Spotlight Article

Syst. Biol. 73(6):873–885, 2024
© The Author(s) 2024. Published by Oxford University Press on behalf of the Society of Systematic Biologists.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.
https://doi.org/10.1093/sysbio/syae049
Advance Access Publication August 19, 2024

Phylogenomics and Pervasive Genome-Wide Phylogenetic Discordance Among Fin Whales (*Balaenoptera physalus*)

Fabricio Furni[1,*], Eduardo R. Secchi[2], Camilla Speller[3], Daniel DenDanto[4], Christian Ramp[5,6], Finn Larsen[7], Sally Mizroch[8], Jooke Robbins[9], Richard Sears[5], Jorge Urbán R.[10], Martine Bérubé[1,9], and Per J. Palsbøll[1,9,*]

[1]*Marine Evolution and Conservation Group, Groningen Institute of Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands*
[2]*Laboratório de Ecologia e Conservação da Megafauna Marinha, Instituto de Oceanografia, Universidade Federal do Rio Grande-FURG, Rio Grande, Brasil*
[3]*Department of Anthropology, University of British Columbia, Vancouver, Canada*
[4]*Allied Whale, College of the Atlantic, Bar Harbor, ME, USA*
[5]*Mingan Island Cetacean Study Inc., St. Lambert, Quebec, Canada*
[6]*Scottish Oceans Institute, University of St. Andrews, St. Andrews, UK*
[7]*National Institute of Aquatic Resources, Kongens Lyngby, Denmark*
[8]*National Marine Mammal Laboratory, US National Marine Fisheries Service, Seattle, WA, USA*
[9]*Center for Coastal Studies, Provincetown, MA, USA*
[10]*Departamento de Ciencias Marinas y Costeras, Universidad Autónoma de Baja California Sur, La Paz, Baja California Sur, México*
Martine Bérubé and Per J. Palsbøll shared senior authorship.
[*]*Correspondence to be sent to: Marine Evolution and Conservation Group, Groningen Institute of Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands. E-mail: fabriciofurni@palaeome.org, palsboll@gmail.com.*

Received 27 June 2023; reviews returned 8 August 2024; accepted 15 August 2024
Associate Editor: Megan Smith

Abstract.—Phylogenomics has the power to uncover complex phylogenetic scenarios across the genome. In most cases, no single topology is reflected across the entire genome as the phylogenetic signal differs among genomic regions due to processes, such as introgression and incomplete lineage sorting. Baleen whales are among the largest vertebrates on Earth with a high dispersal potential in a relatively unrestricted habitat, the oceans. The fin whale (*Balaenoptera physalus*) is one of the most enigmatic baleen whale species, currently divided into four subspecies. It has been a matter of debate whether phylogeographic patterns explain taxonomic variation in fin whales. Here we present a chromosome-level whole genome analysis of the phylogenetic relationships among fin whales from multiple ocean basins. First, we estimated concatenated and consensus phylogenies for both the mitochondrial and nuclear genomes. The consensus phylogenies based upon the autosomal genome uncovered monophyletic clades associated with each ocean basin, aligning with the current understanding of subspecies division. Nevertheless, discordances were detected in the phylogenies based on the Y chromosome, mitochondrial genome, autosomal genome and X chromosome. Furthermore, we detected signs of introgression and pervasive phylogenetic discordance across the autosomal genome. This complex phylogenetic scenario could be explained by a puzzle of introgressive events, not yet documented in fin whales. Similarly, incomplete lineage sorting and low phylogenetic signal could lead to such phylogenetic discordances. Our study reinforces the pitfalls of relying on concatenated or single locus phylogenies to determine taxonomic relationships below the species level by illustrating the underlying nuances that some phylogenetic approaches may fail to capture. We emphasize the significance of accurate taxonomic delineation in fin whales by exploring crucial information revealed through genome-wide assessments. [Discordance; fin whale; incomplete lineage sorting; introgression; subspecies; whole genomes.]

Whole genome sequences should, in principle, enable the most robust inference of the evolutionary history among taxa and thus the elucidation of taxonomic relationships (Eisen and Fraser 2003; Delsuc et al. 2005; Misof et al. 2014; McGowen et al. 2020). However, several phylogenomic studies have revealed complex patterns underlying such phylogenies, that is, different genomic regions support divergent topologies, which in turn is a product of the species' history (Pamilo and Nei 1988; Maddison 1997; Jeffroy et al. 2006). Such phylogenetic discordances across the genome have emerged in a growing number of studies aimed at inferring taxonomic relationships, especially among closely related taxa (e.g., Chen et al. 2019; Meleshko et al. 2021; de Jong et al. 2023; Rivas-González et al. 2023; Sørensen et al. 2023). A multitude of different processes can result in phylogenetic discordance, such as introgression and incomplete lineage sorting (Pamilo and Nei 1988; Maddison 1997; Jeffroy et al. 2006).

Resolving the phylogenetic relationships among conspecifics, for example, subspecies, is especially challenging, and molecular-based phylogenies are often the only means available in difficult-to-study species lacking obvious geographic, reproductive, or phenotypic barriers. Conspecific populations are also subjected to processes that may disrupt phylogeographic

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025

patterns, such as introgression and retention of ancestral polymorphisms (Pamilo and Nei 1988; Arnold 1993; Kutschera et al. 2014; Ge et al. 2022; de Jong et al. 2023; Rivas-González et al. 2023). Sex-specific and asymmetric gene flow may further enhance phylogeographic discordance resulting in divergent topologies among uni-parentally inherited loci (e.g., mitochondrial and Y chromosome) and autosomal loci (Lyrholm et al. 1999; Petit and Excoffier 2009; Sørensen et al. 2023).

Baleen whales (*Mysticeti*) are gigantic marine mammals undertaking extensive seasonal migrations between high-latitude summer feeding and tropical winter breeding areas (Mackintosh 1946; Lockyer and Brown 1981; Mizroch et al. 1984). Most mysticete species have a cosmopolitan distribution, which is additional evidence of their extensive dispersal ability. Mysticetes' extensive capacity for dispersal, in the comparatively unrestricted pelagic ocean, likely facilitated introgression, even among ocean basins (Palumbi and Baker 1994; Bérubé et al. 1998; Jackson et al. 2014; Stevick et al. 2014; Alter et al. 2015). Despite observations of inter-ocean basin gene flow, some mysticete species have been divided into multiple subspecies. Traditionally, the Northern and Southern Hemisphere mysticete populations were assigned to separate subspecies (Lönnberg 1906; Ichihara 1966; Clarke 2004). In some species, additional "pygmy" subspecies have been identified as well, mostly in the Southern Hemisphere (e.g., the blue whale, *Balaenoptera musculus* spp., Ichihara 1966).

The fin whale (*B. physalus*) is the second largest mysticete and was subjected to intense whaling during the 20th century across the globe (Rocha, Jr. et al. 2015). Based on morphological differences in the vertebrae, Lönnberg (1906) proposed that fin whales in the Southern Hemisphere be assigned to their own subspecies, *B. p. quoyi*. Consequently, fin whales in the Northern Hemisphere were assigned to another subspecies, *B. p. physalus*. Later, a third, smaller form observed in the Southern Hemisphere was nominated as a subspecies, *B. p. patachonica* (Clarke 2004). More recently, Archer et al. (2013, 2019) proposed a taxonomic revision of *B. physalus* spp. based on the phylogeographic pattern observed in a phylogeny estimated from complete mitochondrial genomes and 23 autosomal single nucleotide polymorphisms (SNPs). Based on fixed differences in the mitochondrial DNA, Archer et al. (2019) proposed an additional, new subspecies confined to the North Pacific, *B. p. velifera*, constraining *B. physalus physalus* to the North Atlantic. These intraspecific taxonomic divisions have been the subject of debate, centering on whether subspecies should be recognized based on their phylogeographic pattern (Archer et al. 2019; Cabrera et al. 2019; Pérez-Alvarez et al. 2021; Buss et al. 2023). Cabrera et al. (2019) pointed to the well-known issues in defining subspecies based solely on the phylogenetic topology inferred from the uni-parentally inherited mitochondrial genome. The authors additionally pointed to the effect of low sample sizes which may erroneously result in monophyly, seemingly supporting a distinct phylogeographic division. Such sensitivity of the topology in phylogenies inferred from the mitochondrial genome clearly illustrates the challenges in defining intraspecific phylogenetic relationships in a species with a high dispersal capacity in an environment devoid of physical barriers. The above aspects highlight the need for a robust, genome-wide phylogenetic assessment of fin whales to enhance research into their evolution and systematics, thereby aiding effective conservation of this charismatic species.

To this end, we employed whole nuclear and mitochondrial genome sequences to conduct a phylogenetic assessment of fin whales from different ocean basins. Here we aimed to fill existing gaps and robustly estimate the species' phylogenetic relationships to a chromosome whole-genome level. We subsequently evaluated the concordance among phylogenies inferred from the autosomal chromosomes, sex chromosomes, and the mitochondrial genome. The consensus autosomal genome phylogenetic results revealed strong support for a phylogeographic structure that aligns with the most recently proposed subspecies (see Archer et al. 2019). Furthermore, we found discordances between the phylogeny inferred from the Y chromosome, the mitochondrial genome, and the autosomal genome, along with signs of introgression and pervasive phylogenetic discordance across the autosomal genome. These discordances may be not only due to high levels of incomplete lineage sorting, or low phylogenetic signal but also due to a puzzle of introgression events. Our findings suggest a complex phylogenetic scenario in fin whale genomes and the potential pitfalls of inferring intraspecific phylogenies from a single locus, such as the mitochondrial genome, or a few nuclear SNPs. We emphasize the importance of correct taxonomic delineation in fin whales, which in turn could affect management decisions.

MATERIALS AND METHODS

Sampling and Whole Genome Resequencing

Tissue samples were collected from 3 major ocean basins: the North Atlantic (*B. physalus physalus*), the North Pacific (*B. physalus velifera)*, and the Southern Ocean (*B. p. pathaconica* or *B. p. quoyi)*. In the North Atlantic, tissue samples were collected along the North American eastern sea border ($N = 8$, Gulf of Maine and the Gulf of Saint Lawrence), and Iceland ($N = 1$). In the North Pacific, samples were collected off Kodiak Island, along the eastern sea border of the United States ($N = 6$), and in the Gulf of California, Mexico ($N = 6$). In the Southern Ocean, samples were collected off South Georgia and the eastern Antarctic Peninsula ($N = 7$). A sample of the most closely related species, the humpback whale (*Megaptera novaeangliae*), served as an outgroup. Tissue samples from free-ranging animals were collected as skin biopsies using a crossbow

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025

as described by Palsbøll (1991). Tissue samples were stored in 6M saturated NaCl with 25% DMSO and stored at −80 °C (Amos and Hoelzel 1991). Genomic DNA was extracted using Qiagen DNAEasy™ Blood and Tissue Kit columns (QIAGEN Inc., USA) following the manufacturer's protocol and resuspended in 1XTE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0). DNA concentrations were estimated by fluorometric quantification using a Qubit® 2.0 Fluorometer (Life Technologies™). The integrity of the DNA was assessed by electrophoresis through a 0.7% agarose gel at 175 volts for ~30 min and subsequently visualized by staining with ethidium bromide. Library construction and whole-genome resequencing were conducted using the BGI-SEQ 500 (MGI, China) 100 base pair (bp) paired-reads manufacturer's protocols outsourced to the Beijing Genomics Institute, Europe. In addition to the samples sequenced during this study, published raw reads were retrieved from the National Center of Biotechnology Information (NCBI) from one North Pacific fin whale (NCBI accession code: SRR935201; Yim et al. 2014) and eight Icelandic fin whales (NCBI accession codes: SRR14986187, SRR15013014, SRR15042059, SRR15042061, SRR15042062, SRR15042063, SRR15048183, SRR15082441; Wolf et al. 2022).

Variant Calling

Raw FASTQ reads were mapped against a male blue whale (*Balaenoptera musculus*) genome assembly (NCBI accession code: mBalMus1.pri.v2; Bukhman et al. 2024) using the BWA *mem* (v. 0.7, Li and Durbin 2009) with default parameter settings. Duplicate reads were removed using the *MarkDuplicates* function in PICARD (v. 2.25, http://broadinstitute.github.io/picard/, MIT Broad Institute 2021). The Genome Analysis Toolkit (GATK, v. 4.1, McKenna et al. 2010) was applied to call high-confidence variants from the aligned reads in the following manner: alignment files were recalibrated using the variant sites called with the highest confidence using GATK base recalibration workflow. Variants were first called using GATK HaplotypeCaller in reference confidence mode (-ERC) emitting intermediate genotype files (GVCF) per sample. GATK *GenotypeGVCF* was then used to convert GVCF into variant files. Hard filtering of single nucleotide polymorphisms (SNPs) and indels was conducted using GATK SelectVariants and VariantFiltration (filtering expression for SNPs: "QD < 2.0 || FS > 60.0 || MQ < 40.0 || *MQRankSum* < -12.5 || ReadPosRankSum < -8.0 || SOR > 4.0" / Indels: 'QD < 2.0 || FS > 200.0 || ReadPosRankSum < -20.0 || SOR > 10.0'). After mapping the highest-confidence variants, read alignment files were recalibrated using GATK BaseRecalibrator and ApplyBQSR functions. After recalibration steps, variant calling was performed using GATK *HaplotypeCaller* in ERC mode, retaining all active positions including monomorphic sites (BP_RESOLUTION and MQ > 30). GATK *CombineGVCFs* and *GenotypeGVCF* were used to merge genotypes for all samples. All GATK steps were conducted per chromosome (-L option) and results were combined using the PICARD *GatherVcfs* function.

Mitochondrial Genome Assembly From Whole Genome Sequence Data

Whole mitochondrial genome fasta files were retrieved directly from the recalibrated read alignment files (BAM) using the ANGSD (v. 0.93, Korneliussen et al. 2014) *-doCounts* and *-dofasta* functions, selecting the most frequent allele and a minimum 3-fold read depth. In addition to the mitochondrial genome sequences generated during this study, a total of 161 sequences were retrieved from public repositories (Archer et al. 2013; Cabrera et al. 2019). MAFFT (v. 7.4, Katoh and Standley 2013) alignments were conducted using the default parameters (*--auto*) for 2 datasets: one alignment of all available whole mitochondrial genome data, and a second, containing only the mitochondrial genomes retrieved from the whole genome sequences analyzed in this study. Alignments were checked and trimmed at both ends using JALVIEW (v. 2.11, Waterhouse et al. 2009). IQ-TREE (v. 2.2, Minh et al. 2020) maximum likelihood (ML) phylogenetic analyses were performed on both alignments using the General Time Reversible (GTR) + GAMMA model (Abadi et al. 2019) with 20 initial topology searches and 1000 bootstraps. The humpback whale mitochondrial genome sequence generated during this study, as well as a published mitochondrial sequence, served as outgroups (NCBI accession number: NC006927l; Sasaki et al. 2005).

Concatenated Nuclear Autosomal Genome Phylogeny Estimation

The phylogeny estimated from the whole autosomal genome sequences was based on the 21 fin whale autosomal chromosomes. Low-quality sites were removed (*-MQ*, *-FORMAT/DP* < 3, *exclude-types* indel, *-e* F_MISSING > 0.2, *-m* 2 *-M* 2) using BCFTOOLS (v. 1.15, Danecek et al. 2011), and the remained sites (including monomorphic sites) converted into a PHYLIP-formatted alignment using *vcf2phylip* (v. 2.0, Ortiz 2019). The PHYLIP-based alignment served as input for estimating an ML phylogeny as implemented in IQ-TREE. The GTR + GAMMA model was applied in 20 initial topology searches, followed by 1000 (-B 1000) ultrafast bootstraps, using the humpback whale genome as the outgroup. Y and X chromosome concatenated topologies were obtained from the data restricted to these genome regions using the BCFTOOLS *view* tool. Chromosome-specific alignment generation and phylogenetic estimates were performed as described above for the autosomal genome.

Phylogenies Across the Autosomal Genome

Additional phylogenies were obtained across the nuclear genome for nonoverlapping windows with sizes of 1 M bps, 100k bps, and 50k bps. The windows-based analysis for the Y chromosome was restricted to

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025

windows of 50k bps and 100k bps due to the chromosome size. IQ-TREE ML phylogenies were estimated for each window as described above for concatenated genomes. The coalescent-based framework implemented in the ASTRAL weighted method (v. 1.15, Zhang and Mirarab 2022) was employed to estimate a consensus phylogeny for each set of genomic window sizes. Concordance between the genomic window phylogenies with the ASTRAL consensus, concatenated autosomal, and mitochondrial phylogenies was checked by calculating gene concordance factors (gCFs) using IQ-TREE (Ané et al. 2007). The topology weighting test was performed with topology weighting by iterative sampling of subtrees in TWISST (Martin and Van Belleghem 2017) based on the output phylogenies from the 100k bps window analysis and assigning groups according to their sampling origin.

Inter-Ocean Basin Introgression

The possible effects of introgression were investigated using Patterson's *D*, also known as *D*-statistics or the ABBA-BABA test (Green et al. 2010; Durand et al. 2011). First, global *D*-statistics estimates were conducted using the whole autosomal dataset. The input variant file was first filtered using BCFTOOLS to contain only variable sites (--type snps). *D*-statistics were estimated using D-SUITE (v. 0.4, Malinsky et al. 2021) *Dtrios* using the likely topology inferred using ASTRAL ((North Pacific, (North Atlantic, Southern Ocean)), Outgroup). Significance was assessed using 1000 Jackknife blocks (-k 1000) across the autosomal genome. *D*-statistics were subsequently estimated locally for each of the 100k bps genomic windows using *Dtrios* with default parameters and the most likely topology from the TWISST results as input tree ((North Pacific, (North Atlantic, Southern Ocean)), Outgroup). *D*-statistics significance values for the 100 k bps genomic windows were corrected for multiple tests using the Bonferroni correction method (Bonferroni 1936) implemented in R (v. 4.1, R Core Team 2020).

RESULTS

Resequencing and Alignment Metrics

The final data set comprised 38 whole genome sequences, of which 29 were generated during the present study and nine were retrieved from NCBI. Of 12.2 billion initial raw FASTQ reads, 10.53 billion passed duplicate and quality filters. The average read depth for the mitochondrial genomes was 909 (SD 436.5) and the mean genome coverage was 99.5% (SD 0.09). The final mitochondrial genome sequence alignment comprised 200 samples and 16,404 sites. The mean autosomal read depth was 13 (SD 1.69) and the mean genome coverage was 98.07% (SD 0.17) resulting in 2.25 billion (SD 0.025) sites in total. After filtering, the final whole autosomal genome alignment contained a total of 1.568 billion sites with 9,559,510 alignment patterns (Supplementary Tables S1, S2). For the Y and X chromosomes, the alignment yielded 1,797,065 sites with 46,518 distinctive patterns and 23,828,069 with 232,314 distinctive patterns, respectively.

Whole Mitochondrial and Concatenated Nuclear Genome Phylogenetic Relationships

We detected 23 novel haplotypes among the 198 mitochondrial genomes. The most likely phylogeny inferred from the mitochondrial genomes did not align fully with ocean basins, as shown in earlier work (Archer et al. 2013; Cabrera et al. 2019; Buss et al. 2023). Most individual samples were grouped within ocean-specific clades, but all ocean basins were polyphyletic. One clade, which contained most Southern Ocean samples, also included North Atlantic samples, and another clade included North Pacific and a Southern Ocean sample (Fig. 1). The phylogeny based only on haplotypes from the samples with whole genome data yielded a similar topology. Polyphyly was also evident in this reduced representation phylogeny, where samples from the North Pacific and the North Atlantic clustered together with samples from the Southern Ocean (Fig. 2a).

The phylogeny estimated from the concatenated Y chromosome also contained polyphyletic clades, in which individuals sampled in the Southern Ocean clustered together within samples from the North Atlantic, although sampling was somewhat reduced ($N = 15$) as this assessment was based solely on males (Supplementary Fig. S7). In contrast, the concatenated phylogeny estimated from the autosomal genome clustered samples in well-supported monophyletic clades corresponding to their geographic origin, that is, ocean basins. Similar to the whole mitochondrial-based phylogeny, the Southern Ocean and North Pacific clades were placed as sister groups (Fig. 2b). Within some ocean basins, the additional phylogeographic structure was visible, for example, samples from the Gulf of California all clustered together in a distinct, well-supported sub-clade within the main North Pacific clade, differentiating these samples from the eastern North Pacific samples. The best-fitting phylogeny estimated from the concatenated X chromosome genomes corresponded to that based on the autosomal genome with a clear phylogeographic structure corresponding to the oceanic origin of samples (Supplementary Fig. S5).

Consensus and Phylogenetic Signals Throughout the Nuclear Genome

As a concatenated phylogeny may not capture discrepancies among different loci (Pamilo and Nei 1988; Maddison 1997), we estimated windows-based phylogenies across the nuclear genome. Regions were defined as non-overlapping, fixed windows of 1M, 100k, and 50k bps windows across the 21 autosomal chromosomes for which we estimated a total of 2253, 22,405, and 43,629 maximum likelihood phylogenies, respectively. We first

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025

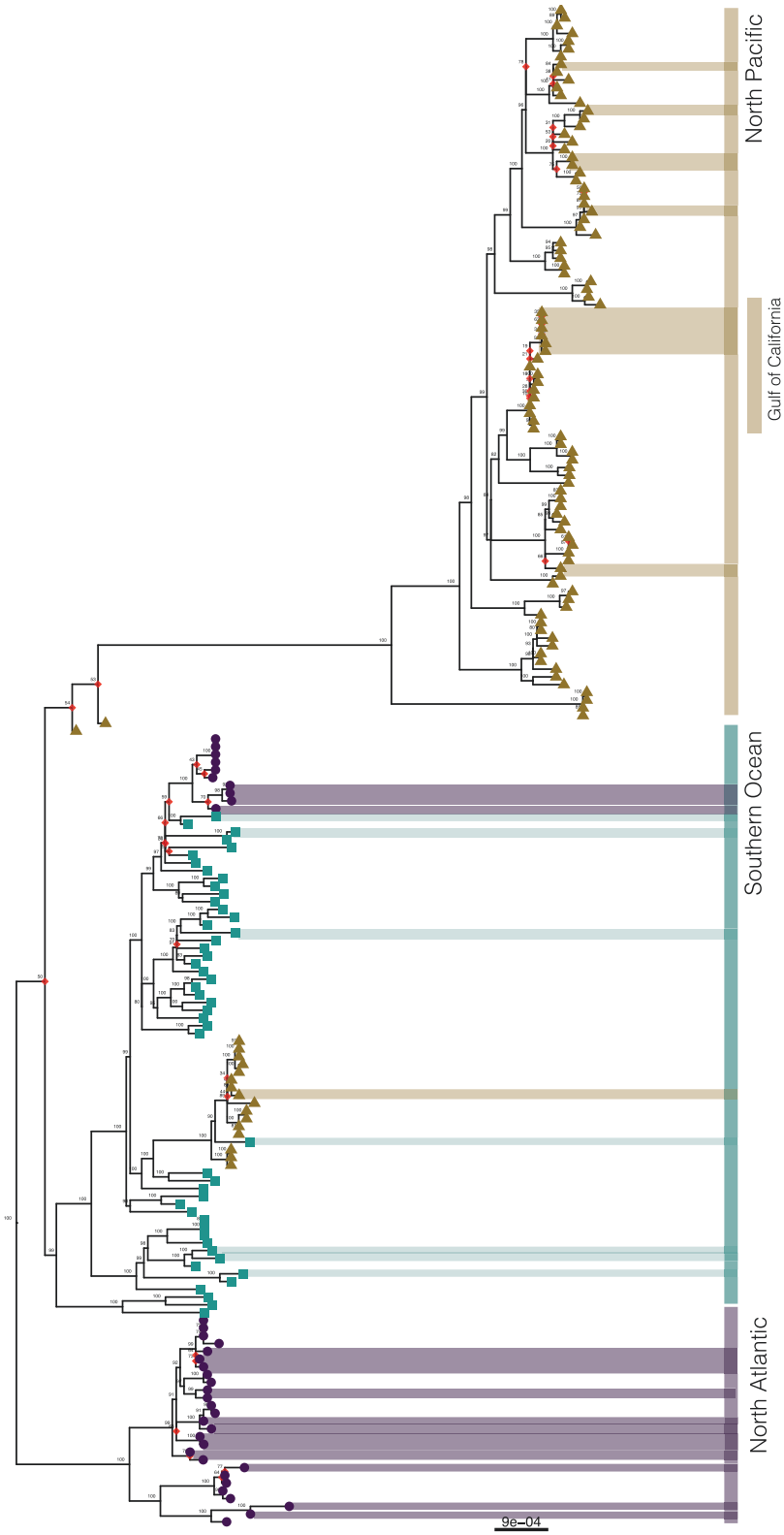

FIGURE 1. ML phylogeny inferred from previously published and new mitochondrial genomes. Horizontal bars indicate the location of the new mitochondrial genome sequences used in this study. Vertical bars indicate the main ocean basin clades, such as the North Pacific Ocean (top), the Southern Ocean (middle), or the North Atlantic Ocean (bottom). Tip shown as circles (North Atlantic), squares (Southern Ocean), and triangles (North Pacific and Gulf of California). Node-coloured squares indicate low bootstrap support (< 80). Bootstrap values are indicated on the right side of the nodes.

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025

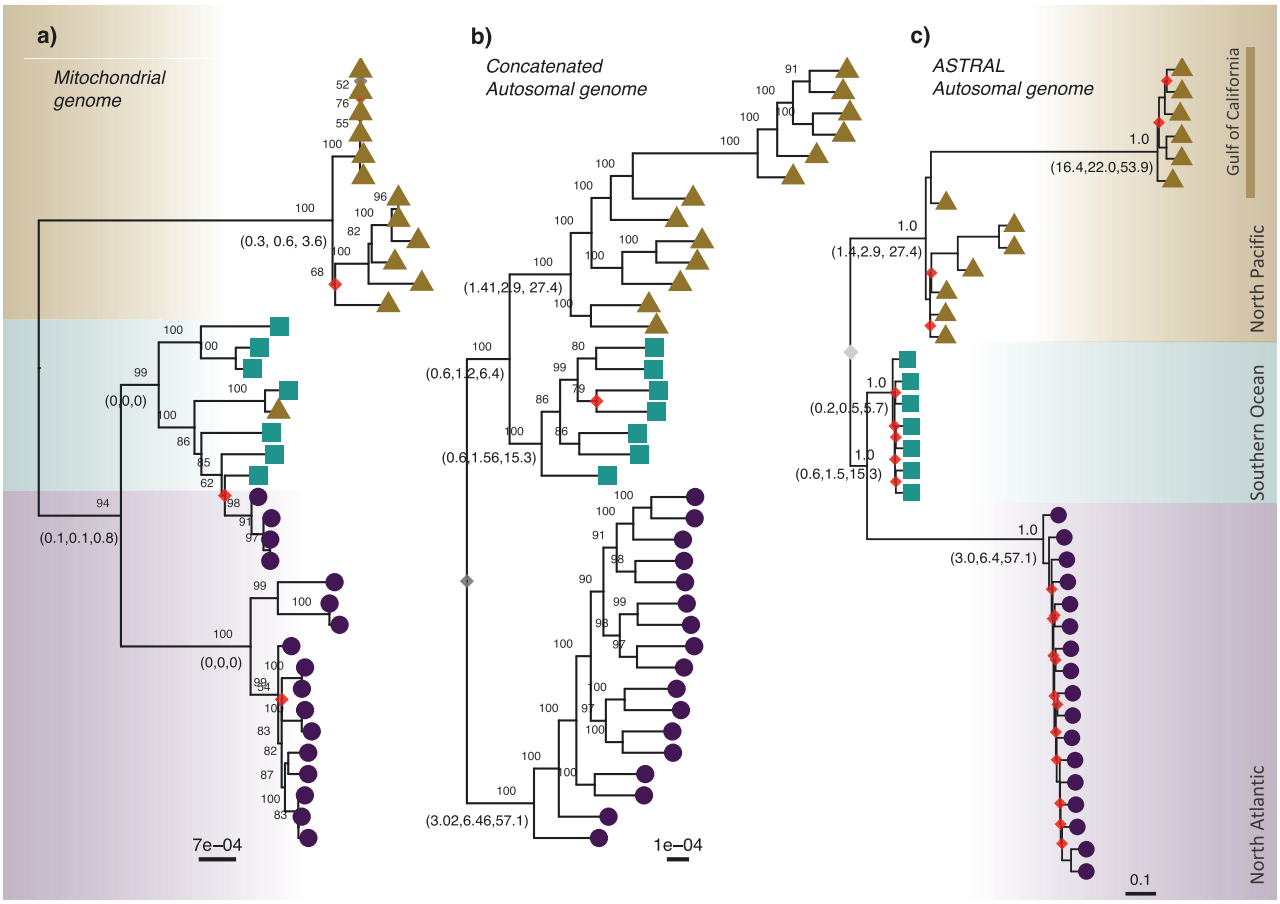

FIGURE 2. Maximum-likelihood (ML) phylogenies estimated from mitochondria and autosome genome. Tip shown as circles (North Atlantic), squares (Southern Ocean), and triangles (North Pacific and Gulf of California) Detailed phylogenies are available in Supplementary Fig. S3, S4. a) ML phylogeny estimated from mitochondrial genomes. b) ML phylogeny estimated from the concatenated autosomal genome. c) ASTRAL consensus phylogeny estimated across autosomal genome windows. Values under the nodes represent support estimated using IQ-TREE fast bootstrap (a and b) or ASTRAL scores (c). Values above the nodes in parenthesis correspond to gCF results for windows of 50k, 100k, and 1M bps, respectively. Squares highlight nodes with low support (i.e., < 80/0.8 IQTREE/ASTRAL).

estimated a heterogeneity-sensitive consensus phylogeny from all the above window-based phylogenies using ASTRAL (Zhang and Mirarab 2022). Irrespective of the window size, all ASTRAL consensus phylogenies converged onto a similar topology (Supplementary Fig. S4) in which clades were monophyletic with respect to the oceanic origin of the samples (Fig. 2c). The clade containing the Southern Ocean samples appeared as a sister clade to the clade with the North Atlantic samples in contrast to the concatenated phylogenies based on the autosomal and the X chromosome. We also performed a similar ASTRAL consensus analysis based on windows across the X and Y chromosomes. The X chromosome ASTRAL consensus phylogenies were also congruent with the ocean basin origin of samples, that is, with an unambiguous phylogeographic structure (Supplementary Fig. S6). However, the ASTRAL consensus phylogenies estimated from the Y chromosome were polyphyletic with respect to samples from the North Atlantic and Southern Ocean (Supplementary Fig. S7).

Topology Weighting and Concordance to Phylogeographic Structure

Due to the incongruence in the deeper relationships among ocean basins in the phylogenies, we performed a topology weighting test using the 100k bps window-based phylogenies across the autosomal chromosomes. The topology with the highest weight value (37%) agreed with the ASTRAL consensus phylogeny inferred from the autosomal chromosomes, in which the North Atlantic and Southern Ocean were sister clades. However, weight values were only slightly lower (31%) for the topology with the Southern Ocean and North Pacific as sister clades, followed by the lowest weighting score (29%) for the topology in which the Southern Ocean and North Pacific were sister clades (Fig. 3a, b).

We assessed concordance, that is, gCF, of the windows-based phylogenies with the ASTRAL consensus and concatenated phylogenies inferred from the entire autosomal genome and sex chromosome. A large fraction of the autosomal window-based phylogenies was

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025

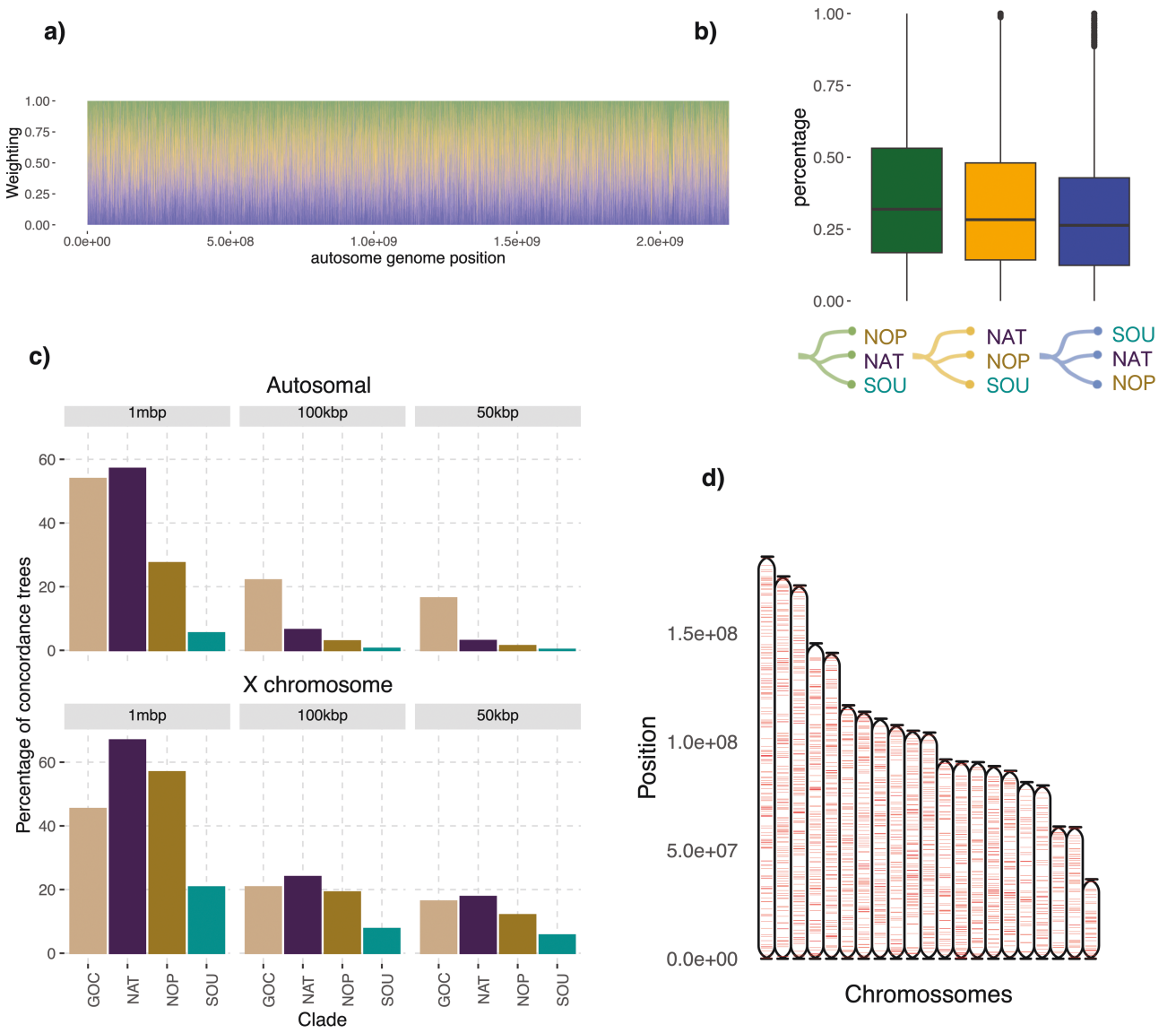

FIGURE 3. a) A genome-wide view of the TWISST topology weights results for three different topologies: top; (SOU,(NAT,NOP)), middle; (NOP,(NAT,SOU)), and bottom; (NAT,(NOP,SOU)). b) TWISST topology weighting results in percent across the 100K bps autosomal window size phylogenies. c) Percentage (based on the gCF results) of window phylogenies supporting nodes with monophyletic clades in terms of ocean basin origin of the samples. Top panel: autosomal genome; Bottom: X chromosome. d) Genomic regions (horizontal bars) showing significant *D*-values across the autosomal genome. GOC: Gulf of California, NAT: North Atlantic, NOP: North Pacific, SOU: Southern Ocean.

discordant with the unambiguous phylogeographic structure observed in the consensus and concatenated phylogenies. The level of windows-based concordance with a phylogeographic structure was positively correlated with window size (Fig. 3c). For instance, 57.1 %, 27.4 %, and only 5.7 % of the phylogenies based on a 1M bps window supported monophyly for all North Atlantic, North Pacific, and the Southern Ocean samples, respectively. Decreasing window sizes resulted in a lower fraction of window-based phylogenies supporting monophyletic clades for the ocean basin samples. Across all three window sizes, the support was lowest for the Southern Ocean, whereas a monophyletic clade for the Gulf of California was relatively well-supported compared to other deeper nodes. The highest degree of concordance with the ASTRAL consensus topology was observed in the X chromosome across all window sizes (Fig. 3c).

Signs of Introgression

Introgression can result in phylogenetic discordances across the genome (Pamilo and Nei 1988; Maddison 1997). We inferred the degree of introgression from global and local estimates of *D*-statistics. Global *D* values indicate introgression between the North Pacific and the Southern Ocean (*D*-statistics = 0.03; *Z*-score = 16.9, *P*-value = $1e^{-16}$). Locally, the inferred *D* values were low and nonsignificant across a large part of the autosomal

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025

genome. Nevertheless, genomic windows with significant positive *D* values accounted for 6.9% of the autosomal genome (Fig. 3d).

Discussion

Here we conducted the first, in-depth, intra-specific phylogenetic assessment of a baleen whale based on whole genome sequences from multiple ocean basins. The concatenated and ASTRAL consensus phylogeny based on the autosomal genome sequences placed samples from each ocean basin in well-supported ocean basin-specific, monophyletic clades (Fig. 2b, c). However, the support across the autosomal genome for such ocean-specific monophyletic clades was much lower among the fin whale genomes from the Southern Ocean compared to the fin whale genomes from both Northern Hemisphere ocean basins (i.e., the North Pacific and the North Atlantic, Fig. 3c). One straightforward explanation for the substantially reduced phylogeographic structure in the Southern Ocean may partially be due to directional, inter-oceanic introgression, that is, sink-source dynamics, after a secondary contact (e.g., introgression from the Southern Hemisphere into the Northern Hemisphere). In addition to introgression, high levels of incomplete lineage sorting and a low phylogenetic signal may also explain the disruption of the unambiguous phylogeographic structure.

Phylogeographic Structure

The topology of the phylogeny inferred from the mitochondrial genomes in this study largely agreed with earlier studies, some of which were complemented with a few nuclear loci (Bérubé et al. 1998, 2002; Archer et al. 2013; Cabrera et al. 2019; Pérez-Alvarez et al. 2021; Buss et al. 2023). In agreement with Cabrera et al. (2019), the new mitochondrial genome haplotypes detected in this study revealed all ocean basins to be polyphyletic or paraphyletic, contrasting to previous work (Archer et al. 2013) and likely a result of a larger sample size. The haploid, non-recombining mitochondrial genomes introduced into a population may persist in the recipient population or disappear altogether due to random genetic drift and thus may go undetected if low frequencies and sample sizes are insufficient to capture rare haplotypes (Funk and Omland 2003; Ballard and Whitlock 2004).

Notwithstanding the above, both the concatenated and ASTRAL consensus phylogenies based on the autosomal genomes uncovered a clear phylogeographic structure in which samples from each ocean basin were monophyletic. This distinct phylogeographic structure aligns with the most recent proposed subspecies in fin whales (Archer et al. 2019), apart from a subdivision of the Southern Ocean fin whales into *B. p. quoyi* and *B. p. Patachonica*, as no genetic samples have yet been assigned unequivocally to either subspecies. We observed some degree of substructure among the more recent nodes, possibly indicating a recent divergence within the ocean basin. For instance, samples from the highly divergent fin whale population in the Gulf of California (Bérubé et al. 1998, 2002; Rivera-León et al. 2019) often formed a distinct subclade within a larger North Pacific clade (Fig. 2a). Bérubé et al. (1998) reported results from a population genetic analysis that suggested some degree of divergence between western and eastern North Atlantic fin whales. The concatenated phylogeny estimated here also supported some degree of subdivision between the western and Icelandic samples in the North Atlantic, albeit with low support in the ASTRAL consensus phylogeny (Supplementary Fig. S3, S4). We also did not detect any well-supported subclades among samples from the Southern Ocean in the phylogenies estimated from the whole autosomal genomes. The samples from the Southern Ocean in our study originated off the Antarctic Peninsula and may not represent the entire Southern Ocean. Nonetheless, our results were in agreement with previous studies focusing on a few autosomal loci with a larger sampling range in the Southern Ocean (Pérez-Alvarez et al. 2021; Buss et al. 2023). Although our ASTRAL consensus and concatenated results align with an unambiguous phylogeographic structure, they also reveal pervasive genome-wide discordances, including differences among sex chromosomes, autosomal, and mitochondrial genomes. Multiple factors might be the underlying mechanism of these discordances, including a byproduct of scenarios with introgression, incomplete lineage sorting, and/or low phylogenetic signal (Pamilo and Nei 1988; Jeffroy et al. 2006).

A Scenario With Introgression

The effective population size (N_E) of the haploid, maternally inherited mitochondrial genome is fourfold lower than that of the autosomal nuclear genome. Consequently, mitochondrial lineages are expected to sort comparatively faster (Rosenberg 2003). If incomplete lineage sorting is not the sole underlying cause of the observed polyphyly inferred from the mitochondrial genomes, introgression may have contributed to the observed polyphyly (with respect to oceanic origin) inferred from the mitochondrial genomes in fin whales (Pamilo and Nei 1988; Maddison 1997). Past paleoclimatic cycles may have facilitated secondary contact between otherwise isolated oceanic populations (Bérubé et al. 1998; Palsbøll et al. 2007; Alter et al. 2012, 2015; Carroll et al. 2019; Cabrera et al. 2022). During glacial periods, inter-oceanic gene flow was likely elevated when the ranges of baleen whales were contracted towards lower latitudes (Bérubé et al. 1998; Carroll et al. 2019; Cabrera et al. 2022). In such cases, introgression of mitochondrial genomes from larger populations into smaller populations will result in sink-source dynamics (Ballard and Whitlock 2004). The work by Cabrera et al. (2019) suggested directional introgression of mitochondrial haplotypes from the much larger fin whale population in the Southern Ocean into the Northern

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025

Hemisphere populations. Signs of introgression were also detected in the autosomal genome, evident by significant *D*-statistic values at 6.9% of the autosomal genomic windows. However, this type of estimation can be sensitive to short branch lengths, for example, recent radiation events and small population sizes, both of which may apply to our study (Zheng and Janke 2018).

We detected phylogenetic discordance among the topologies estimated from the mitochondrial genomes, the Y and X chromosomes, and the autosomal genome. Male-mediated gene flow may yield phylogenetic discordance between sexual and autosomal chromosomes (Yu et al. 2021; Ge et al. 2022; de Jong et al. 2023; Sørensen et al. 2023). Many cetacean species display sex-specific migratory behaviors, such as well-documented maternal-directed philopatry (Clapham and Seipt 1991; Palsbøll et al. 1995; Lyrholm et al. 1999; Engelhaupt et al. 2009; Baker et al. 2013). Specifically, male-mediated gene flow has been suggested in other baleen whales, such as humpback whales (Baker et al. 1986, 2013; Palumbi and Baker 1994; Amaral et al. 2016), and other cetaceans, for example, sperm whales, *Physeter macrocephalus* (Lyrholm et al. 1999). Male-meditated gene flow could partly explain the phylogenetic discordances among the Y chromosome, X chromosome, and autosomal genome in our study. For instance, male-mediated gene flow is consistent with the higher degree of phylogeographic structure in the X chromosome windows-based phylogenies detected in this study, as recently observed in brown bears (*Ursus arctos*; de Jong et al. 2023), and baboons (*Papio sp.*; Sørensen et al. 2023). Furthermore, polyphyletic clades were only detected in the Y chromosome-based consensus and concatenated phylogenies. Although the Y chromosome results should be interpreted with caution, due to the low sample size and nonstandard recombination, the presence of polyphyletic clades in the ocean basin with most samples, the North Atlantic, suggests putative male-mediated gene flow on the species (Supplementary Fig. S7). Altogether, sings of introgression in the autosomal genome, and the discordance found among sex chromosomes, mitochondrial, and autosomal genome phylogenies may indicate a complex introgression puzzle in fin whales, with likely periods when bi-parental gene flow was prominent, and other periods during which male-mediated gene flow might have prevailed (Petit and Excoffier 2009; de Jong et al. 2023; Sørensen et al. 2023). Yet, the observed, pervasive, genome-wide phylogenetic discordances could also in part stem from incomplete lineage sorting and low phylogenetic signal (Maddison 1997; Scornavacca and Galtier 2017; Wang et al. 2018; Rivas-González et al. 2023).

Incomplete Lineage Sorting and Low Phylogenetic Signal

Incomplete lineage sorting and low phylogenetic signal may also explain the phylogenetic discordances across the genomes (Pamilo and Nei 1988; Maddison 1997). For instance, large fractions of the autosomal genome did not support any ocean basin monophyly, nor did they show significant signs of introgression (1M bps: 32%, 100k bps: 84%, 50k bps: 95%). We observed discrepancies in the placement of deeper nodes among the mitochondrial, sex chromosomes, and autosomal genomes. In addition to that, the topology weighting results varied slightly across the autosomal genome, where alternative topologies had only marginally lower values than the best-weighted topology. Therefore, this inability to accurately retrieve the species tree may be due to high levels of incomplete lineage sorting and/or low phylogenetic signal (Kutschera et al. 2014; Wang et al. 2018; Meleshko et al. 2021). Across all estimations, the highest level of discordance was observed in the Southern Ocean. The Southern Ocean fin whales have the largest long-term population size and the highest genetic diversity, according to Cabrera et al. (2022). This results in longer coalescent times, which directly increases the likelihood of incomplete lineage sorting. The high degree of isolation and smaller effective population size (N_E) in the Gulf of California (Rivera-León et al. 2019; Nigenda-Morales et al. 2023) implies comparatively reduced coalescence times and hence a corresponding higher degree of phylogeographic structure, consistent with our results. Processes, such as recombination, are also positively correlated with N_E, which in turn increases the degree of incomplete lineage sorting (Schierup and Hein 2000; Rivas-González et al. 2023).

We also uncovered high levels of topology discordances among the phylogenies estimated from more narrow windows across the nuclear genome. Although the different window sizes (1M, 100k bps, and 50k bps) all converged onto a similar consensus phylogeny, we also observed a positive correlation between window size and concordance with the ASTRAL consensus phylogeny. Such a positive correlation could simply be due to an increase in the phylogenetic signal as more informative sites are included in each window. The positive correlation could also reflect the masked loci-specific, discordant, evolutionary histories with increasing window sizes, as larger windows may reflect the overall species history (Maddison 1997; Jeffroy et al. 2006). Irrespective of the underlying cause, the results illustrate that phylogenies inferred from both single or few loci and concatenated genes, that is, the entire autosomal genome, provide somewhat limited insight into the different processes driving the species' evolutionary history (Pamilo and Nei 1988; Maddison 1997; Jeffroy et al. 2006; Kubatko and Degnan 2007).

Subspecies or Not?

Suppose a phylogeographic structured intraspecific phylogeny is employed as means to define subspecies (Moritz 1994). In such a case, our concatenated and ASTRAL consensus autosomal genome phylogenies suggest at least 3 fin whale subspecies, each specific to their respective ocean basins, aligning with the subspecies proposed by Archer et al. (2019). These are; *B. p. physalus* in the North Atlantic, *B. p. velifera* in the North Pacific, and *B. p. spp.* (i.e., *B. p. quoyi* and *B. p.*

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025

patachonica) in the Southern Ocean. However, if subspecies are defined in this manner, our data suggests the fin whales from the Gulf of California should be assigned to a separate subspecies as well. The Gulf of California fin whales comprise a genetically distinct, small population which have been isolated for a considerable period (Bérubé et al. 1998; Rivera-León et al. 2019; Nigenda-Morales et al. 2023). Taylor et al. (2017) emphasized the benefits of using subspecies level in cetacean conservation, which includes a possible shift from long-term broader to more immediate local actions granted to lower taxonomic units. If that is the case in fin whales, a shift to subspecies could lead to more effective local management efforts in the small population of the Gulf of California (Rivera-León et al. 2019; Nigenda-Morales et al. 2023).

Our results detected the presence of mitochondrial polyphyletic clades, signs of inter-oceanic introgression (Cabrera et. al. 2022), and high levels of discordance across the autosomal genomic windows to the consensus tree. These findings also raise the question of whether the notion of subspecies in fin whales is biologically sensible (Rosenberg 2003; Burbrink et al. 2022). First, these results reinforce the oft-raised issues in defining lower taxonomic units from uni-parental inherited or non-informative loci (e.g., a few dozen SNPs), especially in recently diverged taxa with a high dispersal capacity in an environment devoid of barriers (Funk and Omland 2003; Ballard and Whitlock 2004). Second, as a great part of autosomal window phylogenies failed to accurately capture the species' history, these high levels of discordance (mostly in the Southern Ocean) may indicate recent divergence, hence the lack of taxonomic variation in fin whales (Funk and Omland 2003). Third, while an accurate taxonomic delineation is desirable (e.g., Shirley et al. 2014; Taylor et al. 2017; Devitt et al. 2019), taxonomic inflation can prevent sound conservation strategies (e.g., Zink 2004; Berrilli et al. 2024; Clavero et al. 2024). Regardless, in-depth studies focusing on extensive sampling (especially in the Southern Ocean), robust estimates of divergence times, past and contemporary migration rates, and changes in population sizes will further enhance our understanding of the evolution, hence aid to the systematics of fin whales. In addition, the impact of structural and functional inter- and intra-specific genetic variance on phylogenetic estimates could help improve phylogenetic assessments in the species. For instance, other work has demonstrated the role of introgressed genomic regions in driving local adaptation (Richards and Martin 2017; Jones et al. 2018; Eberlein et al. 2019). Others have demonstrated a positive correlation between highly recombined regions and incomplete lineage sorting (Rivas-González et al. 2023).

Conclusion

Here we shed light on the complex phylogenetic scenario in a baleen whale species with no clear geographical and reproductive boundaries. We demonstrated the potential for whole genome analysis to uncover the complexity underlying a single consensus phylogeny in a conspecific model. The fin whale consensus phylogeny inferred from the autosomal genome revealed a phylogeographic structure that aligned with ocean basins and thus supported the most recent proposal of subspecies in fin whales (Archer et al. 2019). However, we also detected phylogenetic discordances when comparing the topologies estimated from the mitochondrial genomes, the Y chromosome, and among different fixed windows in the autosomal genome, most likely caused by a degree of introgression, high levels of incomplete lineage sorting, and/or low phylogenetic signal. We highlight the importance of an accurate taxonomic delineation in fin whales, showcasing the inherent issues of phylogenetic inferences and taxonomic revisions based solely on organelle genome sequences, or uni-parentally inherited and noninformative loci. Genome-wide assessments hold the potential to uncover details of the processes shaping the phylogenetic relationships among organisms, such as the effect of introgression and incomplete lineage sorting, here in one of the most enigmatic baleen whales.

Supplementary Material

Supplementary material and data available from the Dryad Digital Repository https://dx.doi.org/10.5061/dryad.v6wwpzh24

Funding

This work has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement no. 813383, and the University of Groningen. The National Council for Scientific and Technological Development (CNPq). The European Research Agency is not responsible for any use that may be made of the information contained in this publication.

Acknowledgments

We would like to thank the Center for Information Technology of the University of Groningen for their support and access to the University of Groningen High-Performance Computing cluster. We also would like to thank all the many people and institutions involved in collecting and providing samples. We thank the Brazilian Antarctic Program (PROANTAR) and the Secretary of the Interministerial Commission for the Resources of the Sea (SECIRM) for providing logistical support for researchers from the Marine Megafauna Ecology and Conservation laboratory at the Universidade Federal do Rio Grande-FURG to collect whale samples in the Southern Ocean.

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025

Author Contributions

F.F., P.J.P., and M.B. conceived and designed the study. E.R.S., D.D., C.R., F.L., S.M., J.R., R.S., J.U.R., performed field work and provided samples. M.B. and F.F. conducted laboratory work. F.F. carried out bioinformatic analyses supervised by P.J.P, C.S., and E.R.S. F.F., M.B., and P.J.P interpreted the results with contributions from all authors. F.F., M.B., P.J wrote the manuscript with crucial input from all authors. All authors approved the final version of this manuscript.

Data Availability

Raw sequences produced in this study are available under the NCBI BioProject code PRJNA1145264. Scripts used for this study are available on https://github.com/fabriciofurni/phylo_fin.

REFERENCES

Abadi S., Azouri D., Pupko T., Mayrose I. 2019. Model selection may not be a mandatory step for phylogeny reconstruction. Nat. Commun. 10(1):Article 1. doi: 10.1038/s41467-019-08822-w

Alter S.E., Meyer M., Post K., Czechowski P., Gravlund P., Gaines C., Rosenbaum H.C., Kaschner K., Turvey S.T., van der Plicht J., Shapiro B., Hofreiter M. 2015. Climate impacts on transocean dispersal and habitat in gray whales from the Pleistocene to 2100. Mol. Ecol. 24(7):1510–1522. doi: 10.1111/mec.13121

Alter S.E., Rosenbaum H.C., Postma L.D., Whitridge P., Gaines C., Weber D., Egan M.G., Lindsay M., Amato G., Dueck L., Brownell R.L., Heide-Jørgensen M.-P., Laidre K.L., Caccone G., Hancock B.L. 2012. Gene flow on ice: the role of sea ice and whaling in shaping Holarctic genetic diversity and population differentiation in bowhead whales (*Balaena mysticetus*). Ecol. Evol. 2(11):2895–2911. doi: 10.1002/ece3.397

Amaral A.R., Loo J., Jaris H., Olavarria C., Thiele D., Ensor P., Aguayo A., Rosenbaum H.C. 2016. Population genetic structure among feeding aggregations of humpback whales in the Southern Ocean. Mar Biol 163(6):132. doi:10.1007/s00227-016-2904-0

Amos W., Hoelzel A.R. 1991. Long-term skin preservation of whale skin for DNA analysis. Rep Int Whaling Comm Spec. Issue. 13:99–103.

Ané C., Larget B., Baum D.A., Smith S.D., Rokas A. 2007. Bayesian estimation of concordance among gene trees. Mol. Biol. Evol. 24(2):412–426. doi: 10.1093/molbev/msl170

Archer F.I., Brownell R.L., Hancock-Hanser B.L., Morin P.A., Robertson K.M., Sherman K.K., Calambokidis J., Urbán R J., Rosel P.E., Mizroch S.A., Panigada S., Taylor B.L. 2019. Revision of fin whale *Balaenoptera physalus* (Linnaeus, 1758) subspecies using genetics. J. Mammal. 100(5):1653–1670. doi: 10.1093/jmammal/gyz121

Archer F.I., Morin P.A., Hancock-Hanser B.L., Robertson K.M., Leslie M.S., Bérubé M., Panigada S., Taylor B.L. 2013. Mitogenomic phylogenetics of fin whales (*Balaenoptera physalus* spp.): genetic evidence for revision of subspecies. PLoS One. 8(5):e63396. doi: 10.1371/journal.pone.0063396

Arnold J. 1993. Cytonuclear disequilibria in hybrid zones. Annu. Rev. Ecol. Syst. 24:521–554.

Baker C.S., Steel D., Calambokidis J., Falcone E., González-Peral U., Barlow J., Burdin A.M., Clapham P.J., Ford J.K.B., Gabriele C.M., Mattila D., Rojas-Bracho L., Straley J.M., Taylor B.L., Urbán J., Wade P.R., Weller D., Witteveen B.H., Yamaguchi M. 2013. Strong maternal fidelity and natal philopatry shape genetic structure in North Pacific humpback whales. Mar. Ecol. Prog. Ser. 494:291–306. doi: 10.3354/meps10508

Ballard J.W.O., Whitlock M.C. 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13(4):729–744. doi: 10.1046/j.1365-294X.2003.02063.x

Berrilli E., Gambioli B., Bombi P., Garzia M., Muraro M., Pardo C., Reale M., Sherpa S., Ficetola G.F., Vignoli L., Salvi D. 2024. Subspecies inflation hampers conservation efforts: a case study on wall lizards. Biol. J. Linn. Soc 143(2):blae001. doi: 10.1093/biolinnean/blae001

Bérubé M., Aguilar A., Dendanto D., Larsen F., Sciara G.N.D., Sears R., Sigurjónsson J., Urban-R J., Palsbøll P.J. 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. Mol. Ecol. 7(5):585–599. doi: 10.1046/j.1365-294x.1998.00359.x

Bérubé M., Urbán J., Dizon A.E., Brownell R.L., Palsbøll P.J. 2002. Genetic identification of a small and highly isolated population of fin whales (*Balaenoptera physalus*) in the Sea of Cortez, México. Conserv Genet. 3(2):183–190. doi: 10.1023/A:1015224730394

Bonferroni C. 1936. Teoria statistica delle classi e calcolo delle probabilita. Pubblicazioni Del R Istituto Superiore Di Scienze Economiche e Commericiali Di Firenze. 8:3–62.

Bukhman Y.V., Morin P.A., Meyer S., Chu L.-F., Jacobsen J.K., Antosiewicz-Bourget J., Mamott D., Gonzales M., Argus C., Bolin J., Berres M.E., Fedrigo O., Steill J., Swanson S.A., Jiang P., Rhie A., Formenti G., Phillippy A.M., Harris R.S., Wood J.M.D., Howe K., Kirilenko B.M., Munegowda C., Hiller M., Jain A., Kihara D., Johnston J.S., Ionkov A., Raja K., Toh H., Lang A., Wolf M., Jarvis E.D., Thomson J.A., Chaisson M.J.P., Stewart R., Gaut B. 2024. A High-Quality Blue Whale Genome, Segmental Duplications, and Historical Demography. Molecular Biology and Evolution 41(3). doi:10.1093/molbev/msae036

Burbrink F.T., Crother B.I., Murray C.M., Smith B.T., Ruane S., Myers E.A., Pyron R.A. 2022. Empirical and philosophical problems with the subspecies rank. Ecol. Evol. 12(7):e9069. doi:10.1002/ece3.9069

Buss D.L., Atmore L.M., Zicos M.H., Goodall-Copestake W.P., Brace S., Archer F.I., Baker C.S., Barnes I., Carroll E.L., Hart T., Kitchener A., Sabin R., Sremba A.L., Weir C.R., Jackson J.A. 2023. Historical mitogenomic diversity and population structuring of southern hemisphere fin whales. Genes. 14(5):Article 5. doi: 10.3390/genes14051038

Cabrera A.A., Hoekendijk J.P.A., Aguilar A., Barco S.G., Berrow S., Bloch D., Borrell A., Cunha H.A., Dalla Rosa L., Dias C.P., Gauffier P., Hao W., Landry S., Larsen F., Martín V., Mizroch S., Oosting T., Øien N., Pampoulie C., Bérubé M., Prieto R., Ramp C., Rivera-Léon V., Robbins J., Ryan C., Schall E., Sears R., Silva M.A., Urbán J., Wenzel F.W., Palsbøll P.J., Bérubé M. 2019. Fin whale (*Balaenoptera physalus*) mitogenomics: a cautionary tale of defining sub-species from mitochondrial sequence monophyly. Mol. Phylogenet. Evol. 135:86–97. doi: 10.1016/j.ympev.2019.02.003

Cabrera A.A., Schall E., Bérubé M., Anderwald P., Bachmann L., Berrow S., Best P.B., Clapham P.J., Cunha H.A., Dalla Rosa L., Dias C., Findlay K.P., Haug T., Heide-Jørgensen M.P., Hoelzel A.R., Kovacs K.M., Landry S., Larsen F., Lopes X.M., Palsbøll P.J., Mattila D.K., Oosting T., Pace R.M., Papetti C., Paspati A., Pastene L.A., Prieto R., Ramp C., Robbins J., Sears R., Secchi E.R., Silva M.A., Simon M., Víkingsson G., Wiig O., Øien N., Palsbøll P.J. 2022. Strong and lasting impacts of past global warming on baleen whales and their prey. Global Change Biol. 28(8):2657–2677. doi: 10.1111/gcb.16085

Carroll E.L., Alderman R., Bannister J.L., Bérubé M., Best P.B., Boren L., Baker C.S., Constantine R., Findlay K., Harcourt R., Lemaire L., Palsbøll P.J., Patenaude N.J., Rowntree V.J., Seger J., Steel D., Valenzuela L.O., Watson M., Gaggiotti O.E. 2019. Incorporating non-equilibrium dynamics into demographic history inferences of a migratory marine species. Heredity. 122(1):Article 1. doi: 10.1038/s41437-018-0077-y

Chen L., Qiu Q., Jiang Y., Wang K., Lin Z., Li Z., Bibi F., Yang Y., Wang J., Nie W., Su W., Liu G., Li Q., Fu W., Pan X., Liu C., Yang J., Zhang C., Yin Y., Wang Y., Zhao Y., Zhang C., Wang Z., Qin Y., Liu W., Wang B., Ren Y., Zhang R., Zeng Y., da Fonseca R.R., Wei B., Li R., Wan W., Zhao R., Zhu W., Wang Y., Duan S., Gao Y., Zhang Y.E., Chen C., Hvilsom C., Epps C.W., Chemnick L.G., Dong Y.,

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025

Mirarab S., Siegismund H.R., Ryder O.A., Gilbert M.T.P., Lewin H.A., Zhang G., Heller R., Wang W., Wang W. 2019. Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. Science. 364(6446):eaav6202. doi: 10.1126/science.aav6202

Clapham P.J., Seipt I.E. 1991. Resightings of independent fin whales, *Balaenoptera physalus*, on maternal summer ranges. J. Mammal. 72(4):788–790. doi: 10.2307/1381844

Clarke R. 2004. Pygmy fin whales. Mar. Mamm. Sci. 20(2):329–334. doi: 10.1111/j.1748-7692.2004.tb01161.x

Clavero M., Naves J., Lucena-Perez M., Revilla E. 2024. Taxonomic inflation as a conservation trap for inbred populations. Evol. Appl. 17(5):e13677. doi: 10.1111/eva.13677

Danecek P., Auton A., Abecasis G., Albers C.A., Banks E., DePristo M.A., Handsaker R.E., Lunter G., Marth G.T., Sherry S.T. McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. Bioinformatics. 27(15): 2156–2158. doi: 10.1093/bioinformatics/btr330

de Jong M.J., Niamir A., Wolf M., Kitchener A.C., Lecomte N., Seryodkin I.V., Fain S.R., Hagen S.B., Saarma U., Janke A. 2023. Range-wide whole-genome resequencing of the brown bear reveals drivers of intraspecies divergence. Commun. Biol. 6(1):Article 1. doi: 10.1038/s42003-023-04514-w

Delsuc F., Brinkmann H., Philippe H. 2005. Phylogenomics and the reconstruction of the tree of life. Nat. Rev. Genet. 6(5):361–375. doi: 10.1038/nrg1603

Devitt T.J., Wright A.M., Cannatella D.C., Hillis D.M. 2019. Species delimitation in endangered groundwater salamanders: implications for aquifer management and biodiversity conservation. Proc. Natl. Acad. Sci. U.S.A. 116(7):2624–2633. doi: 10.1073/pnas.1815014116

Durand E.Y., Patterson N., Reich D., Slatkin M. 2011. Testing for ancient admixture between closely related populations. Mol. Biol. Evol. 28(8):2239–2252. doi: 10.1093/molbev/msr048

Eberlein C., Hénault M., Fijarczyk A., Charron G., Bouvier M., Kohn L.M., Anderson J.B., Landry C.R. 2019. Hybridization is a recurrent evolutionary stimulus in wild yeast speciation. Nat. Commun. 10(1):Article 1. doi: 10.1038/s41467-019-08809-7

Eisen J.A., Fraser C.M. 2003. Phylogenomics: intersection of evolution and genomics. Science. 300(5626):1706–1707. doi: 10.1126/science.1086292

Engelhaupt D., Rus Hoelzel A., Nicholson C., Frantzis A., Mesnick S., Gero S., Whitehead H., Rendell L., Miller P., De Stefanis R., Cañadas A., Airoldi S., Mignucci-Giannoni A.A. 2009. Female philopatry in coastal basins and male dispersion across the North Atlantic in a highly mobile marine species, the sperm whale (*Physeter macrocephalus*). Mol. Ecol. 18(20):4193–4205. doi: 10.1111/j.1365-294X.2009.04355.x

Funk D.J., Omland K.E. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34:397–423.

Ge D., Feijó A., Wen Z., Lissovsky A., Zhang D., Cheng J., Yan C., Mu D., Wu X., Xia L., Yang Q. 2022. Ancient introgression underlying the unusual mito-nuclear discordance and coat phenotypic variation in the Moupin pika. Divers. Distrib. 28(12):2593–2609. doi: 10.1111/ddi.13479

Green R.E., Krause J., Briggs A.W., Maricic T., Stenzel U., Kircher M., Patterson N., Li H., Zhai W., Fritz M.H.Y., Hansen N.F., Durand E.Y., Malaspinas A.S., Jensen J.D., Marques-Bonet T., Alkan C., Prufer K., Meyer M., Burbano H.A., Paabo S. 2010. A draft sequence of the neandertal genome. Science. 328(5979):710–722. doi: 10.1126/science.1188021

Ichihara T. 1966. The pygmy blue whale, *Balaenoptera musculus brevicauda*, a new subspecies. Whales, Dolphins, Porpoises. 1:79.

Jackson J.A., Steel D.J., Beerli P., Congdon B.C., Olavarría C., Leslie M.S., Pomilla C., Rosenbaum H., Baker C.S. 2014. Global diversity and oceanic divergence of humpback whales (*Megaptera novaeangliae*). Proc. Biol. Sci. 281(1786):20133222. doi: 10.1098/rspb.2013.3222

Jeffroy O., Brinkmann H., Delsuc F., Philippe H. 2006. Phylogenomics: the beginning of incongruence? Trends Genet. 22(4):225–231. doi: 10.1016/j.tig.2006.02.003

Jones M.R., Mills L.S., Alves P.C., Callahan C.M., Alves J.M., Lafferty D.J.R., Jiggins F.M., Jensen J.D., Melo-Ferreira J., Good J.M. 2018. Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. Science 360(6395):1355–1358. doi: 10.1126/science.aar5273

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30(4):772–780. doi: 10.1093/molbev/mst010

Korneliussen T.S., Albrechtsen A., Nielsen R. 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinf. 15(1):356. doi: 10.1186/s12859-014-0356-4

Kubatko L.S., Degnan J.H. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56(1):17–24. doi: 10.1080/10635150601146041

Kutschera V.E., Bidon T., Hailer F., Rodi J.L., Fain S.R., Janke A. 2014. Bears in a forest of gene trees: phylogenetic inference is complicated by incomplete lineage sorting and gene flow. Mol. Biol. Evol. 31(8):2004–2017. doi: 10.1093/molbev/msu186

Li H., Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 25(14):1754–1760.

Lockyer C.H., Brown S.G. 1981. The migration of whales. In: Aidley D.J., Animal migration, Vol. 13. NY: Cambridge University Press. p. 105–137.

Lönnberg, E. 1906. Contributions to the Fauna of South Georgia: Taxonomic and biological notes on vertebrates, Vol. 40. Uppsala & Stockholm: Almqvist & Wiksell.

Lyrholm T., Leimar O., Johanneson B., Gyllensten U. 1999. Sex-biased dispersal in sperm whales: contrasting mitochondrial and nuclear genetic structure of global populations. Proc R Soc Lond B Biol Sci. 266(1417):347–354. doi: 10.1098/rspb.1999.0644

Mackintosh N.A. 1946. The natural history of whalebone whales. Biol. Rev. 21(2):60–74. doi: 10.1111/j.1469-185X.1946.tb00453.x

Maddison W.P. 1997. Gene trees in species trees. Syst. Biol. 46(3):523–536. doi: 10.1093/sysbio/46.3.523

Malinsky M., Matschiner M., Svardal H. 2021. Dsuite—Fast D-statistics and related admixture evidence from VCF files. Mol. Ecol. Resour. 21(2):584–595. doi: 10.1111/1755-0998.13265

Martin S.H., Van Belleghem S.M. 2017. Exploring evolutionary relationships across the genome using topology weighting. Genetics 206(1):429–438. doi: 10.1534/genetics.116.194720

McGowen M.R., Tsagkogeorga G., Álvarez-Carretero S., dos Reis M., Struebig M., Deaville R., Jepson P.D., Jarman S., Polanowski A., Morin P.A., Rossiter S.J. 2020. Phylogenomic resolution of the cetacean tree of life using target sequence capture. Syst. Biol. 69(3):479–501. doi: 10.1093/sysbio/syz068

McKenna A., Hanna M., Banks E., Sivachenko A., Cibulskis K., Kernytsky A., Garimella K., Altshuler D., Gabriel S., Daly M., DePristo M.A. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20(9):1297–1303. doi: 10.1101/gr.107524.110

Meleshko O., Martin M.D., Korneliussen T.S., Schröck C., Lamkowski P., Schmutz J., Healey A., Piatkowski B.T., Shaw A.J., Weston D.J., Flatberg K.I., Szövényi P., Hassel K., Stenøien H.K. 2021. Extensive genome-wide phylogenetic discordance is due to incomplete lineage sorting and not ongoing introgression in a rapidly radiated bryophyte genus. Mol. Biol. Evol. 38(7):2750–2766. doi: 10.1093/molbev/msab063

Minh B.Q., Schmidt H.A., Chernomor O., Schrempf D., Woodhams M.D., von Haeseler A., Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37(5):1530–1534. doi: 10.1093/molbev/msaa015

Misof B., Liu S., Meusemann K., Peters R.S., Donath A., Mayer C., Frandsen P.B., Ware J., Flouri T., Beutel R.G., Niehuis O., Petersen M., Izquierdo-Carrasco F., Wappler T., Rust J., Aberer A.J., Aspöck U., Aspöck H., Bartel D., Blanke A., Berger S., Böhm A., Buckley T.R., Calcott B., Chen J., Friedrich F., Fukui M., Fujita M., Greve C., Grobe P., Gu S., Huang Y., Jermiin L.S., Kawahara A.Y., Krogmann L., Kubiak M., Lanfear R., Letsch H., Li Y., Li Z., Li J., Lu H., Machida R., Mashimo Y., Kapli P., McKenna D.D., Meng G., Nakagaki Y., Navarrete-Heredia J.L., Ott M., Ou Y., Pass G., Podsiadlowski L., Pohl H., von Reumont B.M., Schütte K., Sekiya

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025

K., Shimizu S., Slipinski A., Stamatakis A., Song W., Su X., Szucsich N.U., Tan M., Tan X., Tang M., Tang J., Timelthaler G., Tomizuka S., Trautwein M., Tong X., Uchifune T., Walzl M.G., Wiegmann B.M., Wilbrandt J., Wipfler B., Wong T.K.F., Wu Q., Wu G., Xie Y., Yang S., Yang Q., Yeates D.K., Yoshizawa K., Zhang Q., Zhang R., Zhang W., Zhang Y., Zhao J., Zhou C., Zhou L., Ziesmann T., Zou S., Li Y., Xu X., Zhang Y., Yang H., Wang J., Wang J., Kjer K.M., Zhou X., Zhou X. 2014. Phylogenomics resolves the timing and pattern of insect evolution. Science 346(6210):763–767. doi: 10.1126/science.1257570

Mizroch S.A., Rice D.W., Breiwick J.M. 1984. The fin whale, *Balaenoptera physalus*. Mar. Fish. Rev. 46(4):20–24.

Moritz C. 1994. Defining 'Evolutionarily Significant Units' for conservation. Trends Ecol. Evolut. 9(10):373–375. doi: 10.1016/0169-5347(94)90057-4

Nigenda-Morales, S.F., Lin, M., Nuñez-Valencia, P.G., Kyriazis, C.C., Beichman, A.C., Robinson, J.A., Ragsdale, A.P., Urbán, R.J., Archer, F.I., Viloria-Gómora, L., Pérez-Álvarez, M.J., Poulin, E., Lohmueller, K.E., Moreno-Estrada, A., Wayne, R.K. 2023. The genomic footprint of whaling and isolation in fin whale populations. Nat. Commun. 14(1):Article 1. doi: 10.1038/s41467-023-40052-z

Ortiz E.M. 2019. *vcf2phylip v2.0: Convert a VCF matrix into several matrix formats for phylogenetic analysis*. [Computer software]. Zenodo. doi: 10.5281/zenodo.2540861

Palsbøll P.J. 1991. Sampling of skin biopsies from free-ranging large cetaceans in West Greenland: development of new biopsy tips and bolt designs. Rep. Int. Whal. Commn, special issue. 13:17–19.

Palsbøll P.J., Bérubé M., Larsen F. 2007. Could genetic diversity in eastern North Pacific gray whales reflect global historic abundance? Proc. Natl. Acad. Sci. U.S.A. 104(52):E2; author reply E3–E2; author reply E2. doi: 10.1073/pnas.0710072105

Palsbøll P.J., Clapham P.J., Mattila D.K., Larsen F., Sears R., Siegismund H.R., Sigurjónsson J., Vasquez O., Arctander P. 1995. Distribution of mtDNA haplotypes in North Atlantic humpback whales: the influence of behaviour on population structure. Mar. Ecol. Prog. Ser. 116(1/3):1–10.

Palumbi S.R., Baker C.S. 1994. Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. Mol. Biol. Evol. 11(3):426–435. doi: 10.1093/oxfordjournals.molbev.a040115

Pamilo P., Nei M. 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5(5):568–583. doi: 10.1093/oxfordjournals.molbev.a040517

Pérez-Alvarez M., Kraft S., Segovia N.I., Olavarría C., Nigenda-Morales S., Urbán R.J., Viloria-Gómora L., Archer F., Moraga R., Sepúlveda M., Santos-Carvallo M., Pavez G., Poulin E. 2021. Contrasting phylogeographic patterns among northern and southern hemisphere fin whale populations with new data from the Southern Pacific. Front. Mar. Sci. 8:630233. https://www.frontiersin.org/article/10.3389/fmars.2021.630233

Petit R.J., Excoffier L. 2009. Gene flow and species delimitation. Trends Ecol. Evol. 24(7):386–393. doi: 10.1016/j.tree.2009.02.011

R Core Team. (2020). R: A language and environment for statistical computing (4.1) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Richards E.J., Martin C.H. 2017. Adaptive introgression from distant Caribbean islands contributed to the diversification of a microendemic adaptive radiation of trophic specialist pupfishes. PLoS Genet. 13(8):e1006919. doi: 10.1371/journal.pgen.1006919

Rivas-González I., Rousselle M., Li F., Zhou L., Dutheil J.Y., Munch K., Shao Y., Wu D., Schierup M.H., Zhang G. 2023. Pervasive incomplete lineage sorting illuminates speciation and selection in primates. Science 380(6648):eabn4409. doi: 10.1126/science.abn4409

Rivera-León V.E., Urbán J., Mizroch S., Brownell R.L., Oosting T., Hao W., Palsbøll P.J., Bérubé M. 2019. Long-term isolation at a low effective population size greatly reduced genetic diversity in Gulf of California fin whales. Sci. Rep. 9(1):Article 1. doi: 10.1038/s41598-019-48700-5

Rocha R.C. Jr., Clapham P.J., Ivashchenko Y. 2015. Emptying the oceans: a summary of industrial whaling catches in the 20th century. Mar. Fish. Rev. 76(4):37–48. doi: 10.7755/MFR.76.4.3

Rosenberg N.A. 2003. The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly, and polyphyly in a coalescent model. Evolution 57(7):1465–1477. doi: 10.1111/j.0014-3820.2003.tb00355.x

Sasaki T., Nikaido M., Hamilton H., Goto M., Kato H., Kanda N., Pastene L.A., Cao Y., Fordyce R.E., Hasegawa M., Okada N. 2005. Mitochondrial phylogenetics and evolution of mysticete whales. Syst. Biol. 54(1):77–90. doi: 10.1080/10635150590905939

Schierup M.H., Hein J. 2000. Consequences of recombination on traditional phylogenetic analysis. Genetics 156(2):879–891. doi: 10.1093/genetics/156.2.879

Scornavacca C., Galtier N. 2017. Incomplete lineage sorting in mammalian phylogenomics. Syst. Biol. 66(1):112–120. doi: 10.1093/sysbio/syw082

Shirley M.H., Vliet K.A., Carr A.N., Austin J.D. 2014. Rigorous approaches to species delimitation have significant implications for African crocodilian systematics and conservation. Proc. Biol. Sci. 281(1776):20132483. doi: 10.1098/rspb.2013.2483

Sørensen E.F., Harris R.A., Zhang L., Raveendran M., Kuderna L.F.K., Walker J.A., Storer J.M., Kuhlwilm M., Fontsere C., Seshadri L., Bergey C.M., Burrell A.S., Bergman J., Phillips-Conroy J.E., Shiferaw F., Chiou K.L., Chuma I.S., Keyyu J.D., Fischer J., Rogers J., Salvi S., Doddapaneni H., Schierup M.H., Batzer M.A., Jolly C.J., Knauf S., Zinner D., Farh K.K.-H., Marques-Bonet T., Munch K., Roos C., Rogers J. 2023. Genome-wide coancestry reveals details of ancient and recent male-driven reticulation in baboons. Science 380(6648):eabn8153. doi: 10.1126/science.abn8153

Stevick P., Allen J., Engel M.H., Félix F., Haase B., Neves M. 2014. First record of inter-oceanic movement of a humpback whale between Atlantic and Pacific breeding grounds off South America. J. Cetacean Res. Manag. 13:159–162.

Taylor B.L., Archer F.I., Martien K.K., Rosel P.E., Hancock-Hanser B.L., Lang A.R., Leslie M.S., Mesnick S.L., Morin P.A., Pease V.L., Perrin W.F., Robertson K.M., Parsons K.M., Viricel A., Vollmer N.L., Cipriano F., Reeves R.R., Krützen M., Baker C.S. 2017. Guidelines and quantitative standards to improve consistency in cetacean subspecies and species delimitation relying on molecular genetic data. Mar. Mamm. Sci. 33((S1):132–155. doi: 10.1111/mms.12411

Wang K., Lenstra J.A., Liu L., Hu Q., Ma T., Qiu Q., Liu J. 2018. Incomplete lineage sorting rather than hybridization explains the inconsistent phylogeny of the wisent. Commun. Biol. 1(1):Article 1. doi: 10.1038/s42003-018-0176-6

Waterhouse A.M., Procter J.B., Martin D.M.A., Clamp M., Barton G.J. 2009. Jalview Version 2—A multiple sequence alignment editor and analysis workbench. Bioinformatics 25(9):1189–1191. doi: 10.1093/bioinformatics/btp033

Wolf M., de Jong M., Halldórsson S.D., Árnason U., Janke A. 2022. Genomic impact of whaling in North Atlantic fin whales. Mol. Biol. Evol. 39(5):msac094. doi: 10.1093/molbev/msac094

Yim H.-S., Cho Y.S., Guang X., Kang S.G., Jeong J.-Y., Cha S.-S., Oh H.-M., Lee J.-H., Yang E.C., Kwon K.K., Kim Y.J., Kim T.W., Kim W., Jeon J.H., Kim S.-J., Choi D.H., Jho S., Kim H.-M., Ko J., Kim H., Shin Y.-A., Jung H.-J., Zheng Y., Wang Z., Chen Y., Chen M., Jiang A., Li E., Zhang S., Hou H., Kim T.H., Yu L., Liu S., Ahn K., Cooper J., Park S.-G., Hong C.P., Jin W., Kim H.-S., Park C., Lee K., Chun S., Morin P.A., O'Brien S.J., Lee H., Kimura J., Moon D.Y., Manica A., Edwards J., Kim B.C., Kim S., Wang J., Bhak J., Lee H.S., Lee J.-H., Lee J.-H. 2014. Minke whale genome and aquatic adaptation in cetaceans. Nat. Genet. 46(1):88–92. doi: 10.1038/ng.2835

Yu H., Xing Y.-T., Meng H., He B., Li W.-J., Qi X.-Z., Zhao J.-Y., Zhuang Y., Xu X., Yamaguchi N., Driscoll C.A., O'Brien S.J., Luo S.-J. 2021. Genomic evidence for the Chinese mountain cat as a wildcat conspecific (*Felis silvestris bieti*) and its introgression to domestic cats. Sci. Adv. 7(26):eabg0221. doi: 10.1126/sciadv.abg0221

Zhang C., Mirarab S. 2022. Weighting by gene tree uncertainty improves accuracy of quartet-based species trees. Mol. Biol. Evol. 39(12):msac215. doi: 10.1093/molbev/msac215

Zheng Y., Janke A. 2018. Gene flow analysis method, the D-statistic, is robust in a wide parameter space. BMC Bioinf. 19(1):10. doi: 10.1186/s12859-017-2002-4

Zink R.M. 2004. The role of subspecies in obscuring avian biological diversity and misleading conservation policy. Proc. Biol. Sci. 271(1539):561–564. doi: 10.1098/rspb.2003.2617

Downloaded from https://academic.oup.com/sysbio/article/73/6/873/7735851 by National Oceanic & Atmospheric Administration user on 18 March 2025