

TECHNICAL ADVANCES

# When are genetic methods useful for estimating contemporary abundance and detecting population trends?

DAVID A. TALLMON,\* DAVE GREGOVICH,+ ROBIN S. WAPLES,‡ C. SCOTT BAKER,§ JENNIFER JACKSON,§ BARBARA L. TAYLOR,+ ERIC ARCHER,+ KAREN K. MARTIEN,+ FRED W. ALLENDORF¶ and MICHAEL K. SCHWARTZ\*\*

\*Biology and Marine Biology Program, University of Alaska Southeast, 11120 Glacier Highway, Juneau, AK 99801, USA, +Southwest Fisheries Science Center, 8604 La Jolla Shores Drive, La Jolla, CA 92037-1508, USA, ‡Northwest Fisheries Science Center, 2725 Montlake Blvd. East, Seattle, WA 98112, USA, §Marine Mammal Institute and Department of Fisheries and Wildlife, Oregon State University, 2030 SE Marine Science Dr., Newport, OR 97365, USA, ¶Division of Biological Sciences, University of Montana, Missoula, MT 59801, USA, \*\*USDA Forest Service, Rocky Mountain Research Station, 800 E. Beckwith Ave., Missoula, MT 59801, USA

## Abstract

The utility of microsatellite markers for inferring population size and trend has not been rigorously examined, even though these markers are commonly used to monitor the demography of natural populations. We assessed the ability of a linkage disequilibrium estimator of effective population size ( $N_e$ ) and a simple capture-recapture estimator of abundance ( $N$ ) to quantify the size and trend of stable or declining populations (true  $N = 100$ – $10,000$ ), using simulated Wright–Fisher populations. Neither method accurately or precisely estimated abundance at sample sizes of  $S = 30$  individuals, regardless of true  $N$ . However, if larger samples of  $S = 60$  or  $120$  individuals were collected, these methods provided useful insights into abundance and trends for populations of  $N = 100$ – $500$ . At small population sizes ( $N = 100$  or  $250$ ), precision of the  $N_e$  estimates was improved slightly more by a doubling of loci sampled than by a doubling of individuals sampled. In general, monitoring  $N_e$  proved a more robust means of identifying stable and declining populations than monitoring  $N$  over most of the parameter space we explored, and performance of the  $N_e$  estimator is further enhanced if the  $N_e/N$  ratio is low. However, at the largest population size ( $N = 10,000$ ),  $N$  estimation outperformed  $N_e$ . Both methods generally required  $\geq 5$  generations to pass between sampling events to correctly identify population trend.

**Keywords:** abundance, effective population size, genetic monitoring, population size, population trend

Received 20 September 2009; revision received 14 December 2009; accepted 23 December 2009

## Introduction

Genetic markers have become increasingly popular as a means to gain insights into the demography of wild populations. For over a decade, it has been generally acknowledged that genetic markers can provide insights into recent demographic and genetic changes from

tissues obtained invasively or non-invasively from species that are common, rare or cryptic (Schwartz *et al.* 1998, 2007; England & Luikart 1999; Lukacs & Burnham 2005b). Because genetic markers can provide adequate data to model population abundance in situations where conventional capture-recapture or other techniques do not, there has been great enthusiasm surrounding their use for assessing and monitoring abundance. Given the popularity of genetic markers to obtain demographic insights from a wide variety of different animal species and populations, it is somewhat surprising that there has

Correspondence: David A. Tallmon, Fax: 1(907)796 6447; E-mail: david.tallmon@uas.alaska.edu

been no systematic examination of the performance of genetic marker-based estimators of effective population size ( $N_e$ ) and abundance ( $N$ ). As a consequence, it is unclear under what conditions managers or scientists might be able to detect demographic trends or what they might do to increase their ability to obtain useful insights into demography.

Genetic markers have been used successfully to obtain insights into contemporary demography of wild animal populations in a variety of ways (Bellemain *et al.* 2005; Aspi *et al.* 2006; Goossens *et al.* 2006; Kendall *et al.* 2008; Robinson *et al.* 2009), but nearly all methods require a set of multilocus genotypes collected from a randomly sampled set of individuals (Mills *et al.* 2000). However, the differences in precision and power of genotypes for estimating both  $N_e$  and  $N$  under a common sampling design have not previously been evaluated using simulations with known  $N_e$  and  $N$ . In general, researchers tend to collect samples of genotypes to estimate either  $N_e$  or  $N$ , depending upon their expertise or the basic biology of their focal species.

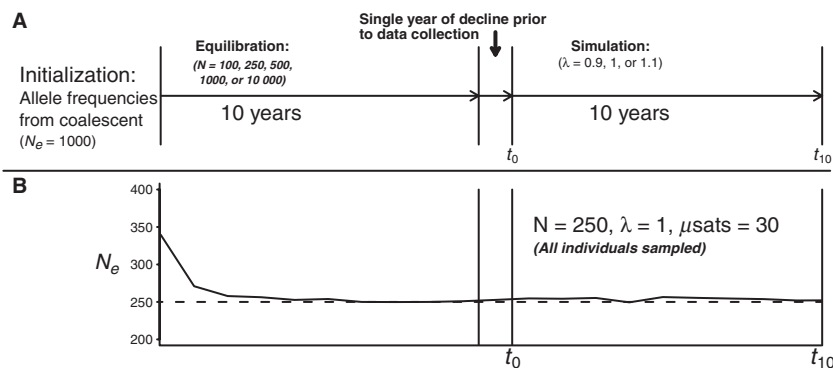
Here, we focus on use of a sample of multilocus microsatellite genotypes to estimate  $N_e$  and  $N$  and gain insights into contemporary population demography. Related coalescent methods that can be used to provide insights into long-term or historical evolutionary processes are reviewed elsewhere (Kuhner 2008). Our focus is on the relative performance of  $\hat{N}_e$  and  $\hat{N}$  estimators under a consistent set of simulated biological and sampling conditions used in typical population genetics studies.  $\hat{N}_e$  and  $\hat{N}$  are often used in a conservation context to assess population status or extinction risk. If samples are obtained from different cohorts or generations of a population, these samples can be used in a variety of ways to infer trends in  $N_e$  or  $N$ . Although one might wish to obtain genetic data for reasons other than, or in addition

to, insights into abundance and trends in abundance (Lukacs & Burnham 2005b), here we focus on how limited genetic data and  $\hat{N}_e$  or  $\hat{N}$  might be used to infer current  $N$  and population growth rate ( $\lambda$ ).

We examine the performance of two genetic methods for detecting trends in population abundance under an array of sampling conditions and simple population dynamics. The first method uses genetic markers to identify (mark) individuals for a traditional Lincoln-Petersen capture-recapture estimate of changes in  $N$  over time. The second method uses genetic markers to detect change in  $N_e$  based on the magnitude of gametic (linkage) disequilibrium in the samples. We examine the ability of these methods to provide useful information about  $N$  and  $\lambda$  in a population of known, simulated demography with Wright-Fisher (W-F) mating and discrete generations. In a stable population with W-F mating,  $N_e = N$ , so our simulations provide useful conditions under which to directly compare  $N_e$  and  $N$  estimators without the confounding demographic complexity found in more complicated mating schemes. Our examination includes a variety of  $N$  and  $\lambda$  values, as well as sampling efforts of individuals ( $S$ ) and loci ( $L$ ) typical for population genetics studies. From these varied demographic and sampling conditions, we evaluate and discuss the performance of the two methods to provide useful and reliable insights into population abundance and trends in a W-F population and identify conditions where one method is clearly better than the other.

## Materials and methods

We simulated the evolution of populations of known abundance ( $N$ ) and growth rate ( $\lambda$ ) using simulation methods developed by Martien *et al.* (2009). A standard initialization and equilibration phase was used for every



**Fig. 1** The component phases (initialization, equilibration and simulation) of each of the replicate simulations used in this study (A), along with an example of the corresponding behaviour of median  $\hat{N}_e$  for a stable population (B).

simulation (Fig. 1A). First, SimCoal (Laval & Excoffier 2004) was used to create multilocus allele frequencies for 100 replicate populations with historic  $N_e = 1000$ . For each replicate, a close approximation of a W-F population of size  $N$  was created with the R package Rmetasim (Strand 2002). Initial genotypes for each individual in the population were drawn from the multilocus allele frequency distribution generated by SimCoal. Each population then went through a 10-generation equilibration phase of W-F mating at size  $N = 100, 250, 500, 1000$  or  $10000$ , which allowed each population to come into Hardy–Weinberg proportions and stable levels of gametic disequilibrium (Fig. 1B), while retaining plausible levels of genetic variation. Three to four generations are generally sufficient to reach asymptotic levels of gametic disequilibrium after initialization (Waples 2006).

After 10 generations of equilibration, we simulated different population dynamics and sampling conditions. The simulated populations ( $N = 100, 250, 500, 1000, 10000$ ) followed a deterministic growth rate ( $\lambda = 0.9$  or  $1.0$ ) for one generation, starting at generation  $t_{-1}$ , and then data collection began at generation  $t_0$  as the populations followed the same  $\lambda$  for 10 generations of the simulation phase (Fig. 1A). This initiation of population growth at generation  $t_{-1}$  allowed us to more directly compare the  $N_e$  and  $N$  estimators, because single sample linkage disequilibrium  $N_e$  estimates reflect the number of parents in the generation preceding a sample (Waples 2005). In declining populations ( $\lambda = 0.9$ ),  $N_e$  values will be biased upwards by the larger  $N_e$  in previous generations, but the impact should not be large relative to the effects of  $N_e$  in the immediately preceding generation (Waples 2005). For the stable populations ( $\lambda = 1.0$ ) we simulated,  $N_e \cong N$  each generation apart from demographic fluctuations (Waples & Faulkner 2009). Mutation was included throughout the Rmetasim simulations of the equilibration and simulation phases at a rate of  $2 \times 10^{-3}$ /locus per generation, based upon a survey of published mutation rates.

During the simulation phase, samples of loci ( $L = 15$  or  $30$ ) and individuals ( $S = 30, 60$  or  $120$ ) were collected at specified times ( $t = 0, 1, 5$  or  $10$ ) and used to estimate  $\hat{N}_e$ ,  $\hat{N}$  and  $\hat{\lambda}$ . For  $N = 10,000$ , we used larger samples ( $S = 240$  or  $480$ ). From each sample,  $\hat{N}_e$  and  $\hat{N}$  were obtained from individual genotypes using gametic disequilibrium and Chapman-corrected LP estimators, respectively. Estimates of  $\hat{N}_e$  depend on the theoretical relationship between  $r^2$  (a measure of gametic disequilibrium) and  $N_e$  (Hill 1981). We used the program LDNe (Waples & Do 2008), which implements the bias-correction method developed by (Waples 2006), to obtain  $\hat{N}_e$  from each sample of  $S$  individuals. For LDNe, we used the criterion  $P_{\text{crit}} = 0.02$  (alleles with frequency  $< 0.02$  are excluded), which generally provides a good balance between precision and bias (Waples & Do 2009). Confi-

dence intervals (CIs) for  $\hat{N}_e$  are based upon the chi-square approximation implemented by LDNe (Equation 12 in (Waples 2006)). To obtain  $\hat{N}$  and keep sampling effort equal for the two methods, each sample was split evenly between a set of initially captured individuals ( $S/2$ ), which were genotyped and returned to the population, followed immediately by a random sampling of a second set of  $S/2$  individuals. This sampling approach mimics that most likely to be used in a population genetics study. The  $\hat{N}$  values and lognormal confidence CIs for the LP estimator were calculated following Seber (1982), and assume each individual is uniquely and accurately identified from its genotype.

To first assess the performance of these estimators in stable populations, we compared the bias and precision of these two estimators at time  $t_0$  for each  $N$ . Next, we examined how well  $\hat{N}_e$  or  $\hat{N}$  obtained at regular time intervals can be used to detect change in a population that is either declining or stable over a 1–10 generation period. For each method,  $\lambda$  was estimated as the slope of a linear regression on the log transforms of the point estimates of abundance at  $t_0$  and  $t_i$ . We recorded the proportion of times  $\hat{N}_e$  or  $\hat{N}$  estimates taken from sequentially collected samples correctly identified  $\hat{\lambda} < 0.95$  when true  $\lambda = 0.9$ , and  $\hat{\lambda} > 0.95$  when true  $\lambda = 1.0$ . In other words, we posed the simple decision rule where a manager might take action if a population was thought to be declining at least 5% per generation ( $\lambda = 0.95$ ). For our two growth scenarios, correct decisions would be: when  $\lambda = 0.9$  the proportion identified as  $\lambda < 0.95$  and when  $\lambda = 1.0$  the proportion identified as  $> 0.95$ . For simplicity, we discuss only the case using 30 loci for the population trend analysis.

In a population that conforms closely to a W-F population,  $N_e \cong N$ . However, this is rarely the case in real populations. Recent surveys of natural populations have reported median  $N_e/N$  ratios of 0.14 (Palstra & Ruzzante 2008) and 0.11 (Frankham 1995). Arguably, then, it is more realistic to compare  $\hat{N}_e$  performance at a given  $N$  to  $\hat{N}$  performance at a much larger  $N$ , because  $N_e$  is usually much less than  $N$  in real populations. To achieve this comparison, we compared the accuracy and precision of  $\hat{N}$  when true  $N = 1000$  with that of the  $N_e$  estimator when true  $N$  (and  $N_e$ ) was 100, 250 or 500. By using these values for the  $N_e$  estimator, we bracketed the median values of  $N_e/N$  found in the literature for natural populations and provide insight into the most effective way to monitor a population that has a  $N_e$  that is 10–50% of  $N$ .

## Results

Population size during the equilibration phase affected the initial level of genetic diversity to begin each simulation phase. Following theoretical predictions, the primary

**Table 1** Mean expected heterozygosity ( $H_e$ ) and number of alleles per locus (# Alleles) present in the initial ( $t_0$ ) and final ( $t_{10}$ ) generation of 100 replicate simulations of different population sizes ( $N$ ) and growth rates ( $\lambda$ ). In all simulations, samples of  $S = 30, 60$  or  $120$  individuals were collected at  $t = 0, 1, 5$  and  $10$  generations and genotyped at  $L = 15$  or  $30$  loci

N	Initial variation ( $t_0$ )		Final variation ( $t_{10}$ )			
	$H_e$	# Alleles	$\lambda = 1.0$		$\lambda = 0.9$	
			$H_e$	# Alleles	$H_e$	# Alleles
100	0.72	6.3	0.69	5.7	0.66	4.8
250	0.74	7.2	0.73	7.0	0.72	6.3
500	0.75	7.7	0.74	7.6	0.74	7.1
1000	0.75	8.1	0.75	8.2	0.75	7.7
10000	0.76	9.3	0.76	9.5	0.76	9.0

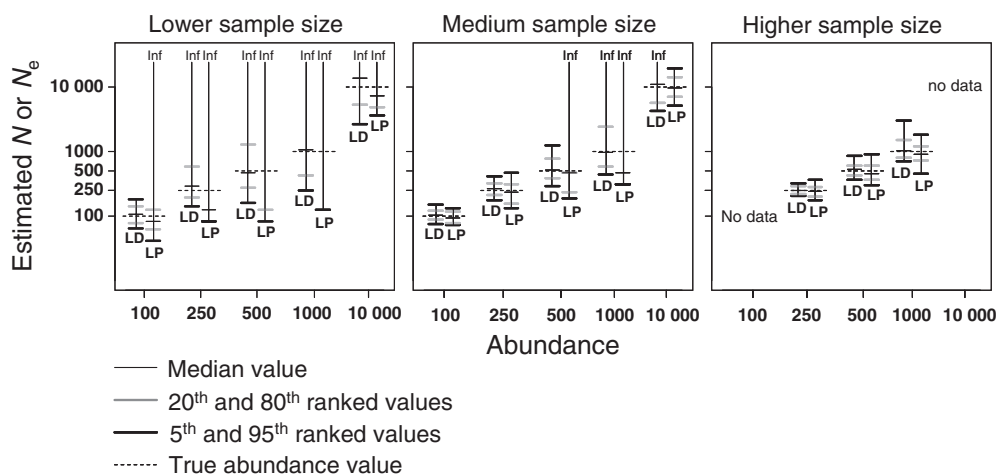
impact of population size is on the number of alleles per locus, with less impact on initial heterozygosity (Table 1). Mean number of alleles per locus at generation  $t_0$  varied from 6.3 to 9.3, whereas heterozygosity varied only from 0.72 to 0.76, for  $N = 100$  and  $N = 10000$ , respectively. Genetic variation was lost in declining populations following theoretical expectations, with smaller populations losing genetic variation more rapidly than larger ones. However, in some simulations of large stable populations, genetic variation actually increased over time from generation  $t_0$ – $t_{10}$  because mutation created new alleles that were not lost via genetic drift.

*Abundance*

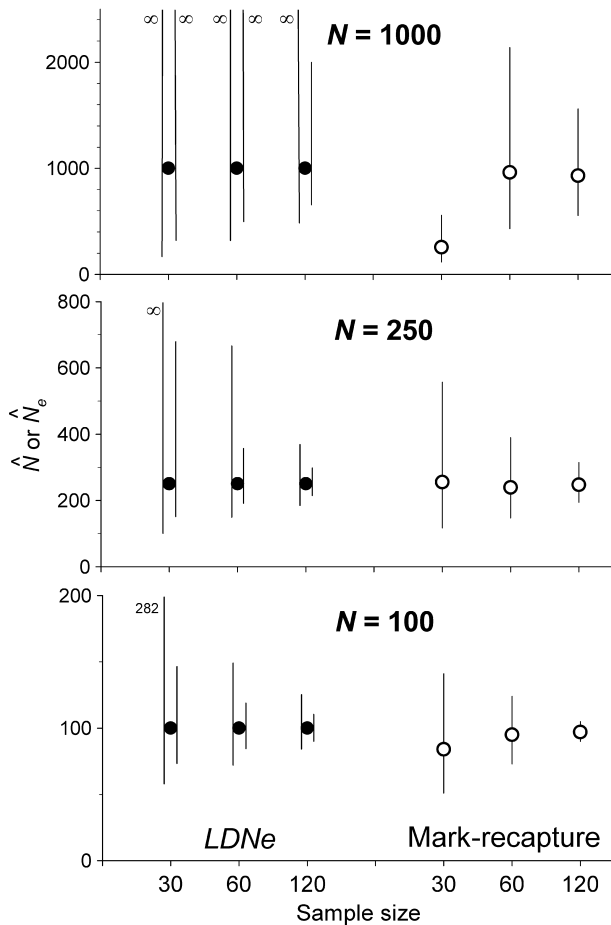
The ability of  $\hat{N}_e$  and  $\hat{N}$  to provide insight into abundance varies considerably with the number of individuals

sampled (Fig. 2). Although the bias of both estimators decreases with increasing numbers of individuals sampled,  $\hat{N}$  tends to be consistently negatively biased, whereas  $\hat{N}_e$  is consistently slightly upwardly biased. In general,  $\hat{N}$  shows a larger absolute bias than  $\hat{N}_e$  under most conditions. However, with moderate ( $S = 60$ ) to large ( $S = 120$ ) sample sizes for a population genetics study, both  $\hat{N}_e$  and  $\hat{N}$  are reasonably unbiased.

The precision of  $\hat{N}_e$  and  $\hat{N}$  in stable populations is also informative and is a function of the true abundance and number of individuals and loci sampled (Fig. 3). With only 30 individuals sampled and genotyped at 15 loci, both  $\hat{N}_e$  and  $\hat{N}$  have very large CIs, whether  $N = 100$  or 1000. In several cases of small  $S$  and large  $N$ , the  $\hat{N}_e$  upper CIs are indistinguishable from infinity. However, the CIs decrease very rapidly with increased  $S$  for both estimators across the range of abundances investigated. These results are consistent with median point estimates shown in Fig. 2. Interestingly, the relative precision of the two estimators changes with the number of loci genotyped. That is, at  $N = 100$  or  $250$ , the  $\hat{N}$  CIs are usually tighter than the  $\hat{N}_e$  CIs at  $L = 15$  for a given number of individuals, but this relationship is reversed if  $L = 30$ . This is a consequence of our assumption that 15 loci are sufficient for perfect identification of individuals in capture-recapture. Therefore, increasing  $L$  does not affect precision of  $\hat{N}$ , whereas the width of  $\hat{N}_e$  CIs decreases rapidly as the number of loci, and hence pairs of alleles used to estimate linkage disequilibrium, increases. At  $N = 100$  or  $250$ , the precision of  $\hat{N}_e$  improves slightly more with a doubling of  $L$  than with a doubling of  $S$ . At  $N \geq 500$ , it becomes difficult to obtain finite CIs for  $\hat{N}_e$  unless sampling effort is relatively large ( $S = 120$  and  $L = 30$ ), but the  $\hat{N}$  CIs are more reasonable. At such large  $N$ , the signal from genetic



**Fig. 2** Accuracy of  $\hat{N}_e$  and  $\hat{N}$  for a range of true abundances ( $N$ ) and sampling efforts. Shown are the 5th, 20th, 50th, 80th and 95th ranked values from 100 replicate simulations. Lower, medium and higher sample sizes were  $S = 30, 60$  and  $120$ , respectively, except for  $N = 10000$ , for which  $S = 240$  and  $480$  were lower and medium values, respectively. For  $N = 100$ , population size was insufficient to samples of  $S = 120$ .



**Fig. 3** Confidence intervals for the  $\hat{N}_e$  and  $\hat{N}$  estimators for 3 different values of  $N$ . For each  $\hat{N}_e$  point estimate, the vertical line to the left assumes  $L = 15$  loci and the line to the right assumes  $L = 30$  loci, both using the mean number of alleles observed in the simulated data for that value of  $N$  (see Table 1). For  $\hat{N}$ , the sample size is the sum of the numbers of individuals collected in the two time intervals, which are assumed to be equal. This illustration assumes that the point estimate is equal to  $N$  ( $\hat{N}_e$ ; filled circles) or is as close to  $N$  as possible given the sample size ( $\hat{N}$ ; open circles). Small numbers or symbols show exact values for upper or lower bounds of confidence intervals that are beyond the scale shown.

drift is weak, so unless sampling effort is considerable (or the ratio  $N_e/N$  is low; see below), the genetic estimates are not very useful. Overall, the results suggest a sample of  $S = 60$  individuals genotyped at  $L = 15$  loci would be a useful sampling target to have reasonable expectations of finite CIs for  $N = 100$ – $250$ . With a sample of  $S = 120$ , reasonably precise  $\hat{N}$  and  $\hat{N}_e$  can be obtained for abundances up to  $N = 500$ . Interestingly, at the largest abundance  $N = 10000$ , a sample of  $S = 480$  individuals was sufficient to obtain reasonably accurate and precise  $\hat{N}$ , but not  $\hat{N}_e$ . Again, at such a large population size, there is little

genetic drift and so genetic methods are unlikely to work well unless  $N_e \ll N$ .

#### Population trends

In the vast majority of cases,  $\hat{N}_e$  correctly identifies both declining and stable populations more frequently than does  $\hat{N}$  (Table 2). However, the performance of both methods is greatly affected by population abundance and sampling effort, as described earlier for abundance inferences. Accurate identification of stable and declining populations is also strongly influenced by the number of generations that pass between samples. As more time passes, the amount of signal from each population increases and so does the proportion of simulations in which population trend is correctly identified. At the largest abundance ( $N = 10000$ ),  $\hat{N}$  outperforms  $\hat{N}_e$ , but can only detect population decline  $> 70\%$  of the time if at least five generations have passed and  $S = 480$ .

The influence of time and sampling effort, as well as the relative performance of  $\hat{N}_e$  and  $\hat{N}$ , can be seen clearly from the distribution of  $\hat{\lambda}$  in simulations for which  $N = 250$  at  $t_0$  (Fig. 4). The distributions of  $\hat{\lambda}$  estimated from  $\hat{N}_e$  (Panel A) or  $\hat{N}$  (Panel B) are flat, with few defined peaks or evidence of a central tendency, if samples are taken only a generation apart. In contrast, if sampling effort is  $S \geq 60$  and  $t \geq 5$  generations have passed between sampling events, the methods perform fairly well and a well-defined peak emerges with the centre of the  $\hat{\lambda}$  distribution near true  $\lambda$ . Under these biological and sampling conditions, at least 70% of both the  $\hat{N}_e$ - and  $\hat{N}$ -based  $\hat{\lambda}$ s correctly identify population decline or stability (Table 2). In the best cases of large sampling effort ( $S = 120$ ) and maximal time between sampling events (10 generations) at  $N = 250$ , over 90% of the  $\hat{\lambda}$ s correctly identify population trend as either stable or declining.

Although both methods show promise at inferring population trends in simulations where true  $N \leq 250$ , particularly if samples are collected several generations apart, they are much less effective at identifying population trends where initial true  $N \geq 500$  (Table 2). With only one generation between samples, the methods correctly identify population trends for  $N \leq 500 < 60\%$  of the time, regardless of the number of individuals used in this study. At  $N = 500$ , five generations must pass and  $S = 60$  or 120 individuals must be sampled to correctly identify stable or declining population trends  $> 70\%$  of the time using LDNe  $\hat{N}_e$ . However, at the highest levels of sampling effort ( $S = 120$ ) and  $N = 500$ , both  $\hat{N}_e$  and  $\hat{N}$  correctly identify declining and stable populations 88% of the time or better over 10 generations. At  $N = 1000$ , only when sampling effort is high ( $S = 120$ ) and 10 generations pass between samples does either method

**Table 2** Proportion of times population trend was correctly identified in declining or stable populations using the abundance ( $\hat{N}$ ) or effective size methods ( $\hat{N}_e$ ) under a range of initial abundances ( $N$ ), time between sampling periods and number of individuals sampled ( $S$ )

		Declining population						Stable population					
		Gens 0–1		Gens 0–5		Gens 0–10		Gens 0–1		Gens 0–5		Gens 0–10	
		$\hat{N}_e$	$\hat{N}$	$\hat{N}_e$	$\hat{N}$	$\hat{N}_e$	$\hat{N}$	$\hat{N}_e$	$\hat{N}$	$\hat{N}_e$	$\hat{N}$	$\hat{N}_e$	$\hat{N}$
$N = 100$	$S = 30$	57	47	75	61	95	77	55	56	73	55	86	73
	$S = 60$	66	56	–	–	–	–	62	63	89	81	97	97
$N = 250$	$S = 30$	55	14	61	25	78	35	50	26	53	28	60	38
	$S = 60$	55	49	77	71	91	87	63	50	73	71	93	86
	$S = 120$	62	61	90	85	–	–	59	49	87	83	98	94
$N = 500$	$S = 30$	39	5	45	5	57	4	36	9	52	4	52	14
	$S = 60$	46	25	65	44	89	57	50	45	66	41	82	57
	$S = 120$	53	59	83	75	100	96	59	53	73	73	94	88
$N = 1000$	$S = 30$	21	0	26	1	35	1	24	2	25	0	34	4
	$S = 60$	36	11	42	26	58	21	46	21	57	23	56	34
	$S = 120$	49	42	67	65	90	74	51	45	63	56	78	76
$N = 10\,000$	$S = 240$	30	20	25	41	33	41	29	34	27	45	32	39
	$S = 480$	44	51	44	73	71	88	38	51	40	64	56	84

Cases where one or more estimates were infinite were categorized as incorrectly identified. For some combinations of parameters, there were insufficient numbers of individuals to meet sampling size requirements (–).

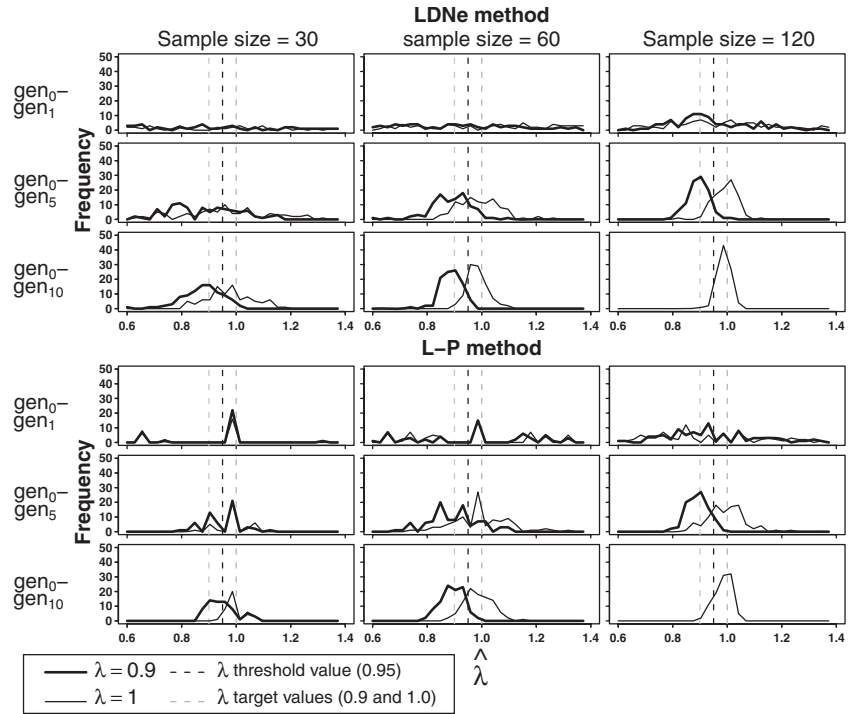
correctly identify population trend > 70% of the time. At this initial abundance,  $\hat{N}$ -based  $\hat{\lambda}$ s are particularly poor for all but the greatest sampling effort and time between sampling events.  $\hat{N}$  performs worse than  $\hat{N}_e$  for nearly all combinations of parameters when initial  $N = 500$  or  $1000$ . In contrast, at  $N = 10000$  and  $S = 240$  or  $480$ , the  $\hat{N}$  estimator outperforms  $\hat{N}_e$ . However,  $\hat{N}$  correctly identifies population trend in more than half of the replicates only when 5–10 generations have passed and  $S = 480$  individuals are sampled.

#### When $N_e < N$

More realistic situations with  $N_e < N$  can be evaluated by comparing the accuracy and precision of  $\hat{N}_e$  at  $N = 100$ – $500$  to that for  $\hat{N}$  at  $N = 1000$ , and this reveals several benefits of estimating  $N_e$  instead of  $N$ . For example, with medium sampling efforts, the  $\hat{N}_e$  estimates are far more likely to be finite and are more tightly clustered around the true value at  $N = 100, 250$  or  $500$  than are the  $\hat{N}$  estimates at  $N = 1000$  (Fig. 2). It is also worthwhile to compare the proportion of times the  $\hat{N}_e$  estimator correctly identifies trend when the initial  $N_e$  value is 0.10 or 0.25 the initial  $N$  value.  $\hat{N}$  correctly identifies a stable population of  $N = 1000$  over 50% of the time only at the largest sample size ( $S = 120$ ; Table 2). By comparison, at  $N = 100$  and  $S = 60$  or  $N = 250$  and  $S = 120$ , using  $\hat{N}_e$  provides a ~30% better chance of correctly identifying a stable population (Table 2).

## Discussion

Some useful guidelines emerge from our simulations of the use of multilocus genotypes to infer population abundance and trends. Under certain sampling and biological conditions,  $\hat{N}_e$  and  $\hat{N}$  can provide useful insights into demography. In general,  $\hat{N}_e$  estimated by LDNe performed better than  $\hat{N}$  estimated by LP for trend detection over most of the sampling and biological conditions we simulated, even when we assumed  $N_e/N = 1$ . However, there are some formidable limitations that should be recognized if using  $\hat{N}_e$  or  $\hat{N}$  to infer population status or trend. The primary limitation is that with genotypes from only  $S = 30$  individuals,  $\hat{N}_e$  and  $\hat{N}$  are likely to be biased and imprecise, whether  $N = 100$  or  $1000$ . Our simulations indicate that to have a reasonable chance of making useful inferences about abundance from  $\hat{N}_e$  or  $\hat{N}$ , 60 or more individuals should be sampled. At small to moderate population sizes ( $N < 500$ ),  $\hat{N}_e$  is more precise and accurate than  $\hat{N}$  for making inferences about population status under the conditions we simulated. The slight positive bias and strong precision of  $\hat{N}_e$  at small abundances are consistent with simulations by others (Waples & Do 2009). At larger population sizes ( $N = 500$ – $1000$ ),  $\hat{N}$  is more precise than  $\hat{N}_e$  under most of the conditions we simulated, assuming  $N = N_e$ . At large  $N = 10000$ ,  $\hat{N}_e$  is not very useful unless  $N_e \ll N$ , because genetic drift and drift-induced gametic disequilibrium are almost nonexistent. At this



**Fig. 4** Distribution of the  $\hat{N}_e$  (Panel A) and  $\hat{N}$  (Panel B) based  $\hat{\lambda}$  estimates of population growth rate ( $\lambda = 0.9$  or  $1.0$ ) under a range of sampling conditions for  $N = 250$  at  $t_0$  for sampling intervals of 1, 5 or 10 generations. Thick solid lines indicate distributions for declining populations and thin solid lines indicate distributions for stable populations.

population size,  $\hat{N}$  can be fairly precise and accurate if a large sample ( $S = 480$ ) can be obtained.

To successfully identify population trend in a population of initial size  $N = 100$ – $500$ , samples of 60 individuals should be taken  $\geq 5$  generations apart. Samples taken 1 generation apart or samples of 30 individuals will rarely provide accurate insights into population trend, which is disappointing but not too surprising. In real populations, it would be difficult to know whether a change in abundance over one generation, even if estimated without error, was because of natural fluctuations or something more dramatic. Obviously, with more time between samples and larger samples, our ability to accurately identify stable or declining populations improves. Increased time between samples increases the signal in the data, and increased sampling effort increases the signal to noise ratio in the samples. A useful rule of thumb might be to obtain samples of at least 60 individuals more than a generation apart to monitor populations of 100–500 individuals.

*Practical considerations*

Several important considerations should be addressed when designing a real-world study around these simulation results. First, although the rate of population decline we modelled was moderately strong ( $\lambda = 0.9$ ), this is a per generation decline. In real populations of long-lived species, there may be a much smaller annual decline that

translates into an equivalent per generation decline. Second, we used only moderately polymorphic microsatellite loci in our simulations. It may be possible to increase statistical power to infer abundance or population trend by targeting loci with the greatest amount of diversity. That is, using highly polymorphic loci will provide more alleles to estimate linkage disequilibrium and to obtain unique genotypes for abundance estimation. However, this benefit of high polymorphism should be tempered by considerations of genotyping errors that should be addressed with laboratory (McKelvey & Schwartz 2004) and modelling (Lukacs & Burnham 2005a) efforts.

Our simulations directly compared  $N_e$  and  $N$  estimators under identical sampling conditions. However, there are some important departures from these conditions in real populations that should help researchers studying real populations. For example, we have simulated a W-F-like population in which  $N_e$  is close to  $N$ . In many natural populations,  $N_e < N$  (Frankham 1995; Palstra & Ruzzante 2008). Therefore,  $N_e$  may provide much more precise and useful estimates for making demographic inferences, as our simulations showed. On one hand,  $N_e$  is not  $N$ , and there may be reasons to track  $N$  directly because it may be more directly related to short-term management guidelines, considerations, triggers or thresholds. A perhaps more important caveat is that changes in  $N_e$  could reflect changes in the  $N_e/N$  ratio as a result of altered mating system or age structure instead of (or in addition to) changes in  $N$  (Palstra & Ruzzante 2008). On the other

hand,  $N_e$  provides useful insights into the potential for loss of genetic diversity and evolutionary potential.

In populations with high fecundity and Type III survivorship, it is frequently easy to obtain large samples of particularly abundant stage classes, such as juveniles. Large samples will increase the precision of  $N_e$  estimators and may favour their use for monitoring natural populations with such life histories. We will investigate this situation in a forthcoming article.

In addition, we used sample sizes typical of population genetics studies ( $S = 30\text{--}120$  individuals) in our simulations. This meant that, in many instances, a fairly small proportion of the simulated population was sampled, which translates into a low probability of detection in capture-recapture parlance. If probabilities of detection are low, capture-recapture estimators are imprecise and model selection algorithms will be inaccurate (Menkins & Anderson 1988; McKelvey & Pearson 2001). In contrast, in situations where a higher probability of detection is possible because of the characteristics of the species studied or the sampling design, more sophisticated capture-recapture methods (Lukacs and Burnham 2005a, b; Boulanger *et al.* 2006; Schofield & Barker 2008) than the simple LP method used here can be applied, and more precise and accurate estimates can be obtained. This highlights the importance of understanding the underlying biology and sampling limitations of a target population.

Finally, our interest here was in comparing the relative performance of simple  $N_e$  and  $N$  estimates to make inferences about population demography, so we compared and contrasted LDNe and LP. However, these methods and others (Schofield & Barker 2008; Tallmon *et al.* 2008; Wang 2009; Waples & Do 2009) could be used on the same data sets to obtain more information about contemporary population demography. In addition, temporal  $N_e$  estimators could be used in many instances when three or more genetic samples are obtained across cohorts or generations. Obviously, it would be wise to use as much information as can be extracted from multilocus genotype data sets by combining insights from  $N_e$  and  $N$  for any population, while also considering the assumptions that come with each approach. Future efforts that address how information on  $N_e$  and  $N$  can be used together to maximize the inferences about population status and trends would be especially helpful, perhaps through the use of open population models (Lukacs & Burnham 2005b) that incorporate recruitment and survival in  $N$  estimates combined with one sample and temporal  $N_e$  estimators.

## Acknowledgements

This work was conducted as part of the Genetic Monitoring (GeM) Working Group jointly supported by the National Evolu-

tionary Synthesis Center (NSF #EF-0423641) and the National Center for Ecological Analysis and Synthesis, a centre funded by NSF (NSF #EF-0553768), the University of California, Santa Barbara, and the State of California. We thank all members of the NCEAS/NESCent Genetic Monitoring Working Group who contributed useful input and ideas to this effort. We also thank Paul Lukacs and anonymous reviewers for helpful comments on an earlier draft of this manuscript.

## References

- Aspi J, Roininen E, Ruokonen M, Kojola I, Vila C (2006) Genetic diversity, population structure, effective population size and demographic history of the Finnish wolf population. *Molecular Ecology*, **15**, 1561–1576.
- Bellemain E, Swenson JE, Tallmon DA, Brunberg S, Taberlet P (2005) Estimating population size of elusive animals with DNA from hunter-collected feces: comparing four methods for brown bears. *Conservation Biology*, **19**, 150–161.
- Boulanger J, Proctor M, Himmer S *et al.* (2006) An empirical test of DNA mark-recapture sampling strategies for grizzly bears. *Ursus*, **17**, 149–158.
- England PE, Luikart G (1999) Statistical analysis of microsatellite data. *Trends in Ecology and Evolution*, **14**, 253–256.
- Frankham R (1995) Effective population size/adult population size ratios in wildlife: a review. *Genetical Research*, Cambridge, **66**, 95–107.
- Goossens B, Chikhi L, Ancrenaz M *et al.* (2006) Genetic signature of anthropogenic population collapse in orang-utans. *Public Library of Science Biology* **4**, e25. doi:10.1371/journal.pbio.0040025.
- Hill WG (1981) Estimation of effective population size from data on linkage disequilibrium. *Genetical Research*, Cambridge, **38**, 209–216.
- Kendall KC, Stetz JB, Waits LP, Boulanger JB, Paetkau D (2008) Grizzly bear density in Glacier National Park, Montana. *Journal of Wildlife Management*, **72**, 1693–1705.
- Kuhner MK (2008) Coalescent genealogy samplers: windows into population history. *Trends in Ecology and Evolution*, **24**, 86–93.
- Laval G, Excoffier L (2004) SIMCOAL 2.0: a program to simulate genomic diversity over large recombining regions in a subdivided population with a complex history. *Bioinformatics*, **20**, 2485–2487.
- Lukacs PM, Burnham KP (2005a) Estimating population size from DNA-based closed capture-recapture data incorporating genotyping error. *Journal of Wildlife Management*, **29**, 396–403.
- Lukacs PM, Burnham KP (2005b) Review of capture-recapture methods applicable to noninvasive genetic sampling. *Molecular Ecology*, **14**, 3909–3919.
- Martien KK, Gregovich D, Bravington MV *et al.* (2009) TOSSM: an R package for assessing performance of genetic analytical methods in a management context. *Molecular Ecology Resources*, **9**, 1456–1459.
- McKelvey KS, Pearson DE (2001) Population estimation with sparse data: the role of estimators versus indices revisited. *Canadian Journal of Zoology*, **79**, 1754–1765.
- McKelvey KS, Schwartz MK (2004) Genetic errors associated with population estimation using non-invasive molecular tag-

- ging: problems and new solutions. *Journal of Wildlife Management*, **68**, 439–448.
- Menkins GE, Anderson SH (1988) Estimation of small mammal population size. *Ecology*, **69**, 1952–1959.
- Mills LS, Citta JJ, Lair KP, Schwartz MK, Tallmon DA (2000) Estimating animal abundance using noninvasive data sampling: promise and pitfalls. *Ecological Applications*, **10**, 283–294.
- Palstra FP, Ruzzante DE (2008) Genetics estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, **17**, 3428–3447.
- Robinson SJ, Waits LP, Martin ID (2009) Estimating abundance of American black bears using DNA-based capture-mark-recapture models. *Ursus*, **20**, 1–11.
- Schofield MR, Barker RJ (2008) A unified capture-recapture framework. *Journal of Agricultural, Biological, and Environmental Statistics*, **13**, 458–477.
- Schwartz MK, Tallmon DA, Luikart G (1998) Review of DNA-based census and effective population size estimators. *Animal Conservation*, **1**, 293–299.
- Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution*, **22**, 25–33.
- Seber GAF (1982) *The Estimation of Animal Abundance and Related Parameters*, 2nd edn. Macmillan Publishing Co., Inc., New York, N.Y.
- Strand AE (2002) Metasim 1.0: an individual-based environment for simulating population genetics of complex population dynamics. *Molecular Ecology Notes*, **2**, 373–376.
- Tallmon DA, Koyuk A, Luikart G, Beaumont MA (2008) ONeS-AMP: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources*, **8**, 299–301.
- Wang J (2009) A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Molecular Ecology*, **18**, 2148–2164.
- Waples RS (2005) Genetic estimates of contemporary effective population size: to what time periods to estimates apply? *Molecular Ecology*, **14**, 3335–3352.
- Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, **7**, 167–184.
- Waples RS, Do C (2008) LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, **8**, 753–756.
- Waples RS, Do C (2010) Single linkage disequilibrium estimates of contemporary  $N_e$  using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, doi:10.1111/j.1752-4571.2009.00104.x.
- Waples RS, Faulkner JR (2009) Modeling evolutionary processes in small populations: not as ideal as you think. *Molecular Ecology*, **18**, 1834–1847.