Marine Mammal Science

MARINE MAMMAL SCIENCE, 33(2): 457–479 (April 2017)

DOI: 10.1111/mms.12378

Genetic differentiation between humpback whales (*Megaptera novaeangliae*) from Atlantic and Pacific breeding grounds of South America

Ana Lúcia Cypriano-Souza,[1] Faculdade de Biociências, Pontifícia Universidade Católica do Rio Grande do Sul, Avenida Ipiranga 6681, 90619-900 Porto Alegre, Rio Grande do Sul, Brazil and Projeto Baleia Jubarte/Instituto Baleia Jubarte, Rua Barão do Rio Branco 125, 45900-000 Caravelas, Bahia, Brazil; **Márcia H. Engel,** Projeto Baleia Jubarte/Instituto Baleia Jubarte, Rua Barão do Rio Branco 125, 45900-000 Caravelas, Bahia, Brazil; **Susana Caballero,** Laboratorio de Ecología Molecular de Vertebrados Acuáticos-LEMVA, Departamento de Ciencias Biológicas, Universidad de los Andes, Carrera 1 No. 18A-12, Bogotá, Colombia; **Carlos Olavarría,** School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1010, New Zealand, Centro de Estudios del Cuaternario (CEQUA), Casilla 737, Punta Arenas, Chile, and Centro de Estudios Avanzados en Zonas Aridas (CEAZA), Raúl Bitrán 1305, La Serena, Chile; **Lilián Flórez-González** and **Juan Capella,** Fundación Yubarta, Carrera 24F Oeste No. 3-110, Tejares de San Fernando, Cali, Colombia; **Debbie Steel** and **Angie Sremba,** Marine Mammal Institute, Department of Fisheries and Wildlife, Hatfield Marine Science Center, Oregon State University, 2030 SE Marine Science Drive, Newport, Oregon 97365, U.S.A.; **Anelio Aguayo,** Departamento Científico, Instituto Antártico Chileno, Plaza B. Muñoz Gamero 1055, Punta Arenas, Chile; **Deborah Thiele,** Fenner School of Environment and Society, Australian National University, Linnaeus Way, Canberra, ACT 0200, Australia; **C. Scott Baker,** School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1010, New Zealand and Marine Mammal Institute, Department of Fisheries and Wildlife, Hatfield Marine Science Center, Oregon State University, 2030 SE Marine Science Drive, Newport, Oregon 97365, U.S.A.; **Sandro L. Bonatto,** Faculdade de Biociências, Pontifícia Universidade Católica do Rio Grande do Sul, Avenida Ipiranga 6681, 90619-900 Porto Alegre, Rio Grande do Sul, Brazil.

Abstract

Humpback whales wintering in tropical waters along the Atlantic and Pacific coasts of the South American continent are thought to represent distinct populations or "stocks." Here we present the first analysis of genetic differentiation and estimates of gene flow between these breeding stocks, based on both mitochondrial DNA (mtDNA) control region sequences (465 bp) and 16 microsatellite loci from samples collected off Brazil (n = 277) and Colombia (n = 148), as well as feeding areas near the western Antarctic Peninsula (n = 86). We found significant differentiation between Brazilian and Colombian breeding grounds at both mtDNA (F_{ST} = 0.058) and microsatellite (F_{ST} = 0.011) markers and corroborated previous studies showing genetic similarity between humpbacks from Colombia and those from Antarctic Peninsula feeding areas. Estimates of long-term gene flow between Brazil and

Colombia were low to moderate, asymmetrical, and mostly mediated by males. Assignment procedures detected some cases of interchange and individuals of admixed ancestry between breeding grounds, indicating limited mixing of individuals between these stocks. Overall, results highlight the differentiation of humpback whale breeding populations with adjacent feeding grounds. This appears to be a remarkable example of fidelity to seasonal habitat in the absence of any contemporary barriers.

Key words: *Megaptera novaeangliae*, population genetic structure, microsatellites, mtDNA, migration, individual assignment.

Humpback whales (*Megaptera novaeangliae* Borowski, 1781) are widely distributed throughout the major ocean basins, where they undertake long-distance seasonal migrations between high-latitude summer feeding grounds and low-latitude winter breeding grounds (Dawbin 1966). Dispersal or interbreeding of whales from different ocean basins is rare, resulting in highly significant genetic differentiation between the populations from these basins (Jackson *et al*. 2014). In the Northern Hemisphere, humpback whales from the Atlantic and Pacific Oceans show strong maternal fidelity to segregated feeding grounds and natal philopatry to breeding grounds (Palsbøll *et al*. 1995, Stevick *et al*. 2006*a*, Wenzel *et al.* 2009, Baker *et al.* 2013). In contrast, the feeding grounds in the Southern Hemisphere are distributed mainly throughout a broad circumpolar area in the Southern Ocean (Matthews 1937), providing the potential for longitudinal movements of individuals from different breeding populations during the feeding season. These movements may increase the possibility of interchange among these populations if the individuals switch between the breeding grounds. The International Whaling Commission (IWC) recognizes six humpback whale management feeding areas (I–VI) (Donovan 1991), seven breeding stocks (A–G), and several substocks in the Southern Hemisphere (IWC 2015).

The humpback whale breeding grounds off the eastern and western coasts of South America are considered breeding stocks A (BSA) and G (BSG), respectively (Fig. 1). In the southwestern Atlantic Ocean, whales from BSA are mainly found along the Brazilian coast from approximately 5°S to 23°S, with the Abrolhos Bank (16°40'–19°30'S and 37°25'–39°45'W) being the main mating and calving area for this population (Andriolo *et al*. 2010). At this breeding ground, humpback whales have also been observed near oceanic islands, such as the Fernando de Noronha Archipelago and Trindade Island (Lodi 1994, Wedekin *et al*. 2014). In the eastern Pacific Ocean, humpback whales from BSG occur from the northern coast of Peru (5°S) to Costa Rica (12°N), with a very low density of whales found around the Galápagos Archipelago, located 1,000 km west of the Ecuador coast (Acevedo-Gutiérrez and Smultea 1995; Flórez-González *et al*. 1998; Félix *et al*. 2011, 2012). This breeding ground covers a wide range of approximately 3,000 km of coast, with at least five separate humpback whale calving areas, in which individual movement between several of them indicate some degree of interchange among these areas (Flórez-González *et al*. 1998, Rasmussen *et al*. 2007). Although its breeding area has a large extension located off Central America in the Northern Hemisphere, BSG follows the migration pattern of the other Southern Hemisphere humpback whale populations.

Comparisons of photo-identification catalogues and mitochondrial DNA (mtDNA) data have supported a migratory connection between the feeding areas

[1]Corresponding author (e-mail: anacypriano@yahoo.com.br).

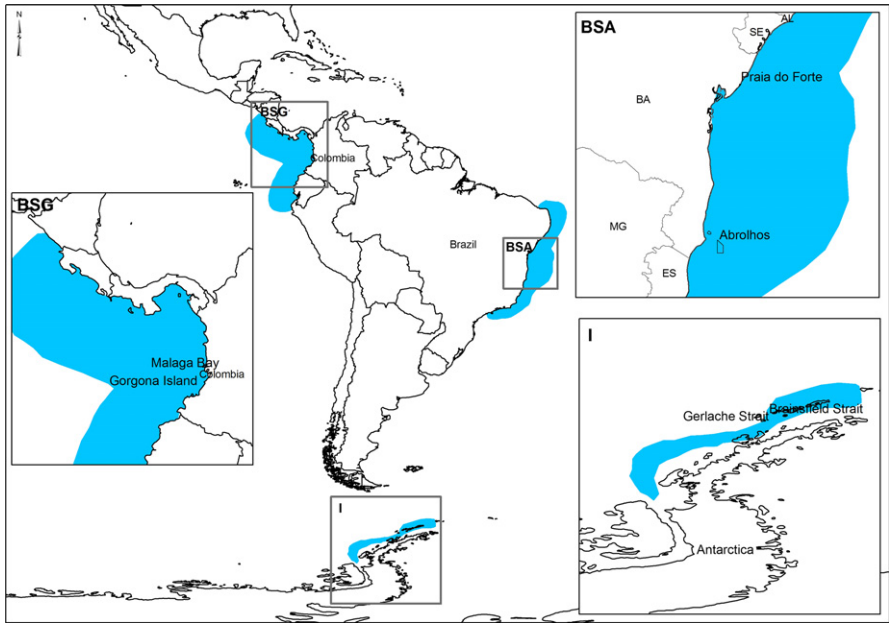

Figure 1. Map showing the geographic distribution (blue) of each South American breeding ground, Breeding stock A (BSA) and Breeding stock G (BSG), and the feeding areas around Antarctic Peninsula (Area I). Sampling locations in Brazil (Abrolhos Bank and Praia do Forte), Colombia (Malaga Bay and Gorgona Island) and Antarctic Peninsula (Brainsfield Strait and Gerlache Strait) are shown (boxed regions).

around the western Antarctic Peninsula (AP) (~60°S, 64°W) and BSG (Stone *et al*. 1990, Caballero *et al*. 2001, Garrigue *et al*. 2002, Stevick *et al*. 2004, Olavarría *et al*. 2006, Rasmussen *et al*. 2007, Félix *et al*. 2012). Further, the Magellan Strait (MS) and the Corcovado Gulf (CG), located off southwestern South America, are recognized as two other feeding areas for BSG (Gibbons *et al*. 2003, Acevedo *et al*. 2007, Capella *et al*. 2008, Hucke-Gaete *et al*. 2013). Nevertheless, photo-identification and genetic data have shown that these feeding aggregations are distinct from the AP feeding areas, suggesting segregation of maternal lineages in these feeding areas of BSG (Olavarría *et al*. 2006; Acevedo *et al*. 2007, 2013), similarly to what was observed for the humpback whale populations from the North Atlantic and North Pacific Oceans, and from the west coast of Africa (Stevick *et al*. 2006*a*, Baker *et al*. 2013, Carvalho *et al*. 2014).

The historical hypothesis of a migratory link between whales wintering off Brazil and those found in Antarctic Peninsula feeding areas is not supported given the lack of photographic matches (Stevick *et al*. 2004) and the significant mtDNA differentiation (Engel *et al*. 2008) between these two regions. On the other hand, a migratory connection between BSA and feeding areas around South Georgia and South Sandwich Islands (between 54°S and 60°S, and 33°W and 22°W) in the Scotia Sea, had been suggested by Slijper (1962, 1965) and Mackintosh (1965). This connection was first supported by satellite telemetry (Zerbini *et al*. 2006, 2011*a*) and further corroborated through individual photo-identification (Stevick *et al*. 2006*b*, Engel and

Martin 2009). In addition, two whales sampled in 2006 near the island of South Georgia presented mtDNA haplotypes found in the Brazilian breeding grounds (Engel *et al*. 2008), one of which showed a putative parent-offspring relationship with a female sampled off Abrolhos Bank in 2001 (Cypriano-Souza *et al*. 2010).

Previous studies based on photo-identification and mtDNA genetic data support the differentiation between humpback whale breeding grounds from the western and eastern coast of South America (Stevick *et al*. 2004, Olavarría *et al*. 2007, Engel *et al*. 2008), but the small number of photo-identification records and the maternal inheritance of the mtDNA limit the detection of important processes, such as a low male-mediated gene flow. Furthermore, there is evidence of some movement between these areas, since a female accompanied by a calf sighted in Ecuador (BSG) was resighted travelling in a pair in Abrolhos Bank (BSA), representing the first record of a humpback whale on both the Pacific and Atlantic breeding grounds of South America (Stevick *et al*. 2013).

Here we assess the degree of population genetic structure and the gene flow between the humpback whales from South American breeding grounds, and confirm the relationship of Colombian breeding grounds to the Antarctic Peninsula feeding grounds, based on the analysis of mtDNA control region haplotypes (465 bp) and microsatellite genotypes (16 loci). We present the first analysis of population differentiation using microsatellite genotypes in humpback whales from Brazil and Colombia and the first use of genotypes to assess individual assignment and potential interchange between breeding grounds.

Methods

Sample Collection, DNA Extraction and Sex Determination

A total of 511 skin samples of humpback whales were analyzed here for at least one marker (see Table S1), collected from two South American breeding grounds (Brazil, $n = 277$ and Colombia, $n = 148$) during the winter breeding season (July–November), and during the austral summer at the Antarctic Peninsula feeding grounds (AP, $n = 86$) (Fig. 1). Samples from Brazil (Abrolhos Bank and Praia do Forte) were collected by Instituto Baleia Jubarte from 1997 to 2011, while samples from Colombia (Gorgona Island and Malaga Bay, Colombian Pacific coast) were collected by Fundación Yubarta from 1991 to 1999. Several research groups collected samples from AP: the Chilean Antarctic Institute (INACH) from 1996 to 1999, Southern Ocean Global Ecosystems Dynamics (SO-GLOBEC) in 2002, and IDCR surveys by the IWC (three samples used in this study) in 1990 and 1994. Most of the samples were obtained by biopsy dart (Lambertsen 1987) but a few samples from Brazil were collected from stranded whales ($n = 9$). Of the 511 samples, 83 that were collected off Brazil, between 2006 and 2011, have not been previously used in published studies. Biopsy sampling of humpbacks from Brazil followed the same guidelines performed in previous studies (Engel *et al*. 2008, Cypriano-Souza *et al*. 2010). Samples were preserved in 70% ethanol and were stored at –20°C prior to DNA extraction. Total genomic DNA of the previously unreported Brazilian humpback samples ($n = 83$) was extracted using the DNeasy Blood and Tissue Kit (QIAGEN) according to manufacturer's protocol.

Sex determination of the 83 Brazilian samples newly reported was carried out by PCR amplification followed by *Taq*I digestion of the genes ZFX and ZFY, as

described in previous study for the remaining samples (n = 194, reported in Cypriano-Souza *et al*. 2010). Samples from Colombia (n = 148) and Antarctic Peninsula (n = 86) were sexed by amplification of a SRY gene and a ZFX positive control by Olavarría *et al*. (2006, 2007). The sex ratio for each population was compared to an expected 1:1 ratio using an exact binominal test.

Mitochondrial DNA Control Region Sequencing and Diversity

Mitochondrial DNA control region sequences of humpback whales from Colombia (BSG, n = 148) and AP (n = 86) feeding ground were described previously (GenBank accession numbers DQ768307 to DQ768421 from Olavarría *et al*. 2006, 2007). Sequences from Brazilian (BSA, n = 176) humpbacks were described in two publications (GenBank accession numbers AY329844 to AY330096 from Engel *et al*. 2008, and GQ913691 to GQ913852 from Rosenbaum *et al*. 2009). Given previous discrepancies in some of these published haplotypes, here we validate haplotypes of BSA humpbacks by resequencing and reviewing 59 of the 66 haplotypes described by Rosenbaum *et al*. (2009). To accomplish this, we resequenced 59 samples (the only available) representing these haplotypes. Additionally, the remaining haplotypes were validated by reviewing the original (of high quality) electropherograms (for more details of the validation process see Appendix S1).

For resequencing, a fragment of approximately 800 bp of the mtDNA control region was amplified using the primers M13Dlp1.5 (tPro whale, 5′-TGTAAAACGACGGCCAGTTCACCCAAAGCTGRARTTCTA-3′, Baker *et al*. 1998*a*) and Dlp8G (5′-GGAGTACTATGTCTGTAACCA-3′, Pichler *et al*. 2001). PCR conditions followed the same guidelines described in Olavarría *et al*. (2007). PCR products were cleaned using ExoSap-IT (USB) and sequenced in one direction using the primer M13Dlp1.5 with BigDye Dye Terminator Chemistry v3.1 (Applied Biosystems). Sequencing reactions were cleaned using CleanSEQ (Agencourt) to remove unincorporated primers and read on an ABI 3730XL (Applied Biosystems) at Hatfield Marine Science Center (HMSC) of Oregon State University (OSU) in Newport, Oregon, USA. Fifty-eight of the 59 Brazilian samples provided clean sequences. Sequences were aligned, manually adjusted and all variable positions were confirmed visually from the electropherograms using SEQUENCHER 5.0 (Gene Codes Corporation). Sequences were trimmed to a consensus length of 464 bp and the different haplotypes were determined using DnaSP 5.10.1 (Librado and Rozas 2009).

Molecular diversity indices, such as haplotype (h) and nucleotide (π) diversities, and haplotype frequency were estimated for each population using the program ARLEQUIN 3.5 (Excoffier and Lische 2010). Pairwise differences in haplotype and nucleotide diversities between populations were tested by a permutation procedure (10,000 permutations, with significance set at $\alpha = 0.05$) based on an R script (*genetic_diversity_diffs* 1.0.0, Alexander 2015). ARLEQUIN was also used to estimate Tajima's D (Tajima 1989) neutrality test for each population with 1,000 bootstrap replicates.

Microsatellite Genotyping and Diversity

Samples were genotyped at 16 published microsatellite loci (14 dinucleotides: EV1, EV14, EV21, EV37, EV94, EV96, EV104 (Valsecchi and Amos 1996), GT23, GT211, GT575 (Bérubé *et al*. 2000), 464/465 (Schlötterer *et al*. 1991), RW4-10, RW31, RW48 (Waldick *et al*. 1999), and two tetranucleotides: GATA28 and

GATA417 (Palsbøll *et al*. 1997)). Samples from Colombia ($n = 130$) and AP ($n = 77$) had been genotyped for all loci on an ABI 3730XL (Applied Biosystems, Foster City, CA), and these genotypes were provided by South Pacific Whale Research Consortium (SPWR) (Steel *et al*. 2008) for a collaborative comparison between populations. We have previously genotyped 134 samples ($n = 51$ from Cypriano-Souza *et al*. 2010; $n = 83$ from ALC-S, unpublished data) from Brazil, for seven of the 16 loci on a MegaBACE 1000 automated sequencer (GE Healthcare) at Pontifical Catholic University of Rio Grande do Sul, Brazil. The remaining nine loci for this population were genotyped on an ABI 3730XL at HMSC or by Macrogen Inc. (Seoul, South Korea). To calibrate the microsatellite allele sizes of these two data sets genotyped at different machines, some samples were included in different runs, as detailed in Appendix S2.

For the samples genotyped at HMSC and Macrogen, PCRs were carried out in 10 μL with the following conditions: 1.5–4.0 mM of $MgCl_2$ (different concentrations depending on the locus), 200 μM of dNTPs, 0.4 μM of each primer, bovine serum albumin, 0.25 units of Platinum *Taq* DNA polymerase (Invitrogen), 1X PCR buffer (Invitrogen), and 1 μL of DNA (~50 ng). All loci were amplified in separate reactions using the following thermocycle profile: initial denaturation at 94°C for 3 min, 35–40 cycles of denaturation at 94°C for 30 s, annealing at 50°C for 30 s, and extension at 72°C for 40 s, followed by a final extension at 72°C for 10–30 min depending on the locus. Negative controls were run at the PCR step to control for exogenous contamination. Amplicons were pooled in four sets of up to five loci (set 1: EV1, EV14, EV21, and EV104; set 2: GT211, GT575, GATA417, GATA28, and RW4-10; set 3: RW31, RW48, EV37, EV94, and EV96; set 4: 464/465 and GT23) and co-loaded on an ABI 3730XL, and alleles sizes in base pairs (bp) were determined using the LIZ-500 size standard run in each lane. Microsatellite alleles were visualized and scored using GENEMAPPER 4.0 (Applied Biosystems) or Peak Scanner 1.0 (Applied Biosystems).

The program MICRO-CHECKER 2.2.3 (Van Oosterhout *et al*. 2004) was used to check for possible null alleles, large allele dropout, and scoring errors due to stutter peaks. Analyses for matching genotypes (replicate samples and recaptures) and estimates of the probability of identity (the Hardy-Weinberg [HW]$P_{(ID)}$, and the more conservative, $P_{(ID)sib}$ (Waits *et al*. 2001)) were performed by GENECAP 1.4 (Wilberg and Dreher 2004). Pairs of matching genotypes were subsequently compared for sex and mtDNA haplotype data (when available) to support the identification of replicates and/or recaptures. Replicate samples within regions were excluded from further analyses.

Genetic diversity was estimated as the number of alleles per locus (K), observed and expected heterozygosities (H_O and H_E, respectively) under Hardy-Weinberg assumptions (Nei 1978) for each locus in each population using ARLEQUIN 3.5. FSTAT 2.9.3 (Goudet 2002) was used to estimate allelic richness (AR) and Weir and Cockerham's (1984) measure of F_{IS}. Global and population-specific tests for deviations from Hardy-Weinberg equilibrium (HWE) (Guo and Thompson 1992) and linkage disequilibrium were carried out using the program ARLEQUIN, corrected for simultaneous comparisons with the sequential Bonferroni test (Rice 1989).

Population Genetic Structure and Assignment

Population genetic differentiation for mtDNA control region sequences and for microsatellites was estimated with pairwise F-statistics (F_{ST}, Weir and Cockerham

1984) and tested by permutation procedure in ARLEQUIN (10,000 permutations, with significance set at $\alpha = 0.05$). Population genetic structure was also assessed for the whole data set with a Bayesian clustering approach implemented in STRUCTURE 2.3.3 (Pritchard *et al*. 2007). First, 10 independent runs were performed for each K (number of cluster) between one and five with no prior information on sampling location, using the admixture and correlated allele frequencies model. Burn-in and length of simulation were set at 500,000 and 1,000,000 iterations, respectively. In a second set of analyses, the sampling location prior was used to assist the identification of clustering, as suggested for cases of subtle population structure (Hubisz *et al*. 2009). The results generated were processed in STRUCTURE HARVESTER 0.6.93 (Earl and vonHoldt 2012), a web-based program that determines the optimal number of clusters using the Evanno method (Evanno *et al*. 2005). The multiple results generated by STRUCTURE were summarized in CLUMPP (Jakobsson and Rosenberg 2007), and graphically displayed by DISTRUCT (Rosenberg 2004). Additionally, we ran a STRUCTURE analysis with only Colombia and AP samples using the same parameters as above. In a third STRUCTURE analysis, the "Use PopInfo" option ($G = 0$) was applied to assess individual assignment, possible migrants (individuals born in a population other than the one in which they were sampled), and individuals with admixed ancestry (individuals descended from a recent migrant). The length of simulation and burn-in were the same as described above. Individuals with a membership coefficient $q > 0.8$ were considered residents for the population from which they were sampled, while those with $0.2 < q > 0.8$ were considered to be potentially admixed (neither resident nor migrant), and those with $q < 0.2$ (or $q > 0.8$ for a another population) from the sampled population were considered migrants (*e.g*., Lecis *et al*. 2006, Bergl and Vigilant 2007).

Additionally, the Bayesian assignment method in GENECLASS 2.2.2 (Piry *et al*. 2004) was used to assess assignment/exclusion of individuals to the predefined sampling populations. This procedure is complementary to STRUCTURE analysis since it does not assume that all potential source populations have been sampled (Manel *et al*. 2002, Pearse and Crandall 2004). The program was also used to detect first-generation migrants (individuals sampled in a population other than the one in which they were genetically assigned) using the Bayesian criterion of Rannala and Mountain (1997) in combination with a Monte Carlo resampling algorithm with 10,000 simulated individuals and an alpha of 0.01 to determine the critical value of the L_{home}/L_{max} likelihood test statistics, beyond which individuals were assumed to be F1 migrants (Paetkau *et al*. 2004). Each individual was assigned to the population in which the likelihood of its genotype is the highest, comparing the likelihood of the population where the individual was sampled with those of the available populations $-\text{Log}(L_{home}/L_{max})$.

Gene Flow and Effective Population Size

The magnitude and direction of recent migration were estimated using a Bayesian method implemented in BayesAss 3.0.3 (Wilson and Rannala 2003). Analyses were conducted for 10 independent MCMC runs of 10^7 steps recorded every 200 iterations, with the first 10^6 repetitions discarded as burn-in. To reach the recommended acceptance rates between 20% and 40%, the values of parameters such as migration rates (Δ_M), allele frequencies (Δ_A) and inbreeding coefficient (Δ_F) were adjusted to 0.1, 0.2, and 0.15, respectively. Trace files were examined for convergence.

A Bayesian coalescent approach implemented in MIGRATE-N 3.6.4 (Beerli 2008) was used to estimate the mutation-scaled population size θ ($4N_e\mu$ or $2N_{ef}\mu$) and the mutation-scaled immigration rates M (m/μ) based on both microsatellite and mtDNA data. Analyses were conducted for 10 replicate searches, with 10,000 genealogies recorded every 200th genealogy estimate after a burn-in of 500,000 steps, retaining a total of 20 million parameter values. The effective number of immigrants per generation (N_em) was estimated by multiplying θ and M (Beerli 2008).

Data Archiving

In accordance with the guidelines of the Joint Data Archiving Policy (Whitlock *et al*. 2010), the data underlying the primary conclusions of this article have been submitted to a public archive. The mtDNA sequences (including haplotype names and GenBank accession numbers published in previous studies, see Table S1) and microsatellite genotypes have been deposited in the Dryad repository. Also, the input files (parmfile and infile) to MIGRATE were deposited in the Dryad.

Results

Sex Determination

Sex was determined successfully for 118 (82 males and 36 females) individuals from Colombia, 82 (39 males and 43 females) individuals from AP, and 264 (141 males and 123 females) individuals from Brazil (after replicates were removed, see below) (Table S1). The sex ratio of the Colombian humpback whales was significantly skewed toward males (2.27:1, $P < 0.0001$) but the sex ratio did not differ significantly from the expected 1:1 for the whales sampled in Brazil (1.14:1, $P = 0.295$) or the AP (1.10:1, $P = 0.741$).

mtDNA Diversity

After the validation process (see Appendix S1) and removal of replicate samples (see below and Table S1), we obtained 54 haplotypes defined by 60 variable sites for the mtDNA control region alignment (464 bp) from 158 sequences from the Brazilian population. For the populations of Colombia and AP, the mtDNA control region alignment from 130 and 82 sequences, described 27 and 21 haplotypes defined by 41 and 40 variable sites, respectively. A total of 77 haplotypes, defined by 67 variable sites were found in the sequences' alignment (in 465 bp) from 370 individuals across all three populations (Table 1). Of these 77 haplotypes, seven were shared between Brazil and Colombia, 18 between Colombia and AP and four among all three areas. The haplotype frequencies for each area are shown in Table S2.

Haplotype diversities (h) and nucleotide diversities (π) are shown in Table 1. The permutation test confirmed a significantly higher haplotype diversity for Brazil compared to Colombia ($P < 0.001$) or the AP ($P < 0.001$) but no significant difference between Colombia and AP ($P = 0.760$). Similarly, the permutation test confirmed a significantly higher nucleotide diversity for Brazil compared to Colombia ($P = 0.024$) and AP ($P = 0.013$), but not between Colombia and AP ($P = 0.633$). The neutrality test was near zero and nonsignificant for each population (data not shown).

Microsatellite Diversity

A total of 337 samples (Brazil, $n = 130$; Colombia, $n = 130$; and AP, $n = 77$) were genotyped successfully for at least 11 microsatellite loci, of which 321 (95.2%) were genotyped for all 16 loci (See Table S1). The loci were adequate for individual identification since the probabilities of identity were very low (HW $P_{(ID)} = 2.01 \times 10^{-18}$, and $P_{(ID)sib} = 4.65 \times 10^{-7}$), indicating that even related individuals would have a low probability of sharing identical genotypes. Based on genotype identity and accessory information, such as sex and mtDNA haplotype matches (when available), 23 genotypes were identified as replicate samples, of which 15 were resamplings of the animals in the same year and three were resightings (recapture) in different years along the Colombia coast, four were resamplings of the animals in the same day (individual biopsied twice within a group) in AP, and one previously reported match between Colombia and the AP (Steel *et al.* 2008). No match was found between the two breeding grounds (Brazil and Colombia) or between Brazil and AP. Thus, genotypes were assigned to 314 different individuals (Brazil, $n = 130$; Colombia, $n = 112$; and AP, $n = 72$) that were included in further analyses (Table 2 and S1).

No evidence for stuttering or large allele dropout was seen at loci for all samples pooled or for each population separately. No evidence of null alleles was found for all samples pooled and for Brazil and AP separately, but there was indication of null alleles for the loci 464/465 ($P < 0.01$) and EV94 ($P < 0.05$) in Colombia. However, deviation from Hardy-Weinberg equilibrium (HWE) was detected only at the EV96 locus in Colombia after Bonferroni correction. This locus was retained in the analyses since its removal did not affect the test results (data not shown). Finally, there was no

Table 1. Summary of sample size and mtDNA diversity of the humpback whales sampled at two breeding grounds (Brazil and Colombia) and one feeding area (Antarctic Peninsula). Includes number of haplotypes (H), haplotype (h) and nucleotide (π) diversities. For h and π, standard deviations are shown in parentheses.

Stock	Region	Sample size	H	h (SD)	π (SD)
BSA	Southwestern Atlantic/Brazil	158	54	0.973 (0.004)	2.00 (0.047)
BSG	Southeastern Pacific/Columbia	130	27	0.906 (0.015)	1.80 (0.052)
BSG	Antarctic Peninsula	82	21	0.902 (0.020)	1.80 (0.095)
	Overall	370	77	0.962 (0.005)	1.90 (0.032)

Table 2. Microsatellite diversity for humpback whales sampled at the two breeding grounds (Brazil and Colombia) and one feeding area (Antarctic Peninsula). Number of individuals (n), number of alleles (K), allelic richness (AR), observed (H_O) and expected (H_E) heterozygosity, inbreeding coefficient (F_{IS}) and probabilities of identity ($P_{(ID)}$ and $P_{(ID)sib}$). Duplicate samples were removed from the analysis.

Region	n	K	AR	H_O	H_E	F_{IS}	$P_{(ID)}$	$P_{(ID)sib}$
Brazil	130	10.2	9	0.745	0.74	0.007	1.30×10^{-18}	4.30×10^{-7}
Colombia	112	8.9	8.1	0.699	0.729	0.041	1.02×10^{-17}	5.90×10^{-7}
AP	72	8.4	8	0.704	0.712	0.011	8.18×10^{-17}	9.40×10^{-7}
Overall	314	10.8	8.8	0.719	0.735	0.019	2.01×10^{-18}	4.65×10^{-7}

significant linkage disequilibrium among loci for all samples pooled or within each population after Bonferroni correction.

The nuclear diversity for each population analyzed separately is showed in Table S3. All loci were polymorphic within each population, with the number of alleles per locus ranging from 4 (EV1 for all populations) to 17 (EV37 and GATA417 for Brazil), and from 4 (EV1) to 20 (EV37) for all samples pooled (data not shown) with a mean of 10.8. The total allelic richness estimated for a minimum sample size of 46 individuals was 8.8, ranging from 8.0 (AP) to 9.0 (Brazil). The mean observed heterozygosity for the total sample was 0.719, ranging from 0.699 (Colombia) to 0.745 (Brazil) (Table 2). F_{IS} values were low for all loci (below 0.19) (Table S3) and were not significant when tested for each population separately or for the total sample (Table 2).

Population Genetic Structure

The pairwise tests of differentiation for mtDNA and microsatellites showed very similar and significant differences between Brazil *vs*. Colombia and Brazil *vs*. AP, but no difference between Colombia *vs*. AP (Table 3). The F_{ST} values were several-fold higher for the mtDNA than for the bi-parental microsatellite markers. Finally, the F_{ST} values between Brazil *vs*. Colombia + AP for mtDNA (F_{ST} = 0.058) and microsatellites (F_{ST} = 0.011) are almost identical to values obtained (Table 3) for the comparison between Brazil and the other two areas separately.

The results of the STRUCTURE analysis, with and without location priors, were consistent with the tests of differentiation. The highest posterior likelihood was for K = 2 clusters corresponding to the individuals from Brazil and from Colombia + AP (mean Ln P(D) = –16,402.0, Fig. 2). This was also corroborated by the ΔK method of Evanno *et al*. (2005) (Fig. S1). There is no evidence for any cryptic structure between Colombia and AP samples, as seen in the plot of proportional membership for K = 3 (Fig. 2) as well as in STRUCTURE run only with Colombia and AP samples (Fig. S2).

Identifying Recent Dispersal and Recent Gene Flow

In the STRUCTURE analysis using the PopInfo option assuming two populations (Brazil and Colombia + AP), most (96.2%) individuals had a high probability of being residents ($q > 0.8$) of the population in which they were sampled. However, six individuals (four males and two unknown sex) did not assign clearly to the region

Table 3. Differentiation, expressed as F_{ST}, between humpback whales from Brazil, Colombia and AP based on mtDNA control region haplotypes (below diagonal) and microsatellite loci (above diagonal). n = sample size used for mtDNA haplotypes (vertical) and microsatellite loci (horizontal).

Region	n	Brazil 130	Colombia 112	AP 72
Brazil	158	—	**0.0115**	**0.0111**
Columbia	130	**0.0572**	—	0.0006
AP	82	**0.0589**	0.0012	—

Significant P-values are highlighted in bold ($P < 0.001$).

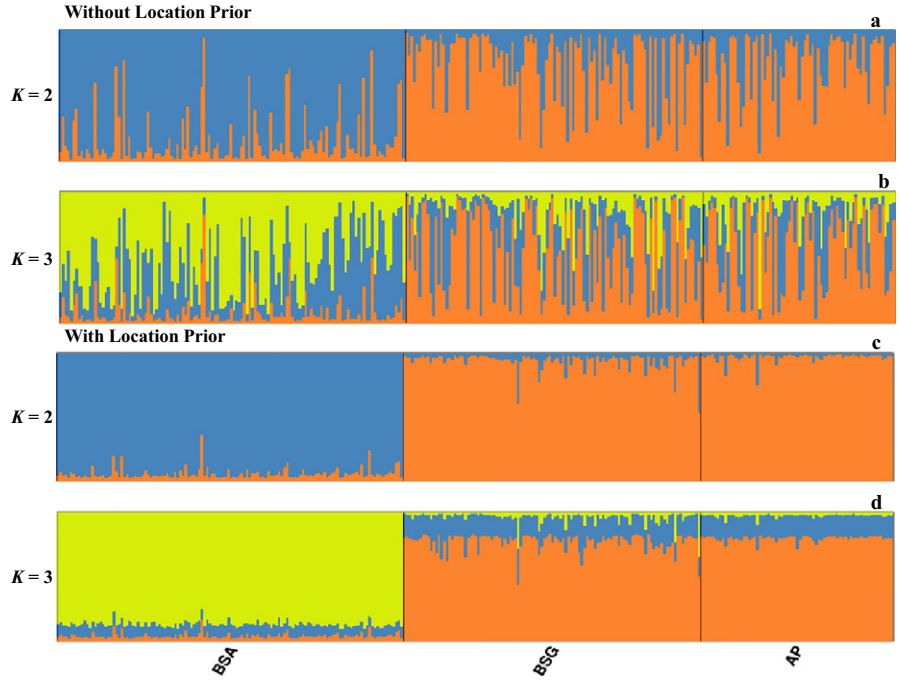

Figure 2. STRUCTURE bar plot of the proportional membership of each individual of humpback whale in each cluster for $K = 2$ (a) and $K = 3$ (b) without the sampling location prior, and for $K = 2$ (c) and $K = 3$ (d) with the location prior. Each individual is represented by a vertical bar broken into colored segments with the length indicating the coefficient of membership to each cluster. BSA = breeding stock A (Brazil), BSG = breeding stock G (Colombia) and AP = Antarctic Peninsula. Black line represents the boundary between individuals sampled in each of the three regions.

from which they were sampled (Table 4). These individuals were considered as likely migrants ($q < 0.2$ for the population in which they were sampled or $q > 0.8$ for the other population) or individuals with recent admixed ancestry ($0.2 < q > 0.8$ for the population from which they were sampled) (Fig. 3).

Using GENECLASS, there were four individuals considered to be likely first-generation migrants (Table 4). Three of these individuals (BR386, BR670 and Mno92Co020) were also identified as potential migrants ($q > 0.8$ for the other population) or individuals with recent admixed ancestry ($0.2 < q < 0.8$) by STRUCTURE. Two of these were males sampled in Brazil and considered to be first-generation migrants from Colombia and one (unknown sex) sampled in Colombia was considered a first-generation migrant from Brazil. Only four of the seven individuals identified as potential migrants or with admixed ancestry have mtDNA haplotypes available. While two haplotypes (SP14 and SP62) are common at both breeding grounds and at AP feeding ground, one haplotype (SP52) is specific for Colombia and AP. The latter was found in a whale sampled in AP and assigned to Brazil (Table 4).

Finally, the estimates of recent gene flow in BayesAss using microsatellites suggested a directional bias in migration, with a moderate migration from Colombia to

Table 4. Potential migrants and individuals with admixed ancestry identified by STRUCTURE and/or GENECLASS.

Individuals	Sex	mtDNA haplotype original/matching	Region		STRUCTURE *q*-values (Brazil/ Colombia)	GENECLASS		
			Sampled	Assigned		$-\text{Log}(L_{home}/L_{max})$	*P*	[−log(L)] (Brazil/ Colombia)
BR205_02	Male	—	Brazil	Colombia	0.539/0.461	1.374	0.0065#	20.443/19.069
BR386_06	Male	—	Brazil	Colombia	0.075/**0.925****	3.163	0.0002#	22.342/**19.179**
BR670_10	Male	—	Brazil	Colombia	0.282/**0.718***	2.222	0.0020#	24.071/**21.849**
Mno92Co020	—	SP62/HBA104	Colombia	Brazil	**0.793**/0.207*	4.003	0.0002#	**21.652**/25.685
Mno96Co032	Male	SP14	Colombia	Brazil	0.684/0.316*	2.005	0.0106	21.041/23.095
Mno99Co004	—	SP62/HBA104	Colombia	Brazil	0.919/0.081**	1.749	0.0196	23.196/24.946
Mno96AP019	Male	SP52	AP	Brazil	0.585/0.415*	1.577	0.0245	19.882/21.399

Note: In STRUCTURE, two individuals were identified as likely migrants (**), *i.e.*, individuals sampled in a population other than the one in which they were genetically assigned, and four as likely admixed (*), *i.e.*, individuals descended from a recent migrant. In GENECLASS, four individuals were identified as F1 migrants (#). Individuals identified as migrants or F1 migrants by both methods and the most likely source population for each individual are shown in bold. Significant *P*-value ($P < 0.01$) to detect F1 migrants. See Figure 3 for membership coefficients generated by STRUCTURE.

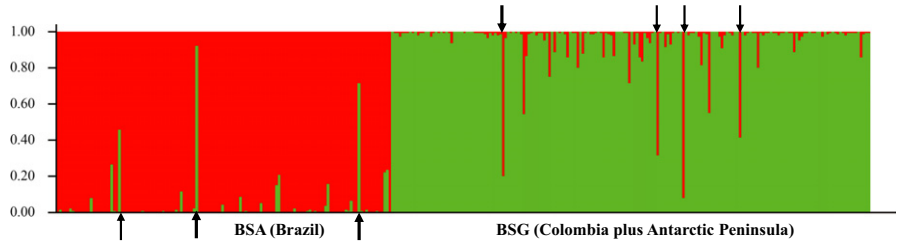

Figure 3. Proportional membership (q) of individuals to the Breeding stock A (Brazil) or Breeding stock G (Colombia plus AP) using STRUCTURE with the prior population information (assuming $K = 2$) incorporated. Arrows indicate individuals with admixed ancestry ($0.2 < q > 0.8$ the population in which they were sampled) and likely migrants ($q < 0.2$ for the population in which they were sampled or $q > 0.8$ for the other population).

Brazil (16% per generation, SE = 5%) and low migration in the opposite direction, from Brazil to Colombia (2% per generation, SE = 1.4%).

Long-term Gene Flow and Effective Population Size

Coalescent-based estimates of long-term gene flow (Nm) based on microsatellite data were moderate (Fig. 4), with the point estimates suggesting a greater gene flow from Colombia to Brazil (35 immigrants per generation) than that from Brazil to Colombia (23 immigrants per generation), although the confidence intervals were wide and overlapped. By comparison, the long-term migration rates based on mtDNA data were very low (Fig. 4) and asymmetric, with greater gene flow from Brazil to Colombia (eight females per generation) than that from Colombia to Brazil (one female per generation). These estimates suggest an interchange between these breeding grounds of around 1–1.5 individuals per year, and of one female every 3–21 yr (assuming a generation time of 21.5 yr, Taylor *et al.* 2007). Additionally, the coalescent analyses in *migrate-n* estimated a θ for Brazil that is 2- or 5-fold higher than that estimated for Colombia based on microsatellites and mtDNA data, respectively (Fig. 4), with little or no overlap in 95% posterior probabilities.

Discussion

Validation of mtDNA Control Region Haplotypes

The quality control review and standardized matching of mtDNA sequences included here are an important incremental advance to describing the genetic diversity of the Brazilian humpback whales. A review of published mtDNA sequences used in human forensics has shown that large-scale surveys of haplotypes are seldom error free (Bandelt *et al.* 2001, 2002). One of the most common errors is the so-called "phantom mutation" generated in the sequencing process (Bandelt *et al.* 2001). These erroneous variants (*e.g.*, false transversions, insertions, and deletions) can show a mutation pattern that significantly differs from that of true mutations (Bandelt *et al.* 2002). Indeed, the 10 mtDNA haplotypes that were not validated here were singletons and presented mainly transversions at 1–5 polymorphic sites, probably due to misreading of low quality electropherograms. Although these few "singleton" errors

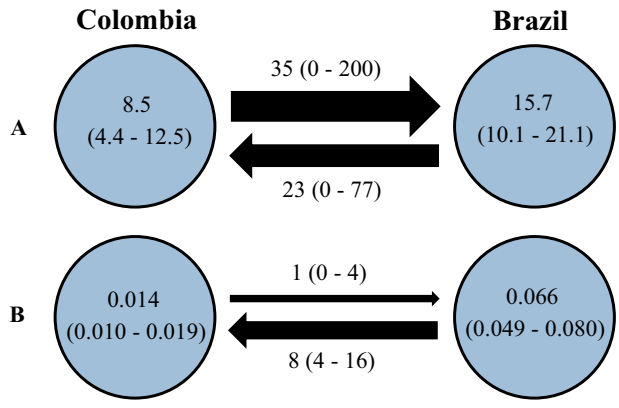

Figure 4. Diagram showing the *migrate-n* coalescent-based estimates of θ (within circles) and long-term gene flow (arrows) between Brazil and Colombia for microsatellite loci (A) and mtDNA control region sequences (B). The width of the arrows corresponds to the numbers of immigrants per generation (Nm) between populations.

did not affect greatly the previous results at population level, correct sequences are essential for comparisons based on shared haplotypes, such as those with historical samples of this species (*e.g.*, Sremba *et al*. 2014).

Genetic Diversity

Our data confirm the high genetic diversity for both mtDNA (Olavarría *et al*. 2007, Engel *et al*. 2008) and microsatellite (Cypriano-Souza *et al*. 2010) markers in the humpback whale populations that winter along western (microsatellites data reported for the first time for Colombia) and eastern coasts of South America. High genetic diversity was also found in other breeding grounds in the Southern Hemisphere (Valsecchi *et al*. 2002, Olavarría *et al*. 2007, Rosenbaum *et al*. 2009, Félix *et al*. 2012), notwithstanding that commercial whaling during the 20th century reduced the humpback whale populations to a small fraction of their preexploitation abundance (Tønnessen and Johnsen 1982). In this Hemisphere, more than 200,000 humpbacks were killed from 1904 to 1972, after accounting for the illegal Soviet whaling (Clapham *et al*. 2009). The absence of a significant reduction of the genetic diversity in these populations is likely due to the relatively brief duration of the population bottleneck (about three generations, assuming a generation time of 21.5 yr, Taylor *et al*. 2007), the longevity of the species (probably greater than 70 yr, Gabrielle *et al*. 2010) and the relatively large minimum absolute population size (*e.g.*, N_{min} = 500 individuals in the late 1950s, for the Brazilian population, Zerbini *et al*. 2011*b*).

The nuclear and mitochondrial genetic diversity (and consequently the θ estimates) of the whales from Colombia (separately or together with those from Antarctic Peninsula) are lower than those from Brazil (Tables 1 and 2, Fig. 4), corroborating previous results in which the BSG has shown the lowest mtDNA diversity among the humpback populations from Southern Hemisphere (Olavarría *et al*. 2007, Rosenbaum *et al*. 2009, Valsecchi *et al*. 2010, Félix *et al*. 2012). The possible reasons for lower genetic diversity of BSG are not known. One feasible explanation is that the long-term effective size of BSG has been lower than other breeding stocks or that it suffered from more intensive exploitation. Lower genetic variability could also be

explained by a relatively recent recolonization of the southeastern Pacific Ocean after the glacial and interglacial oscillations of the climatic conditions of the eastern Pacific (Baker *et al*. 1993, Lambeck *et al*. 2002).

Population Genetic Structure and Breeding Stocks Identification

Our results, based on mtDNA and microsatellite genotypes, supported previous studies showing a close connection between the Colombia breeding grounds and the Antarctic Peninsula feeding areas (Stone *et al*. 1990; Caballero *et al*. 2001; Garrigue *et al*. 2002; Stevick *et al*. 2004; Olavarría *et al*. 2006, 2007; Steel *et al*. 2008). Based on this evidence, both Colombia and the Antarctic Peninsula are seasonal habitats for a single breeding stock, BSG. Our results also confirm that BSG, as represented by both Colombia and the AP, are significantly different in both mtDNA and microsatellite genotypes from the Brazilian breeding grounds, BSA. This was also evident in the Bayesian clustering analysis done without sampling location prior information (Fig. 2). This is notable since STRUCTURE often fails to detect differences between relatively closer populations of whales without using the sampling locations as priors, even when population genetic structure is evident by standard tests of differentiation (*e.g*., Carroll *et al*. 2011, Schmitt *et al*. 2014).

However, it is important to highlight that BSG is the most differentiated of all South Pacific breeding grounds (mtDNA F_{ST} ~0.058–0.079), even when compared with its nearest neighbors to the west, French Polynesia and the Cook Islands, (Breeding Stock F) (Olavarría *et al*. 2007). The significant genetic differentiation between Brazil and Colombia, as shown in our study supports the previously suggested isolation of the latter breeding grounds. Indeed, six mtDNA haplotypes (SP32, SP60, SP61, SP90, SP98, and SP101) shared among BSG breeding grounds and Antarctic Peninsula and Magellan Strait feeding areas have not been found in any other population from Southern Hemisphere (Félix *et al*. 2012). In addition, BSG together with its corresponding feeding grounds represent the only population in the Southern Hemisphere where haplotypes of the SH clade were not found (Olavarría *et al*. 2007, Félix *et al*. 2012). Although historical or current trans-equatorial gene flow mediated by the overlapping breeding grounds of the two hemispheres could explain the isolation of BSG from other Southern Hemisphere populations (Caballero *et al*. 2001, Olavarría *et al*. 2007), estimates of migration rates between the humpbacks from the North and South Pacific Oceans are low, suggesting little or no contemporary interchange (Jackson *et al*. 2014).

Interchange and Patterns of Gene Flow

The Bayesian analyses of nuclear DNA markers using STRUCTURE, GENECLASS, and BAYESASS provided direct evidence of some dispersal (putative migrants) and some recent gene flow (admixed individuals) between Brazil and Colombia. Recent interchange of individuals between these two breeding grounds is consistent with the record of a female photo-identified in Ecuador (BSG) and later resighted in Abrolhos Bank (BSA) (Stevick *et al*. 2013). Our results are also similar to a recent study that detected some possible migrants between two blue whale (*Balaenoptera musculus*) populations at the southeastern Pacific Ocean but also found significant differences for both mitochondrial and microsatellite markers between these populations (Torres-Florez *et al*. 2014).

Results of MIGRATE-N analyses also support long-term gene flow between Colombia and Brazil, which is much higher at the biparental nuclear markers than at

the maternal mtDNA, a result that was corroborated by the much lower F_{ST} at the former than the latter marker. These results support the hypothesis that the females of migrant baleen whales are more philopatric than males and therefore the gene flow between these breeding areas is mostly meditated by males (Baker *et al*. 1998*b*, Rosenbaum *et al*. 2009, Carroll *et al*. 2011, Félix *et al*. 2012, Torres-Florez *et al*. 2014). In support of the above, in the Bayesian assignment results, all individuals with some evidence of being migrants that were also sexed are males (Table 4).

Males sharing the same feeding area may occasionally migrate to different breeding grounds in different wintering seasons (Darling and Cerchio 1993, Garrigue *et al*. 2002, Constantine *et al*. 2007). This suggests a male mixed-strategy over a lifetime for gametic dispersal, without abandoning maternal fidelity to feeding grounds (Baker *et al*. 2013). On the other hand, the krill density on feeding areas may influence the choice of humpback whales to undertake long-distance longitudinal movements to maximize feeding opportunities (Amaral *et al*. 2016). This, in turn, may increase the possibility of mixing or switching between breeding grounds. Valsecchi *et al*. (2010) have also shown that the humpback whale migratory pattern seems to be more complex than previously thought, suggesting that males should undertake longer longitudinal movements to maximize mating opportunities, while females need to save energy for reproduction. In evidence to the contrary, however, movements between breeding grounds in different oceans have been reported for two females, indicating a behavioral flexibility for this species (Stevick *et al*. 2010, 2013). However, these females would only contribute to mtDNA gene flow if they contributed a female offspring without permanent dispersal or with permanent dispersal but with no breeding success.

Recently, Félix *et al*. (2012) have demonstrated stratification between adjacent calving areas in the BSG, suggesting different migration patterns between sexes, in which females show higher site fidelity than males. Given the distinctiveness of the three feeding areas (Antarctic Peninsula, Magellan Strait, and Corcovado Gulf) of the BSG (Olavarría *et al*. 2006, Acevedo *et al*. 2013), the heterogeneity between the wintering areas of this stock, and a higher similarity between Magellan Strait and the northernmost breeding areas (Acevedo *et al*. 2007, Capella *et al*. 2008), an alternate explanation would be that whales feeding around the Antarctic Peninsula may mate and calve off Ecuador and Colombia, while Magellan Strait humpbacks may breed mainly off Panama and Costa Rica. However, nothing is yet known about the migratory movements and population structure of the humpback whales feeding off Corcovado Gulf. On the other hand, the different migration patterns between sexes could be a temporal segregation instead of a spatial one, where whales arrive at the breeding ground at different times since they had migrated from different feeding grounds similar to that suggested for the latitudinal gradations found in the breeding stock B (BSB) and other breeding grounds from North Pacific Ocean (Baker *et al*. 2013, Carvalho *et al*. 2014).

Finally, the results suggest the migration between Brazil and Colombia is asymmetrical, but may not be consistent between sexes. While estimates with the biparental nuclear marker suggest that gene flow (mostly from males) was higher from Colombia to Brazil (eastward bias) than that from the other direction, for the mtDNA the long-term gene flow (exclusively from females) is much higher from Brazil to Colombia (westward bias) than that from the other direction. Likewise, a recent genetic study[2] has shown asymmetrical gene flow between breeding grounds

[2]Personal communication from H. C. Rosenbaum, Wildlife Conservation Society, 2300 Southern Boulevard, Bronx, NY 10460, U.S.A., 2 June 2016.

in the Southern Hemisphere, with the westward bias predominating in the long-term, maternally directed gene flow. However, our migration rates based on mtDNA data were very low suggesting that very little female interchange should occur between Brazil and Colombia. In addition, most of these estimates present large confidence intervals that mostly overlap, suggesting that only additional data could test whether this asymmetrical sex-dependent migration is a real pattern.

Conclusion

The genetic differentiation between Brazilian and Colombian breeding grounds observed in our study, together with low to moderate levels of recent and long-term asymmetrical gene flow, indicate a relatively strong differentiation for humpback whale populations with adjacent feeding grounds. However, although these feeding areas are geographically adjacent, there is no evidence of known matches between the feeding areas of BSA, in the Scotia Sea, and those of BSG, around the Antarctic Peninsula. Our results are also consistent with the differentiation of mtDNA among other humpback whale breeding grounds in the Southern Hemisphere (Olavarría *et al*. 2007), suggesting a remarkable example of fidelity to seasonal habitat in the absence of any contemporary barriers. Therefore, we strongly recommend that these breeding stocks (BSA and BSG) should be stated as separate management units. However, further research is required to address other remaining questions about South American breeding grounds. How long have these populations been isolated? What is the level of the distinctiveness between the breeding areas of the BSG and the connections with their respective feeding areas (Antarctic Peninsula, Magellan Strait, and Corcovado Gulf)? What are the genetic characteristics of the feeding areas (around South Georgia and South Sandwich Islands) for BSA? Additional genetic sampling from other BSG calving areas and the BSA feeding areas, as well as in a larger temporal scale during a breeding season will improve our knowledge of the population genetic structure and connectivity between these migratory populations of South America.

Acknowledgments

This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) to ALC-S; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) to SLB; Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) to SLB; Instituto Chico Mendes de Conservação da Biodiversidade/Ministério do Meio Ambiente (ICMBIO/MMA); and Petróleo Brasileiro S.A (PETROBRAS). Special support was given by CAPES for the scholarship award to ALC-S to conduct part of her doctoral research as a visiting student at Oregon State University (OSU). Biopsy samples were collected off Brazil under permit SISBIO 21489-1. This survey was made possible through the generous collaboration of members of the South Pacific Whale Research Consortium. We thank the staff of Projeto Baleia Jubarte/Instituto Baleia Jubarte and of Fundación Yubarta for collecting Brazilian and Colombian humpback whale samples, respectively, in the field. Also, thanks to field scientists of INACH, SO-GLOBEC program and IWC that collected the samples in the Antarctic Peninsula. We thank the colleagues of the Laboratório de Biologia Genômica e Molecular/PUCRS for their help in lab activities, especially to Priscilla Mena Zamberlan and Tiago Ferraz da Silva for laboratory assistance. We are also grateful for laboratory support from Alana Alexander, Renee Gibb, Rebecca Hamner, and Beth Slikas during the work at Cetacean Conservation and Genomics Laboratory, Oregon State University. We thank Guilherme Maurutto

for preparing the map with the study areas. Special thanks to Jennifer Jackson for providing electropherograms of the mtDNA control region sequences for the Brazilian humpback whale samples collected in 1997 and 1998. Finally, we also want to thank the editors Dr. Patricia Rosel and Dr. Daryl J. Boness, and the anonymous reviewers for their valuable and helpful comments.

Literature Cited

Acevedo-Gutiérrez, A., and M. A. Smultea. 1995. First records of humpback whales including calves at Golfo Dulce and Isla del Coco, Costa Rica, suggesting geographical overlap of northern and southern hemisphere populations. Marine Mammal Science 11:554–560.

Acevedo, J., K. Rasmussen, F. Félix, *et al*. 2007. Migratory destination of humpback whales from the Magellan Strait feeding ground, Southeast Pacific. Marine Mammal Science 23:453–463.

Acevedo, J., D. Haro, L. Dalla Rosa, *et al*. 2013. Evidence of spatial structuring of eastern South Pacific humpback whale feeding grounds. Endangered Species Research Journal 22:33–38.

Alexander, A. 2015. genetic_diversity_diffs 1.0.0. Available at https://github.com/laninsky/genetic_diversity_diffs.

Amaral, A. R., J. Loo, H. Jaris, *et al*. 2016. Population genetic structure among feeding aggregations of humpback whales in the Southern Ocean. Marine Biology 163:1–13.

Andriolo, A., P. G. Kinas, M. H. Engel and C. C. A. Martins. 2010. Humpback whale population estimates and distribution along the Brazilian breeding ground. Endangered Species Research 11:233–243.

Baker, C. S., A. Perry, J. L. Bannister, *et al*. 1993. Abundant mitochondrial DNA variation and worldwide population structure in humpback whales. Proceedings of the National Academy of Science of the United States of America 90:8239–8243.

Baker, C. S., L. Flórez-González, B. Abernethy, H. C. Rosenbaum, R. W. Slade, J. Capella and J. L. Bannister. 1998*a*. Mitochondrial DNA variation and maternal gene flow among humpback whales of the Southern Hemisphere. Marine Mammal Science 14:721–737.

Baker, C. S., L. Medranogonzalez, J. Calambokidis, *et al*. 1998*b*. Population structure of nuclear and mitochondrial DNA variation among humpback whales in the North Pacific. Molecular Ecology 7:695–707.

Baker, C. S., D. Steel, J. Calambokidis, *et al*. 2013. Strong maternal fidelity and natal philopatry shape genetic structure in North Pacific humpback whales. Marine Ecology Progress Series 494:291–306.

Bandelt, H. J., M. Richards and V. Macaulay. 2001. Detecting errors in mtDNA data by phylogenetic analysis. International Journal Legal Medicine 115:64–69.

Bandelt, H. J., L. Quintana-Murci, A. Salas and V. Macaulay. 2002. The fingerprint of phantom mutations in mitochondrial DNA data. The American Journal of Human Genetics 71:1150–1160.

Beerli, P. 2008. Migrate version 3.0 - a maximum likelihood and Bayesian estimator of gene flow using the coalescent. Available at http://popgen.sc.fsu.edu/Migrate/Migrate-n.html.

Bergl, R. A., and L. Vigilant. 2007. Genetic analysis reveals population structure and recent migration within the highly fragmented range of the Cross River gorilla (*Gorilla gorilla diehli*). Molecular Ecology 16:501–516.

Bérubé, M., H. Jorgensen, R. McEwing and P. J. Palsbøll. 2000. Polymorphic di-nucleotide microsatellite loci isolated from the humpback whale, *Megaptera novaeangliae*. Molecular Ecology 9:2181–2183.

Caballero, S., H. Hamilton, C. Jaramillo, *et al*. 2001. Genetic characterization of the Colombian Pacific coast humpback whale population using RAPD and mitochondrial DNA sequences. Memoirs of the Queensland Museum 47:459–464.

Capella, J., J. Gibbons, L. Flórez-González, M. Llano, C. Valladares, V. Sabaj and Y. Vilina. 2008. Migratory round-trip of individually identified humpback whales at the Strait of Magellan: Clues on transit times and phylopatry to destinations. Revista Chilena de Historia Natural 81:547–560.

Carroll, E., N. Patenaude, A. Alexander, *et al*. 2011. Population structure and individual movement of southern right whales around New Zealand and Australia. Marine Ecology Progress Series 432:257–268.

Carvalho, I., J. Loo, T. Collins, *et al*. 2014. Does temporal and spatial segregation explain the complex population structure of humpback whales on the coast of West Africa?. Marine Biology 161:805–819.

Clapham, P. J., Y. Mikhalev, W. Franklin, D. Paton, C. S. Baker, Y. V. Ivashchenko and R. L. Brownell, Jr.. 2009. Catches of humpback whales, *Megaptera novaeangliae*, by the Soviet Union and other nations in the Southern Ocean, 1947–1973. Marine Fisheries Review 71:39–43.

Constantine, R., K. Russell, N. Gibbs, S. Childerhouse and C. S. Baker. 2007. Photo-identification of humpback whales (*Megaptera novaeangliae*) in New Zealand waters and their migratory connections to breeding grounds of Oceania. Marine Mammal Science 23:715–720.

Cypriano-Souza, A. L., G. P. Fernández, C. A. V. Lima-Rosa, M. H. Engel and S. L. Bonatto. 2010. Microsatellite genetic characterization of the humpback whale (*Megaptera novaeangliae*) breeding ground off Brazil (Breeding Stock A). Journal of Heredity 101:189–200.

Darling, J. D., and S. Cerchio. 1993. Movement of a humpback whale, *Megaptera novaeangliae*, between Japan and Hawaii. Marine Mammal Science 9:84–89.

Dawbin, W. 1966. The seasonal migratory cycle of the humpback whale. Pages 145–170 *in* K. S. Norris, ed. Whales, dolphins and porpoises. University of California Press, Berkeley, CA.

Donovan, G. P. 1991. A review of IWC stock boundaries. Report of the International Whaling Commission (Special Issue 13):39–68.

Earl, D. A., and B. M. vonHoldt. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resource 4:359–361.

Engel, M. H., N. J. R. Fagundes, H. C. Rosenbaum, *et al*. 2008. Mitochondrial DNA diversity of the Southwestern Atlantic humpback whale (*Megaptera novaeangliae*) breeding area off Brazil, and the potential connections to Antartic feeding areas. Conservation Genetics 9:1253–1262.

Engel, M. H., and A. R. Martin. 2009. Feeding grounds of the western South Atlantic humpback whale population. Marine Mammal Science 25:964–969.

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

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Lynux and Windows. Molecular Ecology Resources 10:564–567.

Félix, F., D. M. Palacios, S. K. Salazar, S. Caballero, B. Haase and J. Falconí. 2011. The 2005 Galápagos Humpback Whale Expedition: A first attempt to assess and characterize the population in the Archipelago. Journal of Cetacean Research Management Special Issue 3:291–299.

Félix, F., S. Caballero and C. Olavarría. 2012. Genetic diversity and population structure of humpback whales (*Megaptera novaeangliae*) from Ecuador based on mitochondrial DNA analyses. Journal of Cetacean Research Management 12:71–77.

Flórez-González, L., J. Capella, B. Haase, G. A. Bravo, F. Félix and T. Gerrodette. 1998. Changes in winter destinations and the northernmost record of southeastern Pacific humpback whales. Marine Mammal Science 14:189–196.

Gabriele, C. M., C. Lockyer, J. M. Straley, C. M. Jurasz and H. Kato. 2010. Sighting history of a naturally marked humpback whale (*Megaptera novaeangliae*) suggests ear plug growth layer groups are deposited annually. Marine Mammal Science 26:443–450.

Garrigue, C., A. Aguayo-Lobo, V. L. U. Amante-Helweg, *et al*. 2002. Movements of humpback whales in Oceania, South Pacific. Journal of Cetacean Research Management 4:255–260.

Gibbons, J., J. Capella and C. Valladares. 2003. Rediscovery of a humpback whale (*Megaptera novaeangliae*) feeding ground in the Straits of Magellan, Chile. Journal of Cetacean Research Management 5:203–208.

Goudet, J. 2002. FSTAT: A program to estimate and test gene diversities and fixation indices. Institute of Ecology, Lausanne University, Lausanne, Switzerland.

Guo, S., and E. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiples alleles. Biometrics 48:361–372.

Hubisz, M. J., D. Falush, M. Stephens and J. K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resource 9:1322–1332.

Hucke-Gaete, R., D. Haro, J. P. Torres-Florez, Y. Montecinos, F. A. Viddi, L. Bedriñana and J. Ruiz. 2013. A historical feeding ground for humpback whales in the Eastern South Pacific revisited: The case of northern Patagonia, Chile. Aquatic Conservation: Marine and Freshwater Ecosystems 23:858–867.

IWC (International Whaling Commission 2015). Report of the Scientific Committee. Annex H: Report of the sub-committee on other Southern Hemisphere whale stocks. Journal of Cetacean Research Management 16:196–221.

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

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

Lambeck, K., T. M. Esat and E.-K. Potter. 2002. Links between climate and sea levels for the past three million years. Nature 419:199–206.

Lambertsen, R. H. 1987. A biopsy system for large whales and its use for cytogenetics. Journal of Mammalogy 68:443–445.

Lecis, R., M. Pierpaoli, Z. S. Biro, L. Szemethy, B. Ragni, F. Vercillo and E. Randi. 2006. Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. Molecular Ecology 15:119–131.

Librado, P., and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.

Lodi, L. 1994. Ocorrências de baleias-jubarte, *Megaptera novaeangliae*, no Arquipélago de Fernando de Noronha, incluindo um resumo de registros de capturas no Nordeste do Brasil [Sightings of humpback whales, *Megaptera novaeangliae*, in the Fernando de Noronha Archipelago, including a summary of catch records in northeastern Brazil]. Biotemas 7:116–123.

Mackintosh, N. A. 1965. The stocks of whales. Fishing News (Books) Ltd., London, U.K.

Manel, S., P. Berthier and G. Luikart. 2002. Detecting wildlife poaching: Identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. Conservation Biology 16:650–659.

Matthews, L. H. 1937. The humpback whale, *Megaptera nodosa*. Discovery Reports 17:7–92.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590.

Olavarría, C., A. Aguayo, J. Acevedo, L. Medrano, D. Thiele and C. S. Baker. 2006. Genetic differentiation between two feeding areas of the Eastern South Pacific humpback whale population: Update on SC/57/SH3. Unpublished paper SC/A06/HW29 presented to the Scientific Committee of the International Whaling Commission.

Olavarría, C., C. S. Baker, C. Garrigue, *et al*. 2007. Population structure of humpback whales throughout the South Pacific and the origin of the eastern Polynesian breeding grounds. Marine Ecology Progress Series 330:257–268.

Paetkau, D., R. Slade, M. Burden and A. Estoup. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: A simulation-based exploration of accuracy and power. Molecular Ecology 13:55–65.

Palsbøll, P. J., P. J. Clapham, D. K. Mattila, *et al*. 1995. Distribution of mtDNA haplotypes in North Atlantic humpback whales: the influence of behaviour on population structure. Marine Ecology Progress Series 116:1–10.

Palsbøll, P. J., M. Bérubé and H. Jorgensen. 1997. Primers for the amplification of tri- and tetramer microsatellite loci in baleen whales. Molecular Ecology 6:893–895.

Pearse, D. E., and K. A. Crandall. 2004. Beyond F_{ST}: Analysis of population genetic data for conservation. Conservation Genetics 5:585–602.

Pichler, F. B., D. Robineau, R. N. P. Goodall, M. A. Meyer, C. Olavarría and C. S. Baker. 2001. Origin and radiation of Southern Hemisphere coastal dolphins (genus *Cephalorhynchus*). Molecular Ecology 10:2215–2223.

Piry, S., A. Alapetite, J. M. Cornuet, D. Paektau, L. Baudouin and A. Estoup. 2004. GeneClass2: A software for genetics assignment and first generation migrant detection. Journal of Heredity 95:536–539.

Pritchard, J. K., X. Wen and D. Falush. 2007. Documentation for Structure software: version 2.2. Available at http://pritch.bsd.uchicago.edu/software.

Rannala, B., and J. L. Mountain. 1997. Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences of the United States of America 94:9197–9221.

Rasmussen, K., D. Palacios, J. Calambokidis, *et al*. 2007. Southern Hemisphere humpback whales wintering off Central America: Insights from water temperature into the longest mammalian migration. Biology Letters 3:302–305.

Rice, W. R. 1989. Analyzing table of statistical tests. Evolution 43:223–225.

Rosenbaum, H. C., C. Pomilla, M. Mendez, *et al*. 2009. Population structure of humpback whales from their breeding grounds in the South Atlantic and Indian Oceans. PLoS ONE 4:e7318.

Rosenberg, N. A. 2004. DISTRUCT: A program for the graphical display of population structure. Molecular Ecology Notes 4:137–138.

Schlötterer, C., B. Amos and D. Tautz. 1991. Conservation of polymorphic simple sequence loci in cetacean species. Nature 354:63–65.

Schmitt, N. T., M. C. Double, S. N. Jarman, *et al*. 2014. Low levels of genetic differentiation characterize Australian humpback whale (*Megaptera novaeangliae*) populations. Marine Mammal Science 30:221–241.

Slijper, E. J. 1962. Whales. 1st English edition. Hutchinson & Co., London, U.K.

Slijper, E. J. 1965. A hundred years of modern whaling. Netherlands Commission for International Nature Protection, Amsterdam, Netherlands.

Sremba, A. L., A. R. Martin and C. S. Baker. 2014. Species identification and likely catch time period of whale bones from South Georgia. Marine Mammal Science 31:122–132.

Steel, D., C. Garrigue, S. Caballero, *et al*. 2008. Migratory connections between humpback whales from South Pacific breeding grounds and Antarctic feeding areas based on genotype matching. Unpublished paper SC/60/SH13 presented to the Scientific Committee of the International Whaling Commission.

Stevick, P. T., A. Aguayo, J. Allen, *et al*. 2004. A note on the migrations of individually identified humpback whales between the Antarctic Peninsula and South America. Journal of Cetacean Research Management 6:109–113.

Stevick, P. T., J. Allen, P. J. Clapham, *et al*. 2006*a*. Population spatial structuring on the feeding grounds in North Atlantic humpback whales (*Megaptera novaeangliae*). Journal of Zoology 270:244–255.

Stevick, P. T., L. Pacheco de Godoy, M. McOsker, M. H. Engel and A. Allen. 2006*b*. A note on the movement of a humpback whale from Abrolhos Bank, Brazil to South Georgia. Journal of Cetacean Research Management 8:297–300.

Stevick, P. T., M. C. Neves, F. Johansen, M. H. Engel, J. Allen, M. C. C. Marcondes and C. Carlson. 2010. A quarter of a world away: Female humpback whale moves 10,000 km between breeding areas. Biology Letters 7:299–302.

Stevick, P. T., J. M. Allen, M. H. Engel, F. Félix, B. Haase and M. C. Neves. 2013. Inter-oceanic movement of an adult female humpback whale between Pacific and Atlantic breeding grounds off South America. Journal of Cetacean Research Management 13:159–162.

Stone, G. S., L. Flórez-González and S. Katona. 1990. Whale migration record. Nature 346:705–706.

Tajima, F. 1989. Evolutionary relationship of DNA sequences in finite populations. Genetics 105:437–460.

Taylor, B. L., S. J. Chivers, J. Larese and W. F. Perrin. 2007. Generation length and percent mature estimates for IUCN assessments of cetaceans. Administrative Report LJ-0701, National Marine Fisheries Service, Southwest Fisheries Science Center, La Jolla, CA.

Torres-Florez, J. P., R. Hucke-Gaete, R. LeDuc, *et al*. 2014. Blue whale population structure along the eastern South Pacific Ocean: Evidence of more than one population. Molecular Ecology 23:5998–6010.

Tønnessen, J. N., and A. O. Johnsen. 1982. The history of modern whaling. University of California Press, Berkeley, CA.

Valsecchi, E., and W. Amos. 1996. Microsatellite markers for the study of cetacean populations. Molecular Ecology 5:151–156.

Valsecchi, E., P. Hale, P. Corkeron and W. Amos. 2002. Social structure in migrating humpback whales (*Megaptera novaeangliae*). Molecular Ecology 11:507–518.

Valsecchi, E., P. J. Corkeron, P. Galli, W. Sherwin and G. Bertorelle. 2010. Genetic evidence for sex-specific migratory behaviour in western South Pacific humpback whales. Marine Ecology Progress Series 398:275–286.

Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills and P. Shipley. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatelite data. Molecular Ecology Notes 4:535–538.

Waits, L. P., G. Luikart and P. Taberlet. 2001. Estimating the probability of identity among genotypes in natural populations: Cautions and guidelines. Molecular Ecology 10:249–256.

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

Wedekin, L. L., M. Rossi-Santos, C. Baracho, A. L. Cypriano-Souza and P. C. Simões-Lopes. 2014. Cetacean records along a coastal-offshore gradient in the Vitória-Trindade Chain, western South Atlantic Ocean. Brazilian Journal of Biology 74:137–144.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population Structure. Evolution 38:1358–1370.

Wenzel, F. W., J. Allen, S. Berrow, *et al*. 2009. Current knowledge on the distribution and relative abundance of humpback whales (*Megaptera novaeangliae*) off the Cape Verde Islands, Eastern North Atlantic. Aquatic Mammals 35:502–510.

Whitlock, M. C., M. A. McPeek, M. D. Rausher, L. Rieseberg and A. J. Moore. 2010. Data archiving. The American Naturalist 175:145–146.

Wilberg, M. J., and B. P. Dreher. 2004. GENECAP: A program for analysis of multilocus genotype data for non-invasive sampling and capture-recapture population estimation. Molecular Ecology Notes 4:783–785.

Wilson, G. A., and B. Rannala. 2003. Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163:1177–1191.

Zerbini, A. N., A. Andriolo, M. P. Heide-Jørgensen, *et al*. 2006. Satellite-monitored movements of humpback whales *Megaptera novaeangliae* in the Southwest Atlantic Ocean. Marine Ecology Progress Series 313:295–304.

Zerbini, A. N., A. Andriolo, M. Peter, *et al*. 2011*a*. Migration and summer destinations of humpback whales (*Megaptera novaeangliae*) in the western South Atlantic Ocean. Journal of Cetacean Research Management Special Issue 3:113–118.

Zerbini, A. N., E. J. Ward, P. G. Kinas, M. H. Engel and A. Andriolo. 2011*b*. A Bayesian assessment of the conservation status of humpback whales (*Megaptera novaeangliae*) in the western South Atlantic Ocean. Journal of Cetacean Research Management (Special Issue) 3:131–144.

Received: 18 September 2015
Accepted: 5 October 2016

Supporting Information

The following supporting information is available for this article online at http://onlinelibrary.wiley.com/doi/10.1111/mms.12378/suppinfo.

Appendix S1. Description of the validation of mtDNA control region haplotypes from BSA.

Appendix S2. Description of the calibration of the microsatellite allele size in different machines.

Table S1. Summary of the humpback whales studied in two breeding grounds (Brazil and Colombia) and one feeding area (Antarctic Peninsula), including sampling years, number of samples analyzed for each marker (mtDNA and microsatellites), and the final number of individuals sequenced, genotyped and sexed after replicates removal. All samples from Colombia and Antarctic Peninsula were analyzed for both markers, but for the 277 samples from Brazil, 143 were only sequenced, 119 were only genotyped, and 15 were analyzed for both markers.

Table S2. Positions of the 67 variable sites defining 77 mtDNA control region haplotypes (465 bp) detected in 370 humpback whales sampled in two breeding ground (BSA, Brazil, n = 158; BSG, Colombia, n = 130) and one feeding ground (AP, Antarctic Peninsula, n = 82). Position 1 of alignment corresponds to position 22 in Engel *et al*. (2008) and Rosenbaum *et al*. (2009). Dots (.) indicate matches with reference sequence (HBA002 or SP1) and dashes (–) indicate insertion/deletion. The frequencies of haplotypes are shown for each of the three sampling regions. Haplotype nomenclature follows Olavarría *et al*. (2006, 2007), Engel *et al*. (2008) and Rosenbaum *et al*. (2009). GenBank accession numbers of haplotypes previously submitted are: AY329844 to AY330096, GQ913691 to GQ913852, and DQ768307 to DQ768421.

Table S3. Summary of microsatellite diversity for the 16 loci analyzed for humpback whales sampled in the two breeding grounds (Brazil and Colombia) and one feeding area (Antarctic Peninsula, AP). Includes number of individuals (n), number of alleles (K), allelic richness (AR), observed (H_O) and expected (H_E) heterozygosities, and inbreeding coefficient (F_{IS}) for each locus. Probabilities of identity ($P_{(ID)}$ and $P_{(ID)sib}$) for each population is also included. Duplicate samples were removed from the analysis.