

# Review of DNA-based census and effective population size estimators

Michael K. Schwartz<sup>1</sup>, David A. Tallmon<sup>2</sup> and Gordon Luikart<sup>2</sup>

<sup>1</sup>Wildlife Biology Program, School of Forestry, University of Montana, Missoula MT 59812, USA

<sup>2</sup>Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA

(Received 24 June 1997; accepted 15 July 1998)

## Abstract

The detection of reductions in effective population size ( $N_e$ ) or census size ( $N_c$ ) is essential for conservation. Recent developments allow wildlife researchers to obtain genetic material *via* non-invasive sampling techniques that may provide the large sample sizes necessary for precise estimates of  $N_e$  and  $N_c$ . Population genetic theory provides several methods to estimate  $N_e$  from allele frequency data: including temporal change in allele frequencies, gametic disequilibrium and heterozygote excess methods. Modification of capture–mark–recapture methods for use with multi-locus genotype data provides new means for estimating  $N_c$ . The combination of new DNA sampling techniques, polymerase chain reaction-based DNA markers and analytical methods may provide unprecedented power to detect reductions in  $N_e$  and  $N_c$  of endangered populations. However, these genetic methods are largely untested in the field. We review some relatively unexplored, but promising ways that multi-locus genetic data can be used to provide important genetic and demographic information and suggest avenues for further research in this area.

## INTRODUCTION

A primary culprit behind the decline of biodiversity is anthropogenic habitat reduction and fragmentation (Caughley & Gunn, 1996), which isolates populations while simultaneously decreasing local census population sizes ( $N_c$ ). A decrease in  $N_c$  is accompanied by a decrease in the effective population size ( $N_e$ ), defined as the size of an ideal population experiencing the same rate of genetic change as the natural population of interest (Crow & Kimura, 1970). It has been suggested that  $N_e$  may often be as small as 10% of  $N_c$  (Frankham, 1995a).  $N_e$  is important because it determines rates of loss of genetic variation, fixation of deleterious alleles and inbreeding (Wright, 1969). At small population sizes, positive feedback between reduced population size and inbreeding depression accelerates population extirpation (e.g. Newman & Pilson, 1997). Therefore, early detection of an  $N_c$  and/or  $N_e$  reduction is critical, because immediate management actions may be necessary to avoid population endangerment or extinction.

Some authors suggest  $N_e$  estimates be included in ranking the vulnerability of species or populations to

extinction (Mace & Lande, 1991; Allendorf *et al.*, 1997). However,  $N_e$  has proven difficult to estimate for wildlife populations, even though it can be estimated using demographic or genetic methods. Demographic methods incorporate data for important life history parameters, such as variance in reproductive success, into analytical equations to estimate  $N_e$  (e.g. Nunney & Elam, 1994). Yet, the required demographic data are often difficult (Husband & Barrett, 1992) and expensive (Frankham 1995b) to obtain for wild populations, and these demographic estimators frequently overestimate  $N_e$  (Harris & Allendorf, 1989; Frankham, 1995a).

In contrast to the extensive data necessary for demographic estimates of  $N_e$ , statistical advances now allow us to estimate  $N_e$  from one or two large genetic samples. In this review we describe three analytical methods used to estimate  $N_e$  from genetic data. These are the following: (1) gametic disequilibrium, (2) heterozygote excess, and (3) temporal change in allele frequency methods. Each transforms allele frequency data into estimates of  $N_e$ . We evaluate their potential use in conservation biology based on practical and theoretical limitations of each approach. Subsequently, we review a method that has been used to estimate the minimum number of individuals alive and then conclude by reporting novel techniques currently being explored to estimate  $N_c$  from genetic data.

All correspondence to: Michael K. Schwartz.

E-mail: mks@selway.umt.edu

Present Address: Gordon Luikart, Laboratoire de Biologie des Populations d'Altitude, CNRS UMR 5553, Université Joseph Fourier, BP 53, F-38041 Grenoble Cedex 9, France.

### PCR-based molecular markers

Two advances in molecular genetics hold great promise for the application of genetic markers to the estimation of  $N_e$  or  $N_c$  of wildlife populations. These are: (1) the development of highly polymorphic DNA markers, and (2) the ability to amplify these markers with the polymerase chain reaction (PCR) from low-quality, low-quantity DNA samples. In particular, the use of microsatellites, a class of highly variable, single locus, genetic markers has proven valuable in assessing genetic variation at the population level (e.g. Bruford & Wayne, 1993). We focus on the use of these markers for  $N_e$  and  $N_c$  estimation, but also recognize that other nuclear, sex-specific, and mitochondrial DNA markers can yield important genetic and demographic information (Reed *et al.*, 1997; Taberlet, Camarra *et al.*, 1997).

It has recently been shown that microsatellite genotypes can be determined using samples taken from hair, skin, saliva, feather, feces, or urine (for a review, see Morin & Woodruff, 1996). This means non-invasive sampling techniques, in which direct contact (visual or otherwise) between researchers and animals is avoided, may provide sufficient DNA for the genetic and demographic analyses outlined below. Perhaps one of the greatest advantages of non-invasive genetic sampling is that it will provide the large sample sizes needed to make more precise estimates of  $N_e$  or  $N_c$ . Additionally, these markers allow PCR amplification of DNA from museum specimens. For instance, Mundy *et al.* (1997) compared DNA from 30 loggerhead shrikes collected from San Clemente Island in 1915 (museum specimens) to those collected in the 1990s to quantify changes in genetic variation.

However, when relying on microsatellite genotypes amplified from low-quality, low-quantity DNA it is important to acknowledge potential pitfalls, such as (1) allelic dropout (Gerloff *et al.*, 1995; Pemberton *et al.*, 1995; Taberlet, Griffin *et al.*, 1996; Gagneux, Boesch & Woodruff, 1997) and (2) PCR generated 'false-alleles' (Taberlet, Camarra *et al.*, 1997). To minimize these problems several laboratories have recommended repeating extractions (in the case of shed hair; Gagneux *et al.*, 1997) or a 'multiple-tubes approach' where DNA is extracted and subsequently distributed among several tubes to be amplified separately by PCR (Taberlet *et al.*, 1996, 1997).

### ESTIMATES OF $N_e$ FROM GENETIC DATA

Each of the estimators of  $N_e$  is based on the assumption that genetic drift increases when  $N_e$  decreases. Implicit with this assumption is that selection, mutation, population subdivision and migration do not change gene frequencies within a population. Microsatellite markers are less likely to be under selection than isozymes as they do not code for proteins. However, some microsatellite loci may be linked to regions of the genome under selection. Mutation is unlikely to be problematic unless the mutation rates are very high (i.e.  $\gg 10^{-3}$ ). Migration and

substructure assumptions, on the other hand, are likely to be violated in many natural populations (for a thorough discussion, see Waples, 1989, 1991), but are less likely to be violated in small, endangered populations that are often isolated.

Two of these estimators, linkage disequilibrium and heterozygote excess, actually assess  $N_{eb}$  (the effective number of breeding adults).  $N_{eb}$  is nearly equivalent to  $N_e$  in populations with non-overlapping generations.  $N_e$  can be estimated from  $N_{eb}$  in populations with overlapping generations if generation length, age class specific birth and death rates, and the mating system of the species is well known (Waples, 1991; Jorde & Ryman, 1995; Scribner, Arntzen & Burke, 1997).

### One sample: gametic disequilibrium and heterozygote excess methods

The first  $N_e$  estimation method we review is gametic disequilibrium ( $D$ ), also known as linkage disequilibrium. This is simply the non-random association of alleles at different loci.  $D$  can be generated by many factors, including natural selection (Vrijenhoek, Pfeiler & Wetherington, 1992), hybridization or mixing of differentiated gene pools (Forbes & Allendorf, 1991), and genetic drift (Waples, 1991), and can occur between linked and unlinked loci. Here, we assume observed  $D$  is produced by drift in a small population among unlinked loci and focus on how this relationship can be exploited to estimate  $N_e$ . The validity of this assumption will depend upon factors such as the number of loci used, chromosome number and potential for hybridization, and should be carefully considered for each population.

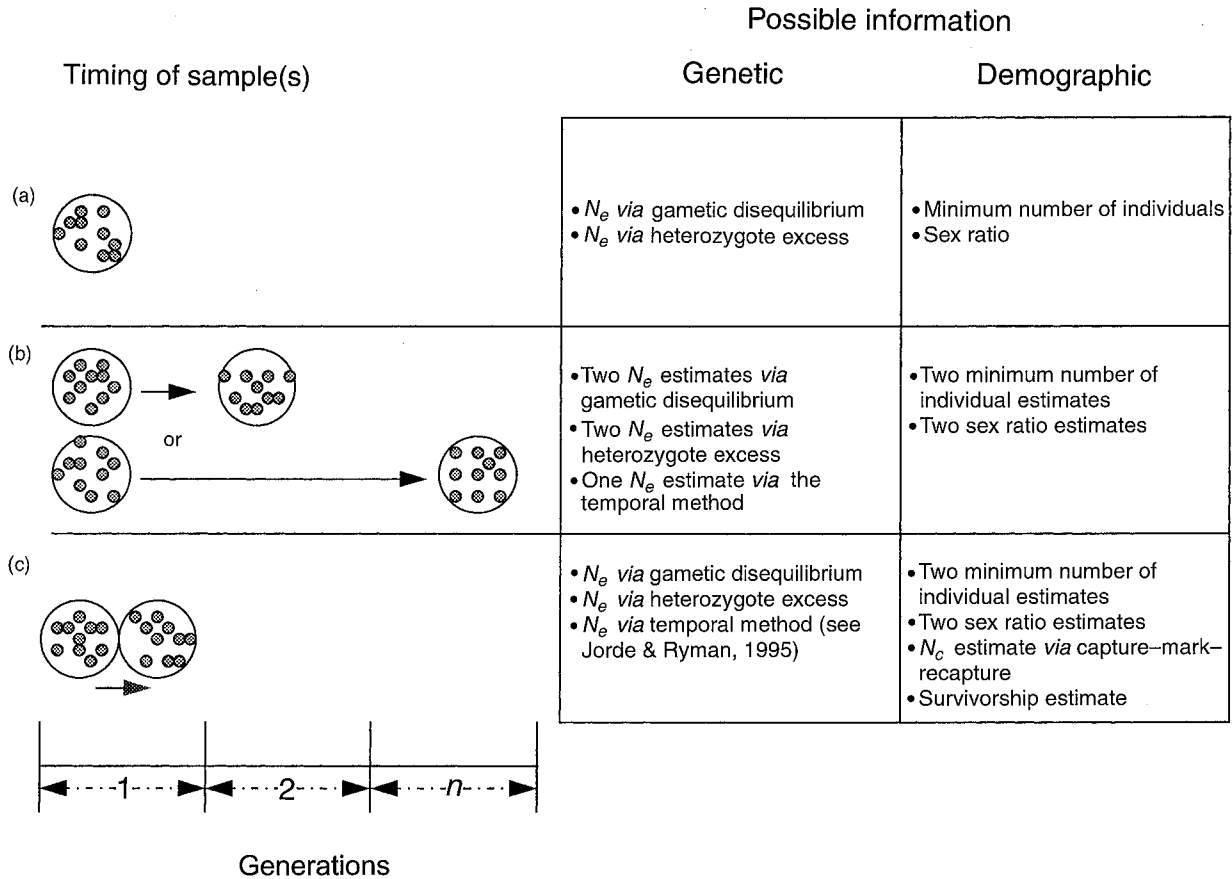
Just as genetic drift in a small population causes allele frequencies to change at individual loci, drift generates non-random associations among alleles at different loci. For example, with two genes,  $D$  is the difference between the observed frequency of co-occurrence of alleles ( $A_1, B_2$ ) at two loci (1 and 2) and that expected based on their allele frequencies ( $p_1, q_2$ ), assuming random association of gametes (from Hill, 1981).

$$D = \text{freq}(A_1, B_2) - p_1q_2 \quad (1)$$

As  $N_e$  decreases, drift plays an increasingly large role in determining allele frequencies, causing observed and expected frequencies of combinations of alleles from different loci to diverge and  $D$  to increase.  $D$  is closely related to the correlation,  $r$ , among alleles at different loci ( $r = D/[(p_1, q_1, p_2, q_2)^{1/2}]$ ). The relationship between  $r$  and  $N_e$  has been determined by Hill (1981) and Waples (1991):

$$N_e = \frac{1}{\left[3\left(r^2 - \frac{1}{S}\right)\right]} \quad (2)$$

where  $S$  is the number of individuals sampled. Using this method, Bartley *et al.* (1992) found the  $N_e$  of an



**Fig. 1.** Summary of genetic and demographic information that can be obtained from one or two population (DNA) samples. Each large open circle represents a population sample. Small, filled circles are individuals from the population sample. (a) One sample. (b) Two samples taken either in sequential generations or after two generations. Waiting two time intervals increases precision in the temporal method. (c) Two samples taken within one time interval.

endangered chinook salmon population to be much smaller than the census size;  $\hat{N}_e = 85.5$  (95% confidence interval (CI): 45 – 266) compared to an  $\hat{N}_c$  value of  $\cong 2000$ . However, these authors also provide an example where  $D$  failed to yield a reasonable estimate of  $\hat{N}_e$  due to an insufficient number of individuals being sampled and too few polymorphic markers.

Despite the obvious advantage that gametic disequilibrium can be used to estimate  $N_e$  with a single sample (Fig. 1), there appears to be a number of drawbacks associated with estimating  $N_e$  from  $D$  (Table 1). Foremost among the disadvantages is that  $>6$  polymorphic loci and  $S > 90$  may be necessary to obtain reasonably precise  $N_e$  estimates (Waples, 1991; Bartley *et al.*, 1992). Recent advances in non-invasive genetic sampling can increase the ability to obtain samples from endangered or secretive organisms, but  $S > 90$  may still be problematic for (1) endangered species with extremely small population sizes (i.e. 90 may exceed  $N_c$ ), and (2) species dispersed over large geographic areas or rugged terrain that will require extensive sampling efforts. However, we emphasize that this method may be most useful with organisms that produce large numbers of offspring, such as anadromous fish or marine invertebrates.

There are other inherent problems in using  $D$  to estimate  $N_e$ . For instance,  $D$  estimates are allele frequency dependent (Hedrick, 1987) and high allelic diversity may not increase power to detect small  $N_e$ , since only a single allele at each locus is independent with this estimator. To overcome the independence problem, allele frequencies at loci with greater than two alleles may be collapsed into two categories (i.e. one allele *versus* all others). A thorough investigation of the accuracy and precision of this method when using various sample sizes, numbers of marker loci and population sizes is needed.

A second method that requires only one sample to estimate  $N_e$  is the heterozygote excess method (Pudovkin, Zaykin & Hedgecock, 1996). This is based on the observation that small populations of dioecious species differ in genotype frequencies between sexes and deviate from overall Hardy-Weinberg (HW) proportions (Falconer & Mackay, 1996, pp. 62–63). In other words, when the breeding population is small, binomial sampling error produces allele frequency differences between male and female breeders, resulting in an excess of heterozygotes in their progeny. By quantifying the magnitude of heterozygote excess in these progeny, one can estimate  $N_e$  as follows:

**Table 1.** Summary of the advantages and disadvantages of using each  $N_e$  and  $N_c$  estimator/enumeration technique

|       | Estimator                              | Advantage   | Disadvantage   |
|-------|--|---|--|
| $N_e$ | Gametic disequilibrium                 | One sample needed   | Large sample size required<br>Greater than six loci required<br>Accuracy and precision for a given sample size and number of loci unknown  |
|       | Heterozygote excess                    | One sample needed   | Large confidence intervals around the estimates<br>Accuracy and precision for a given sample size and number of loci unknown<br>Only useful with breeding systems that have random unions of gametes |
|       | Temporal allele method                 | Sample size, number of loci needed, and power known<br>Bias and accuracy known<br>Many example us use to date | Two samples are needed   |
| $N_c$ | Minimum number of individuals          | One sample is needed  | Number of markers needed is unknown  |
|       | Open and closed capture-mark-recapture | Conventional capture-mark-recapture techniques are well developed and can be modified                         | Number of markers needed is unknown<br>Two samples are needed<br>Few published examples  |

$$\hat{N}_e = \frac{1}{2 \frac{H_{\text{obs}} - H_{\text{exp}}}{H_{\text{exp}}}} + \frac{1}{2 \left( \frac{H_{\text{obs}} - H_{\text{exp}}}{H_{\text{exp}}} + 1 \right)} \quad (3)$$

where  $H_{\text{obs}}$  is the proportion of heterozygotes observed in the sample and  $H_{\text{exp}}$  is the expected heterozygosity from sample allele frequency data given HW proportions (i.e.  $2pq$  for a diallelic locus; Pudovkin *et al.*, 1996, see their appendix). If sample sizes are small  $H_{\text{exp}}$  must be estimated using an unbiased estimator (Nei, 1987), such as  $2pq(2n/2n-1)$ , where  $n$  is the number of individuals sampled. For multiallelic loci, Pudovkin *et al.* (1996) sum the excess of heterozygotes for each allele (e.g.  $2pq + 2pr$  for the  $p$  allele in a triallelic system) and divide this sum by the number of alleles at a locus to estimate  $(H_{\text{obs}} - H_{\text{exp}})/H_{\text{exp}}$ . This value is averaged over all independent loci to estimate  $N_e$ .

Using eqn (3), allele frequencies from a single cohort (i.e. individuals of the same age) can be used to estimate  $N_e$  when the progeny are produced by random union of gametes (Pudovkin *et al.*, 1996). The primary advantage of this method, as with gametic disequilibrium, is that only one sample is required to estimate  $N_e$  (*cf.* temporal allele method section; Fig. 1 and Table 1).

Unfortunately, this method may only be valid when a species has a breeding system where there is a random union of gametes such as a broadcast spawner (Pudovkin *et al.*, 1996). Furthermore, although the method performed well in simulations, the evaluation of Pudovkin *et al.* (1996) did not include sampling error, and most studies to date have used far fewer loci than the 30 loci used in these simulations. Lastly, Pudovkin *et al.* (1996) found large 95% confidence intervals, limiting its usefulness.

### Two samples: the temporal method

The last  $N_e$  estimator we discuss, the temporal method, is based on changes in allele frequencies that occur between two samples taken from a small population. Allele frequencies at neutral loci change more rapidly in small populations than large ones as a result of genetic drift. This inverse relationship between  $N_e$  and the temporal change in allele frequencies ( $F_k$ ) can be characterized by the following equation:

$$\hat{N}_e = \frac{t}{2 \left[ F_k - \frac{1}{2S_0} - \frac{1}{2S_t} \right]} \quad (4)$$

where  $t$  is time in generations between samples,  $S_0$  and  $S_t$  are the number of individuals sampled in generations 0 and  $t$ . The variance in allele frequencies over time,  $F_k$ , is estimated by:

$$\hat{F}_k = \frac{1}{A-1} \sum \frac{(x_i - y_i)^2}{(x_i + y_i)/2} \quad (5)$$

where  $A$  is the number of segregating alleles at the locus and  $x$  and  $y$  are the allele frequencies of the  $i$ th allele in 0 and  $t$  generations, respectively (Nei & Tajima, 1981; Pollack 1983; Waples, 1989). Consequently, one can monitor  $N_e$  by monitoring changes in allele frequencies at multiple loci over two or more generations (see Jorde & Ryman, 1995; Fig. 1).

The temporal method requires samples of approximately 30 individuals in successive generations and >5 polymorphic microsatellite loci to achieve precise estimates of  $N_e$  when  $N_e$  is very small (i.e. four). The method requires 45 individuals and 10 loci when  $N_e$  is 10–20 (Luikart & Cornuet, 1998). However, to best enhance

the ability to detect bottlenecks using the temporal method it is most efficient to increase the number of generations between samples. This allows more drift to occur (Waples, 1989). Therefore, this method may be most useful for long-term monitoring, where several generations have passed between population samples.

To date, the temporal allele method is the most tested of the genetic  $N_e$  estimators and has been used to determine the  $N_e$  of plant, salmonid, shellfish and amphibian populations (Waples, 1989; Hedgecock, Chow & Waples, 1992; Husband & Barret, 1992; Jorde & Ryman 1995; Scribner *et al.*, 1997). Its strengths lie in its accuracy and precision for estimating bottleneck sizes (Luikart & Cornuet, 1998). The temporal method produces biased estimates of  $N_e$  when  $N_e$  is small (Richards & Leberg, 1996), however this bias is negligible unless  $N_e$  is less than approximately 6–10 or  $t$  is large (>5; Luikart & Cornuet, 1998).

The primary weakness of the temporal allele method is that two or more samples are necessary (Table 1). This can be expensive and, by nature, time-consuming. For secretive or elusive species, it may be necessary to pool multiple years of data to obtain a large enough sample size for a first population sample. However, in cases where many museum specimens from a population are available, only one current population sample would be necessary to estimate  $N_e$  with the temporal allele method.

Overall, this is the most promising method for estimating  $N_e$  because it provides increasingly precise estimates of  $N_e$  as the number of alleles per locus increases. However, it is difficult to compare the relative usefulness of the three methods because their performance has never been compared under the same conditions. Computer simulation studies are badly needed to determine (1) the consequences of violating the assumptions of the methods, and (2) the number of loci and individuals that must be sampled to achieve precise estimates of  $N_e$ .

## ESTIMATIONS OF $N_c$ FROM GENETIC DATA

The advent of hypervariable DNA markers allows researchers to use genetic samples from individual animals in a manner analogous to conventional marks such as ear tags. Specifically, each DNA sample can be screened at several loci to identify the multi-locus genotype of the animal (see Fig. 1). Each unique genotype is considered equivalent to 'marking' an animal. The number of loci needed to correctly identify unique individuals is primarily a function of levels of heterozygosity and allelic diversity (Paetkau & Strobeck, 1994). Below we outline how these unique genotypes can be used to estimate  $N_c$ .

### One sample

When resources or feasibility limit sampling efforts to a single population sample, non-invasive genetic samples may provide important rudimentary demographic information. A single intensive sampling effort can provide

information about the minimum number of individuals (i.e. unique genotypes) in a population or social group. For some taxa, sex-specific markers have been developed to determine the minimum number of females and males in a population. For example, Taberlet, Camarra *et al.* (1997) used microsatellites and a sex-specific marker to determine the minimum number of male and female Pyrenean brown bears. Reed *et al.* (1997) used five microsatellite markers and one sex-specific marker on fecal samples to identify the species, number of individuals per species and sex ratios at a mixed species seal 'haul-out'. Genotypes recovered from feces matched blood samples used as controls. Furthermore, sex ratios estimated from the seal fecal sample genotypes agreed with prior observational data. These recent studies suggest samples of hair and feces can be analyzed with PCR-markers to provide demographic information.

### Two or more samples: the most recent application

Estimating population size from multiple samples taken from a population in a single generation is the most undeveloped of the  $N_e$  or  $N_c$  estimation techniques. The general idea is that each unique genotype in a first population sample can be considered a capture and each genotype that re-occurs in subsequent population samples can be considered a recapture. These data are then functionally equivalent to those obtained *via* more conventional trapping/sighting field methods and can be analyzed with standard capture–mark–recapture methods reviewed extensively elsewhere (Seber, 1982; Nichols, 1991; Pollock *et al.*, 1990). The only published use of genetic capture–mark–recapture that we are aware of is an estimation of humpback whale population size in the North Atlantic Ocean (Palsbøll *et al.*, 1997). Palsbøll *et al.* (1997) used six microsatellite markers and sex-specific markers on genetic samples collected in 1992 and 1993 to estimate that there were 4894 males and 2804 females in this population. This estimate was not significantly different from estimates determined from previously used photographic techniques.

It is important to recognize trade-offs and pitfalls, as well as the benefits, that may accompany the use of genetic samples to obtain demographic information (see Table 1). An important trade-off of these methods is that while they may be non-invasive, they will not provide potentially important demographic information, such as age structure, individual health and sexual condition, that can be obtained *via* conventional survey methods. In addition, no one has thoroughly investigated violations of capture–mark–recapture model assumptions and potential problems that may be caused by using genetic data to estimate population size. For example, demographic events that decrease the number of genotypes in a population far below  $N_c$ , such as a bottleneck, or frequent local extinction and recolonization (Hedrick, 1996), may reduce the power to identify individuals and bias demographic estimators. Thus, the probability of finding two identical genotypes for any set of loci and allele frequencies should be calculated as described by

Paetkau & Strobeck (1994). If this probability is high, emphasis should be placed on the development of additional markers to increase the power to detect unique individuals (Blouin *et al.*, 1996; Taberlet, Camarra *et al.*, 1997).

Additional problems with genetic capture-mark-recapture methods is that there may be biases introduced into population estimation using multi-locus genotype data (L. S. Mills, pers. comm.). Furthermore, how samples are collected (e.g. the spatial configuration of hair trap stations) needs to be modeled, to determine how best to modify conventional capture-mark-recapture and line transect methods for use with genotype data (Buckland *et al.*, 1992; Thompson, 1992). With computer simulations and statistical modification, repeated genetic population samples will provide information about population growth rate and survivorship estimates, as do traditional open capture-mark-recapture models (Pollock *et al.*, 1990; Lebreton *et al.*, 1992).

On the positive side, the use of genetic data may be the only way to obtain sufficient demographic information for some species and may be the least risky way to acquire these data (i.e. preventing the need for capturing or immobilizing large animals). Additionally, mark loss and estimator bias associated with behavioral responses to 'trapping' may be minimized with genetic data.

## MONITORING PROGRAMS

The most practical application of genetic estimates of  $N_e$  and  $N_c$  is to enhance the power of monitoring programs to detect changes in population size of threatened or endangered species. Implementing the methods we reviewed above in monitoring programs will be challenging. While there is great promise, there also remain practical and theoretical limitations of each method that must be addressed. Overriding practical constraints to the use of DNA markers for monitoring programs are primarily the costs involved. In addition, both the  $N_e$  and  $N_c$  estimators generally require large sample sizes to obtain reasonable precision, and have not been extensively tested to date.

Consideration of the goals of a monitoring program must dictate how limited resources will be allotted to collect data. In some cases, the sampling protocols required to derive  $N_c$  and  $N_e$  estimates may make it necessary to choose between the two. For example, repeated sampling must occur within a single generation to use capture-mark-recapture methods for estimation of demographic parameters (or genetic 'marks' are lost, and the assumptions of a closed model are violated), whereas the temporal allele method of  $N_e$  estimation will be most useful if sample collections are separated by several generations. Therefore, a limited budget will constrain the type of information and the accuracy of the information that can be obtained. An ideal monitoring program would consist of multiple sampling periods within and across generations to estimate both  $N_c$  and  $N_e$  accurately.

## CONCLUSIONS

We emphasize that although the molecular genetics technology and the  $N_e$  and  $N_c$  estimator models have been developed, the combined application of these to population monitoring remains in its infancy. Research is particularly needed on both the feasibility of different sampling designs and on the statistical power of most methods, especially gametic disequilibrium and heterozygote excess, to detect population changes under the constraints imposed by 'real world' settings. In contrast, the number of loci and individuals required to estimate  $N_e$ , using the temporal allele method, and  $N_c$  are fairly well understood. It is important to use tools such as simulation modeling to explore the quality and quantity of data that can be expected from the use of both non-invasive sampling and these analytical tools before a monitoring plan is implemented. Of course, careful consideration of the natural history of each species must also dictate which analytical method is most appropriate for monitoring efforts and conservation goals.

Recent advances in molecular and theoretical population biology provide an opportunity to more effectively design and conduct biological monitoring efforts. Further efforts are needed to evaluate the usefulness of promising molecular tools for obtaining important demographic and genetic data. While technological advances will never be a panacea for saving biological diversity, we hope to spur efforts to use available tools that have minimum impacts on the species of concern, yet hold great promise for their conservation in the wild.

## Acknowledgements

We thank Kevin Lair, John Citta, John MacKay, Fred W. Allendorf, L. Scott Mills and two anonymous reviewers for reviewing early drafts of the manuscript. This paper benefited from work and conversations with P. Taberlet, F. W. Allendorf and L. S. Mills. While writing this manuscript we were supported by NSF Graduate Research Training Grant (NSF DGE 955 3611) (D.A.T.), the Graduate School at the University of Montana and an NSF-NATO Fellowship (G.L.), McIntire-Stennis Grant and Les Pengelly scholarship (M.K.S.).

## REFERENCES

- Allendorf, F. W., Bayles, D., Bottom, D. I., Currens, K. P., Frissell, C. A., Hankin, D., Lichatowich, J. A., Nehlsen, W., Trotter, P. C. & Williams, T. H. (1997). Prioritizing Pacific salmon stocks for conservation. *Conserv. Biol.* **11**: 140-152.
- Bartley, D., Bagley, M., Gall, G. & Bentley, B. (1992). Use of linkage disequilibrium data to estimate effective size of hatchery and natural fish populations. *Conserv. Biol.* **6**: 365-375.
- Blouin, M. S., Parsons, M., Lacaille, V. & Lotz, S. (1996). Use of microsatellite loci to classify individuals by relatedness. *Mol. Ecol.* **5**: 393-401.
- Bruford, M. W. & Wayne, R. K. (1993). Microsatellites and their application to population genetic studies. *Curr. Opin. Genet. Dev.* **3**: 939-943.
- Buckland, S. T., Anderson, D. R., Burnham, K. P. & Laake, J. L. (1992). *Distance sampling: estimating abundance of biological populations*. London: Chapman and Hall.

- Caughley, G. & Gunn, A. (1996). *Conservation biology in theory and practice*. Oxford: Blackwell Scientific.
- Crow, J. F. & Kimura, M. (1970). *An introduction to population genetics theory*. Minneapolis: Burgess.
- Falconer, D. S. & Mackay, T. F. C. (1996). *Introduction to quantitative genetics* (4th edn). Harlow, Essex: Longman.
- Forbes, S. H. & Allendorf, F. W. (1991). Associations between mitochondrial and nuclear genotypes in cutthroat trout hybrid swarms. *Evolution* **45**: 1332–1349.
- Frankham, R. (1995a). Effective population size/adult population size ratios in wildlife: a review. *Genet. Res. Camb.* **66**: 95–106.
- Frankham, R. (1995b). Conservation genetics. *Annu. Rev. Genet.* **29**: 305–327.
- Gagneux, P., Boesch, C. & Woodruff, D. S. (1997). Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Mol. Ecol.* **6**: 861–868.
- Gerloff, U., Schlötterer, C., Rassmann, K., Rambold, I., Hohmann, G., Fruth, B. & Tautz, D. (1995). Amplification of hypervariable simple sequence repeats (microsatellites) from excremental DNA of wild living bonobos (*Pan paniscus*). *Mol. Ecol.* **4**: 515–518.
- Harris, R. B. & Allendorf, F. W. (1989). Genetically effective population size of large mammals: an assessment of estimators. *Conserv. Biol.* **3**: 181–191.
- Hedgecock, D., Chow, V. & Waples, R. S. (1992). Effective population numbers of shellfish brood stocks estimated from temporal variance in allelic frequencies. *Aquaculture* **108**: 215–232.
- Hedrick, P. W. (1987). Gametic disequilibrium measures: proceed with caution. *Genetics* **117**: 331–341.
- Hedrick, P. W. (1996). Bottleneck(s) or metapopulation in cheetahs. *Conserv. Biol.* **10**: 897–899.
- Hill, W. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genet. Res. Camb.* **38**: 209–216.
- Husband, B. C. & Barrett, S. C. H. (1992). Effective population size and genetic drift in tristylous *Eichlornia paniculata* (Pontederiaceae). *Evolution* **46**: 1875–1890.
- Jorde, P. E. & Ryman, N. (1995). Temporal allele frequency change and estimation of effective population size in populations with overlapping generations. *Genetics* **139**: 1077–1090.
- Lebreton, J.-D., Burnham, K. P., Clobert, J. & Anderson, D. R. (1992). Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecol. Monog.* **62**: 67–118.
- Luikart, G. & Cornuet, J. M. (1998). Temporal change in allele frequencies provide estimates of bottleneck size. *Conserv. Biol.* in press.
- Mace, G. H. & Lande, R. (1991). Assessing extinction threats: towards a re-evaluation of IUCN threatened species categories. *Conserv. Biol.* **5**: 148–157.
- Morin, P. A. & Woodruff, D. S. (1996). Noninvasive genotyping for vertebrate conservation. In *Molecular genetic approaches in conservation*. Smith, T. B. & Wayne, R. K. (Eds). Oxford: Oxford University Press.
- Mundy, N. I., Winchell, C. S., Burr, T. & Woodruff, D. S. (1997). Microsatellite variation and microevolution in the critically endangered San Clemente Island loggerhead shrike (*Lanius ludovicianus mearnsi*). *Proc. R. Soc. Lond. Ser. B* **264**: 869–875.
- Nei, M. (1987). *Molecular evolutionary genetics*. New York: Columbia University Press.
- Nei, M. & Tajima, F. (1981). Genetic drift and estimation of effective population size. *Genetics* **98**: 625–640.
- Newman, D. & Pilson, D. (1997). Increased probability of extinction due to decreased genetic effective population size: experimental population of *Clarkia pulchella*. *Evolution* **51**: 354–362.
- Nichols, J. D. (1991). Capture–recapture models. *Bioscience* **42**: 94–102.
- Nunney, L. & Elam, D. R. (1994). Estimating the effective population size of conserved populations. *Conserv. Biol.* **8**: 175–184.
- Paetkau, D. & Strobeck, C. (1994). Microsatellite analysis of genetic variation in black bear populations. *Mol. Ecol.* **3**: 489–495.
- Palsbøll, P., Allen, J., Bérubé, M., Clapham, P. J., Feddersen, T. P., Hammond, P. S., Hudson, R. R., Jørgensen, H., Katona, S., Larsen, A. H., Larsen, F., Lien, J., Mattila, D. K., Sigurjansson, J., Sears, R., Smith, T., Sponer, R., Stevick, P. & Ølen, N. (1997). Genetic tagging of humpback whales. *Nature, Lond.* **388**: 767–769.
- Pemberton, J. M., Slate, J., Bancroft, D. R. & Barrett, J. A. (1995). Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. *Mol. Ecol.* **4**: 249–252.
- Pollack, E. (1983). A new method for estimating the effective population size from allele frequency changes. *Genetics* **104**: 531–548.
- Pollock, K. H., Nichols, J. D., Brownie, C. & Hines, J. E. (1990). Statistical inference for capture–recapture experiments. *Wildl. Monog.* **197**: 1–97.
- Pudovkin, A. I., Zaykin, D. V. & Hedgecock, D. (1996). On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. *Genetics* **144**: 383–387.
- Reed, J. Z., Tollit, D. J., Thompson, P. M. & Amos, W. (1997). Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal faeces. *Mol. Ecol.* **6**: 225–234.
- Richards, C. & Leberg, P. L. (1996). Temporal changes in allele frequencies and a population's history of severe bottlenecks. *Conserv. Biol.* **10**: 832–839.
- Scribner, K., Arntzen, J. W. & Burke, T. (1997). Effective number of breeding adults in *Bufo bufo* estimated from age-specific variation at minisatellite loci. *Mol. Ecol.* **6**: 701–712.
- Seber, G. A. F. (1982). *The estimation of animal abundance and related parameters*. London: Charles Griffin.
- Taberlet, P., Camarra, J.-J., Griffin, S., Uhres, E., Hanotte, O., Waits, L. P., Dubois-Paganon, C., Burke, T. & Bouvet, J. (1997). Noninvasive genetic tracking of the endangered Pyrenean brown bear population. *Mol. Ecol.* **6**: 869–876.
- Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L. P. & Bouvet, J. (1996). Reliable genotyping of samples with very low DNA quantities using PCR. *Nucl. Acids Res.* **24**: 3189–3194.
- Thompson, S. K. (1992). *Sampling*. New York: John Wiley and Sons, Inc.
- Vrijenhoek, R. C., Pfeiler, E. & Wetherington, J. D. (1992). Balancing selection in a desert stream-dwelling fish, *Poeciliopsis monacha*. *Evolution* **46**: 1642–1657.
- Waples, R. S. (1989). A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* **121**: 379–391.
- Waples, R. S. (1991). Genetic methods for estimating the effective size of cetacean populations. *Rep. Int. Whaling Commn Spec. Issue* **13**: 279–300.
- Wright, S. (1969). *Evolution and the genetics of populations*. Vol. 2: *The theory of gene frequencies*. Chicago: University of Chicago Press.