Lack of nuclear differentiation suggests reproductive connectivity between the 'southern feeding group' and the larger population of eastern North Pacific gray whales, despite previous detection of mitochondrial differences

ANNA M. D'INTINO[1], JAMES D. DARLING[2], JORGE URBÁN R.[3] AND TIMOTHY R. FRASIER[1]

Contact e-mail: timothy.frasier@smu.ca

ABSTRACT

During winter, eastern North Pacific gray whales migrate south to calving grounds in the lagoons of Baja California, and in spring they migrate north to their summer feeding grounds in the Chukchi and Beaufort Seas. Although the majority of the population makes this migration, a small subset of the population known as the 'southern feeding group' ends their northward migration early, spending summers feeding in waters ranging from northern California to southern Alaska. Previous analyses based on photo-ID and mtDNA data indicate that this seasonal substructuring results from maternally-directed site fidelity to different feeding grounds, and that this site fidelity and feeding ground preference is passed from mothers to their offspring. It is currently assumed, but not known, that the individuals of the southern feeding group mate with the rest of the population, and therefore that the eastern North Pacific gray whale represents one interbreeding population. Testing this assumption and understanding how these whales are related to the rest of the population, is key to making appropriate management decisions, which are particularly relevant given the recent increase in potential removals, or threats in the area such as the proposed resumption of aboriginal whaling, and increased oil pipeline development and subsequent vessel traffic. This paper analyses 15 nuclear microsatellite loci in 59 individuals from the southern feeding group and 40 individuals from the calving lagoons (representative of the larger population) to test the hypothesis that the eastern North Pacific gray whale represents one interbreeding population. No indication of population substructuring was found based on these nuclear loci, suggesting that all sampled whales do indeed represent one interbreeding population. Combined, these data from mitochondrial and nuclear markers therefore suggest one interbreeding population that is seasonally subdivided based on maternally-directed site fidelity to different feeding areas.

KEYWORDS: EASTERN NORTH PACIFIC; GRAY WHALE; REPRODUCTION; WHALING–ABORIGINAL; FEEDING GROUNDS; GENETICS; SITE FIDELITY; SEGREGATION

INTRODUCTION

The eastern North Pacific gray whale (*Eschrichtius robustus*) was extensively hunted from the mid-1800s through the early 1900s, reducing the population to just a fraction of its estimated pre-exploitation population size (Butterworth *et al.*, 2002; Henderson, 1984; Reilly, 1992). However, since gaining international protection, the population has steadily increased to roughly 20,000 individuals (Laake *et al.*, 2012; Rugh *et al.*, 2005; Sheldon and Laake, 2002). This recovery resulted in the removal of this population from the US Endangered Species List in 1994, and some data suggest it has recovered to its pre-exploitation population size (Alter *et al.*, 2007; Moore *et al.*, 2001; Rugh *et al.*, 2005; Wade, 2002). Despite this recovery, the population is still faced with numerous threats throughout its range, particularly in the area encompassing the Pacific northwest of the USA and the Pacific southwest of Canada (roughly Northern California through southeast Alaska). Here, there are several proposed activities where informed management will be critical. These include the proposed resumption of gray whale hunts by some aboriginal groups, the development of new oil pipelines off the British Columbia coast and the subsequent increase in tanker traffic for shipping oil to Asia.

One of the primary topics of relevance to management is population structure: how individuals are divided in space and/or time such that some groups may be differentially affected by direct hunting or by non-intentional threats. If such structuring exists, then the different groups often require separate management/conservation consideration because the detrimental effects will not be spread evenly throughout the population, but instead will disproportionally affect the different groups (e.g. Hoelzel, 1998; Taylor, 2005; Wang, 2009). Such localised impacts on structured populations, if not considered, can nullify otherwise well-planned management/conservation initiatives. Indeed, there are several case studies where the effectiveness of conservation actions has been compromised because population structure was not taken into consideration (e.g. Daugherty *et al.*, 1990; Frankham *et al.*, 2002).

Previous studies have detected seasonal population substructuring in the eastern North Pacific gray whale in relation to summer feeding ground use (Frasier *et al.*, 2011). As with many other baleen whales, gray whales show a seasonal migration from low-latitude calving grounds in the winter to high-latitude feeding grounds in the summer. The winter calving grounds for this population are located in the lagoons of Baja California (Findley and Vidal, 2002; Swartz, 1986; Swartz *et al.*, 2006), whereas during the summer the majority of the population feeds in the Bering and Chukchi Seas (Moore and Ljungblad, 1984). However, there is a small

[1]*Saint Mary's University, Department of Biology, 923 Robie Street, Halifax, Nova Scotia, B3H 3C3, Canada.*
[2]*Pacific Wildlife Foundation, Vancouver, British Columbia, V3H 1V6, Canada.*
[3]*Departamento de Biologia Marina, Universidad Autonoma Baja California Sur, La Paz, BCS 23081, Mexico.*

subset of the population—estimated at roughly 200 individuals (Calambokidis *et al.*, 2002) that spends the summer in lower-latitude feeding areas ranging from northern California to southeastern Alaska (Calambokidis *et al.*, 2002; Darling, 1984; Hatler and Darling, 1974; Pike, 1962; Swartz *et al.*, 2006). This group is often referred to as the 'southern feeding group', but is also referred to as the 'Pacific coastal feeding group' by the Scientific Committee of the International Whaling Commission (IWC, 2011; 2013a). Data from photo-ID and mitochondrial DNA (mtDNA) show that this seasonal population substructuring results from maternally-directed site fidelity to different feeding areas, and that this differential use of feeding areas is passed from mothers to offspring (Calambokidis *et al.*, 2002; Calambokidis *et al.*, 2010; Darling, 1984; Frasier *et al.*, 2011). This type of seasonal substructuring is common in baleen whales (e.g. Baker *et al.,* 1990; Malik *et al.,* 1999).

Despite this substructuring during the summer, it has always been assumed, but not known, that individuals from all feeding areas utilise the same mating area(s), and therefore represent one interbreeding population. Unfortunately, there is a lack of clarity regarding where fertilisation likely occurs, and hence where the mating grounds are. Sexual behaviour is frequently observed on the winter calving grounds (e.g. Swartz, 1986), and indeed these are often referred to as 'breeding grounds' (e.g. Alter *et al.*, 2009; Goerlitz *et al.*, 2003; Jones, 1990). However, the limited physiological data available actually suggest that fertilisation most often occurs during the southward migration, prior to arrival at the lagoons (Rice and Wolman, 1971). Thus, there is potential for differential feeding area use to also result in substructuring with respect to reproductive patterns.

There is also evidence that individuals show differential use of the calving grounds, with some females showing fidelity to particular lagoons (Jones, 1990; Goerlitz *et al*., 2003; Alter *et al*., 2009). Combined, there are enough questions regarding the timing and location of mating, as well as potential for differential habitat use of potential 'breeding' grounds, to warrant a full evaluation of the hypothesis of one interbreeding population. This paper uses data from nuclear microsatellite loci to compare genetic data from individuals of the southern feeding group to samples obtained from one of the calving lagoons (Laguna San Ignacio) used here as representatives of the larger population to test the hypothesis that individuals of the southern feeding group interbreed with individuals from the larger population. These data, in combination with previous information regarding structuring of mitochondrial haplotypes, can identify the degree of substructuring of the southern feeding group.

METHODS

Sample collection

Samples used for these analyses were collected over many years as part of other long-term research programmes on eastern North Pacific gray whales: off Vancouver Island, Canada by one of the authors (JDD); and off San Ignacio Bay, Mexico by another (JUR). Specifically, 86 samples from Vancouver Island (representing the southern feeding group, collected from 1996–2010) and 51 samples from San Ignacio Bay (representing the larger population, collected from 1996–1997) were analysed. All samples were collected using a crossbow and a modified bolt, as is common for collecting small skin samples from free-ranging large whales (Lambertsen, 1987; Palsbøll *et al.*, 1991). This method of sample collection has been extensively scrutinised, has proven to be safe, and does not have any short- or long-term impacts on the whales, other than an immediate startle response (Best *et al.*, 2005; Brown *et al.*, 1991). All sample collection procedures were conducted under permits obtained from the relevant governmental agencies. The timing of sample collection overlaps between the two sample sets, as is appropriate for comparison. The longer time for sample collection off Vancouver Island should not bias the results because the long life span and generation time of gray whales suggest that substantial genetic change within this putative population would take longer than the time interval sampled here. Thus, both sample sets should represent comparable genetic pools for analysis.

Genetic analyses

DNA was extracted from ~40mg of tissue from all samples using standard phenol:chloroform procedures as commonly used for whale skin (e.g. Wang *et al.*, 2008). The quantity of DNA obtained from each sample was estimated based on spectrophotometry using a NanoDrop 2000 (Thermo Scientific Inc.). The quality of DNA obtained (i.e. the amount of DNA degradation) was assessed based on electrophoresis through 2.0% agarose gels stained with SYBR Green I (Invitrogen). Sex was determined for each sample based on PCR amplification of a region on the X and Y chromosomes using the primers described in Gilson *et al.* (1998).

To identify useful microsatellite markers for this study, 23 loci were screened for amplification and variability in gray whales (Table 1). These loci were chosen because previous experience showed that they amplify well, and are highly variable, across a wide range of whale species. Specifically, loci were initially tested based on the amplification of two gray whale samples (one from the southern feeding group and one from the larger population) using annealing temperatures of 50, 55 and 60°C. The reactions contained 10ng of template DNA, 1X PCR Buffer (20 mM Tris pH 8.4, 50 mM KCl), 1.5 mM $MgCl_2$, 0.05 U/µl *Taq* DNA polymerase (Invitrogen), 0.2 mM each dNTP (Invitrogen), 0.3 µM each primer, and 0.1 mg/ml of bovine serum albumin (BSA) (Invitrogen). The cycling conditions were as follows: an initial denaturing step of 94°C for 5 minutes; 30 cycles of 94°C for 30 seconds, annealing temperature for 1 minute, and 72°C for 1 minute; followed by a final extension step of 60°C for 45 minutes. All PCR was conducted on Veriti® 96-well thermal cyclers (Applied Biosystems). PCR products were then size-separated and visualised via electrophoresis through 1.5% agarose gels stained with ethidium bromide. Loci that showed a clean PCR product (without amplification of multiple regions) were selected for further development.

For primer pairs that amplified well, the forward primer was re-ordered with one of four fluorescent tags (6FAM-Blue, VIC-Green, PET-Red, or NED-Yellow). Each locus was then screened for variation by amplifying 10 samples using the optimal conditions that had previously been

Table 1

Name, and reference for each of the 23 microsatellite loci tested for amplification and variability in gray whales.

Locus	Reference
EV1*Pm*	Valsecchi and Amos (1996)
EV5*Pm*	Valsecchi and Amos (1996)
EV14*Pm*	Valsecchi and Amos (1996)
EV37*Mn*	Valsecchi and Amos (1996)
EV94*Mn*	Valsecchi and Amos (1996)
EV104*Mn*	Valsecchi and Amos (1996)
FCB1	Buchanan *et al*. (1996)
FCB4	Buchanan *et al*. (1996)
FCB5	Buchanan *et al*. (1996)
FCB14	Buchanan *et al*. (1996)
FCB17	Buchanan *et al*. (1996)
GATA028	Palsbøll *et al.* (1997)
GATA098	Palsbøll *et al.* (1997)
GATA417	Palsbøll *et al.* (1997)
GT023	Bérubé *et al*. (2000)
IGF1	Barendse *et al.* (1994)
RW31	Waldick *et al*. (1999)
RW34	Waldick *et al*. (1999)
RW48	Waldick *et al*. (1999)
SW10	Richard *et al*. (1996)
SW13	Richard *et al*. (1996)
SW19	Richard *et al*. (1996)
TexVet5	Rooney *et al*. (1999)

determined. PCR amplification was carried out using the same conditions as described for testing annealing temperatures. PCR products were de-salted via ethanol precipitation (Irwin *et al.*, 2003), and size-separated and visualised on an ABI 3500xl capillary-based genetic analyser (Applied Biosystems).

Based on these amplification data (i.e. peak height and allele ranges) multiplex reactions where multiple loci are amplified simultaneously in the same PCR were developed through testing the amplification of different combinations of loci. These tests resulted in combinations of loci that minimised the number of reactions that were needed to amplify the variable loci. These protocols were then used to genotype all individuals. Alleles were scored using the GeneMarker software (SoftGenetics). Each 96-well plate of samples contained 2 individuals as 'standards', meaning that these same two individuals were present on all plates, to ensure consistency in genotyping across plates. All samples were scored by two individuals in a double-blind fashion to identify any potential scoring errors (Morin *et al.*, 2010).

Statistical analyses

Once genotypes for each sample had been obtained, the program CERVUS (Marshall *et al.*, 1998) was used to identify any duplicate genotypes (i.e. individuals that had unknowingly been sampled more than once). Based on these data, duplicate individuals were removed. CERVUS was also used to obtain estimates of allele frequencies and to estimate the frequency of null alleles for each locus. Loci were tested for deviations from linkage and Hardy-Weinberg Equilibrium (HWE) using exact tests as implemented in the program GENEPOP (Rousset, 2008).

Population structure was assessed using 'classical' approaches based on estimating differentiation of allele frequencies between pre-defined groups of individuals and Bayesian techniques. Classical methods have the benefit of being far more powerful than available Bayesian methods but have the detriment of requiring pre-defined groups of individuals, which may or may not represent the real biological patterns (e.g. Waples and Gaggiotti, 2006). On the other hand, Bayesian methods allow for simultaneous assessment of the number of groups represented by the sampled individuals and the assignment of individuals to those groups (and therefore do not requiring pre-defined groupings) but suffer from lower power.

The classical assessment of population differentiation was conducted using the program GENEPOP. Here, individuals were categorised as representing either the southern feeding group or the larger population, based on sampling location. Specifically, samples collected off Vancouver Island were classified as the southern feeding group, while samples collected on the calving grounds off Baja California were classified as representing the larger population. Based on this division, estimates of F_{ST} were obtained and exact tests of population differentiation were conducted.

To assess the power to detect population structure using the methods implemented in GENEPOP (exact tests), the simulation program POWSIM (Ryman and Palm, 2006) was used. Specifically, conditions were simulated that would result in differing levels of differentiation (F_{ST} = 0.001, 0.0025, 0.005, 0.01, 0.025, and 0.05), given the characteristics of the loci. There are two biologically realistic scenarios that can result in the same F_{ST} value. First, two populations that are not completely isolated will eventually reach an equilibrium F_{ST} value dependent upon the migration rate between populations (N_m). Second, if two populations have recently become reproductively isolated, they will drift apart, with F_{ST} values increasing with increasing time since divergence. POWSIM obtains desired F_{ST} values under the latter scenario by generating a single simulated population and then splitting it into equally sized populations with complete isolation once the split occurs. In this way, different F_{ST} values are obtained based on the number of generations that have passed since the split. Thus, users obtain estimates of the power to detect different degrees of differentiation by selecting combinations of N_e and t that result in the desired F_{ST} value, where $F_{ST} = 1 - (1 - 1/2N_e)^t$ (e.g. Nei, 1987, p.359). Due to uncertainty regarding N_e for the two putative gray whale populations, power to detect the different F_{ST} values (0.001, 0.0025, 0.005 0.01, 0.02, 0.025, and 0.05) was estimated using N_e values spanning three orders of magnitude (500, 5,000, and 50,000) and using t values for each that would result in the appropriate F_{ST} value. Simulations for each scenario (N_e and F_{ST} value) were conducted 100 times and the proportion of iterations where significant differentiation was detected (using a critical α value of 0.05) was recorded.

Population structure was also assessed without making *a priori* assumptions about the nature of population structure (e.g. how many groups there are, and which individuals represent each group). These analyses were conducted in two different ways: using the programs STRUCTURE (Hubisz *et al.*, 2009; Pritchard *et al.*, 2000), and STRUCTURAMA (Huelsenbeck and Andolfatto, 2007). For the analyses in STRUCTURE, the number of groups, and the membership of individuals within those groups, were estimated based a run length of 1,000,000 Markov Chain Monte Carlo

(MCMC) steps, with 50,000 steps as the burn-in period. The program was run assuming that allele frequencies were correlated between groups, and allowing for admixture (i.e. allowing for individuals to have ancestry in more than one group). The program was run 16 times, testing for 1–4 populations (K = 1–4), with four iterations of each K. The average likelihood over the four iterations for each K was taken as the likelihood for that K.

The program STRUCTURAMA works in a similar manner as STRUCTURE but differs in how the user specifies the number of populations to be tested. With STRUCTURE, the user must explicitly specify the number of populations considered and then run the program independently for each hypothesised number, and subsequently compare the probabilities associated with each. With STRUCTURAMA the number of populations considered can be a random variable within the model (Pella and Masuda, 2006) and therefore the posterior probabilities associated with a range of values for the number of putative populations can be obtained within a single run, without requiring *a priori* specification by the user (Huelsenbeck and Andolfatto, 2007). STRUCTURAMA was run allowing the number of populations to be a random variable with a Dirichlet process prior. The alpha value (which determines the shape of Dirichlet prior, where smaller values result in individuals being distributed across fewer populations and larger values result in individuals being dispersed across more populations) was also treated as a random variable, following a gamma distribution. A shape parameter of 1.0001 and a scale parameter of 0.0001 were initially used, which result in a relatively flat distribution. However, to test the robustness of the subsequent posterior probabilities to the choice of shape and scale values, STRUCTURAMA was also run with the shape and scale values both set to 1, which results in an L-shaped distribution. Each scenario was run three times, to ensure consistency between runs under the same conditions. All analyses were run for 1,000,000 MCMC steps, with 10,000 steps used as the burn-in.

Finally, it has been argued that estimating migration rates between putative populations is more biologically meaningful than simply rejecting panmixia when trying to identify biologically independent 'units' for conservation (e.g. Palsbøll *et al.*, 2006). Estimating migration rates from genetic data is challenging, however, particularly when migration rates are high. This is because demographic independence can occur at migration rates that are high enough to genetically homogenise allele frequencies (e.g. Waples and Gaggiotti, 2006). Regardless, great progress has recently been made in the development of analytical techniques designed to address this issue and better infer biological processes from genetic data. For management, the parameter of interest is estimated contemporary migration rates. BayesAss has become the software commonly used for this purpose (Wilson and Rannala, 2003). However, the approach implemented therein is known to perform poorly when genetic differentiation is low (Faubet *et al.*, 2007). Our results suggest extremely low differentiation between the putative groups and perhaps none (see Results), therefore BayesAss was not appropriate for estimating migration rates for our data. Instead, the programs IMa and MIGRATE were used to jointly estimate migration rates in both directions, as well as effective population sizes of each putative population (Beerli, 2006; Beerli and Felsenstein, 1999; 2001; Hey and Nielsen, 2004).

RESULTS

The average yield of DNA from the extraction process was 2.80 nanograms (ng) of DNA per milligram (mg) of tissue for the samples from Vancouver Island, and 0.42 ng/mg for the samples from San Ignacio Bay. This lower yield from the San Ignacio Bay samples likely results from the fact that these samples were stored in ethanol rather than DMSO, and DNA degradation is known to occur at a higher rate in ethanol than DMSO (e.g. Michaud and Foran, 2011).

The tests of microsatellite amplification and variability resulted in the identification of 15 loci that amplify well in gray whales, and are also variable. Multiplex reaction development resulted in all 15 loci being amplified in 5 reactions (Table 2), which were subsequently used for genotyping all samples.

No genotypes differed between the duplicate scoring personnel, indicating that allele calls were consistent across scorers. Genotypes were considered to be 'full' if individuals were missing data from 3 or fewer loci (i.e. they had data for at least 12 of the 15 loci). Full microsatellite profiles were not obtained for 14 samples, resulting in 123 genotyped samples that were used for subsequent analyses. Analysis of the genotypes identified 24 duplicate sampling events (23 off Vancouver Island, 1 from Laguna San Ignacio). All re-sampling events were from the same location (i.e. there were no cases where one individual was sampled in one location and re-sampled at the other). These duplicates were removed from the data set, resulting in genotype data for 99 individuals (59 from off Vancouver Island, 40 from Laguna San Ignacio). The genotypes did not show any significant deviations from Hardy-Weinberg Equilibrium expectations when analysed independently for each putative population, or when the data were combined into one 'population' (Table 3). Out of the 105 pairwise comparisons for assessing linkage between loci (considering all individuals together), six (5.7%) had p-values <0.05. However, none of these were significant after Bonferroni correction (Hochberg, 1988).

Table 2

Amplification information. Included is the locus name, fluorescent label, and reaction number for all loci. The annealing temperature for all reactions is 55°C.

Locus	Label	Reaction
EV14*Pm*	6FAM	1
EV37*Mn*	VIC	1
FCB14	VIC	1
GATA028	NED	1
FCB5	NED	1
GT023	VIC	2
FCB4	PET	2
EV1*Pm*	NED	2
TexVet5	NED	2
FCB17	6FAM	3
GATA417	PET	3
SW10	NED	3
SW13	6FAM	4
EV94*Mn*	6FAM	5
RW31	VIC	5

Table 3

Characteristics of each locus in each putative population for: (a) the southern feeding group; and (b) the larger population. Included is the locus name, the number of alleles, observed heterozygosity (H_O), expected heterozygosity (H_E, Nei, 1987), the polymorphic information content (*PIC*, Botstein *et al*., 1980), and the p-value for deviation from HWE. No p-values were statistically significant after Bonferroni correction (correction conducted independently for each putative population).

Locus	Alleles	H_O	HE	PIC	P-Value
(a) Southern feeding group					
EV14*Pm*	10	0.828	0.852	0.826	0.688
EV37*Mn*	17	0.845	0.886	0.867	0.204
FCB14	7	0.741	0.808	0.773	0.273
FCB5	4	0.500	0.438	0.402	0.765
GATA028	5	0.780	0.753	0.704	0.437
GT023	6	0.741	0.741	0.688	0.642
EV1*Pm*	3	0.603	0.508	0.385	0.184
TexVet5	5	0.741	0.730	0.678	0.678
FCB4	3	0.143	0.250	0.221	0.008
FCB17	13	0.930	0.907	0.890	0.984
SW10	7	0.776	0.776	0.733	0.832
GATA417	7	0.707	0.723	0.676	0.161
SW13	8	0.706	0.630	0.552	0.062
EV94*Mn*	9	0.831	0.816	0.783	0.458
RW31	9	0.828	0.822	0.790	0.216
(b) Larger population					
EV14*Pm*	10	0.769	0.840	0.809	0.132
EV37*Mn*	15	0.914	0.873	0.848	0.644
FCB14	7	0.759	0.836	0.798	0.033
FCB5	4	0.500	0.489	0.451	0.745
GATA028	5	0.769	0.764	0.715	0.185
GT023	7	0.650	0.685	0.627	0.276
EV1*Pm*	2	0.564	0.498	0.371	0.517
TexVet5	5	0.725	0.683	0.621	0.857
FCB4	2	0.105	0.191	0.171	0.036
FCB17	14	0.972	0.903	0.881	0.996
SW10	7	0.750	0.805	0.766	0.295
GATA417	6	0.700	0.717	0.676	0.182
SW13	5	0.629	0.611	0.530	0.738
EV94*Mn*	9	0.806	0.810	0.770	0.757
RW31	9	0.825	0.815	0.777	0.301

No loci had estimates of null allele frequencies greater than 0.05.

Sex could be determined for 86 of the 99 individuals, with 38 males and 48 females. The DNA was too degraded from the remaining 13 individuals to obtain reliable sex information. Within each region, the sex ratios were 33 females: 24 males and 15 females: 14 males for Vancouver Island and Laguna San Ignacio, respectively.

The 'classic' tests (based on hypothesis testing of pre-defined groupings) did not show any significant signs of genetic differentiation between the genotypes of the southern feeding group and the larger population. Specifically, the F_{ST} estimate was –0.0010, with a *P*-value estimate of 0.489. Simulation analysis showed that the power to detect structure, if it exists, was quite high. Specifically, given the sample sizes and the characteristics of the loci, we would expect to detect population structure over 70% of the time with an F_{ST} value as low as 0.005 (Fig. 1). The STRUCTURE analyses also did not detect any significant genetic differentiation within the data set, with one population (K = 1) having the highest probability (Table 4). Similar results were also obtained with STRUCTURAMA, with the scenario of the data representing one single population having a higher probability than scenarios with any other putative number of populations (Table 5).

Table 4

Results from STRUCTURE analysis. Shown is the estimated natural logarithm (*ln*) of the probability of the data with the number of populations (K) ranging from one to four, and performing four iterations of each K. The bold value indicates the average K with the highest probability.

	K			
Iteration	1	2	3	4
1	–4,219.9	–4,243.6	–4,273.9	–4,276.9
2	–4,220.9	–4,243.6	–4,268.5	–4,272.6
3	–4,220.3	–4,238.9	–4,266.2	–4,257.1
4	–4,220.6	–4,228.6	–4,248.7	–4,263.9
Average	–4,220.4	–4,238.7	–4,264.3	–4,267.6

Table 5

Results from STRUCTURAMA analysis. Shown are the estimated posterior probabilities of the data representing 1–3 populations. The top panel shows these probabilities calculated with the alpha value for the Dirichlet process being a random variable following a gamma distribution with a shape parameter of 1.0001, and a scale parameter of 0.0001. The second panel shows these probabilities calculated when the shape and scale value parameters were both set to 1.

	Number of populations		
Iteration	1	2	3
gamma(1.0001,0.0001)			
1	0.96	0.04	0
2	0.96	0.04	0
3	0.96	0.04	0
gamma(1,1)			
1	0.97	0.03	0
2	0.97	0.03	0
3	0.97	0.03	0

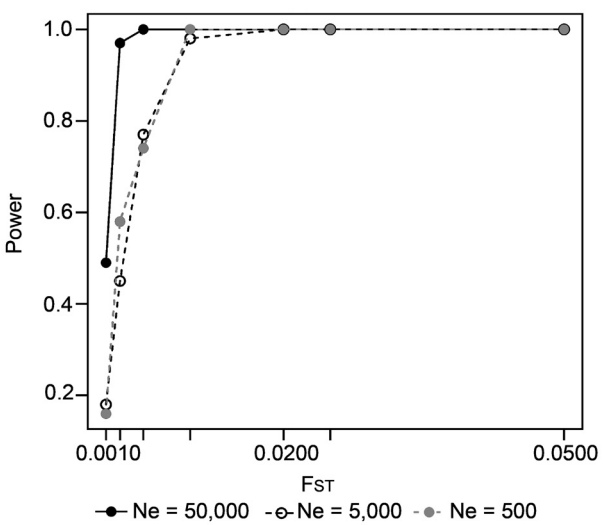

Fig. 1. Results from the POWSIM analyses. Shown is the power of the data (the proportion of simulations where population structure was detected (using a critical α value of 0.05)) under the different scenarios. The F_{ST} values tested were 0.05, 0.025, 0.02, 0.01, 0.005, 0.0025, and 0.001. These F_{ST} values were generated under three scenarios, N_e = 500, 5,000, and 50,000, with the time since divergence (t) varying to result in the desired F_{ST} values.

Despite testing a wide range of options with both IMa and MIGRATE, we were unable to get either program to converge on consistent estimates of migration rates. Our interpretation is that this inability is due to the lack of genetic differentiation of nuclear markers between the putative

groups. This interpretation, as opposed to a lack of information in the data, seems appropriate particularly because we have previously obtained consistent estimates from these same programs and the same individuals but based on mitochondrial data, which showed significant structuring and limited migration consistent with maternally-directed site fidelity (Frasier *et al.*, 2011).

DISCUSSION

The differential recovery of DNA from tissue stored in different solutions merits consideration for future sample storage. Several studies have compared the ability of different solutions (primarily ethanol and DMSO) to preserve DNA over long periods of time (e.g. Michaud and Foran, 2011; Seutin *et al.*, 1991). All such studies indicate that DMSO solutions preserve DNA at a higher quality, and over a longer period of time. However, many researchers and museum staff still use ethanol for the long-term preservation of tissue. The differential yields of DNA obtained here add to the growing amount of data suggesting that DMSO is the desirable storage solution for long-term storage of tissue that may be used as a source of DNA.

The results of all analyses of population structure lead to the same conclusion: a lack of differentiation of nuclear genotypes. The 'classical' tests did not detect significant differences in allele frequencies between whales of the southern feeding group and those sampled in Laguna San Ignacio, and both Bayesian approaches indicated that the probability that all samples originated from one single population was substantially higher than any other alternatives. These data suggest that the whales of the southern feeding group do indeed freely interbreed with whales that utilise other summer feeding grounds. Thus, from these data, it appears that the eastern North Pacific gray whale represents one interbreeding population.

One caveat of our study is that samples representing the larger population were all collected from whales in one of the lagoons Laguna San Ignacio (but spanning several years). Gray whales are not evenly distributed throughout the three known lagoons in winter. Instead, photo-ID data suggest that females show some site fidelity to different lagoons (e.g. Jones, 1990) and genetic data also suggest some structuring (Goerlitz *et al.*, 2003; Alter *et al*., 2009). However, the patterns and degree of structuring between lagoons remains unclear. For example, Alter *et al*. (2009) did not find significant structuring of mitochondrial haplotypes between the lagoons, but found slight but statistically significant differentiation of microsatellite alleles between Laguna San Ignacio and Bahia Magdalena. The authors attribute this pattern to either a high contemporary migration rate, or perhaps to stronger patterns of structuring being erased by whaling and only beginning to accumulate and leave a detectible genetic signature today.

Our justification for using samples from Laguna San Ignacio as representative of the larger population, are three-fold. First, based on numbers of single individuals and mother-calf pairs, Laguna San Ignacio represents the second most populated lagoon, with numbers of whales that are vastly larger than the estimated size of the southern feeding group (Jones and Swartz, 1984; Urban R *et al.*, 2003). Thus, Laguna San Ignacio is clearly used by a larger subset of the population than the southern feeding group. Second, although there are data suggesting that some females show fidelity to specific lagoons, there is also an abundance of data showing that some individuals move freely between lagoons, and that average residence times within the lagoon (for non-mother-calf pairs) are less than a week (Jones and Swartz, 1984; Urbán R *et al*., 2003), suggesting that at least single whales move readily between lagoons. Moreover, some mothers are known to utilise different lagoons in different years (Swartz, 1986). Lastly, even if mothers do show fidelity to specific lagoons, this should not result in corresponding structuring of nuclear gene flow because fertilisation is thought to take place during the southward migration, prior to arrival at the lagoons (Rice and Wolman, 1971). Thus, although it would be ideal to have representative samples from all known lagoons, the available data provide no reason to doubt that the samples from Laguna San Ignacio are representative of the larger population.

These data based on nuclear markers add to previous photo-ID and mtDNA data to provide a more complete picture of the relationship between seasonal habitat use patterns and gene flow throughout the population. The photo-ID and mtDNA data indicate that, during the summer, whales of the southern feeding group represent a seasonal subpopulation, where this differential habitat use is driven by maternally-directed site fidelity to this feeding area that is then passed on to their offspring (Frasier *et al*., 2011). However, the nuclear DNA data suggest that these whales breed with whales that show fidelity to other feeding grounds, and therefore are part of one interbreeding population. Thus, the combined picture is one of seasonal subdivision on summer feeding grounds, but with no such substructuring during the mating season, where all individuals in the eastern North Pacific represent one gene pool.

This result and interpretation is consistent with other data relating to known migration patterns and the timing of fertilisation. Briefly, it is thought that conceptions may occur in December, during the southern migration (Rice and Wolman, 1971). Whales that use the northern feeding grounds migrate through the areas occupied by the southern feeding group suggesting that whales from both known feeding grounds may migrate together the remainder of the way to the winter calving grounds (Darling, 1984). The peak time of migrants passing through the southern feeding area is mid-December to mid-January (Darling, 1984). Thus, the timing of fertilisation coincides with when whales from different feeding grounds become intermingled during their southern migration. This pattern indicates the strong potential for interbreeding regardless of any substructuring that may exist during the summer, or on the winter calving grounds.

Despite the presence of nuclear gene flow between whales from the southern feeding group and the rest of the population, this group still represents a separate management unit that warrants separate consideration with respect to the impacts of proposed threats, such as the resumption of traditional whaling and the impacts of oil distribution. This is the approach being used by the IWC Scientific Committee in examining the potential impacts of hunting (e.g. see IWC,

2013b). The presence of long-term site fidelity to this area that is passed on from mothers to offspring, indicates that these whales represent a seasonal subpopulation. Thus, detrimental impacts (e.g. 'takes') to these whales will not have a 'random' impact on the population at large, but will instead primarily impact these matrilines specifically. The resulting effect on this local subpopulation could be far greater than would be expected under the assumption of a single, unstructured population. Potential impacts could include the loss of knowledge of these feeding areas from this population, and localised extirpation. For example, if the whales that currently show this site fidelity are removed, then this information will be lost, and thus these whales will not likely be replaced by others from the larger population, resulting in localised extirpation. Indeed, the recognition of such seasonal subpopulations as separate management units is recommended, and common, for baleen whales (e.g. Dizon *et al.*, 1997).

ACKNOWLEDGEMENTS

This work was funded by the Canadian Wildlife Foundation, the Faculty of Graduate Studies and Research at Saint Mary's University, the Pacific Wildlife Foundation through donations from Remote Passages Marine Excursions, Jamie's Whaling Station and Ocean Outfitters in Tofino BC, and a grant from Clayoquot Biosphere Trust. We would like to thank Drs. Barbara Taylor and Phillip Morin for reviewing earlier drafts of this manuscript and providing valuable comments.

REFERENCES

Alter, S.E., Flores, S.R., Nigenda, S., Urban, J.R., Rojas Bracho, L. and Palumbi, S.R. 2009. Mitochondrial and nuclear genetic variation across calving lagoons in eastern North Pacific gray whales (*Eschrichtius robustus*). *J. Hered.* 100: 34–46.

Alter, S.E., Rynes, E. and Palumbi, S.R. 2007. DNA evidence for historic population size and past ecosystem impacts of gray whales. *Proc. Natl. Acad. Sci. USA* 104(38): 15162–67.

Baker, C.S., Palumbi, S.R., Lambertsen, R.H., Weinrich, M.T., Calambokidis, J. and O'Brien, S.J. 1990. Influence of seasonal migration on geographic distribution of mitochondrial DNA haplotypes in humpback whales. *Nature* 344(6263): 238–40.

Barendse, W., Armitage, S.M., Kossarek, L.M., Shalom, A., Kirkpatrick, B.W., Ryan, A.M., Clayton, D., Li, L., Neibergs, H.L., Zhang, N., Grosse, W.M., Weiss, J., Creighton, P., McCarthy, F., Ron, M., Teale, A.J., Fried, R., McGraw, R.A., Moore, S.S., Georges, M., Soller, M., Womack, J.E. and Hetzel, D.J.S. 1994. A genetic linkage map of the bovine genome. *Nat. Genet.* 6: 227–35.

Beerli, P. 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22(3): 341–45.

Beerli, P. and Felsenstein, J. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* 152: 763–73.

Beerli, P. and Felsenstein, J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proc. Natl Acad. Sci. USA* 98(8): 4,563–4,68.

Bérubé, M., Jorgensen, H., McEwing, R. and Palsbøll, P.J. 2000. Polymorphic di-nucleotide microsatellite loci isolated from the humpback whale, *Megaptera novaeangliae*. *Mol. Ecol.* 9: 2,181–2,83.

Best, P.B., Reeb, D., Rew, M.B., Palsbøll, P.J., Schaeff, C. and Brandão, A. 2005. Biopsying southern right whales: their reactions and effects on reproduction. *J. Wildl. Manage.* 69: 1171–80.

Botstein, D., White, R.L., Skolnick, M. and Davis, R.W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314.

Brown, M.W., Kraus, S.D. and Gaskin, D.E. 1991. Reaction of North Atlantic right whales (*Eubalaena glacialis*) to skin biopsy sampling for genetic and pollutant analysis. *Rep. int. Whal. Commn (special issue)* 13: 81–89.

Buchanan, F.C., Friesen, M.K., Littlejohn, R.P. and Clayton, J.W. 1996. Microsatellites from the beluga whale *Delphinapterus leucas*. *Mol. Ecol.* 5: 571–75.

Butterworth, D.S., Korrubel, J.L. and Punt, A.E. 2002. What is needed to make a simple density-dependent response population model consistent with data for the eastern North Pacific gray whales? *J. Cetacean Res. Manage.* 4(1): 63–76.

Calambokidis, J., Darling, J.D., Deeke, V., Gearin, P., Gosho, M., Megill, W., Tombach, C.M., Goley, D., Toropova, C. and Gisborne, B. 2002. Abundance, range and movements of a feeding aggregation of gray whales from California to southeastern Alaska. *J. Cetacean Res. Manage.* 4(3): 267–76.

Calambokidis, J., Laake, J.L. and Klimik, A. 2010. Abundance and population structure of seasonal gray whales in the Pacific Northwest 1978–2008. Paper SC/62/BRG32 presented to the IWC Scientific Committee, June 2010, Agadir, Morocco (unpublished). 50pp. [Paper available from the Office of this Journal].

Darling, J.D. 1984. Gray whales off Vancouver Island, British Columbia. pp.267–87. *In*: Jones, M.L., Swartz, S.L. and Leatherwood, S. (eds). *The Gray Whale,* Eschrichtius robustus. Academic Press, Orlando, Florida. xxiv+600pp.

Daugherty, C.H., Cree, A., Hay, J.M. and Thompson, M.B. 1990. Neglected taxonomy and continuing extinction of tuatara (*Sphenodon*). *Nature* 347: 177–79.

Dizon, A.E., Perrin, W.F., Amos, W., Baker, C.S., Chivers, S.J., Costa, A.S., Curry, B.E., Gaggiotti, O., Hoelzel, A.R., Hofman, R., LeDuc, R.G., Loughlin, T.R., Lux, C.R., O'Corry-Crowe, G.M., Rosel, P.E., Rosenberg, A., Scribner, K.T. and Taylor, B.L. 1997. Report of the Workshop on the Analysis of Genetic Data to Address Problems of Stock Identity as Related to Management of Marine Mammals. pp.3–48. *In*: Dizon, A.E., Chivers, S.J. and Perrin, W.F. (eds). *Molecular Genetics of Marine Mammals*. The Society for Marine Mammalogy, Lawrence, KS.

Faubet, P., Waples, R.S. and Gaggiotti, O.E. 2007. Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. *Mol. Ecol.* 16: 1149–66.

Findley, L.T. and Vidal, O. 2002. The gray whale, *Eschrichtius robustus*, at calving sites in the Gulf of California, México. *J. Cetacean Res. Manage.* 4(1): 27–40.

Frankham, R., Ballou, J.D. and Briscoe, D.A. 2002. *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge, UK. 617pp.

Frasier, T.R., Koroscil, S.M., White, B.N. and Darling, J.D. 2011. Assessment of population substructure in relation to summer feeding ground use in the eastern North Pacific gray whale. *Endanger. Species. Res.* 14: 39–48.

Gilson, A., Sylvanen, M., Levine, K. and Banks, J. 1998. Deer gender determination by polymerase chain reaction: validation study and application to tissues, bloodstains and hair forensic samples from California. *Calif. Fish Game* 84(4): 159–69.

Goerlitz, D.S., Urbán R., J., Rojas-Bracho, L., Belson, M. and Schaeff, C.M. 2003. Mitochondrial DNA variation among Eastern North Pacific gray whales (*Eschrichtius robustus*) on winter breeding grounds in Baja California. *Can. J. Zool.* 8: 1965–72.

Hatler, D.F. and Darling, J.D. 1974. Recent observations of the gray whale in British Columbia. *Can. Field-Nat.* 88: 449–59.

Henderson, D.A. 1984. Nineteenth century gray whaling: grounds, catches and kills, practices and depletion of the whale population. pp.159–86. *In*: Jones, M.L., Swartz, S.L. and Leatherwood, S. (eds). *The Gray Whale*, Eschrictius robustus. Academic Press, Inc., Orlando, Florida. xxiv+600pp.

Hey, J. and Nielsen, R. 2004. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc. Natl Acad. Sci. USA* 104: 2785–90.

Hochberg, Y. 1988. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 75: 800–02.

Hoelzel, A.R. 1998. Genetic structure of cetacean populations in sympatry, parapatry, and mixed assemblages: implications for conservation policy. *J. Hered.* 89: 451–58.

Hubisz, M.J., Falush, D., Stephens, M. and Pritchard, J.K. 2009. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol.* 9: 1322–32.

Huelsenbeck, J.P. and Andolfatto, P. 2007. Inference of population structure under a Dirichlet process model. *Genetics* 175: 1787–802.

Irwin, D.L., Mitchelson, K.R. and Findlay, I. 2003. PCR product cleanup methods for capillary electrophoresis. *BioTech.* 34: 932–36.

International Whaling Commission. 2011. Report of the Scientific Committee. Annex E. Report of the Standing Working Group on the Aboriginal Whaling Management Procedure (AWMP). *J. Cetacean Res. Manage. (Suppl.)* 12:143–67.

International Whaling Commission. 2013a. Report of the AWMP Workshop Focusing on the PCFG Gray Whale *Implementation Review*. *J. Cetacean Res. Manage. (Suppl.)* 14:371–84.

International Whaling Commission. 2013b. Report of the Scientific Committee. Annex E. Report of the Standing Working Group on the Aboriginal Whaling Management Procedure (AWMP). *J. Cetacean Res. Manage. (Suppl.)* 14:137–71.

Jones, M.L. 1990. The reproductive cycle in gray whales based on photographic resightings of females in the breeding grounds from 1977–82. *Rep. int. Whal. Commn (special issue)* 12: 177–82.

Jones, M.L. and Swartz, S.L. 1984. Demography and phenology of gray whales and evaluation of whale-watching activities in Laguna San Ignacio, Baja California Sur, Mexico. pp.309–74. *In*: Jones, M.L., Swartz, S.L. and Leatherwood, S. (eds). *The Gray Whale,* Eschrichtius robustus. Academic Press, Inc., Orlando, Florida. xxiv+600pp.

Laake, J.L., Punt, A.E., Hobbs, R., Ferguson, M., Rugh, D. and Breiwick, J. 2012. Gray whale southbound migration surveys 1967–2006: an integrated re-analysis. *J. Cetacean Res. Manage* 12(3): 287–306.

Lambertsen, R.H. 1987. A biopsy system for large whales and its use for cytogenetics. *J. Mammal.* 68(2): 443–45.

Malik, S., Brown, M.W., Kraus, S.D., Knowlton, A.R., Hamilton, P.K. and White, B.N. 1999. Assessment of mitochondrial DNA structuring and nursery use in the North Atlantic right whale (*Eubalaena glacialis*). *Can. J. Zool.* 77(81): 1217–22.

Marshall, T.C., Slate, J., Kruuk, L.E.B. and Pemberton, J.M. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* 7(5): 639–55.

Michaud, C.L. and Foran, D.R. 2011. Simplified field preservation of tissues for subsequent DNA analyses. *J. Forensic. Sci.* 56: 846–52.

Moore, S., Urbán R, J., Perryman, W., Gulland, F., Perez-Cortes, H., Rojas-Bracho, L. and Rowles, T. 2001. Are gray whales hitting 'K' hard? *Mar. Mammal Sci.* 17(4): 954–58.

Moore, S.E. and Ljungblad, D.K. 1984. Gray whales in the Beaufort, Chukchi, and Bering Seas: distribution and sound production. pp.543–59. *In*: Jones, M.L., Swartz, S.L. and Leatherwood, S. (eds). *The Gray Whale,* Eschrichtius robustus. Academic Press Inc., Orlando, Florida. xxiv+600pp.

Morin, P.A., Martien, K.K., Archer, F.I., Cipriano, F., Steel, D., Jackson, J. and Taylor, B.L. 2010. Applied conservation genetics and the need for quality control and reporting of genetic data used in fisheries and wildlife management. *J. Heredity* 101(1): 1–10.

Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York. x+512pp.

Palsbøll, P.J., Bérubé, M. and Allendorf, F.W. 2006. Identification of management units using population genetic data. *Trends Ecol. Evol.* 22(1): 11–16.

Palsbøll, P.J., Bérubé, M., Larsen, A.H. and Jorgensen, H. 1997. Primers for the amplification of tri- and tetramer microsatellite loci in baleen whales. *Mol. Ecol.* 6: 893–95.

Palsbøll, P.J., Larsen, F. and Hansen, E.S. 1991. Sampling of skin biopsies from free-ranging large cetaceans in West Greenland: development of new biopsy tips and bolt designs. *Rep. int. Whal. Commn (special issue)* 13: 71–79.

Pella, J. and Masuda, M. 2006. The Gibbs and split-merge sampler for population analysis from genetic data with incomplete baselines. *Can. J. Fish. Aquat. Sci.* 63: 576–96.

Pike, G.C. 1962. Migration and feeding of the gray whale (*Eschrichtius gibbosus*). *J. Fish. Res. Bd Can.* 19(5): 815–38.

Pritchard, J.K., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–59.

Reilly, S.B. 1992. Population biology and status of eastern Pacific gray whales: recent developments. pp.1,062–74. *In*: McCullough, D.R. and Barrett, R.H. (eds). *Wildlife 2001: Populations*. Elsevier Applied Science Publishers, London. xv+1,163pp.

Rice, D.W. and Wolman, A.A. 1971. *The Life History and Ecology of the Gray Whale* (Eschrichtius robustus). American Society of Mammalogists, Special Publication No. 3, Stillwater, Oklahoma. viii+142pp.

Richard, K.R., Whitehead, H. and Wright, J.M. 1996. Polymorphic microsatellites from sperm whales and their use in the genetic identification of individuals from naturally sloughed pieces of skin. *Mol. Ecol.* 5: 313–15.

Rooney, A.P., Merritt, D.B. and Derr, J.N. 1999. Microsatellite diversity in captive bottlenose dolphins (*Tursiops truncatus*). *J. Hered.* 90: 228–31.

Rousset, F. 2008. GENEPOP'07: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol. Ecol. Notes* 8: 103–06.

Rugh, D.J., Hobbs, R.C., Lerczak, J.A. and Breiwick, J.M. 2005. Estimates of abundance of the Eastern North Pacific stock of gray whales 1997 to 2002. *J. Cetacean Res. Manage.* 7(1): 1–12.

Ryman, N. and Palm, S. 2006. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. *Mol. Ecol. Notes* 6: 600–02.

Seutin, S., White, B.N. and Boag, P.T. 1991. Preservation of avian blood and tissue samples for DNA analysis. *Can. J. Zool.* 69: 82–90.

Shelden, K.E.W. and Laake, J.L. 2002. Comparison of the offshore distribution of southbound migratory gray whales from aerial survey data collected off Granite Canyon, California, 1979–1996. *J. Cetacean Res. Manage.* 4(1): 53–56.

Swartz, S.L. 1986. Gray whale migratory, social and breeding behavior. *Rep. int. Whal. Commn (special issue)* 8: 207–29.

Swartz, S.L., Taylor, B.L. and Rugh, D.J. 2006. Gray whale *Eschrichtius robustus* population and stock identity. *Mammal Rev.* 36: 66–84.

Taylor, B.L. 2005. Identifying units to conserve. pp.149–64. *In*: Reynolds III, J.E., Perrin, W.F., Reeves, R.R., Montgomery, S. and Regan, T.J. (eds). *Marine Mammal Research: Conservation Beyond Crisis*. The Johns Hopkins University Press, Baltimore, Maryland.

Urban R, J., Rojas-Bracho, L., Perez-Cortes, H., Gomez-Gallardo, A., Swartz, S.L., Ludwig, S. and Brownell, R.L. 2003. A review of gray whales *Eschrichtius robustus* on their wintering grounds in Mexican waters. *J. Cetacean Res. Manage.* 5(3): 281–95.

Valsecchi, E. and Amos, W. 1996. Microsatellite markers for the study of cetacean populations. *Mol. Ecol.* 5: 151–56.

Wade, P.R. 2002. A Bayesian stock assessment of the eastern Pacific gray whale using abundance and harvest data from 1967–1996. *J. Cetacean Res. Manage.* 4(1): 85–98.

Waldick, R.C., Brown, M.W. and White, B.N. 1999. Characterization and isolation of microsatellite loci from the endangered North Atlantic right whale. *Mol. Ecol.* 8: 1763–65.

Wang, J.Y. 2009. Stock identity. pp.1115–18. *In*: Perrin, W.F., Würsig, B. and Thewissen, J.G.M. (eds). *Encyclopedia of Marine Mammals*. Academic Press, San Diego, California.

Wang, J.Y., Frasier, T.R., Yang, S.C. and White, B.N. 2008. Detecting recent speciation events: the case of the finless porpoise (genus *Neophocaena*). *Heredity* 101: 145–55.

Waples, R.S. and Gaggiotti, O. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol. Ecol.* 15(6): 1419–39.

Wilson, G.A. and Rannala, B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163(3): 1177–91.