SC/68A/SH/06 Rev1

Genetic diversity and connectivity of southern right whales (Eubalaena australis) found in the Chilean wintering ground and South Georgia/Islas Georgias del Sur feeding ground

Carroll, Gaggiotti, Galletti Vernazzani, Ott, Neveceralova, Vermeulen, McMillan, Andriolo, Baker, Bamford, Best, Cabrera, Calderan, Chirife, Fewster, Flores, Frasier, Freitas, Groch, Hulva, Kennedy, Leaper, Leslie, Moore, Oliveira, Seger, *more listed

INTERNATIONAL
WHALING COMMISSION

GENETIC DIVERSITY AND CONNECTIVITY OF SOUTHERN RIGHT WHALES *(EUBALAENA AUSTRALIS)* FOUND IN THE CHILEAN WINTERING GROUND AND SOUTH GEORGIA/ISLAS GEORGIAS DEL SUR FEEDING GROUND

Emma L. Carroll[1], Oscar E. Gaggiotti[2], Barbara Galletti Vernazzani[3], Paulo H. Ott[4,5], Petra Neveceralova[6], Els Vermeulen[7], Louise F. McMillan[8], Artur Andriolo[9,10], C. Scott Baker[11], Connor Bamford[12] , Peter Best, Elsa Cabrera[3], Susannah Calderan[13], Andrea Chirife[14], Rachel M. Fewster[8], Paulo A. C. Flores[15], Timothy Frasier[16], Thales R. O. Freitas[17] , Karina Groch[18], Pavel Hulva[6,19], Amy Kennedy[20], Russell Leaper[21], Matthew S. Leslie[22], Michael Moore[23], Larissa Oliveira[4,24], Jon Seger[25] , Emilie N. Stepien[26], Luciano O. Valenzuela[25,27], Bradley White[28], Alexandre Zerbini[20], Jennifer A. Jackson[12]

1 School of Biological Sciences, University of Auckland, Auckland, New Zealand
2 School of Biology, University of St Andrews, Scotland, UK
3 Centro de Conservación Cetacea—Casilla 19178 Correo 19, Santiago, Chile
4 Grupo de Estudos de Mamíferos Aquáticos do Rio Grande do Sul, Torres, RS, Brazil
5 Universidade Estadual do Rio Grande do Sul, Osório, RS, Brazil
6 Department of Zoology, Faculty of Science, Charles University
7 Mammal Research Institute Whale Unit, Department of Zoology and Entomology, University of Pretoria, Hatfield, South Africa
8 Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand
9 Departamento de Zoologia, Instituto de Ciências Biológicas, Universidade Federal de Juiz de Fora, Campus Universitário, 36036-900 Juiz de Fora, MG
10 Aqualie Institute, Brazil
11 Marine Mammal Institute and Department of Fisheries and Wildlife, Oregon State University, Newport, OR, 97365 United States
12 British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK
13 Scottish Association for Marine Science, Oban, Argyll, Scotland
14 Instituto de Ciencias Biomédicas (ICB), Universidad Andrés Bello, Chile
15 Área de Proteção Ambiental (Environmental Protection Area) Anhatomirim, ICMBio, MMA, Florianópolis, SC, Brazil
16 Department of Biology, Saint Mary's University, Halifax, Nova Scotia, Canada, B3H 3C3
17 Programa de Pós-Graduação em Genética e Biologia Molecular- Universidade Federal do Rio Grande do Sul (UFRGS)
18 Projeto Baleia Franca/Instituto Australis, Imbituba, Santa Catarina 88780-000, Brazil
19 Department of Biology and Ecology, University of Ostrava
20 National Marine Mammal Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA. 7600 Sand Point Way, NE, Seattle, WA, 98115-6349, USA
21 IFAW. 87-90 Albert Embankment, London, SE1 7UD. UK
22 Biology Department, Swarthmore College, 500 College Avenue, Swarthmore, PA 19081, USA
23 Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, 02543, USA
24 Laboratório de Ecologia de Mamíferos, Universidade do Vale do Rio dos Sinos, Centro de Ciências da Saúde Avenida Unisinos 950 Cristo Rei - Sao Leopoldo 93022-000, RS - Brazil
25 Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA
26 Marine Mammal Research, Department of Bioscience, Aarhus University, Denmark
27 Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Laboratorio de Ecología Evolutiva Humana, UNCPBA, Quequén, Buenos Aires Province, Argentina
28 Department of Biology, Trent University, Peterborough, ON, Canada K9J 7B8

ABSTRACT

Southern right whale (*Eubalaena australis*) population structure can be viewed as a migratory network of winter calving/socialising and summer feeding grounds. Here we investigate the position of the Chile-Peru wintering ground (n = 1) and the South Georgia/Islas Georgias del Sur (SG) feeding ground (n = 15) in the broader migratory network, using mitochondrial DNA (mtDNA) and nuclear microsatellite data (nDNA) from all major wintering grounds. This includes new data from Brazil (n = 60) and South Africa (n = 88), as well as published data from across the species' circumpolar distribution (nDNA = 222; mtDNA = 1327). The single sample from Chile-Peru had a mtDNA haplotype previously only observed in the Indo-Pacific and had a nuclear genotype that appeared admixed between the Indo-Pacific and South Atlantic, based on genetic clustering and assignment algorithms. The SG samples were clearly South Atlantic, based on both genetic differentiation and clustering analyses. As a group, SG was more similar to the Southwest Atlantic wintering grounds (Brazil, Argentina) than to the South African wintering ground, and showed significant genetic differentiation from the latter. However, the weak genetic differentiation amongst the South Atlantic wintering grounds meant that population assignment methods were unable to resolve the likely winter association of the SG samples. This may be overcome using additional loci and/or by limiting comparisons to nursery areas within wintering grounds.

INTRODUCTION

Southern right whale (*Eubalaena australis*) population structure can be viewed as a migratory network of winter calving/socialising and summer feeding grounds. Recent analyses of four major extant wintering grounds showed that there is hierarchical genetic structure in both mitochondrial DNA (mtDNA) and microsatellite loci amongst ocean basins[1], consistent with female philopatry to natal wintering grounds. Correlations between genetic and isotopic variation also support the hypothesis of maternally directed fidelity to summer feeding grounds[2,3]. Connectivity between wintering areas is likely facilitated by migratory corridors and/or shared feeding grounds shared, as indicated by photo-identification genetic markers, stable isotope and tagging studies [3–7].

Here we build on existing work to augment the genotype databases for extant southern right whale wintering grounds by constructing DNA profiles for samples from Brazil and South Africa. These data are combined with previously generated data to create reference genetic datasets with which to understand the position of the Chile-Peru wintering ground and South Georgia/Islas Georgias del Sur (SG) feeding ground in the context of the broader migratory network.

The Chile-Peru southern right whale subpopulation is considered critically endangered by the International Union for the Conservation of Nature[8] and is subject to an International Whaling Commission (IWC) Conservation and Management Plant (CMP)[9], as it appears isolated and small in size. Between 1964 and 2011, there were 108 opportunistic sightings of 179 southern right whales off the coasts of Chile and Peru[10], including 39 calves. Given that female (and male[4,11]) southern right whales show fidelity to wintering grounds, it is possible that Chile-Peru represents distinct genetic lineages. Alternatively, the population could have been extirpated during the whaling era and it could now be undergoing recolonisation from elsewhere. The geographic position of the Chile-Peru population makes this particularly interesting in light of recent genetic analyses showing significant genetic differentiation particularly between the South Atlantic (Argentina, South Africa) and Indo-Pacific (Australia, New Zealand)[1]. The Southeast Pacific Chile-Peru subpopulation is separated from the Southwest Pacific New Zealand subpopulation by a large discontinuity in distribution (approximately 90º to 180ºW) and from the Southwest Atlantic subpopulation by the South American continent. Little is known about the historical or contemporary feeding grounds of the Chile-Peru population that might provide insight into which of these neighbouring populations it may mix with, for example, with Southwest Atlantic whales on shared feeding grounds in summer. Here we conduct the first genetic analysis of the only available sample from this southern right whale subpopulation.

Additionally, we conduct a genetic comparison of the South Atlantic southern right whale wintering grounds (Argentina, Brazil and South Africa) and the summer feeding ground off SG. Southern right whales are the most commonly seen whale species in SG during the austral summer[12,13]. The movement of whales between the Península Valdés nursery area and SG has been shown using photo-identified individuals[7,12] and satellite telemetry work[14]. Previous genetic analyses, based on 10 microsatellite loci and 495 bp of the mtDNA control region, showed that the SG feeding ground and Brazilian and Argentinean wintering grounds were not significantly genetically differentiated[15,16]. Here we expand the analysis to include additional samples from SG and both Southwest and Southeast Atlantic wintering grounds, and increase the number of microsatellite loci examined.

METHODS

Sample collection: Chile

A southern right whale was seen alive on 9 February 2017 in Cocotue Bay, Chile, with clear entanglement scars and lesions, however, poor weather prevented further at-sea observations[17]. Subsequently, a southern right whale carcass was found at Playa Mar Brava, Carelmapu (41º42'S, 73º42'W, see Figure 1), on 16 February, and reported on 18 February 2017. It corresponded to the same individual seen at sea previously and entanglement lesions were confirmed. Skin and blubber samples were collected from the carcass during a necropsy on 19 February 2017[17]. Samples were given to Centro de Conservación Cetacea under the certificate of delivery and custody of marine protected species 2017-54-4 granted by National Fisheries and Aquaculture Service. Samples were transported on dry ice from Chile to the University of St Andrews, Scotland, for analysis.

Sample collection: South Georgia/Islas Georgias del Sur

Sampling in 2018 was conducted from the R/V Song of the Whale, which departed from Stanley in the Falkland Islands/Islas Malvinas on 22 January, sailed to the north coast of SG and returned to Stanley on 21 February (31 days total sailing). Right whales were located using a combination of passive acoustics and visual observations[18]. Visual searching was conducted along vessel tracks which followed bearings to vocalising whales that were received on DIFAR sonobuoys[18]. All data collection was carried out under permit RAP/2017/017.

Three skin biopsy samples were collected from southern right whales around SG using a stainless steel biopsy dart fired from a modified veterinary capture rifle[19] or deployed from a crossbow[20] during the 2018 field season. Samples were stored in ethanol and frozen at -20ºC. An additional 12 skin biopsy samples were available from a 1997 field expedition around SG, as previously described[12].

Sample collection: South Africa

A further 88 skin samples were available from southern right whales sampled on the South African wintering ground that had not previously been analysed. This included 50 skin samples archived at the University of Pretoria and 38 sloughed skin samples collected by P. Neveceralova and P. Hulva from Charles University. The 50 samples held by the University of Pretoria were collected between 1990 and 2013 under permits RES2009/06, RES2011/24, RES2012/89, RES2013/58, V1/9/3/1, VI/19/5/1, and VI/9/5/1. The samples held by Charles University were collected during the austral winter of 2016 and 2017 under research permits RES2016/99 and RES2017/89. This is in addition to the mtDNA control region sequence and microsatellite genotype data from 350 and 46 South African right whales, respectively, previously published[1]. Samples were transferred to the UK under CITES permits: South African export permit 207896 and UK import permit 569921/01(University of Pretoria samples) and under South African export permit 171354 and UK import permit 550270/01 (Charles University samples).

Sample collection: Brazil

Skin samples (n = 60) were collected from southern right whales along the Brazilian coast, primarily at Santa Catarina and Rio Grande do Sul States (between 27º50’S and 29º20’S)[15] using a dart projection system[16,21]. Samples were collected from a small inflatable boat operating from and around the coastline, usually up to 3 - 5 km and rarely as far as 10 km distance from shore, during field work conducted between 1998 and 2007. Field work was undertaken by GEMARS in collaboration with the Projeto Baleia Franca under permits IBAMA # 039/98, 1803/98-23, CMA 004-02, CMA 011-03, SISBIO 12022-1 and exported under CITES permit No. 15BR018985/DF. A small number of samples were collected from carcasses of stranded whales[15]. Samples were placed in 20% DMSO and frozen at -20ºC.

Published data included in the analysis

We augmented the data generated in this study with mtDNA and microsatellite genotype data from Argentina and the Indo-Pacific using published data[1,3] (see Table 1). The mtDNA data for the SG samples collected in 1997 were sourced from Patenaude et al.[6] and therefore we limited our analysis to the 275 bp fragment of mtDNA control region used in this study. Sequencing to a longer fragment length (1000 bp) is planned for these samples in future.

DNA extraction and DNA profile construction

DNA was extracted using standard phenol/chloroform protocols, modified for small sample sizes[22]. We constructed DNA profiles, comprising genetically identified sex, mtDNA haplotype and microsatellite genotype (up to 17 loci), for each sample, following previously described methodology (for more information on sex ID and mtDNA haplotyping methods see Carroll et al. [23] and on microsatellite genotyping see Supplementary Material 1 of Carroll et al.[3]). Each tray of samples was run with between four and seven control samples (New Zealand and/or Australian samples previously genotyped) to ensure consistent binning of alleles and a negative to detect contamination. In previous work, no significant deviations from the Hardy Weinberg Equilibrium were detected using these loci, such as null alleles[1,23]. CERVUS[24] was used to identify and calculate the probability of identity[25] for replicate samples within and between sampling locations.

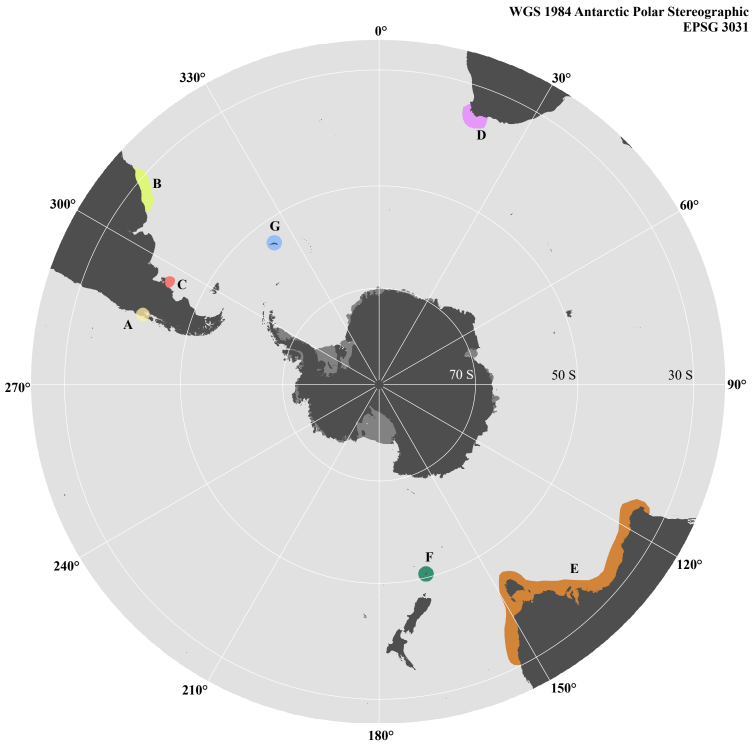

Figure 1: Location of southern right whale sampling sites including wintering grounds A. Chile; B. Brazil; C. Argentina; D. South Africa; E. Australia; F. New Zealand subantarctic and summer feeding ground G. South Georgia.

Estimating genetic diversity and differentiation

Genetic diversity was calculated for all sampling locations except for Chile, due to the single sample from this area. For the microsatellites, we estimated the actual and effective number of alleles and observed and expected heterozygosity using the programme Genodive[26]. For the mtDNA data, we report the number of haplotypes, haplotype diversity and nucleotide diversity, calculated using Arlequin v3.5[27]. To visualise the place of the mtDNA haplotypes found in SG and Chile in the broader dataset, we created a median joining haplotype network[28] using popart[29].

We first compared the SG dataset to samples grouped by ocean basin (Indo-Pacific, comprising Australia and New Zealand, and South Atlantic, comprising Argentina, Brazil and South Africa). The SG samples were significantly different from the Indo-Pacific samples, so we continued analyses using the South Atlantic samples grouped by wintering ground. For microsatellites, differentiation was estimated using F_{ST} calculated in Genepop[30] and Jost's D[31] calculated in Genodive. Significance was assessed using the exact G test in Genepop with 100K dememorizations, 1K batches and 10K iterations per batch. For mtDNA, differentiation was estimated using F_{ST} and ϕ_{ST} calculated in Arlequin. Significance was assessed using permutation tests and, given the small sample size from SG, the exact test of differentiation (1,000,000 Markov chain steps; 1,000,000 dememorization steps, with significance set at $\alpha = 0.05$).

Population structure and assignment of samples

We conducted two distinct population structure and assignment analyses to investigate how the Chile-Peru and SG samples sat within the broader migratory network. The first was using the program STRUCTURE[32], which clusters samples into user-defined k populations and estimates the proportion of each sample's genotype assigning to each cluster. For the STRUCTURE analysis, we started by using all data (published and newly generated) and running the analysis in admixture mode with and without location prior (sampling location). Ten replicates of $k = 1$ to 5 were conducted each with burn-ins of one million iterations and runs of ten million Markov chain Monte Carlo

(MCMC) iterations, and the convergence was assessed by visually inspecting the summary statistics (e.g., F_{ST}). Where structure was detected, the dataset was split and reanalysed to detect any further substructure. We assessed the most likely value of k using the mean log likelihood from across the ten runs summarised with STRUCTURE HARVESTER[33], as well as the method of Evanno et al[34]. We used CLUMPAK[35] to summarise the modes or distinct solutions for each value of k.

We also used the population assignment program GENEPLOT[36] to determine the ability of the data to assign SG samples to reference wintering ground datasets. Given the low level of differentiation seen between wintering grounds, this population assignment method is preferable to a mixed stock analysis (MSA) such as ONCOR[37]. This is because when F_{ST} is low, MSA analyses are biased towards estimating $1/k$, where k is the number of reference populations (ONCOR manual).

GENEPLOT takes the concept and statistical methodology used in GENECLASS2[38] and produces a graphical output that visualises the fit of individuals to reference populations, as well as the genetic variance within these reference populations. GENEPLOT first calculates the allele frequencies for two reference populations, A and B, using the Bayesian method of Rannala and Mountain[39] with the priors of Rannala and Mountain[39] or Baudouin and Lebrun[40] (we chose the latter). The method allows the posterior allele frequencies to be estimated using leave-one-out for individuals from the reference population samples. GENEPLOT then calculates the fit of the reference samples, and the fit of the query samples, into populations A and B by calculating the log genotype probability (LGP) of the individual's genotype arising in each population using the method of Piry et al.[38], approximated for individuals with missing data using the method of McMillan and Fewster[36]. If an individual has a high LGP for population A, it indicates that the genotype of that individual has a high probability of arising from that population, though the individual may have high LGP for more than one population.

The program also calculates the correct assignment probability, which is the probability that a random genotype arising from reference population A will be correctly assigned to A rather than reference population B. This is estimated for all potential genotypes that could arise from population A using the allele frequencies of reference population A. The LGP for these potential genotypes is calculated for both reference populations A and B. The differences in LGP between A and B for each potential genotype are distributed according to the probability of the genotype arising in reference population A, giving a probability density function. The correct assignment probability is the percentage of genotypes arising from A that have a higher LGP with respect to A than to B.

We undertook comparisons of the South Atlantic wintering ground datasets to determine how distinct they are using GENEPLOT, by visually inspecting the displays and calculating pairwise correct assignment probabilities. We then considered the wintering grounds (Argentina, Brazil, combined into Southwest Atlantic and South Africa) reference populations and investigated whether the SG samples were able to be conclusively assigned to these references.

RESULTS

Genotyping success for Chilean and SG southern right whale samples

During the genotyping run for the two sets of samples, one locus failed across all samples (RW31). All three SG samples collected during the 2018 field season successfully amplified (average of 15 of 16 loci) and produced mtDNA and genetically identified sex. Of the 12 samples collected during the 1997 field season, 11 produced DNA of sufficient quality and quantity to genotype. Of these, 8 amplified at a minimum of 11 out of 16 loci (the QC threshold) and were used in subsequent analyses (average of 14.6 loci). At the time of writing this report, the mtDNA sequences of the samples had not been generated and so previously published sequences were used (see below). Therefore a total 11 samples from SG were available for analyses.

Genotyping success for Brazilian and South African southern right whale samples

Of the 60 Brazilian samples, 59 produced genotypes that passed QC and comprised an average of 16.6 loci. Comparison of the genotypes showed that 59 genotypes represented 50 unique individuals. Of the 88 South African samples, 86 produced genotypes that passed QC and comprised an average of 15.5 loci. Comparison of the genotypes showed that the 86 genotypes represented 77 unique individuals.

Comparison of genotypes across all sampling locations revealed one potential match between Argentina and Brazil. This sample matched at 16 of 17 loci and genetically identified sex, but the mtDNA haplotype was different. Effort is underway (1) to check the mtDNA haplotype because the probability of identity (8.0E-20) gives high confidence in the match, and (2) to determine whether there is any additional information about this female. The sample was retained in both Argentinean and Brazilian datasets for wintering-ground based comparisons.

Genetic diversity

All South Atlantic wintering grounds have very similar levels of genetic diversity (Table 1) for both mtDNA and microsatellite data. Chile was not included in this analysis given the sample size of 1.

Table 1: Summary of microsatellite and mtDNA diversity statistics for southern right whale winter nursery/socialising (W) and summer feeding (S) grounds. Sample size (2N), number of alleles ($\mathrm{n_{alleles}}$), effective number of alleles ($\mathrm{Effn_{alleles}}$), observed ($\mathrm{H_{Obs}}$) and expected ($\mathrm{H_{Exp}}$) heterozygosity is reported for the microsatellite loci used in the analysis. Sample size (N), number of haplotypes ($\mathrm{n_{hap}}$), haplotype diversity (*h*) and nucleotide diversity (*n*) reported for the 275 bp fragment of the mtDNA control region analysed. South Africa, Argentina and Brazil are pooled for the South Atlantic dataset.

	2N	$\mathrm{n_{alleles}}$ (±SD)	$\mathrm{Effn_{alleles}}$ (+/-SD)	$\mathrm{H_{Obs}}$ (±SD)	$\mathrm{H_{Exp}}$ (±SD)	N	$\mathrm{n_{hap}}$	*h*±*SD*	*N* (%) ± SD	Source
South Africa-W	248	10.94±1.42	5.18±0.49	0.74±0.03	0.77±0.03	416	38	0.94±0.01	3.0±1.5	Carroll et al. (2018); this study
Argentina-W	92	8.23±0.84	4.73±0.47	0.71±0.04	0.76±0.03	208	27	0.94±0.01	2.8±1.5	Carroll et al. (2018)
Brazil-W	100	9.35±1.11	4.91±0.47	0.75±0.04	0.77±0.03	50	21	0.94±0.02	3.3±1.8	This study
South Atlantic-W	440	11.94±1.65	4.93±0.49	0.73±0.04	0.77±0.03	674	55	0.96±0.02	3.1±1.5	Carroll et al. (2018); this study
SG-S	22	6.50±0.68	4.41±0.51	0.74±0.07	0.77±0.06	11	9	0.95±0.07	2.7±1.6	This study
Indo-Pacific-W	258	9.44±0.86	4.83±0.38	0.76±0.03	0.78±0.02	769	12	0.71±0.01	1.9±1.1	Carroll et al. (2015)

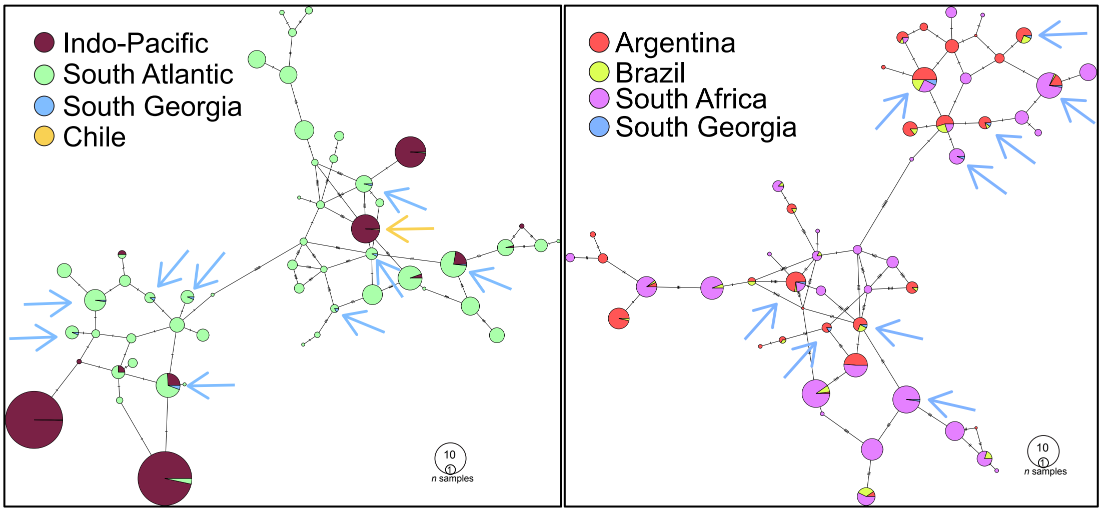

Figure 2: Haplotype network of southern right whale mtDNA control region (275 bp) showing; Left panel: position of Chilean sample (indicated by orange arrow) and South Georgia/Islas Georgias del Sur samples (blue arrows) in the broader southern right whale haplotype network; Right panel: position of SG samples (highlighted by blue arrows) relative to haplotypes found in the South Atlantic wintering ground.

Chilean sample findings

The Chilean sample had a mtDNA haplotype only previously found in the Indo-Pacific (BakHapC; see Figure 2). The Chilean sample also appeared to be admixed between the Indo-Pacific and South Atlantic populations, with a high proportion from the Indo-Pacific, based on the STRUCTURE analysis with and without location prior settings (See Figure 3 and Figures S1 and S2 in the Supplementary Material). The Chilean sample also had a strong fit (high LGP) to both oceans using GENEPLOT (Figure 4). The correct assignment probabilities for the South Atlantic and Indo-Pacific are 0.86 and 0.836, respectively.

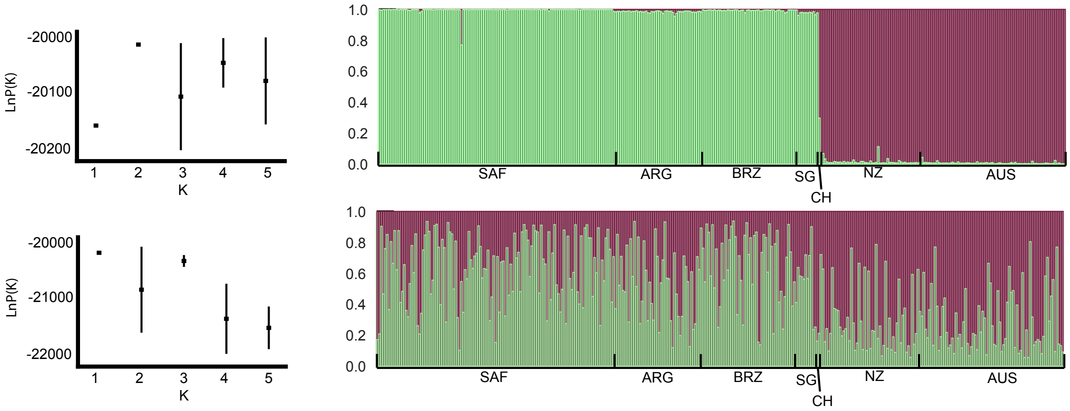

Figure 3: Mean log likelihood (LnP(K)), for K = 1 to 5 (left), and the proportion of each individual's genome that is assigned to each cluster when K = 2 (right) for location prior with admixture setting (top) and standard admixture setting (below) implemented. SAF = South Africa; ARG = Argentina; BRZ = Brazil; SG = South Georgia; CH = Chile; NZ = New Zealand; AUS = Australia.

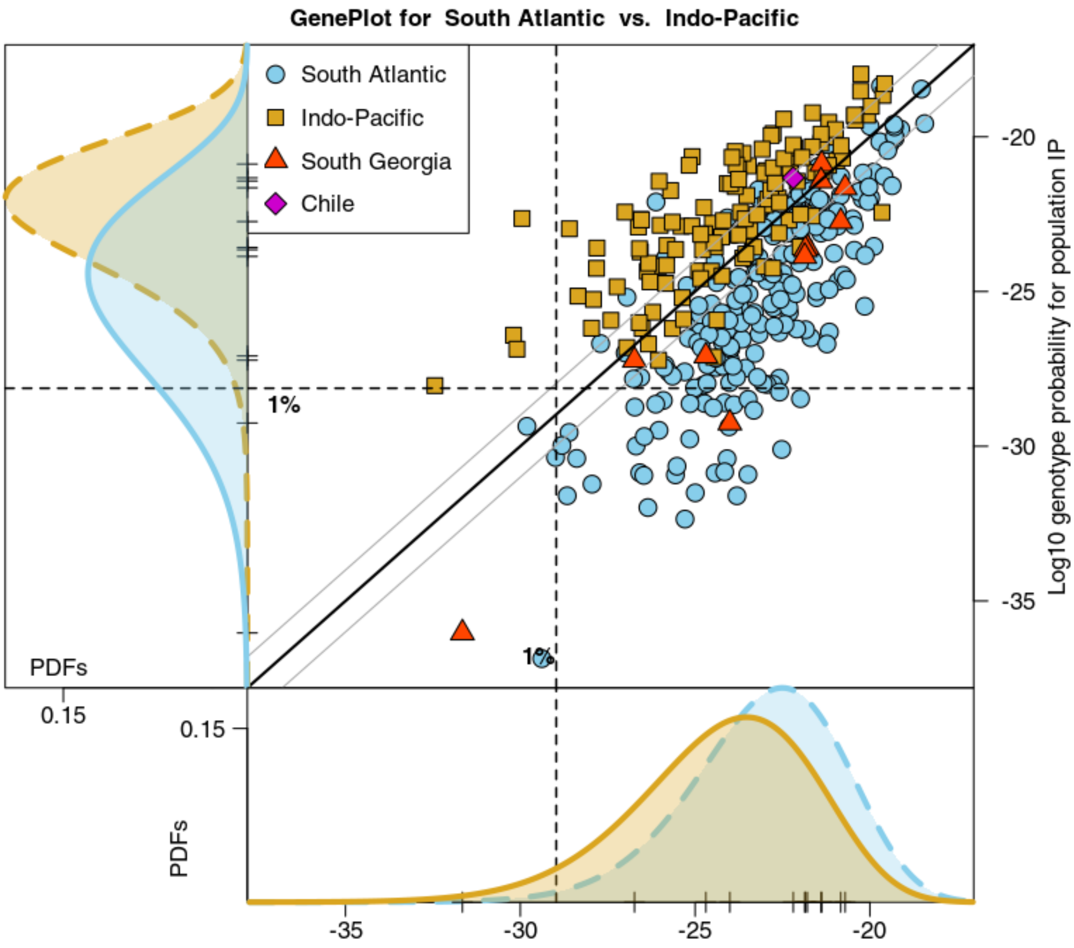

Figure 4: Geneplot showing log genotype probabilities (LGP) for South Atlantic and Indo-Pacific reference and SG and Chilean query southern right whale microsatellite genotypes. The thick centre diagonal line indicates an equal posterior probability for both reference populations; a point lying on that line has the same LGP with respect to both reference populations. The thin diagonal lines indicate where the genotype probability (the inverse-log of the LGP) for one population is 10 times greater than it is for the other population. Also shown are the 1% quantiles; only 1% of all possible genotypes that could arise from the given population would have LGPs below this line. The bottom graph shows the fit of possible South Atlantic (dashed line) and Indo-Pacific (solid line) individuals with respect to the South Atlantic population; the left graph shows the fit of possible Indo-Pacific (dashed line) and South Atlantic (solid line) individuals with respect to the Indo-Pacific population.

South Georgia/Islas Georgias del Sur sample findings

Unsurprisingly, the SG samples are clearly South Atlantic (Argentina, Brazil and South Africa) in origin, based on the mtDNA and microsatellite indices of differentiation (Table 2). Of the 9 mtDNA haplotypes found in SG based on the 275 bp fragment, three were found in all three South Atlantic wintering grounds, four were shared with only the Southwest Atlantic wintering grounds, and two were shared with only the South African wintering ground (Figure 2).

The SG feeding ground appears more genetically similar to the Argentina and Brazilian wintering grounds, as there was no significant genetic differentiation detected (Table 3, $p>0.05$). In contrast, there was significant difference in both mtDNA ($\phi_{ST} = 0.088$, $F_{ST} = 0.018$) and microsatellites (Jost's D = 0.025, $F_{ST} = 0.007$) between SG and South Africa.

Table 2: Estimates of differentiation between the Indo-Pacific and South Atlantic wintering grounds and the SG feeding ground, with mtDNA table showing ϕ_{ST} (top right), F_{ST} (bottom left quadrant) and microsatellites showing Jost's D (top right) and F_{ST} (bottom left quadrant).

A. mtDNA	South Atlantic	SG	Indo-Pacific
South Atlantic		0.036	0.203***§
SG	0.001		0.082***§
Indo-Pacific	0.159***§	0.201***§	
B. microsatellites	South Atlantic	SG	Indo-Pacific
South Atlantic		0.015	0.030***
SG	0.005		0.042***
Indo-Pacific	0.010 ***	0.012***	

significance: *** p<0.001; **p<0.01; *p<0.05 from permutation test; § p<0.001 from exact test of differentiation for mtDNA data

Table 3: Pairwise differentiation indices with mtDNA table showing ϕ_{ST} (top right), F_{ST} (bottom left quadrant) and microsatellites showing Jost's D (top right) and F_{ST} (bottom left quadrant).

A. mtDNA	Argentina	Brazil	South Africa	SG
Argentina		0.031*§	0.087***§	0.000
Brazil	0.023***§		0.022*§	0.010
South Africa	0.048***§	0.034***§		0.088***§
SG	0.007	0.001	0.029*§	
B. microsatellites	Argentina	Brazil	South Africa	SG
Argentina		0.047**	0.008***	0.002
Brazil	0.001**		0.011***	0.006
South Africa	0.003***	0.003***		0.025*
SG	0.004	0.001	0.007*	

significance: *** p<0.001; **p<0.01; *p<0.05 from permutation test; § p<0.001 from exact test of differentiation for mtDNA data

Overall, the STRUCTURE analysis again supported previous findings of weak differentiation between ocean basins (see Supplementary Material). The expanded dataset did suggest that southern right whales had three ancestral populations under the standard admixture setting, compared with two in the previous analyses. However, the South Atlantic and Indo-Pacific show different patterns of admixture of these ancestral populations, such that when k = 3, the samples within ocean basin have similar admixture patterns (Figure S1). Therefore, k = 2 is presented in Figure 3, and k = 3 in Supplementary Material (Figures S1 and S2). Under the location prior setting, the best fitting k = 2 for both methods, closely followed by 4. As seen previously, k = 2 split the samples into ocean basins. Under k = 4, the South Atlantic samples were admixed primarily between three ancestral populations, and southwest and southeast Atlantic samples showed different admixture proportions (Figure S2). Analyses with only the South Atlantic samples did not find evidence of substructure.

Furthermore, the GENEPLOT analyses showed that while there is some genetic distinctiveness amongst the South Atlantic wintering grounds, there is not sufficient resolution with which to assign the SG samples with confidence. This is shown visually in Figures 5 and 6. The correct assignment probabilities calculated on a pairwise manner ranged from a modest 0.75, for the probability that a South African

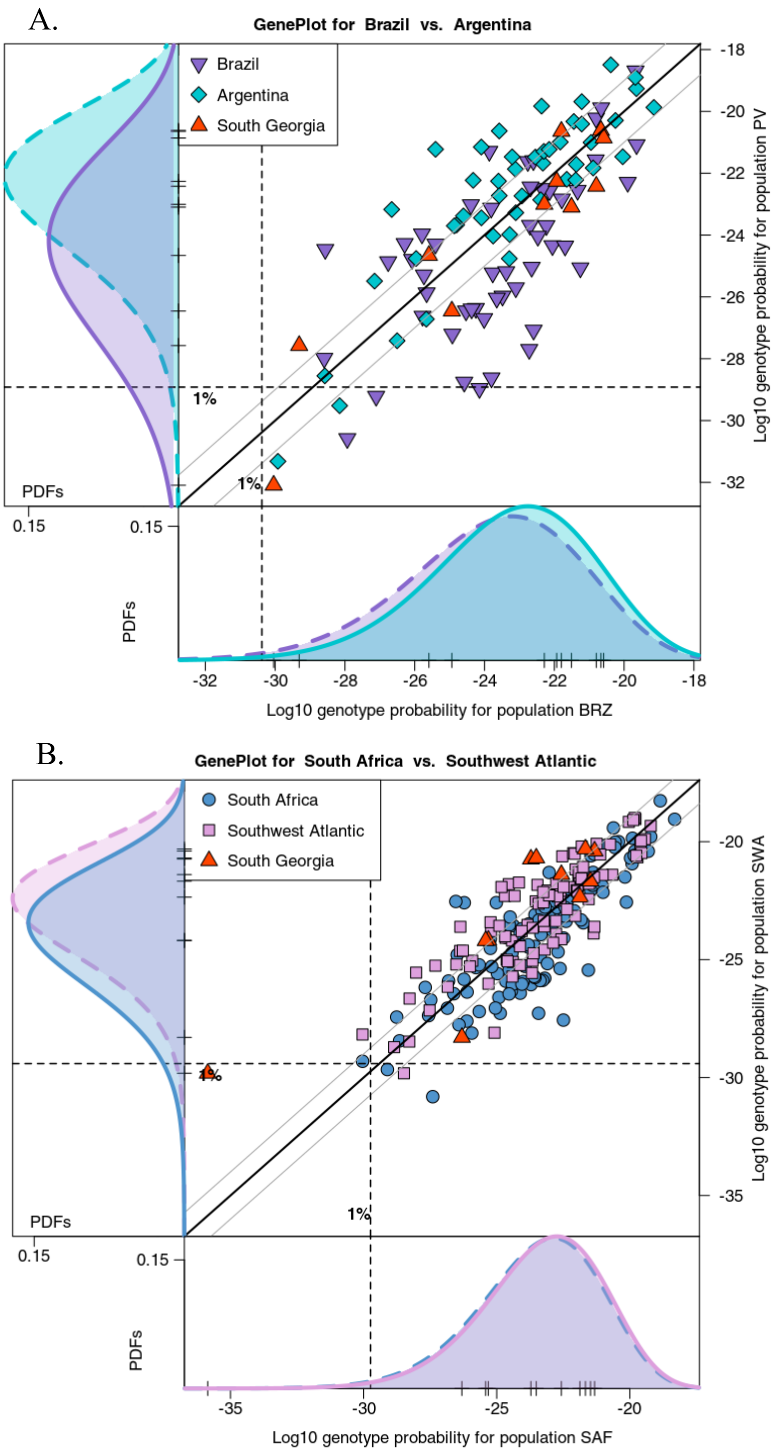

Figure 5: Geneplot showing log genotype probabilities (LGP) for A. Brazilian and Argentinean and B. Southwest Atlantic (Brazil+Argentina) and South African southern right whale microsatellite genotypes. The thick centre diagonal line indicates an equal posterior probability for both reference populations; a point lying on that line has the same LGP with respect to both reference populations. The thin diagonal lines indicate where the genotype probability (the inverse-log of the LGP) for one population is 10 times greater than it is for the other population. Also shown are the 1% quantiles, showing the lowest 1% of the LGP distribution for the corresponding population. The shaded curves left of and below the central plot show the overall distributions of the baseline populations (dashed lines) and the overall fit of the other population into the baseline (solid line).

sample would be a better fit to South Africa than Brazil, to only 0.51 for the probability that a Brazilian sample would be a better fit to Brazil than South Africa (Table S1).

Additionally, all the SG samples had high LGPs for Brazil, Argentina and South Africa, meaning the sampled genotypes have a high probability of arising in any of these regions. The exception was sample 6B, which had a poor fit to all wintering grounds, and in fact, to both the Indo-Pacific and South Atlantic groupings. This sample had rare alleles at 5 loci and will be re-genotyped to investigate if this is an error or simply a whale with rare alleles.

DISCUSSION

Chile-Peru: a stepping stone between the Indo-Pacific and South Atlantic?

A sample size of one means that we are unable to make any firm statements about the genetic identity of the Chile-Peru right whale wintering grounds. However, the findings presented here suggest hypotheses that could be tested with additional data. The single microsatellite genotype appears to be admixed between the South Atlantic and Indo-Pacific, whilst having a maternal lineage (BakHapC) that is only previously found in the Indo-Pacific. This leads to the hypothesis that the Chile-Peru stock could have historically been a stepping stone between the South Atlantic and Indo-Pacific. Alternatively, the genetic characteristics seen today could be the result of demographic processes and resulting genetic drift subsequent to the whaling era. The limited data show no trend in the number sightings or individuals in the Chile-Peru stock over time, although there has been a small increase in the number of calves between 1975 and 2010[9,10], highlighting the importance of continuing to monitor the region and collect additional samples as the species recovers from whaling.

A weakness of this analysis is that the correct assignment probabilities for the ocean basins relative to each other were 0.834 for South Atlantic and 0.861 for Indo-Pacific. Therefore, there is a chance that the Chile-Peru sample could originate from either ocean basin and simply not be strongly assigned using GENEPLOT. Interestingly, the STRUCTURE location prior analysis that could partition samples into ocean basins still found the Chile-Peru sample admixed.

Additional analyses are underway for the Chile-Peru stock sample, including stable isotopes, hormone analyses and additional genomic work.

South Georgia/Islas Georgias del Sur: more Southwest than Southeast Atlantic?

The indices of genetic differentiation suggest that the SG feeding ground is genetically closer to the Southwest Atlantic wintering grounds (Argentina and Brazil) than the South African wintering ground. This is consistent with several lines of existing evidence. Firstly, previous analyses of 10 microsatellite loci from the samples collected in SG in 1997 with samples from the Southwest Atlantic wintering grounds did not find significant differentiation[15]. Secondly, there have been photo-ID matches between SG and Peninsula Valdes[7,12]. In contrast, comparison of photos collected during the SG surveys of 1997 and 2018 and the South African wintering ground catalogue did not produce any matches[12] (J. Jackson et al. in prep). Finally, satellite tagging studies have shown movement from Peninsula Valdes, Argentina, but not from South Africa, to SG[5,14]. Southern right whales wintering in South Africa appear to feed to the east (20°W - 20°E), including Bouvet Island[5], and so may not visit SG.

However, there is also evidence to suggest that there may be connectivity between SG and South Africa. Firstly, there were two samples from SG with mtDNA haplotypes only found in South Africa. Additionally, the SG samples typically had high genotype probabilities in all of the South Atlantic wintering grounds, based on the GENEPLOT analysis. It may be that using additional loci could increase the power of the GENEPLOT analysis, and work is underway to see if the available ddRAD data (E. Carroll, O. Gaggiotti and colleagues unpublished) improve the correct assignment probabilities observed in this study.

ACKNOWLEDGEMENTS

This research was funded by the EU BEST 2.0 medium grant 1594 and UK DARWIN DPLUS grant 057. For the Chilean sample, the authors thank Jose Aviles for his contribution to the sighting, reporting of the stranding, and the necropsy work; the National Fisheries and Aquaculture Service for assistance with the necropsy: Antonio Palma, Mauricio Ulloa, Marisol Romero and especially Hans Ossvald; the Southern Right Whale Health Monitoring Program in Argentina for necropsy protocols and the Global Greengrants Fund and the Pacific Whale Foundation for their financial support. The collection of the Brazilian samples was supported through grants by the Brazilian National Research Council to Paulo H. Ott (CNPq proc. n° 144064/98-7) and Paulo A.C. Flores (CNPq proc. n° 146609/1999-9) and with support from the World Wildlife Fund **(**WWF-Brazil).

REFERENCES

1. Carroll, E. L. *et al.* Incorporating non-equilibrium dynamics into demographic history inferences of a migratory marine species. *Heredity (Edinb).* **122**, 53–68 (2019).
2. Valenzuela, L. O., Sironi, M., Rowntree, V. J. & Seger, J. Isotopic and genetic evidence for culturally inherited site fidelity to feeding grounds in southern right whales (*Eubalaena australis*). *Mol. Ecol.* **18**, 782–791 (2009).
3. Carroll, E. L. *et al.* Cultural traditions across a migratory network shape the genetic structure of southern right whales around Australia and New Zealand. *Sci. Rep.* **5**, 16182 (2015).
4. Rowntree, V. J., Payne, R. & Schell, D. Changing patterns of habitat use by southern right whales (*Eubalaena australis*) on their nursery ground at Península Valdés, Argentina, and in their long-range movements. *J. Cetacean Res. Manag. Spec. Issue* **2**, 133–143 (2001).
5. Mate, B., Best, P., Lagerquist, B. & Winsor, M. Coastal, offshore and migratory movements of South African right whales revealed by satellite telemetry. *Mar. Mammal Sci.* **27**, 455–476 (2011).
6. Patenaude, N. J. *et al.* Mitochondrial DNA diversity and population structure among southern right whales (*Eubalaena australis*). *J. Hered.* **98**, 147–157 (2007).
7. Best, P., Payne, R., Rowntree, V. J., Palazzo, J. & Both, M. Long-range movements of South Atlantic right whales *Eubalaena australis*. *Mar. Mammal Sci.* **9**, 227–234 (1993).
8. Reilly, S. B. *et al.* Eubalaena australis. *IUCN Red List Threat. Species. Available http//dx.doi.org/10.2305/IUCN.UK.2013-1.RLTS.T8153A44230386.en.* (2013).

9. Galletti Vernazzani, B., Arroyo, P., Goya, E. & Palma, A. Revised conservation management plan for eastern South Pacific southern right whale population (*Eubalaena australis*) - submitted by Chile and Peru April 2016. Report SC/66b/BRG/23 submitted to the Scientific Committee of the International Whaling,. (2016).
10. Galletti Vernazzani, B., Cabrera, E. & Brownell, R. L. Eastern South Pacific southern right whale photo-identification catalog reveals behavior and habitat use patterns. *Mar. Mammal Sci.* **30**, 389–398 (2014).
11. Carroll, E. L. *et al.* Accounting for female reproductive cycles in a superpopulation capture – recapture framework. *Ecol. Appl.* **23**, 1677–1690 (2013).
12. Moore, M. J. *et al.* Relative abundance of large whales around South Georgia (1979-1998). *Mar. Mammal Sci.* **15**, 1287–1302 (1999).
13. Richardson, J., Wood, A., Neil, a, Nowacek, D. & Moore, M. Changes in large whale distribution, relative abundance, and species composition around South Georgia from opportunistic sightings: 1992–2011. *Endanger. Species Res.* **19**, 149–156 (2012).
14. Zerbini, A. N. *et al.* Tracking southern right whales through the southwest Atlantic: new insights into migratory routes and feeding grounds. Report SC/66b/BRG26 presented to the Scientific Committee of the International Whaling Commission, Cambridge, UK. (2015).
15. Ott, P. H., Flores, P. A. C., Freitas, T. R. O. & White, B. N. Genetic diversity and population structure of southern right whales, *Eubalaena australis*, from the Atlantic coast of South America. Report SC/12/RW25 presented to the Scientific Committee of the International Whaling Commission, Cambridge, UK. (2011).
16. Ott, P. H. Diversidade genética e estrutura populacional de duas espécies de cetáceos do Atlântico Sul Ocidental: *Pontoporia blainvillei* e *Eubalaena australis*. Ph.D. dissertation, Universidade Federal do Rio Grande do Sul, Brazil. (2002).
17. Galletti Vernazzani, B., Chirife, A., Cabrera, E., Sironi, M. & Brownell Jnr, R. L. Entanglement and death of a critically endangered eastern South Pacific southern right whale (*Eubalaena australis*) in Chile. Report SC/67A/HIM/14 presented to the Scientific Committee of the International Whaling Commission, Cambridge, UK. (2017).
18. Jackson, J. A. *et al.* Sightings and acoustic records of right whales collected in South Georgia (Islas Georgias del Sur) waters January- February 2018. Report SC/67B/SH/20 submitted to the Scientific Committee of the International Whaling Commission, Cambridge, UK. (2018).
19. Krützen, M. *et al.* A biopsy system for small cetaceans; darting success and wound healing in Tursiops spp. *Mar. Mammal Sci.* **18**, 863–878 (2002).
20. Lambertsen, R. A biopsy system for large whales and its use for cytogenetics. *J. Mammal.* **68**, 443–445 (1987).
21. Brown, M., Kraus, S. & Gaskin, D. E. *Reaction of North Atlantic right whales (*Eubalaena glacialis*) to skin biopsy sampling for genetic and pollutant analysis. Report of the International Whaling Commission Special Issue 13:81-90.* (1991).
22. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual 2nd ed.* (Cold Spring Harbor Laboratory Press, 1989).
23. Carroll, E. L. *et al.* Population structure and individual movement of southern

right whales around New Zealand and Australia. *Mar. Ecol. Prog. Ser.* **432**, 257–268 (2011).

24. Kalinowski, S., Taper, M. & Marshall, T. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **9**, 801–888 (2007).
25. Waits, L. P., Luikart, G. & Taberlet, P. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Mol. Ecol.* **10**, 249–256 (2001).
26. Meirmans, P. & van Tienderen, P. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes* **4**, 792–794 (2004).
27. Excoffier, L. & Lischer, H. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
28. Bandelt, H. J., Forster, P. & Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48 (1999).
29. Leigh, J. W. & Bryant, D. Popart: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116 (2015).
30. Rousset, F. Genepop'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol. Ecol. Resour.* **8**, 103–106 (2008).
31. Jost, L. GST and its relatives do not measure differentiation. *Mol. Ecol.* **17**, 4015–4026 (2008).
32. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
33. Earl, D. A. & VonHoldt, B. M. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361 (2012).
34. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).
35. Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. & Mayrose, I. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **15**, 1179–1191 (2015).
36. McMillan, L. F. & Fewster, R. M. Visualizations for genetic assignment analyses using the saddlepoint approximation method. *Biometrics* **73**, 1029–1041 (2017).
37. Kalinowski, S., Manlove, K. & Taper, M. *ONCOR: a computer program for genetic stock identification, v2 http://www.montana.edu/kalinowski/Software/ONCOR. html. Department of Ecology, Montana State University, Bozeman, USA*. (2008).
38. Piry, S. *et al.* GENECLASS2: A software for genetic assignment and first-generation migrant detection. *J. Hered.* **95**, 536–539 (2004).
39. Rannala, B. & Mountain, J. Detecting immigration by using multilocus genotypes. *Proc. Natl. Acad. Sci.* **94**, 9197–9201 (1997).
40. Baudouin, L. & Lebrun, P. An operational Bayesian approach for the identification of sexually reproduced cross-fertilized populations using molecular markers. *ISHS Acta Hortic.* **546**, 81–93 (2001).

SUPPLEMENTARY MATERIAL

Table S1: Estimates of correct assignment probability for pairwise comparisons of South Atlantic wintering grounds based on 17 microsatellite loci and calculated using GENEPLOT. This is the probability that a random genotype from population A will be correctly assigned to population A rather than to population B.

Population A	Population B	Correct assignment probability
Argentina	Brazil	0.66
Argentina	South Africa	0.67
Brazil	Argentina	0.77
Brazil	South Africa	0.51
South Africa	Argentina	0.50
South Africa	Brazil	0.75
South Africa	Southwest Atlantic	0.66
Southwest Atlantic	South Africa	0.61

Supplementary Material 1: Details of STRUCTURE results

Table S2: Best fitting *k* inferred from delta *k* method of Evanno et al. and using mean log likelihood directly from STRUCTURE output. *k* infers the number of ancestral clusters, whereas the observed division highlights how the contemporary samples group given a value of *k*.

Model	Delta *k*	Mean log likelihood	Observed division (*k*)
Admixture without location prior	1 (3 second)	3	(3) Ocean Basin
Admixture with location Prior	2 (4 second)	2 (4 second)	Ocean Basin (2); Southwest and Southeast Atlantic different (4)

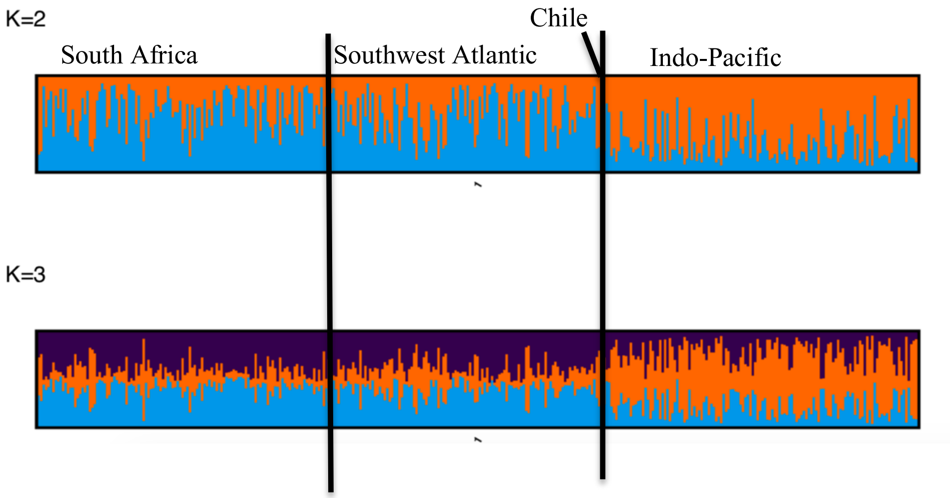

Figure S1: Individual admixture proportions from STRUCTURE analysis using standard admixture settings, showing most common mode identified by CLUMPAK. The *k* = 2 (in Figure 3 with more detail) and *k* = 3 figures show admixture proportion varies by ocean basin.

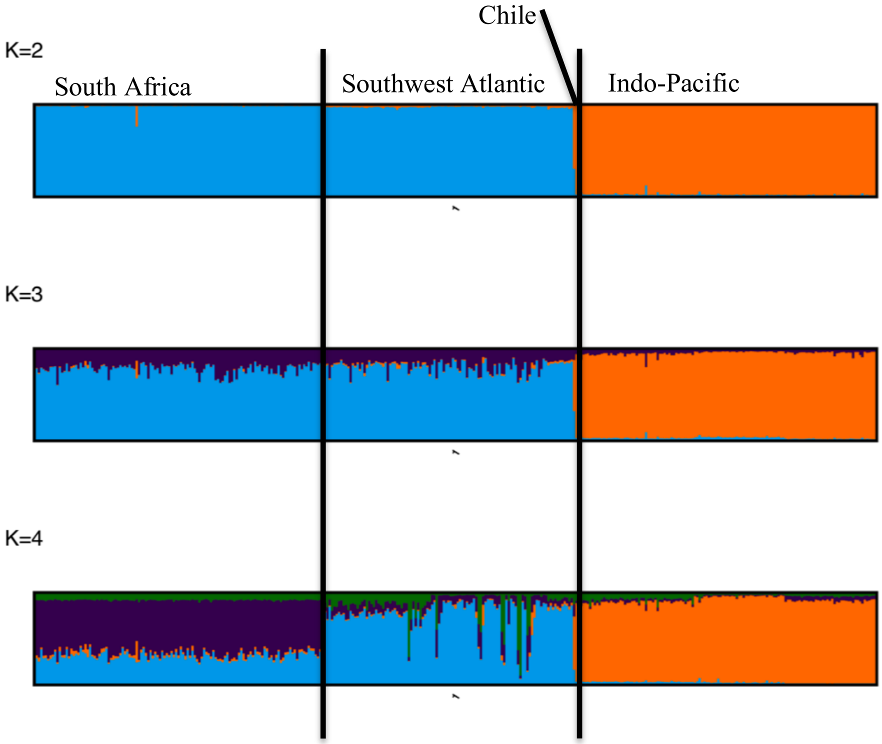

Figure S2: Individual admixture proportions from STRUCTURE analysis using location prior settings, showing most common mode identified by CLUMPAK. The *k* = 2 (in Figure 3 with more detail) and *k* = 3 figures show admixture proportion varies by ocean basin (orange colour shows Indo-Pacific samples), whereas *k* = 4 suggests some division within the South Atlantic.