Journal of Mammalogy, 100(5):1653–1670, 2019
DOI:10.1093/jmammal/gyz121

Revision of fin whale *Balaenoptera physalus* (Linnaeus, 1758) subspecies using genetics

Frederick I. Archer,*, Robert L. Brownell, Jr., Brittany L. Hancock-Hanser, Phillip A. Morin, Kelly M. Robertson, Kathryn K. Sherman, John Calambokidis, Jorge Urbán R., Patricia E. Rosel, Sally A. Mizroch, Simone Panigada, and Barbara L. Taylor

Southwest Fisheries Science Center, 8901 La Jolla Shores Drive, La Jolla, CA 92037, USA (FIA, RLB, BLH-H, PAM, KMR, BLT)
Ocean Associates, Inc., 4007 N. Abingdon St., Arlington, VA 22207, USA (KKS)
Cascadia Research Collective, 218 1/2 W 4th Ave., Olympia, WA 98501, USA (JC)
Departamento de Ciencias Marinas y Costeras, Universidad Autónoma de Baja California Sur, Km 5.5 Carretera al Sur, La Paz, B.C.S. 23080, Mexico (JUR)
National Marine Fisheries Service, Southeast Fisheries Science Center, 646 Cajundome Blvd., Lafayette, LA 70506, USA (PER)
Blue Sea Research, P.O. Box 15805, Seattle, WA 98115, USA (SAM)
Tethys Research Institute, Viale G.B. Gadio 2, 20121 Milano, Italy (SP)

* *Correspondent: eric.archer@noaa.gov*

Three subspecies of fin whales (*Balaenoptera physalus*) are currently recognized, including the northern fin whale (*B. p. physalus*), the southern fin whale (*B. p. quoyi*), and the pygmy fin whale (*B. p. patachonica*). The Northern Hemisphere subspecies encompasses fin whales in both the North Atlantic and North Pacific oceans. A recent analysis of 154 mitogenome sequences of fin whales from these two ocean basins and the Southern Hemisphere suggested that the North Pacific and North Atlantic populations should be treated as different subspecies. Using these mitogenome sequences, in this study, we conduct analyses on a larger mtDNA control region data set, and on 23 single-nucleotide polymorphisms (SNPs) from 144 of the 154 samples in the mitogenome data set. Our results reveal that North Pacific and North Atlantic fin whales can be correctly assigned to their ocean basin with 99% accuracy. Results of the SNP analysis indicate a correct classification rate of 95%, very low rates of gene flow among ocean basins, and that distinct mitogenome matrilines in the North Pacific are interbreeding. These results indicate that North Pacific fin whales should be recognized as a separate subspecies, with the name *B. p. velifera* Cope in Scammon 1869 as the oldest available name.

Key words: diagnosability, mtDNA, North Atlantic, North Pacific, SNPs, Southern Hemisphere, systematics, whales

Most species of large whales are distributed throughout the world's oceans, separated by continental land masses. Migratory patterns between high-latitude feeding grounds and low-latitude breeding grounds in some mysticetes restrict movement across the equator, isolating populations in either hemisphere (Kellogg 1929; Mackintosh 1965). Additionally, in some productive warm-water regions, smaller resident populations can be found in semi-enclosed seas like the Mediterranean and the Gulf of California (Geijer et al. 2016; Jiménez López et al. 2018). These geographical isolating mechanisms promote divergence between populations in different ocean basins and within semi-enclosed seas, driving the development of subspecies and species (Pastene et al. 2007). For example, the more polar-distributed right whales are represented by separate species in the North Atlantic (*Eubalaena glacialis*), North Pacific (*E. japonica*), and Southern Hemisphere (*E. australis*—Rosenbaum et al. 2000). Distinct subspecies of common minke whales (*Balaenoptera acutorostrata* ssp.) and humpback whales (*Megaptera novaengliae* ssp.) are recognized in each ocean basin (Rice 1998; Jackson et al. 2014).

Fin whales (*Balaenoptera physalus*) are found from temperate to subpolar oceans across the world. Edwards et al. (2015) demonstrated a distributional hiatus between the Northern and Southern Hemispheres located approximately between 20° and 30° on either side of the equator that likely serves as a barrier to north-south gene flow. In the Northern Hemisphere, fin whales

tend to occur in offshore temperate to subpolar waters south of approximately 60–70°N (Mizroch et al. 1984, 2009; Edwards et al. 2015), likely precluding migration between the North Atlantic and North Pacific even during periods of warming when the northwest passage might allow movement of arctic species between ocean basins (Dyke et al. 1996).

The species was first named from the eastern North Atlantic by Linnaeus (1758) based on the whales described by Martens (1675). Lönnberg (1906) then recognized a larger subspecies in the Southern Hemisphere (*B. p. quoyi*), which was later confirmed by morphometric examinations of a larger series of specimens by Lockyer and Waters (1986) using data sets from the North Atlantic and the Southern Hemisphere. With the establishment of the Southern Hemisphere subspecies, the Northern Hemisphere fin whales (North Atlantic and North Pacific) became the nominate subspecies, *B. p. physalus*. This trinomial was first used by Tomilin (1946). More recently, Clarke (2004) proposed the smaller, darker midlatitude Southern Hemisphere fin whale he examined ($n = 1$) as a subspecies, and assigned them to *B. p. patachonica* Burmeister 1865, giving it the common name of the "pygmy" fin whale. Although *B. p. patachonica* has been accepted by the Society for Marine Mammalogy (SMM) Committee on Taxonomy (2018) based on Clarke (2004), this proposal has not been critically reviewed and no genetic comparisons have been made.

Previous studies have shown significant genetic (Bérubé et al. 1998, 2002; Hatch 2004) and acoustic (Hatch 2004) differences between North Pacific and North Atlantic fin whales. However, those studies analyzed small sample sizes from restricted localities in the North Pacific and did not include samples from the Southern Hemisphere. More recently, the distinctiveness of North Pacific fin whales was confirmed by a phylogenetic study of 154 full mitogenome sequences from samples collected from the North Atlantic (including the Mediterranean Sea), the Southern Hemisphere, and the North Pacific (Archer et al. 2013). In this analysis, North Atlantic samples formed their own well-supported clade distinct from the North Pacific and Southern Hemisphere samples. Samples from the North Pacific clustered into three separate clades, one of which clustered with two Southern Hemisphere samples, diverging approximately 370 KYA, providing strong evidence for at least one successful trans-equatorial migration event from the Southern Hemisphere to the North Pacific.

While Archer et al.'s (2013) study provided strong evidence that North Pacific and North Atlantic fin whales should be treated as distinct subspecies, the data used in that study were insufficient to support a taxonomic revision of Northern Hemisphere fin whales. The observed phylogenetic patterns left open several questions, including whether the North Pacific mtDNA clades are freely interbreeding with one another, and whether the same pattern of differentiation among ocean basins is supported by the nuclear genome. If both are true, then it is reasonable to consider all North Pacific fin whales, with the exception of those found in the East China Sea (ECS), as part of the same taxon. Otherwise, more research is necessary to understand the identity of the matrilines, the mechanisms that keep them reproductively isolated, and how much dispersal is occurring among ocean basins. To properly delineate units, the results of genetic analyses should be examined in the context of explicitly described species and subspecies concepts and definitions (Rosel et al. 2017b).

Although there is a rich literature of species concepts (Häuser 1987; Zink and Davis 1999; Lee 2003; Sites and Marshall 2004; de Queiroz 2007; Hausdorf 2011), less attention has been paid to practical definitions and delineations of subspecies, especially using genetic data (Taylor et al. 2017b). Historically, subspecies have been recognized as geographically separate and diagnosably, often based on morphology, distinct breeding populations (Mayr and Ashlock 1991; Patten and Unitt 2002). Here, we follow the definitions given by Taylor et al. (2017b:17), wherein a species is defined as "a separately evolving lineage composed of a population or collection of populations," and a subspecies is defined as "a population, or collection of populations that appears to be a separately evolving lineage with discontinuities resulting from geography, ecological specialization, or other forces that restrict gene flow to the point that the population or collection of populations is diagnosably distinct." Based on these definitions and results from a suite of analyses of mitochondrial DNA (mtDNA) control region data sets from 51 pairs of recognized cetacean populations, subspecies, and species (Archer et al. 2017b; Rosel et al. 2017a), Taylor et al. (2017a) developed a set of guidelines for delimiting subspecies and species with genetic data.

These guidelines connect two of the key elements of the subspecies definition to quantifiable metrics. The first, diagnosability, is defined as a "measure of the ability to correctly determine the taxon of a specimen of unknown origin based on a set of distinguishing characteristics" (Archer et al. 2017b:104). A threshold of 95% was set, as populations are more likely to only exhibit differences in frequency-based metrics (Patten 2010; Martien et al. 2017; Rosel et al. 2017b). This threshold recognizes that unlike species, for which the threshold is 100%, some classification error is to be expected as gene flow among subspecies may still be present at a low level (Patten 2010; Archer et al. 2017b; Taylor et al. 2017b).

Ensuring that putative taxa appear to be on different evolutionary trajectories is the second key element that separates subspecies from populations. It is recognized that with enough markers, groups that would not otherwise be considered subspecies or species (e.g., close family units, or in the case of mtDNA, philopatrically distributed maternal groups) could be found to be highly diagnosable (Martien et al. 2017). Rosel et al. (2017a) found that these units could best be distinguished from subspecies and species with d_A, a measure of net nucleotide divergence (Nei et al. 1983). Based on the results of these analyses, Taylor et al. (2017a) set $d_A = 0.004$ as the lower bound for separating subspecies from populations.

Given that no systematically collected morphometric data exist or are likely to be available in the near future for fin whales from all of the ocean basins in which they occur, genetic data are likely to be the only way to sort out the taxonomy of this species, as is the case for most species of large whales

(Taylor et al. 2017b). Thus, in this study, we conduct additional analyses of fin whale mtDNA sequences and nuclear single-nucleotide polymorphism (SNP) loci, genotyped on the same individuals as in the Archer et al. (2013) mitogenome phylogenetic study to address the following five related questions:

1) Are individuals from each mitogenome clade within the North Pacific interbreeding with one another, or do members of each clade also represent genetically distinct taxa?
2) Are patterns of differentiation in nuclear loci similar to those seen in the mitogenome or is there evidence of male-mediated gene flow between ocean basins?
3) Are individuals sufficiently diagnosable to their ocean basin of origin such that the North Atlantic, North Pacific, and Southern Hemisphere should be considered separate taxonomic units?
4) Is divergence among ocean basins sufficiently large to be important on an evolutionary time scale?
5) If the individual ocean basins can be considered separate taxa, are they subspecies or species?

We address these questions and the results of these analyses in light of the standards set forth by Taylor et al. (2017a) for delimiting cetacean subspecies and species with genetic data. From these results, we propose a synthesis of the taxonomy of fin whales at the level of subspecies.

Materials and Methods

mtDNA sequences.—We examined three fin whale mtDNA data sets representing different portions of the mitogenome. The first was the 154 full mitogenome sequences (~16K bp) published in Archer et al. (2013). For compatibility with the Random Forest analyses for diagnosability (described below), we had to exclude ambiguous sites (Ns) from the published mitogenome sequence data. Because a relatively small number of sequences were responsible for a large number of sites with ambiguities, we conducted a multistep filtering process to maximize both sample size and the number of sites available for analysis. Of the 154 mitogenome sequences, 24 had Ns at one site, and 38 had Ns at more than one site. A total of 160 sites that were otherwise variable among sequences had Ns. After examining the distribution of ambiguities in sequences and variable sites, we elected to remove the 10 sequences with the greatest number of Ns (6–274 sites). This left nine sites with Ns that would otherwise be variable, which were then removed, leaving a final data set of 144 mitogenome sequences with 912 variable sites. Nine of the 10 samples removed were from the North Pacific and one was from the Southern Hemisphere. Two of the North Pacific samples were the ones that composed ambiguously placed Clade B in the Archer et al. (2013) phylogenetic tree. Of the remaining seven, four were from Clade A and three were from Clade C. Thus, in this analysis, we only refer to the two well-supported North Pacific mitogenome matrilines from Archer et al. (2013): Clade A ($n = 81$), and Clade C ($n = 14$).

The second data set used the same final 144 samples as in the mitogenome data set, but was restricted to 403 bp of the hypervariable portion of the control region. The third data set was composed of the first 233 bp of this same hypervariable portion of the control region that included an additional 258 samples from the North Pacific sequenced at the Southwest Fisheries Science Center (SWFSC), 319 sequences from the Atlantic and Mediterranean published in Bérubé et al. (1998), 57 sequences from South Georgia Island (remains from commercial whaling in the Southern Ocean) published in Sremba et al. (2015), and three sequences from around South Korea (Kim et al. 2018). Sanger sequencing of the SWFSC samples followed the same protocol given in Archer et al. (2013). All sequences were initially aligned with MAFFT (Katoh and Standley 2013) and refined by eye. Sequences with substitutions that created unique haplotypes or those with unusually large mean pairwise distances were re-called from the original chromatogram.

SNP genotypes.—We genotyped SNPs for 146 of the 154 samples for which mitogenomes were sequenced in Archer et al. (2013). Of the eight samples that were not genotyped, two were from the North Pacific, two from the North Atlantic, and four from the Southern Hemisphere. Eight additional North Pacific samples, for which no mitogenomes were available, were genotyped to make a final data set of 154 genotyped samples. The majority of the samples were biopsies of living animals; however, six North Pacific and two North Atlantic samples were from stranded individuals.

All tissue samples are stored in the National Marine Fisheries Service Marine Mammal and Sea Turtle Research (MMASTR) collection housed at SWFSC or the Marine Mammal Molecular Genetics Laboratory at the Southeast Fisheries Science Center (SEFSC). Samples were stored in a −80°C freezer with no preservative, or fixed in a salt-saturated 20% DMSO solution (Amos 1997) or ethanol and archived in a −20°C freezer. DNA was extracted using various methods. The methods included a silica-based filter purification (Qiaxtractor DX reagents; Qiagen, Valencia, California) following manufacturer's instructions, performed on a JANUS automated workstation (Perkin-Elmer, Waltham, Massachusetts), a sodium chloride protein precipitation (Miller et al. 1988) during which in the initial digestion of the sample, additional proteinase K and overnight digestion at 37°C was sometimes added to assist with breakdown of the skin sample, a phenol–chloroform method (Sambrook et al. 1989), and a DNeasy Blood & Tissue kit (Qiagen) following manufacturer's instructions.

Fin whale SNPs were ascertained from screening 39 loci (described in Aitken et al. 2004) with Sanger sequencing and from next-generation sequencing (NGS) of 50 loci (~80,000 bp) after capture enrichment of pooled, indexed samples as described in Hancock-Hanser et al. (2013). Sixteen of the loci screened with Sanger sequencing included SNPs for which primers could be designed (three loci included two separate SNPs). Eight SNPs were obtained from the sequence reads generated by NGS. Of the 27 SNPs obtained from the two screening methods, 23 were used to genotype the samples (three were removed for excess heterozygosity and one was removed for poor performance in quantitative PCR assays). The 23 SNPs were on 18 loci, with 13 loci having one SNP each and five loci having two SNPs each.

Single-nucleotide polymorphism genotypes were generated as described in Morin and McCarthy (2007) with a few minor modifications, which included using water instead of AE buffer in multiplex dilutions that were 25 µl instead of 20 µl, and each sample was only genotyped once (unless it was part of random replication or a control sample). In addition to using the program Amplifluor Assay Architect (Chemicon Inc., Temecula, California) to design SNP assays, WASP (http://bioinfo.biotec.or.th/WASP/index/wasp) with default parameters except: Ta min = 58, optimal = 62, max = 65, product range = 45–150 was also used. PCR conditions for each SNP assay and PCR primer sequences are given in Supplementary Data SD1 and SD2.

Seven samples were replicated from original DNA as genotype controls on every plate of 96 samples, and eight randomly chosen samples were used for replication. A second person independently called 20% of the assays from each plate. We determined SNP genotyping error rates based on duplicate genotyping of the 15 control and random samples. Samples with < 80% completed genotypes were excluded from all further analyses. Tests for Hardy–Weinberg equilibrium (HWE) as well as linkage disequilibrium were conducted with Genepop v.4.2.1 (Raymond and Rousset 1995), with *P*-values adjusted for multiple comparisons using the method of Benjamini and Yekutieli (2001). We stratified the North Pacific samples in the SNP data set into Clades A and C as defined in Archer et al. (2013). The eight North Pacific SNP genotypes that did not have mitogenome data were assigned to clades based on their 403 bp mtDNA control region sequence, for which there is a single fixed difference between Clades A and C.

The distribution of all samples used in this study along with an estimated distribution of the species from Edwards et al. (2015) is presented in Fig. 1. Because we do not have the ability to independently assign samples from the Southern Hemisphere samples to either *B. p. quoyi* or *B. p. patachonica*, we are unable to examine the status of these two currently recognized subspecies. Thus, we analyzed all samples from this ocean basin as a single stratum. Sample information with ocean basin assignment is given in Supplementary Data SD3. SNP genotypes are archived in the Figshare data repository (https://doi.org/10.6084/m9.figshare.8856092.v1).

Diversity and population structure.—Standard measures of genetic diversity (haplotypic diversity, proportion of unique alleles, number of variable sites, and heterozygosity), as well as measures of population structure (mtDNA: Nei's d_A, χ^2, θ_{ST}, and F_{ST}; SNPs: χ^2, F_{ST}, and F'_{ST}, and Jost's D) were estimated using the *strataG* package (Archer et al. 2017a) in R v3.3.1 (R Core Team 2016). Significance of population structure metrics was estimated using 1,000 permutation replicates. For the SNP analyses, these tests were run on all SNPs as well as all possible data combinations of SNPs combining the single-SNP loci with one SNP from each of the five two-SNP loci. From this latter analysis, we report the range of the test statistics and their permutation *P*-values for comparison with the same values obtained using all SNPs together.

We used a discriminant analysis of principal components (DAPC—Jombart and Ahmed 2011) to assess overlap of SNP genotypes. The DAPC was conducted with the *adegenet* v2.0.1 package in R (Jombart and Ahmed 2011). The first six principal components, accounting for 60% of the variability, were used

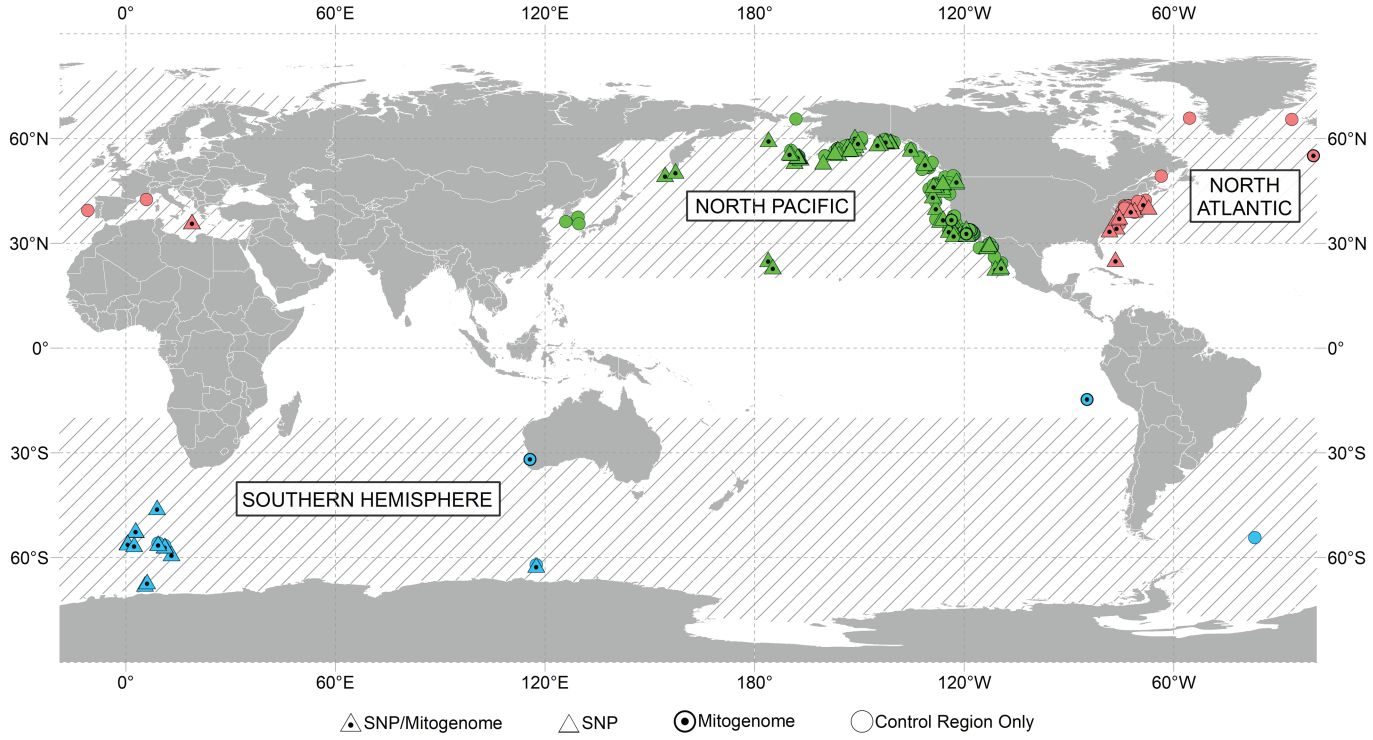

Fig. 1.—Location of samples with control region or full mitogenome sequences and single-nucleotide polymorphism (SNP) genotypes. Hashed region represents estimated global fin whale (*Balaenoptera physalus*) distribution from Edwards et al. (2015).

as classification results from analyses using more components were the same. All other parameters were left to their default settings.

We ran the Bayesian clustering algorithm STRUCTURE (Pritchard et al. 2000) to examine partitioning of the SNP genotypes with respect to ocean basins under the null model of HWE. We estimated membership assignment for one to five clusters (K), with 100 replicates for each value of K. For each replicate, 200,000 burn-in and 50,000 sampling steps were run. The distributions of the log-likelihood for each value of K across the 100 replicates were inspected for stability to confirm convergence. The models were run without admixture and without correlated allele frequencies. For each value of K, group membership estimates across replicates were aggregated using CLUMPP (Jakobsson and Rosenberg 2007). The most likely value of K was evaluated using the delta log-likelihood method of Evanno et al. (2005). We also computed the MedMeaK, MaxMeaK, MedMedK, and MaxMedK estimators of K proposed by Puechmaille (2016) which count the number of clusters to which at least one stratum belongs. We computed these estimators for a range of mean and median membership probabilities for each stratum from 0.5 to 0.9 in steps of 0.1.

Because STRUCTURE has been demonstrated to inaccurately recover group membership with small and unbalanced sample sizes (Kalinowski 2011; Puechmaille 2016; Wang 2017), we also followed the subsampling procedure of Puechmaille (2016). We ran 100 replicates for each value of K using the same parameters as above; however, for each replicate we drew a random 13 genotypes without replacement from the North Pacific and Southern Hemisphere to match the sample size of the North Atlantic, the smallest stratum. We estimated K from these replicates using the same estimators from Evanno et al. (2005) and Puechmaille (2016) as described above.

Diagnosability.—Following the methods described in Archer et al. (2017b), we assessed the diagnosability of ocean basins with a Random Forest analysis using the three mtDNA sequence data sets. The Random Forest analyses were conducted with the *randomForest* package (Liaw and Wiener 2002) in R. The three ocean basins (North Atlantic, North Pacific, and Southern Hemisphere) were used as a priori classes. Individual base pairs for each variable site in the mtDNA sequence were used as independent predictors. For each forest, 10,000 trees were grown. The sample size for each tree in the forest was set to half of the smallest sample size from all ocean basins in the data set under consideration. This was done to avoid classification bias due to uneven sample sizes (Berk 2006; Archer et al. 2017b). Samples for each tree were randomly selected without replacement. All other *randomForest* parameters were left at their defaults.

As a measure of diagnosability, we report the class-specific correct classification estimate. Given that the Random Forest models in this study were built on three classes, this value is the percent of individuals diagnosable with assignment probabilities ≥ 33.33% (PD_{33}). We also report the percent diagnosable with high confidence as PD_{80} and PD_{95} (assignment probabilities ≥ 80% and ≥ 95%, respectively). Central 95% confidence intervals for PD_{33} were calculated using a binomial distribution as described in Archer et al. (2017b).

RESULTS

Mitochondrial DNA.—In all three mtDNA data sets, overall haplotypic diversity as well as diversity within each ocean basin was relatively high (0.925–0.997; Table 1), but no haplotypes were shared among ocean basins. Although a small fraction of the variable sites in the mitogenome were in the full control region (54 of 912), a majority of the mitogenome haplotypes was defined by these control region variable sites. In fact, when the full mitogenome was used, there were only 21 more haplotypes than the 100 defined by the full control region, all of which were in the North Pacific. To increase the number of control region sequences in all ocean basins, the overall sequence was shortened to 233 bp. In this data set there were eight fewer variable sites than in the 403 bp sequences; however, the total number of haplotypes increased by 34. Although 261 new samples were added to the North Pacific, the number of haplotypes in this ocean basin decreased by 12 with the shorter sequence. However, increasing the number of samples in the North Atlantic and Southern Hemisphere increased the number of haplotypes to 43 and 56, respectively.

Table 1.—Mitochondrial DNA data set summaries. Columns are number of individual fin whales (*Balaenoptera physalus*) in each ocean basin (n), number of haplotypes (H), proportion of haplotypes that are unique (Pr(unique)), haplotypic diversity (D), mean nucleotide diversity (π), and the number of variable sites (VS).

Data set	Ocean basin	n	H	Pr(unique)	D	π	VS
16,390 bp Mitogenome	N.Atl	14	12	0.917	0.967	0.0017	96
	N.Pac	88	69	0.768	0.994	0.0041	490
	S.Hem	42	40	0.975	0.997	0.0027	432
	Total	144	121	0.851	0.997	0.0065	912
403 bp Control Region	N.Atl	14	12	0.917	0.967	0.0134	15
	N.Pac	88	48	0.583	0.980	0.0117	35
	S.Hem	42	40	0.975	0.997	0.0125	36
	Total	144	100	0.780	0.992	0.0181	54
233 bp Control Region	N.Atl	333	43	0.465	0.79	0.0125	23
	N.Pac	349	36	0.250	0.815	0.0068	27
	S.Hem	99	56	0.589	0.982	0.0146	32
	Total	781	134	0.463	0.925	0.0215	46

There were 27 fixed differences between the North Atlantic and the North Pacific and 28 between the North Atlantic and Southern Hemisphere in the mitogenome data set (Table 2). Of these, only two (North Atlantic versus North Pacific) and one (North Atlantic versus Southern Hemisphere) were found in the full control region, and none were found in the shorter control region data set. No fixed differences were found between the North Pacific and Southern Hemisphere in any of the data sets. However, there were 62 fixed differences between North Pacific Clade A and the Southern Hemisphere and 12 between North Pacific Clade C and the Southern Hemisphere in the mitogenome data set. Nei's measure of net nucleotide divergence (d_A) ranged from 0.0036 to 0.0072 for the full mitogenome and 0.0096 to 0.0213 for both control region data sets. Significant differentiation ($P \leq 0.05$) was found among all pairs of ocean basins for χ^2, F_{ST}, and θ_{ST} (Table 2).

Diagnosability (PD_{33}) from the Random Forest models was greater than 90% for all three mtDNA data sets (Table 3). The Random Forest models built on the mitogenome and 403 bp control region data set were able to correctly classify all samples from the North Atlantic and Southern Hemisphere. For the 233 bp data set, 97% of the North Atlantic sequences were correctly classified, with the other 3% being misclassified to the Southern Hemisphere. In this data set, all Southern Hemisphere sequences were correctly classified. Between 91% and 98% of the North Pacific samples were correctly classified, with the mitogenome control region data set performing worst, and the full mitogenome data set performing the best. The lowest 95% confidence interval of the diagnosability estimate was 77%, which was for the North Atlantic full mitogenome and 403 bp control region data sets.

The distribution of individual classification probabilities as measured by the fraction of trees in the forest voting for each ocean basin showed that individuals were classified to their ocean basins of origin with high certainty (Table 3; Fig. 2). PD_{80}, or the fraction of samples with classification probabilities ≥ 80% was over 80% for all ocean basins and all data sets. The more stringent value of PD_{95} was 50% or greater in all data sets except for the Southern Hemisphere mitogenome control region, where it was just 2%. However, in the expanded control region data set, a much higher proportion of Southern Hemisphere samples were correctly classified with high probabilities, also reflected by the PD_{80} and PD_{95} values of 93% and 66%, respectively (Table 3). We also examined the spatial distribution of individual assignment probabilities across the North Pacific and found no significant correlation with sampling location (χ^2 P-value > 0.5 for all three Random Forest models). This indicates that the diagnosability of a new North Pacific sample from the North Pacific should be the same across the range for which samples were available.

Single-nucleotide polymorphisms.—Four samples were removed from the SNP data set due to missing genotypes at more than three SNPs (two North Pacific Clade A, one North Pacific Clade C, and one Southern Hemisphere), leaving 150 samples. Of the 23 SNPs available, three were out of HWE in at least one ocean basin, which were also removed. Seven of the remaining 20 SNPs were variable in all strata, with the other 13 SNPs being fixed for a single allele in at least one stratum (Table 4). No ocean basins were fixed for an allele at a SNP that did not occur in at least one other ocean basin. No pairs of loci were found to be significantly in linkage disequilibrium. Mean heterozygosity was positively correlated with sample size, with the North Atlantic having the lowest diversity ($H = 0.11$), and the North Pacific having the greatest diversity ($H = 0.26$ and 0.27 for Clades A and C, respectively).

Estimated differentiation between North Pacific Clades A and C was very low (Jost's D, F_{ST}, and $F'_{ST} \leq 0.0007$) and nonsignificant for all four metrics (Table 5). Therefore, in further analyses, the North Pacific was treated as a single stratum. All three ocean basins were significantly differentiated from one another for χ^2, Jost's D, F_{ST}, and F'_{ST}. Differentiation was greatest between the North Pacific and the North Atlantic ($F_{ST} = 0.17$), and lowest between the North Pacific and Southern Hemisphere ($F_{ST} = 0.10$). The range of estimates of Jost's D, F_{ST}, and F'_{ST} for the 32 tests conducted on reduced data sets of 15 SNPs with one SNP chosen from each multi-SNP locus is small, and the range of P-values were highly significant ($P \leq 0.01$) for all tests, except between North Pacific Clades A and C which were all nonsignificant. Thus, all 20 SNPs were used together in further analyses.

The DAPC analysis produced very high classification scores (85–92%) for all three ocean basins (Table 6). The two misclassifications in the North Atlantic and three in the North Pacific Clade A were classified to the Southern Hemisphere, while one Southern Hemisphere sample was misclassified to the North Atlantic and two were misclassified to the North

Table 2.—Pairwise comparisons of mitochondrial DNA sequences from fin whales (*Balaenoptera physalus*) between ocean basins. Columns are Nei's net nucleotide divergence (d_A), number of fixed differences (FD), and three population structure metrics. For χ^2, only the P-value is given, while for F_{ST} and θ_{ST}, the test statistic is given with P-value in parentheses.

Data set	Ocean basins		d_A	FD	χ^2	F_{ST}	θ_{ST}
16,390 bp Mitogenome	N.Atl	N.Pac	0.0072	27	0.001	0.0184 (0.003)	0.442 (0.001)
	N.Atl	S.Hem	0.0036	28	0.038	0.0174 (0.01)	0.448 (0.001)
	N.Pac	S.Hem	0.0055	0	0.001	0.0046 (0.003)	0.433 (0.001)
403 bp Control Region	N.Atl	N.Pac	0.0154	2	0.001	0.0260 (0.003)	0.314 (0.001)
	N.Atl	S.Hem	0.0120	1	0.043	0.0174 (0.016)	0.365 (0.001)
	N.Pac	S.Hem	0.0096	0	0.001	0.0119 (0.001)	0.282 (0.001)
233 bp Control Region	N.Atl	N.Pac	0.0213	0	0.001	0.198 (0.001)	0.102 (0.001)
	N.Atl	S.Hem	0.0173	0	0.001	0.121 (0.001)	0.182 (0.001)
	N.Pac	S.Hem	0.0112	0	0.001	0.106 (0.001)	0.244 (0.001)

Table 3.—Confusion matrices from Random Forest analyses for each mitochondrial DNA data set from fin whales (*Balaenoptera physalus*) from the North Atlantic (N.Atl), North Pacific (N.Pac), and Southern Hemisphere (S.Hem). Rows are original strata in each data set, followed by three predicted strata. Last three columns are the proportion diagnosable (PD) with assignment probabilities of 33.33%, 80%, and 95%, respectively. For PD_{33}, the 95% *CI* from the binomial distribution is given in parentheses.

Data set	Original	Predicted			PD_{33} (95% *CI*)	PD_{80}	PD_{95}
		N.Atl	N.Pac	S.Hem			
16,390 bp Mitogenome	N.Atl	14	0	0	100 (77–100)	100	100
	N.Pac	0	86	2	98 (92–100)	88	83
	S.Hem	0	0	42	100 (92–100)	98	71
	Overall				99 (95–100)	92	81
403 bp Control Region	N.Atl	14	0	0	100 (77–100)	93	50
	N.Pac	0	80	8	91 (83–96)	83	58
	S.Hem	0	0	42	100 (92–100)	86	2
	Overall				94 (89–98)	85	41
233 bp Control Region	N.Atl	324	0	9	97 (95–99)	97	90
	N.Pac	0	323	26	93 (89–95)	91	90
	S.Hem	0	0	99	100 (96–100)	93	66
	Overall				96 (94–97)	94	87

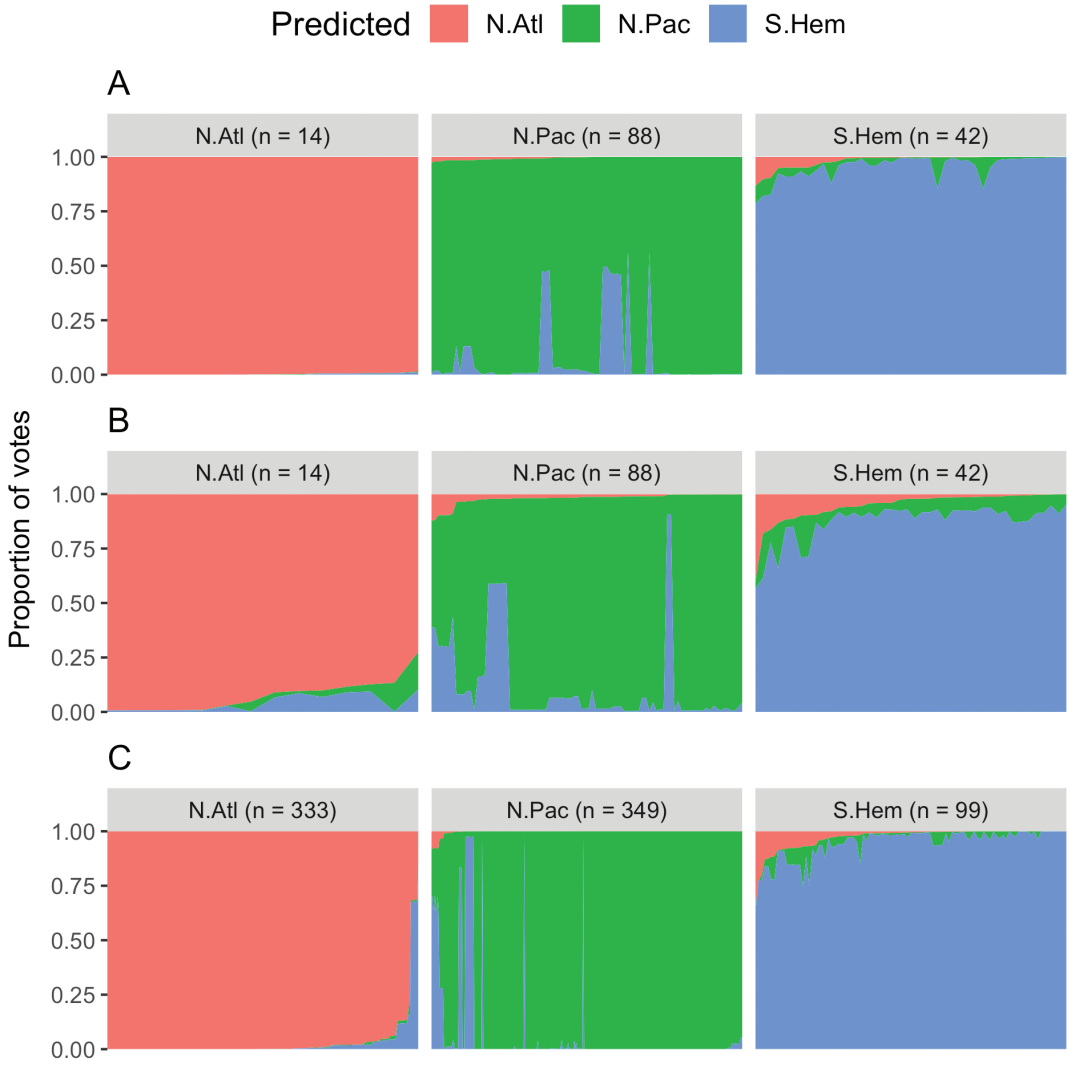

Fig. 2.—Distribution of classification probabilities for individual fin whales (*Balaenoptera physalus*) in each ocean basin from Random Forest models on three data sets: A) 16,390 bp Mitogenome, B) 403 bp Control Region, and C) 233 bp Control Region. Sample sizes are in parentheses. Within each ocean basin, individual samples are sequentially arranged along the x-axis.

Pacific. Mean membership probabilities to the original stratum were 80% for the North Atlantic and Southern Hemisphere and North Pacific, indicating that individuals that were correctly assigned were assigned with large certainty. The distribution of samples on the first two principal component axes (Fig. 3) illustrates the relative distance between the North Pacific and

Table 4.—Summary of single nucleotide polymorphisms (SNPs) from fin whales (*Balaenoptera physalus*) from the North Atlantic (N.Atl), North Pacific (N.Pac), and Southern Hemisphere (S.Hem). For each stratum, the columns are number of samples genotyped (*n*) at each SNP with the proportion in parentheses, the minor allele frequency (MAF), and observed heterozygosity (*H*). Note that the minor allele used to report MAF is minor with respect to all samples to facilitate comparison of frequencies across strata. SNPs on the same locus are grouped together.

SNP	N.Atl			N.Pac Clade A			N.Pac Clade C			S.Hem		
	n (%)	MAF	*H*	*n* (%)	MAF	*H*	*n* (%)	MAF	*H*	*n* (%)	MAF	*H*
BpADH2S156	13 (100)	0	0	85 (100)	0.01	0.02	14 (100)	0.04	0.07	38 (100)	0	0
BpADRBK1Y117	13 (100)	0	0	85 (100)	0.2	0.35	14 (100)	0.21	0.43	38 (100)	0	0
BpCGAS138	12 (92)	0.29	0.38	85 (100)	0.22	0.29	14 (100)	0.14	0.29	38 (100)	0.5	0.47
BpCKY126	13 (100)	0.58	0.38	85 (100)	0.15	0.25	14 (100)	0.32	0.64	37 (97)	0.31	0.5
BpCOL10A1K90	13 (100)	0	0	83 (98)	0.31	0.44	14 (100)	0.21	0.29	38 (100)	0	0
BpCOL3A1M72	13 (100)	0.23	0.15	85 (100)	0.4	0.47	14 (100)	0.36	0.29	36 (95)	0.68	0.34
BpCYP1A1R136	13 (100)	0	0	85 (100)	0	0	14 (100)	0	0	38 (100)	0.01	0.03
BpEPOY143	13 (100)	0	0	85 (100)	0.02	0.04	14 (100)	0	0	38 (100)	0	0
BpESDM147	13 (100)	0	0	85 (100)	0.34	0.47	14 (100)	0.21	0.29	38 (100)	0.3	0.55
BpESDY121	12 (92)	0.75	0.46	84 (99)	0.39	0.51	14 (100)	0.39	0.5	35 (92)	0.44	0.34
BpFESY21	13 (100)	0	0	85 (100)	0.05	0.09	14 (100)	0.04	0.07	38 (100)	0	0
BpFSHBM226	12 (92)	0	0	85 (100)	0.32	0.47	14 (100)	0.39	0.5	38 (100)	0.12	0.24
BpFSHBY120	13 (100)	0	0	85 (100)	0.11	0.18	14 (100)	0.18	0.36	38 (100)	0.12	0.18
BpGLUT2S255	13 (100)	0.08	0.15	85 (100)	0.12	0.2	14 (100)	0.07	0.14	38 (100)	0.32	0.53
BpGLUT2W136	13 (100)	0	0	85 (100)	0.17	0.25	14 (100)	0.18	0.36	38 (100)	0.01	0.03
BpLAPTM4AM476	13 (100)	0.73	0.54	85 (100)	0.07	0.14	14 (100)	0.04	0.07	37 (97)	0.34	0.45
BpRDSR128	13 (100)	0	0	84 (99)	0.09	0.18	14 (100)	0.18	0.36	38 (100)	0	0
BpRDSY373	13 (100)	0.08	0.15	85 (100)	0.16	0.32	13 (93)	0.12	0.21	38 (100)	0.07	0.13
BpRHO1R172	12 (92)	0	0	85 (100)	0.13	0.24	14 (100)	0.11	0.21	38 (100)	0	0
BpTPI1W157	13 (100)	0	0	84 (99)	0.19	0.33	14 (100)	0.32	0.36	38 (100)	0	0
Mean	12.8 (98)	0.14	0.11	84.75 (100)	0.17	0.26	13.95 (100)	0.18	0.27	37.65 (99)	0.16	0.19

Table 5.—Results of pairwise population structure analyses of single-nucleotide polymorphism (SNP) data from fin whales (*Balaenoptera physalus*) from the North Atlantic (N.Atl), North Pacific (N.Pac), and Southern Hemisphere (S.Hem). First column gives pairs of strata with sample sizes in parentheses. χ^2 is the *P*-value for the test, while all other columns give test statistics above and *P*-values below. Values outside of parentheses are the results from the full data set of 20 SNPs. Values in parentheses are the minimum and maximum values from the 32 tests of 15 SNPs with one SNP chosen from each multi-SNP locus.

Strata (*n*)	χ^2	Jost's *D*	F_{ST}	F'_{ST}
N.Pac.A (85) versus N.Pac.C (14)	0.539 (0.2667–0.6344)	0.0007 (0.0006–0.0011) 0.437 (0.159–0.5015)	0.0002 (−0.0018–0.0076) 0.442 (0.1598–0.5495)	0.0003 (−0.0025–0.0099) 0.442 (0.1598–0.5495)
N.Atl (13) versus N.Pac (99)	0.001 (0.001–0.001)	0.0171 (0.0113–0.0204) 0.001 (0.001–0.001)	0.1668 (0.167–0.1954) 0.001 (0.001–0.001)	0.2193 (0.214–0.2529) 0.001 (0.001–0.001)
N.Atl (13) versus S.Hem (38)	0.001 (0.001–0.004)	0.011 (0.0053–0.0132) 0.001 (0.001–0.0011)	0.1447 (0.1093–0.1596) 0.001 (0.001–0.001)	0.1743 (0.1281–0.1884) 0.001 (0.001–0.002)
N.Pac (99) versus S.Hem (38)	0.001 (0.001–0.001)	0.0164 (0.0137–0.0187) 0.001 (0.001–0.001)	0.098 (0.0974–0.1126) 0.001 (0.001–0.001)	0.1288 (0.1253–0.1436) 0.001 (0.001–0.001)

the North Atlantic with both being more similar to the Southern Hemisphere than either Northern Hemisphere ocean basin is to the other.

Using the delta log-likelihood method of Evanno et al. (2005) to interpret the results from the Bayesian clustering analysis, STRUCTURE, the most likely number of genetic clusters (*K*) in the SNP data set is two (Supplementary Data SD4). With $K = 2$, all North Atlantic samples and all but one Southern Hemisphere sample are assigned to one group with high probability, while both North Pacific clades are assigned to the other group (Fig. 4). With $K = 3$, North Atlantic and Southern Hemisphere both share large membership probabilities to one group, while North Pacific samples have their membership probabilities distributed between the other two groups. There was no difference in the probability distribution of group membership between Clades A and C (Supplementary Data SD5).

Additionally, the MedMeaK, MaxMeaK, MedMedK, and MaxMedK estimators (Puechmaille 2016) for the full SNP data set were all equal to two for all group membership thresholds examined. When the North Pacific and Southern Hemisphere were subsampled to equal sample sizes ($n = 13$), MedMeaK and MedMedK were two, and MaxMeaK and MaxMedK were three for a threshold of 0.5. For group membership thresholds greater than 0.5, all four estimators using results from the subsampled data indicated $K = 2$.

Discussion

The results of both the mtDNA and SNP analyses provide strong evidence that fin whales in each ocean basin are sufficiently differentiated to be considered separate subspecies. In the guidelines for delineating cetacean subspecies and species with mtDNA in Taylor et al. (2017a), the first threshold for distinguishing subspecies from populations is that subspecies should exhibit at least 95% diagnosability. The most comparable data set in this study to those used by Taylor et al. (2017a) to set this threshold is the 403 bp control region sequences (n = 144), for which the North Pacific was estimated to be 91% diagnosable, with the other two ocean basins being 100% diagnosable. However, the diagnosability estimates also contain some uncertainty, which is measured by the confidence interval (*CI*). For the North Pacific, the 95% *CI* spans 83–96%, which encompasses the threshold value, as do the confidence intervals for all ocean basins with all data sets. Additionally, diagnosability of the North Pacific is greater (98%) for the longer full mitogenome data set indicating that there are more diagnostic base pairs outside of the highly variable portion of the control region to resolve some of the misclassifications in the shorter sequences. This large and significant diagnosability in the North Pacific is also notable given the paraphyletic nature of sequences in this ocean basin due to the two very disparate mitogenome matrilines, Clades A and C in Archer et al. (2013), demonstrating that neither monophyly nor fixed differences are a requirement for the establishment of diagnostic characters.

The second threshold in the Taylor et al. (2017a) guidelines is that putative subspecies have values of Nei's net divergence (d_A) > 0.004. This measure of genetic distance is used as a proxy for the temporal scale of differentiation to ensure that groups that are highly diagnosable are not merely closely related family groups, and are on separate evolutionary trajectories (Archer et al. 2017b; Rosel et al. 2017a; Taylor et al. 2017a). Between North Atlantic and North Pacific fin whales, d_A was 0.021 and 0.015 for the 233 and 403 bp control region data sets, respectively, an order of magnitude greater than the threshold separating populations and subspecies, and close to or greater than the proposed threshold between subspecies and species (0.02). This large divergence is also reflected in the divergence dates estimated in the mitogenome phylogenetic tree presented in Archer et al. (2013). In that study, it was estimated that the Clade A maternal lineage in the North Pacific diverged

Table 6.—Confusion matrix from discriminant analysis of principal components (DAPC) analysis of single-nucleotide polymorphism (SNP) data set from fin whales (*Balaenoptera physalus*) from the North Atlantic (N.Atl), North Pacific (N.Pac), and Southern Hemisphere (S.Hem). Confidence interval (*CI*) of classification scores is from binomial distribution.

	Predicted			
Original	N.Atl	N.Pac	S.Hem	% Correct (95% *CI*)
N.Atl	11	0	2	85 (55–98)
N.Pac	0	96	3	97 (91–99)
S.Hem	1	2	35	92 (79–98)
Overall				95 (90–98)

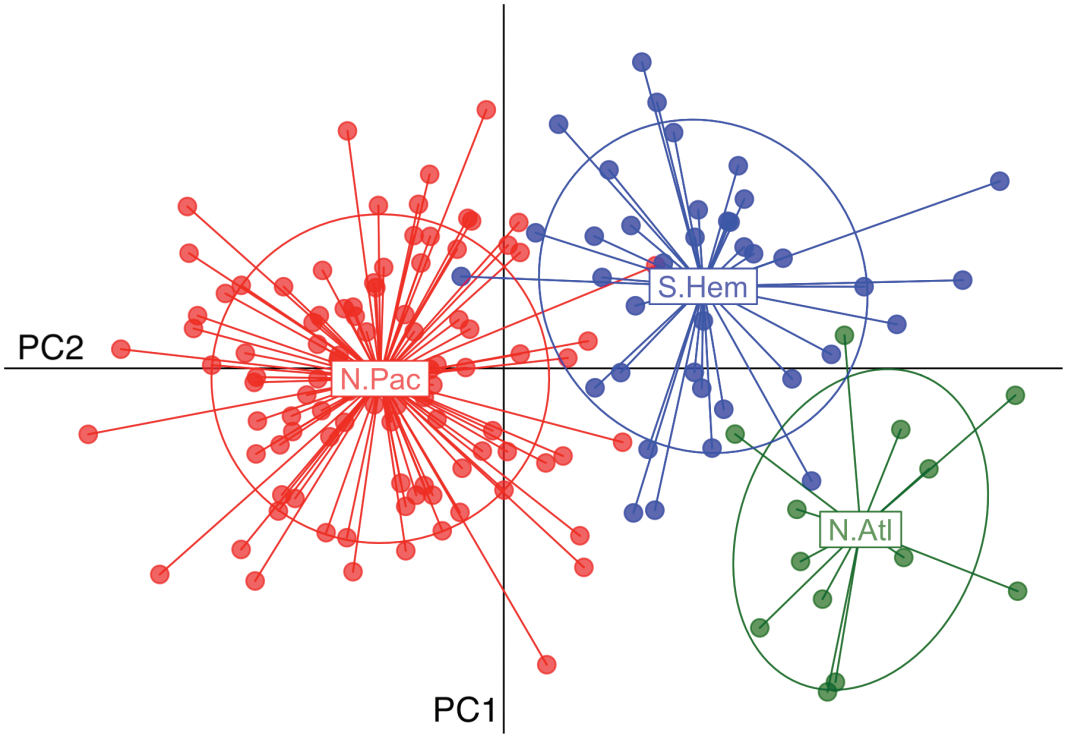

Fig. 3.—Scatterplot of single-nucleotide polymorphism (SNP) genotypes on first and second principal component axes from discriminant analysis of principal components (DAPC) analysis. Inertia ellipses around centroids represent 67% of the bivariate variance. Horizontal and vertical lines mark origin of each axis.

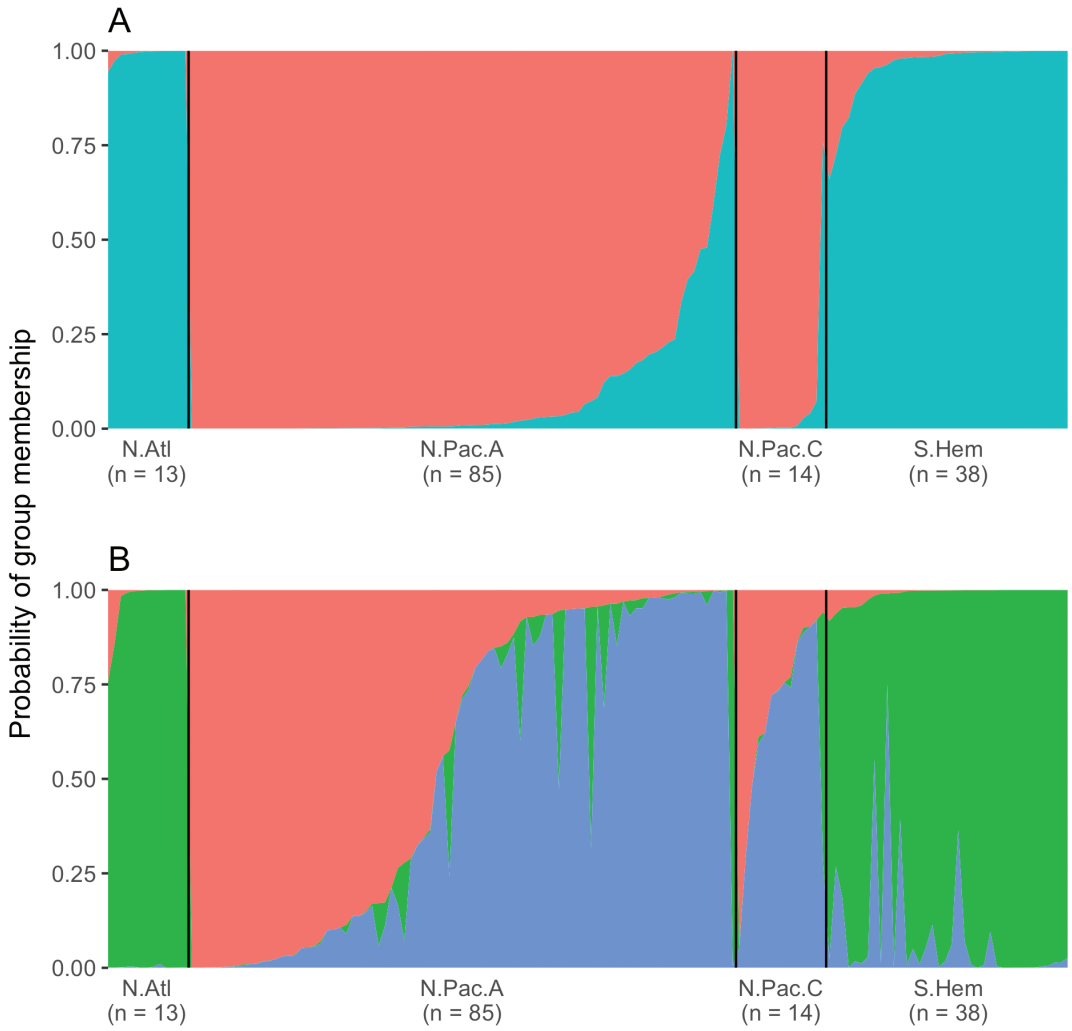

Fig. 4.—Distribution of group membership probabilities (y-axis) for individual fin whales (*Balaenoptera physalus*) in each ocean basin (x-axis) from the STRUCTURE analyses for A) $K = 2$, and B) $K = 3$ groups.

from the rest of the fin whale haplotypes approximately 1.8 MYA, then North Atlantic haplotypes formed a separate clade diverging approximately 1 MYA, and the second North Pacific maternal lineage (Clade C) was diverged from the Southern Hemisphere approximately 300 KYA.

After demonstrating high mitochondrial control region diagnosability (PD > 95%) and sufficient net nucleotide divergence ($d_A > 0.004$) for subspecies delimitation, and near-threshold net nucleotide divergence for species delimitation ($d_A \approx 0.02$), the final threshold for species delimitation is demonstrating that male-mediated dispersal is unlikely to be occurring (Taylor et al. 2017a). The results of the SNP analyses indicate strong nuclear differentiation among ocean basins, consistent with a lack of gene flow between whales of either sex between the North Atlantic and North Pacific and between the Northern and Southern Hemispheres. In fact, the greatest differentiation in both the mitochondrial and SNP analyses of allele frequency differences (F_{ST} and F'_{ST}) was between the North Atlantic and North Pacific. This was mirrored in the DAPC analysis where no samples were misclassified between these two Northern Hemisphere ocean basins.

Additionally, the STRUCTURE analysis indicates that the North Pacific is a distinct genetic cluster with respect to the North Atlantic and Southern Hemisphere. However, it did not find that the North Atlantic and Southern Hemisphere formed distinct genetic clusters. It is well recognized that STRUCTURE has difficulty correctly identifying groups represented by few and unbalanced samples (Kalinowski 2011; Puechmaille 2016; Wang 2017). Subsampling the North Pacific and Southern Hemisphere as suggested by Puechmaille (2016) to help alleviate these issues also identified two as being the most likely number of clusters. Given this finding and the significant differentiation seen in the standard tests of population structure and the DAPC, the lack of resolution in the STRUCTURE analysis between the North Atlantic and Southern Hemisphere most likely stems from the small number of samples representing the North Atlantic ($n = 13$).

Finally, the degree of mitogenomic and nuclear differentiation between the North Pacific and Southern Hemisphere is consistent with a lack of contemporary gene flow. The trans-equatorial migration event from the Southern Hemisphere

that led to North Pacific Clade C approximately 300 KYA Archer et al. (2013) appears to have been the last contact. Given their distribution in temperate to subpolar waters and oceanic warming trends, the likelihood of future north-south exchanges decreases even more. Also, the lack of nuclear differentiation between Clades A and C indicates that these two matrilines are not evidence of current barriers to gene flow within the North Pacific, but rather only reflect a single historical migration event.

The status of the two currently recognized subspecies of fin whales within the Southern Hemisphere cannot be addressed with these data. The pygmy fin whale (*B. p. patachonica*) was originally described from a specimen that stranded at the mouth of Rio Plata, Argentina (Burmeister 1865) as being darker in color than *B. p. quoyi*, and inhabiting more lower-latitude waters in the Southern Ocean (50–60°S) than *B. p. quoyi*. Clarke (2004) also provided observations on this form from the South Atlantic and the eastern tropical Pacific off Ecuador, suggesting they may occur farther north as well. There is no type specimen for the larger "Southern Hemisphere" subspecies (*B. p. quoyi*) that was described from the Falkland Islands–Islas Malvinas (Desmoulins 1822), approximately 1,000 nm to the south of the location of Burmeister's (1865) stranded specimen of *B. patachonica* at the mouth of Rio Plata.

To date, there has been no genetic confirmation of differences between these subspecies, largely because, to our knowledge, no tissue samples have been collected that have been positively assigned to one form or the other in the Southern Hemisphere. Most of the Southern Hemisphere samples in this study were collected during 2 years of the IWC-SOWER cruise (Ensor et al. 2007, 2008) off South Africa. Although Clarke's descriptions of the pygmy form are from waters around South America, some are from similar latitudes as the IWC-SOWER sample (around 60°S), and it is unknown if their ranges overlap. Thus, even tentative assignments of samples are difficult. It is clear that in order to elucidate the taxonomy of Southern Hemisphere fin whales, collection of tissue samples associated with photos or descriptions of color pattern and morphology (sizes) is necessary. Given the mitogenomic diversity of samples from the Southern Hemisphere (Archer et al. 2013), it is possible that our sample contains specimens from both recognized subspecies, or from unrecognized taxonomic structure in the Southern Hemisphere resulting in underestimation of net nucleotide divergence (d_A) between Southern and Northern Hemisphere subspecies. However, the STRUCTURE analysis did not detect more than one group in our Southern Hemisphere samples, so if our Southern Hemisphere samples do contain a mix of *B. p. quoyi* and individuals referred to as pygmy fin whales by Clarke (2004), they are unlikely to be as differentiated from each other as the Southern Hemisphere is to either the North Atlantic or North Pacific.

The current data are consistent with the following three different taxonomic hypotheses: 1) there is one species of fin whale with three subspecies (North Atlantic, North Pacific, and one in the Southern Hemisphere), 2) the North Atlantic is a separate species relative to the North Pacific and Southern Hemisphere combined, which is itself represented by three subspecies, and 3) each ocean basin should be recognized as a separate species with the Southern Hemisphere represented by two subspecies, one of which has yet to be confirmed with genetic analyses.

There are few other sources of information to help distinguish among these hypotheses. Traditionally, morphometric data have been used to demonstrate species-level differentiation. Both North Pacific and North Atlantic fin whales are reported to be smaller than the Southern Hemisphere *B. p. quoyi* (Mizroch et al. 2009). The largest fin whale recorded from the North Pacific was 23 m (74 feet—Nichol and Heise 1992), while the largest confirmed length from the Southern Hemisphere was 25.9 m (85 feet—Mackintosh 1942). Most adult North Pacific fin whales are typically between 15 and 21 m (50–70 feet), with females usually 0.6–1.2 m longer than males (2–4 feet—Nichol and Heise 1992; Clapham et al. 1997). Ichihara (1957) reported that fin whales found in the ECS were of different shapes from those east of the Kamchatka Peninsula (KP) with those off KP having longer heads and shorter tails. Unfortunately, little systematically collected morphometric data from fin whales are available from these three ocean basins to permit a better comparison.

Acoustics is proving to be a useful tool for characterizing regional differences among baleen whales at various taxonomic scales (Payne and Guinee 1983; McDonald et al. 2006). Across their range, fin whales are characterized by a set of short 20-Hz downsweep calls, which are believed to be produced only by males (Watkins et al. 1987; Delarue et al. 2009; Širović et al. 2013). A higher-frequency call (~90–100 Hz) has been described in Southern Hemisphere fin whale calls (Širović et al. 2004, 2009) and is periodically present in the North Atlantic at a centroid frequency of 132 Hz (Simon et al. 2010), but has not been described from the North Pacific (Širović et al. 2013). Additionally, a 40-Hz call has been described in the North Pacific (Watkins 1981; Širović et al. 2013), but does not appear in other ocean basins. Thus, there is some evidence of diagnostic differences among calls from the various ocean basins. However, because most studies tend to focus on the lower-frequency, stereotypical 20-Hz call, a determination of the consistency of these differences in the higher-frequency calls awaits a detailed broadband comparison of global recordings.

Thus, genetic data will likely be the best source of information for clarifying fin whale taxonomy. Until more samples can be obtained from the Southern Hemisphere with known subspecies provenance (correlated with morphology), more samples added from the North Atlantic, particularly the eastern North Atlantic and Mediterranean, and more nuclear loci are genotyped for all ocean basins, we will be unable to definitively say which of the three hypotheses best describes the global taxonomic status of fin whales.

Regardless, it is clear that the North Pacific should not be classified as the same subspecies as the North Atlantic. Thus, we propose recognizing the North Pacific as a separate subspecies until this uncertainty can be resolved. If more evidence is produced to support it, the North Atlantic could be elevated to its own species, leaving the North Pacific a subspecies of

a species including the Southern Hemisphere, or the North Pacific could be elevated to a full species.

The nominate form of the current species, *B. physalus*, was originally described from the eastern North Atlantic (Linnaeus 1758). It was previously believed that the oldest name in the synonymy of *B. physalus* in the North Pacific was *Balaenoptera swinhoei* Gray, 1865. The type specimen for this nominal species was collected from southern Taiwan. However, recent genetic evidence shows that this specimen is actually a North Pacific humpback (*Megaptera novaeangliae kuzira*) making *B. swinhoei* invalid for North Pacific fin whales (Archer et al. 2018). The Gray (1865) specimen, therefore, becomes a junior synonym of *M. n. kuzira*.

The oldest name for the North Pacific fin whale is *Balaenoptera velifera* Cope in Scammon 1869. E. D. Cope originally described this species as "The Finner Whale of the Oregon coasts," but did not associate the description to a specimen that could be designated as a holotype. True (1904) synonymized this species with *B. physalus* based on a comparison of measurements of two specimens from the U.S. West Coast with Atlantic fin whales. As detailed in the subspecies descriptions that follow, neither of these specimens can currently be located; thus, we propose to designate a specimen housed at the Los Angeles County Museum (LACM 54761) as the neotype for the North Pacific subspecies.

Diversification between hemispheres, as well as between the North Atlantic and North Pacific on evolutionary time scales (1–4 Mya) is common in other baleen whale taxa. The three right whale species (*Eubalaena* spp.) have been separated long enough to establish 6–7 diagnostic sites in 292 bp of the mtDNA control region between ocean basins. Cessation of contact between North Atlantic (*E. glacialis*) and North Pacific (*E. japonica*) right whales was estimated at 3.5 Mya with the closing of the isthmus of Panama (Rosenbaum et al. 2000). Antarctic and common minke whales (*Balaenoptera bonarensis* and *B. acutorostrata*) were estimated to have diverged approximately 4.4–4.9 Mya, and the divergence date of the three common minke whale subspecies has been estimated at 1.2 Mya (Pastene et al. 2007). In a recent review of global divergence of humpback whales (*Megaptera novaeangliae*) based on mtDNA control region sequences and eight nuclear loci, Jackson et al. (2014) proposed three oceanic subspecies representing the North Atlantic (*M. n. novaeangliae*), the North Pacific (*M. n. kuzira*), and the Southern Hemisphere (*M. n. australis*). It was estimated that major humpback mtDNA clades diverged in the last million years, radiating in the various ocean basins approximately 200 Kya (Jackson et al. 2014). Although the authors found no mtDNA diagnostic sites, there was significant frequency differentiation among ocean basins and estimates of mtDNA and nDNA migration rates were sufficiently low to indicate that ocean basins were along independent evolutionary trajectories. Additionally, they found lower migration rates between Southern Hemisphere and Northern Hemisphere subspecies, reinforcing the suggestion of the equator as an efficient barrier to gene flow.

Although there have been fewer global genetic studies, several other globally distributed species of baleen whales exhibit patterns more complex than simple partitioning by ocean basin, also suggesting that the true picture of fin whale taxonomy has yet to be fully recognized. Blue whales (*Balaenoptera musculus*) currently comprise of a single Northern Hemisphere subspecies (*B. m. musculus*), and four Southern Hemisphere subspecies: Antarctic (*B. m. intermedia*), northern Indian Ocean (*B. m. inidica*), pygmy (*B. m. brevicauda*), and an unnamed Chilean subspecies (Committee on Taxonomy 2018). While significant genetic differentiation between blue whales on either side of the equator has been detected (Conway 2005; LeDuc et al. 2007, 2017), there has been little phylogeographic resolution in mitochondrial control region sequences and thus no estimates of divergence time between hemispheres or between subspecies in the Southern Hemispheres have been generated. In addition, very limited data are available to compare North Atlantic and North Pacific blue whales. Within Bryde's whales (*B. edeni*), two subspecies are currently recognized: the coastal Eden's whale (*B. e. edeni*) found in the western Pacific and Indian Ocean, with the remainder of the global offshore Bryde's whale form being assigned to *B. e. brydei* (Committee on Taxonomy 2018). However, Rosel and Wilcox (2014) identified a distinct lineage of Bryde's whales in the Gulf of Mexico. In 375 bp of the mtDNA control region, between 25 to 52 fixed differences were found between Gulf of Mexico Bryde's whales and lineages of both *B. e. edeni* and *B. e. brydei*, respectively, with this unique lineage more closely related to *B. e. edeni*. This finding suggests that more taxonomic structure than is currently recognized exists in Bryde's whales, warranting a detailed global review of the taxonomy of the species. With the rapid evolution of high-throughput sequencing technologies (Andrews et al. 2016; Goodwin et al. 2016), we are at a point where data sets to elucidate these complex phylogenies can be generated. The limiting factor will still be ensuring that enough samples from across the range of these species are available.

Systematics

Order Cetartiodactyla Montgelard, Catzefils and Douzery, 1997

Family Balaenopteridae Gray, 1864

Subfamily Balaenopterinae Gray, 1864

Genus *Balaenoptera* Lacépède, 1804

Balaena physalus Linnaeus, 1758:75

Balaenoptera physalus Schlegel 1862:101; first use of current name combination

Balaenoptera physalus physalus (Linnaeus, 1758)

Tomilin 1946 first use as subspecies

North Atlantic fin whale

Type specimen.—No type specimen was designated. First described by Friderich Martens in 1675. Based on these descriptions, Linnaeus described his *Balaena physalus* in 1758. The

first combination of the current name was not until Schlegel (1862) made a connection with the genus *Balaenoptera* described by Lacépède (1804).

Type locality.—"Oceano Europaeo" (northeast Atlantic Ocean) and specifically as the region between Spitzbergen and Europe as it is based on Marten's (1675) "finfisch."

Etymology.—The specific name, "*physalus*" is derived from the Greek "physalos." In the literature, there appear to be several potential derivations. In the first, it is most frequently translated as "puffed up toad" (Leatherwood and Reeves 1983), which likely comes from its use in Lucian (Philopseudes "The Lover of Lies or Cheater" 12). However, Lucian (Diapsids 3) also uses the term to refer to the puff adder, a snake. Additionally, Aelian (Naturae Animalus "On the Character of Animals" 3.18) uses the word to describe a puffer fish of the family Tetraodintidae, and then again in 9.49 to directly refer to a large whale. Therefore, the term may simply mean "one that is puffed up," and may refer to the throat pleats or bloated beached whales. Others believe its origin comes from the root of the Greek word "physa," meaning "blows" and refers to the prominent blows of fin whales, similar to the root of *Physeter*, the sperm whales. Martens (1675:132) describes them: "They know the finn-fish by the... vehement blowing and spouting up of the water...."

Diagnosis.—The species was originally described from the northeast Atlantic Ocean. No morphological justification for the split between Northern (*B. p. physalus*) and Southern Hemisphere (*B. p. quoyi*) fin whales has been offered other than the larger average size of fin whales in the Antarctic as opposed to those in the Northern Hemisphere, supposedly engendered by the cold Antarctic water (Tomilin 1946). If some fin whales are resident in lower latitudes of the Southern Hemisphere (not usually migrating to the Antarctic), and if one accepts Tomilin's hypothesis that larger size is connected to cold water, the lower-latitude whales may not be larger than their northern relatives. All Southern Hemisphere balaenopterids are larger than their Northern Hemisphere counterparts (Tomilin 1957). Indeed, Clarke (2004) noted smaller size of fin whales from lower latitudes in his paper suggesting they belonged to the subspecies *B. p. patachonica* based on *B. patachonica* Burmeister, 1865. In the mitogenome, there are 28 fixed differences between the North Atlantic and Southern Hemisphere, and 27 between the North Atlantic and North Pacific.

Distribution.—The nominate subspecies occurs throughout most of the cooler waters of the North Atlantic basin. In the western North Atlantic, from the U.S. east coast normally around 30°N (to rarely south into the Gulf of Mexico) to the ice edge in the Labrador Sea around 70°N. In the eastern North Atlantic, from approximately 30°N around Las Canarias to approximately 80°N in the Barents Sea and Arctic Ocean (Edwards et al. 2015). Also includes fin whales found in the Mediterranean Sea.

Balaenoptera physalus quoyi (Fischer, 1829)
Southern Hemisphere fin whale

Type specimen.—No type specimen was designated. This subspecies was first described as *Balaena rostrata australis* by Desmoulins (1822). He used the trinomial but referred to the form as a species. The name *Balaena rostrata australis* was preoccupied by *Balaena australis* (= *Eubalaena australis*), which Desmoulins had applied to the southern right whale earlier in the same paper. Fischer (1829) therefore renamed the taxon *Balaena quoyi*, interpreting Desmoulins' form with a trinomial name as being a species. Desmoulins noted that his holotype whale (not collected), observed by Quoy off the Falkland Islands–Malvinas, was 16.7 m (55 feet), which was twice as long as whales of the species *B. rostrata* (minke whales) in the North Atlantic. His descriptions of baleen, dorsal fin, and throat grooves match the fin whale. The Swedish zoologist Einar Lönnberg (1906) showed that a set of external measurements of a specimen of *Balaenoptera quoyii* taken off of South Georgia Island were similar to those of "North Atlantic Finners." However, one measurement, length from snout to posterior margin of dorsal fin as a percent of total body length, was significantly greater in this specimen, which, in addition to differences in coloration led Lönnberg to speculate that this form was actually a subspecies of fin whales. Based on differences in vertebral morphology Lönnberg (1931) more formally proposed that Southern Hemisphere fin whales be recognized as the subspecies, *B. physalus quoyii*, which was later corrected to *B. p. quoyi* by Tomilin (1957). Lönnberg interpreted the name to apply to all fin whales in the Southern Hemisphere.

Type locality.—Falkland Islands–Islas Malvinas.

Etymology.—Named after naturalist and naval surgeon Jean René Constant Quoy, who observed the specimen off the Falkland Islands while on the *Astrolabe* (1826–1829) under the command of Jules Dumont d'Urville when Fischer (1829) provided the new name for Desmoulins' *Balaena rostrata australis*.

Diagnosis.—No morphological justification for the split between Northern (*B. p. physalus*) and Southern Hemisphere (*B. p. quoyi*) fin whales has been offered other than the larger average size of fin whales in the Antarctic as opposed to those in the Northern Hemisphere, supposedly engendered by the cold Antarctic water (Tomilin 1946). If some fin whales are resident in lower latitudes of the Southern Hemisphere (not usually migrating to the Antarctic), and if one accepts Tomilin's hypothesis that larger size is connected to cold water, the lower-latitude whales may not be larger than their northern relatives. Indeed, Clarke (2004) noted smaller size of fin whales from lower latitudes in his paper suggesting they belonged to the subspecies *B. p. patachonica* based on *B. patachonica* Burmeister, 1865. No genetic samples have been positively associated with *B. p. quoyi* or *B. p. patachonica*. There are 28 fixed differences in the mitogenome between the Southern Hemisphere and the North Atlantic. Although there are no fixed differences between the entire North Pacific and the Southern Hemisphere, there are 62 and 12 fixed differences relative to North Pacific Clades A and C, respectively.

Distribution.—The subspecies is presumed to occur throughout the Southern Ocean between approximately 20°S and 80°S (Edwards et al. 2015).

Balaenoptera physalus patachonica Burmeister, 1865
pygmy fin whale

Type specimen.—Burmeister (1865) listed two specimens in his description of *Balaenoptera patachonica*. One a partial but nearly complete skeleton, that was determined to be a juvenile based on the state of the cervical vertebrae. The other specimen was "the hinder part [basicranium], without the jaws." Burmeister (1865) illustrated the skull, several cervical vertebrae, scapula and rib heads from the nearly complete skeleton, as well as providing several measurements. Therefore, we consider this specimen to be the type. Both specimens were deposited in the Museo Argentino de Ciencias Naturales Bernardino Rivadavia (MACN) in Buenos Aires, Argentina. However, only the scapula (Figure 10 in Burmeister 1865) remains in the collection (P. Teta, MACN, pers. comm., 19 January 2016 to FIA).

Type locality.—"some leagues from Buenos Ayres, on the banks of the River Plata" [Argentina].

Etymology.—Refers to the Patagonia region of Argentina.

Diagnosis.—According to Clarke (2004) this whale is small (total length < approximately 20 m) and has a dark body color. Burmeister (1865) reported that they had all black baleen. No genetic samples have been positively associated with *B. p. quoyi* or *B. p. patachonica*. There are 28 fixed differences in the mitogenome between the Southern Hemisphere and the North Atlantic. Although there are no fixed differences between the entire North Pacific and the Southern Hemisphere, there are 62 and 12 fixed differences relative to North Pacific Clades A and C, respectively.

Distribution.—Occurring in the Southern Ocean from around 35°S (Burmeister 1865) to approximately 55°S or further south, and to the Equator along the Pacific coast of South America (Clarke 2004).

Remarks.—Burmeister (1865) presumed his specimen might be the same as *Balaenoptera australis* (Desmoulins 1822, = *B. quoyi* Fischer 1829). However, since he had not seen that specimen, he applied a new name (*Balaenoptera patachonica*) to his specimen out of a sense of caution. J. E. Gray (in Burmeister 1865) agreed that Burmeister's specimen was a fin whale. The only distinctive character, different from other fin whales, that Burmeister (1865) reported for his specimen was its all black baleen, a color normally only found in blue whales. In the Clarke (2004) proposal for a new subspecies bearing the name *B. physalus patachonica*, the only specimen of this putative subspecies that he fully examined was a single whale that had "just attained physical maturity although it was only 19.8 m long" (Clarke 2004:330–331). This whale was noted to be dark in color, recently sexually mature based on ovarian scars, and collected near the ice edge at 63°48′S, 86°37′E. Burmeister's specimen was physically immature and the length of the mandible was 310 cm (10 feet 2 inch). Using the formulae from Pyenson et al. (2013) the mandible length can be converted to a total length of 13.1 m. Burmeister (1865) was unable to describe the body color of his specimen, and although Clarke (2004) lists the black baleen as a feature of his proposed subspecies, he did not comment on the baleen color of his 19.8 m specimen. Given that no positive link was made between Clarke's specimen and Burmeister's specimen, more work needs to be done to fully evaluate the taxonomic status of *B. p. patachonica.*

Balaenoptera physalus velifera Cope in Scammon, 1869
North Pacific fin whale

Type specimen.—No type specimen was designated. True (1904) published a detailed description of two specimens of *B. velifera*, one collected by E. D. Cope (not the type) and only identified as coming from the U.S. West Coast, and the other captured off of San Clemente Island, California. The Cope specimen was a fully articulated skeleton, measuring 18.9 m when mounted, and was bequeathed to the Wistar Institute in Philadelphia. It was hung in the upper exhibition hall sometime between 1898 and 1899 (Jayne 1898, 1899). In 1955, it was sent to the Field Museum in Chicago, where it was intended to replace a right whale skeleton on exhibition (see Supplementary Data SD6 for copy of shipping receipt and Gregg 1956). However, there is no record that it was ever actually hung in an exhibition hall, and the right whale skeleton that it was to replace is still on display as of this writing. The skeleton cannot currently be found in the collection of the Field Museum and does not appear in the records of any other major U.S. natural history museum collection.

The second specimen was exhibited in a field outside of an amusement park in Celoron, New York from July to August, 1896 by Captain C. H. Shirley of the U.S. Marine Aquarium (Jamestown Evening Journal, 13 July 1896 and 31 August 1896). It was reported to be 16.8 m (55 feet) long, weighing 36,287 kg (80,000 lbs), and described as being "fully fleshed." The specimen was reported to have been enroute to the National Museum in Washington, DC, United States (Jamestown Evening Journal, 13 July 1896). However, it does not appear in the records of the Smithsonian's National Museum of Natural History, and its current location is unknown.

Given that no type specimens are available, we here designate as a neotype the skeleton of a whale, nicknamed "Minnie," that stranded on 20 May 1897 at Alamitos Avenue, Long Beach, California (Grobaty 2012). The specimen is stored at the Natural History Museum of Los Angeles County (LACM 54761). Our choice of this specimen as a neotype was based on it being the oldest complete skeleton within the known range of the subspecies, with high-quality associated data, and one for which we could extract diagnostic DNA (see Morin et al. 2006 for extraction methods). From bone powder of this specimen (Southwest Fisheries Science Center LabID 174470), we generated 411 bp of mitochondrial control region sequence (GenBank accession number MK319943) that is a 100% match to nine other North Pacific sequences in this study. A full description, measurements, and photos of the specimen are provided in Supplementary Data SD7 and SD8.

Type locality.—Long Beach, California, United States.

Etymology.—From the Latin "velifer," meaning "sail-bearing." First described by Cope (Scammon 1869:16) as the

"Finner whale of the Oregon coasts." No description of the source of the name has been published, but possibly refers to the large falcate dorsal fin noted by Scammon (1869:52).

Diagnosis.—Balaenoptera physalus velifera is currently not easily distinguishable from other subspecies of *B. physalus* by diagnostic morphological characters. However, it is diagnosable from the North Atlantic by 27 fixed differences in the mitogenome. There are no fixed differences between the entire North Pacific and the Southern Hemisphere; however, there are 62 and 12 fixed differences relative to North Pacific Clades A and C, respectively.

Specimens examined.—Genetics = 349 mtDNA control region, 88 full mitogenome, and 85 for 23 SNP loci. Morphology = 1. See Supplementary Data SD3 for collection information of genetic samples.

Distribution.—The range of *B. p. velifera* extends from the Gulf of California, along the western coast of the United States and British Columbia, Canada into the Gulf of Alaska, and along the Aleutians. They are found in the Bering Sea and into the Chukchi Sea up to approximately 70°N (Edwards et al. 2015) and genetic samples match with others from the along the west coast of North America. In the western Pacific, they are found off of Kamchatka, Okhotsk Sea, and Japan. They also occur in the northern waters of Hawaii, although in lower numbers (Barlow 2006), and genetic samples from this region are diagnosable to this subspecies.

Genetic samples in this study from the western Pacific were sparse. Therefore, we do not include fin whales found in the ECS and the Sea of Japan in the range of *B. p. velifera*. Several lines of evidence indicate that they are significantly different from whales in the western North Pacific (Ichihara 1957; Fujino 1960; Brownell 1981). A linear discriminant function analysis of seven external measurements collected from 276 individuals produced correct classification rates of approximately 93% when ECS fin whales were compared to those off of Kamchatka and the Aleutian Islands (Ichihara 1957). ECS fin whales were noted to have relatively shorter heads, and longer tails. Likewise, in a study of blood immunogenetic markers, Fujino (1960) found ECS fin whales had a significantly greater frequency of the Ju2 phenotype and a lower frequency of Ju1 than those off Kamchatka or the Aleutians. Thus, it is possible that ECS fin whales are more than just a resident population and may even be a subspecies distinct from the rest of the North Pacific. However, further genetic studies focused on this region would help clarify their status.

Acknowledgments

The authors thank D. Janiger of the Natural History Museum of Los Angeles County for sampling bone powder from and facilitating access to LACM 54761. We also thank the command, crew, and marine mammal observers sailing on the NOAA Ships *David Starr Jordan*, *MacArthur*, and *MacArthur II*, and the International Whaling Commission Southern Ocean Whale and Ecosystem Research Programme (IWC-SOWER) for their assistance in obtaining biopsy samples for this project. We appreciate the help provided by Dr. L. Heaney of the Field Museum of Natural History, Chicago, and Dr. C. Potter and Dr. J. Mead of the Smithsonian Institution Natural History Museum in the quest for type specimens. T. J. Moore assisted in producing the sample map.

Supplementary Data

Supplementary data are available at *Journal of Mammalogy* online.

Supplementary Data SD1.—PCR conditions for single-nucleotide polymorphism (SNP) loci.

Supplementary Data SD2.—Multiplex and Amplifluor single-nucleotide polymorphism (SNP) PCR primers.

Supplementary Data SD3.—Genetic samples used in analyses. GenBank accession numbers for sequences published in Bérubé et al. (1998; n = 319 samples, 51 haplotypes) are AF119956–AF120006. Fifty-seven sequences from South Georgia from Sremba et al. (2015) are archived at Dryad (http://doi.org/10.5061/dryad.c8k6q).

Supplementary Data SD4.—Evanno plots for Structure results with 23 single-nucleotide polymorphisms (SNPs).

Supplementary Data SD5.—Distribution of group membership probabilities (y-axis) for individuals in each strata (x-axis) from Structure analyses for A) K = 2, and B) K = 3.

Supplementary Data SD6.—Shipping receipt indicating delivery of specimen labeled *Balaenoptera velifera* from the Wistar Institute to the Chicago Field Museum.

Supplementary Data SD7.—Description and measurements of neotype of *Balaenoptera physalus velifera*, Cope in Scammon 1869.

Supplementary Data SD8.—Images of LACM 54761. Scale bars in photos measure 55 cm.

Literature Cited

Aitken, N., S. Smith, C. Schwarz, and P. A. Morin. 2004. Single nucleotide polymorphism (SNP) discovery in mammals: a targeted-gene approach. Molecular Ecology 13:1423–1431.

Amos, W. 1997. Marine mammal tissue sample collection and preservation for genetic analyses. Pp. 107–113 in Molecular genetics of marine mammals (A. E. Dizon, S. J. Chivers, and W. F. Perrin, eds.). Society for Marine Mammalogy.

Andrews, K. R., J. M. Good, M. R. Miller, G. Luikart, and P. A. Hohenlohe. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews: Genetics 17:81–92.

Archer, F., K. Robertson, R. Sabin, and R. L. J. Brownell. 2018. Taxonomic status of a "finner whale" (*Balaenoptera swinhoei* Gray, 1865) from southern Taiwan. Marine Mammal Science 34:1134–1140.

Archer, F. I., P. E. Adams, and B. B. Schneiders. 2017a. strataG: an R package for manipulating, summarizing and analysing population genetic data. Molecular Ecology Resources 17:5–11.

Archer, F. I., K. K. Martien, and B. L. Taylor. 2017b. Diagnosability of mtDNA with Random Forests: using sequence data to delimit subspecies. Marine Mammal Science 33:101–131.

Archer, F. I., et al. 2013. Mitogenomic phylogenetics of fin whales (*Balaenoptera physalus* spp.): genetic evidence for revision of subspecies. PLoS One 8:e63396.

Barlow, J. 2006. Cetacean abundance in Hawaiian waters estimated from a summer/fall survey in 2002. Marine Mammal Science 22:446–464.

Benjamini, Y., and D. Yekutieli. 2001. The control of the false discovery rate in multiple testing under dependency. The Annals of Statistics 29:1165–1188.

Berk, R. 2006. An introduction to ensemble methods for data analysis. Sociological Methods and Research 34:263–295.

Bérubé, M., et al. 1998. Population genetic structure of North Atlantic, Mediterranean Sea and Sea of Cortez fin whales, *Balaenoptera physalus* (Linnaeus 1758): analysis of mitochondrial and nuclear loci. Molecular Ecology 7:585–599.

Bérubé, M., J. Urban, A. E. Dizon, R. L. Brownell, and P. J. Palsbøll. 2002. Genetic identification of a small and highly isolated population of fin whales (*Balaenoptera physalus*) in the Sea of Cortez, Mexico. Conservation Genetics 3:183–190.

Brownell, R. L., Jr. 1981. Review of coastal whaling by the Republic of Korea. Report of the International Whaling Commission 31:395–402.

Burmeister, H. 1865. Letter on a new species of whale *Balaenoptera patachonica* from Argentina. Proceedings of the Zoological Society, London 13:190–195.

Clapham, P. J., S. Leatherwood, I. Szczepaniak, and R. L. Brownell, Jr. 1997. Catches of humpback and other whales from shore stations at Moss Landing and Trinidad, California 1919-1926. Marine Mammal Science 13:368–394.

Clarke, R. 2004. Pygmy fin whales. Marine Mammal Science 20:329–334.

Committee on Taxonomy. 2018. List of marine mammal species and subspecies. Society for Marine Mammalogy. www.marinemammalscience.org. Accessed 5 June 2019.

Conway, C. A. 2005. Global population structure of blue whales, *Balaenoptera musculus* ssp., based on nuclear genetic variation. Ph.D. dissertation, University of California, Davis.

Delarue, J., S. K. Todd, S. M. Van Parijs, and L. Di Iorio. 2009. Geographic variation in Northwest Atlantic fin whale (*Balaenoptera physalus*) song: implications for stock structure assessment. The Journal of the Acoustical Society of America 125:1774–1782.

Desmoulins, A. 1822. Baleine. Balaena. Mam. Dictionnaire Classique d'Histoire Naturelle 2:155–165.

Dyke, A. S., J. Hooper, and J. M. Savelle. 1996. A history of sea ice in the Canadian Arctic Archipelago based on postglacial remains of the bowhead whale (*Balaena mysticetus*). Arctic 49:235–255.

Edwards, E. F., C. Hall, T. J. Moore, C. Sheredy, and J. V. Redfern. 2015. Global distribution of fin whales *Balaenoptera physalus* in the post-whaling era (1980-2012). Mammal Review 45:197–214.

Ensor, P., H. Komiya, I. Beasley, K. Fukutome, P. Olson, and Y. Tsuda. 2007. Report of the 2006-2007 International Whaling Commission-Southern Ocean Whale and Ecosystem Research (IWC-SOWER) Cruise. International Whaling Commission, Cambridge, United Kingdom.

Ensor, P., K. Minami, L. Morse, P. Olson, and K. Sekiguchi. 2008. Report of the 2007-2008 International Whaling Commission-Southern Ocean Whale and Ecosystem Research (IWC-SOWER) Cruise. International Whaling Commission, Cambridge, United Kingdom.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611–2620.

Fischer, J. B. 1829. Synopsis mammalium. J. G. Cottae, Stuttgardtriae, Germany.

Fujino, K. 1960. Immunogenetic and marking approaches to identifying sub-populations of the North Pacific whales. Scientific Reports of the Whales Research Institute [Tokyo] 15:84–142.

Geijer, C. K. A., G. Notarbartolo Di Sciara, and S. Panigada. 2016. Mysticete migration revisited: are Mediterranean fin whales an anomaly? Mammal Review 46:284–296.

Goodwin, S., J. D. McPherson, and W. R. McCombie. 2016. Coming of age: ten years of next-generation sequencing technologies. Nature Reviews: Genetics 17:333–351.

Gray, J. E. 1864. Notes on the whalebone-whales; with a synopsis of the species. The Annals and Magazine of Natural History 3:345–353.

Gray, J. E. 1865. Short account of part of a skeleton of a finner whale, sent by Mr. Swinhoe from the coast of Formosa. Proceedings of the Scientific Meetings of the Zoological Society of London 1:725–728.

Gregg, C. C. 1956. Report of the Directory to the Board of Trustees for the year 1955. Chicago Natural History Museum, Chicago.

Grobaty, T. 2012. Long beach chronicles. from pioneers to the 1933 Earthquake. The History Press, Charleston, South Carolina.

Hancock-Hanser, B. L., A. Frey, M. S. Leslie, P. H. Dutton, F. I. Archer, and P. A. Morin. 2013. Targeted multiplex next-generation sequencing: advances in techniques of mitochondrial and nuclear DNA sequencing for population genomics. Molecular Ecology Resources 13:254–268.

Hatch, L. T. 2004. Male genes and male song: integrating genetic and acoustic data in defining fin whale, *Balaenoptera physalus*, management units. Ph.D. dissertation, Cornell University, Ithaca, New York.

Hausdorf, B. 2011. Progress toward a general species concept. Evolution 65:923–931.

Häuser, C. L. 1987. The debate about the biological species concept - a review. Journal of Zoological Systematics and Evolutionary Research 25:241–257.

Ichihara, T. 1957. An application of linear discriminant function to external measurements of fin whale. Scientific Reports of the Whales Research Institute [Tokyo] 12:127–189.

Jakobsson, M., and N. A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806.

Jackson, J. A., et al. 2014. Global diversity and oceanic divergence of humpback whales (*Megaptera novaeangliae*). Proceedings of the Royal Society of London, B. Biological Sciences 281:20133222.

Jayne, H. 1898. Annual Report of the Provost to the Board of Trustees from September 1st, 1897, to September 1st, 1898. University of Pennsylvania.

Jayne, H. 1899. Annual Report of the Provost to the Board of Trustees from September 1st, 1898, to September 1st, 1899. University of Pennsylvania.

Jiménez López, M. E., D. M. Palacios, A. Jaramillo Legorreta, J. Urbán R., and B. R. Mate. 2018. Fin whale movements in the Gulf of California, Mexico, from satellite telemetry. PLoS One 14:e0209324.

Jombart, T., and I. Ahmed. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070–3071.

Kalinowski, S. T. 2011. The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. Heredity 106:625–632.

Katoh, K., and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30:772–780.

Kellogg, R. 1929. What is known of the migration of some of the whalebone whales. Smithsonian Institution. Annual Report of the Board of Regents 1928:467–494.

Kim, J. H., H. W. Kim, E.-M. Kim, and H. Sohn. 2018. Occurrence of the Omura's whale (*Balaenoptera omurai*) in Korean waters. Pp. 9 in Paper SC/67B/NH2 presented to the International Whaling Commission Scientific Committee, May 2018 (unpublished).

Lacépède, B. G. 1804. Les cétacés, les quadrupédes ovipares, les serpents et les poissons. Furne, Jouvet et cie, Paris, France.

Leatherwood, S. and R. R. Reeves. 1983. The sierra club handbook of whales and dolphins. Sierra Club Books, San Francisco, California.

Leduc, R. G., et al. 2017. Genetic variation in blue whales in the eastern Pacific: implication for taxonomy and use of common wintering grounds. Molecular Ecology 26:740–751.

Leduc, R. G., et al. 2007. Patterns of genetic variation in Southern Hemisphere blue whales and the use of assignment test to detect mixing on the feeding grounds. Journal of Cetacean Research and Management 9:73–80.

Lee, M. S. 2003. Species concepts and species reality: salvaging a Linnaean rank. Journal of Evolutionary Biology 16:179–188.

Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. R News 2:18–22.

Linnaeus, C. 1758. Systema naturae per regna tria naturae:secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. 10th ed. Laurentius Salvius, Stockholm, Sweden.

Lockyer, C., and T. Waters. 1986. Weights and anatomical measurements of northeastern Atlantic fin (*Balaenoptera physalus*, Linnaeus) and sei (*B. borealis*, Lesson) whales. Marine Mammal Science 2:169–195.

Lönnberg, E. 1906. Contributions to the Fauna of South Georgia. Kungliga Svenska Vetenskapsakademiens Handlingar 40:1–104.

Lönnberg, E. 1931. The skeleton of *Balaenoptera brydei* Ö. Olsen. Arkiv För Zoologi 23:1–23.

Mackintosh, N. A. 1942. The southern stocks of whalebone whales. Discovery Reports 22:197–300.

Mackintosh, N. A. 1965. The stocks of whales. Fishing News, London, United Kingdom.

Martens, F. 1675. Spitzbergische oder Groenlandische Reise Beschreibung gethan im Jahr 1671. Gottfried Schultzens Kosten, Hamburg.

Martien, K. K., et al. 2017. Analytical approaches to subspecies delimitation with genetic data. Marine Mammal Science 33:27–55.

Mayr, E., and P. K. Ashlock. 1991. Principles of systematic zoology. McGraw-Hill, Inc., New York.

Mcdonald, M. A., S. L. Mesnick, and J. A. Hildebrand. 2006. Biogeographic characterisation of blue whale song worldwide: using song to identify populations. Journal of Cetacean Research and Management 81:55–65.

Miller, S. A., D. D. Dykes, and H. F. Polesky. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Research 16:1215.

Mizroch, S. A., D. W. Rice, and J. M. Breiwick. 1984. The fin whale, *Balaenoptera physalus*. Marine Fisheries Review 46:20–24.

Mizroch, S. A., D. W. Rice, D. Zwiefelhofer, J. Waite, and W. L. Perryman. 2009. Distribution and movements of fin whales in the North Pacific Ocean. Mammal Review 39:193–227.

Montgelard, C., F. M. Catzeflis, E. Douzery. 1997. Phylogenetic relationships of artiodactyls and cetaceans as deduced from the comparison of cytochrome B and 12S rRNA mitochondrial sequences. Molecular Biology and Evolution 14:550–559.

Morin, P. A., et al. 2006. Genetic analysis of killer whale (*Orcinus orca*) historical bone and tooth samples to identify western US ecotypes. Marine Mammal Science 22:897–909.

Morin, P. A., and M. Mccarthy. 2007. Highly accurate SNP genotyping from historical and low-quality samples. Molecular Ecology Notes 7:937–946.

Nei, M., F. Tajima, and Y. Tateno. 1983. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. Journal of Molecular Evolution 19:153–170.

Nichol, L., and K. Heise. 1992. The historical occurrence of large whales off the Queen Charlotte Islands. Research Paper Report prepared for South Moresby/Gwaii Haanas National Parks Reserve, Canadian Parks Service, Queen Charlotte City, British Columbia.

Pastene, L. A., et al. 2007. Radiation and speciation of pelagic organisms during periods of global warming: the case of the common minke whale, *Balaenoptera acutorostrata*. Molecular Ecology 16:1481–1495.

Patten, M. A. 2010. Null expectations in subspecies diagnosis. Ornithological Monographs 67:35–41.

Patten, M. A., and P. Unitt. 2002. Diagnosability versus mean differences of sage sparrow subspecies. Auk 119:26–35.

Payne, R., and L. N. Guinee. 1983. Humpback whale (*Megaptera novaeangliae*) songs as an indicator of "stocks". Pp. 333–358 in Communications and behavior of whales (R. Payne, ed.). Westview Press, Boulder, Colorado.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

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 16:608–627.

Pyenson, N. D., J. A. Goldbogen, and R. S. Shadwick. 2013. Mandible allometry in extant and fossil Balaenopteridae (Cetacea: Mammalia): the largest vertebrate skeletal element and its role in rorqual lunge feeding. Biological Journal of the Linnean Society 108:586–599.

de Queiroz, K. 2007. Species concepts and species delimitation. Systematic Biology 56:879–886.

R Core Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Heredity 86:248–249.

Rice, D. W. 1998. Marine mammals of the world: systematics and distribution. Vol. 4. Society for Marine Mammalogy, Lawrence, Kansas.

Rosel, P. E., et al. 2017a. Examining metrics and magnitudes of molecular genetic differentiation used to delimit cetacean subspecies based on mitochondrial DNA control region sequences. Marine Mammal Science 33:76–100.

Rosel, P. E., et al. 2017b. A review of molecular genetic markers and analytical approaches that have been used for delimiting

marine mammal subspecies and species. Marine Mammal Science 33:56–75.

Rosel, P. E., and L. A. Wilcox. 2014. Genetic evidence reveals a unique lineage of Bryde's whales in the northern Gulf of Mexico. Endangered Species Research 25:19–34.

Rosenbaum, H. C., et al. 2000. World-wide genetic differentiation of *Eubalaena*: questioning the number of right whale species. Molecular Ecology 9:1793–1802.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Scammon, C. M. 1869. On the cetaceans of the western coast of north America. Proceedings of the Academy of Natural Sciences of Philadelphia 21:13–63.

Schlegel, H. 1862. De Dieren van Nederland Zoogdieren. A.C. Kruseman, Haarlem, The Netherlands.

Simon, M., K. M. Stafford, K. Beedholm, C. M. Lee, and P. T. Madsen. 2010. Singing behavior of fin whales in the Davis Strait with implications for mating, migration and foraging. The Journal of the Acoustical Society of America 128:3200–3210.

Širović, A., J. A. Hildebrand, S. M. Wiggns, M. A. Mcdonald, S. E. Moore, and D. Thiele. 2004. Seasonality of blue and fin whale calls and the influence of sea ice in the Western Antarctic Peninsula. Deep Sea Research (Part II, Topical Studies in Oceanography) 51:2327–2344.

Širović, A., J. A. Hildebrand, S. M. Wiggns, and D. Thiele. 2009. Blue and fin whale acoustic presence around Antarctica during 2003 and 2004. Marine Mammal Science 25:125–136.

Širović, A., L. N. Williams, S. M. Kerosky, S. M. Wiggins, and J. A. Hildebrand. 2013. Temporal separation of two fin whale call types across the eastern North Pacific. Marine Biology 160:47–57.

Sites, J. W., and J. C. Marshall. 2004. Operational criteria for delimiting species. Annual Review of Ecology Evolution and Systematics 35:199–227.

Sremba, A. L., A. R. Martin, and C. S. Baker. 2015. Species identification and likely catch time period of whale bones from South Georgia. Marine Mammal Science 31:122–132.

Taylor, B. L., et al. 2017a. Guidelines and quantitative standards to improve consistency in cetacean subspecies and species delimitation relying on molecular genetic data. Marine Mammal Science 33:132–155.

Taylor, B. L., et al. 2017b. Why we should develop guidelines and quantitative standards for using genetic data to delimit subspecies for data-poor organisms like cetaceans. Marine Mammal Science 33:12–26.

Tomilin, A. G. 1946. Thermoregulation and the geographical races of cetaceans. Comptes Rendus (Doklady) de l'Académie des Sciences d l'URSS 54:465–468.

Tomilin, A. G. 1957. Kitoobraznye [Cetaceans]. Volume 9 in Zveri SSSR Iprilezhashchickh stran. Zveri vostochnoi Evropy I severnoi Azii [Mammals of U.S.S.R. and adjacent countries. Mammals of eastern Europe and northern Asia] (V. G. Heptner, ed.). Doklady Akademii Nauk, SSSR, Moscow, 756 pp. Translation (by O. Ronen):1967. Mammals of the U.S.S.R. and adjacent countries. Mammals of eastern Europe and adjacent countries [sic]. Israel Program for Scientific Translations, Jerusalem. 717 pp.

True, F. W. 1904. The whalebone whales of the western north Atlantic compared with those occurring in European waters; with some observations on the species of the north Pacific. Smithsonian Contributions to Knowledge 33:332.

Wang, J. 2017. The computer program STRUCTURE for assigning individuals to populations: easy to use but easier to misuse. Molecular Ecology Resources 17:981–990.

Watkins, W. A. 1981. Activities and underwater sounds of fin whales. Scientific Reports of the Whales Research Institute 33:83–117.

Watkins, W. A., P. Tyack, K. E. Moore, and J. E. Bird. 1987. The 20-Hz signals of finback whales (*Balaenoptera physalus*). The Journal of the Acoustical Society of America 82:1901–1912.

Zink, R. M., and J. I. Davis. 1999. New perspectives on the nature of species. Pp. 1505–1518 in Proceedings 22 International Ornithological Congress, Durban (N. J. Adams and R. H. Slotow, eds.). BirdLife South Africa, Johannesburg, South Africa.

Submitted 20 December 2018. Accepted 13 June 2019.

Associate Editor was Ricardo Moratelli.