

Are Antarctic minke whales unusually abundant because of 20th century whaling?

KRISTEN C. RUEGG,* ERIC C. ANDERSON,†‡ C. SCOTT BAKER,§¶ JENNIFER A. JACKSON§ and STEPHEN R. PALUMBI*

*Department of Biology, Hopkins Marine Station, Stanford University, 120 Oceanview Boulevard, Pacific Grove, CA 93950, USA, †Southwest Fisheries Science Center, National Marine Fisheries Service, 110 Shaffer Road, Santa Cruz, CA 95060, USA, ‡Department of Applied Math and Statistics, University of California, Santa Cruz, CA 95060, USA, §Marine Mammal Institute, Hatfield Marine Science Center, Oregon State University, 2030 SE Marine Science Drive, Newport, OR 97365, USA, ¶School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand

Abstract

Severe declines in megafauna worldwide illuminate the role of top predators in ecosystem structure. In the Antarctic, the Krill Surplus Hypothesis posits that the killing of more than 2 million large whales led to competitive release for smaller krill-eating species like the Antarctic minke whale. If true, the current size of the Antarctic minke whale population may be unusually high as an indirect result of whaling. Here, we estimate the long-term population size of the Antarctic minke whale prior to whaling by sequencing 11 nuclear genetic markers from 52 modern samples purchased in Japanese meat markets. We use coalescent simulations to explore the potential influence of population substructure and find that even though our samples are drawn from a limited geographic area, our estimate reflects ocean-wide genetic diversity. Using Bayesian estimates of the mutation rate and coalescent-based analyses of genetic diversity across loci, we calculate the long-term population size of the Antarctic minke whale to be 670 000 individuals (95% confidence interval: 374 000–1 150 000). Our estimate of long-term abundance is similar to, or greater than, contemporary abundance estimates, suggesting that managing Antarctic ecosystems under the assumption that Antarctic minke whales are unusually abundant is not warranted.

Keywords: Antarctic marine ecosystem, Antarctic minke whale, coalescent modelling, competitive release, effective population size, krill surplus hypothesis

Received 15 July 2009; revision received 21 October 2009; accepted 29 October 2009

Introduction

Ecologists have long debated the relative roles of top-down (consumer-driven) and bottom-up (resource-driven) forces in shaping natural communities (Power 1992; Frank *et al.* 2007). Trophic cascades stemming from removal of top predators provide compelling support for consumer-driven control of food webs across a diversity of ecosystems (Pace *et al.* 1999). While most examples of cascades are from small-scale, simple food webs, recent studies suggest that cascades may be

occurring in larger, more complex marine ecosystems (Estes *et al.* 1998; Frank *et al.* 2005). Teasing apart the effects of top predator removal requires knowledge of an ecosystem before and after their extirpation – a challenging situation in today's oceans, where over-exploitation has eliminated much of the once-abundant megafauna (Myers & Worm 2003).

The commercial hunting of approximately 2 million whales (Fig. 1a, b) from the Southern Ocean in the early 1900s (Clapham & Baker 2002) provides an opportunity to investigate the ecological consequences of top predator removal on an oceanic scale. The hunted whales would have consumed as much as 150 million tonnes of krill annually (Laws 1977), leading some

Correspondence: Kristen C. Ruegg, Fax: 831-655-6215; E-mail: kruegg@stanford.edu

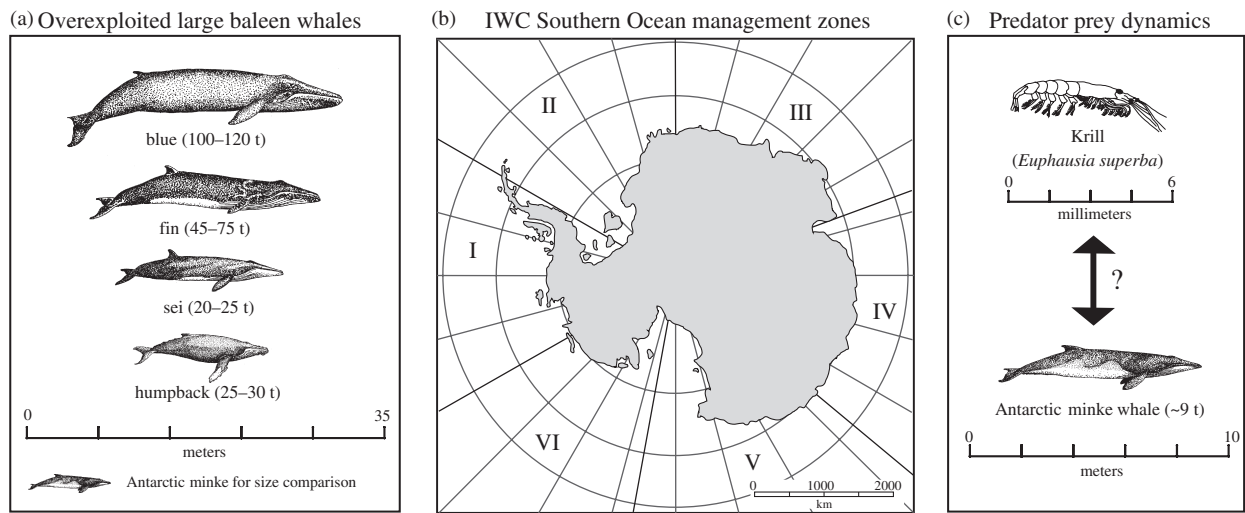


Fig. 1 The Southern Ocean ecosystem: (a) four main species of baleen whale hunted in the early 1900s. The figure illustrates that a blue whale in the Antarctic measuring up to 29 m and weighing between 100 and 120 metric tonnes (t) is approximately three times longer and more than 11 times heavier than an Antarctic minke whale, measuring up to 9 m and weighing up to 9 t. (b) Southern Ocean management zones as defined by the IWC. (c) Krill, the main food source of baleen whales. The question mark refers to the unknown relationship between krill abundance and minke whale abundance in the Southern Ocean.

authors to suggest that their removal led to competitive release for smaller krill-eating organisms (reviewed in Ballance *et al.* 2006) – a top-down hypothesis often referred to as the ‘Krill Surplus Hypothesis’. However, krill populations in the Southern Ocean are currently thought to be regulated by recruitment, which is positively correlated with sea ice cover (Loeb *et al.* 1997; Nicol *et al.* 2000) – a bottom-up explanation that has profound implications for krill and krill-dependent species amidst rising global temperatures (Croxall *et al.* 2002; Nicol *et al.* 2008). Both hypotheses may be correct; larger baleen whales may have regulated krill abundance until they were removed and replaced by bottom-up regulatory forces (Ballance *et al.* 2006). However, the absence of a prewhaling baseline makes it difficult to distinguish between the long-term influences of top-down vs. bottom-up forces in the Antarctic Marine Ecosystem.

One species hypothesized to have benefited from a krill surplus is the Antarctic minke whale, *Balaenoptera bonaerensis* (Burmeister 1867) (Fig. 1c). While there is no direct evidence to support the Krill Surplus Hypothesis for Antarctic minke whales, population size increases during the latter half of the 20th century have been inferred from hypothesized decreases in the age at sexual maturity (Thomson & Butterworth 1999), and modelled increases in minke whale recruitment (Butterworth *et al.* 1999). Some suggest that minke whales in the Southern Ocean increased in population size approximately 8-fold after the removal of the large

baleen whales (blue, fin, sei, humpback; Fig. 1a) (Ohsumi 1979; ICR 2006), but this assessment has since been challenged by more recent Antarctic ecosystem models suggesting a 3-fold increase (Mori & Butterworth 2006). Regardless of the magnitude of increase, it has been suggested that, at their present level, Antarctic minke whales may be inhibiting the recovery of other overexploited whale species and reducing human food resources through competition (Ohsumi 1979; Morishita & Goodman 2001; ICR 2006). While there is a lack of firm data on prewhaling population sizes, and the extent to which competition regulates whale populations (Gales *et al.* 2005), some agencies advocate culling minke whales as a way to reduce competition with fisheries and to support the recovery of other overexploited whale species (ICR 2006). An estimate of the average size of the Antarctic minke whale population prior to human disturbance may shed light on the extent to which the Krill Surplus Hypothesis is necessary to explain present abundance.

Recently, scientists have employed genetic data to assay past population sizes of baleen whales and other species (Roman & Palumbi 2003; Shapiro *et al.* 2004; Alter *et al.* 2007; Atkinson *et al.* 2008), based upon the relationship between genetic diversity (θ) and long-term effective population size (N_e) ($\theta = 4N_e\mu$ where μ is the average mutation rate). Initial reconstructions of the long-term population size of whales using genetic data were limited by reliance on a single locus and incomplete oceanic sampling (Lubick 2003; Roman & Palumbi

2003; Holt & Mitchell 2004). Following these initial reconstructions of long-term population size, several authors proposed improvements (Clapham *et al.* 2005) including: (i) using multiple unlinked nuclear loci, (ii) testing for deviations from mutation-drift equilibrium, (iii) estimating species-specific mutation rates, (iv) estimating overall variance in abundance, and (v) testing the effect of un-sampled populations. More recent efforts to reconstruct the long-term population size of grey whales have helped to overcome many of these initial limitations (Alter & Palumbi 2007; Alter *et al.* 2007), however, additional challenges arise when using high-diversity nuclear sequence data.

A significant hurdle in working with high-diversity nuclear sequence data is incorporating the uncertainty that results from unresolved gametic phase. Gametic phase is defined as the original combination of alleles that an individual received from each of its parents. Resolving gametic phase becomes more difficult as sequence diversity increases. Separating the constituent alleles in heterozygous individuals has been traditionally performed through the use of laboratory methods (cloning, single-strand conformation polymorphism) (reviewed in Zhang & Hewitt 2003). However, as diploid sequencing becomes easier and the sheer amount of nuclear sequence data increases, traditional laboratory methods for resolving phase become increasingly time-consuming and costly. Computational methods, such as the Bayesian software PHASE (Stephens *et al.* 2001), that infer haplotypes from sequence data provide an alternative to laboratory techniques and have been shown to produce accurate results (Harrigan *et al.* 2008). However, despite their accuracy, computational methods cannot always resolve haplotypes with a high degree of certainty. A method for incorporating this uncertainty in final estimates of population parameters, such as effective size, is needed.

To assess the likelihood of a post-whaling competitive release in the Antarctic minke whale, we estimate their long-term population size through a coalescent-based analysis of genetic diversity across eleven unlinked nuclear markers. We build upon the methodological improvements described in Alter *et al.* (2007) by describing a general method for capturing uncertainty in estimates of effective size due to unresolved gametic phase in highly heterozygous sequences and by developing a method for incorporating mutation rate variation among loci in coalescent analyses of effective size. We also use coalescent simulations to determine the potential effect of population substructure and limited geographic sampling on our estimate of whole-ocean Antarctic minke whale genetic diversity.

Materials and methods

Sample collection and sequencing

Fifty-two whale meat samples were purchased from Japanese meat markets and copies of their genomes were amplified using Whole Genome Amplification (Lasken & Egholm 2003). The sampled individuals were originally killed in Antarctic Management Areas III, IV, V and VIW (Fig. 1b) by the Japanese Whale Research Program under the Special Permit in the Antarctic (JARPA). Eleven nuclear loci (Table 1; Table S1) were amplified from these samples and sequenced using standard PCR and sequencing protocols (sequences identical to those in Jackson *et al.* (2009) are catalogued under GenBank accession nos GQ407272–GQ408882; new sequences can be found under GenBank accession nos GU144923–GU145028). Individuals were sequenced in both directions when possible and all variable sites were checked by eye using Sequencher ver. 4.8 (Gene Codes Corporation). Despite multiple attempts, not all individuals sequenced successfully for every locus, resulting in variation in the final sample sizes (mean 40, range: 20–52). PHASE 2.1 (Stephens *et al.* 2001) was used to reconstruct gametic phase, defined as the original allele combination that an individual received from each of its parents, using a burn-in of 10 000 iterations and run length of 10 000 iterations. To ensure that each sample was unique, we confirmed that none of the samples had identical

Table 1 Summary statistics for 11 loci sequenced in Southern Ocean minke whales

Intron	N	N_S	N_H	π	Tajima's D	Fu's F_s
ACTA	34	9	7	0.00255	-1.09	-1.27
BTN	32	6	6	0.00120	0.50	0.54
CAT	38	13	13	0.00685	0.78	-1.81
CHRNA	49	16	16	0.01352	0.02	-3.95
CP	20	24	19	0.00717	-0.04	-0.92
ESD	47	23	25	0.00572	0.01	-16.17*
FGG	41	13	12	0.00084	-0.01	-0.35
GBA	44	4	5	0.00061	-1.58	-4.66
LAC	45	12	12	0.00198	-0.06	-1.22
PTH	42	4	4	0.00144	-1.06	-1.11
RHO	52	3	4	0.00074	-1.37	-3.59

*Number in bold refers to a significant deviation from zero as determined by 95% confidence intervals generated using coalescent simulations in DNAsp (Rozas *et al.* 2003).

N = number of individuals, N_S = number of polymorphic sites, N_H = number of distinct haplotypes as determined by PHASE, ver. 2.1 (Stephens *et al.* 2001), π = nucleotide diversity (Tamura & Nei 1993).

sequences at all loci. Using Arlequin ver. 3.0 (Excoffier *et al.* 2005) we found no significant linkage disequilibrium among loci after correcting for multiple comparisons; therefore we considered the loci independent.

Testing for equilibrium, neutrality and substructure

To determine, if our sequences were evolving in a manner consistent with equilibrium and neutrality, Tajima's *D* (Tajima 1989) and Fu's *F_s* (Fu 1997) tests were performed using DnaSP (Rozas *et al.* 2003). We also used DnaSP to calculate the minimum number of recombination events in the sample (Hudson & Kaplan 1985) and found that 6 of the 11 introns showed evidence of recombination. Therefore, coalescent simulations ($n = 10\,000$) incorporating the per gene recombination parameter (*R*) were used to generate 95% confidence intervals (CI) for both Tajima's *D* and Fu's *F_s* statistics.

Population subdivision can increase coalescence time between genes and inflate estimates of genetic diversity. While preliminary reports have suggested that there may be population substructure within Antarctic minke whales ($\phi_{st} = 0.0090$, $P = 0.0025$) (Pastene *et al.* 1996), these distinctions are weak. To further investigate the possibility of population substructure, we estimated the most likely number of populations (*K*) within our data set using the program Structure ver. 2.2 (Pritchard *et al.* 2000). To avoid the potentially confounding effect of background linkage disequilibrium between closely linked sites within a locus, the maximum *a posteriori* haplotypes from PHASE at each locus were recoded as alleles at each locus. We performed three independent runs at each *K* value ($K = 1, 2$, and 3) using a burn-in period of 100 000 iterations and a run length of 500 000. The best *K* value was selected as the one giving the highest average log probability of the data [$\ln P(X|K)$].

Estimating θ , accounting for interlocus variation in mutation rate and uncertain gametic phase

LAMARC ver. 2.1.3 was used to simultaneously estimate θ , while incorporating recombination in the model (Kuhner 2006). In contrast to summary statistic estimates of θ (θ_s , θ_π , etc.), LAMARC accounts for uncertainty in the data by integrating over the space of possible genealogies using Markov chain Monte Carlo simulations. Initial runs with PHASE indicated that 8 of the 11 loci contained some sites, where the gametic phase could not be resolved with high confidence (probability threshold <90%). While LAMARC has an option for entering data as 'phase unknown,' initial tests indicated that inputting samples from PHASE's posterior distribution produced tighter convergence across runs. Therefore, to account for the uncertainty in

the data resulting from unknown gametic phase, LAMARC was run on 10 realizations from PHASE's posterior distribution for each of the 11 loci.

To accommodate interlocus variation in mutation rate, we implemented LAMARC's gamma model for mutation rate variation. The gamma model option models individual locus mutation rates as being drawn from a gamma distribution with mean one and a shape parameter estimated from the data. Comparing the distribution of variation in estimates of the individual locus substitution rates with the gamma distribution estimated by LAMARC confirmed that the gamma model provides a good fit to the data (see Fig. S1). For 9 of the 11 loci we applied the best fitting mutation models according to the phylogenetic analysis by Jackson *et al.* (2009); for the remaining two loci, we used the mutation models inferred by Alter *et al.* (2007) (Table S1). For recombination rate, we used a flat prior on a log scale from 1E-05 to 10. For θ , we used a flat prior from 1E-05 to 0.4, achieving nearly identical results on both the log and the linear scale.

We used LAMARC's Bayesian option and achieved excellent mixing and concordance between replicate runs. Because the current implementation of LAMARC allows the gamma model only within the likelihood framework, we developed our own extension of LAMARC (called GUFBUL-Gamma Updating For Bayesians Using LAMARC) that allows the gamma model to be applied in a Bayesian framework (see Supporting Information). For the final analysis, each PHASE realization was run in LAMARC three times with different random number seeds using 150 000 iterations of burn-in and 600 000 iterations after burn-in, taking samples every 20 iterations. The final θ values were obtained by allowing LAMARC to combine information across the 10 alternate PHASE realizations and the three runs as 30 separate replicate runs for a total of 18 million Markov chain Monte Carlo iterations after burn in.

Simulating effects of substructure and limited geographic sampling on θ

Our samples were originally collected from a restricted geographic region and the extent to which this sampling bias influenced our estimate of θ depends upon the degree of population structure. To investigate the relationship between population substructure, limited geographic sampling, and θ , we simulated seven populations in an Antarctic ring, joined by stepping stone migration with migration rates ranging between 1.25 and 50 migrants per population per generation. Using the program makesamples (Hudson 2002), we set $\theta = 4N_e$ to 3.75; this corresponds to a per nucleotide $\theta = 0.0071$ in a sequence of average length for our

study. We simulated two scenarios: (i) single sub-population sampling: 84 sequences sampled from a single subpopulation, and (ii) multi-population sampling: 12 sequences sampled from each of the seven populations. We simulated 50 000 coalescent trees with the infinite sites mutation model at each migration rate and under each of the two sampling scenarios. From each replicate, we estimated θ using the number of segregating sites, θ_s and θ based upon pairwise differences θ_π .

Calculating census population size from θ

The conversion of θ into effective population size (N_e) is based upon the relationship $\theta = 4N_e\mu$ where μ is the average mutation rate. To calculate an average μ for Antarctic minke whales, and to estimate uncertainty surrounding our estimate, we sampled with replacement from among 11 previously published individual locus mutation rates for the 11 loci in our study; two of the individual locus mutation rates were from Alter *et al.* (2007), while nine were taken from a Bayesian analysis of baleen whale phylogeny and fossil history (Jackson *et al.* 2009). For each re-sampled locus, a sample mutation rate was drawn from the posterior distribution of the estimated mutation rate or uniformly from the 95% CI on the mutation rate.

To convert μ from units of mutations per base pair per year into mutations per base pair per generation, we estimated a generation time for Antarctic minke whales. The average age of sexually mature individuals can be used as a proxy for generation length, assuming fecundity is constant with age (Roman & Palumbi 2003). Using 7 years as the age at sexual maturity (Klinowska 1991), we calculated the average age of sexually mature individuals (from 7 to 53 years old) using commercial and JARPA catch records from 43 236 individuals as reported in Table 1 of Butterworth *et al.* (1999). There was considerable variation in generation length across years, sample areas, and sample methods (commercial and JARPA catches from area IV and area V) (Fig. S2); to more accurately reflect uncertainty in our estimate we sampled uniformly from between the lower and upper bounds of year-to-year and area-to-area estimates (from 14.60 to 21 years).

To convert N_e to census population size (N_c) requires knowledge of the ratio of mature adults to the effective number of adults (N_{mature}/N_e) and the proportion of juveniles in the population. We based our estimate of N_{mature}/N_e on eqn 1 in Nunney & Elam (1994): $N_e = N/(2-T^{-1})$, where T = generation length. To approximate juvenile abundance, we estimated the ratio of total population size to total adults based upon age structure information as reported in Table 2 of Kato *et al.* (1990, 1991). Again, we sampled with replacement

Table 2 Posterior mean theta (θ) estimated using LAMARC

Marker	θ	Min	Max
ACTA	0.0061	0.0023	0.0119
BTN	0.0013	0.0005	0.0026
CAT	0.0105	0.0051	0.0187
CHRNA	0.0201	0.0102	0.0350
CP	0.0081	0.0043	0.0145
ESD	0.0145	0.0081	0.0232
FGG	0.0044	0.0021	0.0076
GBA	0.0042	0.0013	0.0093
LAC	0.0065	0.0028	0.0120
PTH	0.0043	0.0012	0.0102
RHO	0.0051	0.0014	0.0118
OVERALL	0.0071	0.0045	0.0112

1 million times to generate CI around our estimate of the ratio of total population size to total adults.

Results

Tests for equilibrium, neutrality and substructure

Among eleven nuclear loci, nucleotide diversity averaged 0.00387 (range: 0.00074–0.01352), with an average of 11.2 haplotypes per locus (range: 4–25) (Table 1). These values are higher than for other baleen whales: for example, nucleotide diversity is nearly four times higher than that seen in grey whales (Alter *et al.* 2007), reflecting higher heterozygosity for all six loci for which direct comparisons can be made. The Structure analysis did not reveal any hidden population structure in our data, suggesting the most likely number of populations (K) was $K = 1$ ($\ln P(X|K) = -1427$) with $K = 2$ ($\ln P(X|K) = -1431$) and $K = 3$ ($\ln P(X|K) = -1437$) being less likely.

The results of Tajima's D and Fu's F_s tests were consistent with neutrality and equilibrium (Table 1): no Tajima's D values and only one of eleven Fu's F_s tests was significantly different from zero (Table 1). These results differ from those of Pastene *et al.* (2007) and Alter & Palumbi (2009) who found evidence of departure from equilibrium conditions based upon mtDNA.

Estimate of θ , while accounting for uncertainty in gametic phase and interlocus variation in mutation rate

From locus to locus, estimates of θ varied from 0.0010 to 0.0174 (Table 2; Fig. S3), presumably reflecting variation among loci in mutation rate and coalescent history. Our method for incorporating uncertainty in gametic phase by taking ten evenly spaced realizations from PHASE's posterior distribution and running them as

separate samples in LAMARC resulted in tight convergence across runs despite differences in phasing (Fig. S3). Furthermore, our Bayesian framework for implementing the gamma model in LAMARC was considerably more efficient than the likelihood framework, reducing computation time from several weeks to several days. Incorporating information across all 11 loci and alternative phases, we estimated the posterior mean θ to be 0.0071 (95% CI: 0.0045–0.0112) (Table 2).

Effects of substructure and limited geographic sampling on θ

Simulations designed to investigate the relationship between limited geographic sampling and potential population substructure within Antarctic minke whales indicated that population structure would not significantly increase θ unless migration between sub-populations was so low that the expected ϕ_{st} would be >0.10 (an order of magnitude higher than the previously calculated ϕ_{st} for Antarctic minke whales) or $N_m < 2.5$ (Fig. 2). Furthermore, our simulated θ differed little regardless of whether samples were drawn from a single subpopulation or drawn evenly from across all sub-populations, indicating that even though our samples are from a limited geographic area, our θ estimate reflects ocean-wide genetic diversity.

Estimate of census population size from θ

Based upon individual locus mutation rates from Alter *et al.* (2007) and a Bayesian analysis of baleen whale phylogeny and fossil history (Jackson *et al.* 2009), the average mutation rate was estimated to be 4.54×10^{-10} /bp/year (95% CI: 3.50×10^{-10} to 5.75×10^{-10}). Our estimate was reduced from the slightly higher

average rate used by Alter *et al.* (2007) of 4.8×10^{-10} /bp/year. Using the average age of sexually mature individuals as a proxy for generation length, the average generation length was estimated to be 17.65 years. Sampling uniformly from within lower and upper bounds of year-to-year and area-to-area estimates for generation length resulted in an age range between 14.60 and 21 years. Using θ estimated from LAMARC and our multi-locus mutation rate in mutations per base pair per generation, we calculated the effective size of the Antarctic minke whale population to be 199 849 (95% CI: 140 519–349 736).

To convert from effective population size into census population size, we incorporated juvenile abundance and variation in reproductive success. We estimated juvenile abundance or the ratio of total population size to total adults to be 100:67 or 1.48 (95% CI: 1.39–1.59). We approximated variation in reproductive success or the ratio N_{mature}/N_e to be ~ 2 based upon eqn 1 in Nunney & Elam (1994). Multiplying the product of the two above ratios by our estimate of effective population size gives an estimate of census population size of 671 000 individuals. Bootstrap re-sampling across the variation in mutation rate, generation time, the ratio of total population size to total adults and from the posterior distribution of effective size to estimate variation in abundance yielded 95% CI from 374 000 to 1 150 000 (Fig. 3).

Discussion

Comparison between survey-based and genetic estimates of population size

Here, we show that Antarctic minke whales are not unusually abundant as a result of 20th century whaling.

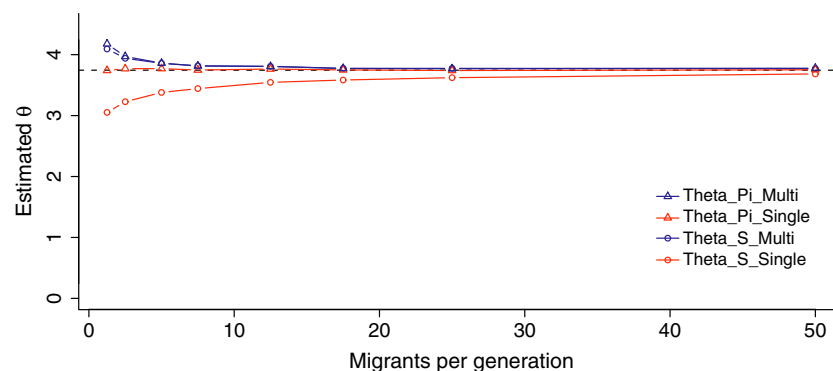


Fig. 2 Population sub-structure simulation results demonstrating that θ estimated using segregating sites and pairwise differences differed little regardless of whether samples were drawn from a single sub-population or drawn evenly from across seven sub-populations. Simulations are described in the supplemental information text. The horizontal dotted black line at 3.75 shows the expected value of θ under panmixia (same number of individuals, but no population substructure). Blue lines show results for the multipopulation sampling scenario; red the single population sampling scenario. The results show that across a wide range of migration rates, estimates of θ are close to the value expected under panmixia.

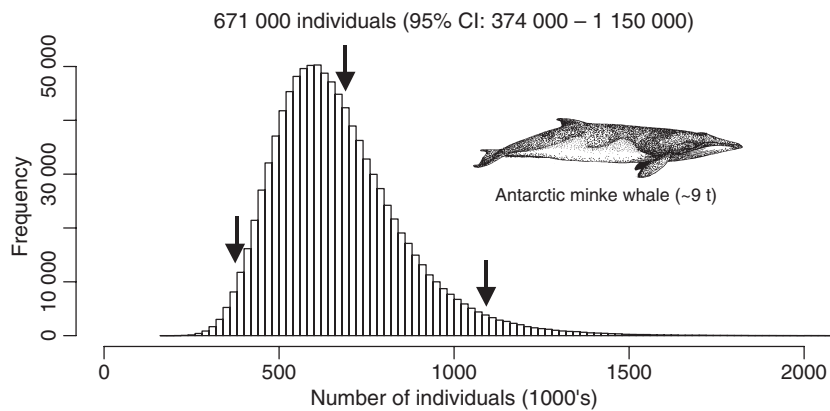


Fig. 3 Distribution of long-term census population size estimates, taking account of uncertainty in θ , mutation rate, generation time, and the ratio of total population size to total adults. The arrows represent the mean value and the upper and lower 95% confidence intervals.

Our genetic estimate of long-term population size for the Antarctic minke whale is 671 000 individuals (95% CI: 374 000–1 150 000) and calculations based upon coalescent theory confirm that our estimate predates any purported population expansion facilitated by removal of the large baleen whales. Rather than being unusually large, our long-term genetic estimate spans the range of estimates from three circumpolar surveys conducted under the supervision of the IWC: 608 000 (CV = 0.089) for the years 1978–1984; 766 000 (CV = 0.091) from 1985 to 1991 (Branch & Butterworth 2001); and an unpublished report to the IWC that suggested 338 000 (CV = 0.079) for the years 1991–2004 (Branch 2006). Differences between various survey-based abundance estimates remain controversial (Branch & Butterworth 2001; Clapham *et al.* 2007) and a consensus regarding present abundance is pending. Some survey-based estimates are considered minimum estimates because they do not include whales missed on the track-line, whales north of the survey region, and whales inside the pack ice (Branch & Butterworth 2001). However, as our long-term evolutionary estimate of abundance encompasses the range of most contemporary estimates of abundance we conclude that competitive release as predicted by the Krill Surplus Hypothesis is not required to explain current Antarctic minke whale abundance.

Accounting for uncertainty in gametic phase

We implement a method for capturing uncertainty in estimates of effective size due to unresolved gametic phase that may be generally useful to researchers working with high-diversity nuclear sequence data. To account for uncertainty due to unknown phase, we sampled from across the range of possible allele combinations generated by PHASE and then used these sample allele combinations to estimate θ in LAMARC. Our results showed strong convergence among different simulations of the genealogy, even when heterozygosity

was high and phase was uncertain (Fig. S3). In addition, a comparison between the method that we employ and LAMARC's method for inferring phase (Kuhner 2006) indicated that our method provided tighter convergence across runs (data not shown). Our results suggest that using samples from PHASE's posterior distribution can produce reliable conclusions and highlights a general method for incorporating uncertainty due to phase in coalescent analyses of population parameters in addition to the method already provided by LAMARC. Differences between the two methods' ability to accurately predict phase would make an interesting topic for future research.

Accounting for uncertainty in estimates of long-term population size

There are a number of events that may decouple the relationship between genetic diversity and long-term population size, including deviations from neutrality, past hybridization and population sub-structure (Clapham *et al.* 2005; Alter *et al.* 2007). Population size changes and/or selection may increase or decrease diversity relative to neutral expectation and, as a result, such events may artificially increase or decrease estimates of long-term population size from genetic data. We have attempted to avoid the complicating effects of departures from neutrality by sequencing 11 loci that were found to be consistent with neutrality according to Tajima's D and Fu's F_s tests (Tajima 1989; Fu 1997; see Supporting information for further discussion). However, we acknowledge that the very idea of neutral molecular evolution in mammals has been recently brought into question (Chamary *et al.* 2006) and the role that this potential paradigm shift may play in interpreting our results provides an interesting area for future study.

Major, past hybridization events may increase diversity and inflate estimates of long-term population size.

Recent genetic evidence found that Antarctic minke whales are reciprocally monophyletic with their closest living relative, the common minke whale, at the rapidly evolving mtDNA control region (Pastene *et al.* 2007; see Supporting information for further discussion), suggesting that a major, recent hybridization event is unlikely. However, additional sequencing of common minke whale samples at the same slowly evolving nuclear loci as those sequenced in this study would further test the extent to which hybridization could have influenced our estimate of genetic diversity.

Undetected population structure can also increase estimates of diversity and inflate estimates of long-term population size from genetic data (Alter *et al.* 2007; Atkinson *et al.* 2008). In the current case, tests for population structure within the Antarctic minke whale did not reveal any hidden population substructure. Furthermore, our simulations suggest that population structure would not significantly increase diversity (θ) unless migration between subpopulations was so low that the expected $\phi_{st} > 0.10$ (Fig. 2; see Supporting information for further discussion). The amount of structure needed to influence our estimate of θ is an order of magnitude greater than the amount of structure detected by previous estimates of population structure within the Antarctic minke whale (Pastene *et al.* 1996). Therefore, our results and those of previous studies suggest that population structure is unlikely to have significantly influenced our estimate of long-term population size.

In addition to factors that may decouple the relationship between genetic diversity and long-term population size, there are a number of uncertainties surrounding the calculation of long-term population size that cannot be captured by our CI. These include general problems in the field of ecology and evolutionary biology such as attaining accurate estimates of mutation rates (Ho *et al.* 2005; Emerson 2007) and the ratio of N_{mature}/N_e (Nunney 1991, 1993; Nunney & Elam 1994; Frankham 1995). In both of these cases, we have chosen values that most closely reflect the current state of understanding in the field, while acknowledging the role that these uncertainties play in our final estimate of long-term population size (see Supporting information for further discussion).

Implications for the Krill Surplus Hypothesis

While our results suggest that competitive release is not necessary to explain current abundance in Antarctic minke whales, they do not allow us to reject unequivocally some level of increase as a result of a krill surplus. This is due to the fact that our estimate of diversity is affected by the harmonic mean of population size across time, and therefore cannot be attributed to any particular

point in history. It is possible that Antarctic minke whale abundance was abnormally low just prior to whaling and that a krill surplus returned them closer to their long-term average, though there are no data to suggest that this was the case. For Antarctic minke whales to have increased by ~ 8 -fold, as predicted Ohsumi (1979), the prewhaling population size would have to have been significantly lower than the lower bound surrounding the long-term average (671 000 individuals; 95% CI: 374 000–1 150 000) and a krill surplus would have returned them close to or greater than the estimated mean population size in <100 years. Alternatively, for Antarctic minke whales to have increased by ~ 3 -fold, as predicted by the Antarctic ecosystem model of Mori & Butterworth (2006), the prewhaling population size would have to have been only slightly less than the lower bound on the long-term average (~ 319 000) and current size would have to be ~ 957 000, a number greater than all recently published estimates. While the Mori & Butterworth (2006) model may be biologically feasible, it is dependent on the assumption of top-down forcing (competition) and a number of input parameters, such as an initial 1780 abundance of 319 000 individuals (an estimate not based upon empirical data). As a result, there remains no direct evidence for competitive release within the Antarctic minke whale as a result of a krill surplus. An interesting area for future research would be to re-run the Mori & Butterworth (2006) ecosystem model with our genetic estimate of long-term abundance as a prior distribution for initial abundance and determine the extent to which support for the Krill Surplus Hypothesis depends upon on the initial abundance parameter.

Trophic cascades and the Antarctic marine ecosystem

If the Krill Surplus Hypothesis is not valid for minke whales, why then would the removal of ~ 2 million large baleen whales fail to result in competitive release for minke whales? One possibility that has been mentioned (Kawamura 1978), but not thoroughly investigated is that minke whales are not resource-limited because krill abundance exceeds the demands of krill-dependent predators in the Antarctic Marine Ecosystem. Another possibility is that, as the smallest baleen whale in the world, minke whales do not use krill in the same way and at the same time as whales that are between 3 and 11 times heavier (Fig. 1a). Niche specialization would make competitive release less likely and recent evidence indicates minke and humpback whales partition food resources by depth within the water column, krill size, and aggregation area (Friedlaender *et al.* 2009).

It is widely accepted that the removal of ~ 2 million large baleen whales had a profound impact on the Antarctic Marine Ecosystem, but over a half-century later,

direct evidence for a post-whaling competitive release across species remains elusive (reviewed by Ainley *et al.* 2007; Ballance *et al.* 2006; Nicol *et al.* 2007). In some species, population size appears to be limited by factors other than krill abundance. For example, recent evidence indicates population levels in Adelie penguins may be set by the availability of sea ice in addition to food availability (Fraser *et al.* 1992; Forcada *et al.* 2006). For minke whales, mechanisms that limit population size are not well understood and both top-down and bottom-up forces remain possibilities. Our results add to a growing body of literature suggesting that top-down and bottom-up forces must be considered concurrently when attempting to explain forces regulating populations within the Antarctic Marine Ecosystem (Ainley *et al.* 2007; Nicol *et al.* 2007). These data highlight the need for caution when making management recommendations to hunt Antarctic minke whales based upon the assumption they are unusually abundant and in direct competition with other recovering whale species.

Acknowledgements

We thank E. Archer, E. Alter, T. Branch, B. Brownell, J. Cooke, D. Demaster, A. Lang, K. Martien, P. Morin, M. Pespeni, M. Pinsky, R. Waples, and J. Wiedenmann for comments on early versions of this manuscript. We thank N. Funahashi and the International Fund for Animal Welfare for market sample collection. This work was supported by a grant from the Marsden Fund of the New Zealand Royal Society (UOA309) and the Lenfest Ocean Program (#2004-001492-023).

References

- Ainley D, Ballard G, Ackley S *et al.* (2007) Opinion: paradigm lost, or is top-down forcing no longer significant in the Antarctic marine ecosystem? *Antarctic Science*, **19**, 283–290.
- Alter SE, Palumbi SR (2007) Could genetic diversity in eastern North Pacific gray whales reflect global historic abundance? *Proceedings of the National Academy of Sciences of the United States of America*, **104**, E3–E4.
- Alter SE, Palumbi SR (2009) Comparing evolutionary patterns and variability in the mitochondrial control region and cytochrome B in three species of baleen whales. *Journal of Molecular Evolution*, **68**, 97–111.
- Alter SE, Rynes E, Palumbi SR (2007) DNA evidence for historic population size and past ecosystem impacts of gray whales. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 15162–15167.
- Atkinson QD, Gray RD, Drummond AJ (2008) mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. *Molecular Biology and Evolution*, **25**, 468–474.
- Ballance LT, Pitman RL, Hewitt RP *et al.* (2006) The removal of large whales from the Southern Ocean: evidence for long-term ecosystem effects? In: *Whales, Whaling and Ocean Ecosystem* (eds Estes JA, Demaster DP, Doak DF, Williams TM, Brownell RLJ), pp. 215–230. University of California Press, Berkeley.
- Branch TA (2006) Abundance estimates for Antarctic minke whales from three completed circumpolar sets of surveys, 1978/79 to 2003/2004. Paper presented to IWC Scientific Committee SC/58/1A18.
- Branch TA, Butterworth DS (2001) Southern Hemisphere minke whales: standardized abundance estimates from the 1978/79 to 1997/98 IDCR-SOWER surveys. *Journal of Cetacean Research and Management*, **3**, 143–174.
- Butterworth DS, Punt AE, Geromont HF, Kato H, Fujise Y (1999) Inferences on the dynamics of Southern Hemisphere minke whales from ADAPT analyses of catch-at-age information. *Journal of Cetacean Research and Management*, **1**, 11–32.
- Chamary JV, Parmley JL, Hurst LD (2006) Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nature Reviews Genetics*, **7**, 98–108.
- Clapham PJ, Baker CS (2002) Modern whaling. In: *Encyclopedia of Marine Mammal* (eds Perrin WF, Wursig B, Thewissen GGM), pp. 1328–1332. Academic Press, New York.
- Clapham PJ, Palsboll PJ, Pastene LA, Smith T, Walloe L (2005) Annex S: estimating pre-whaling abundance. *Journal of Cetacean Research and Management*, **7**(Suppl.), 386–387.
- Clapham PJ, Childerhouse S, Gales NJ *et al.* (2007) The whaling issue: conservation, confusion, and casuistry. *Marine Policy*, **31**, 314–319.
- Croxall JP, Trathan PN, Murphy EJ (2002) Environmental change and Antarctic seabird populations. *Science*, **297**, 1510–1514.
- Emerson BC (2007) Alarm bells for the molecular clock? No support for Ho *et al.*'s model of time-dependent molecular rate estimates *Systematic Biology*, **56**, 337–345.
- Estes JA, Tinker MT, Williams TM, Doak DF (1998) Killer whale predation on sea otters linking oceanic and nearshore ecosystems. *Science*, **282**, 473–476.
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics (Online)*, **1**, 47–50.
- Forcada J, Trathan P, Reid K, Murphy E, Croxall J (2006) Contrasting population changes in sympatric penguin species in association with climate warming. *Global Change Biology*, **12**, 411–423.
- Frank KT, Petrie B, Choi JS, Leggett WC (2005) Trophic cascades in formerly cod-dominated ecosystem. *Science*, **308**, 1621–1623.
- Frank K, Petrie B, Shackell N (2007) The ups and downs of trophic control in continental shelf ecosystems. *Trends in Ecology and Evolution*, **22**, 236–242.
- Frankham R (1995) Effective population size/adult population size ratios in wildlife: a review. *Genetic Research*, **66**, 96–107.
- Fraser WR, Trivelpiece WZ, Ainley DG, Trivelpiece SG (1992) Increases in Antarctic penguin populations: reduced competition with whales or a loss of sea ice due to environmental warming? *Polar Biology*, **11**, 525–531.
- Friedlaender AS, Lawson GL, Halpin PN (2009) Evidence of resource partitioning and niche separation between humpback and minke whales in Antarctica. *Marine Mammal Science*, **25**, 402–415.

- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.
- Gales NJ, Kasuya T, Clapham PJ, Brownell RL Jr (2005) Japan's whaling plan under scrutiny: useful science or unregulated commercial whaling? *Nature*, **435**, 883–884.
- Harrigan RJ, Mazza ME, Sorenson MD (2008) Computational vs. cloning: evaluation of two methods for haplotype determination. *Molecular Ecology Resources*, **8**, 1239–1248.
- Ho S, Phillips MJ, Cooper A, Drummond AJ (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution*, **22**, 1561–1568.
- Holt SJ, Mitchell ED (2004) Counting whales in the North Atlantic. *Science*, **303**, 39–40.
- Hudson RR (2002) Generating samples under a Wright-Fisher neutral model. *Bioinformatics*, **18**, 337–338.
- Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147–164.
- ICR (2006) Why whale research?. Available at: <http://icrwhale.org/04-B-jen.pdf>.
- Jackson JA, Baker CS, Vant M *et al.* (2009) Big and slow: phylogenetic estimates of molecular evolution in baleen whales (Suborder Mysticeti). *Molecular Biology and Evolution*, **11**, 2427–2440.
- Kato H, Kishino H, Fujise Y (1990) Some analysis on age composition and segregation of Southern minke whales using samples obtained from the Japanese feasibility study in 1987/88. *Reports of the International Whaling Commission*, **40**, 249–256.
- Kato H, Fujise Y, Kishino H (1991) Age structure and segregation of Southern minke whales by the data obtained during Japanese research take in 1988/89. *Reports of the International Whaling Commission*, **41**, 287–292.
- Kawamura A (1978) An interim consideration on a possible interspecific relation in southern baleen whales from the viewpoint of their food habits. *Reports of the International Whaling Commission*, **28**, 411–420.
- Klinowska M (1991) Minke Whale. In: *Dolphins, Porpoises and Whales of the World: The IUCN Red Data Book* (ed. Klinowska M), p. 379. IUCN, Gland, Switzerland.
- Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics*, **22**, 768–770.
- Lasken RS, Egholm M (2003) Whole genome amplification: abundant supplies of DNA from precious samples or clinical specimens. *Trends Biotechnology*, **21**, 531–535.
- Laws RM (1977) Seals and whales of the Southern Ocean. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **279**, 81–96.
- Loeb V, Siegel V, Holm-Hansen O *et al.* (1997) Effects of sea-ice extent and krill or salp dominance on the Antarctic food web. *Nature*, **387**, 897–900.
- Lubick N (2003) New count of old whales adds up to big debates. *Science*, **301**, 451.
- Mori M, Butterworth DS (2006) A first step towards modeling the krill-predator dynamics of the Antarctic ecosystem. *CCAMLR Science*, **13**, 217–277.
- Morishita J, Goodman D (2001) Competition between fisheries and marine mammals-feeding marine mammals at the expense of food for humans. In: *Proceedings of the Third World Fisheries Congress: Feeding the World With Fish in the Next Millennium the Balance Between Production and Environment*, Symposium 38 (eds Phillips B, Megrey BA, Zhou Y). pp. 403–408. American Fisheries Society, Bethesda, MD.
- Myers RA, Worm B (2003) Rapid worldwide depletion of predatory fish communities. *Nature*, **423**, 280–283.
- Nicol S, Pauly T, Bindoff NL *et al.* (2000) Ocean circulation off east Antarctica affects ecosystem structure and sea-ice extent. *Nature*, **406**, 504–507.
- Nicol S, Croxall J, Trathan P, Gales NJ, Murphy E (2007) Paradigm misplaced? Antarctic marine ecosystems are affected by climate change as well as biological processes and harvesting *Antarctic Science*, **19**, 291–295.
- Nicol S, Worby A, Leaper R (2008) Changes in the Antarctic sea ice ecosystem: potential effects on krill and baleen whales. *Marine Freshwater Research*, **59**, 361–382.
- Nunney L (1991) The influence of age structure and fecundity on effective population size. *Proceedings of the Royal Society of London. Series B, Biological Science*, **246**, 71–76.
- Nunney L (1993) The influence of mating system and overlapping generations on effective population size. *Evolution*, **47**, 1329–1341.
- Nunney L, Elam DR (1994) Estimating the effective population size of conserved populations. *Conservation Biology*, **8**, 175–184.
- Ohsumi S (1979) Population assessment of the Antarctic minke whale. *Reports of the International Whaling Commission*, **29**, 407–420.
- Pace ML, Cole JJ, Carpenter SR, Kitchell JF (1999) Trophic cascades revealed in diverse ecosystems. *Trends in Ecology and Evolution*, **14**, 483–488.
- Pastene LA, Goto M, Itoh S, Numachi K (1996) Spatial and temporal patterns of mitochondrial DNA variation in minke whales from Antarctic areas IV and V. *Reports of the International Whaling Commission*, **46**, 305–314.
- Pastene LA, Goto M, Kanda N *et al.* (2007) Radiation and speciation of palagic organisms during periods of global warming: the case of the common minke whale, *Balaenoptera acutorostrata*. *Molecular Ecology*, **16**, 1481–1495.
- Power ME (1992) Top-down and bottom-up forces in food webs – do plants have primacy? *Ecology*, **73**, 733–746.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Roman J, Palumbi SR (2003) Whales before whaling in the North Atlantic. *Science*, **301**, 508–510.
- Rozas J, Sanchez-Delbarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.
- Shapiro B, Drummond AJ, Rambaut A *et al.* (2004) Rise and fall of the Beringian steppe bison. *Science*, **306**, 1561–1565.
- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.
- Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.

- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.
- Thomson RB, Butterworth DS (1999) Has the age at transition of southern hemisphere minke whales declined over recent decades? *Marine Mammal Science*, **15**, 661–682.
- Zhang DX, Hewitt GM (2003) Nuclear DNA analysis in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, **6**, 563–584.

This work is part of a long-term study on the genetic diversity and historical ecology of whale populations. The multi-disciplinary team includes expertise in molecular ecology, evolutionary genetics, statistical genetics, the dynamics of whale populations and the history of whaling. Several authors have acted as national delegates or invited participants to the Scientific Committee of the International Whaling Commission.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Nuclear locus description

Fig. S1 Comparison of variation in individual locus mutation rates with the gamma distribution estimated by LAMARC. Bars represent a histogram of mean mutation rate obtained by bootstrap re-sampling over individual locus mutation rates from Alter *et al.* (2007) and Jackson *et al.* (2009). Blue line is a histogram of 50 000 means of 11 random variables simulated from the gamma distribution with the shape parameter estimated in LAMARC.

Fig. S2 The figure illustrates considerable variation in generation length across years and areas in the Antarctic minke whale. Generation length was estimated as the average age of sexually mature individuals from commercial and JARPA catch-at-age matrices (see Table 1 in Butterworth *et al.* 1999). Blue dots indicate catches in area IV and red dots indicate catches in area V (Fig. 1).

Fig. S3 Convergence of LAMARC-estimated θ at 11 nuclear loci using 10 different realizations from PHASE's posterior distribution across three separate LAMARC runs. Each curve represents the estimated posterior distribution of θ at each locus. The colours distinguish between the different LAMARC runs and the 10 curves represent alternate PHASE realizations.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.