INVITED REVIEWS AND META-ANALYSES
Inferring recent historic abundance from current genetic diversity

PER J. PALSBOLL,*† M. ZACHARIAH PEERY,‡ MORTEN T. OLSEN,§ STEVEN R. BEISSINGER¶ and MARTINE BERUBE*†
*Marine Evolution and Conservation, Centre of Evolutionary and Ecological Studies, University of Groningen, PO Box 11103 CC, Groningen, The Netherlands, †Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA, ‡Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, 1630 Linden Drive, Madison, WI 53703, USA, §Department of Genetics, Microbiology and Toxicology, Stockholm University, SE-106 91, Stockholm, Sweden, ¶Department of Environmental Science, Policy and Management, University of California Berkeley, 130 Mulford Hall, Berkeley, CA 94720-3110, USA

Abstract
Recent historic abundance is an elusive parameter of great importance for conserving endangered species and understanding the pre-anthropogenic state of the biosphere. The number of studies that have used population genetic theory to estimate recent historic abundance from contemporary levels of genetic diversity has grown rapidly over the last two decades. Such assessments often yield unexpectedly large estimates of historic abundance. We review the underlying theory and common practices of estimating recent historic abundance from contemporary genetic diversity, and critically evaluate the potential issues at various estimation steps. A general issue of mismatched spatio-temporal scales between the estimation itself and the objective of the estimation emerged from our assessment; genetic diversity–based estimates of recent historic abundance represent long-term averages, whereas the objective typically is an estimate of recent abundance for a specific population. Currently, the most promising approach to estimate the difference between recent historic and contemporary abundance requires that genetic data be collected from samples of similar spatial and temporal duration. Novel genome-enabled inference methods may be able to utilize additional information of dense genome-wide distributions of markers, such as of identity-by-descent tracts, to infer recent historic abundance from contemporary samples only.

Keywords: coalescent, conservation, mutation–migration–drift equilibrium, population genetic inference

Received 7 April 2012; revision received 1 September 2012; accepted 17 September 2012

Introduction
Abundance, the number of individuals in a population, is a fundamental parameter in evolution, ecology and conservation. Although there are a variety of methods to estimate current abundance (e.g. Lebreton et al. 1992; Schwarz & Seber 1999; Lukacs & Burnham 2005; Thomas et al. 2010), options to estimate recent historic abundance are far more restricted, limiting assessments of long-term trends in abundance (e.g. Koschinski 2002; Smith & Reeves 2003; O’Connell et al. 2007; Clucas 2011; Lund et al. 2011). When a population undergoes demographic change, the reduction in genetic diversity is typically subject to a time delay because the rates of processes that alter genetic diversity (i.e. random genetic drift, mutation and selection) change more slowly than demography. This ‘time lag’ is apparent in the current global human population, which harbours much lower levels of genetic diversity than expected from the current abundance, reflecting its recent expansion from a much smaller ancestral population...
Examples of genetic diversity–based estimations of recent historic abundance

During the last two decades, a number of studies have employed estimates of genetic diversity in contemporary populations to infer recent historic abundance of endangered species. Here, we review three high-profile examples that applied this approach to whales (Roman & Palumbi 2003), wolves (Leonard et al. 2005) and extinct moas (Gemmell et al. 2004) to illustrate the general approach and the inferences drawn from the results.

The large whales have long served as an example of anthropogenic endangerment due to ill-managed exploitation. Abundance of most populations of large whales across the globe was reduced significantly by whaling during the 19th and 20th century. However, the degree of the reduction in abundance caused by whaling remains elusive, as pre-whaling abundances are essentially unknown. In 2003, Roman and Palumbi used previously published mitochondrial control region DNA sequences (published by Andersen et al. 2003; Baker et al. 1993; Palsbøll et al. 1995; Bérubé et al. 1998) to infer the historic abundance of humpback whales (Megaptera novaeangliae), fin whales (Balaenoptera physalus) and minke whales (Balaenoptera acutorostrata) in the North Atlantic from current levels of genetic diversity (Roman & Palumbi 2003). Their genetic diversity–based estimates of recent historic abundance for all three species were substantially higher than the estimates of current abundance, in one case by as much as ~24 times (humpback whale). Roman & Palumbi (2003) inferred their genetic diversity–based abundance estimates to represent recent historic (and not current) abundances of these three species, that is, the abundance just prior to the onset of whaling. The genetic diversity–based estimate of pre-whaling abundance in the humpback whale was also substantially higher than other pre-whaling estimates based upon nongenetic data (e.g. backcasting from models using estimated catches, demographic parameters and current abundance estimates). The upper bound from these nongenetic estimates of pre-whaling abundance in the North Atlantic was 25 800 humpback whales (Punt et al. 2006), which was ~10 times lower than the genetic diversity–based estimate at 240 000 humpback whales reached by Roman & Palumbi (2003). Implications of the genetic diversity–based estimates of pre-whaling abundance are profound. If the recent historic abundances in the North Atlantic of 240 000 humpback whales, 360 000 fin whales and 265 000 minke whales (Roman & Palumbi 2003) are correct, they suggest that large whales in the North Atlantic have recovered to a much lower degree than currently assumed. High genetic diversity–based estimates of abundance, inferred as representing recent historic abundance, have also been reported for minke whales in the Southern Ocean (Ruegg et al. 2010), eastern North Pacific Ocean grey whales, Eschrichtius robustus (Alter et al. 2007), and indirectly in Southern Ocean blue whales (Balaenoptera musculus intermedia) (Sremba...
et al. 2012). A larger historic biomass of whales also implies a much greater historic biomass of the prey base, suggesting a significant baseline shift in ocean productivity during the last century (Jackson 2001).

Leonard et al. (2005) reached similar conclusions from their genetic diversity-based analysis of historic abundance for the North American grey wolf, Canis lupus. The grey wolf is a textbook example of a terrestrial carnivore subjected to intense eradication programs during the last century, resulting in a greatly reduced abundance but to an unknown extent. Leonard et al. (2005) sequenced the mitochondrial control region in grey wolf samples from 90- to 150-year-old specimens assumed to represent the pre-exploitation population in the contiguous U.S. Abundance from these historic samples was estimated at 380,000 wolves, much higher than the contemporary abundance in North America, which has been estimated at 60,000 to 70,000 individuals. As was the case for the large whales, Leonard et al. (2005) suggested that the conservation goals for U.S. grey wolves be adjusted upward to reflect their genetic diversity-based estimates of recent historic abundance. The ecological ramifications of the findings by Leonard et al. (2005) are similar to those of Roman & Palumbi (2003); more wolves imply more abundant prey and suitable habitat in the recent historic past.

Our last example of genetic diversity-based estimates of recent historic abundance is for the moa, Dinornis spp., an extinct flightless ratite bird that was endemic to New Zealand. Nongenetic analyses have suggested that abundance prior to human settlement (1000 to 6000 years BP) was at 159,000 individuals (Gemmell et al. 2004). Based upon previously published DNA sequences (Huynen et al. 2003), Gemmell et al. (2004) arrived at genetic diversity-based estimates of 300,000 to 1.4 million individuals between 1000 and 6000 years BP and 3–12 million individuals >6000 years BP. Gemmell et al. (2004) interpreted their findings to suggest that the demise of Dinornis spp. pre-dates human settlement and the Moa’s drastic pre-human settlement decline probably was caused by diseases and epizootics introduced to the islands by immigrant birds.

These three examples, which are typical for the majority of genetic diversity-based estimates of abundance, illustrate two common features: (i) genetic diversity-based estimates of recent historic abundance are generally much larger than estimates of both the current abundance and the presumed historic abundance and (ii) genetic diversity-based estimates are usually interpreted as representing abundance at (or until) a specific point in time and often for a specific population, as well. Some genetic diversity-based estimates of recent historic abundance have since been cited as the de facto historic abundance in policy-oriented literature discussing current conservation and management policies (e.g. Myers & Worm 2005; Bolster 2006; Pinnegar & Engelhard 2008; Lotze & Worm 2009; Bonebrake et al. 2010; Pershing et al. 2010).

Estimating recent historic abundance from contemporary levels of genetic diversity

The popularity of the genetic diversity-based approach is due to the simple and straightforward relationship between the number of mutations (denoted \( \theta \)) expected for any two random gene copies as the product of the per-generation mutation rate (\( \mu \)) and the effective population size of the targeted genome (\( N_e \)) in a single idealized population. Population genetic theory predicts the following expectation,

\[
\theta = x N_e \mu
\]

where the value of \( x \) depends upon the ploidy and mode of inheritance for the locus under investigation (Ewens 1972; Watterson 1984; Hudson 1990). For a diploid, autosomal and Mendelian-inherited locus \( x \) equals 4 and \( N_e \) the effective population size of both sexes. In the case of a haploid, uniparentally inherited locus (such as the maternally transmitted vertebrate mitochondrial genome) \( x \) equals 2. In this latter case, \( N_e \) denotes the effective population size of the relevant sex (e.g. females in the case of the vertebrate mitochondrial genome). Consequently, if the value of \( \mu \) is known, then it is, in principle, possible to infer the effective population size from an estimate of \( \theta \) when the assumptions underlying the above relationship between \( \theta, N_e \) and \( \mu \) are upheld. The final estimate of abundance (i.e. census population size, \( N_c \)) is inferred from the estimate of \( N_e \) by correction for the difference between \( N_c \) and \( N_e \). Putting it all together, we obtain the following relationship,

\[
\text{Historic } N_c = \left[ \frac{N_c}{N_e} \right] \frac{\theta}{x \mu} \quad (2)
\]

The rationale for treating \( N_c \) as an estimate of recent historic (and not current) abundance (even though \( N_c \) is derived from the current observed genetic diversity) is due to the observation that a recent reduction in abundance has little, if any, effect on some measures of genetic diversity such as nucleotide diversity (\( \pi \), Tajima 1989), heterozygosity (Cornuet & Luikart 1996) and the range of allele sizes [for simple tandem repeated (STR), loci, Garza & Williamson 2001]. Figure 1 depicts the rate of decline in \( \theta \) (estimated as \( \pi \)) following an instantaneous 95% reduction in the effective size of a population when \( \theta = 6.0 \). At 200 and 2000 generations since the reduction to 5% of the original population size, the mean estimate of \( \theta \) is only reduced by 7% and 50%,
respectively. The estimate of $\theta$ is thus much higher than the expected value of 0.3 for the current abundance. Most genetic diversity–based estimations of recent historic abundance are concerned with much more recent timescales (i.e. <20 generations).

In Table 1, we have outlined the overall estimation procedure and each step necessary to go from an estimate of contemporary genetic diversity (i.e. $\theta$) to an estimate of historic $N_c$ using eqn 2. There are three main components of eqn 2: $\theta$, $\mu$ and finally the ratio of $N_e$ to $N_c$, each of which is estimated separately, using either population genetic, phylogenetic, fossil or contemporary demographic data as well as combinations thereof. In the following section, we briefly outline how each of these estimations is conducted. We then provide a more detailed discussion of some of the potential issues in each of these estimations. Finally, we conclude with a more general discussion of the applicability of the final ‘historic’ abundance estimate as well as alternative means to obtain more reliable estimates of recent historic abundance from genetic data.

**Inferring $N_c$ from an estimate of $\theta$**

An estimate of $\theta$ can be obtained analytically in cases of a single isolated, unstructured population assuming a simple mutation model (Fig. 2, models A.1), such as the infinite site mutation model (Fig. 2, models A.1), such as the infinite site mutation model (i.e. no nucleotide position is subject to recurrent mutations, Kimura 1969). In this

<table>
<thead>
<tr>
<th>Parameter in eqn 2</th>
<th>Temporal scale of estimation</th>
<th>Data</th>
<th>Analysis</th>
<th>Estimate of</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>Interspecific</td>
<td>Interspecific phylogenetic data</td>
<td>e.g. MODELTEST (Posada &amp; Crandall 1998)</td>
<td>Mutation model parameters</td>
</tr>
<tr>
<td></td>
<td>Interspecific</td>
<td>Interspecific fossil dates and phylogenetic data</td>
<td>e.g. BEAST (Drummond &amp; Rambaut 2007)</td>
<td>Annual mutation rate</td>
</tr>
<tr>
<td></td>
<td>Contemporary</td>
<td>Survival and reproductive rates</td>
<td></td>
<td>Generation time</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Intraspecific</td>
<td>Intraspecific population genetic data</td>
<td>e.g. LAMARC (Kuhner et al. 2004), MIGRATE (Beerli &amp; Felsenstein 2001), MNa2 (Hey 2010b), BATWING (Wilson et al. 2003)</td>
<td>Mutation model parameters $\theta$ (estimation requires specification of mutation model)</td>
</tr>
<tr>
<td></td>
<td>Intraspecific</td>
<td>Intraspecific population genetic data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_e/N_c$</td>
<td>Contemporary</td>
<td>Sex-specific reproductive and mortality rates</td>
<td>e.g. MARK (White &amp; Burnham 1999)</td>
<td>Contemporary $N_e$</td>
</tr>
<tr>
<td></td>
<td>Contemporary</td>
<td>Population abundance data</td>
<td>e.g. MARK (White &amp; Burnham 1999), DISTANCE (Thomas et al. 2010)</td>
<td>Contemporary $N_e$</td>
</tr>
</tbody>
</table>
Fig. 2 Population models underlying the most common approaches to infer local genetic diversity. Absence of a symbol implies that the parameter is not specified in the model. Equal-sized population circles (under the heading $\theta$) or migration rate arrows (under the heading $m$) denote the assumption of identical population sizes and/or migration rates among populations. Different symbol sizes denote that a rate may differ among and in populations. Although an exponential growth rate ($g$) is depicted, the growth rates may be (for some implementations) linear or instantaneous. Different curves for $g$ in the same model imply that $g$ is estimated separately for each population. $T$ denotes population divergence time. Examples of approaches for the different population models are as follows: Model A.1: Genetic diversity summary statistics, such as $\pi$ (Nei & Li 1979), number of segregating sites (Tajima 1989) and the variance in number of repeats at STR loci (Di Rienzo et al. 2007), MSVAR (Beaumont 1999), FLUCTUATE (Kuhner et al. 1998) and $\tau$ (Rogers & Harpending 1992). Model B.1: Likelihood or approximate Bayesian computation (ABC) methods with only two populations with symmetric gene flow and population sizes, such as M2V (Nielsen & Wakeley 2001). Model B.2 Likelihood or ABC methods with two or more populations, such as MIGRATE-N (Beerli & Felsenstein 2001). Model B.3: Likelihood or ABC methods with two or more populations, such as LAMARC (Kuhner et al. 2004). Model B.4: Likelihood or ABC methods with two or more populations, such as 1Ma2 (Hey 2010a,b) and BATWING (in this implementation divergence times are fixed priors and growth rates identical across populations, Wilson et al. 2003).
**GENETIC ESTIMATION OF HISTORIC ABUNDANCE** 27

**Deriving \( N_e \) from an estimate of \( N_c \)**

The estimate of ‘abundance’ derived from eqn 1 is the effective population size (i.e. \( N_e \)). The effective population size in population genetics is defined as the size of an ‘ideal’ population with the observed rate of genetic drift (Wright 1931, 1969) and does thus not represent census population size, the parameter of interest in terms of historic and contemporary abundance. In ecology, the effective population size is usually thought of as the number of breeding individuals that successfully transmit their genes to the next generation (Frankham 1995) and thus should equal the ‘genetic’ effective population size. In practice, however, the exact relationship is rarely known. Consequently, most studies apply a generic ratio representing a range of estimates (often between 1:4–1:10) obtained from a variety of species (e.g. Frankham 1995; Roman & Palumbi 2003).

Several studies that have inferred recent historic abundance from contemporary genetic diversity in vertebrates were based solely or partly upon DNA sequences from the mitochondrial genome (Hoelzel et al. 1993; Kokko et al. 1999; Roman & Palumbi 2003; Gemmell et al. 2004; Leonard et al. 2005; Alter et al. 2007). Because the mitochondrial genome in many species is haploid and maternally inherited, a correction of \( N_c \) to \( N_e \) for sex ratio (and ploidy) is required as well. The sex ratio and ploidy is usually well described in contemporary populations.

**The main issues**

**Estimating \( \theta \): Temporal issues**

Natural populations are subjected to changing ecological and environmental conditions (e.g. Willerslev et al. 1999, 2003; Shapiro et al. 2004; Campos et al. 2010a,b; Lorenzen et al. 2011). As a result, abundance, migration rates and selective agents change over time and space within and among natural populations. The highly variable dynamics experienced by natural populations contrast with the constancy of the in silico populations upon which genetic diversity–based estimators of recent historic abundance are based (see Fig. 2). These highly simplified assumptions are necessary to make the estimation of \( \theta \) tractable and computable. Accordingly, most in silico populations are not subject to changes in rates in time or space but instead have constant growth (\( \lambda \)), migration (\( m \)) and mutation rates. The consequence of this difference between the in vivo and in silico populations is that the single and final estimate of \( \theta \) (as well as growth and migration rates, depending on the estimation procedure) represents the most likely, weighted-mean value of \( \theta \) (and, when applicable, growth and migration rates) over the tempo-spatial range covered by the estimation procedure. In other words, the estimate of \( \theta \) is not necessarily an estimate of \( \theta \) that applies to a specific point in time and place. We deliberately use the rather vague, term ‘weighted mean’ because it is not clear what kind of mean the final estimate of \( \theta \) (and \( N_e \)) represents. In a single isolated population, current \( N_e \) inferred from genetic diversity is the harmonic \( N_e \) (and thus heavily influenced by past bottlenecks). However, in the case of populations subject to both fluctuating demographic and genetic rates, such as migration and mutation rates, this may not necessarily be the case, which is why we use a ‘weighted mean’ instead of the median.

The time frame of an estimate of \( \theta \) is the time that has elapsed since the most recent common ancestor, which is \( 2N_e \) and \( 4N_e \) generations for haploid and diploid loci, respectively, where \( N_e \) differs depending on the mode of inheritance. The time to the most recent common ancestor is thus considered in many cases. One example is the aforementioned estimate of pre-whaling abundance in North Atlantic whales by Roman & Palumbi (2003). Their estimate of \( \theta \) (and \( \mu \)) suggested an effective female population size in North Atlantic humpback whales at -34 000. Assuming a generation time of 24 years (Barlow & Clapham 1997), the expected time to the most recent common ancestor is 1.6 Myr. Roman & Palumbi (2003) used the coalescent-based MCMC-likelihood method implemented in MIGRATE (Beerli & Felsenstein 2001) to estimate \( \theta \), which assumes constant migration rates among populations and no population growth (i.e. constant population sizes). Values for \( mN_e \) (i.e. the number if migrants per generation) and \( \theta \) estimated in this manner thus represent the mean values for which the observed data are most likely over a period of approximately 1.6 Myr.

An illustrative example is an assessment of bias for the estimation of \( mN_e \) and \( \theta \) due to recent population divergence conducted by Beerli (2010). The MCMC coalescent-based estimation procedure implemented in MIGRATE-n ignores population divergence time, which essentially equals a change of \( mN_e \) from \( \infty \) to a lower value at some point in the past. Using simulated data with \( \theta \) at 0.0025 and varying population divergence times and migration rates, Beerli’s assessment revealed that the effect of ignoring population divergence time prior to \( N_e / 2 \) generations ago on \( \theta \) was negligible, whereas this was not the case for more recent population divergence times. Using the estimates in Roman & Palumbi (2003) as an illustrative example, Beerli’s preliminary assessment translates into an ‘effect’ time at 408 000 years. The very long time frame involved in this and similar estimations implies that, although it is, in principle, possible to model and thus evaluate most biologically feasible deviations of assumptions causing bias, our knowledge of the relevant
population dynamics during such long time frames is very limited and insufficient to do so.

Estimating $\theta$: Spatial issues

Migrants remove or add genetic diversity to or from local populations, which must be accounted for when the objective is an estimate of local (as opposed to global) abundance from estimates of genetic diversity. Hudson (1990) demonstrated using coalescent-based simulations that the mean estimate of $\theta$ in a sample of gene copies collected in a single population is equal to the sum of $\theta$ for all populations combined. This result applies for the case of high ($4mN_e = 1000$) and low ($4mN_e = 0.2$) migration rate (see Fig. 3). In other words, when the level of gene flow is high, all populations are essentially part of the same global population and this is reflected in the level of genetic diversity in each population. In contrast, when $mN_e$ is low, immigrant gene copies are highly divergent from the resident gene copies in the receiving population due to the combined action of mutation and genetic drift. As a result, each immigration event will result in a substantial increase in local genetic diversity (from which $N_e$ is inferred) that does not correspond to the relative demographic increase in local abundance due to immigrant individuals. Consequently, when migration is likely to occur, even at low values of $mN_e$, an estimation of local $N_e$ from $\theta$ must account for migration in order for the estimate to reflect local rather than global abundance. The coalescent-based MCMC approaches implemented in MIGRATE-n (Beerli & Felsenstein 2001), BATWING (Wilson et al. 2003), LAMARC (Kuhner et al. 2004) or ima2 (Hey & Nielsen 2004) all yield joint estimates of $\theta$ and migration rates in a full matrix of populations, which may also be modelled in an ABC framework.

Relevant, but nonsampled populations are known as ghost populations (Beerli 2004; Slatkin 2005). The most extreme example of ghost populations would be extinct populations, which in most cases will be unknown and unavailable for sampling. The example (from simulated data) depicted in Fig. 4 illustrates that the effects on local genetic diversity due to past migration to and from extinct ghost populations may linger for thousands of years. Even for the species that we probably know best, Homo sapiens, new studies constantly change our recent migration history, such as revising the initial immigration of humans into East Asia from 25 000–38 000 BP to 50 000 BP (Rasmussen et al. 2011). Our knowledge of the past in natural populations of nonmodel species is far poorer. Ghost populations may be an issue even for large, relatively easy-to-observe species, such as baleen whales, which are and have been subject of intense human attention. Several recent observations in the baleen whales discussed below illustrate that the presence of such ghost populations is a real phenomenon.

![Fig. 3](image1.png)

**Fig. 3** The effect of migration rates on local levels of genetic diversity ($\theta$). Simulations were conducted to emulate the findings presented in Hudson (1990). Data were generated using coalescent-based simulation software ms (Hudson 2002). Sequence data were generated under an Island model with three populations (each with $\theta = 6$) connected by different levels of gene flow. A total of 10 000 data sets were generated for each migration value. Bars denote the 95% confidence interval of the estimate of $\theta$ estimated from a sample of 200 gene copies collected from the target population.

![Fig. 4](image2.png)

**Fig. 4** The lingering effect on local genetic diversity of past connectivity. Data were generated using coalescent-based simulation software ms (Hudson 2002). Data were simulated for two populations each at $\theta = 6$ connected by a migration rate of 0.1, 1.0 or 10, respectively, prior to generation zero, after which gene flow was zero. For each estimate, 100 data sets were generated under an infinite sites mutation model, and the value of $\theta$ for the single extant population was estimated as $\pi$ (Nei & Tajima 1981).
abundance estimated to be ~20,000 individuals (Alter et al. 2007). Even though the grey whale type specimen is from the North Atlantic, where several grey whale fossils have been located, grey whales were believed to have gone extinct in the North Atlantic during the 1600s (Reeves & Mitchell 1988). Alter et al. (2007) dismissed any effects of gene flow into the North Pacific from the North Atlantic because migration between the North Atlantic and Pacific Ocean was deemed infeasible during the last 114,000 years due to the density of sea ice in the Arctic Ocean. However, during the summer of 2010, a live grey whale was observed in the Mediterranean Sea (Gazo & Chicote 2010). The population origin of the sighted grey whale is unresolved, but the sighting suggests that migration between the North Pacific and North Atlantic possibly supports recent work by Funder et al. (2011) and Antoniades et al. (2011), suggesting much lower levels of sea ice in the Arctic Ocean 4000–8000 years before present compared to contemporary sea ice levels. A similar observation was made in the common minke whale (B. bonaerensis). Ruegg et al. (2010) estimated pre-whaling abundance of the common minke whale in the Southern Hemisphere, the only ocean in which this species had been observed, thus greatly simplifying the estimation because population structure could be reasonably ignored (Ruegg et al. 2010). In 2011, Glover et al. (2010) reported two tissue samples collected from North Atlantic minke whales with mitochondrial DNA sequences matching the common minke whale, one of which appeared to be a first-generation hybrid between the two minke whale subspecies (Glover et al. 2010). Phylogeographic analyses in humpback whales and fin whales also suggest occasional trans-equatorial and trans-oceanic migrations (Baker et al. 1993; Palsbøll et al. 1995; Bérubé et al. 1998).

The above examples illustrate clearly the incompleteness in our basic knowledge in aspects that may be critical to inferring recent historic abundance from current levels of genetic diversity, even in large seemingly readily detected and intensively studied organisms.

**Issues in estimating μ**

An estimate of μ, the generational mutation rate, is required to make inferences of abundance from estimates of current genetic diversity (eqn 1). Studies published to date have relied upon phylogenetic estimates of μ, that is, interspecific calibrations (e.g. Roman & Palumbi 2003; Gemmell et al. 2004; Leonard et al. 2005; Alter et al. 2007; Ruegg et al. 2010). This approach is common in taxonomic studies relying on known fossil dates as interspecific time calibration points to get annual mutation rates. In the case of taxonomy, the objective (i.e. interspecific divergence times) and the estimation method (i.e. phylogenetic calibration of mutation rates using known interspecific time points) are both on similar temporal scales (e.g. Hasegawa et al. 1985; Ball & Avise 1992; Brunner et al. 2001; Thorne & Kishino 2002; Barker et al. 2004; Yang & Rannala 2006; Meredith et al. 2011). When estimating recent historic abundance, however, there is a substantial difference between the temporal time frame of the estimation (i.e. of μ) and the objective (i.e. recent historic abundance). The possible effect of this difference in time frames is exacerbated by the fact that relatively rapidly evolving loci are utilized in intraspecific studies.

A case in point is the mitochondrial control region, arguably the most common and widespread DNA sequence employed in intraspecific assessments of genetic diversity in nonmodel vertebrates (Moritz 1994; Palumbi et al. 2009). The mitochondrial control region exhibits highly heterogeneous mutation rates among sites, with so-called hyper-variable regions dominating in the 5′ end (Aquadro & Greenberg 1983; Hoelzel 1993; Ho et al. 2011). A general discrepancy has emerged when comparing the mutation rates obtained in the traditional manner (i.e. the phylogenetic approach) to the rates estimated from sequences collected in pedigrees of closely related individuals (Howell et al. 2003a,b; 2005; Irwin et al. 2009; Loogvali et al. 2009; Phillips et al. 2009; Goto et al. 2011; Klutsch et al. 2011) as well as temporarily spaced samples, in so-called ancient DNA analyses (Ho et al. 2007, 2008). In general, the mutation rates derived from pedigrees or ancient DNA studies have been ~10 times higher than the phylogenetic-derived rates (e.g. Millar et al. 2008). Possible explanations for this discrepancy are several and the subject of considerable current debate (Ho & Larson 2006; Emerson 2007; Navascues & Emerson 2009; Ho & Lanfear 2010; Ho et al. 2011). While the phylogenetic-derived mutation rate may be underestimated due to recurrent mutations (Santos et al. 2005; Phillips et al. 2009), the pedigree-derived mutation rate may be overestimated due to transient mutations that are lost almost immediately (Irwin et al. 2009; Klutsch et al. 2011). Ho and colleagues have published a series of studies addressing this particular aspect (Ho et al. 2005, 2007, 2011; Ho & Larson 2006; Ho & Lanfear 2010), arguing that mutation rates are subject to temporal rate variation due to temporal variation in selection intensity (e.g. Loogvali et al. 2009). Other authors have utilized other slower-evolving regions in the mitochondrial genome as a means to conduct an ad hoc correction for recurrent substitutions in the mitochondrial control region (Alter & Palumbi 2009; Phillips et al. 2009).

Irrespective of the underlying cause, it is becoming increasingly clear that phylogenetic methods tend to underestimate μ by as much as one order of magni-
tude, which consequently implies a corresponding overestimation of recent historic abundance, all other factors being equal (Ho et al. 2008, 2011). The above-mentioned \textit{ad hoc} corrections yielded new, increased estimates of $\mu$ resulting in a corresponding lower estimate of recent historic abundance. For instance, Alter & Palumbi (2009) estimated $\mu$ for the mitochondrial control region at 2.6 times higher than previous estimates, implying a revised genetic diversity–based estimate of recent historic abundance of humpback whales in the North Atlantic from 240 000 to 90 000–100 000 individuals.

The use of temporally spaced samples (i.e. ancient and modern DNA samples) enables the estimation of the mutation rate over a more appropriate time frame (Drummond et al. 2002), but assumptions of current approaches (Fig. 2A.2) are very basic (e.g. ignoring gene flow), which may lead to incorrect inferences (Navascues & Emerson 2009).

Natural selection, which potentially affects the observed substitution rate as well as $N_e$, alters the rate of genetic drift relative to neutral expectations. If the degree of selection is similar across all lineages used in the estimation of $\mu$, then a phylogenetic estimation of $\mu$ should, in principle, be unbiased. However, if the timing and strength of selection is unequal among the lineages (e.g. species or populations) employed in the estimation of $\mu$ and $\theta$, then the resulting estimate of $\theta$ may be biased (Galtier & Boursot 2000; Torroni et al. 2001) unless the rate estimate is corrected for fluctuations in the timing and strength of selection (Tamura & Kumar 2002; Soares et al. 2009). Although genetic diversity–based estimates of recent historic abundance to date have assumed all loci to be selectively neutral, it is likely that many loci employed are subject to selection —both the presumed neutral mitochondrial control region (Casane et al. 1997; Howell et al. 2004, 2007; Haney et al. 2010) and nuclear loci, which often are introns (e.g. Palumbi & Baker 1994; Bouck & Vision 2007; Jackson et al. 2009; Li et al. 2010; Gante et al. 2011; Kemppainen et al. 2011).

**Issues in estimating demographic parameters**

Estimates of both generation time and the ratio of effective to census population size are needed to estimate recent historic abundance from contemporary genetic diversity. Generation time is needed to translate annual mutation rates into generational mutation rates, whereas the ratio of effective to census population size is needed to convert estimates of $N_e$ (derived from the estimates of $\theta$ and $\mu$) into estimates of abundance. Generation time is not a fixed entity but can change over time and space due to the variation in demographic and environmental processes, such as population density and habitat changes. These fluctuations, in turn, alter the age at first reproduction, overall reproductive output and mortality rates (Deevey 1947; Caswell 2001). Hence, a full-fledged estimation of generation time would require a complete Leslie matrix or life table with detailed demographic data on reproductive and mortality rates, which are unavailable in many non-model species. In addition, demographic rates vary over time and across the species’ range. For example, both age distribution and age-specific fecundity rates are unknown in most species of baleen whales. Some species appear to have very long life spans of more than one hundred years (Garde et al. 2007; Arrigoni et al. 2011; George et al. 2011). Even in some of the best studied species, such as humpback whales, with shorter life spans (such as 30–50 years), estimates of mean age at maturity differ by a factor of two (12 and 24 years, respectively) depending on the methods used to determine age (Chittleborough 1965; Barlow & Clapham 1997). The difference between these two ‘generation times’ results in a doubling of $\mu$, which in turn reduces the resulting abundance estimate by 50%.

In order to convert an estimate of $N_e$ to abundance (i.e. census population size or $N_c$), the ratio of the two population size measures must be known. Census population size is a relatively straightforward entity (but difficult to obtain in many cases), whereas the same is not the case for the effective population size which (in this context) is a population genetic entity. Hence, in order to infer $N_e$ from an estimate of $N_c$, a rigorous estimate of the ratio of current effective population size is warranted. Indeed, in the absence of temporally-spaced genetic samples, estimating $N_e$ requires values for reproductive variance among individuals, as age of reproduction, and age-specific reproductive and survival rates (Waples 2002).

**Is it feasible to infer recent historic abundance from current levels of genetic diversity?**

All estimation relies upon a model, in other words, a set of explicit (or implicit), often highly simplified assumptions. Such simplification consists of reducing the number of nuisance parameters (i.e. parameters that are not of interest per se but which must be accounted for), such as $g$, $\mu$ and $mN_e$. These simplifications are necessary to render a problem analytically (or computationally) tractable and to achieve a reasonably precise estimate. Ideally, a model is simplified by reducing the number of parameters to only those that are essential (Haldane 1964). Which parameters to retain depend upon the objective of the estimation and targeted level of precision (i.e. each nuisance parameter fixed at a single value results in an...
increased precision). Population genetic inference methods commonly used in ecology and conservation are often effective (compared to many other methods used in evolution, ecology and conservation) because they are based on highly simplified assumptions, but such simplifications can also become an Achilles Heel.

The population genetic inference methods that currently are employed to estimate recent historic abundance from current levels of genetic diversity were not developed with this specific objective in mind. Most of the parameter values necessary to infer recent historic abundance from current genetic diversity using eqns 1 and 2 are usually poorly known and fixed at a single value during the estimation (e.g. the mutation, migration and growth rates and generation time). The consequence is inference methods that yield estimates which can best be described as long-term and spatially broad weighted means. The very long time span often covered in the estimation of $\theta$ (and $\mu$) greatly increases the spatial scale of the estimation due to the effect of past and present gene flow upon local levels of genetic diversity, necessitating assumptions (and fixed nuisance parameter values) for a multitude of cellular and population processes far back in time (see Fig. 5).

These observations beg the key question: Does a genetic diversity–based estimate of $N$ represent the abundance during recent historic time? Although there is no single universal answer to this question, we believe the answer in the majority of cases is probably ‘no’. Exceptions are those few (if any) cases where the targeted population, those populations to which it is connected by migration, and the employed genetic markers behave as assumed by the estimation procedure and do so across the entire spatial and temporal scale covered by the estimation. This is rarely, if ever, the case for natural populations. Thus, an assessment of the extent of bias due to biologically plausible deviations from underlying assumptions should be undertaken. Indeed, some deviations (or combinations thereof) may have little effect on the final estimate of recent historic abundance, whereas other deviations may substantially bias the final estimate.

The reduction of nuisance parameters also implies that the uncertainty in the final abundance estimate likely is underestimated. Thus, most genetic diversity–based estimates of recent historic abundance appear overly precise.

Fig. 5 Illustration of the different temporal and spatial scales involved in a typical estimation of recent historic abundance of a local population. The depicted genealogy is estimated from the mitochondrial DNA sequence data used by Roman & Palumbi (2003) to estimate the recent historic abundance in North Atlantic humpback whales. The data were originally published by Baker et al. (1993) and Palsbøll et al. (1995). The mtDNA lineages depicted in grey were obtained from whales sampled in a different ocean basin than that which the specific clade was assigned to. These branches were inferred as the results of rare interoceanic gene flow (Baker et al. 1993; Palsbøll et al. 1995). The genealogy is the majority-rule consensus tree estimated from 1000 bootstrapped samples and an unrooted neighbour-joining tree with a transition/transversion ratio of 20, as implemented in the **Phyloree** 3.69 computer package (Felsenstein 1993). The abbreviation yr denotes years and Myr, million years.
In conclusion, only under the most restrictive of circumstances—namely when populations under study behave as assumed by models employed in the estimation—is a genetically derived estimate of recent historic abundance likely to be correct. While increasing the amount of data (loci and/or samples) will produce a more precise ‘abundance’ estimate, such an estimate is, in most cases, unlikely to represent local abundance at a specific recent time period as all the above-mentioned simplifying assumptions apply across all loci and samples.

**Look before leaping to conclusions**

The highly simplistic assumptions underlying most population genetic inference methods that are applied to infer recent historic local abundance are unlikely to apply to most natural populations, and hence, one may conclude that estimating recent historic abundance from contemporary genetic diversity is ill-advised. However, before dismissing such results altogether, it is worthwhile to conduct an assessment of biologically relevant deviations from the underlying assumptions and their effect upon the final estimate of recent historic abundance. If the effect of such deviations is small, then the final abundance estimate derived from an assessment of contemporary genetic diversity may be sound and reasonably accurate. Assessing the effects of different parameter values is relatively straightforward for some parameters (e.g. generation time and $\mu$) when there is a relatively simple relationship with $N_c$. Figure 6 is an example of such an assessment, which illustrates the large impact of changes in $\mu$ relative to generation time. Given the extensive body of population genetic simulation software available (e.g. Hudson 2002; Peng & Kimmel 2005; Excoffier & Foll 2011), it has become relatively straightforward to simulate most population histories and models under a variety of mutation models, thereby generating *in silico* genetic data from which to estimate the ‘historic’ abundance (reviewed in Hoban *et al.* 2012). Such simulations may not only serve to assess the effect of deviations from the assumptions underlying the estimation, but serve equally to plan a study by specifying the amount of data required to achieve a reasonably correct and precise estimate of historic abundance at the targeted tempo-spatial scale.

An estimate of recent historic abundance inferred from contemporary levels of genetic diversity may also provide the means for predicting a range of statistical characteristics for other genetic data, such as complementary STR data (e.g. genetic bottleneck tests or skyline plots). Additional tests will strengthen the conclusion or reveal inconsistencies. In undertaking additional testing, it is crucial to ensure that different tests or estimates are truly independent. This is often not the case for genetic analyses, which in many cases are based upon related data or related aspects of those data, and hence, care should be undertaken in the analysis and interpretation of such additional estimations (e.g. Alter *et al.* 2012).

Returning to our three introductory examples—large whales in the North Atlantic (2003), grey wolves in North America (Leonard *et al.* 2005) and New Zealand...
moas (Gemmell et al. 2004), we ask whether the estimates presented are representative of recent historic abundance and if so are they likely to be accurate? Our review of present knowledge and the analytical approaches used indicates numerous potential issues for all three cases. A full-fledged, well-supported assessment of the precision and the degree and direction of bias in the estimates of historic abundance for each of these cases is beyond the scope of this review, but some general comments are possible. All three studies were based upon a single locus (mitochondrial DNA) with the added uncertainties single locus estimates entail. The three studies differ in temporal and spatial scales aiming at estimating abundance at a specific place and time. The study in moas, by Gemmell et al. (2004), aimed at a broad temporal and spatial scale and is thus likely less susceptible to some of the temporal and spatial issues we have outlined above as the objective was a global long-term estimate. Leonard et al.’s (2005) study of grey wolves in North America was also inferred as representing the historic abundance at a ‘global scale’, and hence, spatial issues are likely less of a concern, although immigration to and from Asia (across the Bering Sea) and Greenland/Scandinavia via the Arctic might occur. However, the estimate was viewed as representing abundance prior to the anthropogenic extermination of grey wolves. For reasons explained above, this is unlikely to be the case. The third study, the abundance of whales in the North Atlantic Ocean at the onset of whaling (Roman & Palumbi 2003), presented the genetic diversity–based estimate of recent historic abundance as representing the local abundance for a specific population (the North Atlantic) and narrow time period (i.e. just prior to the onset of whaling). In this latter case, all the temporal and spatial issues we have outlined in this review apply to this study, in addition to the issues raised about estimates of demographic and mutation rates.

Nevertheless, each of the three studies represents one hypothesis for past abundance, and large discrepancies among different estimates (from genetic as well as non-genetic data) highlight the incompleteness of our current understanding of past abundances. Rigour would be added to the conclusions if estimates derived from different sources and methods converged. Any single estimate of past abundance based only on current genetic diversity should be viewed as tentative.

How then may we estimate recent historic abundance from genetic data?

Significant technical advances in the field of genomics now permit the production of massive amounts of genetic data from natural populations at an ever decreasing effort and cost. Accordingly, limitations in estimating recent historic abundance lie mainly in the analytical approach, in particularly the abovementioned discrepancy between the tempo-spatial scale of the objective and the estimation.

One possible solution is to resolve the discrepancy in the spatial and temporal scales between the objective and the estimation. This may be carried out either (i) by expanding the scope of the objective or inference or (ii) by narrowing the tempo-spatial scale of the estimation to match that of the objective. Expanding the scope of the estimation of abundance to the long-term global average (rather than local) abundance may be sufficient for some objectives (e.g. Gemmell et al. 2004; Leonard et al. 2005; Ruegg et al. 2010), but not others (Roman & Palumbi 2003). When the objective is a long-term and global estimate of abundance, we believe that current genetic approaches can yield reasonably accurate estimates provided that estimates of \( \mu \) and the necessary demographic parameters are sound.

The temporal scale of the inference may be narrowed by estimating the relative change in \( \theta \) for each intercoalescent interval in a population sample of orthologous gene copies which generates so-called Bayesian skyline/ride plots (Drummond et al. 2005; Drummond & Rambaut 2007; de Bruyn et al. 2011; Wu & Drummond 2011). This approach may be conducted with contemporary samples only, but benefits greatly by the inclusion of ‘ancient’ DNA (aDNA) obtained from carbon-dated samples (e.g. Shapiro et al. 2004; Campos et al. 2010a,b; Lorenzen et al. 2011). The aDNA samples add intraspecific calibration dates (to complement interspecific fossil calibration dates), enabling an estimate of the mutation rate based upon both inter- and intraspecific time calibration points (Ho et al. 2008). A key obstacle in terms of inferring recent historic abundance from Bayesian skyline plots is the low information content of the most recent coalescent intervals (Ho & Shapiro 2011). Accordingly, Bayesian skyline plots but will often fail to detect very recent (i.e. 5–15 generations ago) demographic changes (but see Hoffman et al. 2011) and the precision of the estimate of \( \theta \) (and consequently \( N_e \) is low (e.g. Hoffman et al. 2011 who estimated the 95% credible interval of pre-exploitation abundance at 0.8–2.3 million individuals). Hence, Bayesian skyline plots are better suited to detect older and prolonged changes in \( N_e \) (Fontaine et al. 2010; Herman & Searle 2011; Hope et al. 2011; Jezkova et al. 2011; Liu et al. 2011; Morgan et al. 2011; Mourier et al. 2012). This also appears to be the case with an equivalent approach developed for STR loci, MSVAr (Beaumont 1999; Bourke et al. 2010; Girod et al. 2011).

However, Bayesian skyline plots (and estimations using MSVAr) are based upon a single population model and thus ignore population structure (and changes in...
migration rates), which may bias results substantially (Chikhi et al. 2010). Translating changes in genetic diversity estimated from Bayesian skyline plots into census abundance requires the use of the same conversion factors, such as the generation time and the ratio of effective to census population size. Thus, while Bayesian skyline (using aDNA) may deal better with issues surrounding mutation rates, they are subject to many of the same issues affecting more standard genetic diversity-based estimators of abundance. In addition, taking advantage of aDNA samples requires historic samples from the relevant areas and the technical skills and a laboratory set-up to work with degraded DNA (Cooper & Poinar 2000).

Another manner by which to narrow the temporal and spatial scale is to employ samples from around the (historic) time point when an estimate of abundance is required. In such case, an estimate of \( N_e \) may be inferred from the degree of linkage disequilibrium (LD) in a random population sample (Hare et al. 2011; Waples et al. 2011). Individuals in small populations share more recent ancestors and therefore display increased levels of linkage across the genome, which may be employed to estimate the \( N_e \) of the parental generation (reviewed in Hare et al. 2011). In an ideal population, that is, with discrete generations, it is relatively straightforward to determine the point in time that such an estimate of \( N_e \) applies. The point in time is less certain for species with overlapping generations and poorly known demographic structure (e.g. Hare et al. 2011), but the window is restricted to a small number of generations. Most current implementations to estimate \( N_e \) from LD are based upon a single ideal population model and thus ignore migration, although some methods do attempt to adjust for movements among populations (Vitalis & Couvet 2001). Waples & England (2011) assessed the impact of violation of assumptions fluctuations on the final estimate \( N_e \) due to migration and structured samples, when estimators ignore these processes. Not surprisingly, both processes may have substantial effects upon the final estimate of \( N_e \). For instance, generational migration rates above 5–10% (which translates into very low annual migration rates in long-lived species) strongly influence the estimate of \( N_e \) (Waples & England 2011). The authors suggest that immigrants may be identified and removed by other means, such as assignment methods (Paetkau et al. 1995; Piry et al. 2004), although the rigour of individual multilocus assignments decreases with reduced divergence (Waples & Gaggiotti 2006; Hall et al. 2009), and consequently, this kind of adjustment may not be infeasible at such high rates of migration (i.e. above 5–10% per generation). The advantage of \( N_e \) estimates from LD is the very narrow time scale of the estimation (i.e. the last few generations depending upon the approach and data used), which greatly reduces the uncertainty and possible influence from other processes (e.g. migration and mutation, and variations in those and demographic parameters) that may bias the final abundance estimate. However, in most cases, the ‘historic’ abundance of interest will be farther back in time than one or two generations, which then implies the need for acquiring ‘historic’ samples along with all the technical difficulties associated with collecting genetic data from such samples. Last, but not least, the estimation of \( N_e \) from LD requires substantial sample sizes in most cases but especially for large undepleted populations (Waples & Do 2008), which may altogether prevent estimating \( N_e \) in this manner from a few rare historical samples.

**Novel opportunities in the age of genomics and ABC?**

Perhaps the most exciting prospects for estimating recent historic abundance arise from novel opportunities offered by massive parallel sequencing technologies as well as new analytical approaches enabling the assessment of more complex models. These new technologies provide the opportunity to increase the number of loci by several orders of magnitude. As we have pointed out above, adding more data will increase the precision of the final abundance estimate, but for many inference methods the estimate will represent a long-term mean of abundance and not the local abundance at a specific recent time period. However, the greatly increased genome-wide density of data adds new possibilities in terms of estimating the degree of linkage among markers along chromosomes provided that a (partial) genome backbone is available. Such data enable the estimation of the lengths and distribution of linkage tracts in each sample. The smaller the population, the longer the linkage tracts [also known as identity-by-descent (IBD) tracts], because most individuals share a very recent common ancestor, resulting in low recombination rates. The distribution of the lengths of IBD tracts in an individual’s genome reflects past and present population processes, such as selection, migration and effective population size. Recently, Pool & Nielsen (2009) showed how changes in migration rates affect the distribution of lengths of ‘migrant’ tracts. The same should be the case for changes in population size, which will be reflected in the genome-wide distribution of IBD tract lengths (Chapman & Thompson 2003). The degree of LD will be highest for small \( N_e \)'s, which implies longer IBD tracts and greater precision at low \( N_e \)'s, whereas the opposite will be the case for large \( N_e \)'s (Chapman & Thompson 2003). Accordingly, the utility of the genome IBD trac-
lengths approach will depend, in part, upon the size of the historic population with low precision at large historic population sizes (Pool et al. 2010).

Although these new technological advances offer great potential, the speed and computational resources required to make inferences under sufficiently realistic population and mutation models is an associated and purely practical issue. Even the current (comparatively low coverage) multilocus DNA sequence data set (e.g. Alter et al. 2007; Ruegg et al. 2010) poses a computational challenge for likelihood-based estimation methods. The adoption of massive parallel sequencing technologies (Davey et al. 2011) will lead to much more data per specimen. Increasing the amount of data usually increases the precision and rigour of the estimation. However, likelihood MCMC approaches commonly employed to estimate $\theta$ from population genetic data can quickly become computationally infeasible as the amount of data increases (Wegmann et al. 2009). Sequential Markov coalescent-based (McVean & Cardin 2005; Excoffier & Foll 2011) or ABC (Beaumont et al. 2002; Hamilton et al. 2005; Wegmann et al. 2010) approaches appear to constitute a promising alternative to the likelihood approaches (Lopes & Boessenkool 2010). The computational efficiency afforded by these non-likelihood approaches will make it feasible to utilize more of the information contained in the large amounts of data generated by these parallel sequencing methodologies, such as IBD tracts and the allele frequency spectrum which may be used to estimate demographic changes (Nielsen 2000; Polanski & Kimmel 2003; Caicedo et al. 2007; Gusev et al. 2009; Gutenkunst et al. 2009; Crisci et al. 2012). However, even though non-likelihood-based approaches provide additional flexibility (and thus are able to accommodate more complex models), such estimations need careful consideration and a realistic number of nuisance parameters (Bertorelle et al. 2010).

Acknowledgements

PJP was in part supported by a Hrdy Fellowship in Conservation Biology from Harvard University during the writing of the manuscript. SRB was supported by NSF (DEB-1051342 and CNH 1115069). We are indebted to Drs. Oscar Gaggiotti, Allen-dorf, Luis Pastene and Finn Larsen as well as two anonymous reviewers for helpful comments which greatly improved the manuscript.

References


optera acutorostrata from Greenland, the North East Atlantic and the North Sea probably reflects different ecological regions. Marine Ecology Progress Series, 247, 263–280.


Baker CS, Perry A, Bannister JL et al. (1993) Abundant mito-


Beerli P (2010) Violation of assumptions, or are your migration estimates wrong when the populations split in the recent past? http://popgen.fs.fsu.edu/Migrate/ Blog/Entries/2010/8/15_Violation_of_assumptions_or_are_your_migration_estimates_wrong_when_the_populations_split_in_the_recent_past.html, Florida State University.


© 2012 Blackwell Publishing Ltd


The authors share an interest in the application of population genetic data and inference methods to conservation in addition to good food, wine and the Cal Bears.