally available. Therefore, F must be estimated by $\hat{F}$, which is also affected by random error in computing sample allele frequencies. Methods for computing $\hat{F}$ are discussed in the Appendix.

Under certain sampling schemes (plan II of Nei and Tajima, 1981; Pollak, 1983; Waples, 1989), the approximate expectation of $\hat{F}$ is given by

$$\begin{aligned} E(\hat{F}) &\approx t/(2N_e) + 1/(2S_0) + 1/(2S_t) \\ &= t/(2N_e) + 1/(\tilde{S}), \end{aligned} \quad (2)$$

where S_0 and S_t are the number of individuals sampled in generations 0 and t and $\tilde{S}$ is the harmonic mean of S_0 and S_t. Rearrangement of (2) leads to the following estimator of N_e:

$$\hat{N}_e = t/[2(\hat{F}-1/\tilde{S})] \quad (3)$$

In the above, $1/\tilde{S}$ represents the expected contribution to $\hat{F}$ caused by random error in drawing the samples. The term $(\hat{F} - 1/\tilde{S})$ in (3) can thus be interpreted as an adjusted value of $\hat{F}$, or that portion of $\hat{F}$ attributable to genetic drift.

Application

The temporal method is designed to estimate the effective size of a closed population or sub-population. Important assumptions are that (1) mutation is unimportant, (2) the alleles considered are selectively neutral and not in linkage disequilibrium with other loci subject to selection, (3) the samples for genetic analysis are randomly drawn from the population as a whole and (4) there is no immigration from neighbouring sub-populations. Provided these assumptions are met, estimates of N_e computed from (3) will have little bias (Waples, 1989).

Given the time scale available for most research projects, it is probably safe to assume that mutation has little effect on $\hat{F}$. The assumption of strict neutrality for all alleles is unlikely to be met, but a good deal of evidence suggests that the bulk of electrophoretically detectable variation is under at most very weak selection. Furthermore, selection of constant intensity has little effect on $\hat{F}$ unless N_e is very large or a large number of generations are considered (Pollak, 1983; Mueller *et al.*, 1985).

Nevertheless, it is prudent to exclude data for loci that may be influenced by selection (or those closely linked to selected loci). The test proposed by Lewontin and Krakauer (1973), which compares the variance of single locus $\hat{F}$ values with the variance expected under neutrality, can help to identify loci for which allele frequency change is larger than can reasonably be attributed to drift. Although several authors have pointed out problems with the Lewontin and Krakauer test as applied to F_{ST} values for geographic populations, its use with temporal data for a single population is appropriate (Lewontin and Krakauer, 1975; Gaines and Whittam, 1980; Kimura, 1983).

Failure to meet either the third or fourth assumption can seriously bias $\hat{F}$ (and hence $\hat{N}_e$). For many cetacean species, the lack of detailed biological information may make it difficult to determine whether these assumptions have been met. For example, in a truly random sample, each individual has an equal opportunity of being included. This assumption is violated to the extent that certain ages or sexes are more likely than others to appear in the sample. Sampling efforts that concentrate on selected pods or groups of cetaceans will also fail to accurately reflect attributes of the population as a whole. In general, these types of nonrandom sampling should lead to $\hat{F}$ values that are biased upwards, causing the estimate of N_e to be too low.

Population subdivision can also create problems for use of the temporal method, in two different ways. First, if the subdivisions are not easily recognised, as might be the case with some cetaceans (Hammond, 1987), samples taken at different times may include different proportions of the various sub-populations. In this case, $\hat{F}$ will measure interpopulation differences as well as genetic changes within a population, with a resulting tendency to underestimate N_e. Second, even if sampling is confined to a single sub-population, the estimate of N_e can still be biased if migration from another sub-population with different allele frequencies occurs between the two samples.

If the above assumptions are met, then the primary factor determining the usefulness of the temporal method is precision of the estimate $\hat{N}_e$. Precision is an important consideration because the variance associated with single-locus $\hat{F}$ values is typically large. Fortunately, the distribution of $\hat{F}$ is known to closely approximate the chi square distribution (Lewontin and Krakauer, 1973; Nei and Tajima, 1981; Waples, 1989), so it is easy to evaluate precision of $\hat{F}$ (or $\hat{N}_e$) values computed from actual data. The number of degrees of freedom associated with the estimate $\hat{F}$ equals the total number of independent alleles (n) sampled over all loci [$n = \Sigma(L_i-1)$, where L_i is the number of alleles at the i^{th} locus]. As n increases, confidence limits for $\hat{F}$ narrow, and these translate into narrower confidence limits for $\hat{N}_e$ (Waples, 1989; see Appendix). Note that the number of loci used does not necessarily have to be large if very polymorphic systems are available.

Precision also depends on sample size and elapsed time between samples. Simulation results show that in general, estimates with reasonable precision cannot be expected unless at least 50 to 100 individuals are sampled each time period and the number of independent alleles surveyed is 10 to 20 or more. However, Waples (1989) showed that increasing n, S, or t by the same proportion leads to a similar increase in precision, so some tradeoffs in experimental design are possible. For example, a shortage of polymorphic loci can be compensated for in part by increasing the number of individuals sampled.

Even if the samples of individuals and alleles are adequate, however, the temporal method generally will not produce a precise estimate of N_e unless several generations elapse between samples (Nei and Tajima, 1981; Waples, 1989). There is thus a difficulty inherent in applying the temporal method to long-lived organisms such as cetaceans. For many cetaceans, a sampling interval of 5 to 10 years (a long time for most research projects) would span only part of a generation. A fractional value of t might be used in (3) to estimate N_e in such cases, but the estimate would be biased unless the age structure

were stable. Furthermore, the variance of the point estimate would be so large that it would provide little useful information, unless a very large number of alleles could be sampled.

Fig. 2 illustrates the importance to the temporal method of sampling several generations of genetic drift. The contribution to $\hat{F}$ from genetic drift is approximately $1/(2N_e)$ per generation and thus increases almost linearly with time between samples, whereas the

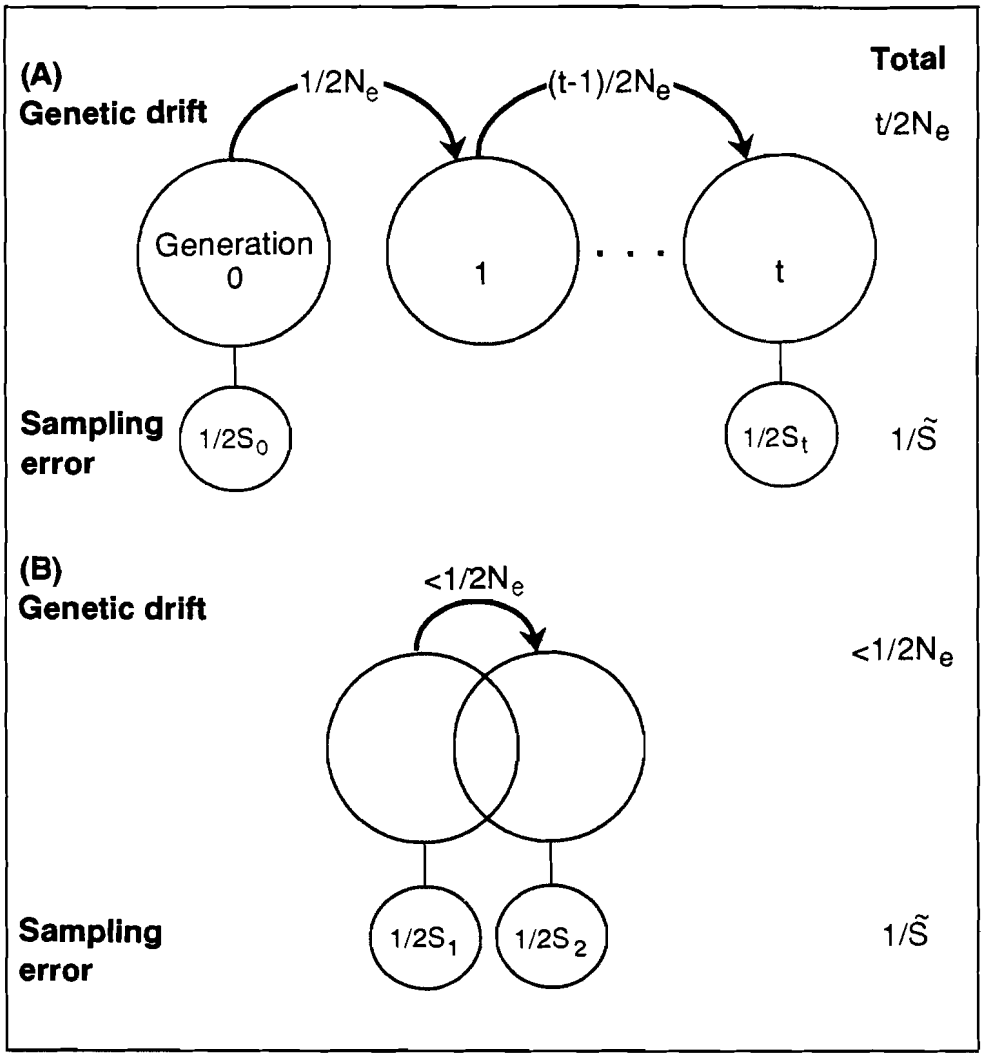

Fig. 2. The temporal method for estimating N_e from the standardised variance of allele frequency change ($\hat{F}$). Ideally, samples are well-separated in time (A) so that they measure the cumulative effects of several (t) generations of genetic drift. The total contribution to $\hat{F}$ from genetic drift $[t/(2N_e)]$ may be large relative to sampling error ($1/\tilde{S}$, where $\tilde{S}$ is the harmonic mean of S_0 and S_t). For long-lived organisms with overlapping generations (such as cetaceans), a realistic sampling interval might span only part of a generation (B), in which case the weak signal from genetic drift may be swamped by sampling error.

contribution from sampling error ($1/\tilde{S}$) does not change. Over several generations, the cumulative effects of genetic drift constitute a signal that can be strong relative to the noise created by sampling error. If sampling is at subgeneration intervals, only a portion of the population has changed and the signal due to genetic drift [$<1/(2N_e)$] is weak relative to the signal form sampling error (still $1/\tilde{S}$; see Fig. 2).

This problem can be overcome if there is another measurable sampling process that has a signal with magnitude determined by N_e. This proved to be the case with Pacific salmon, *Oncorhynchus* spp. (Waples, 1990b). In these species, the population as a whole (effective size N_e per generation, which is typically 3 to 5 years) is fragmented into discrete, semelparous (breeding only once per lifetime), yearly breeding populations (effective size N_b per year). A comparison of allele frequencies in samples taken in consecutive years does not strictly measure genetic drift in the population as a whole, but it does measure a process (selection of the N_b effective breeders each year) that is a direct function of effective size.

For organisms (such as cetaceans) with overlapping generations, a somewhat similar fragmentation of the population occurs if we consider only those individuals that produce offspring in a given year. If sampled individuals can be aged and grouped into cohorts, then the magnitude of allele frequency change among cohorts should, in theory, provide information about the effective number of breeders (N_b) producing each cohort. The expected value of F for a comparison of cohorts depends on life history features of the organism being studied. For example, many baleen whales have a two- or three-year reproductive cycle (Lockyer, 1984), which suggests that different portions of the population reproduce in consecutive years. Alternatively, essentially the same individuals may reproduce each year, or that portion of the population that breeds in any year may be determined in a random fashion. The expectation of F for these scenarios is discussed in the Appendix.

There are some potential problems, however, in the practical application of this method. First, cetacean reproductive patterns might not exactly fit any of the scenarios discussed in the Appendix. For example, although a cyclical reproductive cycle may be realistic for females of many species of baleen whales (Lockyer, 1984), this pattern seems less likely to apply to males. For a given species, it may only be possible to provide upper and lower limits for E(F), rather than an exact value. Second, it is not clear how to relate the concept of N_b (effective number per year) to the more familiar concept of N_e (effective number per generation). Waples (1990a) found that in Pacific salmon, N_e is directly proportional to N_b, but the relationship should not be so simple for cetaceans. Pacific salmon are semelparous, so reproductive contribution in any given brood year is synonymous with lifetime reproductive contribution. This is not the case with cetaceans, which may produce offspring in many different years. Iteroparity introduces correlations of allele frequencies among cohorts, making it difficult to calculate N_e from N_b. Nevertheless, information about the effective number of breeders per year may be useful, particularly if data necessary to estimate effective size by other methods are not available.

Gametic disequilibrium

Approach

Gametic disequilibrium (or linkage disequilibrium, denoted by D) is the nonrandom association of alleles at different gene loci. If we consider a pair of loci, with p the frequency of allele A at the first locus, q the frequency of allele B at the second locus and AB the gamete having both alleles A and B, then by definition $D = \text{freq}(AB) - pq$. That is, D measures the deviation of the observed frequency of AB gametes from that expected given random mating and independent assortment. If, as is often the case, gametic

frequencies cannot be directly monitored, they can be estimated by the iterative method of Hill (1974) or the simpler method of Burroughs (cited in Weir, 1979). Gametic disequilibrium can arise from a variety of factors, including physical linkage, epistatic selection, genetic hitchhiking, migration (or population admixture) and random drift in finite populations.

In the decade or so following the realisation that natural populations generally harbor a great deal of electrophoretically-detectable genetic variation, most studies of gametic disequilibrium were designed to look for epistatic interactions between individual pairs of loci. In general such efforts were not successful (Hill, 1981; Smit-McBride *et al.*, 1988), suggesting that this type of selection may not be strongly affecting these allozyme polymorphisms. Brown (1975) however, pointed out that in analysing an individual pair of loci, even moderately strong linkage disequilibrium can escape detection unless very large samples are taken.

Hill (1981) considered the potential of using data from linkage disequilibrium to estimate effective population size and came to a conclusion similar to Brown's: even fairly large samples may produce an estimate with a large standard error. This is not necessarily true, however, if data for many gene loci are available. Focusing on individual pairs of loci is appropriate if an evaluation of epistasis or linkage groups is the goal, but this practice fails to take advantage of a property of neutral loci: all multilocus associations are affected by genetic drift to the same extent and therefore provide information relevant to an estimate of effective population size. Combining data for many locus-pairs thus provides the most powerful method for estimating N_e. In evaluating the power of multilocus gametic disequilibrium analysis to detect mixtures of gene pools (Waples and Smouse, 1990), we found that the multilocus test used was very sensitive to population size. This result suggested that it might be worthwhile to reconsider the potential usefulness of the disequilibrium approach to estimating N_e.

Although Hill (1981) showed that tightly linked loci provide the greatest precision in estimating N_e, I consider here only unlinked loci because data on recombination fractions often are not available. The correlation of allele frequencies at a pair of loci is defined as $r = D/[p(1-p)q(1-q)]^{\frac{1}{2}}$. In an infinite population at equilibrium with random mating, no selection or migration and no linkage, r is 0. In a finite population, random processes will in general cause r to differ from 0, and the variance of r [$E(r^2)$] is a function of N_e (Weir and Hill, 1980; Hill, 1981) for unlinked loci, $E(r^2) \approx 1/(3N_e)$. Estimating the parametric r^2 by $\hat{r}^2$ using data for a sample of S individuals introduces an additional source of sampling error, so the expectation of $\hat{r}^2$ for unlinked loci is given (after Hill, 1981) by

$$E(\hat{r}^2) \approx 1/(3N_e) + 1/S, \quad (5)$$

which can be rearranged to yield

$$\hat{N}_e = 1/[3(\hat{r}^2 - 1/S)] \quad (6)$$

As is the case with $\hat{F}$, a single data point (i.e. for a single pair of loci) cannot be expected to yield a precise estimate of N_e. However, if information for 8 to 10 or more unlinked gene loci is available and all possible pairs are considered, the number of useful data points [J(J-1)/2 pairs for J loci] is substantial. Hill (1981) showed that for a single pair of loci, the variance of r^2 was about twice the mean, as would be expected for a variable distributed as chi square with one degree of freedom. He also showed that if data for multiple loci are used, the variance of r^2 is inversely proportional to the quantity J(J-1)/2. Simulations I have performed using up to 15 loci confirm that $\bar{r}^2$ (the mean of the $\hat{r}^2$ values over all pairs of loci) is distributed approximately as chi square with J(J-1)/2 degrees of freedom. Precision of $\hat{N}_e$ thus increases rapidly with the number of loci used.

One drawback to the use of tightly linked loci to estimate N_e is that considerable time is necessary for recombination to erode existing disequlibria. Therefore, current D or r values for tightly linked loci may contain information about effective population size in the distant past. Residual disequilibria are much less a problem at unlinked loci; for such loci, existing disequilibria decay at a rate of one-half per generation. (Use of such loci to estimate N_e depends on the generation of new disequilibria each generation by genetic drift.) Simulation results (unpublished data) verify that for unlinked loci, D and r values are little influenced by disequilibrium levels in previous generations. This is particularly true if population size decreases, in which case the signal from the small size in the current generation overwhelms the residual disequilibria from previous generations.

Application

The assumptions relevant to the disequilibrium method are similar to those of the temporal method. Strong selection at loci being considered or closely linked loci could bias the estimate of N_e but, as mentioned previously, evidence from natural populations does not suggest that epistasis strongly affects D values for most pairs of gene loci detected by electrophoresis. Balancing selection at individual gene loci apparently does not affect the mean or variance of r (Felsenstein, 1974; Avery, 1978). The caveats about random sampling and population subdivision mentioned in connection with the temporal method also apply here. If data on linkage are available, equations (5) and (6) can be modified slightly according to Hill (1981). In the absence of such data, the presumption that the loci considered are unlinked is probably not unreasonable, given that cetaceans typically have over 20 pairs of chromosomes (e.g. Duffield and Wells, this volume).

The disequilibrium method makes the additional assumption of random mating, not required by the temporal method. Although allele frequencies (as monitored in the temporal method) are not affected by mating structure *per se*, genotypic frequencies are, and D (and r) will be larger to the extent that only certain pairs of individuals in the parental generation unite to form gametes. Weir and Hill (1980) explored the effects of mating structure on linkage disequilibrium and found that for unlinked loci in dioecious organisms, the most extreme departure of $E(\hat{r}^2)$ from that given in (5) occurs if matings are strictly monogamous. Even in that event, however, the expectation of $\hat{r}^2$ $[5/(12N_e) + 1/S]$ is still close to that given in (5). As permanent pair bonds do not appear to be formed in most cetacean species (Lockyer, 1984), the effect on $E(\hat{r}^2)$ from this factor would appear to be small.

Sampling from age-structured populations introduces additional possibilities for nonrandom mating effects. The complication is that individuals of different ages are not the product of a single episode of random mating in the population; rather, they are produced by a breeding population that changes gradually in composition (and allele frequency) over time. Treating such a population as a single sample will tend to produce gametic disequilibrium, for the same reason that including multiple gene pools in a single sample will produce an apparent deficit of single-locus heterozygotes (the Wahlund effect).

This problem can be ignored on the assumption that the effect will be small, but this may not be true in small populations or if age structure is changing rapidly. A better approach, if enough demographic information were available, would be to compute an appropriate adjustment for this effect, as Waples (1990a) did to account for the slight reduction in heterozygotes in Pacific salmon populations caused by uniting individuals from several different brood years. Another alternative is to avoid the problem by considering data for individual cohorts, as suggested above for the temporal method. A separate estimate of effective size (N_b as described previously) can be obtained for each cohort, and the

harmonic mean of these estimates provides an estimate of the harmonic mean effective number of breeders per year (N_b). This approach has the disadvantage (also discussed above) that N_b cannot easily be converted into an estimate of N_e.

Because reasonable precision can be expected from the disequilibrium method only by using data for a number of gene loci, it is important that a sufficient number of polymorphic systems are available. Although low average heterozygosities have been reported for a number of large mammal species, including some cetaceans, this is not the case with all cetaceans (Wada and Numachi, this volume), and a sufficient number of variable loci may be found in some cases. The maximum amount of information can be attained by considering all alleles at very polymorphic systems, but doing so would require knowledge of the covariance structure of multiallele, multilocus systems. At present, analysis should probably be restricted to a single allele at each locus. Loci with alleles at very high frequency should probably not be used, as the expectation of r in this case may not be zero (Hill, 1981; Hedrick, 1987). An exact guideline cannot be given at this time, but results obtained by Waples and Smouse (1990) suggest that loci with alleles at frequency >0.95 should probably not be used, and a lower threshold may be advisable if sample sizes are small.

Comments

The two methods discussed above share many features. Both provide an estimate of N_e (or N_b) at the time of sampling. Both are affected in a similar way (if at all) by violations of basic assumptions. Both require substantial samples of individuals and gene loci for reasonable precision, and for both methods the key estimator is distributed as chi square, so precision can be evaluated and confidence limits placed on point estimates.

It seems reasonable, therefore, that maximum information about current effective population size can be obtained by a combination of the two methods, provided that the relevant assumptions are met. Simulation results (unpublished data) indicate that $\hat{f}^2$ and $\hat{F}$ are at most very weakly correlated, so the two methods provide essentially independent information about effective size. Since genotypic data must be gathered to generate allele frequencies for use with the temporal method, the data necessary for use of the disequilibrium method will already be available. An appropriate strategy is to compute $\hat{N}_e$ or $\hat{N}_b$ separately for each method and use the harmonic mean as the final point estimate. (The process of averaging individual $\hat{F}$ or $\hat{f}^2$ values before arriving at a mean value to use in (3) or (6) is equivalent to taking the harmonic mean of $\hat{N}_e$ values computed for each individual $\hat{f}^2$ or $\hat{F}$.)

Another feature the two methods share is that they can produce estimates of N_e (or N_b) that are negative. This occurs if $\hat{F}$ or $\hat{f}^2$ is smaller than would be expected to result from sampling error alone (i.e. smaller than $1/\tilde{S}$). In this event, sampling error alone can account for all of the allele frequency change (or disequilibrium) without invoking genetic drift, and the appropriate estimate of N_e is infinity (Laurie-Ahlberg and Weir, 1979; Hill, 1981). This result is unlikely in small populations if adequate samples of individuals and genes are used, but it is not unusual for large populations, particularly if sample size is limited (Waples, 1989).

The last point illustrates another feature that the two methods share: for a given sample size, each is best suited to the analysis of small populations. The explanation for this is twofold. First, both measure processes that are inversely proportional to effective size, and the term $1/N_e$ changes most dramatically for small values of N_e. Because $1/N_e$ differs very little for populations of, say, size N_e = 1,000 and N_e = 5,000, the two methods will often have trouble discerning large populations from very large ones (cf. Fig. 1). Second, in large populations, the contribution to $\hat{f}^2$ or $\hat{F}$ from sampling error ($1/\tilde{S}$) may be large in

comparison with the signal from genetic drift, thus creating considerable noise in the analysis. On the other hand, in small populations, the genetic consequences of drift are substantial and allow much more precision for both methods. In this respect, analysis of individual cohorts is an advantage. Since N_b per year is less than N_e per generation, the signal from random processes related to effective size is proportionally stronger for individual cohorts than it is for the population as a whole. Waples (1990b) found a similar effect in the Pacific salmon model when individual brood years were compared.

The difficulty the two methods have in estimating effective size of large populations may be a limitation in some contexts, but in the field of conservation biology it is generally more important to be able to identify populations that have suffered reductions in size (and therefore may have small N_e). For large populations, the temporal method and the disequilibrium method may only be able to indicate that the effective size is large enough that inbreeding and genetic drift are unlikely to cause problems in the near future.

Another feature the two methods share is that they both have been used primarily with data provided by protein electrophoresis. This need not be the case, however. Increasingly, DNA techniques are capable of isolating single-copy nuclear genes that conform to Mendelian models of inheritance. It seems likely that data for such systems will be used with these indirect methods in the near future.

METHODS THAT ESTIMATE LONG-TERM N_e

In contrast to the methods described above, which monitor the stochastic effects of genetic drift over periods of a few generations or less, methods designed to estimate long-term N_e depend on equilibrium processes, which are also strongly affected by mutation rates.

Levels of heterozygosity

Approach

Kimura and Crow (1964) showed that under selective neutrality, the expected heterozygosity (H) at equilibrium is a function of effective population size and the mutation rate to neutral alleles (μ):

$$E(H) = \frac{4N_e\mu}{1 + 4N_e\mu} \tag{7}$$

In (7), N_e represents the long-term effective size for the species as a whole, rather than a local population. Long-term effective size is the harmonic mean of N_e values in each generation if those values change over time. Equation (7) was derived for the infinite-allele model, in which each new mutation occurs at a different site. An analogous formula is available for the charge-state model of Ohta and Kimura (1973) which in theory is more applicable to the analysis of electrophoretic data. However, Nei and Graur (1984) argued that results presented by Fuerst and Ferrell (1980) and McCommas (1983) show that empirical data for protein electrophoresis are more compatible with expectations from the infinite allele model.

Equation (7) can be rearranged to provide an estimator of effective size:

$$\hat{N}_e \approx \frac{H}{4\mu(1\text{-}H)} \tag{8}$$

Application

A number of factors argue for considerable caution in using heterozygosity as an indication of effective population size. First, the estimate can be no more accurate than

the estimate of mutation rate, and there is considerable uncertainty about the appropriate value for μ. In any event, it is clear that the mutation rate to electrophoretically detectable alleles is not the same for all gene loci (Hartl, 1980; Selander and Whittam, 1983; Hills, 1984). Second, empirical data do not support the relationship shown in (7). In examining data for a large number of species, Nei and Graur (1984) found a positive correlation between heterozygosity and estimated population size, but the relationship between the two variables was substantially different than that predicted by (7). In particular, species thought to have very large current population sizes all had much lower H values than predicted from (7). It is possible that all of the species with large current population sizes have experienced severe bottlenecks in the past, possibly causing long-term N_e (and expected heterozygosity) to be much lower than the current size would suggest (Kimura, 1983). However, another likely explanation is that many (perhaps most) electrophoretically-detectable mutations are slightly deleterious and therefore do not reach frequencies at which they would make an appreciable contribution to average heterozygosity (Ohta and Kimura, 1975). If true, this would cause estimates of effective size from (8) to be biased downward.

A third difficulty with using equation (8) to estimate N_e is that, even at equilibrium, there is a large inter-locus stochastic variance associated with the expected heterozygosity (Stewart, 1976; Archie, 1985). This means that species with identical long-term N_e can have very different levels of heterozygosity simply by chance, particularly if the number of loci surveyed is not large. Finally, the assumption of selective neutrality is a crucial one for an equilibrium model in which selection can be very weak and still markedly affect the result. The plethora of selection models that have been proposed to explain genetic polymorphisms and the variety of efforts that have been made to relate observed heterozygosities to environmental or life-history variables emphasise the lack of consensus in the scientific community regarding the applicability of the neutrality assumption with respect to heterozygosities.

Nucleotide diversity

Approach

Another approach to estimating long-term N_e, focuses on nucleotide diversity (p), which is equivalent to heterozygosity at the DNA nucleotide level (Nei, 1987). Nucleotide diversity in a population is the average proportion of nucleotide differences between homologous nucleomorphs (sections of DNA) from two different individuals. If the rate of nucleotide sequence divergence (s) is known, p can be used to estimate the average time in generations (t_x) since DNA sequences from two randomly-chosen individuals shared a common ancestor. This latter quantity is significant because the expected value of t_x is equal to the effective population size (Tajima, 1983). In large populations, there is ample opportunity for mutations to accumulate, and the resulting genetic diversity means that the common ancestor for any given pair of DNA sequences will generally be found in the distant past. In small populations, on the other hand, random extinction constantly prunes the number of existing gene lineages, so that a recent common ancestry for any given pair is more likely (Avise *et al.*, 1984).

If s is the estimate of the percent substitution rate per million years and g is the generation length in years, then the following formula can be used to estimate N_e (after Wilson *et al.*, 1985):

$$N_e = E(t_x) \approx 10^8 p/(sg) \quad (9)$$

Equation (9) is appropriate if recombination is absent or very rare, so this approach is suitable for use with mtDNA data (Tajima, 1983; Avise *et al.*, 1988). Because mtDNA is

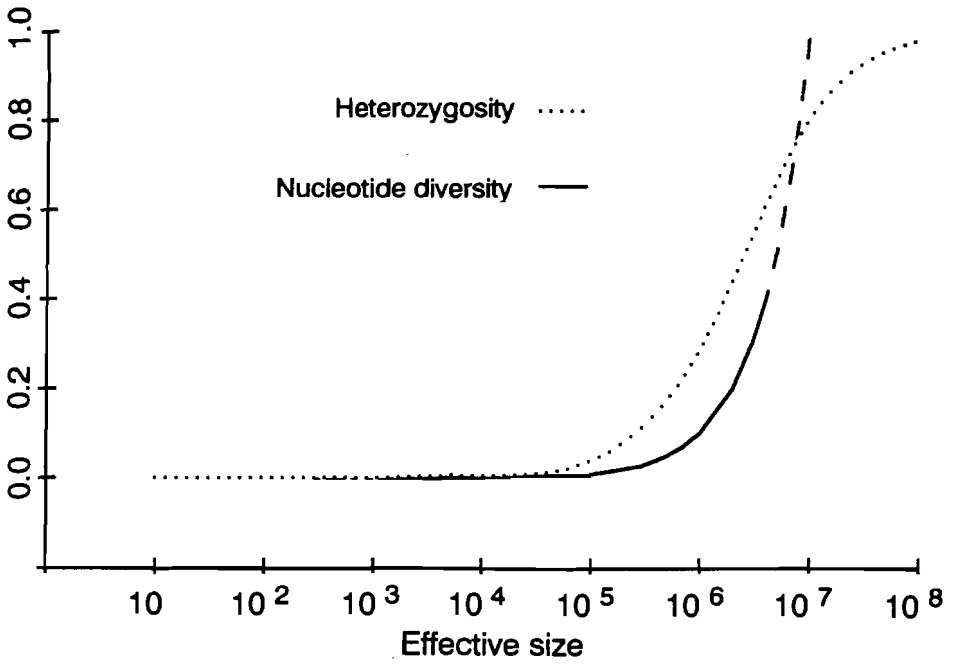

Fig. 3. Theoretical relationship at equilibrium between effective population size and two measures of gene diversity. Expected values of average heterozygosity (H) and nucleotide diversity (p) are both very small unless population size is quite large. In contrast to the temporal and disequilibrium methods, therefore, these indices are best suited to providing information about the effective size of large populations or entire species ($10^7 > N_e > 10^5$). Curve for H assumes $\mu = 10^{-7}$ per gene per generation; that for p assumes a generation length of 5 years and a constant divergence rate of s = 0.02 per site per million years. Upper part of curve for p is shown as a broken line because if p is already large, new mutations are more likely to occur at sites that already differ, thus reducing the measured rate of divergence. The true approach of p to the limiting value of 1.0 must be asymptotic; as shown for heterozygosity.

maternally inherited, (9) actually produces an estimate of the female effective population size; the term $N_{f(e)}$ has been used by some authors in this context. Nei and Li (1979) and Nei (1987) discuss methods for calculating p. The substitution rate has been estimated for a number of groups of organisms using DNA distance data and divergence times estimated from the fossil record (see Hoelzel and Dover, this volume, for discussion).

Application

As do most of the methods discussed here, the nucleotide method assumes selective neutrality and a closed population. The effects of selection on the estimate of N_e have not been evaluated in detail; however, Tajima (1983) pointed out that slightly deleterious mutations should not strongly affect the average number of nucleotide differences between individuals, suggesting that this type of selection might not appreciably bias the estimate of N_e.

In theory, this method could be used to estimate N_e for either a single population or a species, but most individual populations probably are not large enough and do not remain isolated long enough for sufficient mutations to accumulate for the method to be useful (Avise *et al.*, 1988; see also Fig. 3). As was the case with the temporal and disequilibrium

methods, population subdivision or migration between sub-populations may bias the estimate. The difference, however, is that whereas the only processes of interest to the former two models are those in the current (or very recent) generations, historical structuring of the population can affect the estimate of N_e in the nucleotide method (see below). Because of this complication, Avise *et al.* (1988) used this method only on species believed to have experienced consistently high levels of gene flow (i.e. those that have approached panmixia for the species as a whole) throughout their recent history.

Clearly, it is important that the estimate of substitution rate be accurate for the method to provide reliable information about effective population size. Results discussed by Hoelzel and Dover (this volume) show that there is considerable variability in estimated substitution rates among groups of animals and, within groups, among different portions of the genome. It therefore does not seem possible (or advisible) to suggest a single, generic value for use in all cases. If possible, independent calibrations should be performed using the relevant organisms and portions of the genome (cf. Hoelzel and Dover, this volume). If a 'conventional' rate of substitution is used, the analysis should also be done using a broader range of rates to assess robustness of the results (cf. Avise *et al.*, 1988).

The question of precision of the $\hat{N}_e$ estimates derived from the nucleotide method has not been addressed directly. However, Tajima (1983) used data from Aquadro and Greenberg (1983) to show that the variance of $\hat{p}$ is quite large. Furthermore, the stochastic variance represents a substantial proportion of the total variance, the remainder being due to error in taking a sample for analysis. This suggests that the potential to increase precision of $\hat{N}_e$ by taking larger samples is limited.

Other approaches

Lethal mutation rate

Nei, 1968 suggested a method for estimating effective size of an isolated population from the allelism rates of lethal chromosomes. It seems unlikely that sufficient data will be available from cetacean populations for this method to be useful in the foreseeable future. Yamazaki *et al.* (1986) give an example of the use of this approach with *Drosophila*.

Genetic differentiation and gene flow

A large number of authors have used allele frequency data to estimate the parameter mN_e from the relationship $F_{ST} \approx 1/(1 + 4mN_e)$ (Wright, 1943). In this formula F_{ST} is the standardised variance of allele frequency among sub-populations and m is the proportion of individuals that migrate between subpopulations each generation. The parameter mN_e can also be estimated using the private allele method of Slatkin (1985b). Larson *et al.* (1984) estimated mN_e for several species of salamanders using F_{ST} values and calculated a maximum value for m (m*) from the relationship $I = m^*/(m^* + \mu)$ (Nei, 1975), where I is the average value of Nei's (1972) genetic identity for all possible pairwise comparisons of populations. Larson *et al.* (1984) obtained a minimum estimate of N_e by dividing the mN_e estimate from F_{ST} by m*. Because F_{ST} values may take a long time to reach equilibrium in subdivided populations (Slatkin, 1985a), the resulting estimate should be interpreted as one of the long-term effective population size.

Assumptions associated with this method are that the alleles studied are selectively neutral, that $m >> \mu$ and that an island model of population structure (and its attendant assumptions; Wright, 1943) is appropriate. Results discussed by Slatkin (1985a) indicate that if migration actually corresponds to a stepping-stone model, the estimate of mN_e using the above method will be biased downwards. Robustness of the estimate is difficult

to evaluate; it seems likely that any point estimate would have quite a large variance, given that the parameters mN_e, m^* and μ all must be estimated.

Comments

The two methods (heterozygosity, nucleotide differences) that provide an estimate of long-term N_e share some features that differ from those of the models discussed previously. Accuracy of both methods depends heavily on how well the mutation or substitution rate can be estimated. The long-term methods also differ from those discussed previously with respect to the type of populations they are best suited to analyse. Whereas the methods designed to estimate current N_e provide much greater precision with small populations, the reverse is true for the long-term methods. As shown in Fig. 3 the expected heterozygosity at equilibrium is very similar for all populations with effective size less than about 10^5, and a similar insensitivity to small population size applies to the method based on average nucleotide divergence.

Of the two methods, that based on nucleotide differences in mtDNA seems better suited to estimating effective size because it apparently is less affected by slightly deleterious mutations.

METHODS THAT PROVIDE INFORMATION ABOUT POPULATION BOTTLENECKS

Population bottlenecks lead to non-equilibrium states, and the genetic signatures of these events can in theory be detected in a variety of ways. The two methods discussed here may provide insight into recent changes in population size by comparing genetic parameters that respond in different ways to population changes.

Number of alleles versus heterozygosity

Approach

Maruyama and Fuerst (1984; 1985) considered two measures of genetic variability (average heterozygosity and the number of alleles at a locus) and compared their expected values for initially large populations that suffered a severe, temporary bottleneck and subsequently increased again in size. At equilibrium in a steady-state population, expected heterozygosity is given by (7) and comparable formulae are available for the number of alleles expected to occur one, two ...j times in a sample of j alleles. In such a steady-state population, one could use H values obtained empirically to estimate the parameter $4N_eu$ using a variation of (7) (cf. Nei, 1975), and this estimate should allow one to make predictions about the number of alleles expected in the population. Changes from the steady-state values, however, do not occur at the same rate for the two measures when population size changes. During a bottleneck, heterozygosity is lost at the rate of $1/(2N_e)$ per generation regardless of the initial level of variation. In contrast, alleles are lost at a rate that depends on $4N_eu$ (Maruyama and Fuerst, 1985), meaning that losses are most rapid in large populations with high initial allelic diversity. If such a population were sampled following a bottleneck, the number of alleles observed would be less than predicted from the $4N_eu$ value consistent with the observed heterozygosity.

In an increasing population, conversely, the number of alleles increases more rapidly than does heterozygosity, causing an apparent excess of alleles in relation to heterozygosity. Therefore, the perceived deficit of alleles due to a population bottleneck is transitory if the bottleneck also is temporary. In fact, the excess caused by population increase is greater than the deficit caused by population bottleneck, so on average there

will be a net excess of alleles in a population that has recently undergone a temporary bottleneck followed by a restoration of the original size (Maruyama and Fuerst, 1985).

Segregating sites and nucleotide diversity

Approach

Tajima (1989) considered a pair of measures of DNA variation that behave in a fashion similar to those just described above. The average number of nucleotide differences between a pair of DNA sequences and number of sites that are segregating (polymorphic) in a population are analogous to heterozygosity and the number of alleles, respectively, in the model of Maruyama and Fuerst. Tajima found that in non-equilibrium populations, the number of segregating sites is more strongly influenced by current N_e, whereas the average number of nucleotide differences is more strongly affected by original population size. In a population that has experienced a bottleneck and then recovered, the effect on the number of nucleotide differences will be greater

Comments

Neither of the approaches discussed in this section were developed for the purpose of estimating effective population size. Rather, the motivation for these studies was to identify factors that might cause certain population genetic parameters to depart from expectations given by neutral models. Consequently, there has been no direct evaluation of the usefulness of these approaches for the purposes considered here. A brief description of them has been included because they may provide insight into the occurrence of recent population bottlenecks, which is a topic of considerable interest to cetacean biologists.

Several other approaches, or variations on those discussed above, have been used to draw inferences about historical population size. Wilson *et al.* (1985) pointed out that a single breeding pair contains four copies of the nuclear genome but only a single transferable copy of mitochondrial DNA, suggesting that bottlenecks effects are more likely to be detected with mtDNA. They summarised mtDNA data for a number of species that are consistent with recent bottlenecks. Rozas *et al.* (1990) found that presumed bottlenecks in *Drosophilia subobscura* were more easily detected from analysis of mtDNA data than from allozyme data. O'Brien and Evermann (1988) found evidence for reduced levels of genetic diversity (measured by allozymes, 2D gels, DNA markers and tissue grafts) in a number of species thought to have experienced recent population crashes. Choudhary and Singh (1987) compared genetic parameters for two closely related *Drosophila* species and concluded that founder effects and population bottlenecks had played a significant role in shaping the genetic composition of *D. melanogaster*. A similar comparative approach may be possible for some closely related pairs or groups of cetacean species.

DISCUSSION

Limiting factors

The usefulness of indirect methods for estimating the effective size of cetacean populations can be evaluated with respect to two questions: (1) Does the method, on average, provide an estimate that is accurate (unbiased)? and (2) Does the method provide an estimate that is precise enough (small enough variance) that it provides biologically meaningful information?

Accuracy of the estimators provided by the various models depends primarily on how well the assumptions of the model are met. The most serious potential sources of bias for

the two methods that estimate current effective size are probably nonrandom sampling and population subdivision. Cetaceans are diverse enough in life-history features that the degree to which these factors will be important must be evaluated on a case-by-case basis. Methods providing estimates of long-term N_e may be more sensitive to natural selection and will provide biased estimates if the estimate of mutation rate or DNA base substitution rate is inaccurate. These latter methods also assume conditions of genetic equilibrium and may be biased if there have been recent changes in population size. Under certain conditions, comparison of observed and expected values for different genetic parameters can provide insight into such changes.

In many situations, precision of the point estimate of N_e may be the limiting factor for the practical application of the indirect methods. This can occur even for a method that provides an unbiased estimate (i.e. one that on average gives the correct value) if the confidence limits associated with any single estimate are so large that it is not very useful. All indirect methods are subject to a variance which derives from two factors. First, any genetic process that provides information about effective population size does so because it measures a random process whose magnitude or rate is determined by N_e. Being random, these processes are predictable only in a statistical sense. For example, in a population of $N_e = 50$ and $P = 0.5$, allele frequency might change by 0.04 in one generation, but it might also remain the same or change by twice as much (0.08). Although these different outcomes can result from random drift in the same population, each would provide a very different estimate of N_e using the temporal method.

This type of uncertainty, termed 'stochastic variance' by Tajima (1983), is intrinsic to all the indirect methods. It can only be reduced if there are multiple events that can be measured, each a random process but together characteristic of a population of a particular effective size. This explains why it is important in the temporal and disequilibrium methods to sample a large number of gene loci. Each independent allele frequency (or pair of alleles) provides information about a random process that reflects effective size. Similarly, allele frequencies measured several generations apart in the temporal method provide information about several episodes of genetic drift.

As noted by Tajima (1983) there is a large stochastic variance inherent in the nucleotide difference method for estimating long-term N_e. Increasing the number of restriction sites (or bases) measured reduces this variance but only to a certain point. It is also important to have data for haplotype relationships at many different gene loci, but at present the only suitable molecule for analysis in animals is mtDNA (Avise *et al.*, 1988). Ball *et al.* (1990) discuss some of the potential problems in using data from a single gene lineage to draw inferences about a population or species.

The second source of variance in estimates produced by the indirect methods is sampling error in estimating genetic parameters for the population. In contrast to the stochastic variance described above, this source of variability can be reduced by taking large samples. For the two methods that estimate current N_e, the ratio $S:N_e$ is an indication of the relative importance of sampling error (Nei and Tajima, 1981; Waples, 1989). If S/N_e is small, noise from sampling may swamp the signal from genetic drift. For a given sample size, then, these methods provide maximum precision for the study of small populations, as we have seen.

Together, the two sources of variance may be large enough that the estimate of effective size from an indirect method provides little biologically useful information. On the other hand, it often may be possible to reduce these sources of uncertainty to manageable proportions. Furthermore, effective population size is such an important parameter in biology, and so difficult to measure directly, that even an estimate that is within an order of magnitude of the true value may prove useful (Nei and Tajima, 1981).

Recommendations

Because most cetacean species are long-lived, it will be difficult to sample enough generations of genetic drift for standard applications of the temporal method to provide much precision in estimating current N_e. For species for which a number of unlinked, polymorphic gene loci can be resolved, the disequilibrium method may provide biologically meaningful information about effective population size. A variation of these methods, in which genetic characteristics are monitored in individual cohorts, provides a means of estimating the effective number of breeders each year (N_b). Although difficulties remain in relating N_b to N_e in cetaceans, this approach may still provide more insight into effective size than is possible with other methods. Monitoring both allele frequency change and gametic disequilibrium is recommended if efforts are made to estimate N_b. For this approach to be useful, it will be necessary to sample a sizeable number of individuals (so that the number in each cohort is not too small), age them accurately and have data for a reasonable number of polymorphic systems.

Data to be used in estimating N_e or N_b can be gathered in conjunction with other genetic studies (e.g. those designed to evaluate stock structure). As protein and mtDNA data often provide complementary insights into population subdivision, both types of data should probably be gathered routinely. Estimates of long-term N_e provided by mtDNA data can similarly complement the information about current N_e provided by the temporal and disequilibrium methods. In certain circumstances, other approaches discussed briefly here may provide insight into the dynamics of historical changes in population size.

ACKNOWLEDGEMENTS

I thank John Avise and two anonymous reviewers for helpful comments on an earlier draft of the manuscript.

REFERENCES

Aquadro, C.F. and Greenberg, B.D. 1983. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequence from seven individuals. *Genetics* 103:287–312.

Archie, J.W. 1985. Statistical analysis of heterozygosity data: independent sample comparisons. *Evolution* 39:623–37.

Avery, P.J. 1978. The effects of finite population size on models of linked overdominant loci. *Genet. Res.* 31:239–54.

Avise, J.C., Neigel, J.E. and Arnold, J. 1984. Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *J. Mol. Evol.* 20:99–105.

Avise, J.C., Ball, R.M. and Arnold, J. 1988. Current versus historical population sizes in vertebrate species with high gene flow: A comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. *Mol. Biol. Evol.* 5:331–44.

Ball, R.M., Neigel, J.E. and Avise, J.C. 1990. Gene genealogies within the organismal pedigrees of random-mating populations. *Evolution* 44:360–70.

Brown, A.H.D. 1975. Sample sizes required to detect linkage disequilibrium between two or three loci. *Theor. Popul. Biol.* 8:184–201.

Choudhary, M. and Singh, R.S. 1987. Historical effective size and the level of genetic diversity in *Drosophila melanogaster* and *Drosophila pseudoobscura*. *Biochem. Genet.* 25:41–51.

Crow, J.F. and Denniston, C. 1988. Inbreeding and variance effective population numbers. *Evolution* 42:482–95.

Duffield, D.A. and Wells, R.S. The combined application of chromosome, protein and molecular data for the investigation of social unit structure and dynamics in *Tursiops truncatus*. (Published in this volume.)

Felsenstein, J. 1974. Uncorrelated genetic drift of gene frequencies and linkage disequilibrium in some models of linked overdominant polymorphisms. *Genet. Res.* 24:281–94.

Fuerst, P.A. and Ferrell, R.E. 1980. The step-wise mutation model: an experimental evaluation utilising hemoglobin variants. *Genetics* 94:185–201.

Gaines, M.S. and Whittam, T. 1980. Genetic changes in fluctuating vole populations: selective vs. nonselective forces. *Genetics* 96:767–78.

Hammond, P.S. 1987. Techniques for estimating the size of whale populations. pp. 225–45. *In:* S. Harris (ed.) *Mammal Population Studies*. Clarendon Press, Oxford. i-xvi + 350pp.

Hartl, D. 1980. *Principles of Population Genetics*. Sinauer Associates, Sunderland, Mass. 480pp.

Hedrick, P.W. 1987. Gametic disequilibrium measures: proceed with caution. *Genetics* 117:331–41.

Hill, W.G. 1974. Estimation of linkage disequilibrium in randomly mating populations. *Heredity* 33:229–39.

Hill, W.G. 1981. Estimation of effective population size from data on linkage disequilibrium. *Genet. Res.* 38:209–16.

Hills, D.M. 1984. Misuse and modification of Nei's genetic distance. *Syst. Zool.* 33:238–40.

Hoelzel, A.R. and Dover, G.A. Mitochondrial D-loop DNA variation within and between populations of the minke whale (*Balaenoptera acutorostrata*). (Published in this volume.)

Kimura, M. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, New York. i-xv + 367pp.

Kimura, M. and Crow, J.F. 1964. The number of alleles that can be maintained in a finite population. *Genetics* 49:725–38.

Krimbas, C.B. and Tsakas, S. 1971. The genetics of *Dacus oleae* V. Changes of esterase polymorphism in a natural population following insecticide control-selection or drift? *Evolution* 25:454–60.

Larson, A., Wake, D.B. and Yanev, K.P. 1984. Measuring gene flow among populations having high levels of genetic fragmentation. *Genetics* 106:293–308.

Laurie-Ahlberg, C.C. and Weir, B.S. 1979. Allozyme variation and linkage disequilibrium in some laboratory populations of *Drosophila melanogaster*. *Genetics* 92:1295–314.

Lewontin, R.C. and Krakauer, J. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of protein polymorphisms. *Genetics* 74:175–95.

Lewontin, R.C. and Krakauer, J. 1975. Testing the heterogeneity of F values. *Genetics* 80:397–8.

Lockyer, C. 1984. Review of baleen whale (Mysticeti) reproduction and implications for management. *Rep. int. Whal. Commn* (special issue 6):27–50.

Maruyama, T. and Fuerst, P.A. 1984. Population bottlenecks and nonequilibrium models in population genetics. I. Allele numbers when populations evolve from zero variability. *Genetics* 108:745–63.

Maruyama, T. and Fuerst, P.A. 1985. Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics* 111:675–89.

McCommas, S.A. 1983. A taxonomic approach to evaluation of the charge state model using twelve species of sea anemone. *Genetics* 103:741–52.

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

Nei, M. 1968. The frequency distribution of lethal chromosomes in finite populations. *Proc. NAS (U.S.)* 60:517–24.

Nei, M. 1972. Genetic distance between populations. *Am. Nat.* 106:283–92.

Nei, M. 1975. *Molecular Population Genetics and Evolution*. North-Holland, Amsterdam. i-xiii 283pp.

Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York. i-x 511pp.

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

Nei, M. and Li, W.H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. NAS (U.S.)* 76:5269–73.

Nei, M. and Tajima, F. 1981. Genetic drift and estimation of effective population size. *Genetics* 98:625–40.

O'Brien, S.J. and Evermann, J.F. 1988. Interactive influence of infectious disease and genetic diversity in natural populations. *Trends Ecol. Evol.* 3:254–9.

Ohta, T. and Kimura, M. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genet. Res.* 22:201–4.

Ohta, T. and Kimura, M. 1975. Theoretical analysis of electrophoretically detectable polymorphisms: Models of very slightly deleterious mutations. *Am. Nat.* 109:137–45.

Pollak, E. 1983. A new method for estimating the effective population size from allele frequency changes. *Genetics* 104:531–48.

Rozas, J., Hernández, M., Cabrera, V.M. and Prevosti, A. 1990. Colonisation of America by *Drosophila subobscura*: effect of the founder event on mitochondrial DNA polymorphism. *Mol. Biol. Evol.* 7:103–9.

Selander, R.K. and Whittam, T.S. 1983. Protein polymorphism and the genetic structure of populations. pp. 89–114. *In:* M. Nei and R.K. Koehn (eds.) *Evolution of Genes and Proteins*. Sinauer, Sunderland, Maryland. i-xiii + 331pp.
Slatkin, M. 1985a. Gene flow in natural populations. *Annu. Rev. Ecol. Syst.* 16:393–430.
Slatkin, M. 1985b. Rare alleles as indicators of gene flow. *Evolution* 39:53–65.
Smit-McBride, Z., Moya, A. and Ayala, F.J. 1988. Linkage disequilibrium in natural and experimental populations of *Drosophila melanogaster*. *Genetics* 120:1043–51.
Stewart, F.M. 1976. Variability in the amount of heterozygosity maintained by neutral mutations. *Theor. Popul. Biol.* 9:188–201.
Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105:437–60.
Tajima, F. 1989. The effect of change in population size on DNA polymorphism. *Genetics* 123:597-601.
Wada, S. and Numachi, K. Allozyme analyses of genetic differentiation among the populations and species of the *Balaenoptera*. (Published in this volume.)
Waples, R.S. 1989. A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* 121:379–91.
Waples, R.S. 1990a. Conservation genetics of Pacific salmon. II. Effective population size and the rate of loss of genetic variability. *J. Heredity* 81:267–76.
Waples, R.S. 1990b. Conservation genetics of Pacific salmon. III. Estimating effective population size. *J. Heredity* 81:277–89.
Waples, R.S. and Smouse, P.E. 1990. Gametic disequilibrium analysis as a means of identifying mixtures of salmon populations. *Am. Fish. Soc. Symp.* 7:439–58.
Weir, B.S. 1979. Inferences about linkage disequilibrium. *Biometrics* 35:235–54.
Weir, B.S. and Hill, W.G. 1980. Effect of mating structure on variation in linkage disequilibrium. *Genetics* 95:477–88.
Wilson, A.C., Cann, R.L., Carr, S.M., George, M., Gyllensten, U.B., Helm-Bychowski, K.M., Higuchi, R.G., Palumbi, S.R., Prager, E.M., Sage, R.D. and Stoneking, M. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc.* 26:375–400.
Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159.
Wright, S. 1938. Size of population and breeding structure in relation to evolution. *Science* 87:430-1.
Wright, S. 1943. Isolation by distance. *Genetics* 28:114–38.
Yamazaki, T., Choo, J.K., Watanabe, T.K. and Takahata, N. 1986. Gene flow in natural populations of *Drosophila melanogaster* with special reference to lethal allelism rates and protein variation. *Genetics* 113:73–89.

[*Appendix overleaf*]

APPENDIX

Computation of $\hat{F}$

Several methods have been suggested for computing $\hat{F}$; those presented by Nei and Tajima (1981) and Pollak (1983) have seen the most use. As they lead to very similar results, only the latter measure is discussed here. Consider a sample taken in generation 0 and surveyed for genetic variation at a gene locus with L alleles, and denote the sample allele frequencies by X_{01}, X_{02}, ... X_{0L} ($\Sigma X_{0i} = 1$). Assume a second sample is taken at time t, yielding X_{t1}, X_{t2}, ... X_{tL}. Pollak's measure is

$$\hat{F} = \frac{1}{L-1} \sum_{i=t}^{L} \frac{(X_{0i} - X_{ti})^2}{(X_{0i} + X_{ti})/2} \qquad \text{(A1)}$$

The following formula can be used to combine $\hat{F}$ values ($\hat{F}_j$) from several individual loci to arrive at a single mean value: mean $\hat{F} = \Sigma(L_j - 1)\hat{F}_j]/\Sigma(L_j - 1)$. In this scheme, each single locus $\hat{F}$ value (as calculated above) is weighted by the number of independent alleles (L-1 per locus) used in its computation.

Confidence interval for $\hat{N}_e$

Waples (1989) showed that the correct formula for the 1-α confidence limits to an estimate $\hat{F}$ are given by

$$\text{(1-}\alpha\text{) Confidence interval for } \hat{F} = \frac{n\hat{F}}{\chi^2\alpha/2[n]}, \frac{n\hat{F}}{\chi^2 1\text{-}\alpha/2[n]} \qquad \text{(A2)}$$

where n is the number of degrees of freedom associated with $\hat{F}$ and $\chi^2_{\alpha/2[n]}$ (for example) is the critical $\alpha/2$ chi square value for n degrees of freedom. Once confidence limits for $\hat{F}$ are obtained, they can be inserted in the appropriate equation (e.g. Equation 3) to obtain confidence limits to $\hat{N}_e$. A similar approach can be used with the disequilibrium method by using $\hat{r}^2$ instead of $\hat{F}$ in (A2).

The expectation of F in the overlapping generation model

Consider a population with overlapping generations, N adults of reproductive age and allele frequency P at a gene locus of interest. Assume that, for one of the reasons described below, not all adults reproduce successfully every year, and denote by N_i the number breeding in year i. We can also define an effective number of breeders in a given year i (N_{bi}), which in general will not be the same as N_i because of departures from a 1:1 sex ratio and non-binomial variance in reproductive success. N_{bi} is analogous to N_e, except that it refers to the effective number per year rather than per generation. I will explore the expectation of F for a comparison of allele frequencies (P_1, P_2) in the effective breeding population in consecutive years under a variety of different reproductive scenarios.

Random fluctuation in the breeding population

If only a portion of the population breeds successfully in any given year, and the choice of which individuals breed is random and independent across years, then E(F) can be computed as follows. Choosing $2N_b$ genes for the effective breeding population each year is a hypergeometric sampling process, and the variance of $\mathbf{P}_1$ [$V(P_1) = E(P\text{-}\mathbf{P}_1)^2$] is

$$V(P_1) = \frac{P(1\text{-}P)}{2N_b}\left[\frac{2N_b-1}{2N-1}\right]$$

$$= \frac{P(1\text{-}P)}{N_b}\left[\frac{N-N_b}{2N-1}\right] \quad \text{(A3)}$$

$V(P_2)$ is also given by (A2). The variance of the difference in allele frequency $[V(P_1 - P_2) = E(P_1 - P_2)^2]$ is equal to $V(P_1) + V(P_2) - 2Cov(P_1, P_2)$ is the covariance of P_1 and P_2. Since random fluctuation implies independence in the breeding composition across years, the covariance is 0 and

$$V(P_1 - P_2) = \frac{2P(1\text{-}P)}{N_b}\left[\frac{N-N_b}{2N-1}\right] \quad \text{(A4)}$$

If the effective number of breeders differs in the two years being compared then N_b in the above (and the equations that follow) is the harmonic mean of N_{bi} in the two years. The approximate expectation of F is obtained by dividing (A4) by P(1-P):

$$\underset{\text{(random)}}{E(F)} \approx \frac{2}{N_b}\left[\frac{N-N_b}{2N-1}\right] \approx \frac{1}{N_b} - \frac{1}{N} \quad \text{(A5)}$$

Note that with random fluctuations of the breeding population each year, the expectation of F depends on total population size (N) as well as on N_b. However, E(F) does not depend on the number actually participating in breeding in a given year (N_i).

Cyclical reproduction

If reproduction is cyclical, so that individuals reproducing in one year do not do so the next year, $V(P_1)$ and $V(P_2)$ are again as shown in (A2). $V(P_1 - P_2)$, however, differs from the previous model because in the cyclical model the makeup of the breeding population is not independent over time; instead, breeders in consecutive years are mutually exclusive. Waples (1989, Appendix) showed that in this case

$Cov(P_1, P_2) = -P(1\text{-}P)/(2N - 1)$, leading to

$$V(P_1 - P_2) = \frac{2P(1\text{-}P)}{N_b}\left[\frac{N-N_b}{2N-1}\right] + \frac{2P(1\text{-}P)}{2N-1}$$

$$\frac{2P(1\text{-}P)}{N_b}\left[\frac{N}{2N-1}\right] \approx \frac{P(1\text{-}P)}{N_b}$$

and

$$\underset{\text{(cyclical)}}{E(F)} \approx \frac{1}{N_b} \quad \text{(A6)}$$

Note that with cyclical reproduction, E(F) is independent of population size (N).

A mixed reproduction model

If female reproduction is cyclical (owing to gestation time and lactation) but male reproduction is not, there is a mixture of reproductive modes. Such a pattern may be

typical of many cetacean populations. If the same males participate in breeding to the same degree in two consecutive years, then the expected change in allele frequency in the breeding population over a one year period is half that for the cyclical model, and the square of the allele frequency difference $[E(P1 - P2)^2]$ is one-fourth as large. Therefore,

$$E(F)_{\text{mixed}} \approx \frac{1}{4N_b} \qquad \text{(A7)}$$

Almost certainly, however, there will be some variation (perhaps random) in male reproductive success among years. Under such conditions, E(F) is intermediate to that given by (A5) and (A6).

Sampling

The above results apply to allele frequencies in the effective breeding population in consecutive years. A useful way to estimate allele frequencies in the effective breeding population is to sample individuals from the cohorts that they produce (i.e. their progeny). Such samples can be considered binomial samples from the breeding population, so sampling error can be expected to add approximately $1/\tilde{S}$ to the quantities shown in (A5), (A6) and (A7).

Comments

Equations (A5), (A6) and (A7), adjusted to account for sampling as described above, can be used to obtain estimates of the effective number of breeders per year (N_b). However, the difficulties (discussed in the text) in relating N_b to the effective number per generation (N_e) should be kept in mind.