Marine Mammal Science

Population structure of North Pacific gray whales in light of trans-Pacific movements

Journal:	*Marine Mammal Science*
Manuscript ID	MMSCI-5048.R2
Manuscript Type:	Article
Date Submitted by the Author:	26-Jul-2021
Complete List of Authors:	Lang, Aimee; Ocean Associates, Inc., On contract to the Marine Mammal and Turtle Division, Southwest Fisheries Science Center Weller, David; NOAA Fisheries, Southwest Fisheries Science Center Burdin, Alexandr; Kamchatka Branch of Pacific Institute of Geography of Russian Academy of Sciences Robertson, Kelly; National Marine Fisheries Service, Southwest Fisheries Science Center Sychenko, Olga; CetaEcoResearch Urban R., Jorge; Universidad Autónoma de Baja California Sur, Departamento de Ciencias Marinas y Costeras Martínez Aguilar, Sergio; Universidad Autónoma de Baja California Sur, Departamento de Ciencias Marinas y Costeras Pease, Victoria; Southwest Fisheries Science Center LeDuc, Richard; NOAA SWFSC, Marine Mammal & Turtle Division Litovka, Dennis; Chukotka Governor and Government Office Burkanov, Vladimir; Alaska Fisheries Science Center, National Marine Mammal Laboratory; (2) Kamchatka Branch of the Pacific Institute of Geography Russian Academy of Sciences, Brownell Jr., Robert; NOAA/NMFS/Southwest Fisheries Science Center, 34500 Highway 1
Keywords:	Eschrichtius robustus, Okhotsk Sea, microsatellites, mitochondrial DNA, population structure

Received: 28 August 2020 | Accepted: 27 July 2021

Running head: LANG et al.

ARTICLE

Population structure of North Pacific gray whales in light of trans-Pacific movements

Aimée R. Lang[1] | David W. Weller[2] | Alexander M. Burdin[3] | Kelly Robertson[2] | Olga Sychenko[4] | Jorge Urbán R.[5] | Sergio Martínez-Aguilar[5] | Victoria L. Pease[2] | Richard G. LeDuc[2] | Dennis I. Litovka[6] | Vladimir N. Burkanov[3,7] | Robert L. Brownell, Jr.[2]

[1] Ocean Associates, Inc., Arlington, VA; on contract to the Marine Mammal & Turtle Division, Southwest Fisheries Science Center, NOAA-Fisheries, La Jolla, CA

[2] Marine Mammal & Turtle Division, Southwest Fisheries Science Center, NOAA-Fisheries, La Jolla, CA

[3] Kamchatka Branch of the Pacific Geographical Institute, Russian Academy of Sciences, Petropavlovsk-Kamchatsky, Kamchatka, Russia

[4] CetaEcoResearch, Lacey, WA

[5] Departamento de Ciencias Marinas y Costeras, Universidad Autónoma de Baja California Sur, La Paz, B.C.S., Mexico

[6] Chukotka Governor and Government Office, Anadyr, Chukotka, Russia

[7] Marine Mammal Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanographic and

Atmospheric Administration, Seattle, WA

Correspondence

Aimée R. Lang, Southwest Fisheries Science Center, 8901 La Jolla Shores Drive, La Jolla, CA 92037.

Email: aimee.lang@noaa.gov

Abstract

Recent findings that some gray whales that feed off Sakhalin Island (SI), Russia, in the western North Pacific (WNP) overwinter in the eastern North Pacific (ENP) indicate that population structure in this species is more complex than originally thought. We generated mitochondrial DNA (mtDNA) control region sequences and microsatellite genotypes (n = 12 loci) from 156 whales sampled off SI and compared them to available data from 106 ENP whales. Significant mitochondrial and nuclear genetic differentiation between the SI and ENP whales was found. Genetic cluster analysis identified two groups among the SI whales, one of which was genetically similar to ENP whales. Photographs collected from the biopsied SI whales showed that both groups comprised whales known to migrate to the ENP, suggesting that the clustering pattern was not reflective of some SI whales interbreeding while overwintering in the WNP. Instead, the genetic differentiation observed between the SI and ENP whales may be due to assortative mating of SI whales while west of eastern migratory routes. The rare but continued reports of gray whales off the coasts of Japan and China, however, confirm that some gray whales overwinter in the WNP and highlight the need to collect additional data from these whales.

KEYWORDS

Eschrichtius robustus, gray whale, microsatellites, mitochondrial DNA, population structure

1 | INTRODUCTION

In the 20th century alone, almost three million whales were killed as part of commercial whaling (Rocha et al., 2014), with some populations reduced to 1% or less of their estimated prewhaling abundance (e.g., the Antarctic blue whale; Branch et al., 2007). Today, the status of baleen whale populations varies widely. While some populations, such as the eastern North Pacific right whale (*Eubalaena japonica*), may number only in the tens of animals (Wade et al., 2011), others, including humpback whales (*Megaptera novaeangliae*) in many parts of their Southern Hemisphere range (e.g., eastern Australia; Noad et al., 2019), are considered to be at or near preexploitation numbers (Thomas et al., 2016). Monitoring of these populations has provided some of the first opportunities to learn about patterns of increase and recovery in depleted large whale populations.

In 1994, the eastern North Pacific (ENP) gray whale (*Eschrichtius robustus*) population became the first baleen whale population to be removed from the U.S. List of Endangered and Threatened Wildlife and Plants (U.S. Federal Register, 1994). During summer and early fall, most of the whales in this population feed in the Bering, Beaufort, and Chukchi Seas (hereafter referred to as the Northern feeding ground, or NFG). A small number of ENP gray whales (<250), identified as the

Pacific Coast Feeding Group (PCFG), show fidelity to feeding grounds located farther south between Northern California and southeastern Alaska (Calambokidis et al., 2002, 2017; Darling, 1984; Gilmore, 1960; Hatler & Darling, 1974; Pike, 1962). Whales from both feeding grounds migrate south along the west coast of North America to wintering areas in the lagoons and coastal waters off Baja California, Mexico (Rice & Wolman, 1971). Like many other baleen whale populations, ENP gray whales were decimated by commercial whaling, primarily during the 19th and early 20th centuries (Henderson, 1984; Reeves, 1984). Following protection from commercial whaling in 1938, the ENP population began to recover, and shore-based counts of migrating gray whales off California indicated that the population numbered ~21,000 whales by the late 1980s (Buckland et al., 1993). The most recent estimate of abundance for this population is ~27,000 whales (Durban et al., 2015, 2017).

In the western North Pacific (WNP), gray whales historically fed in the Okhotsk Sea and used the coastal waters of southeastern Russia, the Korean Peninsula, and Japan to migrate to WNP wintering ground(s) thought to be in the coastal waters of southern China (reviewed in Weller et al., 2002b). Modern commercial whaling off Korea by Japan began around 1900, peaked in the 1910s and then declined rapidly in the 1920s and

1930s (Kato & Kasuya, 2002), at which point some considered the population to have been extirpated (Bowen, 1974; Mizue, 1951). In the late 1970s, however, Brownell and Chun (1977) proposed that a small relict population of western gray whales remained extant based on: (1) catch records showing that gray whales were taken in Korean waters through at least 1966, (2) the sighting of a female with a calf in Korean waters in 1968, and (3) sporadic sightings of small numbers of whales in the Okhotsk Sea between the late 1960s and 1970s.

Aerial and vessel surveys conducted in the Okhotsk Sea during summer and fall months between 1979 and 1989 identified aggregations of feeding gray whales off the northeastern coast of Sakhalin Island (SI), Russia (Berzin, 1990; Berzin et al., 1988, 1991; Blokhin, 1996; Blokhin et al., 1985), an area which, by the mid-1990s, had become a site for extensive offshore oil and gas development. In 1995, concern for the potential impact of such activities on gray whales feeding off SI led to the initiation of a long-term monitoring effort by Russian and U.S. scientists (Weller et al., 1999). Photo-identification data collected as part of these efforts, which have been conducted annually since 1997 and continue to date as part of the Russian Gray Whale Project. Also, data collected by industry-funded efforts since 2002, indicate that a small number of whales

demonstrate strong fidelity to this feeding area both within and between years (Bradford, 2011; Bradford et al., 2008; Bröker et al., 2020; Burdin et al., 2020; Weller et al., 1999, 2002b). Use of the SI feeding ground is driven in part by matrilineal fidelity, as many of the whales first identified as calves with their mothers have been sighted on the SI feeding ground in subsequent seasons. Consistent with these observations, a model-based assessment shows that, at least in recent years, the group of whales feeding off SI is demographically self-contained, such that most or all of the animals recruited into the population are the calves of SI mothers (Cooke, 2017; Cooke et al., 2013). While the number of whales using the area was estimated to be <100 in the early years of the study (Bradford et al., 2008; Cooke, 2018), recent assessment indicated that the number of whales feeding off SI numbered ~191 whales of age one or older in 2018 and has been growing at about 3.4%–4.8% over the past 20 years (Cooke et al., 2019). Although WNP gray whales were initially listed as a Critically Endangered subpopulation by the IUCN (Reilly et al., 2008), that designation was recently revised to Endangered given evidence that the number of mature individuals now exceeds 50 (Cooke et al., 2018).

Genetic analyses based on maternally inherited mitochondrial DNA (mtDNA), support recognition of ENP and WNP

gray whales as separate stocks. Significant mtDNA genetic differentiation was identified between whales sampled in the ENP (*n* = 120), primarily along the migratory route, and whales (*n* = 45) sampled off SI (LeDuc et al., 2002), and haplotypic diversity was markedly reduced among the SI whales, consistent with expectations for a small remnant population. While the LeDuc et al. (2002) study was based on analysis of the mtDNA control region alone (~520 bps of sequence), Meschersky et al. (2012) analyzed additional regions of the mitochondrion (cytochrome B, ND2, and the control region, totaling >2,700 bps) and found similar results, with genetic differentiation observed when the whales sampled off SI were compared to NFG whales sampled in the Bering Sea off the Chukotka and Koryak coasts of Russia (Meschersky et al., 2015). All the mitochondrial haplotypes identified among Sakhalin whales sampled in this study were also found among samples collected from whales off Chukotka, Russia. A subsequent study by Brüniche-Olsen et al. (2020), in which full mitogenomes were sequenced, found that four of the nine mitogenome haplotypes found among Sakhalin whales were also identified among whales sampled in Mexico. Similar to the earlier studies, no phylogeographic structure was apparent when the relationships among haplotypes were examined, which the authors interpreted as suggesting either female-

mediated gene flow or recent divergence.

Initially, WNP whales were presumed to feed only off the coast of SI, in a nearshore and an offshore area (Bröker et al., 2020; Meier et al., 2007; Weller et al., 2002a, 2003). However, photo-identification studies off southern and eastern Kamchatka indicated that approximately half of the whales identified in that region are whales known to use the SI feeding ground (Burdin et al., 2011; Tyurneva et al., 2010; Vertyankin et al., 2004). The population affiliation of the whales off Kamchatka that were not observed off SI is unknown; these whales could be whales that use the SI feeding ground but have not been observed there or they could be whales of ENP origin.

While the photo-identification work conducted off the southeastern coast of Kamchatka raised the possibility that ENP and WNP gray whales may mix in this area, until recently it was presumed that the ENP and WNP populations remained largely reproductively isolated based on the presumed use of separate migratory routes and wintering grounds on each side of the North Pacific. Recent results from tagging, photo-identification, and genetic studies have, however, changed the scientific perspective about this premise. In 2010 and 2011, three whales were satellite-tagged off SI, and all three migrated toward the eastern North Pacific, with one animal retaining its tag for a

full migratory cycle between SI, east and south to Mexico, and then back to SI (Mate et al., 2015). When combined with subsequent comparisons of whales photographed off SI with those photographed on eastern migratory routes (Weller et al., 2012) and in the Mexican lagoons (summarized in Urbán R. et al., 2019), as well as with genetic recaptures of whales sampled off SI and those sampled in the ENP (Lang, 2010), a total of 53 whales have now been identified both in the WNP off SI and in the ENP.

While these movements indicate that a proportion of the animals feeding off SI have, at least once, migrated to the ENP, a model-based assessment suggests that 20%–55% of SI whales do not overwinter in the ENP wintering grounds off Mexico (Cooke et al., 2019). In addition, a small number of records of gray whales off the coast of Japan (n = 22; Nakamura et al., 2018) and China (n = 2; Wang et al., 2015; Zhao, 1997; Zhu, 2012; Zhu & Yue, 1998) have been reported over the last two decades. The majority of the records from Japanese waters are from the months of March to May, when ENP whales are known to be migrating north up the west coast of North America. Although little is known about the population identity of the Japan sightings, high quality photographs of two individuals identified them as whales first sighted as calves on the SI feeding ground. One of these

whales was entrapped and died in a set net off the Pacific coast of Honshu in January 2007 (Weller et al., 2008), while the other whale was recorded as a calf off SI in August 2014, sighted as a yearling off the Pacific coast of Japan in March 2015, returned to feed off SI in August 2015, and was resighted off Japan in January and February 2016 (Weller et al., 2016).

The trans-Pacific movements documented for some of the SI whales raises the question of whether gene flow is occurring with whales that feed on the NFG and/or the PCFG. Brüniche-Olsen et al. (2018) genotyped single nucleotide polymorphisms (n = 84 autosomal SNPs) in 55 individuals sampled off SI and 111 individuals sampled in the Mexican wintering lagoons. Comparison of SNP allele frequencies between these two groups revealed genetic differentiation (F_{ST} = 0.039, p < .001), indicating that mating between the two groups was not random. Cluster analyses supported the presence of two stocks of gray whales in the NP, with some of the whales sampled in the WNP grouped with the cluster comprised primarily of ENP whales, which is consistent with what would be expected if some SI whales breed in the WNP and others in the ENP. Subsequently, four regions of the mtDNA (~3,500 bps, including the control region and three protein coding regions) were sequenced in these same Sakhalin whales (Brykov et al., 2019). When the sequences of the three protein

coding regions were concatenated, three of the 11 haplotypes found among the Sakhalin whales that clustered in the WNP group were also found among the Sakhalin whales that clustered in the primarily ENP group, and when the relationships among these haplotypes were visualized in a haplotype network no phylogeographic pattern was apparent. The authors noted that this pattern is not consistent with long-term isolation during the Pleistocene and that it indicated that the two groups may have diverged recently, possibly since the end of commercial whaling in the early 20th century or, more likely, post-Pleistocene.

Individuals that appeared to be admixed in the Brüniche-Olsen et al. (2018) study were present in both ENP and WNP, suggesting that gene flow between the two stocks is occurring. However, gray whales are thought to mate primarily while on their southbound migration (Rice & Wolman, 1971). If the two genetic clusters detected among the SI whales sampled in the Brüniche-Olsen et al. (2018) study represent some whales that remain in the WNP year-round and other whales that overwinter in the ENP, it is not clear where breeding between these two groups would occur. In addition, finding that some of the whales sampled in the Mexican lagoons appeared to be admixed was surprising given that Sakhalin whales and any of their offspring

should comprise only a very small proportion of the whales that utilize that wintering ground.

Here we generate microsatellite genotypes and mtDNA control region sequences from samples collected from whales feeding off SI. We combine this data set with previously published data generated from NFG whales (Lang et al., 2014) and conduct a suite of genetic analyses aimed at providing further insight into the nature and extent of connectivity between whales using the primary ENP feeding ground and those feeding in the WNP. Importantly, all the SI samples analyzed here are linked to photographically identified individuals (Burdin et al., 2019; Weller et al., 1999, 2002b), allowing any patterns revealed in the genetic data to be interpreted in the context of what is known about the trans-Pacific movements documented for some Sakhalin whales. Finally, we analyze samples (*n* = 24 samples representing 16 individuals) collected from whales feeding off the southeastern coast of Kamchatka to evaluate their genetic similarity to the Sakhalin and NFG samples.

2 | METHODS

2.1 | Sample collection

Skin samples (*n* = 198) were collected via biopsy darting of whales on the feeding grounds off the northeastern coast of SI, Russia. All samples were collected between July and September,

and all except one are linked to a photographically identified whale. The majority of the samples (*n* = 175) were collected between 1995 and 2007, and these samples represent 84.6% of all whales (*n* = 169 individuals) identified on the Sakhalin feeding ground during this period. Following a gap in sampling effort, 23 biopsies were collected from the SI feeding ground during the summer of 2010 and 2011. In addition, 24 samples were collected via biopsy darting of whales off the southeastern coast of Kamchatka during the summer months (June–August) of 2004 and 2010–2011. The locations where samples were collected are shown in Figure 1, and sample details are provided in Table S1.

2.2 | DNA extraction, mtDNA sequencing, and genetic sex determination

DNA extraction and mtDNA control region sequencing of the samples (*n* = 45) collected from the SI feeding ground between 1995 and 1999 are described in LeDuc et al. (2002). Where needed, DNA from these samples was reextracted as described below.

A variety of common extraction methods were used to extract genomic DNA from the tissue samples: (1) standard phenol/chloroform extractions (modified from Sambrook et al., 1989), (2) sodium chloride protein precipitation (Miller et al., 1988), and (3) silica-based filter purification (Qiagen). A 523-

bp segment of the 5′ end of the hyper-variable mitochondrial DNA (mtDNA) control region was amplified from the extracted DNA using the polymerase chain reaction (PCR) and the primers H00034 (Rosel et al., 1995) and L15812 (Chivers et al., 2005). The reaction was carried out in a 25 µl final volume using the protocol described in Lang et al. (2014). Standard techniques (Palumbi et al., 1991; Saiki et al., 1988) were used to sequence both strands of the amplified DNA on an Applied Biosystems (ABI) (Waltham, MA) model 377, 3100, or 3730 Sequencer. Sequences were aligned in Geneious 11.1.3 (Kearse et al., 2012) using the MAFFT algorithm (v7.388, Katoh et al., 2002, Katoh & Standley, 2013).

The sex of each sample was determined via amplification and Real-Time PCR (MX3000p; Stratagene, Inc., San Diego, CA) of the zinc finger genes (ZFX and ZFY) as described in Morin et al. (2005). Samples from one male and one female for which sex had been determined via physical examination of a stranded animal were included as positive controls. Two amplification products were obtained in males, and a single product identified females.

2.3 | Microsatellite genotyping

Twelve previously published polymorphic microsatellite loci isolated from other cetacean species were used to genotype all samples (Table S2): EV14, EV37, and EV94 (Valsecchi & Amos, 1996); Gata028, Gata098, Gata417, and Gt023 (Palsbøll et al.,

1997); RW31 and RW48 (Waldick et al., 1999); and SW10, SW13, and SW19 (Richard et al., 1996). Forward primers were fluorescently labeled, and, with the exception of loci Gata098 and EV37, reverse primers were modified and tailed (Brownstein et al., 1996) to reduce allelic stutter partway through the study (described in Table S2). Extracted DNA was amplified in a 25 µl final reaction volume containing ~100 ng of DNA, 1 × PCR buffer (50 mM KCl, 10 mM Tris-HCl, ph8.3 and 1.5 mM $MgCl_2$), 0.6 mM dNTPs, 0.3 µM primers, and 0.5 units of Taq DNA polymerase (New England BioLabs, Inc., Ipswich, MA). The PCR cycling profile included 90°C for 2.5 min, followed by 35 cycles of 94°C for 45 s, 1 min at the optimal annealing temperature (Table S2), and 72°C for 1.5 min, followed by a final extension of 72°C for 5min. Only one locus was amplified per reaction, and each product was assessed electrophoretically on a 2% agarose gel for size and quality before genotyping. Sakhalin samples collected prior to 2002 were genotyped at the "original" six loci using an ABI 377 Genetic Analyzer (Table S2), while genotyping of the remaining samples was conducted using an ABI 3100 or 3730 Genetic Analyzer. GeneMapper v.4.0 (Applied Biosystems) was used along with an internal size standard (GeneScan-500 ROX, ABI) to determine allele sizes.

Data generated from the SI and Kamchatka samples were

combined with previously published data (mtDNA control region sequences, sex, and genotypes at the same 12 microsatellite loci) generated from 106 samples collected from whales north of the Aleutian Islands during summer and fall (Lang et al., 2014). This sample set included whales killed by native hunters in the Chukotka Peninsula region of Russia (n = 71); whales biopsied off Utqiagvik, Alaska (n = 12) and between Cape Navarin and Cape Olyutorskiy, Russia, south of the Koryak Mountains (n = 17); and stranded whales from the Bering Sea (n = 4) and Utqiavik, Alaska (n = 2) (Figure 1). This region represents the NFG that is used by the majority of the ENP whales. Protocols and quality control measures used in generating these data are described by Lang et al. (2014). All data were produced in the same lab (SWFSC) and the NFG data were produced in tandem (i.e., used identical protocols and equipment) with the data generated from all Kamchatka samples and the SI samples collected in 2010 and 2011. Thus, the calibration of data sets used to ensure consistency between the earlier (1995–2007) and more recent (2010–2011) SI data (described below) also ensured consistency with the NFG data set.

Extensive quality control/quality analysis was conducted to ensure that no biases were introduced for genotypes generated across different ABI instruments and, in some cases, with tailed

and untailed primer sets. Allele binning was manually checked after scoring loci, and normalization of allele sizes across different platforms or primer sets (tailed/untailed) was conducted manually and then visualized graphically. Over 20% of the samples that were genotyped on the ABI 377 were regenotyped on the ABI 3100 or 3730 instruments to ensure consistency across data sets. Following recommended quality control guidelines described in Morin et al. (2010), a subset of samples (comprising ≥10% of all individuals genotyped) were rerun at random and used to estimate per-allele error rates.

2.4 | Data analysis

2.4.1 | Sample stratification

Samples were grouped based on the geographic sampling location (Figure 1), representing the three feeding grounds (SI, Kamchatka, and NFG) for analyses requiring a priori stratification. For some analyses, the SI and Kamchatka samples were combined to form a WNP regional stratum. For individuals that were sampled off both SI and Kamchatka, the sample collected off Kamchatka was removed from this combined WNP regional stratum.

In addition, for some analyses (measures of genetic diversity and genetic differentiation between strata) the SI stratum was subdivided into those whales that were first

identified as calves on the SI feeding ground versus whales that were first identified as noncalves. The SI stratum was also subdivided into those whales that have been identified in both the ENP and SI (SI-ENP) and those whales recorded off SI but not in the ENP (SI only). Identification of whales as SI-ENP was based on Urbán R. et al. (2019), which provides the most recent summary of whales known to transit between SI and the ENP based on photographic comparisons, genetic matching, and satellite tagging studies. The probability of photographically identifying a Sakhalin whale amongst the much larger ENP population is small, and thus it is likely that at least some of the SI only whales also visit the ENP.

2.4.2 | Genetic diversity

For the mtDNA sequence data, jModelTest 2.1.10 (Darriba et al., 2012; Posada 2008) was used to identify the appropriate model of nucleotide substitution for the data. Haplotypic (*h*) and nucleotide (π) diversity for each stratum were calculated using the R package *strataG* (Archer et al., 2017). The software PopArt (Leigh & Bryant, 2015) was used with the default parameters to generate a median joining tree showing the relationship between haplotypes.

The R package *genepop* (Rousset, 2008) was used to test the microsatellite loci for deviations from Hardy-Weinberg

equilibrium (HWE) using both the probability test (Guo & Thompson, 1992) and the test for heterozygote deficiency (Rousset & Raymond, 1995) with the default values for the Markov chain parameters (10,000 dememorization steps, 20 batches, 5,000 iterations/batch). The *genepop* package was also used to test for linkage disequilibrium (LD) for each pair of loci and for the presence of null alleles. The R package *strataG* (Archer et al., 2017) was used to calculate the proportion of homozygous loci per individual and identify samples with unusual (outlier) levels of homozygosity. *StrataG* was also used to perform a jackknife analysis to assess the effect of individual samples on significant deviations from Hardy Weinberg equilibrium, which allows rare homozygous genotypes and influential samples to be identified and rechecked.

The program GenAlEx v6.5 (Peakall & Smouse, 2012) was used to calculate the probability of identity, defined as the probability that two randomly chosen individuals would share the same multi-locus genotype under both the assumption of Hardy-Weinberg equilibrium (PID_{HW}; Paetkau & Strobeck, 1994) and under the more conservative assumption that full siblings might be present in the data set (PID_{SIB}; Waits et al., 2001). *StrataG* (Archer et al., 2017) was used to identify samples with genotypes that matched at 80% or more of the loci and were thus

likely to represent resampling of the same individual. The mtDNA haplotype and sex of identified matches were compared to confirm or dispute putative duplicate individuals. For samples that were not perfect matches, genotypes were rechecked for scoring errors and regenotyped, if necessary, to resolve mismatches. These regenotyped samples provided additional replicates with which to ensure consistency across the data set.

The R package *diveRsity* (Keenan et al., 2013) was used to generate measures of diversity for the microsatellite data set, including the number of alleles per locus, allelic richness, observed and expected heterozygosity, and F_{IS}. Private alleles were identified in the R package *strataG* (Archer et al., 2017). The R package *related* (Pew et al., 2015) was used to estimate relatedness between pairs of individuals based on the allele frequencies of the combined data set. Although relatedness coefficients were calculated based on all seven estimators included in related, the results reported here are based on the Queller and Goodnight (1989) estimator.

2.4.3 | Population structure analyses

Pairwise estimates of genetic divergence between strata were generated using the *strataG* package (Archer et al., 2017). For mtDNA control region sequences, F_{ST}, Φ_{ST}, and χ^2 were calculated. For microsatellite genotypes, F_{ST} (Weir & Cockerham, 1984),

normalized F_{ST} (F'_{ST}) (Hedrick, 2005), and a χ^2 test were used to assess genetic differentiation. For both data sets, *p*-values were computed using 5,000 permutations of each data set. Sex differences in genetic differentiation were assessed by subdividing each stratum by sex. For the Sakhalin samples, these sex-specific analyses excluded whales first identified as calves with their mothers, which may not have been reproductively mature for all or part of the study.

Isolation by distance (IBD), in which individuals that are geographically close tend to be genetically more similar than individuals that are far apart due to limited dispersal, can lead to the false detection of population structure and may influence Bayesian clustering algorithms (Frantz et al., 2009; Meirmans, 2012; Schwartz & McKelvey, 2009). To explore whether IBD was present within our data set, we subdivided the NFG samples by sampling location, such that three smaller regions were represented: Utqiagvik, Alaska (n = 14); the Chukotka Peninsula (n = 71), Russia; and the Koryak region of Russia (n = 17). Four additional samples were collected as strandings from isolated regions of the Bering Sea and were not included in this analysis. Matrices of pairwise genetic differentiation (mtDNA F_{ST}, microsatellite F_{ST}) between these six strata were generated in strataG (Archer et al., 2017). Geographic distance between

sites was measured in Google Earth by generating paths closely following the coastline. A Mantel test, which tests the null hypothesis that the genetic distances are not linearly correlated with the geographic distances, was performed in R using the *ecodist* package (Goslee & Urban, 2020).

Two methods were used to evaluate population structure among gray whale feeding grounds without a priori assumptions. First, we used the Bayesian model-based clustering approach implemented in STRUCTURE v2.3.4 (Pritchard et al., 2000) to estimate the number of genetic clusters present in our data. STRUCTURE uses a Bayesian algorithm to cluster individuals into groups based on genetic similarity, such that the identified groups are in Hardy Weinberg and linkage equilibrium. We used a model of admixture with correlated allele and did not include information on the geographic location of sampling. We first analyzed the full data set containing all genotyped individuals and then analyzed the NFG, SI feeding ground, and SI+Kamchatka data sets separately. Values of K ranging from one to eight were tested. In all cases, five independent runs at each K used to check for consistency among runs. A burn-in length of 100,000 followed by 500,000 MCMC iterations was used. All other parameters were left at the default values of the program.

The results of each run for a given K were summarized in

STRUCTURE HARVESTER v0.6.94 (Earl, 2012). The most likely number of clusters (K) was evaluated using both the maximum estimated mean log likelihood of the data, LnP(D), and the ad hoc statistic ΔK (Evanno et al., 2005), which estimates the rate of change in the log probability of data between successive K values but which does not allow assessment of K is 1. The optimal value of K is identified as the K at which a sharp drop in the likelihood value occurs or a peak in ΔK.

Uneven sampling of strata can result in STRUCTURE wrongly inferring the number of genetic clusters present in a data set (Fogelqvist et al., 2010; Puechmaille, 2016). To evaluate this effect, we repeated the STRUCTURE analysis using ten data sets comprised of all individuals from the NFG stratum (n = 105) and a subset of 105 randomly chosen individuals from the SI stratum. All other run parameters were identical to those used in the analysis of the full data set.

Our data set included samples from 69 whales first identified as calves (initially via photographs and confirmed genetically) for which the mother had also been sampled. Only four mother-calf pairs were sampled together, indicating that in most cases our sampling was random (i.e., the probability of sampling the calf was independent of the probability of sampling the mother). However, both simulation-based and empirical

studies have shown that STRUCTURE may overestimate the number of clusters present in a data set when a large number of kin are sampled (Anderson & Dunham, 2008; Rodríguez-Ramilo & Wang, 2012). Although debate exists over when and whether it is appropriate to purge related individuals from genetic data sets (Wang, 2018; Waples & Anderson, 2017), we reran the STRUCTURE analysis as outlined above retaining all known mothers from mother-calf sampled pairs, but removing all the calves from the Sakhalin data set.

To further explore population genetic structure without model-based assumptions, a discriminant analysis of principle components (DAPC; Jombart et al., 2010) was performed on the microsatellite data set using the R package *adegenet* (Jombart et al., 2008). This method does not make assumptions about the cause of structure (i.e., island model versus IBD), and, unlike other clustering approaches (e.g., STRUCTURE; Pritchard et al., 2000), does not assume that identified clusters are in HWE or gametic disequilibrium. We first ran sequential K-means clustering (the "find.clusters" function) and used the Bayesian information criterion (BIC) to identify the most likely number of clusters in the data in the absence of a priori geographical stratifications. We also ran the DAPC with information on geographical strata (feeding ground location) specified. In both

cases, the number of principal components (PCs) to retain was determined using alpha-score optimization (the "optim.a.score" function). Scatter plots were used to visualize the differences between clusters, with inertial ellipses drawn to encompass 67% of the cloud of points representing each cluster.

2.4.4 | Genetic assignment of individuals

The R package *assignPop* (Chen et al., 2018) was used to (1) assess the accuracy of self-assignment of the SI and NFG samples to their stratum of origin, and (2) to assign whales sampled off Kamchatka to either of the other two strata, based on multilocus microsatellite genotypes. We used the Monte Carlo resampling cross validation procedure (Xu & Liang, 2001) to split the data representing the SI and NFG strata into training and test groups and then test the predictive accuracy of the training data by resampling over 100 iterations. This allowed us to assess the reliability of the "baseline data" to accurately assign individuals to a source population. To avoid sample size biasing the assignment results, the size of the training data sets representing both strata was set to 84, 94, and 100 individuals, which correlated with 80%, 90%, and 95% of the smallest sample size (i.e., n = 105 for the NFG stratum). For each training set, we also assessed self-assignment accuracy using 80%, 90% and 100% of the loci, with (where needed) loci selected based on F_{ST}.

The algorithm works by first reducing dimensionality using a principal components analysis (PCA) and then using the output of the PCA to build a machine learning classification function. After evaluating options for this function, we chose the support vector machine learning function as it performed best on the data set. We then used the "assign.X" function to assign individuals sampled off Kamchatka to the baseline strata (NFG and SI) using the random forest model.

2.4.5 | Estimation of effective size

Contemporary effective population size (N_e) was estimated from the microsatellite genotype data using the bias-corrected version of the linkage disequilibrium method (Hill, 1981; Waples, 2006; Waples & Do, 2008), as implemented in NeEstimator v2.1 (Do et al., 2014), which has been shown to be one of the most robust and accurate single sample estimators of Ne (Gilbert & Whitlock, 2015; Wang, 2016) As recommended in Waples and Do (2010), a minimum allele frequency cutoff of 0.02 was used for all strata except Kamchatka, for which a cutoff of 0.05 was used to ensure that the critical value fell between $1/2n$ and $1/n$, where n is the number of samples representing the stratum. The jackknife approach was used to estimate 95% confidence intervals.

3 | RESULTS

3.1 | Quality control analysis of microsatellite data

The per-allele error rate for the microsatellite genotype data was estimated as 0.0086 based on random replication of 18 or more samples. Among SI samples, no loci were identified as being out of HWE based on either the probability test or the test for heterozygote deficiency. The test for heterozygote deficiency identified one locus as being out of HWE in the NFG stratum (RW48, p = .009), and the probability test identified one locus as being out of HWE in the Kamchatka stratum (EV37, p = .043). Tests for null alleles identified locus SW10 and SW19 as having null allele frequencies >0.05 in the Kamchatka stratum, likely resulting from the small number of samples representing this stratum. None of the other strata had null allele frequencies >0.05 for any locus. No samples were identified as being overly influential on significant deviations from HWE in the jackknife analysis.

Thirteen of 66 loci pairs were identified as being out of linkage equilibrium within the SI stratum, while only three loci pairs were identified in the Kamchatka stratum (Table S3). When the two feeding strata were combined, 10 loci pairs were out of linkage equilibrium in the WNP regional stratum. Three loci pairs were out of linkage equilibria in the NFG stratum. No loci were identified as being out of HWE or in LD in all three

feeding strata, and thus all loci were retained in the analysis.

The probability that two randomly chosen individuals would share the same multi-locus genotype under the assumption of Hardy-Weinberg equilibrium (PID_{HW}) was estimated at 2.9×10^{-12}, while the estimate allowing for the presence of full siblings in the data set was 3.1×10^{-5}. Within the SI stratum, 42 samples (from 32 individuals) were identified as duplicates (having come from an individual that had previously been sampled) based on identical genotypes. An additional eight samples (from six individuals) were identified as genetic duplicates within the Kamchatka stratum. Sample pairs identified as duplicates also had the same sex and mtDNA haplotype. In all cases, these duplicate sampling events were confirmed by examining the photo-identification records associated with each biopsy, and one of the samples was removed from subsequent analyses, leaving the sample sizes for SI and Kamchatka at 156 and 16 individuals, respectively. Six individuals were sampled in both SI and Kamchatka. For analyses in which these two strata were grouped to form a regional WNP stratum, one of each duplicate pair was removed prior to further analysis, leaving that stratum represented by 166 individuals. For the NFG stratum, within-stratum duplicates were removed from the data set as part of the analyses conducted in Lang et al. (2014).

3.2 | Sex ratios

Within the NFG stratum, 41% of the individuals sampled were males. Within the WNP feeding strata, the proportion of males was markedly higher, with 57% and 63% of the individuals sampled off Sakhalin and Kamchatka, respectively, being males, and 57% of all the individuals sampled in the WNP (Sakhalin + Kamchatka) being males.

3.3 | Genetic diversity

A total of 34 mtDNA control region haplotypes were identified among the NP gray whale samples (Table 1). Thirty-two of these haplotypes were found in the NFG, while 22 haplotypes, including two not sampled in the ENP, were represented among SI whales. Seven of the nine haplotypes identified among Kamchatka whales were also identified among SI whales, while the other two haplotypes were found in the NFG stratum. All haplotypes were previously identified as part of other studies (Lang et al., 2014; LeDuc et al., 2002).

The median joining network shows the relationship among mtDNA haplotypes and their frequency in each stratum (Figure 2). MtDNA haplotypes identified among animals feeding off Sakhalin are dispersed throughout the network, and no phylogeographic pattern is apparent.

Haplotype diversity in the Sakhalin stratum (h = 0.760) was

markedly lower than that found in the NFG stratum (h = 0.952), while the diversity identified among the Kamchatka whales was intermediate (h = 0.883) (Table 1). Diversity was even lower when only those whales first identified as calves off Sakhalin were considered (h = 0.711 for calves, h = 0.804 for whales not first identified as calves, i.e., noncalves). Sixty-nine percent of sampled SI whales carried one of two haplotypes (Table S4). One of these haplotypes (Hapid001, found in 36% of sampled SI whales) was also one of the most common haplotypes in the other two feeding strata (10% of sampled whales in the NFG stratum, 19% of sampled Kamchatka whales), while the second haplotype common among Sakhalin whales (Hapid002, in 33% of sampled SI whales) was found in only low frequencies (3%) in the NFG whales, although it was also common in the Kamchatka stratum (31% of sampled whales).

The distribution of haplotypes among SI whales also differed from the distribution among NFG whales with respect to the "singleton" haplotypes. Of the 11 haplotypes found in only a single SI whale, all were found in samples genetically determined to be males. Although a greater proportion of the whales sampled off SI were males (57%), finding that all singleton haplotypes were carried by males contrasts with what was seen in the other two strata, where singleton haplotypes

were carried by roughly similar numbers of males and females (50%, Kamchatka; 39%, NFG).

Eight haplotypes were found among SI whales known to be reproductive females, and known mother-calf pairs comprised 70% and 73% of the whales carrying the two most common haplotypes (Hapid001 and Hapid002, respectively) in that stratum. With one exception, all haplotypes that were identified in more than two sampled animals in the SI stratum are composed of at least one known mother-calf pair. Off Kamchatka, all of the sampled females carried different haplotypes.

Measures of microsatellite diversity for each stratum across the twelve loci are shown in Table 1 and Table S5. In general, nuclear diversity was similar but slightly lower in the SI and Kamchatka strata when compared to that found among the NFG stratum. The SI stratum contained seven private alleles (i.e., alleles found only in that feeding stratum), whereas all of the microsatellite alleles found in the Kamchatka stratum were found in at least one other strata. The NFG stratum contained 13 alleles that were not found in either of the two WNP strata.

Although the F_{IS} value for the NFG stratum was positive (F_{IS} = 0.014), measures for the WNP strata were both slightly negative (F_{IS} = −0.022, Kamchatka; F_{IS} = −0.01, SI). Mean relatedness was

higher within the Sakhalin whales (r = 0.043) than within each of the other two feeding strata (r = −0.016, NFG; r = −0.034, Kamchatka). Within SI, mean relatedness among calves (r = 0.064) and among noncalf females (r = 0.049) was higher than that seen among noncalf males (r = 0.017).

When the SI stratum was split into those whales recorded in the ENP (SI-ENP) versus those not seen in the ENP (SI only), measures of genetic diversity in both these subgroups were similar to each other and to that seen in the combined SI stratum (Table 1).

3.4 | Population structure

The results of pairwise comparisons between strata are shown in Table 2. For both the mtDNA and the microsatellite comparisons, the magnitude of differentiation between the NFG and Kamchatka ($F_{STmtDNA}$ = 0.027, p = .026; $F_{STmsats}$ = 0.015, p = .003) was markedly higher than that observed between SI and Kamchatka (mtDNA: $F_{STmtDNA}$ = 0.001, p = .355; $F_{STmsats}$ = 0.001, p = .355). The greatest genetic differentiation was observed between NFG stratum and the SI individuals that were first identified as calves ($F_{STmtDNA}$ = 0.116, p < .001; $F_{STmsats}$= 0.021, p < .001), with lower levels of differentiation observed in the comparison of NFG with the SI noncalves ($F_{STmtDNA}$ = 0.064, p < .001; $F_{STmsats}$ = 0.012, p < .001). No significant genetic differentiation was

found when the SI stratum was split into whales first identified as calves and those first identified as noncalves ($F_{STmtDNA}$ = 0.002, p = .291; $F_{STmsats}$ < 0.001, p = .471) or when it was split into those whales that have been seen off Sakhalin and in the ENP, and those not seen in the ENP (i.e., Sakhalin only; $F_{STmtDNA}$ = 0.021, p = .051; $F_{STmsats}$ = −0.002, p = .824). The magnitude of differentiation between the NFG and the Sakhalin only individuals ($F_{STmtDNA}$ = 0.100, p = .001; $F_{STmsats}$ = 0.018, p = .001) was greater than that seen between the NFG and the Sakhalin-ENP whales ($F_{STmtDNA}$ = 0.073, p = .001; $F_{STmsats}$ = 0.008, p = .009). When the NFG and SI strata were subdivided by sex, the magnitude of differentiation was higher among females ($F_{STmtDNA}$ = 0.069, p = .001; $F_{STmsats}$ = 0.018, p = .001) than males ($F_{STmtDNA}$ = 0.060, p = .001; $F_{STmsats}$ = 0.008, p = .009).

The Mantel test did not detect statistically significant associations between geographic distances and genetic differentiation (mtDNA, Mantel's r = 0.43, p = .173; microsatellites, Mantel's r = 0.56, p = .14; Figure S1), suggesting that IBD is not driving the observed pattern of genetic differentiation. Of note, however, our sampling of the NFG was limited to three primary areas and did not encompass all areas gray whales are known to feed; as such, our data may have had low power to detect IBD if it were present (Perez et al.,

2018).

When the full microsatellite data set was analyzed in STRUCTURE using a model without a priori information on the geographic location of sampling, both the ΔK and the mean LnP(K) were maximized at K = 2 (Table 3). The pattern of individual assignments is discussed below. The most likely number of clusters present in the data remained the same when whales identified as the calves of sampled Sakhalin mothers were removed (Table S6). When the NFG stratum was analyzed separately, ΔK supported the presence of two clusters, although mean LnP(K) was maximized at K = 1. Given that ΔK cannot be used to evaluate the likelihood that K = 1, we interpreted these results as supporting a single cluster within the NFG samples. Both ΔK and the mean LnP(K) were maximized at K = 3 when only the SI samples were analyzed, as well as when the combined SI+Kamchatka data set was analyzed (Table S6). The assignment of individuals to clusters was generally similar to the results for the full data set, where the third SI cluster was represented primarily by whales that did not have strong (≥80%) assignment to either cluster in the full data set analysis.

When the STRUCTURE runs were replicated using all NFG samples and ten randomly chosen subsamples (n = 105) from the Sakhalin stratum, the ΔK criterion identified the optimal number

of clusters as K = 2 for eight of the ten subsampled data sets (Table S7). However, when the mean LnP(K) was used as the criterion, K = 2 was chosen in only 50% of the subsample runs, while K = 1 was chosen as optimal in the remaining runs. This highlights the patterns seen in the full data set analysis; if the SI subsample contains a higher proportion of whales that assign strongly to the ENP cluster, then only a single cluster was detected.

When individuals were assigned to the cluster in which they had a 50% or greater assignment probability in the absence of a location prior, the majority of whales sampled off Sakhalin (61%) and Kamchatka (56%) assigned to a cluster comprised primarily of WNP whales, while most of the whales sampled in the NFG stratum (91%) assigned to a cluster comprised primarily of ENP whales (Figure 3). When only "strong" (≥80%) assignments were considered, less than half of the whales sampled off SI (39%) and Kamchatka (38%) assigned to the cluster dominated by WNP whales, while 59% of the whales sampled on the NFG assigned strongly to the ENP cluster. The proportion of whales that were 'mis-assigned' (strongly assigned to a cluster containing primarily whales from the other side of the North Pacific) differed for the ENP and WNP strata. In Kamchatka and SI, respectively, 12.5% (n = 2) and 18% (n = 28, including 15 males

first identified as noncalves, four females first identified as noncalves, and nine whales first identified as calves) of the whales strongly assigned to the ENP cluster, while only 1% (*n* = 1) of whales sampled on the NFG assigned strongly to the predominantly WNP cluster. If each cluster were defined as containing only whales that strongly assigned to that cluster, 67% of the whales in one cluster were sampled in the ENP, while 99% of the whales assigned to the other cluster were sampled in the WNP.

When STRUCTURE was run with the same parameters but incorporating a location prior, all except one whale sampled on the NFG assigned strongly to the ENP cluster (Figure 3). The one remaining whale assigned in approximately equal proportions to the ENP and WNP clusters. This whale was a male sampled off Barrow, AK in the summer of 2010. Of the whales sampled off Sakhalin, 62% (*n* = 96) assigned to the WNP cluster using a threshold of $Q \geq 0.5$; 69 of these assigned to the western cluster using an 80% threshold. Thirty-eight percent (*n* = 60) of the SI whales assigned to the ENP cluster at the 50% threshold; approximately half of those whales (*n* = 27) remained assigned to the ENP cluster using the more stringent 80% cutoff.

If only whales strongly assigned (under the model incorporating the location prior) to one of the two clusters are

considered, 45% of the WNP cluster are whales first identified as calves, while noncalf females make up 30% and noncalf males 25%) of the cluster. In contrast, a large proportion of the whales making up the ENP cluster are noncalf males (59%) with noncalf females making up only 11% of the cluster and whales first identified as calves making up 30% of the cluster. Seventeen of the known reproductive females assigned strongly to the WNP cluster, while only two assigned strongly to the ENP cluster. Six of the 11 males that carry haplotypes not found among other Sakhalin whales assign strongly to the ENP cluster, while the other five do not assign strongly to either cluster.

Of the whales (n = 36) sampled off Sakhalin that are known from photo-identification, tagging, or genetic studies to have traveled to the ENP (Lang, 2010; Mate et al., 2015; Urban R. et al., 2012, 2013, 2019; Weller et al., 2012), 50% (n = 18) assigned to the ENP cluster, nine of which were strongly assigned, while the other half (n = 18) assigned to the WNP cluster (12 of which were strongly assigned).

When the DAPC was run without incorporating geographic information, the most likely number of clusters based on BIC was between two and five (Figure S2a). The alpha scores for these solutions were relatively high, ranging from 0.42 to 0.67, and maximized when 12–13 PCs were retained (Figure S2b).

Scatterplots of the first two discriminant functions showed some overlap between the clouds of points representing each cluster. However, all three clusters comprised individuals from all three feeding grounds (Figure S3). Examination of the individual sighting histories and known relationships of the SI whales did not reveal clear patterns. Although there was a tendency for whales first identified as calves to group with their mothers, this was not always the case, with 12 of the 69 calves with known and sampled mothers grouping into different clusters. SI whales known to have migrated into the ENP were represented in all three clusters.

When a priori information on feeding ground affiliation was incorporated into the DAPC, the alpha score was small (α = 0.15), indicating relatively low discriminatory power, and was maximized when 19 PCs were retained. When visualized on a scatter plot, the ellipses encompassing the cloud of points for each feeding ground stratum overlapped, and none of the samples were tightly grouped (Figure 4).

The accuracy of self-assignment of Sakhalin and NFG samples to their strata of origin was similar across the range of parameters tested in assignPop. When the training data were selected from among all Sakhalin samples, the highest median accuracy across parameter sets was 64.3%, only slightly better

than what would be expected if assignment were random (Figure 5). The highest median accuracy for the NFG samples was 68.2%, although the minimum nonoutliers (e.g., the whiskers in the boxplot) fall below 50%. The highest median accuracy increased slightly when the training data representing Sakhalin were drawn from only those whales that have not been recorded in the ENP. Six of the whales sampled off Kamchatka assigned to the NFG while 10 assigned to Sakhalin (Figure 6). Of those, five of the six Sakhalin assignments had high (≥.80) probabilities while only two had high probabilities of belonging to the NFG. Using the same set of baseline samples, 15 of the SI whales known to have visited the ENP assigned to the NFG while 19 assigned to SI. Eleven and ten of those, respectively, were strong assignments.

3.5 | Estimates of effective size

The N_e estimate for the NFG stratum (N_e = 1,027, 95% CI [263.4, ∞]) is lower than would be expected based on the census size (~27,000, Durban et al., 2015, 2017), and the upper bound of the confidence interval includes infinity (Table 1). This is likely due to the small number of samples, relative to total abundance, representing that stratum, as simulations suggest that N_e estimates may be biased and imprecise when sample sizes are below 30 and the number of samples relative to N_e is small (<10%)

(Waples, 2006; Waples & Do, 2010). While an upper bound was assigned to the 95% CI for the Kamchatka stratum (N_e = 29.4, 95% CI [12.6, 676.6]), the number of samples representing this stratum is below the criterion noted above, and the confidence intervals surrounding the estimate were broad. Given these limitations, the estimates for the NFG and Kamchatka strata were considered unreliable and are not discussed further.

When all samples representing the SI stratum were included in the analysis, N_e was estimated at 80 whales (95% CI [61.9, 107.7]; Table 1). When the SI whales that have been identified in the ENP were removed from the SI data set and analyzed separately, N_e for the group of whales seen on both sides of the North Pacific (n = 34) was 51 whales (95% CI [26.5, 179.8]) while the estimate for the remaining SI whales was 70 whales (95% CI [44.4, 121.9]). For all strata, the samples were collected from age-structured populations (i.e., more than one cohort was included in the stratum), and thus the resulting estimate can be interpreted as the number of breeders (N_b) that produced the cohorts from which the samples were taken (Waples & Do, 2010). Although restricting the analysis to any single cohort reduced the sample sizes to levels that would not produce meaningful estimates, when only whales first identified as calves between 1995 and 2011 were included, the N_e estimate (N_e =

85, 95% CI [58.7, 136.6]) was lower than that derived from including only samples from noncalf whales (N_e = 104, 95% CI [58.1, 276.6]), although the lower bounds were similar. Although not shown on the table, the estimate of Ne when the SI and Kamchatka samples were combined ($N_{e\text{-}WNP}$ = 82, 95% CI [64.3, 107.4]) was also similar to the estimate derived from Sakhalin alone.

4 | DISCUSSION

Although some of the baleen whale populations decimated by commercial whaling remain depleted, others are at or near prewhaling levels, providing some of the first insights into baleen whale recovery patterns. When research was initiated off Sakhalin Island, Russia, in the mid-1990s, the gray whales feeding there were presumed to represent a remnant of the population of gray whales that was historically hunted off the coasts of Japan and Korea. More recent findings showing that some of the whales that feed off SI migrate to and overwinter in the ENP (Mate et al., 2015; Urbán R. et al., 2019; Weller et al., 2012) changed this perception, raising the possibility that the recovery of the ENP gray whale may have also played a role in the recolonization of the SI feeding ground. At the same time, rare but continued sightings of gray whales off the coast of Japan and China (Nakamura et al., 2019; Wang et al., 2015;

Zhao et al., 2017; Zhu, 2012; Zhu & Yue, 1998) suggest that some whales, including two first identified as calves with their mothers on the SI feeding ground (Weller et al., 2008, 2016), overwinter in the WNP. In a range-wide review of the status of NP gray whales conducted by the International Whaling Commission (2019), evaluation of this and other available information led to the identification of two primary hypotheses[1] regarding the identity of the gray whales feeding off SI, one of which assumes that all of the whales using the SI feeding ground are whales that overwinter in the ENP and a second that assumes that the SI feeding ground is used by some whales of eastern origin but also by some whales that remain in the WNP year-round. Here we used molecular genetic data from samples collected off SI, Kamchatka, and the NFG to evaluate these two hypotheses.

4.1 | Evaluating demographic independence

Several independent lines of evidence indicate that recruitment of whales using the Sakhalin feeding ground is largely driven by matrilineal fidelity. The results of our mtDNA genetic analyses are consistent with the patterns identified in previous genetic studies (LeDuc et al., 2002; Meschersky et al., 2015) and

[1] Of note, while these hypotheses were considered the most plausible, four additional hypotheses have also been considered plausible and are included as sensitivity tests in the IWC Scientific Committee's assessment of NP gray whales (IWC, 2018).

provide further support for the demographic independence of the group of whales feeding off Sakhalin. As with the two previous studies, the genetic signal of matrilineal fidelity among the whales sampled off SI is apparent both in the differences in mtDNA haplotype frequencies between strata and in the distribution of haplotypes among Sakhalin individuals. The majority (69%) of whales sampled on the SI feeding ground, including 21 of the 29 reproductive females that have been biopsied (of 34 identified through 2017, Burdin et al., 2018), carry one of the two most common haplotypes. This haplotype distribution is reflected in the reduced haplotype diversity found among the SI whales, and suggests that recruitment into the SI feeding ground is largely driven by matrilineal fidelity. While still markedly low compared to ENP whales, the haplotype diversity found among our expanded SI sample set (h = 0.77), which includes whales sampled through 2011, has increased since the LeDuc et al. (2002) study (h = 0.70) that was based on the 45 individuals sampled prior to 2000. This increased haplotype diversity is being driven by the eight sampled reproductive females that do not have one of the two common haplotypes, suggesting that the haplotype diversity among whales feeding off SI could continue to increase in the future. However, 11 haplotypes are represented by only a single individual, all of

which are males. Because males do not pass on their mtDNA, the total number of haplotypes found among the whales feeding off SI is likely to decrease in the future unless additional immigration of whales from the ENP occurs.

A high proportion (91%) of the mtDNA control region haplotypes found among sampled SI whales have also been identified in whales sampled on the NFG. The extent to which these shared haplotypes can be attributed to the contemporary movements of animals between the ENP and WNP versus to shared ancestry is not clear. Although there is debate about the spatio-temporal scale for which such estimates are applicable (Palsbøll et al., 2013), the genetic estimate of historic abundance of gray whales in the North Pacific is large (~96,000; Alter et al., 2007). While the extent of connectivity between gray whales in historic times is unknown, it has been hypothesized that increased interchange of gray whales between the eastern and western North Pacific could have occurred during the Little Ice Age (~1,300–1,850), when increased sea ice and decreased sea levels may have blocked access to the Bering Sea and shifted their distribution southward (Alter et al., 2007; Swartz et al., 2006). Even in the absence of contemporary connectivity, the combination of historically high abundance, evolutionarily recent migration, and a long generation time

would likely result in some proportion of shared haplotypes between gray whales in the eastern versus the western North Pacific, as not enough time may have passed for new haplotypes to evolve or existing haplotypes to be removed via drift. Under such a scenario, the population of gray whales that was subjected to intensive commercial whaling in the WNP between 1900 and the 1960s would have carried many of the same haplotypes found in the ENP whales. However, if the WNP population that was extant in the first part of the 19th century was reduced to such low numbers that some thought it extinct, it is unlikely that such a large number (n = 22) of mtDNA haplotypes would have been retained in the WNP without any immigration of ENP whales. For example, the Okhotsk Sea bowhead whale (*Balaena mysticetus*) population, which is thought to number approximately 220 whales (Cooke et al., 2017a), contains four mtDNA control region haplotypes, all of which are also found in the much larger Bering-Chukchi-Beaufort Sea stock of bowheads inhabiting the ENP (Alter et al., 2012; LeDuc et al., 2005). Thus, while we cannot disentangle to what extent the high proportion of haplotypes shared between the ENP and WNP whales is due to contemporary movements versus common ancestry, the number of haplotypes found among the SI whales is not consistent with what would be expected in a small remnant population with

little immigration from other groups.

The gray whale PCFG, which is defined by the International Whaling Commission as whales seen in two or more years during the feeding season (June through November) within the region extending from northern California through northern British Columbia (roughly 41°N to 52°N; IWC, 2011), also exhibits differences in mitochondrial DNA haplotype frequencies when compared to NFG whales (F_{ST} = 0.012, p = .0045; Lang et al., 2014). These differences are consistent with recruitment into the PCFG being driven, at least in part, by matrilineal fidelity to the feeding area, a finding which is supported by long-term photo-identification records (Calambokidis & Pérez, 2017b). The PCFG is estimated to contain approximately 240 whales (Calambokidis et al., 2017), making it similar in size to the number of whales utilizing the Sakhalin feeding ground. Despite both groups being comprised of only a small number of individuals, the magnitude of differentiation found between the PCFG and the NFG whales (F_{ST} = 0.012, p = .005) is markedly lower than that estimated here between the SI and NFG whales (F_{ST} = 0.093, p < .001). This difference suggests that the degree of dispersal into the PCFG is higher than that into the SI feeding ground and/or that the PCFG colonized the southern feeding ground more recently. The different locations of these two

feeding grounds with respect to the ENP migratory route may, at least in part, drive the differences in patterns seen between the two areas. While the SI feeding ground is approximately 2000 km west of the NFG, the PCFG feeding area is located along the ENP migratory route. Thus, some recruitment into the PCFG by whales that previously fed further north could occur if whales migrating through the area opportunistically identify prey resources of sufficient quality that they decide to remain in the area during that season and then to return in subsequent years.

4.2 | Relationship of Kamchatka feeding ground to other feeding grounds

Genetic differentiation based on both mtDNA and nuclear genetic data suggest that the Kamchatka feeding ground is closely connected to the Sakhalin feeding ground. This finding is consistent with the results of photo-identification comparisons, where half of the whales photographed off Kamchatka have also been recorded off SI (39 of 78 individuals identified off Kamchatka; Tyurneva et al., 2010). It is also supported by genetic duplicate samples identifying six of the 16 whales sampled off Kamchatka as having been also sampled off the Sakhalin feeding ground.

4.3 | Mixing versus admixture

The differences in the location of the two southern gray whale feeding grounds (i.e., SI and the PCFG) relative to the ENP migratory path may drive the different patterns seen in the nuclear comparisons of those two areas with the NFG. Comparison of microsatellite allele frequencies between the PCFG (n=71) and the same set of samples used here to represent the NFG did not support genetic differentiation (F_{ST} = 0.000, p = .527; Lang et al., 2014), indicating that PCFG whales likely interbreed with NFG whales while on migration or in or near the lagoons. In contrast, nuclear genetic differentiation was identified between Sakhalin and NFG whales (F_{ST} = 0.016, p < .001), indicating that the Sakhalin whales are not mating randomly with the larger group of NFG whales. The mechanism driving this assortative mating is unclear and could be generated in at least two different ways. First, if a subset of the SI gray whales (potentially representing a remnant of the "historic" western gray whale population that migrated past Japan and Korea in the early to mid-1900s) remain in the WNP year-round and interbreed only with each other, then differences could be generated between Sakhalin and the NFG. Under this hypothesis, the Sakhalin feeding ground would be used by two separate breeding stocks, one from the ENP and one from the WNP, and would represent a mixed-stock feeding aggregation. This hypothesis

would be consistent with (1) the recently documented movements of whales between SI and the ENP (Mate et al., 2015; Urbán R. et al., 2019; Weller et al., 2012); (2) a model-based assessment indicating that 20%–55% of the SI whales do not use ENP wintering grounds (Cooke et al., 2019); and (3) the contemporary winter and spring records of gray whales off Japan and China (Nakamura et al., 2018; Wang et al., 2015; Zhao et al., 2017; Zhu, 2012; Zhu & Yue 1998), at least two of which are known to have been brought to Sakhalin by their mothers as calves (Weller et al., 2008, 2016). Secondly, a signal of genetic differentiation could be generated if most of the Sakhalin whales overwinter in the ENP but primarily interbreed with each other. Both scenarios involve whales of eastern origin either colonizing a new feeding ground, or, if the Sakhalin feeding ground were used historically, recolonizing a previously used area. However, under the latter scenario, only a small number of whales that descended from the population hunted off Japan and Korea in the early 1900s would use the SI feeding ground, although they could be extant and feeding in unknown areas.

Differentiating between these two hypotheses based on the results of our genetic analyses is difficult given the lack of samples from WNP breeding/wintering areas. The DAPC analysis with no a priori information on sampling location clustered the

samples into three somewhat defined groups, but there was little geographic concordance in the assignment of samples to clusters, nor did the grouping reveal any patterns with respect to Sakhalin whales known to travel to Mexico. Some weak evidence against the mixed-stock hypothesis can be derived from the fact that our analysis of microsatellite allele frequencies off SI did not identify any loci as being out of HW equilibrium, nor did we find positive F_{IS} values. Both signals would be expected under a "Wahlund effect" (Wahlund, 1928), which results from the mixing of two distinct stocks on the feeding ground (Waples, 2015). However, the power of this test depends on the amount of genetic differentiation between stocks as well as how evenly each stock is represented in the sample set. Although subject to these same caveats, the near lack of private microsatellite alleles is also inconsistent with expectations based on two breeding stocks, one of which had been reduced to very low levels, using the SI feeding ground. The number of loci pairs out of LD was markedly higher in the Sakhalin group than in the other groups. However, the significance of this is difficult to determine, as such a signal could correlate with mixture and/or admixture between two stocks (Nei & Li, 1973; Sinnock, 1975), but could also result from colonization of the area by a relatively small number of individuals (Slatkin, 1994).

Perhaps the strongest argument against the mixed-stock hypothesis is the STRUCTURE results, which provided evidence for admixture (gene flow) between the two distinct genetic clusters identified among Sakhalin whales (Figure 3). Under the mixed-stock hypothesis, two distinct clusters representing a western breeding stock and an eastern breeding stock should be present. However, under this hypothesis the two groups would not breed until each was on migratory routes and/or wintering grounds on opposite sides of the North Pacific, and thus there would be no opportunity for admixture between the two groups to be generated. Under that hypothesis, it would also be hard to explain why whales known to migrate to the ENP are found in both clusters. However, both results could be explained under the scenario where all or most of the whales using the Sakhalin feeding ground overwinter in the ENP. Much of what is known about the reproductive cycle of gray whales is based on specimens (n = 316) collected off the coast of central California during scientific permit whaling between 1959 and 1969 (Rice & Wolman, 1971). Except for near-term pregnant females, all other mature females that were collected during the southbound migration showed signs of recent ovulation. While not all these females may have conceived, very few of the early pregnant or nonpregnant females collected on the northbound

migration showed evidence of multiple ovulations, suggesting most females do conceive during their first round. Back-calculation of fetal growth rates suggested that conception occurs primarily during a 3-week period from late November to early December (November 27–December 13, Rice & Wolman, 1971). The median (peak) sighting date for the southbound migration in the ENP was estimated to be December 12 for Unimak Pass, Alaska, in 1998/1999 (Rugh et al., 2001), suggesting that many animals from the NFG are north of the Aleutians during the first mating period. Of the three Sakhalin whales that were tagged before they began migrating east, one remained off Sakhalin until December 10 and the other two remained there until November 24 (Mate et al., 2015). This indicates that at least some and perhaps all animals making the journey between Sakhalin Island and Mexico would be relatively far west during the first mating period. Gray whales regularly make low frequency moan calls while migrating (Burnham et al., 2018; Guazzo et al., 2017); these calls are thought to maintain cohesion between groups and may provide the means by which Sakhalin whales remain in contact with each other while migrating east. Thus, if the first mating period occurs while Sakhalin whales are in proximity to each other, but not to whales from other feeding areas, while on the westernmost portion of their migration, some degree of

reproductive isolation could develop between Sakhalin whales and those feeding in areas to the east even if they shared a common wintering destination.

There is also evidence to suggest that whales from the same feeding ground may preferentially associate and migrate with each other. Of the six Sakhalin whales that have been photographed during the northbound migration, all sightings occurred on only two days, with three whales sighted as part of a single group on one day and the other three whales sighted in two groups in close proximity to each other on a single day (Weller et al., 2012). In addition, of the 15 cases where PCFG whales have been photographed in a group of two or more whales while migrating past southern California, nine of the groups included more than one PCFG whale, with five cases where between three and five PCFG whales were part of the same group (Calambokidis & Perez, 2017a). Six of these nine groups were migrating south, and the group photographed with at least five PCFG whales was photographed off the Palos Verdes Peninsula on December 12, 2013 (i.e., during the time period in which conception is thought to occur). Thus, even if mating occurred after Sakhalin whales had joined the migratory path used by NFG and PCFG whales, such associations could provide greater opportunities for Sakhalin whales to mate with each other rather

than with whales from other feeding areas.

The relatively small estimate of the contemporary effective population size of the Sakhalin gray whales is also generally inconsistent with the mixed-stock hypothesis. Given that our estimates were based on samples belonging to multiple age classes, they should be interpreted as the number of breeding whales contributing to the cohorts in the samples (Robinson & Moyer, 2013; Waples et al., 2014). If a substantial proportion of the whales feeding off Sakhalin overwinter in the ENP and breed at random with the very large pool of potential mates there, we would expect the estimate of N_e to be markedly higher and more similar to that estimated for the NFG samples. Under a scenario where all or most of the Sakhalin whales overwinter in the ENP, however, the N_e estimate would remain small if most of the mating was occurring, as hypothesized above, early in the migration with other Sakhalin whales.

Two examples of baleen whale populations for which both N_e and N_c have been calculated are the Okhotsk Sea bowhead whales and North Pacific right whales. For the bowheads, N_e was estimated at 112 whales (95% CI [79, 183]; Morin et al., 2012), while N_c, as estimated from a genetic mark-recapture study, was estimated at ~220 whales in 2016 (*CV* 0.22; Cooke et al., 2017a). For the very small NP right whale population, N_e was estimated to

be 12 whales (95% CI [2.9, 75.0]; LeDuc et al., 2012) while the census size was estimated at 28–31 whales based on photographic mark-recapture data (Wade et al., 2011). Based on estimates from these two baleen whale populations, the ratio of N_c to N_e is approximately 2.0–3.5, which is similar to ratios used in other baleen whale studies in which this ratio is used to calculate historic population sizes (Alter et al., 2007; Roman & Palumbi, 2003; Ruegg et al., 2010, 2013). If a similar ratio were applied to the estimate of N_e for the Sakhalin feeding whales, the census size would be estimated at 200–300 whales. This estimate is similar to model-based estimates of the noncalf abundance of western breeding and/or feeding whales in 2015, which ranged from ~200 to 290 depending on the underlying stock structure model (Cooke, 2018). Although based on a number of assumptions, this rough calculation suggests that it is possible that all breeding individuals could be accounted for off Sakhalin and Kamchatka.

It is generally presumed that gray whales demonstrate natal philopatry to wintering areas, which is supported by records of some females returning to the same Mexican nursing lagoons in multiple years (Jones, 1990; Martínez A. et al., 2016; Urban R. et al., 2003). It is possible, however, that gray whales may exhibit behavioral plasticity in wintering ground affiliation,

which has been documented in humpback whales (e.g., Pomilla & Rosenbaum, 2005; Salden et al., 1999; Stevick et al., 2011, 2016). If so, two alternative hypotheses are plausible. First, as whales that traditionally overwintered in the ENP continue to show fidelity to the Sakhalin feeding ground, some may have also begun to explore and potentially recolonize historically used WNP migratory routes and wintering areas. These whales would presumably interbreed with each other, allowing population structure between the SI and NFG whales to develop. This hypothesis would be consistent with the recent increase in the number of gray whales recorded off Japan (Nakamura et al., 2019), coincidental with increases in the number of gray whales feeding off Sakhalin (by 3.4%–4.8% over the past 20 years, Cooke et al., 2019). A second possibility is that whales that historically used WNP wintering grounds have followed whales of eastern origin that feed off Sakhalin to the ENP wintering grounds. In this case the genetic results would be explained by ongoing homogenization of the formerly separate eastern and western gene pools.

4.5 | Caveats

The strongest evidence against the mixed-stock hypothesis is the STRUCTURE results, which indicate interbreeding between the two distinct genetic clusters identified among Sakhalin whales, each

of which contained some individuals known to have migrated between Sakhalin and the ENP. However, the assignment accuracy of STRUCTURE is known to be low when F_{ST} is less than 0.05 (Latch et al., 2006), as was the case in our comparison of SI and the NFG. In addition, the results of STRUCTURE analyses can be impacted by several different factors, including the inclusion of close relatives (Anderson & Dunham, 2008; Rodríguez-Ramilo & Wang, 2012) and unbalanced sampling strategies (Fogelqvist et al., 2010; Puechmaille, 2016). While the STRUCTURE results supported the presence of two genetic clusters when we removed one of each known mother-calf offspring, when we subsampled the Sakhalin stratum to ensure that it was represented by the same number of samples as the NFG stratum, mixed results were obtained, with some runs supporting two or more clusters but others finding evidence of only a single cluster. This pattern was opposite of that typically observed when sample sizes are uneven, as usually the less sampled group is fit as a mix of multiple groups. In our case, however, the less sampled group (the NFG) represented a relatively distinct cluster while the more heavily sampled SI group was represented as a mix of two groups, one of which was associated with the NFG. However, such a result is consistent with what might be expected if the SI feeding ground is used by some whales that are genetically

undistinguishable from the NFG, as runs where more of those individuals were included in the sample set would result in a higher probability of K = 1.

Furthermore, Lawson et al. (2018) showed that in cases where populations have complex histories, STRUCTURE may arrive at the same "solution" under multiple different population histories, such that a barplot indicative of recent admixture could look very similar to one representing a recent bottleneck or admixture with a "ghost" (unsampled) population. Comparing our results to those of other studies further demonstrates that STRUCTURE barplots can show similar patterns even when the underlying connectivity between the groups being analyzed is thought to be quite different. For example, in the absence of a priori information on location of sampling, a STRUCTURE bar plot based on analysis of microsatellite data generated from humpback whales sampled in the North Atlantic and Southern Hemisphere closely resembled the bar plot generated here for North Pacific gray whales, with a substantial number of humpback whales sampled in the South Atlantic demonstrating a genetic signature consistent with admixture with North Atlantic whales (Ruegg et al., 2013). However, humpback whales in the North Atlantic and Southern Hemisphere are considered separate subspecies, with only a single shared mtDNA control region haplotype and

estimated long-term migration rates based on nuclear intron data indicating fewer than two migrants per generation between ocean basins (Jackson et al., 2014).

A second caveat in our evaluation of stock structure hypotheses is that the power to detect the presence of two isolated breeding stocks within a sample set depends on the degree of population differentiation and the evenness with which the two groups are sampled. If only a small proportion of the individuals sampled off Sakhalin are whales that remain in the WNP year-round, it is unlikely that the analyses presented here would have detected the presence of two breeding stocks on the SI feeding ground. If, in addition, the eastern and western populations of gray whales were connected by some degree of gene flow in the evolutionarily recent past, the difficulty of discriminating between our two hypotheses would be further increased. Both scenarios are plausible.

4.6 | Implications

If most of the whales feeding off Sakhalin represent whales with recent ancestry rooted in the ENP, the rare, but continuing, sightings of gray whales off Japan and China during winter and spring (Nakamura et al., 2019; Nambu et al., 2010; Wang, 1985; Wang et al., 2015; Zhu & Yue, 1998), as well as estimates that 20%–55% of the SI whales do not utilize ENP wintering grounds

(Cooke et al., 2019), indicate that some whales are remaining in the WNP year-round. This group of whales, which is of unknown origin but may include the last remnants of the population of gray whales that was historically hunted off Japan and Korea, faces multiple threats to its persistence, including but not limited to the risk of mortality due to entanglement in coastal net fisheries off Japan (Nakamura et al., 2019; Weller et al., 2008), China (Wang et al., 2015), and Sakhalin Island (Lowry et al., 2018); exposure to potentially harmful activities associated with oil and gas development in the Okhotsk Sea; and possible ship strikes while migrating coastal waters of Japan and Korea, with substantial nearshore industrialization (Weller et al., 2002b). Obtaining additional information on the distribution, movements and origin of these whales is critical to understanding their significance to the conservation of gray whales in the North Pacific.

4.7 | Summary

While other scenarios are possible, here we suggest that the genetic structure observed in our data is primarily driven by interbreeding of Sakhalin whales with each other while on migration to the ENP. Although under this scenario most of the Sakhalin whales would not represent descendants of whales historically hunted off Japan and Korea, both the lack of random

mating between SI and NFG whales and the strong evidence that continued use of the SI feeding ground is driven largely or entirely by internal recruitment indicate that management of the whales that feed off Sakhalin as a separate stock should continue.

While the results presented here are derived from genetic analyses of gray whale samples, the interpretation of those results relied heavily on our ability to link biopsies of individuals to photographically identified whales. Similar to the previous work by Brüniche-Olsen et al. (2018), we identified two genetic clusters among the whales sampled off SI. While the simplest explanation is that the two clusters represent two stocks that use migratory routes and wintering grounds on different sides of the North Pacific, examination of the sighting histories of individuals in each cluster revealed that both contained whales known to travel between SI and the ENP. This insight provides the first indication that the structure revealed in the genetic data may instead be driven by assortative mating among whales traveling to a common wintering ground destination from different high latitude feeding areas, similar to that hypothesized for North Atlantic humpback whales based on differences in wintering ground occupancy patterns among whales from different feeding grounds (Stevick et al.,

2003). It also highlights the value of combining data from different sources, including genetic, photographic, and tagging studies, to provide biological context when assessing complex patterns of population structure from molecular genetic data.

Continued genetic and photographic monitoring of the whales feeding off SI is needed both to better understand contemporary patterns of connectivity and to track future changes. Empirical studies of terrestrial carnivores have shown that natural (re)colonizations can give rise to relatively rapid changes in the magnitude of population structure, in some cases leading to increased admixture between previously differentiated groups (i.e., Finnish brown bears *Ursus arctos* over ~1.5 generations; Hagen et al., 2015) while in others resulting in increased genetic structure (i.e., Canadian fishers, *Pekania pennanti*, over ~5 generations; Greenhorn et al., 2018). Whether colonization of new or formerly used habitats leads to additional population structure or increased homogenization presumably relates to the processes driving the range expansion. For gray whales, it seems likely that the colonization of the SI feeding ground by ENP gray whales was driven at least in part by increases in abundance decades ago following their protection, similar to patterns seen among some Southern Hemisphere right whale populations leading to reclaimed historical calving and

feeding grounds as their numbers have increased in recent years (Arias et al., 2018; Carroll et al., 2014; Charlton et al., 2019; Roux et al., 2015). Although little is known about the whales feeding off Sakhalin prior to 1995, model-based estimates of recruitment suggest that the Sakhalin feeding ground has been largely or entirely closed to immigration in recent years (Cooke et al., 2017b).

ACKNOWLEDGMENTS

We thank B. Adams, R. Andrews, S. Blokhin, A. Bradford, J. Calambokidis, J. C. George, J. Herreman, Y. Ivashchenko, J. Jacobsen, H. W. Kim, S. Oliver, A. Perez, S. Reeve, J. Scordino, M. Sidorenko, B. Taylor, G. Tsidulko, and B. Würsig for their assistance with sample collection and/or their contribution to this project. Support and funding for different aspects of the research program have been provided by (in alphabetical order): Alaska SeaLife Center; Animal Welfare Institute; Exxon Neftegas Limited; the International Fund for Animal Welfare; the International Whaling Commission; the Marine Mammal Commission; the Marine Mammal Research Program at Texas A&M University at Galveston; the National Fish and Wildlife Foundation; the National Research Council Postdoctoral Fellowship program, North Pacific Wildlife Consulting, LLC; the Northwest Regional Office of NOAA National Marine Fisheries; the NOAA Dr. Nancy Foster

Scholarship Program; Ocean Park Conservation Foundation Hong Kong; Sakhalin Energy Investment Company; the U.S. Environmental Protection Agency; and the Washington Cooperative Fish and Wildlife Research Unit. All data for the analyses were generated in the SWFSC Genetics Laboratory, with assistance from A. Bowman, J. Hyde, and J. Minich., G. Serra-Valente, and N. Beaulieu assisted with the import and archiving of samples. We also thank K. Parsons, Steve Stone, and Nancy Young and three anonymous reviewers for their thoughtful reviews of this manuscript. This project was conducted as part of the Marine Mammal Project under Area V: Protection of Nature and the Organization of Reserves within the U.S.-Russia Agreement on Cooperation in the Field of Environmental Protection.

DATA ACCESSIBILITY: The genetic data will be made available on Dryad upon acceptance.

REFERENCES

Alter, S. E., Rosenbaum, H. C., Postma, L. D., Whitridge, P., Gaines, C., Weber, D., Egan, M. G., Lindsay, M., Amato, G., Dueck, L., Brownell, R. L., Jr., Heide-Jorgensen, M.-P., Laidre, K. L., Caccone, G., & Hancock, B. L. (2012). Gene flow on ice: the role of sea ice and whaling in shaping Holarctic genetic diversity and population differentiation in bowhead whales (*Balaena mysticetus*). *Ecology and*

Evolution, *2*(11), 2895–2911. https://doi.org/10.1002/ece3.397

Alter, S. E., Rynes, E., & Palumbi, S. R. (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*(38), 15162–15167. https://doi.org/10.1073/pnas.0706056104

Anderson, E., & Dunham, K. (2008). The influence of family groups on inferences made with the program Structure. *Molecular Ecology Resources*, *8*(6), 1219–1229. https://doi.org/10.1111/j.1755-0998.2008.02355.x

Archer, F. I., Adams, P. E., Schneiders, B. B. (2017). stratag: An R package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources*, *17*(1), 5–11. https://doi.org/10.1111/1755-0998.12559

Arias, M., Coscarella, M. A., Romero, M. A., Sueyro, N., Svendsen, G. M., Crespo, E. A., & González, R. A. (2018). Southern right whale *Eubalaena australis* in Golfo San Matías (Patagonia, Argentina): Evidence of recolonisation. *PLoS ONE*, *13*(12), e0207524. https://doi.org/10.1371/journal.pone.0207524

Berzin, A. A. (1990). Gray whales of the Okhotsk-Korean population in the Sea of Okhotsk. *Paper SC/A90/G28*

presented to the Scientific Committee of the International Whaling Commission.

Berzin, A. A., Vladimirov, V. A., & Doroshenko, N. V. (1988). Results of aerial surveys to study the distribution and abundance of cetaceans in the coastal waters of the Sea of Okhotsk in 1986–1987. In N. S. Chernysheva (Ed.), *Nauchno-issledovatel'skier a'boty PO morskim mlekopitayushchim severnoi chasti Tikhogo okeana v 1986–1987* (pp. 18–25). All-Union Scientific Research Institute of Fisheries and Oceanography (VNIRO), Moscow, Russia (In Russian).

Berzin, A. A., Vladimirov, V. A., & Doroshenko, N. V. (1991). Results of aerial surveys to study the distribution and abundance of whales in the Sea of Okhotsk in 1988–1990. In L. A. Popov (Ed.), *Nauchno-issledovatel'skie ra'boty PO morskim mlekopitayushchim severnoi chasti Tikhogo okeana v 1989–1990* (pp. 6–17). All-Union Scientific Research Institute of Fisheries and Oceanography (VNIRO), Moscow, Russia (In Russian).

Blokhin, S. A. (1996). Distribution, abundance and behavior of gray whales (*Eschrichtius robustus*) of American and Asian populations in regions of their summer location nearshore of the Far East. *Izvestiya Tikhookeanskogo Nauchno-issledovatel'skogo Rybokhozyaistvennogo Tsentra*, *12*(1), 36–

53 (In Russian).

Blokhin, S., Maminov, M., & Kosygin, G. (1985). On the Korean-Okhotsk population of gray whales. *Report of the International Whaling Commission*, *35*(375), 76.

Bowen, S. L. (1974). Probable extinction of the Korean stock of gray whale (*Eschrichtius robustus*). *Journal of Mammalogy*, *55*(1), 208–209. https://doi.org/10.2307/1379272

Bradford, A. L. (2011). *Population characteristics of the critically endangered western gray whale* [Doctoral dissertation]. University of Washington.

Bradford, A. L., Weller, D. W., Wade, P. T., Burdin, A. M., & Brownell, R. L., Jr. (2008). Population abundance and growth rate of western gray whales *Eschrichtius robustus*. *Endangered Species Research*, *6*, 1–14. https://doi.org/10.3354/esr00106

Branch, T. A., Stafford, K. M., Palacios, D. M., Allison, C., Bannister, J. L., Burton, C. L. K., Cabrera, E., Carlson, C. A., Vernazzani, B. G., Gill, P. C., Hucke-Gaete, R., Jenner, K. C. S., Jenner, M. N. M., Matsuoka, K., Mikhalev, Y. A., Miyashita, T., Morrice, M. G., Nishiwaki, S., Sturrock, V. J., … Warneke, R. M. (2007). Past and present distribution, densities and movements of blue whales Balaenoptera musculus in the Southern Hemisphere and

northern Indian Ocean. *Mammal Review*, *37*(2), 116–175. https://doi.org/10.1111/j.1365-2907.2007.00106.x

Bröker, K., Tyurneva, O., Gailey, G., Dupont, J., Yakovlev, Y., Vertyankin, V., Sychenko, O., Shevtsov, E., Drozdov, K. M. (2020). Site-fidelity and spatial movements of western North Pacific gray whales (*Eschrichtius robustus*) on their summer range off Sakhalin, Russia. *PLoS ONE*, *15*(8), e0236649. https://doi.org/10.1371/journal.pone.0236649

Brownell, R. L., Jr., Chun, C. (1977). Probable existence of the Korean stock of gray whale (*Eschrichtius robustus*). *Journal of Mammalogy*, *58*(2), 237–239. https://doi.org/10.2307/1379584

Brownstein, M. J., Carpten, J. D., Smith, J. R. (1996). Modulation of nontemplated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. *Biotechniques*, *20*(6), 1004–1010. https://doi.org/10.2144/96206st01

Brüniche-Olsen, A., Bickham, J. W., Godard-Codding, C. A., Brykov, V., Urban R, J., & Dewoody, J. A. (2020). Recent demographic history and population structure of gray whales inferred from complete mitochondrial genomes. *Paper SC/SC68b/SDDNA2 presented to the Scientific Committee of the International Whaling Commission*.

Brüniche-Olsen, A., Urban, R. J., Vertyankin, V. V., Godard-Codding, C., Bickham, J., & Dewoody, J. A. (2018). Genetic data reveal mixed-stock aggregations of gray whales in the North Pacific Ocean. *Biology Letters*, *14*, 20180399. http://dx.doi.org/10.1098/rsbl.2018.0399

Brykov, V. A., Efimova, K. V., Brüniche-Olsen, A., Dewoody, J. A., & Bickham, J. W. (2019). Population structure of Sakhalin gray whales (*Eschrichtius robustus*) revealed by DNA sequences of four mtDNA genes. In R. D. Bradley, H. H. Genoways, D. J. Schmidly, L. C. Bradley (Eds.), *From field to laboratory: A memorial volume in honor of Robert J. Baker* (pp. 227–240). Museum of Texas Tech University, Lubbock, Texas.

Buckland, S. T., Breiwick, J. M., Cattanach, K. L., & Laake, J. L. (1993). Estimated population size of the California gray whale. *Marine Mammal Science*, *9*(3), 235–249. https://doi.org/10.1111/j.1748-7692.1993.tb00452.x

Burdin, A. M., Bradford, A. L., Tsidulko, G. A., & Sidorenko, M. (2011). Status of western gray whales off northeastern Sakhalin Island and eastern Kamchatka, Russia in 2010. *Paper SC/63/BRG8 presented to the Scientific Committee of the International Whaling Commission.*

Burdin, A. M., Sychenko, O., & Matvey, M. (2018). Gray whale

research off northeastern Sakhalin Island and eastern Kamchatka, Russia, in 2017. *Paper SC/67b/CMP07 presented to the Scientific Committee of the International Whaling Commission.*

Burdin, A. M., Sychenko, O., & Mamaev, M. (2019). Gray whale research off northeastern Sakhalin Island and eastern Kamchatka, Russia, in 2018. *Paper SC/68a/CMP16 presented to the Scientific Committee of the International Whaling Commission.*

Burdin, A. M., Sychenko, O., & Mamaev, M. (2020). Status of gray whales off northeastern Sakhalin Island and eastern Kamchatka, Russia, in 2019. *Paper SC/68b/CMP24 submitted to the Scientific Committee of the International Whaling Commission.*

Burnham, R., Duffus, D., Mouy, X. (2018). Gray whale (*Eschrichtius robustus*) call types recorded during migration off the west coast of Vancouver Island. *Frontiers in Marine Science*, *5*(329. https://doi.org/10.3389/fmars.2018.00329

Calambokidis, J., Darling, J. D., Deecke, V., Gearin, P., Gosho, M., Megill, W., Tombach, C. M., Goley, D., Toropova, C., Gisborne, B. (2002). Abundance, range and movements of a feeding aggregation of gray whales (*Eschrichtius robustus*)

from California to southeastern Alaska in 1998. *Journal of Cetacean Research and Management*, *4*(3), 267–276.

Calambokidis, J., Laake, J., Perez, A. (2017). Updated analysis of abundance and population structure of seasonal gray whales in the Pacific Northwest, 1996–2015. *Paper SC/A17/GW05 presented to the Scientific Committee of the International Whaling Commission.*

Calambokidis, J., & Perez, A. (2017a). Association of PCFG gray whales on migration. *Paper SC/A17/GW02 presented to the Scientific Committee of the International Whaling Commission*.

Calambokidis, J., & Pérez, A. (2017b). Sightings and follow-up of mothers and calves in the PCFG and implications for internal recruitment. *Paper SC/A17/GW04 presented to the Scientific Committee of the International Whaling Commission*.

Carroll, E. L., Rayment, W. J., Alexander, A. M., Baker, C. S., Patenaude, N. L., Steel, D., Constantine, R., Cole, R., Boren, L. J., & Childerhouse, S. (2014). Reestablishment of former wintering grounds by New Zealand southern right whales. *Marine Mammal Science*, *30*(1), 206–220. https://doi.org/10.1111/mms.12031

Charlton, C., Ward, R., McCauley, R. D., Brownell, R. L., Jr.,

Guggenheimer, S., Salgado Kent, C. P., & Bannister, J. L. (2019). Southern right whales (*Eubalaena australis*) return to a former wintering calving ground: Fowlers Bay, South Australia. *Marine Mammal Science*, *35*(4), 1438–1462. https://doi.org/10.1111/mms.12611

Chen, K.-Y., Marschall, E. A., Sovic, M. G., Fries, A. C., Gibbs, H. L., & Ludsin, S. A. (2018). assignPOP: An R package for population assignment using genetic, non-genetic, or integrated data in a machine-learning framework. *Methods in Ecology and Evolution*, *9*, 439–446. https://doi.org/10.1111/2041-210X.12897

Chivers, S. J., Leduc, R. G., Robertson, K. M., Barros, N. B., & Dizon, A. E. (2005). Genetic variation of *Kogia* spp. with preliminary evidence for two species of *Kogia sima*. *Marine Mammal Science*, *21*(4), 619–634. https://doi.org/10.1111/j.1748-7692.2005.tb01255.x

Cooke, J. G. (2017). Updated assessment of the Sakhalin gray whale population and its relationship to gray whales in other areas. *Paper WGWAP 18/24 presented to the Western Gray Whale Advisory Panel on 15-17 November 2017*.

Cooke, J. G. (2018). *Western Gray Whale population assessment update with reference to historic range and recovery prospects*. Western Gray Whale Advisory Panel 19th Meeting,

November 14-16, 2018.

Cooke, J. G., Shpak, O. V., Meschersky, I. G., Burdin, A. M., Maclean, S. A., Chichkina, A. N., Paramonov, A. Y., Rozhnov, V. V., & Brownell, R. L., Jr. (2017a). Updated estimates of population and trend for Okhotsk Sea bowhead whales. *Paper SC/67a/NH10 presented to the Scientific Committee of the International Whaling Commission.*

Cooke, J. G., Sychenko, O., Burdin, A. M., Weller, D. W., Bradford, A. L., Lang, A. R., Brownell, R. L., Jr. (2019). Population assessment update for Sakhalin gray whales. *Paper SC/68a/CMP21 presented to the Scientific Committee of the International Whaling Commission.*

Cooke, J. G., Taylor, B. L., Reeves, R., Brownell, R. L., Jr. (2018). *Eschrichtius robustus* western subpopulation. *IUCN Red List of Threatened Species, 2018*, e.T8099A50345475. http://dx.doi.org/10.2305/IUCN.UK.2018-2.RLTS.T8099A50345475.en

Cooke, J. G., Weller, D. W., Bradford, A. L., Sychenko, O., Burdin, A. M., & Brownell, R. L., Jr. (2013). Population assessment of the Sakhalin gray whale aggregation. Paper SC/65/BRG27 presented to the Scientific Committee of the International Whaling Commission.

Cooke, J. G., Weller, D. W., Bradford, A. L., Sychenko, O.,

Burdin, A. M., Lang, A. R., & Brownell, R. L., Jr. (2017b). Population assessment update for Sakhalin gray whales, with reference to stock identity. *Paper SC/67a/NH11 presented to the Scientific Committee of the International Whaling Commission.*

Darling, J. D. (1984). Gray whales (*Eschrichtius robustus*) off Vancouver Island, British Columbia. In M. L. Jones, S. L. Swartz, & S. Leatherwood (Eds.), *The gray whale* (pp 267–287). Academic Press.

Darriba, D., Taboada, G. L., Doallo, R., Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, *9*(8), 772. https://doi.org/10.1038/nmeth.2109

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NEESTIMATOR v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, *14*(1), 209–214. https://doi.org/10.1111/1755-0998.12157

Durban, J., Weller, D., Lang, A., & Perryman, W. (2015). Estimating gray whale abundance from shore-based counts using a multilevel Bayesian model. *Journal of Cetacean Research & Management*, *15*, 61–68.

Durban, J. W., Weller, D. W., Perryman, W. L. (2017). Gray whale abundance estimates from shore-based counts off California in 2014/15 and 2015/2016. *Paper SC/A17/GW6 presented to the Scientific Committee of the International Whaling Commission.*

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359-361.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fogelqvist, J., Niittyvuopio, A., Agren, J., Savolainen, O., & Lascoux, M. (2010). Cryptic population genetic structure: the number of inferred clusters depends on sample size. *Molecular Ecology Resources*, *10*, 314–323. https://doi.org/10.1111/j.1755-0998.2009.02756.x

Frantz, A. C., Cellina, S., Krier, A., Schley, L., & Burke, T. (2009). Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *Journal of Applied Ecology*, *46*(2), 493–505. https://doi.org/10.1111/j.1365-

2664.2008.01606.x

Gilbert, K. J., & Whitlock, M. C. (2015). Evaluating methods for estimating local effective population size with and without migration. *Evolution*, *69*(8), 2154–2166. https://doi.org/10.1111/evo.12713

Gilmore, R. M. (1960). A census of the California gray whale. *U. S. Fish and Wildlife Service Special Scientific Report*, *342*, 1–30.

Goslee, S., & Urban, D. (2020). Package 'ecodist'. https://cran.r-project.org/web/packages/ecodist/ecodist.pdf

Greenhorn, J. E., Bowman, J., Wilson, P. J. (2018). Genetic monitoring suggests increasing structure following recolonization by fishers. *Journal of Wildlife Management*, *82*(7), 1403–1416. https://doi.org/10.1002/jwmg.21495

Guazzo, R. A., Helbe, T. A., D'Spain, G. L., Weller, D. W., Wiggins, S. M., & Hildebrand, J. A. (2017). Migratory behavior of eastern North Pacific gray whales tracked using a hydrophone array. *PLoS ONE*, *12*(10), e0185585. https://doi.org/10.1371/journal.pone.0185585

Guo, S. W., & Thompson, E. A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, *48*, 361–372. https://doi.org/10.2307/2532296

Hagen, S. B., Kopatz, A., Aspi, J., Kojola, I., & Eiken, H. G.

(2015). Evidence of rapid change in genetic structure and diversity during range expansion in a recovering large terrestrial carnivore. *Proceedings of the Royal Society of London B: Biological Sciences*, *282*(1807), 20150092. https://doi.org/10.1098/rspb.2015.0092

Hatler, D. F., & Darling, J. D. (1974). Recent observations of the gray whale *Eschrichtius robustus* in British Columbia. *Canadian Field-Naturalist*, *88*, 449–460.

Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution*, *59*(8), 1633–1638. https://doi.org/10.1111/j.0014-3820.2005.tb01814.x

Henderson, D. A. (1984). Nineteenth century gray whaling: Grounds, catches and kills, practices and depletions of the whale population. In M. L. Jones, S. L. Swartz, & S. Leatherwood (Eds.), *The gray whale* Eschrichtius robustus (pp. 159–186). Academic Press.

Hill, W. G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetical Research*, *38*(3), 209–216. https://doi.org/10.1017/S0016672300020553

International Whaling Commission. (2011). Report of the Scientific Committee, Annex E: Report of the Standing Working Group on the Aboriginal Whaling Management Plan. *Journal of Cetacean Research and Management*, *12*(Suppl.),

143–167.

International Whaling Commission. (2018). Report of the Fourth Rangewide Workshop on the Status of North Pacific Gray Whales. *Journal of Cetacean Research & Management*, *19*, (Suppl.), 521–536.

International Whaling Commission (2019). Report of the Fifth Rangewide Workshop on the Status of North Pacific Gray Whales. *Journal of Cetacean Research & Management*, *20*(Suppl.), 569–599.

Jackson, J. A., Steel, D. J., Beerli, P., Congdon, B. C., Olavarria, C., Leslie, M. S., Pomilla, C., Rosenbaum, H., & Baker, C. S. (2014). Global diversity and oceanic divergence of humpback whales (*Megaptera novaeangliae*). *Proceedings of the Royal Society B: Biological Sciences*, *281*(1786), 20133222. https://doi.org/10.1098/rspb.2013.3222

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), 1. https://doi.org/10.1186/1471-2156-11-94

Jombart, T., Devillard, S., Dufour, A., & Pontier, D. (2008). Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity*, *101*(1), 92–103. https://doi.org/10.1038/hdy.2008.34

Jones, M. L. (1990). Reproductive cycle in gray whales based on photographic resightings of females on the breeding grounds from 1977 – 1982. *Reports of the International Whaling Commission, Special Issue 12*, 177–182.

Kato, H., & Kasuya, T. (2002). Some analyses on the modern whaling catch history of the western North Pacific stock of gray whales (*Eschrichtius robustus*), with special reference to the Ulsan whaling ground1. *Journal of Cetacean Research & Management*, *4*(3), 277–282.

Katoh, K., Misawa, K., Kuma, K.-I., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*(14), 3059–3066. https://doi.org/10.1093/nar/gkf436

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*, 1647–1649.

https://doi.org/10.1093/bioinformatics/bts199

Keenan, K., Mcginnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, *4*(8), 782–788. https://doi.org/10.1111/2041-210X.12067

Lang, A. R. (2010). *The population genetics of gray whales* (Eschrichtius robustus) *in the North Pacific* [Doctoral dissertation]. University of California San Diego.

Lang, A. R., Calambokidis, J., Scordino, J., Pease, V. L., Klimek, A., Burkanov, V. N., Gearin, P., Litovka, D. I., Robertson, K. M., Mate, B. R., Jacobsen, J. K., & Taylor, B. L. (2014). Assessment of genetic structure among eastern North Pacific gray whales on their feeding grounds. *Marine Mammal Science*, *30*(4), 1473–1493. https://doi.org/10.1111/mms.12129

Latch, E. K., Dharmarajan, G., Glaubitz, J. C., & Rhodes, O. E. (2006). Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics*, 7, 295–302. https://doi/org/10.1007/s10592-005-9098-1

Lawson, D. J., Van Dorp, L., & Falush, D. (2018). A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nature Communications*, *9*(1), 1–11. https://doi.org/10.1038/s41467-018-05257-7

Leduc, R., Dizon, A., Burdin, A., Blokhin, S., George, J., & Brownell, R. (2005). Genetic analyses (mtDNA and microsatellites) of Okhotsk and Bering/Chukchi/Beaufort Seas populations of bowhead whales. *Journal of Cetacean Research and Management*, 7(2), 107–111.

LeDuc, R., Taylor, B., Martien, K., Robertson, K., Pitman, R., Salinas, J., Burdin, A., Kennedy, A., Wade, P., & Clapham, P. (2012). Genetic analysis of right whales in the eastern North Pacific confirms severe extirpation risk. *Endangered Species Research*, *18*(2), 163–167. https://doi.org/10.3354/esr00440

LeDuc, R. G., Weller, D. W., Hyde, J., Burdin, A. M., Rosel, P. E., Brownell, R. L. J., Würsig, B., & Dizon, A. E. (2002). Genetic differences between western and eastern gray whales (*Eschrichtius robustus*). *Journal of Cetacean Research and Management*, *4*, 1–5.

Leigh, J. W., & Bryant, D. (2015). popart: full feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116. https://doi.org/10.1111/2041-

210X.12410

Lowry, L. F., Burkanov, V. N., Altukhov, A., Weller, D. W., & Reeves, R. R. (2018). Entanglement risk to western gray whales from commercial fisheries in the Russian Far East. *Endangered Species Research*, *37*, 133–148. https://doi.org/10.3354/esr00914

Martínez A., S., Swartz, S., Urbán R., J., & Rosales-Nanduca, H. (2016). The age of living gray whales (*Eschrichtius robustus*) estimated from photographic identification data. *Paper SC/66b/BRG20 presented to the Scientific Committee of the International Whaling Commission.*

Mate, B. R., Ilyashencko, V. Y., Bradford, A. L., Vertyankin, V. V., Tsidulko, G. A., Rozhnov, V. V., & Irvine, L. M. (2015). Critically endangered western gray whales migrate to the eastern North Pacific. *Biology Letters*, *11*(4), 20150071. https://doi.org/10.1098/rsbl.2015.0071

Meier, S. K., Yazvenko, S. B., Blokhin, S. A., Wainwright, P., Maminov, M. K., Yakovlev, Y. M., & Newcomer, M. W. (2007). Distribution and abundance of western gray whales off northeastern Sakhalin Island, Russia, 2001–2003. *Environmental Monitoring and Assessment*, *134*(1), 107–136. https://doi.org/10.1007/s10661-007-9811-2

Meirmans, P. G. (2012). The trouble with isolation by distance.

Molecular Ecology, *21*(12), 2839–2846. https://doi.org/10.1111/j.1365-294X.2012.05578.x

Meschersky, I. G., Kuleshova, M. A., Litovka, D. I., Burkanov, V. N., Andrews, R. D., Tsidulko, G. A., Ilyashencko, V. Y., & Rozhnov, V. V. (2012). *Mitochondrial lines composition of gray whale* (Eschrichtius robustus) *in Russian Far Eastern seas: the control region and protein-coding fragments*. Seventh International Conference on Marine Mammals of the Holarctic, Suzdal, Russia.

Meschersky, I., Kuleshova, M., Litovka, D., Burkanov, V., Andrews, R., Tsidulko, G., Rozhnov, V., & Ilyashenko, V. Y. (2015). Occurrence and distribution of mitochondrial lineages of gray whales (*Eschrichtius robustus*) in Russian Far Eastern seas. *Biology Bulletin*, *42*(1), 34–42.

Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out protocol for extracting DNA from human nucleated cells. *Nucleic Acid Research*, *16*(3), 1215. https://doi.org/10.1093/nar/16.3.1215

Mizue, K. (1951). Gray whales in the east sea area of Korea. *Scientific Reports of the Whale Research Institute, Tokyo*, *5*, 71–79.

Morin, P. A., Archer, F. I., Pease, V. L., Hancock-Hanser, B. L., Robertson, K. M., Huebinger, R. M., Martien, K. K.,

Bickham, J. W., George, J. C., Postma, L. D., & Taylor, B. L. (2012). Empirical comparison of single nucleotide polymorphisms and microsatellites for population and demographic analyses of bowhead whales. *Endangered Species Research*, *19*, 129–147. https://doi.org/10.3354/esr00459

Morin, P. A., Martien, K. K., Archer, F. I., Cipriano, F., Steel, D., Jackson, J., & Taylor, B. L. (2010). Applied conservation genetics and the need for quality control and reporting of genetic data used in fisheries and wildlife management. *Journal of Heredity*, *101*(1), 1–10. https://doi.org/10.1093/jhered/esp107

Morin, P. A., Nestler, A., Rubio-Cisneros, N. T., Robertson, K. M., & Mesnick, S. L. (2005). Interfamilial characterization of a region of the ZFX and ZFY genes facilitates sex determination in cetaceans and other mammals. *Molecular Ecology*, *14*(10), 3275–3286. https://doi.org/10.1111/j.1365-294X.2005.02651.x

Nakamura, G., Yoshida, H., Morita, H., Konishi, K., Mogoe, T., Miyashita, T., Kamizawa, Y., & Kato, H. (2018). Status report of conservation and researches on the western North Pacific gray whales in Japan, May 2017 - April 2018. *Paper SC/67b/CMP12 presented to the Scientific Committee of the International Whaling Commission.*

Nakamura, G., Yoshida, H., Terai, R., Konishi, K., Isodo, T., Nishiwaki, S., Taru, H., Suzuki, S., Kato, H. (2019). Status report of conservation and researches on the western North Pacific gray whales in Japan, May 2018 - April 2019. *Paper SC/68a/CMP2 presented to the Scientific Committee of the International Whaling Commission.*

Nambu, H., Ishikawa, H., & Yamada, T. K. (2010). Records of the western gray whale *Eschrichtius robustus*: its distribution and migration. *Japan Cetology*, *20*, 21–29. https://doi.org/10.5181/cetology.0.20_21

Nei, M., Li, W. (1973). Linkage disequilibrium in subdivided populations. *Genetics*, *75*(1), 213–219.

Noad, M. J., Kniest, E., & Dunlop, R. A. (2019). Boom to bust? Implications for the continued rapid growth of the eastern Australian humpback whale population despite recovery. *Population Ecology*, *61*(2), 198–209. https://doi.org/10.1002/1438-390X.1014

Paetkau, D., & Strobeck, C. (1994). Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology*, *3*(5), 489–495. https://doi.org/10.1111/j.1365-294X.1994.tb00127.x

Palsbøll, P. J., Allen, J., Berube, M., Clapham, P. J., Feddersen, T. P., Hammond, P. S., Hudson, R. R., Jorgensen,

H., Katona, S., Larsen, A. H., Larsen, F., Lien, J., Mattila, D. K., Sigurjonsson, J., Sears, R., Smith, T., Sponer, R., Stevick, P., & Oien, N. (1997). Genetic tagging of humpback whales. *Nature*, *388*(6644), 767–769. https://doi.org/10.1038/42005

Palsbøll, P. J., Peery, M. Z., Olsen, M. T., Beissinger, S. R., & Bérubé, M. (2013). Inferring recent historic abundance from current genetic diversity. *Molecular Ecology*, *22*(1), 22–40. https://doi.org/10.1111/mec.12094

Palumbi, S. R., Martin, A. P., Romero, S., Mcmillan, W. O., Stice, L., & Grawboski, G. (1991). *The simple fool's guide to PCR version 2.0*. University of Hawaii.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, *28*, 2537-2539. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Perez, M. F., Franco, F. F., Bombonato, J. R., Bonatelli, I. a. S., Khan, G., Romeiro-Brito, M., Fegies, A. C., Ribeiro, P. M., Silva, G. a. R., & Moraes, E. M. (2018). Assessing population structure in the face of isolation by distance: Are we neglecting the problem? *Diversity and Distributions*, *24*(12), 1883–1889. https://doi.org/10.1111/ddi.12816

Pew, J., Muir, P. H., Wang, J., & Frasier, T. R. (2015).

related: an R package for analysing pairwise relatedness from codominant molecular markers. *Molecular Ecology Resources*, *15*(3), 557–561. https://doi.org/10.1111/1755-0998.12323

Pike, G. C. (1962). Migration and feeding of the gray whale (*Eschrichtius gibbosus*). *Journal of the Fisheries Research Board of Canada*, *19*(5), 815–838. https://doi.org/10.1139/f62-051

Pomilla, C., & Rosenbaum, H. C. (2005). Against the current: an inter-oceanic whale migration event. *Biology Letters*, *1*(4), 476–479. https://doi.org/10.1098/rsbl.2005.0351

Posada, D. (2008). jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, *25*(7), 1253–1256. https://doi.org/10.1093/molbev/msn083

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

Puechmaille, S. J. (2016). The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular Ecology Resources*, *16*(3), 608–627. https://doi.org/10.1111/1755-0998.12512

Queller, D. C., & Goodnight, K. F. (1989). Estimating relatedness using genetic markers. *Evolution*, *43*(2), 258–275. https://doi.org/10.1111/j.1558-5646.1989.tb04226.x

Reeves, R. R. (1984). Modern commercial pelagic whaling for gray whales. In M. L. Jones, S. L. Swartz, & S. Leatherwood (Eds.), *The gray whale* Eschrichtius robustus (pp. 187–200). Academic Press.

Reilly, S. B., Bannister, J. L., Best, P. B., Brown, M., Al., E. (2008). *Eschrichtius robustus* (western subpopulation). *IUCN Red List of Threatened Species* (Version 2011.1). https://www.iucnredlist.org/details/8097/0

Rice, D. W., & Wolman, A. A. (1971). The life history and ecology of the gray whale (*Eschrichtius robustus*). *The American Society of Mammalogists, Special Publication No. 3*.

Richard, K. R., Dillon, M. C., Whitehead, H., & Wright, J. M. (1996). Patterns of kinship in groups of free-living sperm whales (*Physeter macrocephalus*) revealed by multiple molecular genetic analyses. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(16), 8792–8795. https://doi.org/10.1073/pnas.93.16.8792

Robinson, J. D., & Moyer, G. R. (2013). Linkage disequilibrium and effective population size when generations overlap.

Evolutionary Applications, *6*(2), 290–302. https://doi.org/10.1111/j.1752-4571.2012.00289.x

Rocha, R. C., Clapham , P. J., & Ivashchenko, Y. V. (2014). Emptying the oceans: a summary of industrial whaling catches in the 20th century. *Marine Fisheries Review*, *76*(4), 37–48. doi: dx.doi.org/10.7755/MFR.76.4.3

Rodríguez-Ramilo, S. T., & Wang, J. (2012). The effect of close relatives on unsupervised Bayesian clustering algorithms in population genetic structure analysis. *Molecular Ecology Resources*, *12*(5), 873–884. https://doi.org/10.1111/j.1755-0998.2012.03156.x

Roman, J., & Palumbi, S. R. (2003). Whales before whaling in the North Atlantic. *Science*, *301*(5632), 508–510. https://doi.org/10.1126/science.1084524

Rosel, P. E., Dizon, A. E., & Haygood, M. G. (1995). Variability of the mitochondrial control region in populations of the harbor porpoise, *Phocoena phocoena*, on interoceanic and regional scales. *Canadian Journal of Fisheries and Aquatic Sciences*, *52*(6), 1210–1219. https://doi.org/10.1139/f95-118

Rousset, F. (2008). Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, *8*(1), 103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Rousset, F., & Raymond, M. (1995). Testing heterozygote excess and deficiency. *Genetics*, *140*(4), 1413–1419. https://doi.org/10.1093/genetics/140.4.1413

Roux, J.-P., Braby, R. J., & Best, P. B. (2015). Does disappearance mean extirpation? The case of right whales off Namibia. *Marine Mammal Science*, *31*(3), 1132–1152. https://doi.org/10.1111/mms.12213

Ruegg, K. C., Anderson, E. C., Baker, C. S., Vant, M., Jackson, J. A., & Palumbi, S. R. (2010). Are Antarctic minke whales unusually abundant because of 20th century whaling? *Molecular Ecology*, *19*(2), 281–291. https://doi.org/10.1111/j.1365-294X.2009.04447.x

Ruegg, K., Rosenbaum, H. C., Anderson, E. C., Engel, M., Rothschild, A., Baker, C. S., & Palumbi, S. R. (2013). Long-term population size of the North Atlantic humpback whale within the context of worldwide population structure. *Conservation Genetics*, *14*(1), 103–114. https://doi.org/10.1007/s10592-012-0432-0

Rugh, D. J., Sheldon, K. E. W., & Shulman-Janiger, A. (2001). Timing of the gray whale southbound migration. *Journal of Cetacean Research and Management*, *3*, 31–39.

Saiki, R. K., Gelfand, D. H., Stoffle, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., & Erlich, H. A.

(1988). Primer-directed amplification of DNA with a thermostable DNA polymerase. *Science*, *239*(4839), 487–491. https://doi.org/10.1126/science.239.4839.487

Salden, D. R., Herman, L. M., Yamaguchi, M., & Sato, F. (1999). Multiple visits of individual humpback whales (*Megaptera novaeangliae)* between the Hawaiian and Japanese winter grounds. *Canadian Journal of Zoology*, 77(3), 504–508. https://doi.org/10.1139/z99-005

Sambrook, J., Fritsch, E., & Maniatis, T. (1989). *Molecular cloning: a laboratory manual* (2nd ed.). Cold Spring Harbor Laboratory Press.

Schwartz, M. K., McKelvey, K. S. (2009). Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*, *10*(2), 441. https://doi.org/10.1007/s10592-008-9622-1

Sinnock, P. (1975). The Wahlund effect for the two-locus model. *American Naturalist*, *109*(969), 565–570. https://doi.org/10.1086/283027

Slatkin, M. (1994). Linkage disequilibrium in growing and stable populations. *Genetics*, *137*(1), 331–336.

Stevick, P. T., Allen, J., Berube, M., Clapham, P. J., Katona, S. K., Larsen, F., Lien, J., Mattila, D. K., Palsbøll, P. J., Robbins, J., Sigurjonsson, J., Smith, T. D., Oien, N.,

& Hammond, P. S. (2003). Segregation of migration by feeding ground origin in North Atlantic humpback whales (*Megaptera novaeangliae*). *Journal of Zoology*, *259*(3), 231–237. https://doi.org/10.1017/s0952836902003151

Stevick, P. T., Berrow, S. D., Bérubé, M., Bouveret, L., Broms, F., Jann, B., Kennedy, A., López Suárez, P., Meunier, M., Ryan, C., & Wenzel, F. (2016). There and back again: multiple and return exchange of humpback whales between breeding habitats separated by an ocean basin. *Journal of the Marine Biological Association of the United Kingdom*, *96*(4), 885–890. https://doi.org/10.1017/S0025315416000321

Stevick, P. T., Neves, M. C., Johansen, F., Engel, M. H., Allen, J., Marcondes, M. C., & Carlson, C. (2011). A quarter of a world away: female humpback whale moves 10 000 km between breeding areas. *Biology Letters*, 7(2), 299–302. https://doi.org/10.1098/rsbl.2010.0717

Swartz, S. L., Taylor, B. L., & Rugh, D. J. (2006). Gray whale *Eschrichtius robustus* population and stock identity. *Mammal Review*, 36(1), 66–84. https://doi.org/10.1111/j.1365-2907.2006.00082.x

Thomas, P. O., Reeves, R. R., Brownell Jr, R. L. (2016). Status of the world's baleen whales. *Marine Mammal Science*, *32*(2), 682–734. https://doi.org/10.1111/mms.12281

Tyurneva, O. Y., Yakovlev, Y. M., Vertyankin, V. V., & Selin, N. I. (2010). The peculiarities of foraging migrations of the Korean-Okhotsk gray whale (*Eschrichtius robustus*) population in Russian waters of the Far Eastern seas. *Russian Journal of Marine Biology*, *36*(2), 117–124. https://doi.org/10.1134/s1063074010020069

Urban R., J., Rojas-Bracho, L., Perez-Cortes, H., Gomez-Gallardo, A., Swartz, S. L., Ludwig, S., & Brownell, R. L., Jr. (2003). A review of gray whales (*Eschrichtius robustus*) on their wintering grounds in Mexican waters. *Journal of Cetacean Research and Management*, *5*, 281–295.

Urbán R., J., Weller, D. W., Martínez A., S., Tyurneva, O., Bradford, A., Burdin, A. M., Lang, A. R., Swartz, S., Sychenko, O., Viloria-Gomora, L., & Yakovlev, Y. (2019). New information on the gray whale migratory movements between the western and eastern North Pacific. *Paper SC/68a/CMP11 presented to the Scientific Committee of the International Whaling Commission.*

Urbán, R., J., Weller, D. W., Tyurneva, O., Swartz, S., Bradford, A. L., Yakovlev, Y., Sychenko, O., Rosales N., H., Martinez A., S., Burdin, A. M., & Gomez-Gallardo Unzueta, A. (2012). Report on the photographic comparison of the western and Mexican gray whale catalogues. *Paper*

SC/64/BRG13 presented to the Scientific Committee of the International Whaling Commission.

Urbán R., J., Weller, D., Tyurneva, O., Swartz, S., Bradford, A., Yakovlev, Y., Sychenko, O., Rosales N., H., Martínez A., S., Burdin, A. M., & Gómez-Gallardo U., A. (2013). Report on the photographic comparison of the Sakhalin Island and Kamchatka Peninsula with the Mexican gray whale catalogues. *Paper SC/65a/BRG04 presented to the Scientific Committee of the International Whaling Commission.*

U.S. Federal Register. (1994). Endangered and threatened wildlife and plants; Remove the Eastern North Pacific Population of the Gray Whale from the List of Endangered Wildlife. Final rule. *Federal Register*, *59*(115), 31094–31095 (16 June 1994). Fish & Wildlife Service, Department of the Interior and National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Department of Commerce.

Valsecchi, E., & Amos, W. (1996). Microsatellite markers for the study of cetacean populations. *Molecular Ecology*, *5*(1), 151–156. https:/doi.org/10.1111/j.1365-294X.1996.tb00301.x

Vertyankin, V. V., Nikulin, V. C., Bednykh, A. M., & Kononov, A. P. (2004). Sighting of gray whales (*Eschrichtius robustus*) near southern Kamchatka. In M. M. Belkovivh (Ed.), *Marine*

mammals of the Holarctic (pp. 126–128). KMK Scientific Press.

Wade, P. R., Kennedy, A., Leduc, R., Barlow, J., Carretta, J., Shelden, K., Perryman, W., Pitman, R., Robertson, K., Rone, B., Salinas, J. C., Zerbini, A., Brownell, R. L., Jr., & Clapham, P. J. (2011). The world's smallest whale population? *Biology Letters*, 7(1), 83–85. https://doi.org/10.1098/rsbl.2010.0477

Wahlund, S. (1928). Zusammensetzung von populationen und korrelation-serscheiunungen von standpunkt der vererbungslehre aus betrachtet [Composition of populations and correlation phenomena considered from the position of heredity theory]. *Hereditas*, *11*, 65–106.

Waits, L. P., Luikart, G., & Taberlet, P. (2001). Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*, *10*(1), 249–256. https://doi.org/10.1046/j.1365-294X.2001.01185.x

Waldick, R. C., Brown, M. R., & White, B. N. (1999). Characterization and isolation of microsatellite loci from the endangered North Atlantic right whale. *Molecular Ecology*, *8*(10), 1763–1765.

Wang, J. (2016). A comparison of single sample estimators of

effective population sizes from genetic marker data. *Molecular Ecology*, *25*(19), 4692–4711. https://doi.org/10.1111/mec.13725

Wang, J. (2018). Effects of sampling close relatives on some elementary population genetics analyses. *Molecular Ecology Resources*, *18*(1), 41–54. https://doi.org/10.1111/1755-0998.12708

Wang, P. L. (1985). *Distribution of cetaceans in Chinese waters*. (NOAA Administrative Report LJ-85-24). U.S. Department of Commerce.

Wang, X., Xu, M., Wu, F., Weller, D. W., Miao, X., Lang, A. R., & Zhu, Q. (2015). Insights from a gray whale (*Eschrichtius robustus*) bycaught in the Taiwan Strait off China in 2011. *Aquatic Mammals*, *41*(3), 327–332. https://doi.org/10.1578/AM.41.3.2015.327

Waples, R. S. (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, 7(2), 167–184. https://doi.org/10.1007/s10592-005-9100-y

Waples, R. S. (2015). Testing for Hardy–Weinberg proportions: Have we lost the plot? *Journal of Heredity*, *106*(1), 1–19. https://doi.org/10.1093/jhered/esu062

Waples, R. S., & Anderson, E. C. (2017). Purging putative

siblings from population genetic data sets: a cautionary view. *Molecular Ecology*, *26*(5), 1211–1224. https://doi.org/10.1111/mec.14022

Waples, R. S., Antao, T., & Luikart, G. (2014). Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics*, *197*(2), 769–780. https://doi.org/10.1534/genetics.114.164822

Waples, R. S., & Do, C. (2008). LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, *8*(4), 753–956. https://doi.org/10.1111/j.1755-0998.2007.02061.x

Waples, R. S., & Do, C. (2010). Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, *3*(3), 244–262. https://doi.org/10.1111/j.1752-4571.2009.00104.x

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*, 1358–1370. https://doi.org/10.1146/annurev.genet.36.050802.093940

Weller, D. W., Bradford, A. L., Burdin, A. M., Miyashita, A. M., Kariya, T., Trukhin, A. M., Maclean, S. A., Vladimirov, V. A., & Doroshenko, N. V. (2002a). Photographic recaptures of

western gray whales in the Okhotsk Sea. *Paper SC/54/BRG13 presented to the Scientific Committee of the International Whaling Commission.*

Weller, D. W., Bradford, A. L., Kato, H., Bando, T., Otani, S., Burdin, A. M., & Brownell, R. L., Jr. (2008). A photographic match of a western gray whale between Sakhalin Island, Russia, and Honshu, Japan: the first link between the feeding ground and a migratory corridor. *Journal of Cetacean Research & Management*, *10*, 89–91.

Weller, D. W., Burdin, A. M., Würsig, B., Taylor, B. L., & Brownell, R. L., Jr. (2002b). The western Pacific gray whale: A review of past exploitation, current status, and potential threats. *Journal of Cetacean Research and Management*, *4*, 7–12.

Weller, D. W., Burdin, A. M., Würsig, B., Taylor, B. L., & Brownell, R. L., Jr. (2003). Summer sightings of western gray whales in the Okhotsk and western Bering Seas. *Paper SC/55/BRG9 presented to the Scientific Committee of the International Whaling Commission.*

Weller, D. W., Klimek, A., Bradford, A. L., Calambokidis, J., Lang, A. R., Gisborne, B., Burdin, A. M., Szaniszlo, W., Urban, J., Gomez-Gallardo Unzueta, A., Swartz, S., & Brownell, R. L., Jr. (2012). Movements of gray whales

between the western and eastern North Pacific. *Endangered Species Research*, *18*, 193–199. https://doi.org/10.3354/esr00447

Weller, D. W., Takanawa, N., Ohizumi, H., Funahashi, N., Sychenko, O., Burdin, A. M., Lang, A. R., & Brownell, R. L., Jr. (2016). Gray whale migration in the western North Pacific: further support for a Russia-Japan connection. *Paper SC/66b/BRG16 presented to the Scientific Committee of the International Whaling Commission.*

Weller, D. W., Würsig, B., Bradford, A. L., Burdin, A. M., Blokhin, S. A., Minakuchi, H., & Brownell, R. L., Jr. (1999). Gray whales (*Eschrichtius robustus*) off Sakhalin Island, Russia: Seasonal and annual patterns of occurrence. *Marine Mammal Science*, *15*(4), 1208–1227. https://doi.org/10.1111/j.1748-7692.1999.tb00886.x

Xu, Q.-S., & Liang, Y.-Z. (2001). Monte-Carlo cross validation. *Chemometrics and Intelligent Laboratory Systems*, *56*, 1–11.

Zhao, L., Zhu, Q., Miao, X., Xu, M., Wu, F., Dai, Y., Tao, C., Mou, J., & Wang, X. (2017). An overview of cetacean strandings, bycatches and rescues along the western coast of the Taiwan Strait, China: 2010–2015. *Acta Oceanologica Sinica*, *36*(12), 31–36. https://doi.org/10.1007/s13131-017-1100-5

Zhao, Y. (1997). The grey whale stranded at the Liaoning coast in the north of the Yellow Sea. *Fisheries Science*, *16*(3), 8–10.

Zhu, Q. (2012). China gray whale. *Cetoken Newsletter*, *29*, 1–8.

Zhu, Q., & Yue, H. (1998). Strandings and sightings of the western Pacific stock of the gray whale (*Eschrichtius robustus*) in Chinese coastal waters. *Paper SC/50/AS5 presented to the Scientific Committee of the International Whaling Commission.*

TABLE 1 Measures of the genetic diversity found among strata: (a) mtDNA control region sequence diversity, including the number of individuals sequenced (*n*), the number of unique haplotypes identified (*k*), the haplotype diversity (*h*), and the nucleotide diversity (π); (b) nuclear genetic diversity based on the genotypes at 12 microsatellite loci, including the number of individuals genotyped (*n*), the proportion of individuals and loci successfully genotyped (*P*), the allelic richness (*Ar*), the expected heterozygosity (H_e), observed heterozygosity (H_o), the inbreeding coefficient, the number of private alleles (summed over all loci, *p*), relatedness (*r*) with 95% confidence intervals, and the estimate of effective population size (N_e) with 95% jackknife intervals.

(a)

Region	Strata	mtDNA			
		n	***k***	***H***	π
ENP	NFG	103	32	0.952	0.014
	Females	61	25	0.946	0.014
	Males	42	21	0.958	0.014
WNP	Kamchatka	16	9	0.883	0.020
	Sakhalin	156	22	0.760	0.017
	Noncalves only[a]	81	20	0.804	0.018
	Female noncalves	37	10	0.776	0.018
	Male noncalves	44	15	0.833	0.018
	Calves[b]	75	11	0.711	0.017
	Sakhalin only[c]	122	19	0.749	0.018
	Sakhalin-ENP[d]	34	10	0.775	0.016
	All WNP combined	166	24	0.775	0.018

(b)

Region	Strata	Microsatellites								
		n	***P***	***Ar***	***He***	***Ho***	F_{IS}	***p***	***r* [95% CI]**	N_e **(95% jackknife)**
ENP	NFG	105	0.998	6.172	0.729	0.715	0.014	13	−0.016 [−0.020, −0.011]	1,027 (263.4–∞)
	Females	62	0.997	4.517	0.729	0.729	−0.008	4	−0.010 [−0.018, −0.003]	99 (46– 985)
	Males	43	1.000	4.462	0.724	0.696	0.03	6	−0.012 [−0.023, −0.001]	∞ (82.9–∞)
WNP	Kamchatka	16	0.995	5.5	0.677	0.661	−0.01	0	−0.034 [−0.088, 0.021]	29 (12.6– 676.6)
	Sakhalin	154	0.986	5.771	0.688	0.702	−0.022	7	0.043 [0.040, 0.047]	80 (61.9– 107.7)
	noncalves only[a]	80	0.991	5.981	0.699	0.719	−0.034	3	0.023 [0.017, 0.030]	104 (58.1– 276.6)
	Female noncalves	36	0.995	6.704	0.686	0.694	−0.018	3	0.049 [0.035, 0.064]	58 (21.3–∞)
	Male noncalves	41	0.995	6.971	0.704	0.734	−0.056	0	0.017 [0.006, 0.028]	82 (39.2– 570.0)
	Calves[b]	74	0.981	5.53	0.676	0.684	−0.016	2	0.064 [0.058, 0.071]	85 (58.7– 136.6)
	Sakhalin only[c]	122	0.986	5.67	0.685	0.697	−0.019	2	0.049 [0.045, 0.053]	70 (44.4– 121.9)
	Sakhalin-ENP[d]	34	0.985	6.12	0.702	0.692	−0.045	2	0.016 [0.001, 0.032]	51 (26.5– 179.8)
	All WNP combined	164	0.987	8.333	0.690	0.701	−0.017	7	0.040 [0.037, 0.043]	82 (64.3– 107.4)

[a] Sakhalin whales that were >1 year old when they were first photographically identified.

[b] Sakhalin whales that were first photographically identified as calves (whales <1 year old).

[c] Whales that are known to utilize the Sakhalin feeding ground but have not been recorded in the eastern North Pacific.

[d] Whales that are known to utilize the Sakhalin feeding ground and have also been identified on the ENP migratory route and/or wintering ground.

TABLE 2 Results of pairwise comparisons across strata using (a) mtDNA control region sequences and (b) 12 microsatellite loci.

(a)

Comparison	χ^2 *p*-value	F_{ST} (*p*-value)	ϕ_{ST} (*p*-value)
Sakhalin (*n* = 156) vs. Kamchatka (*n* = 16)	0.100	0.001 (.355)	−0.001 (.369)
NFG (*n* = 103) vs. Kamchatka (*n* = 16)	0.253	0.027 (.026)	0.020 (.150)
NFG (*n* = 103) vs. Sakhalin all (*n* = 156)	0.000	0.093 (<.001)	0.090 (<.001)
NFG (*n* = 103) vs. Sakhalin noncalves[a] (*n* = 81)	0.000	0.064 (<.001)	0.058 (.001)
NFG (*n* = 103) vs. Sakhalin calves[b] (*n* = 75)	0.000	0.116 (<.001)	0.069 (.001)
NFG females (*n* = 61) vs. Sakhalin noncalf females (*n* = 37)	0.003	0.069 (.001)	0.045 (.014)
NFG males (*n* = 42) vs. Sakhalin noncalf males (*n* = 44)	0.001	0.060 (.001)	0.072 (.002)
NFG (*n* = 103) vs. Sakhalin only[c] (*n* = 122)	0.001	0.100 (.001)	0.141 (.001)
NFG (*n* = 103) vs. Sakhalin-ENP[d] (*n* = 34)	0.003	0.073 (.001)	0.082 (.001)
Sakhalin only[c] (*n* = 122) vs. Sakhalin-ENP[d] (*n* = 34)	0.126	0.021 (.051)	0.017 (.131)
Sakhalin noncalves[b] (*n* = 81) vs. Sakhalin calves[a] (*n* = 75)	0.237	0.002 (.291)	0.000 (.341)

(b)

Comparison	χ^2 *p*-value	F_{ST} (*p*-value)	F'_{ST} (*p*-value)

Sakhalin (*n* = 156) vs. Kamchatka (*n* = 16)	0.723	0.001 (.348)	0.004 (.35)
NFG (*n* = 105) vs. Kamchatka (*n* = 16)	0.009	0.015 (.003)	0.051 (.003)
NFG (*n* = 105) vs. Sakhalin all (*n* = 156)	0.000	0.016 (<.001)	0.057 (<.001)
NFG (*n* = 105) vs. Sakhalin noncalves[a] (*n* = 81)	0.000	0.012 (<.001)	0.042 (<.001)
NFG (*n* = 105) vs. Sakhalin calves[b] (*n* = 75)	0.000	0.021 (<.001)	0.070 (<.001)
NFG females (*n* = 62) vs. Sakhalin noncalf females (*n* = 37)	0.000	0.027 (<.001)	0.095 (<.001)
NFG males (*n* = 43) vs. Sakhalin noncalf males (*n* = 44)	0.003	0.008 (.009)	0.028 (.009)
NFG (*n* = 105) vs. Sakhalin only[c] (*n* = 122)	0.001	0.018 (.001)	0.062 (.001)
NFG (*n* = 105) vs. Sakhalin-ENP[d] (*n* = 34)	0.001	0.008 (.004)	0.028 (.004)
Sakhalin only[c](*n* = 122) vs. Sakhalin-ENP[d] (*n* = 34)	0.368	−0.002 (.824)	−0.007 (.828)
Sakhalin noncalves[b] (*n* = 81) vs. Sakhalin calves[a] (*n* = 75)	0.319	0.000 (.471)	0.000 (.471)

[a] Sakhalin whales that were >1 year old when they were first photographically identified.

[b] Sakhalin whales that were first photographically identified as calves (whales <1 year old).

[c] Whales that are known to utilize the Sakhalin feeding ground but have not been recorded in the eastern North Pacific.

[d] Whales that are known to utilize the Sakhalin feeding ground and have also been

identified on the ENP migratory route and/or wintering ground.

[5048]-108

TABLE 3 Results of STRUCTURE clustering analysis using a model of admixture with correlated allele frequencies. No a priori information on the geographic location of sampling was included. Values in bold indicate the optimal number of clusters identified by STRUCTURE using the two criteria described in the text.

K	**Reps**	**Mean LnP(K)**	***SD* LnP(K)**	**Ln′(K)**	**Ln″(K)**	**Delta K**
1	5	−10,127.84	0.26	NA	NA	NA
2	5	**−9,973.22**	1.56	154.62	603.22	**386.81**
3	5	−10,421.82	68.57	−448.60	621.14	9.06
4	5	−10,249.28	60.64	172.54	428.96	7.07
5	5	−10,505.70	134.16	−256.42	235.22	1.75
6	5	−10,997.34	159.07	−491.64	16.90	0.11
7	5	−11,505.88	253.37	−508.54	472.72	1.87
8	5	−11,541.70	271.43	−35.82	NA	NA

FIGURE 1 Locations where samples were collected. Key areas mentioned in the text are labeled, including Utqiagvik, Alaska (UTQ); the Chukotka Peninsula, Russia (CHK); the region between Cape Navarin and Cape Olyutorskiy, Russia, south of the Koryak Mountains (KYK); the Kamchatka Peninsula, Russia (KAM); and Sakhalin Island, Russia (SI).

FIGURE 2 Median-joining network showing relationships among the mtDNA haplotypes. The numbers next to the nodes correspond to the haplotype IDs listed in Table S4. The size of the nodes is proportional to the frequencies of the haplotypes, and each node is colored to indicate the fraction of individuals with that haplotype from each stratum. The small black diamonds(unlabeled) indicate haplotypes that were inferred by the program but were not found among our samples. The length of lines connecting nodes is proportional to the inferred number of mutations separating haplotypes; hash marks are used to represent the number of mutational events.

FIGURE 3 STRUCTURE barplots based on a model of admixture with correlated allele frequencies (a) K = 2, 3, and 4 for a model with no *a priori* information on geographic location of sampling; and (b) K = 2 when information on geographic location of sampling (i.e., locprior = 1) is incorporated. Each vertical bar represents a single individual, and is

shaded based on the proportional membership (Q value) of individual whales to each of the inferred genetic clusters. Regions where samples were collected are arranged from east to west. For the Northern Feeding Ground stratum, the regions include Utquivik, AK (UTQ), the Chukotka Peninsula in Russia (CHK), and the region south of the Russian Koryak Mountains (KYK). In addition, four samples were collected from stranded whales in isolated regions of the Bering Sea, and these are grouped together under the BS group. The Western North Pacific stratum includes Kamchatka Peninsula (KAM) and Sakhalin Island (SI), Russia.

FIGURE 4 Scatter plot based on Discriminant Analysis of Principal Components with a priori information on stratum of origin incorporated and assuming that three clusters are present. Eigenvalues of the analysis are displayed in inset. Individuals are represented by dots, and the color of the dots denotes each individual's stratum of origin. Inertial ellipses encompassing 67% of individuals within each stratum are shown.

FIGURE 5 Self-assignment accuracies estimated via Monte-Carlo cross-validation and support vector machine methods with three levels of training loci and training data sets generated by random sampling of (a) all NFG and Sakhalin

samples, and (b) All NFG samples but only those Sakhalin samples representing individuals that have not been recorded in the ENP. Except where all loci were used, the training loci were selected based on those with the highest F_{ST} values. The line within the boxplot shows the median and the top and bottom edges represent the 25th and 75th percentiles. The ends of whiskers are the minimum and maximum of non-outliers, and outliers are shown as black circles. The horizontal red line indicates the null population assignment rate, which for two populations is 50%.

FIGURE 6 Membership probabilities of (a) the whales sampled off Kamchatka, and (b) the whales sampled off Sakhalin that have been recorded in the ENP. The baseline Sakhalin stratum included only those sampled Sakhalin whales that have not been recorded in the ENP. Membership probabilities were estimated using the random forest machine learning algorithm.

TABLE 1 Measures of the genetic diversity found among strata: (a) mtDNA control region sequence diversity, including the number of individuals sequenced (*n*), the number of unique haplotypes identified (*k*), the haplotype diversity (*h*), and the nucleotide diversity (π); (b) nuclear genetic diversity based on the genotypes at 12 microsatellite loci, including the number of individuals genotyped (*n*), the proportion of individuals and loci successfully genotyped (*P*), the allelic richness (*Ar*), the expected heterozygosity (H_e), observed heterozygosity (H_o), the inbreeding coefficient, the number of private alleles (summed over all loci, *p*), relatedness (*r*) with 95% confidence intervals, and the estimate of effective population size (N_e) with 95% jackknife intervals.

(a)

Region	**Strata**	**mtDNA**			
		n	***k***	***H***	π
ENP	NFG	103	32	0.952	0.014
	Females	61	25	0.946	0.014
	Males	42	21	0.958	0.014
WNP	Kamchatka	16	9	0.883	0.020
	Sakhalin	156	22	0.760	0.017

				[5048]-113
Noncalves only[a]	81	20	0.804	0.018
Female noncalves	37	10	0.776	0.018
Male noncalves	44	15	0.833	0.018
Calves[b]	75	11	0.711	0.017
Sakhalin only[c]	122	19	0.749	0.018
			Sakhalin-ENP[d]	34 10
All WNP combined	166	24	0.775	0.018

(b)

Region	Strata	Microsatellites							
		n	***P***	***Ar***	***He***	***Ho***	F_{IS}	***p***	***r* [95% CI]**
ENP	NFG	105	0.998	6.172	0.729	0.715	0.014	13	−0.016 [−0.020, −0.011]
	Females	62	0.997	4.517	0.729	0.729	−0.008	4	−0.010 [−0.018, −0.003]
	Males	43	1.000	4.462	0.724	0.696	0.03	6	−0.012 [−0.023, −0.001]
WNP	Kamchatka	16	0.995	5.5	0.677	0.661	−0.01	0	−0.034 [−0.088, 0.021]
	Sakhalin	154	0.986	5.771	0.688	0.702	−0.022	7	0.043 [0.040, 0.047]
	noncalves only[a]	80	0.991	5.981	0.699	0.719	−0.034	3	0.023 [0.017, 0.030]
	Female noncalves	36	0.995	6.704	0.686	0.694	−0.018	3	0.049 [0.035, 0.064]
	Male noncalves	41	0.995	6.971	0.704	0.734	−0.056	0	0.017 [0.006, 0.028]
	Calves[b]	74	0.981	5.53	0.676	0.684	−0.016	2	0.064 [0.058, 0.071]

Sakhalin only[c]	122	0.986	5.67	0.685	0.697	−0.019	2	0.049 [0.045, 0.053]
Sakhalin-ENP[d]	34	0.985	6.12	0.702	0.692	−0.045	2	0.016 [0.001, 0.032]
All WNP combined	164	0.987	8.333	0.690	0.701	−0.017	7	0.040 [0.037, 0.043]

[a] Sakhalin whales that were >1-year old when they were first photographically identified.

[b] Sakhalin whales that were first photographically identified as calves (whales <1-year old).

[c] Whales that are known to utilize the Sakhalin feeding ground but have not been recorded in the eastern North Pacific.

[d] Whales that are known to utilize the Sakhalin feeding ground and have also been identified on the ENP migratory route and/or wintering ground.

TABLE 2 Results of pairwise comparisons across strata using (a) mtDNA control region sequences and (b) 12 microsatellite loci.

(a)

Comparison	χ^2 *p*-value	F_{ST} (*p*-value)	Φ_{ST} (*p*-value)
Sakhalin (*n* = 156) vs. Kamchatka (*n* = 16)	0.100	0.001 (.355)	−0.001 (.369)
NFG (*n* = 103) vs. Kamchatka (*n* = 16)	0.253	0.027 (.026)	0.020 (.150)
NFG (*n* = 103) vs. Sakhalin all (*n* = 156)	0.000	0.093 (<.001)	0.090 (<.001)
NFG (*n* = 103) vs. Sakhalin noncalves[a] (*n* = 81)	0.000	0.064 (<.001)	0.058 (.001)
NFG (*n* = 103) vs. Sakhalin calves[b] (*n* = 75)	0.000	0.116 (<.001)	0.069 (.001)
NFG females (*n* = 61) vs. Sakhalin noncalf females (*n* = 37)	0.003	0.069 (.001)	0.045 (.014)
NFG males (*n* = 42) vs. Sakhalin noncalf males (*n* = 44)	0.001	0.060 (.001)	0.072 (.002)
NFG (*n* = 103) vs. Sakhalin only[c] (*n* = 122)	0.001	0.100 (.001)	0.141 (.001)
NFG (*n* = 103) vs. Sakhalin-ENP[d] (*n* = 34)	0.003	0.073 (.001)	0.082 (.001)
Sakhalin only[c] (*n* = 122) vs. Sakhalin-ENP[d] (*n* = 34)	0.126	0.021 (.051)	0.017 (.131)
Sakhalin noncalves[b] (*n* = 81) vs. Sakhalin calves[a] (*n* = 75)	0.237	0.002 (.291)	0.000 (.341)

(b)

Comparison	χ^2 *p*-value	F_{ST} (*p*-value)	F'_{ST} (*p*-value)
Sakhalin (*n* = 156) vs. Kamchatka (*n* = 16)	0.723	0.001 (.348)	0.004 (.35)
NFG (*n* = 105) vs. Kamchatka (*n* =	0.009	0.015	0.051

16)		(.003)	(.003)
NFG (*n* = 105) vs. Sakhalin all (*n* = 156)	0.000	0.016 (<0.001)	0.057 (<.001)
NFG (*n* = 105) vs. Sakhalin noncalves[a] (*n* = 81)	0.000	0.012 (<.001)	0.042 (<.001)
NFG (*n* = 105) vs. Sakhalin calves[b] (*n* = 75)	0.000	0.021 (<.001)	0.070 (<.001)
NFG females (*n* = 62) vs. Sakhalin noncalf females (*n* = 37)	0.000	0.027 (<.001)	0.095 (<.001)
NFG males (*n* = 43) vs. Sakhalin noncalf males (*n* = 44)	0.003	0.008 (.009)	0.028 (.009)
NFG (*n* = 105) vs. Sakhalin only[c] (*n* = 122)	0.001	0.018 (.001)	0.062 (.001)
NFG (*n* = 105) vs. Sakhalin-ENP[d] (*n* = 34)	0.001	0.008 (.004)	0.028 (.004)
Sakhalin only[c](*n* = 122) vs. Sakhalin-ENP[d] (*n* = 34)	0.368	−0.002 (.824)	−0.007 (.828)
Sakhalin noncalves[b] (*n* = 81) vs. Sakhalin calves[a] (*n* = 75)	0.319	0.000 (.471)	0.000 (.471)

[a] Sakhalin whales that were >1 year old when they were first photographically identified.

[b] Sakhalin whales that were first photographically identified as calves (whales <1 year old).

[c] Whales that are known to utilize the Sakhalin feeding ground but have not been recorded in the eastern North Pacific.

[d] Whales that are known to utilize the Sakhalin feeding ground and have also been identified on the ENP migratory route and/or wintering ground.

TABLE 3 Results of STRUCTURE clustering analysis using a model of admixture with correlated allele frequencies. No a

priori information on the geographic location of sampling was included. Values in bold indicate the optimal number of clusters identified by STRUCTURE using the two criteria described in the text.

K	Reps	Mean LnP(K)	SD LnP(K)	Ln′(K)	Ln″(K)	Delta K
1	5	−10,127.84	0.26	NA	NA	NA
2	5	**−9,973.22**	1.56	154.62	603.22	**386.81**
3	5	−10,421.82	68.57	−448.6 0	621.14	9.06
4	5	−10,249.28	60.64	172.54	428.96	7.07
5	5	−10,505.70	134.16	−256.4 2	235.22	1.75
6	5	−10,997.34	159.07	−491.6 4	16.90	0.11
7	5	−11,505.88	253.37	−508.5 4	472.72	1.87
8	5	−11,541.70	271.43	−35.82	NA	NA

FIGURE 1 Locations where samples were collected. Key areas mentioned in the text are labeled, including Utqiagvik, Alaska (UTQ); the Chukotka Peninsula, Russia (CHK); the region between Cape Navarin and Cape Olyutorskiy, Russia, south of the Koryak Mountains (KYK); the Kamchatka Peninsula, Russia (KAM); and Sakhalin Island, Russia (SI).

FIGURE 2 Median-joining network showing relationships among the mtDNA haplotypes. The numbers next to the nodes correspond to the haplotype IDs listed in Table S4. The size of the nodes is proportional to the frequencies of the haplotypes, and each node is colored to indicate the fraction of individuals with that haplotype from each strata. The small black diamonds(unlabeled) indicate haplotypes that were inferred by the program but were not found among our samples. The length of lines connecting nodes is proportional to the inferred number of mutations separating haplotypes; hash marks are used to represent the number of mutational events.

FIGURE 3 STRUCTURE barplots based on a model of admixture with correlated allele frequencies (a) K = 2, 3, and 4 for a model with no *a priori* information on geographic location of sampling; and (b) K = 2 when information on geographic location of sampling (i.e., locprior = 1) is incorporated. Each vertical bar represents a single individual, and is

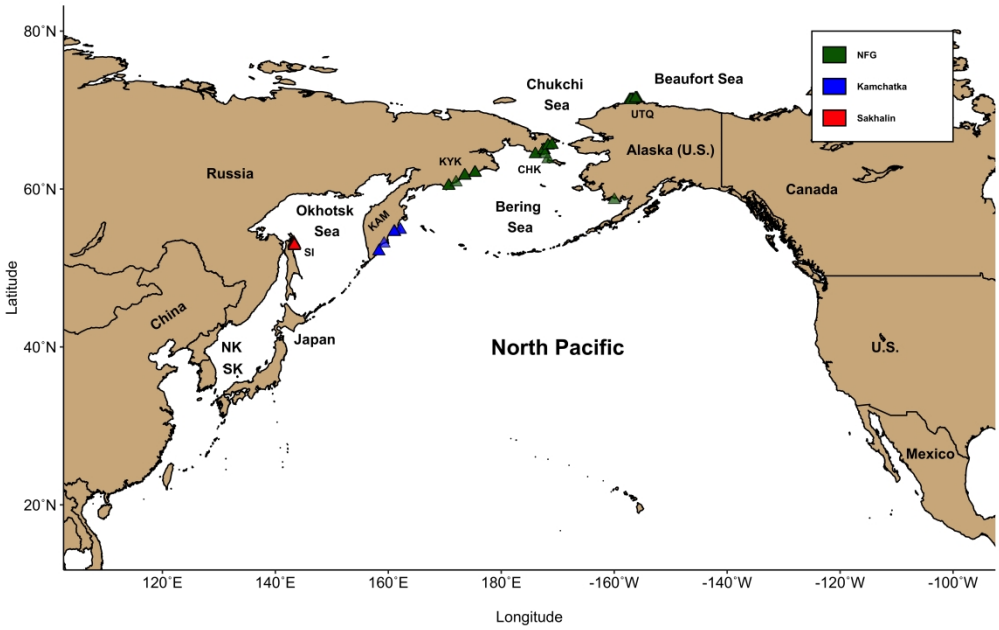
shaded based on the proportional membership (Q value) of individual whales to each of the inferred genetic clusters. Regions where samples were collected are arranged from east to west. For the Northern Feeding Ground stratum, the regions include Utquivik, AK (UTQ), the Chukotka Peninsula in Russia (CHK), and the region south of the Russian Koryak Mountains (KYK). In addition, four samples were collected from stranded whales in isolated regions of the Bering Sea, and these are grouped together under the BS group. The Western North Pacific stratum includes Kamchatka Peninsula (KAM) and Sakhalin Island (SI), Russia.

FIGURE 4 Scatter plot based on Discriminant Analysis of Principal Components with a priori information on stratum of origin incorporated and assuming that three clusters are present. Eigenvalues of the analysis are displayed in inset. Individuals are represented by dots, and the color of the dots denotes each individual's stratum of origin. Inertial ellipses encompassing 67% of individuals within each stratum are shown.

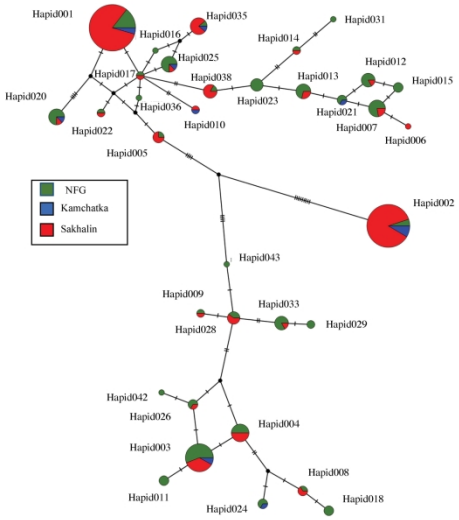
FIGURE 5 Self-assignment accuracies estimated via Monte-Carlo cross-validation and support vector machine methods with three levels of training loci and training data sets generated by random sampling of (a) all NFG and Sakhalin

samples, and (b) All NFG samples but only those Sakhalin samples representing individuals that have not been recorded in the ENP. Except where all loci were used, the training loci were selected based on those with the highest F_{ST} values. The line within the boxplot shows the median and the top and bottom edges represent the 25th and 75th percentiles. The ends of whiskers are the minimum and maximum of non-outliers, and outliers are shown as black circles. The horizontal red line indicates the null population assignment rate, which for two populations is 50%.

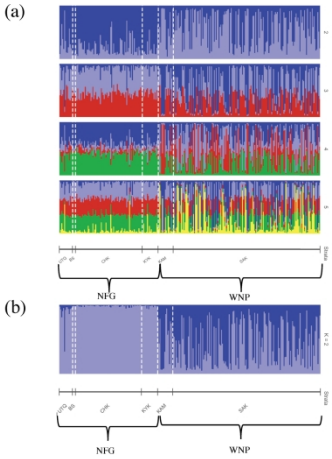
FIGURE 6 Membership probabilities of (a) the whales sampled off Kamchatka, and (b) the whales sampled off Sakhalin that have been recorded in the ENP. The baseline Sakhalin stratum included only those sampled Sakhalin whales that have not been recorded in the ENP. Membership probabilities were estimated using the random forest machine learning algorithm.



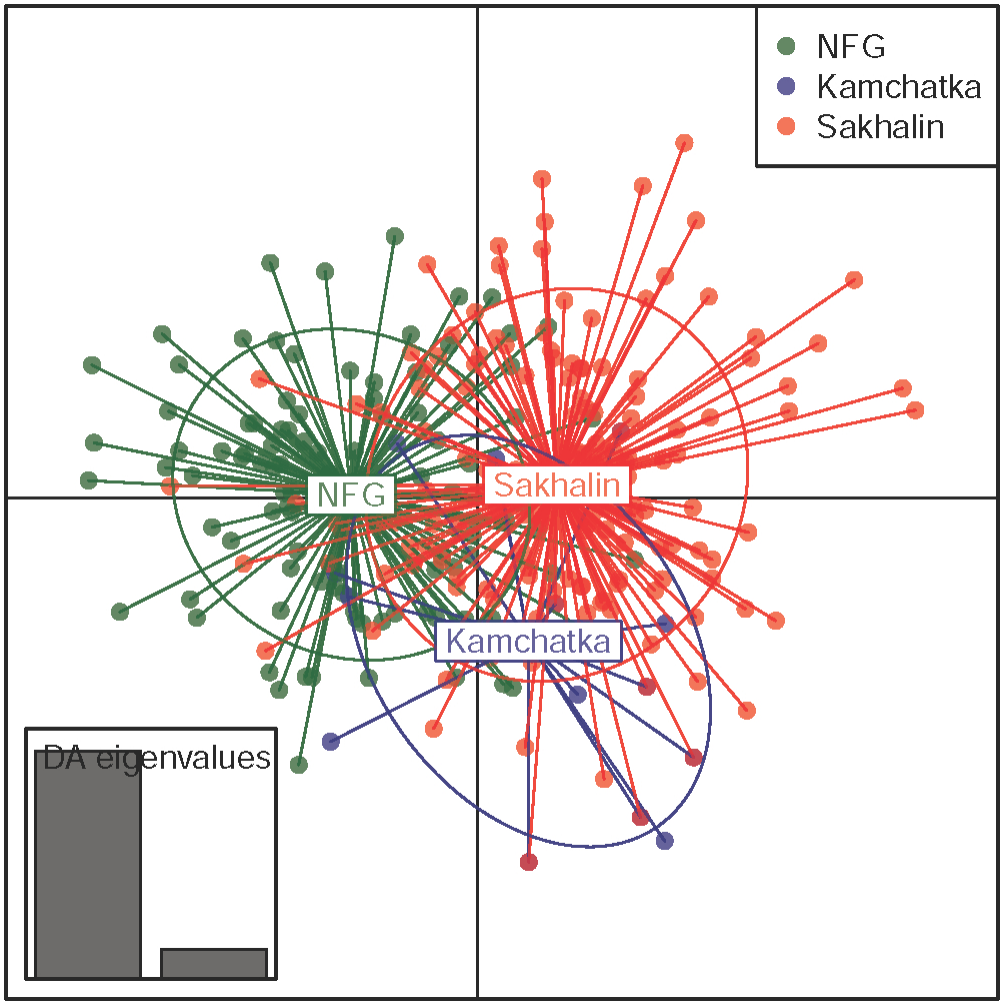
2317x1501mm (72 x 72 DPI)



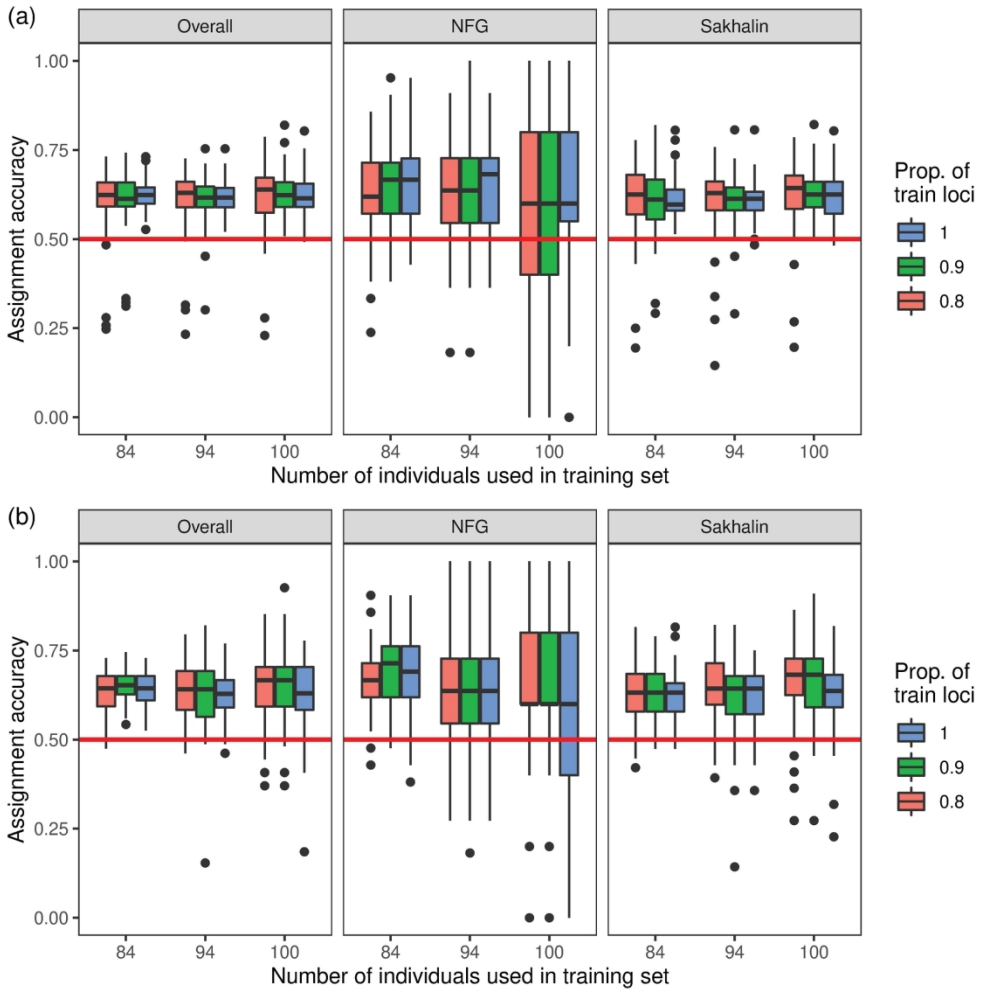
338x190mm (300 x 300 DPI)



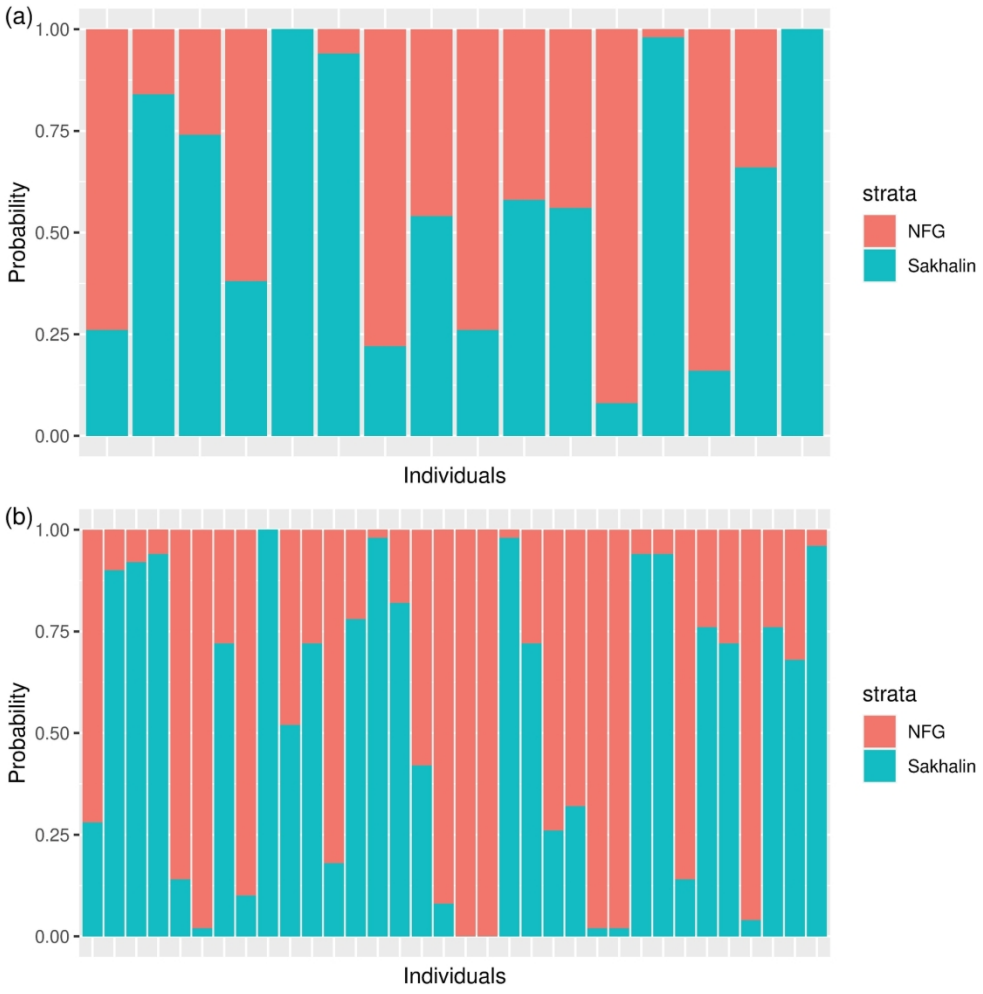
338x190mm (300 x 300 DPI)



127x127mm (200 x 200 DPI)



177x177mm (300 x 300 DPI)



177x177mm (300 x 300 DPI)

TABLE S1 Metadata associated with samples of gray whales collected for microsatellite and mtDNA analysis, including the SWFSC identification number, stratum to which the sample belongs, genetically determined sex, mtDNA haplotype (Genbank Accession number), year of sample collection, and whether the sample was identified as a "duplicate" (i.e., a sample that was removed because it was determined to represent the same individual as another sample in the data set).

Labid	Strata	Genetic sex	MtDNA haplotype	Year	Comment
5110	SAK	F	AF326790	1995	duplicate
5111	SAK	M	AF326790	1995	duplicate
5112	SAK	M	AF326790	1995	
5113	SAK	F	AF326789	1995	duplicate
5114	SAK	F	AF326789	1995	duplicate
5115	SAK	F	AF326789	1995	
5116	SAK	M	AF326789	1995	
5117	SAK	F	AF326789	1995	duplicate
5118	SAK	M	AF326789	1995	duplicate
12182	SAK	F	AF326790	1998	
12183	SAK	F	AF326790	1998	
12184	SAK	F	AF326790	1998	
12185	SAK	M	AF326792	1998	
12186	SAK	M	AF326790	1998	
12187	SAK	F	AF326793	1998	
12188	SAK	M	AF326789	1998	
12189	SAK	F	AF326790	1998	duplicate
12190	SAK	F	AF326789	1998	
12191	SAK	F	AF326790	1998	duplicate
12192	SAK	F	AF326789	1998	
12193	SAK	F	AF326790	1998	
12194	SAK	F	AF326789	1998	duplicate
12195	SAK	M	AF326792	1998	duplicate
12196	SAK	F	AF326796	1998	
12197	SAK	F	AF326790	1998	
12198	SAK	M	AF326789	1998	
12199	SAK	F	AF326790	1998	duplicate
12200	SAK	M	AF326797	1998	
12201	SAK	F	AF326790	1998	duplicate
15151	SAK	F	AF326790	1999	
15152	SAK	M	AF326791	1999	
15153	SAK	F	AF326789	1999	
15154	SAK	F	AF326789	1999	
15155	SAK	F	AF326790	1999	
15156	SAK	F	AF326789	1999	

15157	SAK	F	AF326789	1999	
15158	SAK	M	AF326789	1999	
15159	SAK	M	AF326794	1999	
15160	SAK	F	AF326789	1999	
15161	SAK	M	AF326792	1999	
15162	SAK	F	AF326790	1999	
15163	SAK	M	AF326795	1999	
15164	SAK	M	AF326789	1999	
15165	SAK	F	AF326789	1999	
15166	SAK	M	AF326789	1999	
15167	SAK	M	AF326790	1999	
15168	SAK	M	AF326790	1999	
15169	SAK	M	AF326790	1999	
15170	SAK	M	AF326789	1999	
15171	SAK	M	AF326791	1999	
15172	SAK	M	AF326789	1999	
15173	SAK	M	AF326789	1999	
15174	SAK	M	AF326798	1999	
15175	SAK	M	AF326789	1999	
15176	SAK	M	AF326789	1999	
15177	SAK	F	AF326790	1999	
15178	SAK	M	AF326790	1999	
19043	SAK	F	AF326796	2000	duplicate
19044	SAK	M	AF326789	2000	
19045	SAK	M	AF326790	2000	
19046	SAK	F	AF326790	2000	
19047	SAK	F	AF326791	2000	
19048	SAK	M	AF326790	2000	
19049	SAK	F	AF326796	2000	duplicate
19050	SAK	M	AF326790	2000	
19051	SAK	F	KC917326	2000	
19052	SAK	M	AF326814	2000	
19053	SAK	M	AF326789	2000	
19054	SAK	M	AF326789	2000	
19055	SAK	M	AF326790	2000	
19056	SAK	M	AF326792	2000	
19057	SAK	M	AF326790	2000	
19058	SAK	F	AF326789	2000	
19059	SAK	M	AF326791	2000	
19060	SAK	M	AF326821	2000	

19061	SAK	F	AF326789	2000	
19062	SAK	F	AF326790	2000	
19063	SAK	M	AF326792	2000	
32754	SAK	M	AF326790	2001	
32755	SAK	F	AF326790	2001	
32756	SAK	M	AF326789	2001	
32757	SAK	F	AF326789	2001	
32758	SAK	M	AF326790	2001	
32759	SAK	F	AF326789	2001	
32760	SAK	M	AF326808	2001	
32761	SAK	F	KC917326	2001	
32762	SAK	M	AF326792	2001	
32763	SAK	F	AF326790	2001	
32764	SAK	M	AF326794	2001	duplicate
32765	SAK	F	AF326789	2001	duplicate
32766	SAK	F	AF326790	2001	duplicate
32767	SAK	F	AF326790	2001	duplicate
32768	SAK	M	AF326789	2001	duplicate
32769	SAK	F	AF326790	2001	duplicate
32770	SAK	M	AF326791	2001	duplicate
32771	SAK	F	AF326796	2001	duplicate
32772	SAK	M	AF326797	2001	duplicate
32773	SAK	F	AF326790	2001	duplicate
32774	SAK	F	AF326816	2001	
32775	SAK	M	AF326792	2001	duplicate
32776	SAK	M	AF326789	2001	duplicate
32777	SAK	M	AF326791	2001	duplicate
32778	SAK	M	AF326789	2001	
32779	SAK	M	AF326791	2001	
32780	SAK	F	AF326790	2001	duplicate
32781	SAK	F	AF326789	2001	
32782	SAK	M	AF326800	2001	
32783	SAK	F	AF326790	2002	
32784	SAK	F	AF326793	2002	
32785	SAK	M	AF326790	2002	
32786	SAK	M	AF326790	2002	
32787	SAK	F	AF326823	2002	
32788	SAK	M	AF326823	2002	
32789	SAK	M	AF326789	2002	
32790	SAK	M	AF326789	2002	

32791	SAK	M	AF326789	2002	
32792	SAK	M	AF326814	2002	duplicate
32793	SAK	F	AF326789	2002	
32794	SAK	F	AF326789	2002	
32795	SAK	M	AF326790	2002	duplicate
32796	SAK	M	AF326791	2002	
32797	SAK	M	AF326816	2002	
32798	SAK	M	AF326790	2002	
42207	SAK	F	AF326790	2003	
42208	SAK	M	AF326813	2003	duplicate
42209	SAK	M	AF326789	2003	
42210	SAK	F	AF326801	2003	
42211	SAK	M	AF326790	2003	
42212	SAK	F	AF326789	2003	
42213	SAK	M	AF326790	2003	
42214	SAK	U	AF326789	2003	
42215	SAK	F	AF326795	2003	
42216	SAK	F	AF326789	2003	
42217	SAK	M	AF326790	2003	
42218	SAK	M	AF326823	2003	
42219	SAK	M	AF326791	2003	
42220	SAK	M	AF326802	2003	
42221	SAK	M	AF326801	2003	
50724	SAK	F	AF326790	2005	
50725	SAK	F	AF326790	2004	
50726	SAK	M	AF326796	2004	
50727	SAK	M	AF326790	2005	duplicate
50728	SAK	F	AF326823	2004	
50730	SAK	M	AF326790	2004	
50731	SAK	F	AF326789	2004	
50732	SAK	F	AF326789	2004	
50733	SAK	M	AF326823	2004	
50734	SAK	F	AF326789	2004	
50735	SAK	M	AF326793	2004	
50736	SAK	F	AF326823	2005	
50737	SAK	M	AF326790	2005	
50738	SAK	M	AF326789	2005	
50739	SAK	M	AF326790	2005	
50740	SAK	F	AF326791	2005	
50741	SAK	M	AF326789	2005	

53768	KAM	F	AF326789	2004	duplicate
68984	SAK	F	AF326790	2006	
68985	SAK	F	AF326789	2006	duplicate
68986	SAK	F	AF326791	2006	
68987	SAK	M	AF326789	2006	
68988	SAK	F	AF326789	2006	
68989	SAK	M	AF326810	2006	
72878	SAK	F	AF326790	2007	
72879	SAK	F	AF326790	2007	
72880	SAK	M	KC917326	2007	
72881	SAK	M	AF326789	2007	
72882	SAK	F	AF326789	2007	
72883	SAK	M	AF326789	2007	
72884	SAK	M	AF326790	2007	
72885	SAK	F	AF326790	2007	
72886	SAK	F	AF326816	2007	
72887	SAK	M	AF326805	2007	
72888	SAK	M	AF326789	2007	
72889	SAK	M	AF326823	2007	
72890	SAK	M	AF326789	2007	
72891	SAK	M	AF326813	2007	
100761	KAM	F	AF326809	2004	duplicate
100762	KAM	F	AF326789	2004	duplicate
100763	KAM	F	AF326809	2004	
100764	KAM	M	AF326813	2010	duplicate
100765	KAM	M	AF326790	2010	duplicate
100766	KAM	M	AF326790	2010	
100767	KAM	M	AF326813	2010	
100789	KAM	F	AF326789	2010	
100790	KAM	M	AF326790	2010	
100791	KAM	M	AF326789	2010	
100792	KAM	F	AF326790	2010	
100793	KAM	M	AF326790	2010	
100794	KAM	M	AF326813	2010	duplicate
100795	KAM	M	AF326789	2010	
100796	KAM	F	AF326798	2010	
112375	KAM	F	AF326808	2011	
112376	KAM	M	AF326791	2011	
112377	KAM	M	AF326790	2011	duplicate
112378	KAM	M	AF326823	2011	

112379	KAM	M	AF326791	2011	duplicate
112380	KAM	M	AF326812	2011	
112381	KAM	F	AF326791	2011	
112382	KAM	M	AF326790	2011	
112383	SAK	M	AF326790	2010	
112384	SAK	F	AF326790	2010	duplicate
112385	SAK	M	AF326789	2010	duplicate
112386	SAK	F	AF326790	2010	
112387	SAK	F	AF326790	2010	duplicate
112388	SAK	M	AF326789	2010	
112389	SAK	F	AF326790	2010	duplicate
112390	SAK	F	AF326790	2010	duplicate
112391	SAK	M	AF326791	2010	duplicate
112392	SAK	M	AF326790	2011	
112393	SAK	F	AF326790	2011	
112394	SAK	M	AF326790	2011	
112395	SAK	F	AF326790	2011	
112396	SAK	F	AF326790	2011	
112397	SAK	F	AF326789	2011	
112398	SAK	M	KC917326	2011	
112399	SAK	M	AF326789	2011	duplicate
112400	SAK	F	KC917326	2011	
112401	SAK	F	AF326790	2011	duplicate
112402	SAK	F	AF326789	2011	
112403	SAK	F	AF326789	2011	
112404	SAK	M	AF326789	2011	Duplicate
112405	SAK	M	AF326789	2011	

TABLE S2 Microsatellite loci used in the study. Includes the species for which primers were initially designed, size of repeats, annealing temperature (T_a), and reference listing primer sequences. For loci ending with a "t", genotyping of all samples was conducted using a reverse primer whose sequence was tailed (i.e., the sequence GTTTCTT was added to the on the 5′ end; Brownstein et al., 1996) to reduce allelic stutter. With the exception of EV37 and Gata098, the reverse primer of the remaining loci was tailed partway through the study; see description below the table.

Locus	Source species	Repeat size (bp)	T_a (°C)	Reference
EV14t	*Physeter macrocephalus*	2	55	Valsecchi and Amos (1996)
EV37[a]	*Megaptera novaeangliae*	2	55	Valsecchi and Amos (1996)
EV94[a,b]	*Megaptera novaeangliae*	2	52	Valsecchi and Amos (1996)
Gata028[a]	*Megaptera novaeangliae*	4	54	Palsbøll et al. (1997)
Gata098[a]	*Megaptera novaeangliae*	4	52	Palsbøll et al. (1997)
Gata417[a,c]	*Megaptera novaeangliae*	4	54	Palsbøll et al. (1997)
Gt023[a,c]	*Megaptera novaeangliae*	2	54	Palsbøll et al. (1997)
RW31[c]	*Eubalaena glacialis*	2	54	Waldic et al. (1999)
RW48[c]	*Eubalaena glacialis*	2	55	Waldick et al. (1999)
SW10t	*Physeter macrocephalus*	Complex (2, 4)	55	Richard et al. (1996)
SW13t	*Physeter macrocephalus*	2	55	Richard et al. (1996)
SW19t	*Physeter macrocephalus*	2	55	Richard et al. (1996)

[a] One of the six original loci that was used to genotype the Sakhalin samples collected prior to 2002 on an ABI 377 instrument.
[b] The Kamchatka samples and the Sakhalin samples collected in 2002 and later were genotyped with a tailed reverse primer.
[c] The Kamchatka samples and the samples collected from Sakhalin whales in 2010 and 2011 were genotyped with a tailed reverse primer.

TABLE S3 Measures of linkage disequilibrium between pairs of microsatellite loci. *P*-values <.05 are highlighted in bold.

Locus.1	Locus.2	North *p*	Kamchatka *p*	Sakhalin *p*
EV14	EV37	.791	.775	.065
EV14	EV94	.294	.093	.220
EV14	Gata028	.054	.839	**.015**
EV14	Gata098	.128	**.031**	.158
EV14	Gata417	.154	.054	.120
EV14	Gt023	.063	1.000	.245
EV14	RW31	.899	.080	.209
EV14	RW48	.124	.777	.396
EV14	SW10	.797	1.000	.768
EV14	SW13	.611	.395	.350
EV14	SW19	.927	.773	.224
EV37	EV94	.776	.637	.543
EV37	Gata028	.548	.683	**.044**
EV37	Gata098	**.007**	.300	**.000**
EV37	Gata417	.071	.489	.652
EV37	Gt023	.263	.576	**.040**
EV37	RW31	.808	.657	.530
EV37	RW48	.803	.537	.153
EV37	SW10	.793	.630	.123
EV37	SW13	.717	**.022**	.624
EV37	SW19	.314	.270	**.002**
EV94	Gata028	.875	.270	.468
EV94	Gata098	.989	.092	.140
EV94	Gata417	.771	.183	.793
EV94	Gt023	.572	1.000	.772
EV94	RW31	.948	1.000	**.023**
EV94	RW48	.122	.544	.610
EV94	SW10	.559	.066	.448
EV94	SW13	.538	1.000	.835
EV94	SW19	**.001**	.656	.181
Gata028	Gata098	.672	.309	.556
Gata028	Gata417	.566	.944	.288
Gata028	Gt023	.318	.715	.726
Gata028	RW31	.510	.718	**.004**
Gata028	RW48	.376	1.000	**.010**
Gata028	SW10	.991	.767	.541
Gata028	SW13	.933	.339	.153
Gata028	SW19	**.043**	.699	.266
Gata098	Gata417	.648	.299	.105
Gata098	Gt023	.941	.872	.771
Gata098	RW31	.539	.645	.834
Gata098	RW48	.362	.111	.134
Gata098	SW10	.335	.731	**.036**
Gata098	SW13	.188	.277	.097
Gata098	SW19	.613	.157	.059

Gata417	Gt023	.686	.187	.310
Gata417	RW31	.221	1.000	**.014**
Gata417	RW48	.426	**.021**	.179
Gata417	SW10	.974	1.000	.418
Gata417	SW13	.165	.392	.274
Gata417	SW19	.611	.082	.096
Gt023	RW31	.827	1.000	.162
Gt023	RW48	.847	.133	.388
Gt023	SW10	.553	1.000	.296
Gt023	SW13	.969	.690	**.000**
Gt023	SW19	.945	.579	.544
RW31	RW48	.535	1.000	.902
RW31	SW10	.770	1.000	.060
RW31	SW13	.400	.876	**.001**
RW31	SW19	.985	.616	**.000**
RW48	SW10	.264	.649	.248
RW48	SW13	.196	.263	.590
RW48	SW19	.987	.152	.306
SW10	SW13	.456	1.000	.446
SW10	SW19	.940	.707	.260
SW13	SW19	.827	.700	.470

TABLE S4 The mtDNA haplotypes identified in the study, their corresponding NCBI accession numbers, and the number of individuals with each haplotype in each stratum.

Hapid	GenBank accession #	NFG (*n* = 103)	Sakhalin (*n* = 56)	Kamchatka (*n* = 16)
Hapid001	AF326789	10	56	3
Hapid002	AF326790	3	51	5
Hapid003	AF326791	14	9	2
Hapid004	AF326792	5	5	0
Hapid005	AF326793	1	3	0
Hapid006	AF326794	0	1	0
Hapid007	AF326795	7	2	0
Hapid008	AF326796	1	2	0
Hapid009	AF326797	1	1	0
Hapid010	AF326798	0	1	1
Hapid011	AF326799	3	0	0
Hapid012	AF326800	5	1	0
Hapid013	AF326801	5	2	0
Hapid014	AF326802	1	1	0
Hapid015	AF326803	3	0	0
Hapid016	AF326804	1	0	0
Hapid017	AF326805	1	1	0
Hapid018	AF326806	3	0	0
Hapid020	AF326808	6	1	1
Hapid021	AF326809	2	0	1
Hapid022	AF326810	1	1	0
Hapid023	AF326811	5	0	0
Hapid024	AF326812	2	0	1
Hapid025	AF326813	6	1	1
Hapid026	AF326814	2	1	0
Hapid028	AF326816	2	3	0
Hapid029	AF326817	2	0	0
Hapid031	AF326819	1	0	0
Hapid033	AF326821	5	1	0
Hapid035	AF326823	1	7	1
Hapid036	AF326824	1	0	0
Hapid038	KC917326	1	5	0
Hapid042	KC917327	1	0	0
Hapid043	KC917328	1	0	0

TABLE S5 Measures of the number of individuals genotyped (*n*), number of alleles (*n*A), and allelic richness (Ar) per locus for each of the three feeding strata.

Locus	North			Kamchatka			Sakhalin		
	n	***n*A**	**Ar**	***n***	***n*A**	**Ar**	***n***	***n*A**	**Ar**
EV14	105	8	6.8	16	7	7	154	9	6.3
EV37	105	16	11.0	15	9	9	155	17	10.0
EV94	105	11	7.3	16	6	6	154	9	6.2
Gata028	105	7	5.2	16	5	5	155	5	4.8
Gata098	105	9	5.3	16	5	5	155	7	4.9
Gata417	104	6	5.6	16	5	5	156	7	4.6
Gt023	105	8	5.0	16	4	4	156	8	5.3
RW31	105	10	7.5	16	7	7	142	9	6.9
RW48	105	5	3.7	16	3	3	156	5	3.8
SW10	105	9	6.3	16	6	6	155	9	6.4
SW13	105	7	4.6	16	4	4	155	8	5.0
SW19	104	9	5.7	16	5	5	153	7	5.1
Overall	104.8	8.8	6.2	15.9	5.5	5.5	153.8	8.3	5.8

TABLE S6 Results of STRUCTURE clustering analysis using a model of admixture with correlated allele frequencies and different subsets of the data. No a priori information on the geographic location of sampling was included. Values in bold indicate the optimal number of clusters identified by STRUCTURE using the two criteria described in the text.

(a) Results when only the Sakhalin samples were analyzed.

K	Reps	Mean LnP(K)	Delta K
1	5	−5,436.54	NA
2	5	−5,355.36	3.97
3	5	**−5,288.28**	**141.42**
4	5	−5,555.02	0.29
5	5	−5,831.52	1.90
6	5	−5,686.96	3.01
7	5	−5,880.86	1.78
8	5	−5,704.98	NA

(b) Results when only samples from Sakhalin noncalf whales were analyzed (i.e., whales first photographically identified as calves were removed).

K	Reps	Mean LnP(K)	Delta K
1	5	−7,550.84	NA
2	5	**−7,539.10**	**39.21**
3	5	−7,662.66	2.49
4	5	−7,904.84	0.32
5	5	−8,192.56	0.75
6	5	−8,371.54	1.39
7	5	−8,314.82	0.19
8	5	−8,299.08	NA

(c) Results when only the Sakhalin and Kamchatka samples (with duplicate samples removed) were analyzed.

K	Reps	Mean LnP(K)	Delta K
1	5	−2,440.72	NA
2	5	**−2,125.54**	**1,756.34**
3	5	−2,214.70	5.46
4	5	−2,253.68	1.47
5	5	−2,230.60	13.79
6	5	−2,265.82	1.11
7	5	−2,311.46	0.56
8	5	−2,345.54	NA

(d) Results when only the samples collected from whales on the Northern Feeding Ground were analyzed.

K	Reps	Mean LnP(K)	Delta K
1	5	**−4,009.48**	NA
2	5	−4,019.74	2.53
3	5	−4,054.66	0.09
4	5	−4,086.34	0.76
5	5	−4,089.22	1.03
6	5	−4,133.92	1.11
7	5	−4,147.98	0.78
8	5	−4,109.00	NA

TABLE S7 Results of the NFG and Sakhalin sample sets when the Sakhalin sample set was randomly subsampled to include the same number of individuals as the NFG (n = 105).

Subsample.1

K	Reps	Mean LnP(K)	Delta K
1	3	−7,808.13	NA
2	3	**−7,762.07**	**112.24**
3	3	−8,038.63	0.39
4	3	−8,331.53	0.82
5	3	−8,575.17	NA

Subsample.2

K	Reps	Mean LnP(K)	Delta K
1	3	−7,763.57	NA
2	3	**−7,710.70**	**122.13**
3	3	−7,870.77	3.06
4	3	−8,519.10	10.40
5	3	−8,291.10	NA

Subsample.3

K	Reps	Mean LnP(K)	Delta K
1	3	−7,802.90	NA
2	3	**−7,791.23**	**35.13**
3	3	−7,905.30	4.10
4	3	−8,324.73	4.88
5	3	−8,514.13	NA

Subsample.4

K	Reps	Mean LnP(K)	Delta K
1	3	**−7,808.00**	NA
2	3	−7,811.77	**18.46**
3	3	−8,062.23	0.71
4	3	−8,360.90	3.26
5	3	−8,909.90	NA

Subsample.5

K	Reps	Mean LnP(K)	Delta K
1	3	−7,740.43	NA
2	3	**−7,693.27**	**74.48**
3	3	−8,109.57	7.03
4	3	−8,167.97	0.48
5	3	−8,289.37	NA

Subsample.6

K	Reps	Mean LnP(K)	Delta K
1	3	**−7,782.63**	NA
2	3	−7,879.20	**4.42**
3	3	−8,097.30	0.11
4	3	−8,302.00	0.04
5	3	−8,509.93	NA

Subsample.7

K	Reps	Mean LnP(K)	Delta K
1	2	−7,767.50	NA
2	3	**−7,753.43**	**47.56**
3	3	−7,945.27	5.21
4	3	−8,298.17	3.54
5	3	−9,006.87	NA

Subsample.8

K	Reps	Mean LnP(K)	Delta K
1	3	**−7,779.93**	NA
2	3	−7,800.23	7.76
3	3	−7,951.83	8.50
4	3	−8,299.10	**9.75**
5	3	−8,222.90	NA

Subsample.9

K	Reps	Mean LnP(K)	Delta K
1	3	**−7,740.13**	NA
2	3	−7,758.17	**15.12**
3	3	−7,988.27	1.40
4	3	−7,948.67	2.57
5	3	−8,136.47	NA

Subsample.10

K	Reps	Mean LnP(K)	Delta K
1	3	**−7,762.53**	NA
2	3	−7,830.30	2.55
3	3	−8,007.73	**6.09**
4	3	−8,377.13	4.97
5	3	−8,250.80	NA

FIGURE S1 Relationship between pairwise genetic differentiation between regions and geographic distance between sampling sites: (A) genetic differentiation based on mtDNA haplotype frequencies, (B) genetic differentiation based on microsatellite allele frequencies. Points are colored based on whether the comparison was between two sites within the same stratum or sites from different strata. See Figure 1 for a map of sampling sites corresponding to the abbreviations used here.

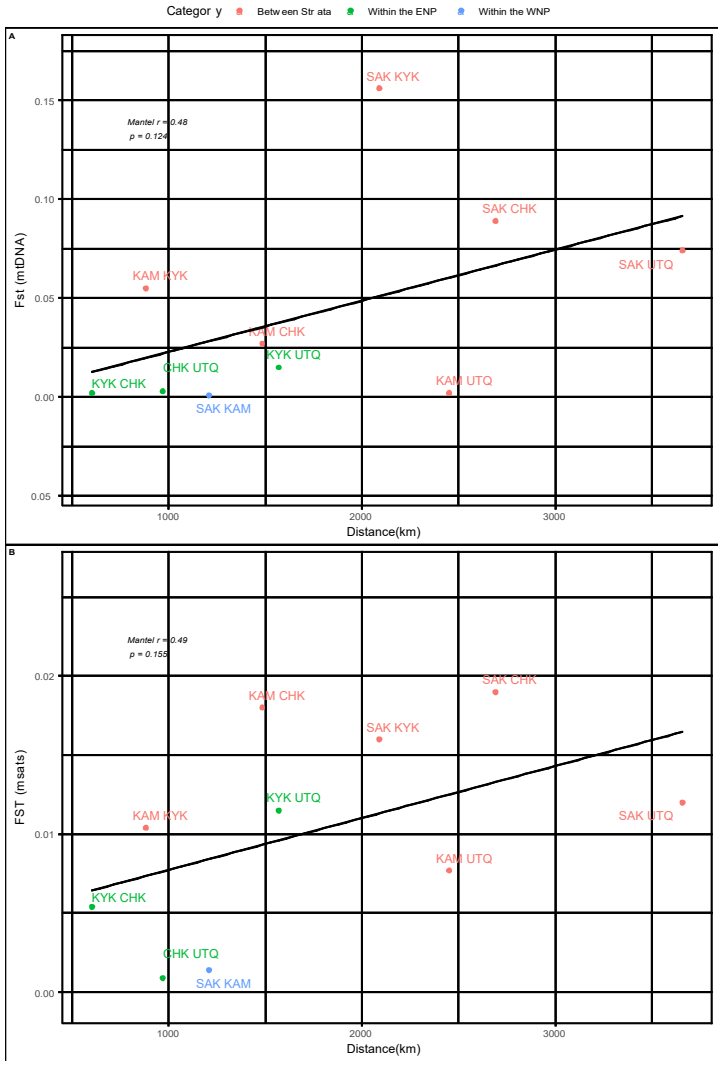
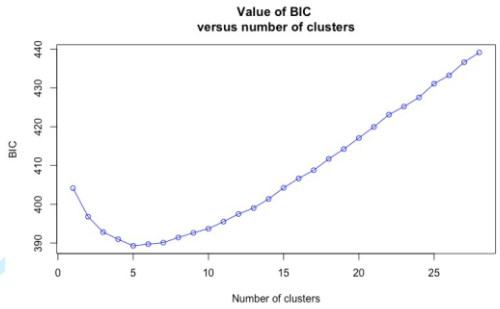


Figure S2. Results of K-means clustering and optimization of the alpha score for the discriminant analysis of principal components run on the gray whale microsatellite data: (a) the BIC values for increasing numbers of clusters (K); and (b) the estimated a-score, which is a measure of the proportion of assignments to the a priori defined groups versus to random clusters.

(a)



(b)

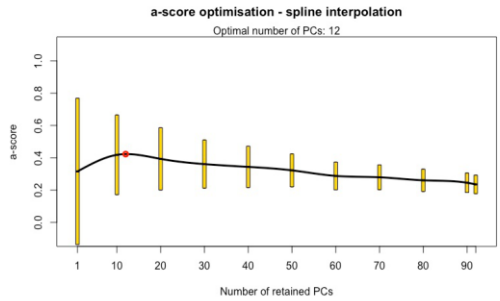


FIGURE S3 Scatter plot based on discriminant analysis of principal components without a priori information on stratum of origin incorporated and assuming that three clusters are present. Individuals are represented by dots, and the color of the dots denotes each individual's stratum of origin. Inertial ellipses encompassing 67% of individuals of the assigned cluster.