ESTIMATING ANIMAL ABUNDANCE USING NONINVASIVE DNA SAMPLING: PROMISE AND PITFALLS

L. Scott Mills,[1,3] John J. Citta,[1] Kevin P. Lair,[1] Michael K. Schwartz,[1] and David A. Tallmon[2]

[1]*Wildlife Biology Program, School of Forestry, University of Montana, Missoula, Montana 59812 USA*
[2]*Division of Biological Sciences, University of Montana, Missoula, Montana 59812 USA*

Abstract. Advances in molecular biology offer promise to the study of demographic characteristics of rare or hard-to-capture species, because individuals can now be identified through noninvasive sampling such as fecal collection or hair snags. However, individual genotyping using such methods currently leads to a novel problem that we call a ''shadow effect,'' because some animals not captured previously are believed to be recaptures due to their DNA profile being an indistinguishable shadow of previously captured animals. We evaluate the impact of the shadow effect on the two methods most commonly used in applied population ecology to estimate the size of closed populations: Lincoln-Petersen and multiple-recapture estimators in program CAPTURE. We find that the shadow effect can cause a negative bias in the estimates of both the number of different animals and the number of different genotypes. Furthermore, with Lincoln-Petersen estimators, the shadow effect can cause estimated confidence intervals to decrease even as bias increases. Because the bias arises from heterogeneity in apparent ''capture'' probabilities for animals with genetic shadows vs. those without, a model in program CAPTURE that is robust to capture heterogeneity ($M_{h-jackknife}$) does not underestimate the number of genotypes in the population and only slightly underestimates the total number of individuals. As the shadow effect increases, CAPTURE is better able to correctly identify heterogeneity in capture probability and to pick $M_{h-jackknife}$, so that the higher levels of shadow effect have less bias than medium levels. The shadow effect will occur in all estimates of demographic rates (including survival) that use DNA sampling to determine individual identity, but it can be minimized by increasing the number of individual loci sampled.

Key words: demography; DNA; genetic marker; Lincoln-Petersen estimator; mark–recapture; microsatellite; minimum number alive; monitoring; population estimation; probability of identity; rare species; trend analysis.

INTRODUCTION

Evaluation of population dynamics of rare or elusive species has long been plagued by small sample size. Even important parameters such as population size or survival rates have been virtually impossible to estimate using mark–recapture or transect methods, due to the high capture probabilities and/or large numbers of animals required for reasonable estimates. The limitations have been especially acute for carnivores and endangered species, whose low densities have left ecologists relying on raw counts or indices with unknown relationships to true population size. The lack of reliable estimates for these species has forced ecologists to consider how indirect measures, such as photographic ''capture'' (Mace et al. 1994, Karanth and Nichols 1998) or ingestion of biomarkers (Garshelis and Visser 1997), could be used to obtain reliable estimates of detection probabilities and, thus, of population size and associated confidence intervals.

Recently, a technical and conceptual revolution has

occurred in ecological research with the availability and ease of use of molecular techniques (Haig 1998, Parker et al. 1998). Now, tiny samples of DNA can be obtained from individuals in noninvasive ways to provide a potentially unique genetic profile. The possibilities seem endless, including the acquisition of samples via hair snags, feces, regurgitated pellets, or other creatively collected bits of tissue (reviews by Morin and Woodruff 1996, Kohn and Wayne 1997). On the heels of these molecular discoveries, a growing wave of enthusiasm seems about to break, advocating the use of these noninvasive molecular techniques in mark–recapture models (see Kohn and Wayne 1997, Levy 1999). For example, noninvasive sampling of DNA to estimate population size has been proposed or carried out for bears (*Ursus arctos* and *U. americanus*; Woods et al. 1996), humpback whales (*Megaptera novaeangliae*; Palsboll et al. 1997), coyotes (*Canis latrans*; Kohn et al. 1999), and a suite of mid-level carnivores in North America (*Lynx canadensis, Gulo gulo, Martes pennanti,* and *M. americana*; Foran et al. 1997).

Indeed, the use of these unobtrusive or noninvasive approaches may offer advantages over conventional mark–recapture techniques, including increased cap-

ture probability, decreased tag loss, and the potential to minimize the effects of capture and marking. Before ecologists embrace this new technology and its applications, however, it seems wise to carefully consider potential limitations (Snow and Parker 1998, Taberlet and Waits 1998).

The premise of using molecular approaches to estimate demographic rates rests on the same fundamental assumptions that apply to mark–recapture methodology, including unique marking of animals. In this case, the sample of hair or tissue is assumed to yield a unique DNA profile that ''marks'' an individual; subsequent samplings (the ''recapture'' sessions) either match samples from the first session, implying a ''recaptured'' animal, or do not match, implying a newly captured animal. However, the uniqueness of these profiles is probabilistic, meaning that different animals can sometimes share the same profile.

This potential to share genetic profiles among different individuals creates a potential bias that does not typically occur in traditional mark–recapture studies: the possibility that a sample will identify as a ''recapture'' an animal that has not, in fact, been previously captured. By convention (see Paetkau and Strobeck 1994, Palsboll et al. 1997), the proportion of the population having genotypes that cannot be distinguished from at least one other individual is defined as the probability of identity (PI). For example, in a population with a PI of 0.10, 10% of the individuals have genetic matches in the population. A nonzero PI, coupled with the number of individuals (two or more) that share each genotype, thus introduces what we call a ''shadow effect'' into mark–recapture estimates: animals that have not been captured before are believed to be recaptures because their DNA profile is an indistinguishable shadow of those of previously captured animals. Although the probability of identity can be made almost infinitesimally small in human forensics and crime applications (see National Research Council 1996), the same may not be true for ecological studies, in which populations are often small and geographically structured, and in which logistics and finances limit the ability to minimize PI. Note that PI is based on the same principle as the probability of detection calculated in behavioral studies, in which the goal is to maximize the probability of excluding a putative father when he is not the actual father (Westneat et al. 1987). It is unknown how the shadow effect that arises from a nonzero PI could affect demographic estimates of vertebrate populations.

We will not consider technical aspects of choosing a molecular technique or obtaining a genetic profile (see Parker et al. 1998, Taberlet et al. 1999). Rather, we first review how noninvasive DNA sampling is, or would be, used to estimate population size in a mark–recapture framework, and then we evaluate the potential of the shadow effect to influence estimates of population size in actual field studies. We conclude that,

although the approach has limitations, there are straightforward ways to address the limitations, so that noninvasive DNA techniques will continue to be a useful tool to address population-level questions.

BACKGROUND: ESTIMATING POPULATION SIZE VIA NONINVASIVE DNA SAMPLING

Probability of identity: quantifying the match between genotype and individual

The most well-known use of DNA to assay individual profiles is in the area of human and wildlife forensics, in which the absence of a genetic match positively excludes the compared samples as coming from the same individual, whereas the presence of a genetic match suggests, but does not prove, that samples come from the same individual (reviews by Avise et al. 1995, National Research Council 1996). The consequences of a false inclusion, that is, claiming a match or identity between samples when they really come from different individuals, can be severe in forensics cases. Therefore, great attention and money are spent on the theory and practice of minimizing the probability of identity (PI) of genetic profiles between different individuals, so that a match can imply ''beyond a reasonable doubt'' that different samples come from the same individual (Avise 1994, Roeder 1994). A flurry of debate has centered on how factors such as sampling, inbreeding, and population structure could increase the PI in any given case, potentially by many orders of magnitude (Cohen 1990, Lewontin and Hartl 1991, Nichols and Balding 1991, Donnelly 1995). Nevertheless, the consensus seems to be that, in forensic applications, a battery of analytical solutions, coupled with a willingness to conduct extensive sampling both within and outside the population from which forensic samples are collected, will result in PI values that are close enough to zero that a genetic profile can be considered unique to one individual (see Roeder 1994, National Research Council 1996).

The same cannot necessarily be said for the case of noninvasive sampling of ecological populations, in which samples are typically in poor condition and the ability to build reference databases to account for population structure is limited. One of the molecular markers of choice for individual identification in ecological studies is microsatellites; these codominant markers rely on the polymerase chain reaction to amplify DNA in small or degraded samples, are highly variable, and are easily interpreted in terms of allele frequencies (Bruford et al. 1996, Jarne and Lagoda 1996, Parker et al. 1998).

The genotype identity derived from a microsatellite profile equals the individual identity only if the probability of identity is zero. The traditional calculation of probability of identity using microsatellite data is based on a simple formula that assumes Hardy-Weinberg equilibrium. For a single locus with multiple al-

leles, the PI can be calculated as the sum of squares of expected frequencies of all possible genotypes (Paetkau and Strobeck 1994, Paetkau et al. 1995, 1998):

$$PI_{single locus} = \sum_i x_i^4 + \sum_i \sum_{j>i} (2x_i x_j)^2 \qquad (1)$$

where x_i and x_j are the frequencies of the ith and jth alleles, respectively. The smallest PI values will occur when there are many alleles of approximately equal frequency, and the largest PI values will occur when genetic variation is low, as may be expected for species that have undergone recent bottlenecks. As an example, using Eq. 1 for three alleles of equal frequency (0.333), PI would be $(0.0369 + 0.148 = 0.184)$; by comparison, for a locus with three alleles of unequal frequency (0.9, 0.05, 0.05) PI would be $(0.656 + 0.0162 = 0.672)$. The PI calculated with this formula will be biased low for many wild populations, because it does not account for subdivision within populations or philopatry of relatives (Avise et al. 1995; Taberlet and Luikart 1999). Paetkau et al. (1998) present a modified formula that accounts for small sample sizes, although L. P. Waits, G. Luikart, J.-M. Cornuet, and P. Taberlet (*unpublished data*) find that the calculated values differ very little using the corrected vs. uncorrected formula.

The estimated overall PI for a given population is calculated as the product of the PI values for all loci (Paetkau et al. 1995, Parker et al. 1998):

$$PI_{overall} = \prod (PI_{single locus}). \qquad (2)$$

This formula assumes that loci are independent; if they are not, the overall PI will again be biased low (Donnelly 1995).

Levels of PI in wild populations

Given that a shadow effect arising from a nonzero PI could be introduced into demographic estimators, what levels of PI are likely in wild populations? The PI, or proportion of the population whose genetic profile will be indistinguishable from at least one different individual, could, in theory, range from 0 to 1.0. It is not known what values of PI are most likely in field studies, because only a handful of ecological studies have calculated PI (see Appendix). To obtain reasonable PI values for our analysis, we reviewed microsatellite studies of wild birds and mammals, reasoning that these taxa are currently the primary targets of DNA sampling to estimate population size. We determined the average PI per locus for each of two populations in 10 published studies (Appendix) that either provided a PI or that provided sufficient information on microsatellite allele frequencies so that we could calculate it using Eq. 1. We only included studies that examined more than one population, so that we could bracket "high" and "low" PI values within studies.

For these populations, the highest average estimate of PI per locus was 0.51 and the lowest was 0.061 (Appendix). The number of polymorphic loci in these studies ranged from two to 15. The overall PI expected for a population having a given number of loci and PI per locus can be calculated from Eq. 2. For example, if the PI per locus were 0.51 for each of three loci, overall PI would be 0.13; if PI were 0.061 for three loci, the overall PI would equal 0.00023. As previously stated, the probability of identity in real populations is probably underestimated by using the traditional approach based on Eqs. 1 and 2 (see Donnelly 1995, Taberlet and Luikart 1999). Given the range of PI and number of loci typically used (Appendix), and given that the calculated PI values are probably biased small both for single loci and for overall PI, we will consider overall PI values ranging from 0 to 0.2.

Having established reasonable levels of PI, next we must identify how PI is manifest in capture–recapture studies. The genetic matches described by a nonzero PI are genetically indistinguishable shadows that will result in the capture sample appearing to consist of fewer individuals than it actually does. This discrepancy between the true sample size and apparent sample size, based on the number of genetically unique individuals, will depend on both the proportion of shadows (PI) and how those shadows are distributed in the population. For a given PI, there are a number of ways in which the shadow individuals can be distributed with respect to which individuals are genetically indistinguishable from one another. Consider two extremes in how a PI of 0.10 could be distributed among individuals in a sample of 100 different animals. First, shadows could be distributed pairwise, so that there are five pairs of indistinguishable individuals. Within a pair, the two individuals are genetically identical, but each pair is genetically distinguishable from other pairs. An enumeration of the 100 animals would indicate 95 unique genotypes. There would be very different distribution if all 10 shadows shared a common genotype and were therefore genetically indistinguishable from one another. In this case, enumeration of the 100 animals would indicate 91 different genotypes, because genetic analysis would reveal only one animal for the 10 shadows. It is not known how genotypic shadows are actually distributed in real populations, but the number of animals sharing a genotype presumably would increase as genetic drift, inbreeding, or philopatry increased.

Estimating population size in the presence of a shadow effect

Once individual genetic profiles are secured, standard mark–recapture estimators can be applied to the "individual" genotype data. Comprehensive reviews of closed-population estimators are provided elsewhere (e.g., Otis et al. 1978, Nichols 1992, Yoccoz et al. 1993). We consider the impact of the shadow effect arising from a nonzero PI, using the two most common techniques of population size estimation for closed populations: the Lincoln-Petersen estimator and mul-

tiple-recapture models in program CAPTURE. Next, we briefly describe these techniques as they would be used to estimate population size using ''capture'' and ''recapture'' of DNA samples, not animals.

Lincoln-Petersen.—For two-sample Lincoln-Petersen (LP) sampling, the maximum likelihood estimates for population size, $\hat{N}$ and its variance [Var($\hat{N}$)] (Bailey 1951, Seber 1982) are:

$$\hat{N} = \frac{(n_1)(n_2)}{(m_2)} \qquad (4)$$

$$\text{Var}(\hat{N}) = \frac{n_1^2 n_2 (n_2 - m_2)}{m_2^3} \qquad (5)$$

where n_1 is the number of different genotypes in the first sample, n_2 the number in the second sample, and m_2 the number of genotypes in the second sample that are recaptures from the first sample. Although the LP estimator uses only two-capture samples, multiple samples can be collapsed into single-capture and recapture sessions. Note that field studies should use the LP equations corrected for small sample size (Chapman 1951, Pollock et al. 1990); in our evaluation, however, we will rely on large sample size approximations of $E(\hat{N})$ and $E(\text{Var}(\hat{N}))$, based on Eqs. 4 and 5.

In the case of no shadow effect (PI = 0), for any capture probability (p) in session one (p_1) or two (p_2) for all u animals present in the population, the expected values of each of the three LP sufficient statistics are:

$$E(n_1) = up_1$$
$$E(m_2) = up_1p_2$$
$$E(n_2) = up_2 = E(m_2) + p_2[u(1 - p_1)]. \qquad (6)$$

To incorporate the shadow effect arising from a nonzero PI, consider that a population of true size N contains u animals with unique genotypes and (PI $\times$ N) animals with shadows (i.e., that share their genotype with at least one other animal). The number of different genotypes that are shared among the (PI $\times$ N) animals depends on the number of animals that share each genotype (g), such that

s = number of shared genotypes

  = (number animals with shadows)/g = (PI $\times$ N)/g.

A population of N individuals, therefore, has $u + s$ different genotypes and $u + (g \times s)$ different animals. As an example, consider the case in which shadows are distributed pairwise ($g = 2$), total population size (N) is 100, and PI = 0.10. There would be 90 animals with unique genotypes ($u = 90$) and five shared genotypes ($s = 5$) common among 10 animals, leading to 95 different genotypes ($u + s$) in the population of 100 different animals ($u + (g \times s)$).

For the case in which genotypic shadows occur among pairs of animals ($g = 2$), a particular sample session could ''capture'' either, both, or neither indi-

vidual making up each pair of shadows. Thus, the LP-sufficient statistics must include both the expected number of unique genotypes (Eq. 6) and the expected number of animals that have shadows. The expected values of the LP statistics would then be:

$$E(n_1) = up_1 + s[2p_1(1 - p_1) + p_1^2] = up_1 + s(2p_1 - p_1^2)$$

$$E(m_2) = up_1p_2 + s(2p_1 - p_1^2)(2p_2 - p_2^2)$$

$$E(n_2) = E(m_2) + p_2[u(1 - p_1)] + s(2p_2 - p_2^2)(1 - p_1^2). \qquad (7)$$

As the number of individuals sharing each shadow increases ($g > 2$), the binomial probability terms associated with capture of shadows expand.

From the expected LP statistics, we calculated expected $\hat{N}$ (Eq. 4), Var($\hat{N}$) (Eq. 5), and standard error of the estimate (SE($\hat{N}$) = [Var($\hat{N}$)]$^{1/2}$; see Pollock et al. 1990) for shadow effects created by PI values ranging from 0 to 0.2 and for shadows distributed as pairs and as fours ($g = 2$, $g = 4$). In addition to evaluating population estimates ($\hat{N}$) under the shadow effect with different true population sizes ($N = 50$ and $N = 100$) and capture probabilities ($p = 0.25$ and $p = 0.5$), we also considered the relative bias of the estimate [($\hat{N} - N$)/SE($\hat{N}$)].

Multiple mark–recapture models in program CAPTURE.—If multiple mark–recapture samples are available, population size can be estimated with a number of models contained within program CAPTURE (Table 1; Otis et al. 1978, White et al. 1982, Rexstad and Burnham 1991). Capture data are a series of 1's and 0's; for DNA analysis, each genotype would have 0's for sample sessions in which it was not detected, and 1's for samples in which it was detected. Although population closure must still hold, the benefit of CAPTURE is its ability to estimate population size in the face of combinations of unequal trappability arising from trap response, individual capture heterogeneity, and changes in capture probability with time (Table 1). To select the proper estimator, CAPTURE contains a model selection procedure that performs chi-square tests to determine the forms of unequal trappability that probably exist in the data (behavior, heterogeneity, and time, plus multiple combinations of these), and computes normalized discriminant function scores such that the most likely model has a selection criterion of 1.00 (see Otis et al. 1978:56–60). Although the user can choose any model for estimation, in practice, the recommended models are those with a model selection criteria of 1.00.

We used CAPTURE's simulation routine to examine the effects of nonzero PI on: (1) mean model selection criterion; (2) the model chosen for estimation of population size; and (3) the mean relative bias of selected models. Random capture histories (0's and 1's) were created for 100 animals, replicated 20 times under M_0 (null model: no variation in capture probability) with

TABLE 1. Population size estimators contained within program CAPTURE (Otis et al. 1978, White et al. 1982, Rexstad and Burnham 1991) and used in analysis of shadow effects.

Model	Type of variation	Estimator
M_0	No variation in capture probabilities; assumes all animals are equally likely to be captured.	null
M_h	Each animal has a different, but constant, capture probability.	jackknife
M_b	Recapture of all individuals is affected by first capture; this is the ''trap happy'' or ''trap shy'' response.	Zippin
M_{bh}	Each animal has a different probability of capture, whereas all animals alter capture probability after first capture.	generalized removal
M_t	Capture probability varies over time; this variation is assumed to be the same for all animals.	Darroch
M_{th}	Each animal has a different capture probability and capture probabilities vary with time.	M_{th}-Chao
M_{tb}	Capture probabilities vary over time and after the first capture.	M_{tb}-Burnham
M_{tbh}	Each animal has a different probability of capture, whereas capture probability changes over time and after first capture.	no estimator available

six capture sessions, and a daily capture probability of 0.5. Capture histories of all 0's or all 1's were possible. Although actual field studies will often have fewer individuals, fewer capture sessions, and/or a lower capture probability, we wanted to limit the effects of inadequate sampling in our analysis of nonzero PI. It is well known that the model selection procedure within CAPTURE performs poorly when the daily capture probability, the total population size, or the number of trap sessions is low (Otis et al. 1978, White et al. 1982, Menkins and Anderson 1988). Because all of these variables are interrelated and dependent upon sources of capture variation, there are no clear rules of thumb. However, White et al. (1982) show that, with population sizes of 50 and fewer than 10 capture sessions, a capture probability of ~0.5 is necessary for CAPTURE's estimation and testing methods to perform well. Hence, our simulation parameters should be sufficient for CAPTURE to select a reasonable model, in the absence of the shadow effect.

We simulated the shadow effect with two levels of PI (0.08 and 0.20) and two levels of the number of individuals sharing a shadow genotype ($g = 2$, $g = 4$) by systematically combining the randomly generated capture histories. For example, for a PI of 0.20 and four individuals per genotype ($g = 4$), 20 animals in total have shadows (PI $\times$ N = 0.2 $\times$ 100 = 20), so we combined the first four capture histories to create the first shadow genotype, then the second four, and so on, until the five shared genotypes were created [s = (PI $\times$ N)/g = 20/4 = 5]. In this way, the 20 capture histories were combined for the five scenarios: PI = 0.0 (capture histories not combined; no shadows); PI = 0.08, $g = 2$; PI = 0.08, $g = 4$; PI = 0.20, $g = 2$; and PI = 0.20, $g = 4$.

To examine the effects of nonzero PI on the model selection criterion, we ran CAPTURE with each of the five levels of shadow effects and recorded the mean model selection criterion for each model. Although the

model selection criteria may change little, the actual model selected (the model whose selection criterion is 1.00) may vary greatly, thereby affecting which model the biologist actually uses to estimate population size. To determine the effects of shadow effects on the model selected, we recorded the number of times each model was chosen in each of the scenarios across the 20 replicates. Because competing models sometimes have equal selection criteria, we sometimes have >20 ''most likely'' models within 20 simulations of each scenario.

Finally, we evaluated the relative bias of each selected model. Even if program CAPTURE selects a different model when PI is nonzero, this does not necessarily indicate that the model selection procedure is failing. Biologists are more interested in whether or not the model selected can correctly estimate the population size. To address this, we estimated the population size using the preferred model in CAPTURE for each scenario, and calculated the mean and standard error of the relative bias in the estimator over the 20 simulations. When CAPTURE selected the heterogeneity model (M_h), we used the default heterogeneity estimator ($M_{h\text{-jackknife}}$). If capture probabilities were very low (approximately <0.2), Chao's (1988) alternative estimator for M_h or the approach of Rosenberg et al. (1995) might be more appropriate.

RESULTS: HOW GENETIC ''SHADOWS'' MIGHT AFFECT POPULATION ESTIMATES

Lincoln-Petersen (LP)

The shadow effect causes the Lincoln-Petersen estimator to be negatively biased. The bias arises because the shadow effect creates heterogeneity in apparent capture probability between animals with and without shadows, and heterogeneity leads to negative bias in LP (see Seber 1982, Pollock et al. 1990). If each individual animal has the same capture probability (as we assume here), then animals with unambiguous ge-
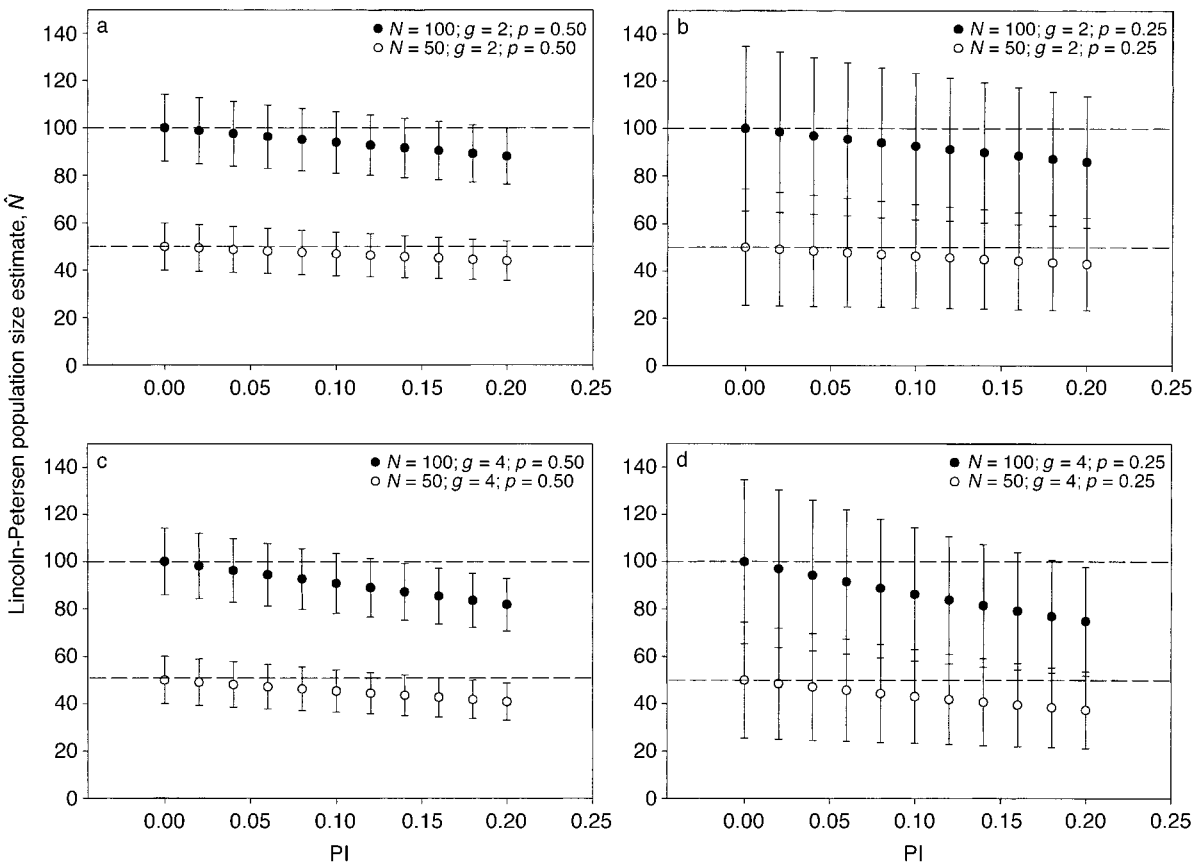
FIG. 1. Lincoln-Petersen estimate of population size ($\hat{N}$) and associated standard error for varying levels of probability of identity (PI) for two true population sizes ($N = 50$ and $N = 100$). Two levels of capture probability ($p = 0.25$, $p = 0.5$) and two levels of shadow effect, number of individuals per shared genotype ($g = 2$, $g = 4$) are considered in each panel. The dotted lines represent the true population sizes of 100 and 50.

notypes (u) will have a lower capture probability than the animals with shadows (gs), whose capture probability includes the probability of catching only one individual, or an indistinguishable shadow, or both the individual and its shadow, or more than one shadow if $g > 2$. This heterogeneity in capture between the (u) animals and the (gs) shadow animals increases when g increases, because the probability of at least one shadow being captured increases. With heterogeneity induced by the shadow effect, the LP ratio (m_2/n_2) overestimates the true proportion marked in the population, making the estimate of both population size and number of distinct genotypes too small. Unfortunately, one will almost never know what g is, so the amount of true bias will be hard to determine. Nevertheless, these results indicate that, with a shadow effect (i.e., $s > 0$), the LP estimator can be expected to underestimate the number of distinct genotypes in the population ($u + s$), as well as the number of animals ($u + gs = N$), such that

$$\hat{N} < u + s < u + gs = N.$$

One way to evaluate the importance of the shadow effect for ecological applications is to observe the change in $\hat{N}$ and relative bias across varying degrees of shadow effect (nonzero PI and g) for population sizes of 50 and 100 and capture probabilities of 0.25 and 0.5. When shadows are distributed pairwise ($g = 2$), $\hat{N}$ becomes negatively biased as PI increases (Fig. 1a, b). The estimated standard error, and therefore the estimated 95% confidence interval, also becomes smaller as PI increases (Fig. 1a, b). An increase in the number of animals sharing a genotype ($g = 4$) exacerbates both the absolute bias and the change in SE, with greater negative absolute bias and smaller standard errors when $g = 4$ (Fig. 1c, d) than when $g = 2$ (Fig. 1a, b).

Relative bias of the LP estimator is also affected by the shadow effect (Fig. 2). For either true population size ($N = 50$ or $N = 100$), a larger capture probability leads to a greater relative bias. Likewise, at a given capture probability (0.25 or 0.5), relative bias is greater when population size is larger (100 vs. 50). Thus, we have the unexpected finding that relative bias of LP under the shadow effect increases with increasing capture probability and true population size. Relative bias is worse with $g = 4$ (Fig. 2b) than with $g = 2$ (Fig. 2a).
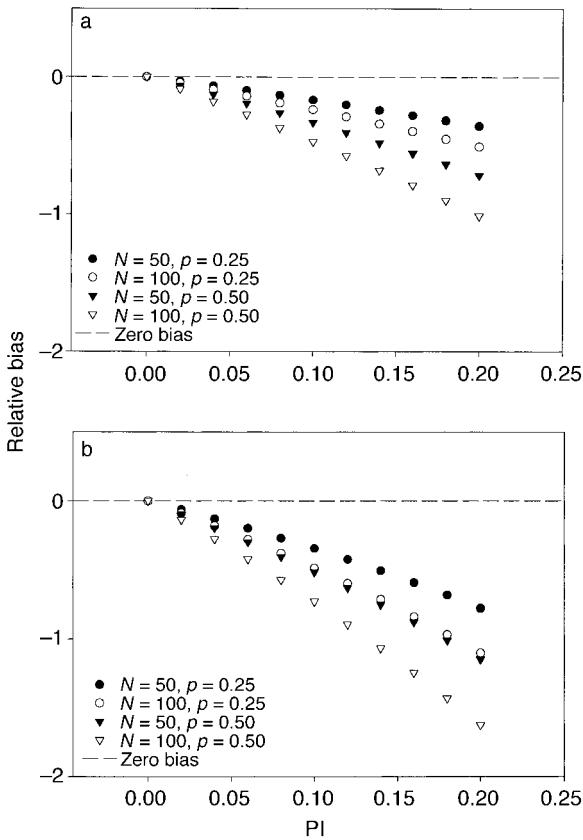
FIG. 2. Relative bias [$(\hat{N} - N)/\text{SE}_{\hat{N}}$] of Lincoln-Petersen estimates with varying levels of probability of identity (PI) and for different combinations of true population size (N) and capture probability, p. In panel (a), the number of animals per shared genotype is $g = 2$; in panel (b), $g = 4$.
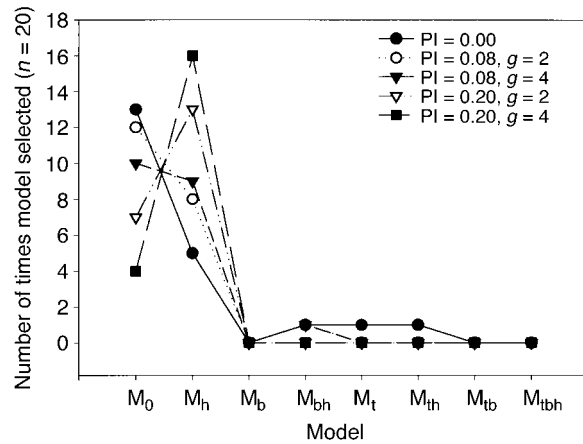

FIG. 3. Number of times a particular model in CAPTURE was chosen when the program was presented with 20 replicates of 100 randomly generated individual capture histories that were combined according to the magnitude of the shadow effect. The shadow effect is a function of PI (probability of identity) and g (number of individuals per shared genotype). Connecting lines are used to make it easier to follow changes in the number of times selected.

Multiple-recapture models (program CAPTURE)

The null model (M_0) has a high average model selection criteria across all levels of shadow effect, an expected result because M_0 was used to generate the simulations of capture data. Nevertheless, at higher levels of the shadow effect (PI = 0.20, $g = 2$ or 4), model $M_{\text{h-jackknife}}$ has the highest overall mean selection criterion across the 20 replicates (Table 2). The model with the highest mean selection criterion may not be the most appropriate model in any given replicate. In this case, however, $M_{\text{h-jackknife}}$ does become the most likely

model (selection criterion = 1.00) for the majority of replicates as the shadow effect increases (Fig. 3). Neither the increase in the mean selection criteria for $M_{\text{h-jackknife}}$ nor the fact that $M_{\text{h-jackknife}}$ is chosen more as the shadow effect increases is particularly surprising, given the fact that the shadow effect creates heterogeneity in the apparent ''capture'' probabilities for animals with distinct vs. shared genotypes. In other words, when the PI or number of individuals sharing a genotype is large, CAPTURE correctly identifies the shadow effect as a form of heterogeneity.

The key practical question is whether the identification of heterogeneity by CAPTURE, and use of a model robust to heterogeneity ($M_{\text{h-jackknife}}$), leads to unbiased estimates of population size, or at least the number of genotypes, in the face of shadow effects. As the shadow effect increases, the mean relative bias of population size estimate tends to increase, regardless of which model is chosen (Fig. 4). However, model $M_{\text{h-jackknife}}$ performed better than M_0 with respect to estimating true population size, with 95% confidence intervals (CI) for model $M_{\text{h-jackknife}}$ including the true pop-

TABLE 2. Average values of model selection criteria for 20 sets of simulated individual capture histories in CAPTURE, under five scenarios of varying degree of shadow effects.

Scenario		Model and selection criteria							
PI	g	M_0	M_h	M_b	M_{bh}	M_t	M_{th}	M_{tb}	M_{tbh}
0.00	⋯	0.95	0.87	0.35	0.61	0.05	0.49	0.37	0.72
0.08	2	0.92	0.91	0.32	0.58	0.02	0.47	0.35	0.67
0.08	4	0.92	0.91	0.34	0.58	0.01	0.46	0.35	0.66
0.20	2	0.89	0.95	0.34	0.56	0.00	0.46	0.34	0.66
0.20	4	0.85	0.97	0.34	0.54	0.00	0.43	0.33	0.64

Note: PI refers to the probability of identity, g is the number of individuals that share a shadow genotype, and the models are as described in Table 1.
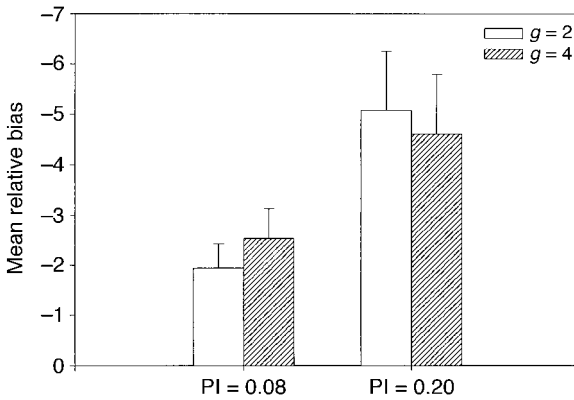
FIG. 4. Mean ($+1$ SE) relative bias [$(\hat{N} - N)/\mathrm{SE}_{\hat{N}}$] of the chosen estimator in CAPTURE for 20 replicates of 100 randomly generated individual capture histories that were combined according to the magnitude of the shadow effect. The shadow effect is as in Fig. 3.

ulation size for all of our simulations. In contrast, the 95% CI for M_0 never contained the true population size at the highest levels of shadow effect. Thus, the heterogeneity model gives an estimate of population size that is less negatively biased than model M_0. The heterogeneity model slightly overestimated the number of genotypes ($u + s$) for conditions that we considered, whereas M_0 tended to underestimate it.

Interestingly, at the highest shadow effect that we considered (PI = 0.2, $g = 4$), the relative bias of population size across the 20 replicates was actually less than it was at a slightly lower shadow effect (PI = 0.2, $g = 2$; Fig. 4). Relative bias is decreased because CAPTURE was better able to detect heterogeneity at higher levels of the shadow effect (Fig. 3). Because the less biased estimator ($M_{h\text{-jackknife}}$) was chosen more often at these high levels of the shadow effect, population estimates had lower relative bias.

DISCUSSION

Problems associated with estimating demographic rates in general, and population size in particular, have attracted sustained attention in the ecological literature (e.g., Marten 1970, Menkens and Anderson 1988, Minta and Mangel 1989, Lebreton et al. 1992, Wileyto et al. 1994, Rosenberg et al. 1995, Karanth and Nichols 1998). The difficulties are especially acute for taxa that are difficult to monitor, yet are under intense management scrutiny (i.e., Kendall et al. 1992, Zielinski and Stauffer 1996). Recently, emerging molecular techniques have been shown to effectively identify individuals in the field (Haig 1998), an advance that leads logically to application in estimating demographic variables. Although accounts advocating noninvasive sampling to estimate population size are only now beginning to be published in journals (Palsboll et al. 1997, Foran et al. 1997), it is difficult to find an applied ecology program or government agency that is not actively interested in these approaches (see, for example,

British Columbia Resources Inventory Committee 1997).

Genetic shadows are different individual organisms that are indistinguishable using noninvasive genetic analysis. Our results indicate that shadows can lead to error in population estimates beyond that arising from traditional mark–recapture studies. Therefore, studies using noninvasive sampling of organisms (via hair snags, feces, etc.) still must account for traditional concerns such as heterogeneity of capture probability, trap response, and change in capture probability with time (Pollock et al. 1990), as well as this novel "shadow effect," which manifests as a type of capture heterogeneity. The shadow effect can affect the population size estimate, the variance around the estimate, and the relative bias.

The effect of shadows on population size estimates will tend to be a negative bias. In the case of two-sample Lincoln-Petersen (LP) estimates, even the estimate of the number of different genotypes will tend to be biased low. The negative bias arises from the creation of an apparent heterogeneity of capture of animals of different genotypes. Therefore, the size of the bias depends on the size of the shadow effect, which, in turn, depends on both the PI and the number of indistinguishable genotypes for that PI. In some cases that are likely to be encountered in field studies, the negative bias is substantial. However, the bias is small in other cases; in fact, a strong argument might be made that the estimates are impressively accurate, given the logistical hurdles required to estimate population sizes for rare organisms such as carnivores!

The estimate of variance under the LP method is also affected by genetic shadows. An increased shadow effect arising from either increasing PI or increasing the number of animals sharing a shadow genotype (Fig. 1) improves the apparent precision by decreasing the estimated standard errors around the estimate, leading to the uncomfortable situation whereby our "certainty" in the estimate increases (smaller estimated 95% CI), even as our estimate gets progressively worse (more bias) at higher levels of the shadow effect. Therefore, the shadow effect results in the largest population with the largest capture probability actually having the *largest* estimated relative bias, and the smallest population with the lowest capture probability having the *smallest* estimated relative bias (Fig 2). The fact that the increased relative bias and increased apparent precision are exacerbated by increasing sample size through increased relative capture probability or true population size does not argue for decreasing the capture probability or sample size. Rather, our results should be taken as a warning that the shadow effect may not be minimized by the traditional solutions of increased sample size or capture probability.

Program CAPTURE, which uses multiple mark–recapture data to identify and use population estimation models robust to violations of the equal catchability

assumptions, is also affected by the presence of shadows. In our simulations using CAPTURE, the number of different genotypes was estimated well, but the parameter of greatest interest in ecological studies, population size, was estimated with a negative bias under the shadow effect. As with LP, the negative bias tended to increase as the shadow effect increased (Fig. 4). Interestingly, at our highest level of the shadow effect, the relative bias dropped because CAPTURE was better able to identify the heterogeneity created by the shadow effect and to appropriately recommend the less biased $M_{h\text{-jackknife}}$ model (Fig. 3). This result implies that moderate levels of shadow effect may be more damaging to population estimates in CAPTURE than high levels; at high levels of the shadow effect, CAPTURE more successfully identifies the heterogeneity and recommends the estimator with less negative bias.

It is not clear how a user of CAPTURE would distinguish between true capture heterogeneity, which arises from different animals being more or less likely to be noninvasively sampled, vs. the heterogeneity detected as a result of shadows being functionally more likely to be ''captured'' than individuals without genetic shadows. Note that the shadow effect does not manifest as a positive (trap happy) trap response because the probability of capture increases for all trap events, not just after the first capture. Of course, the putative ''capture heterogeneity'' resulting from the shadow effect will interact with other sources of variation expected in any mark–recapture study.

We did not consider all available closed-population estimators. For example, ''mark–resight'' approaches (Arnason et al. 1991, Gardner and Mangel 1996, White 1996) might be based on captures of animals (physically marking and obtaining a DNA sample) followed by ''resighting'' (noninvasive DNA collection). Such approaches have the advantage of minimizing bias by making sources of capture heterogeneity independent (Seber 1982), but the shadow effect will continue to affect the second capture session and, therefore, the estimate. Similarly, although we did not examine more complex estimators of demographic rates, such as open-population models or estimates of survival or recruitment, it is probably true that these will also be impacted by the shadow effect if noninvasive genetic sampling is used. For example, shadow effects are likely to positively bias survival estimates, because mortality of an individual could be disguised by its surviving shadow(s).

We emphasize that, in any given case, the actual effect of genetic shadows on $\hat{N}$ could be far greater or less than the limited sets of conditions that we considered. For example, both PI and the number of shadows sharing a genotype (g) might be greater than the ranges that we considered if siblings tend to be caught, or if inbreeding has occurred due to assortative mating or small population size. Shadow effects may be most problematic in taxa such as carnivores, where the very
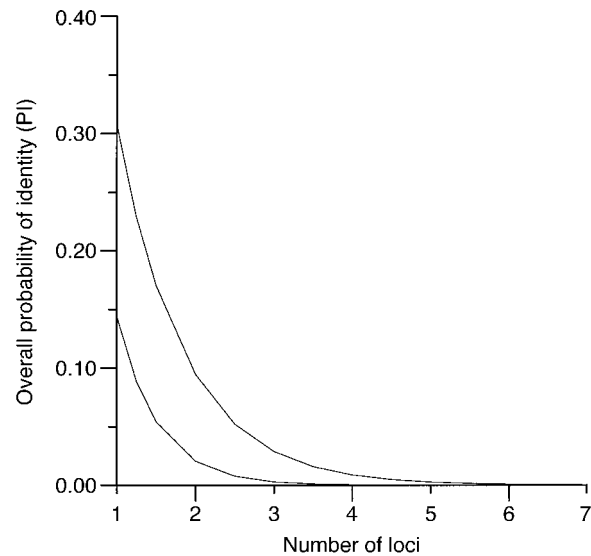

Fig. 5. Overall probability of identity (PI) of individuals in a population as the number of loci changes. The top and bottom lines are based, respectively, on the average PI across populations having the highest or lowest PI in each of the 10 studies surveyed in the Appendix (mean PI per locus is 0.31 for the upper line and 0.14 for the lower line). Loci are assumed to be independent in their impact on PI, as per Eq. 2.

attributes that force us to consider noninvasive sampling (e.g., low population density) also lead to decreased heterozygosity, increased PI, and increased number of individuals per shared genotype. In an analysis using actual wolf and bear data sets, L. P. Waits, G. Luikart, J.-M. Cornuet, and P. Taberlet (*unpublished data*) found that social structure can lead to an actual PI up to three orders of magnitude larger than that calculated using the traditional formulas of Eqs. 1 and 2 (see also Taberlet and Luikart 1999).

On the other hand, PI could certainly be reduced, most readily if more loci are used. The assumption of independence among loci leads to exponential decrease in PI with more loci (Eq. 2). Fig. 5 shows the change in PI using varying number of loci, based on the mean PI per locus calculated from the ''high PI'' and ''low PI'' populations in the Appendix. With seven independent loci or more, the PI becomes very small, and the problems that we discuss here are likely to be minimized.

Finally, the problems that the shadow effect could create for population estimates are not escaped by the use of an index such as ''enumeration'' or ''Minimum Number Alive,'' MNA (see Krebs 1966, Krebs et al. 1969). Although the MNA index is simple and widely used, it is biased under virtually all field conditions (see Hilborn et al. 1976, Nichols and Pollock 1983). The shadow effect will further bias the MNA index by an amount equal to the number of shared genotypes ($s = (\text{PI} \times N)/g$). Population measures using noninvasive DNA sampling should therefore be based on estimators (such as mark–recapture models) and not on biased

indices with unknown detectability that are further biased by the shadow effect.

In summary, all demographic estimates using noninvasive DNA sampling could be affected by a shadow effect. The bias in population size estimates typically will be negative, but the nearly infinite array of possible magnitudes and manifestations of the probability of identity prevent quantitative generalizations about the specific effects of shadows. Our results do not diminish the value of noninvasive genetic techniques for uses such as estimating the identity of species, sex ratio, or presence/absence metrics. Understanding the potential effects of DNA shadows on estimating demographic parameters moves applied ecologists one step closer to confidently embracing the promise of molecular tools for gaining insights into the demography of seldom-seen animals.

Literature Cited

Allen, P. J., W. Amos, P. P. Pomeroy, and S. D. Twiss. 1995. Microsatellite variation in grey seals (*Halichoerus grypus*) shows evidence of genetic differentiation between two British breeding colonies. Molecular Ecology **4**:653–662.

Arnason, A. N., C. J. Schwartz, and J. M. Gerrard. 1991. Estimating closed population size and number of marked animals from sighting data. Journal of Wildlife Management **55**:716–730.

Avise, J. C. 1994. Molecular markers, natural history and evolution. Chapman and Hall, London, UK.

Avise, J. C., S. M. Haig, O. A. Ryder, M. Lynch, and C. J. Geyer. 1995. Descriptive genetic studies: applications in population management and conservation biology. Pages 183–244 *in* J. D. Ballou, M. Gilpin, and T. J. Foose, editors. Population management for survival and recovery. Columbia University Press, New York, New York, USA.

Bailey, N. T. J. 1951. On estimating the size of mobile populations from recapture data. Biometrika **38**:293–306.

Boyce, W. M., P. W. Hedrick, N. E. Muggli-Cockett, S. Kalinowski, M. C. T. Penedo, and R. R. Ramey II. 1997. Genetic variation of a major histocompatibility complex and microsatellite loci: a comparison in bighorn sheep. Genetics **145**:421–433.

British Columbia Resources Inventory Committee. 1997. Draft Standardized Inventory Methodologies for components of British Columbia's biodiversity. Victoria, British Columbia, Canada.

Bruford, M. W., D. J. Cheesman, T. Coote, H. A. A. Green, S. A. Haines, C. O'Ryan, and T. R. Williams. 1996. Microsatellites and their application to conservation genetics.

Pages 278–297 *in* T. B. Smith and R. K. Wayne, editors. Molecular genetic approaches to conservation. Oxford University Press, Oxford, UK.

Chao, A. 1988. Estimating animal abundance with capture frequency data. Journal of Wildlife Management **52**:295–300.

Chapman, D. G. 1951. Some properties of the hypergeometric distribution with applications to zoological censuses. University of California Publications in Statistics **1**:131–160.

Cohen, J. E. 1990. DNA fingerprinting for forensic identification: potential effects on data interpretation of subpopulation heterogeneity and band number variability. American Journal of Human Genetics **46**:358–368.

Donnelly, P. 1995. Nonindependence of matches at different loci in DNA profiles: quantifying the effect of close relatives on the match probability. Heredity **75**:26–34.

Foran, D. R., S. C. Minta, and K. S. Heinemeyer. 1997. DNA-based analysis of hair to identify species and individuals for population research and monitoring. Wildlife Society Bulletin **25**:840–847.

Gardner, S. N., and M. Mangel. 1996. Mark–resight population estimation with imperfect observations. Ecology **77**:880–884.

Garshelis, D. L., and L. G. Visser. 1997. Enumerating megapopulations of wild bears with an ingested biomarker. Journal of Wildlife Management **61**:466–480.

Haig, S. 1998. Molecular contributions to conservation. Ecology **79**:413–425.

Hilborn, R., J. A. Redfield, and C. J. Krebs. 1976. On the reliability of enumeration for mark and recapture census of voles. Canadian Journal of Zoology **54**:1019–1024.

Houlden, B. A., P. England, and W. B. Sherwin. 1996. Paternity exclusion in koalas using hypervariable microsatellites. Journal of Heredity **87**:149–151.

Jarne, P., and P. J. L. Lagoda. 1996. Microsatellites from molecules to populations and back. Trends in Ecology and Evolution **11**:424–429.

Karanth, K. U., and J. D. Nichols. 1998. Estimation of tiger densities in India using photographic captures and recaptures. Ecology **79**:2852–2862.

Kendall, K. C., L. H. Metzgar, D. A. Patterson, and B. M. Steele. 1992. Power of sign surveys to monitor population trends. Ecological Applications **2**:422–430.

Kohn, M. H., and R. K. Wayne. 1997. Facts from feces revisited. Trends in Ecology and Evolution **12**:223–227.

Kohn, M. H., E. C. York, D. A. Kamradt, G. Haught, R. M. Sauvajot, and R. K. Wayne. 1999. Estimating population size by genotyping faeces. Proceedings of the Royal Society of London, Series B **266**:1–7.

Krebs, C. J. 1966. Fluctuating populations of *Microtus californicus.* Ecological Monographs **36**:239–273.

Krebs, C. J., B. L. Keller, and R. H. Tamarin. 1969. *Microtus* population biology: demographic changes in fluctuating populations of *M. ochrogaster* and *M. pennsylvanicus* in southern Indiana. Ecology **50**:587–607.

Lade, J. A., N. D. Murray, C. A. Marks, and N. A. Robinson. 1996. Microsatellite differentiation between Phillip Island and mainland Australian populations of the red fox *Vulpes vulpes.* Molecular Ecology **5**:81–87.

Lebreton, J. D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monographs **62**:67–118.

Levy, S. 1999. Geneticists go wild. BioScience **49**:5–8.

Lewontin, R. C., and D. L. Hartl. 1991. Population genetics in forensic DNA typing. Science **254**:1745–1750.

Mace, R. D., S. C. Minta, T. L. Manley, and K. E. Aune. 1994. Estimating grizzly bear population size using camera sightings. Wildlife Society Bulletin **22**:74–83.

Marten, G. G. 1970. A regression method for mark–recapture estimation of population size with unequal catchability. Ecology **51**:291–295.

Menkens, G. E., Jr., and S. H. Anderson. 1988. Estimation of small-mammal population size. Ecology **69**:1952–1959.

Minta, S., and M. Mangel. 1989. A simple population estimate based on simulation for capture–recapture and capture–resight data. Ecology **70**:1738–1751.

Morin, P. A., and D. S. Woodruff. 1996. Noninvasive genotyping for vertebrate conservation. Pages 298–313 *in* T. B. Smith and R. K. Wayne, editors. Molecular genetic approaches in conservation. Oxford Press, New York, New York, USA.

Moritz, C., A. Heideman, E. Geffen, and P. McRae. 1997. Genetic population structure of the greater bilby *Macrotis lagotis,* a marsupial in decline. Molecular Ecology **6**:925–936.

Mundy, N. I., C. S. Winchell, T. Burr, and D. S. Woodruff. 1997. Microsatellite variation and microevolution in the critically endangered San Clemente Island Loggerhead Shrike. Proceedings of the National Academy of Sciences (USA) **264**:869–875.

National Research Council. 1996. The evaluation of forensic DNA evidence. National Academy Press, Washington, D.C., USA.

Nichols, J. D. 1992. Capture–recapture models: using marked animals to study population dynamics. BioScience **42**:94–102.

Nichols, J. D., and K. H. Pollock. 1983. Estimation methodology in contemporary small mammal capture–recapture studies. Journal of Mammalogy **64**:253–260.

Nichols, R. A., and D. J. Balding. 1991. Effects of population structure on DNA fingerprint analysis in forensic science. Heredity **66**:297–302.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference for capture data on closed animal populations. Wildlife Monographs **62**:1–135.

Paetkau, D., W. Calvert, I. Stirling, and C. Strobeck. 1995. Microsatellite analysis of population structure in Canadian polar bears. Molecular Ecology **4**:347–354.

Paetkau, D., and C. Strobeck. 1994. Microsatellite analysis of genetic variation in black bear populations. Molecular Ecology **3**:489–495.

Paetkau, D., L. P. Waits, P. L. Clarkson, L. Craighead, E. Vyse, R. Ward, and C. Strobeck. 1998. Variation in genetic diversity across the range of North American brown bears. Conservation Biology **12**:418–429.

Palsboll, P. J., et al. 1997. Genetic tagging of humpback whales. Nature **388**:767–769.

Parker, P. G., A. A. Snow, M. D. Schug, G. C. Booton, and P. A. Fuerst. 1998. What molecules can tell us about populations: choosing and using a molecular marker. Ecology **79**:361–382.

Pollock, K. H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical inference for capture–recapture experiments. Wildlife Monographs **107**:1–97.

Pope, L. C., A. Sharp, and C. Moritz. 1996. Population struc-ture of the yellow-footed rock-wallaby *Petrogale xanthopus* (Gray, 1854) inferred from mtDNA sequences and micro-satellite loci. Molecular Ecology **5**:629–640.

Rexstad, E., and K. Burnham. 1991. Users' guide for interactive program CAPTURE: abundance estimation of closed animal populations. Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University, Fort Collins, Colorado, USA.

Roeder, K. 1994. DNA fingerprinting: a review of the controversy. Statistical Science **9**:222–278.

Rosenberg, D. K., W. S. Overton, and R. G. Anthony. 1995. Estimation of animal abundance when capture probabilities are low and heterogeneous. Journal of Wildlife Management **59**:252–261.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. Macmillan Publishing, New York, New York, USA.

Snow, A. A., and P. G. Parker. 1998. Molecular markers for population biology. Ecology **79**:359–360.

Taberlet, P., and G. Luikart. 1999. Non-invasive sampling and individual identification. Proceedings of the Royal Society of London Series B **14**:323–327.

Taberlet, P., and L. P. Waits. 1998. Non-invasive genetic sampling (letter to editor). Trends in Ecology and Evolution **13**:26–27.

Taberlet, P., L. P. Waits, and G. Luikart. 1999. Non-invasive genetic sampling: look before you leap. Trends in Ecology and Evolution **14**:323–327.

Taylor, A. C., W. B. Sherwin, and R. K. Wayne. 1994. Genetic variation of microsatellite loci in a bottlenecked species: the northern hairy-nosed wombat *Lasiorhinus krefftii.* Molecular Ecology **3**:277–290.

Westneat, D. F., P. C. Frederick, and R. H. Wiley. 1987. The use of genetic markers to estimate the frequency of successful alternative reproductive tactics. Behavioral Ecology and Sociobiology **21**:35–45.

White, G. C. 1996. NOREMARK: population estimation from mark–resighting surveys. Wildlife Society Bulletin **24**:50–52.

White, G. C., D. R. Anderson, K. P. Burnham, and D. L. Otis. 1982. Capture–recapture and removal methods for sampling closed populations. Los Alamos National Laboratory, LA 8787-NERP, Los Alamos, New Mexico, USA.

Wileyto, E. P., W. J. Ewens, and M. A. Mullen. 1994. Markov-recapture population estimates: a tool for improving interpretation of trapping experiments. Ecology **75**:1109–1117.

Woods, J. G., B. McLellan, D. Paetkau, C. Strobeck, and M. Proctor. 1996. DNA fingerprinting applied to mark–recapture bear studies. International Bear News **5**:9–10.

Yoccoz, N. G., H. Steen, R. A. Ims, and N. C. Stenseth. 1993. Estimating demographic parameters and the population size: an updated methodological survey. Pages 565–587 *in* N. C. Stenseth and R. I. Ims, editors. The biology of lemmings. Academic Press, San Diego, California, USA.

Zielinski, W. J., and H. B. Stauffer. 1996. Monitoring *Martes* populations in California: survey design and power analysis. Ecological Applications **6**:1254–1267.

APPENDIX

Studies of non-captive birds and mammals that used microsatellite DNA and provided sufficient information to determine the average probability of identity (PI) per amplified locus.

Taxa (citation) and location	No. polymorphic loci	Mean PI per locus†	Heterozygosity‡
Loggerhead Shrike, *Lanius ludovicianus* (Mundy et al. 1997)			
San Clemente Island, *L. l. mearnsi*	6	0.45	0.39
Perris mainland, *L. l. gambeli*	7	0.28	0.56
Polar bear, *Ursus maritimus* (Paetkau et al. 1995)			
South Beaufort Sea	8	0.15§	0.64
Davis Strait	8	0.18§	0.61
Black bear, *Ursus americanus* (Paetkau and Strobeck 1994)			
Terra Nova	4	0.51§	0.36
Banff	4	0.068§	0.80
Grey seal, *Holichoerus grypus* (Allen et al. 1995)			
North Rona	8	0.11	0.74
Isle of May	8	0.11	0.73
Hairy-nosed wombat, *Laxiorhinus* spp. (Taylor et al. 1994)			
NHN, *L. krefftii*	9	0.38	0.47
SHN combined, *L. latifrons*	15	0.17	0.68
Bighorn sheep, *Ovis canadensis* (Boyce et al. 1997)			
San Andres	2	0.35	0.49
Eagle	3	0.11	0.75
Bilby, *Macrotis lagotis* (Moritz et al. 1997)			
Queensland	9	0.11	0.73
Northern Territories	9	0.078	0.79
Koala, *Phascolarctos cinereus* (Houlden et al. 1996)			
Kangaroo Island	4	0.39	0.48
Nowendoc	6	0.061	0.82
Yellow-footed rock wallaby, *Petrogale xanthopus* (Pope et al. 1996)			
Lisburne	4	0.18	0.66
Hill of Knowledge	4	0.12	0.72
Red fox, *Vulpes vulpes* (Lade et al. 1996)			
Philip Island	7	0.45	0.39
Mornington	7	0.26	0.58

Notes: PI represents the proportion of animals in the population that would be expected to share a genotype with at least one other animal. We only included studies with at least two populations to bracket ''high'' and ''low'' PI values for each study.

† Mean PI is the geometric mean of the overall PI across all polymorphic loci for a particular population (see Eq. 2).

‡ Heterozygosity is the arithmetic mean heterozygosity across loci, and is inversely correlated with PI (Eq. 1).

§ Provided by the authors (who used Eq. 1). For all other studies, we calculated PI from allele frequencies of each locus using Eq. 1 (text).