Does temporal and spatial segregation explain the complex population structure of humpback whales on the coast of West Africa?

Inês Carvalho, Jacqueline Loo, Timothy Collins, Jaco Barendse, Cristina Pomilla, Matthew S. Leslie, Solange Ngouessono, Peter B. Best, Howard C. Rosenbaum

Abstract Humpback whales (*Megaptera novaeangliae*) in the Southeastern Atlantic Ocean (International Whaling Commission 'Breeding Stock B'—BSB) are distributed from the Gulf of Guinea to Western South Africa. Genetic data suggest that this stock may be sub-structured, but it remains unknown if this is due to reproductive segregation. This paper evaluates the spatial and temporal population structure of BSB humpback whales using a combination of maternally and bi-parentally inherited markers. The genetic differentiation that we identify in this study could be due to a combination of (1) spatial and/or temporal segregation on breeding grounds in the greater Gulf of Guinea, (2) the possibility of maternally inherited site fidelity to specific feeding grounds and (3) the use of two generalized but exclusive migratory routes (coastal and offshore) between feeding and breeding areas. Further, photo-identification and genetic sampling efforts in other areas of the Sub-Saharan Western Africa winter range and targeted deployment of satellite tags would help to clarify some of the apparent complexity in the population structure of animals biopsied in this region.

Introduction

Understanding the processes by which marine populations diverge or remain homogeneous remains a major challenge

I. Carvalho
Faculdade de Ciências e Tecnologia, Universidade do Algarve, Campus Gambelas, 8000-139 Faro, Portugal
e-mail: carvalho.inesc@gmail.com

I. Carvalho · J. Loo · C. Pomilla · M. S. Leslie · H. C. Rosenbaum
Sackler Institute for Comparative Genomics, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA
e-mail: hrosenbaum@wcs.org

I. Carvalho · J. Loo · T. Collins · S. Ngouessono · H. C. Rosenbaum
Ocean Giants Program, Wildlife Conservation Society, 2300 Southern Blvd., Bronx, NY 10460-1099, USA

J. Loo
Department of Biology, New York University, 100 Washington Square, New York, NY 10012, USA

T. Collins
Environment Society of Oman, PO Box 3955, PC 112 Ruwi, Sultanate of Oman

J. Barendse · P. B. Best
Mammal Research Institute, University of Pretoria, c/o Iziko South African Museum, P.O. Box 61, Cape Town 8000, South Africa

J. Barendse
Sustainability Research Unit/SANParks, Nelson Mandela Metropolitan University, Private Bag X6530, George 6530, South Africa

C. Pomilla
Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

M. S. Leslie
Scripps Institution of Oceanography, University of California San Diego, 9500 Gilman Dr. Mailcode 0202, La Jolla, CA 92038, USA

S. Ngouessono
Agence Nationale des Parcs Nationaux, Batterie 4, BP 20379, Libreville, Gabon

in evolutionary biology. The accurate description of population genetic structure over a species geographic range has great importance for effective management and conservation and can inform us about the past and present evolutionary dynamics of a species. In terrestrial environments, populations are frequently well defined and distinct from each other, often physically separated by barriers to mixing and interbreeding (e.g. Avise 2000). Spatial relationships between populations are a consequence of historical biogeography, evolutionary relatedness, and ongoing gene flow. Differences in dispersal ability and geographic distribution are important influences on patterns of phylogeographic structure. Generally, species that have little dispersal potential exhibit pronounced spatial genetic structure, while those that have great dispersal potential show less spatial structure (Ward et al. 1994; Waples 1998). In the marine environment, physical barriers to dispersal are often absent, and one may expect species with high dispersal capacities to demonstrate less population structure (Palumbi 1994). However, recent genetic studies have revealed population genetic structure in highly mobile marine species (e.g. Shaw et al. 1999; Chow et al. 2000; Knutsen et al. 2003; Bowen and Karl 2007; André et al. 2011). Differences in behaviour within species, such as the timing of migration, the level of site fidelity and sex-specific philopatry, can result in the accumulation of genetic differentiation even in highly mobile marine species. Information on the spatial and temporal patterns of population structure can improve our understanding of underlying intrinsic and extrinsic processes (e.g. Palsbøll et al. 2004; Fontaine et al. 2007).

Population sub-structure is found in a wide variety of species and ecosystems and can be of elementary importance for conservation and management (Ruckstuhl and Neuhaus 2005). Many cetacean species have worldwide distributions but are composed of distinct populations inhabiting different oceanic regions and hemispheres (Rice 1998). In this group, sex or age segregation is broadly observed (Chittleborough 1965; Brown et al. 1995; Lyrholm et al. 1999; Whitehead 2003; Loseto et al. 2006), and most assessments of these point to adaptive advantages, including social structure, environmental constraints or habitat requirements associated with particular life-history events (Laidre et al. 2009).

Southern Hemisphere humpback whales (*Megaptera novaeangliae*) undertake annual migrations between high-latitude waters, where they feed during the summer, and low-latitude waters, where they breed during winter months (Dawbin 1966). In the Southern Ocean, their winter distribution is divided into six longitudinal sectors surrounding the Antarctic, termed Areas I–VI (Mackintosh 1942) and used for sub-population identity by the International Whaling Commission (IWC). Seasonal distribution data and a limited number of 'discovery tag' recoveries between these high-latitude feeding areas and low-latitude breeding grounds led to the designation of seven breeding grounds termed breeding stocks A–G (IWC 1998). In the eastern South Atlantic Ocean (breeding stock B), whales migrate past the west coast of South Africa (Matthews 1938) to wintering areas in what we term here the 'greater Gulf of Guinea', including Angola, the Congo's and Gabon (Mackintosh 1942; Budker and Collignon 1952; Walsh et al. 2000). The most recent available genetic data (mitochondrial DNA and microsatellites) suggest that there is some stock structure in this region: whales that overwinter in the greater Gulf of Guinea, here called the Tropical Southeast Atlantic (TSA), differ genetically from those sampled off the West coast of South Africa (WSA; Rosenbaum et al. 2009; Pomilla 2005). However, as the latter area appears to serve primarily as a migratory corridor and a feeding site, the origins of this difference, as well as the actual breeding site for a potential second sub-stock, are unknown (IWC 1998; Rosenbaum et al. 2009; Barendse et al. 2010). Previous molecular studies using mitochondrial DNA only found indications of population sub-structuring, based on haplotype frequencies and low gene flow rates between the two regions (Rosenbaum et al. 2009). Nuclear DNA analysis (Pomilla 2005) supported in part this division, but showed some connectivity between the two areas, reported as two genetic matches between Gabon (TSA) and WSA: further, photographic matching has resulted in a total of five matches between these two regions (Barendse et al. 2011). On the West South African coast, recent observations indicate a regular and extended presence of whales as late as March including direct observations of feeding on euphausiids during the austral spring and summer months. Data from early twentieth century catches during the summer season off Namibia suggest that this area, besides functioning as a migratory corridor, also serves as an important feeding ground for some members of the population (Best et al. 1995; Findlay 2000; Barendse et al. 2010, 2011, 2013).

A number of questions regarding the definition of Breeding Stock B, the population structure and the precise geographic area it utilizes, remain unanswered. Possible explanations could be attributed to temporal and/or spatial structuring of the stock. Evidence of temporal and geographic sub-structure has been found in other wintering grounds (Smultea 1994; Medrano-González et al. 1995; Dawbin 1997). There can also be significant differences in the timing of arrival and departure of whales from a breeding area, based on their reproductive status (Dawbin 1997), and the temporal distribution of matrilines at breeding grounds (Medrano-González et al. 1995).

Here we present data on microsatellite and mitochondrial DNA variation in order to evaluate how the temporal and spatial distribution of humpback whales along the West coast of Africa can influence underlying population

3

structure. Using a considerably larger sample size than previously published and a combination of maternally and biparentally inherited markers, we provide additional evidence for the complexity of population structure and gene flow of humpback whales on the West coast of Africa.

Materials and methods

Sample collection, DNA extraction and sex determination

Samples were collected from humpback whales in multiple years at four sampling sites: São Tomé and Príncipe (2004–2005, $N = 5$), Gabon (2000–2006, $N = 1,696$), Angola

(Cabinda Province; 1998, $N = 13$), and WSA (1990–2009, $N = 304$) for a total of 2,018 samples (Fig. 1). The present data set includes the mitochondrial DNA and microsatellite data (Gabon data between 2000 and 2002, Angola 1998, and WSA between 1990 and 2003), used in Rosenbaum et al. (2009) and Pomilla (2005), respectively.

Skin tissues were mostly obtained using biopsy darts (Lambertsen 1987), but also included sloughed skin and samples from stranded specimens. Samples were preserved in 95 % Ethanol or salt saturated 20 % Dimethyl Sulfoxide solution (DMSO) and later stored at −20 °C until processed. Total genomic DNA was extracted from the tissue samples using proteinase K digestion, followed by a standard Phenol/Chloroform extraction method (Sambrook et al.

Fig. 1 Area map depicting the sampling localities (*bold*) and the sampling sizes (*N*) included in the genetic analyses
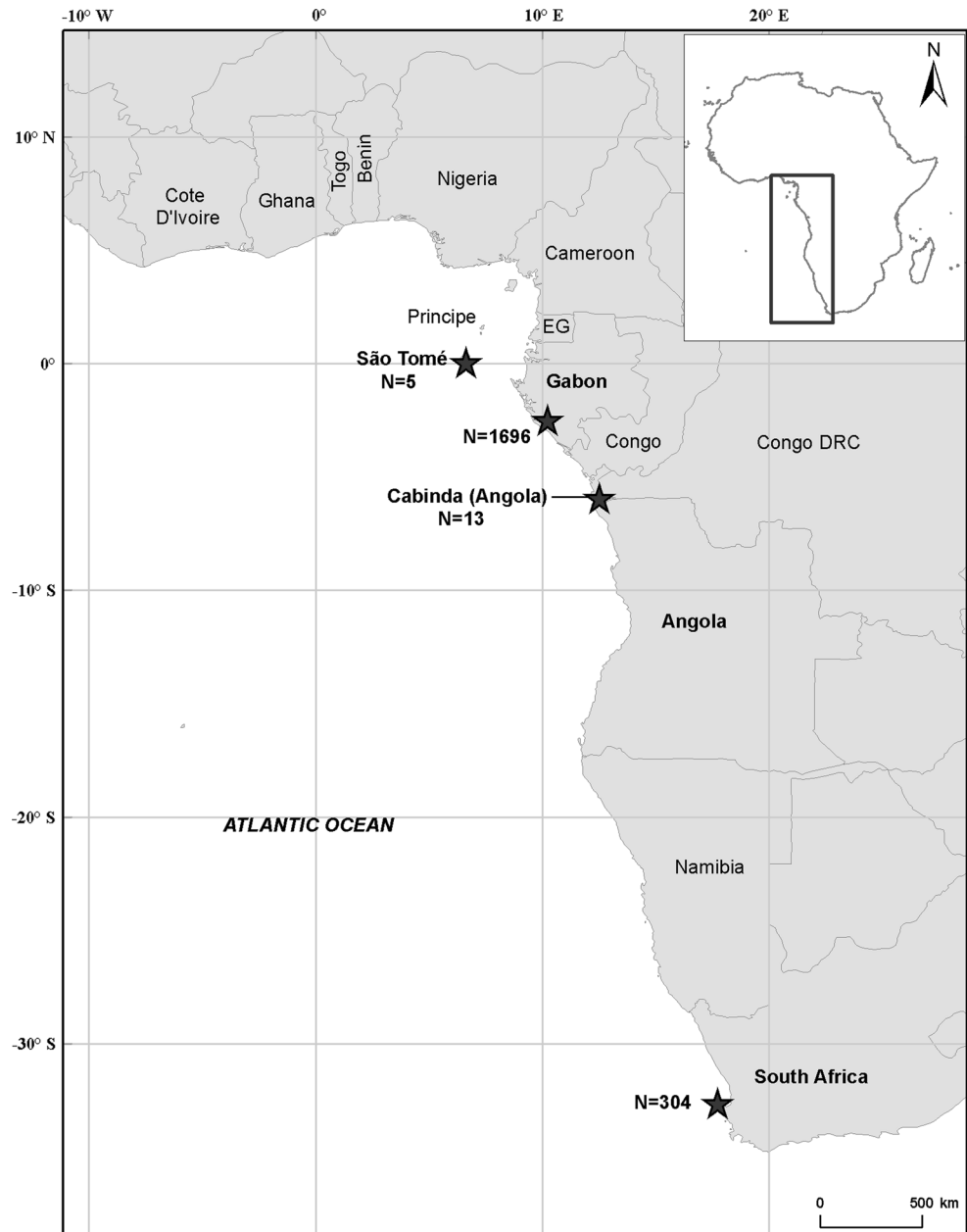
1989) or using QIAamp Tissue Kit (QiaGen) following the manufacturer's protocol.

Sex determination was either carried out by Polymerase Chain Reaction (PCR) amplification followed by TaqI restriction endonuclease digestion of the ZFX/ZFY region of the sex chromosomes (Palsbøll et al. 1992), or using multiplex PCR amplification of the ZFX/ZFY sex linked gene (Berubé and Palsbøll 1996).

Mitochondrial DNA and microsatellite molecular analyses

A 520 bp segment within the mtDNA control region was amplified by PCR, with primers Dlp 1.5 and Dlp 5 as described by Baker et al. (1993). Reactions of 25 µL total volume, containing 50 mM KCl, 10 mM Tris–HCl pH 8.8, 2.5 mM $MgCl_2$, 200 µM of each dNTP, 1.0 µM of each primer, and 0.05 U/µl Taq Gold polymerase, were conducted either on a Perkin–Elmer Thermocycler or an Eppendorf Gradient Mastercycler, following standard PCR procedures (94 °C for 4 min, followed by 30 cycles of 94 °C denaturing for 45 s, 54 °C annealing for 45 s, and 72 °C extension for 45 s). Amplified PCR products were cycle sequenced with dye-labelled terminators in both directions. Reactions were run on a 3700 or 3730 DNA Analyzer (Applied Biosystems, Inc.). PCR amplifications included negative control reactions to check for exogenous contamination.

Ten microsatellite loci, proven to be polymorphic in humpback whales, were used in this study: GATA028, GATA053, GATA417 (Palsbøll et al. 1997), 199/200, 417/418, 464/465 (Schlötterer et al. 1991), EV1Pm, EV37Mn, EV94Mn, EV96Mn (Valsecchi and Amos 1996). The 5′-end of the forward primer from each locus was labelled with a fluorescent tag (HEX, 6-FAM, and TET, Qiagen-Operon; NED, Applied Biosystems, Inc). PCRs were carried out in a 20 or 10 µl volume with the following conditions: 50 mM KCl, 10 mM Tris–HCl pH 8.8, 2.5–3.5 mM $MgCl_2$, 200 µM of each dNTP, 0.4 µM of each primer, and 0.025 U/µl Taq Gold polymerase (Perkin–Elmer). Amplifications were completed in either a Perkin–Elmer 9600 Thermocycler or an Eppendorf Gradient Mastercycler, after optimization of published annealing temperatures and profiles. PCR products were loaded with the addition of an internal standard ladder (GS600 LIZ, ABI) on a 3700 or 3730 DNA Analyzer (Applied Biosystems, Inc). Microsatellite alleles were identified by their sizes in base pairs using the software GENEMAPPER 4.0 software (ABI).

Data analysis

Error checking and duplicate samples

Specific guidelines were used during laboratory work and scoring procedures to reduce genotyping errors. Automation was introduced whenever possible during PCR setup and manipulation of genomic DNA or PCR products. Negative controls were run at the PCR step to control for exogenous contamination. Two reference samples of known allele size were added to each amplification and subsequent analyses to standardize scoring. Scoring was automated in GENEMAPPER, and allele sizing was systematically checked by hand. Samples that yielded ambiguous allele peaks were repeated a second time. Genotyping error was checked by re-amplifying and re-typing 15 % of the total, chosen at random. In order to detect errors in our dataset, such as identifying possible non-amplified alleles (null alleles), large allele dropout and scoring errors due to stutter peaks, we used the programs DROPOUT 1.3 (McKelvey and Schwartz 2005) and MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004). Overall, 10 cases of allele dropout were detected and resolved by duplicate genotyping.

Duplicate samples, within each population and year, were detected from microsatellite genotype identity using MS_Tools (Park 2001) and were subsequently excluded. The probability of different individuals in each population sharing the same genotype by chance (Probability of Identity, P_{ID}) was estimated using GENALEX 6.5 (Peakall and Smouse 2012). Samples which had information of less than 8 loci were excluded from subsequent analysis.

Diversity estimates

From the 520 bp mtDNA control region fragment, a 477 bp consensus region, containing the majority of variable nucleotide positions in the mtDNA control region of humpback whales, was examined for each sample (Baker et al. 1993). Sequences were aligned and assembled using SEQUENCER 4.5 (Gene Codes, Inc). DNA sequence variation patterns were characterized into mtDNA haplotype definitions following the nomenclature developed in Rosenbaum et al. (2002) and Rosenbaum et al. (2009). Matching of sequences to haplotypes was done using DNASP 5.0 (Rozas et al. 2003). Levels of mtDNA polymorphism within the sample set were measured by estimating haplotype (h) and nucleotide (π) diversities (Nei 1987), using ARLEQUIN 3.5 (Excoffier and Lischer 2010).

For microsatellite data, genetic diversity was measured as the observed heterozygosity (Ho) and expected heterozygosity (He) under Hardy–Weinberg assumptions (Nei 1987) using GENALEX software. Evaluation of possible deviations from Hardy–Weinberg (HW) equilibrium and linkage disequilibrium (LD) was performed using Fisher's Exact test and Markov chain methods (iterations per batch set at 10,000; Guo and Thompson 1992) implemented in GENEPOP 4.0 (Raymond and Rousset 1995). Significance levels ($P = 0.05$) for departure from HW and LD

were corrected for multiple comparisons with a Bonferroni correction (Rice 1989). Allelic and private allelic richness within each population (and by season) were estimated using rarefaction implemented using the software HP-Rare 1.1 (Kalinowski, 2005), to account for variance in samples size among regions.

Analysis of population structure

For spatial structure analysis, we tested the TSA (combined samples from Gabon, São Tomé, and Angola) and WSA (samples from the West coast of South Africa), using population divisions suggested by Pomilla et al. (2006) and Rosenbaum et al. (2009).

For the temporal analysis, data were partitioned according to seasonal humpback whale presence in each area, taking into account the data available for these regions and the sampling months. As such, only Gabon and WSA data were suitable for use. In Gabon, the whales are present during the breeding season from July to November (Rosenbaum and Collins 2006). Whales sampled off the coast of Gabon were divided into 'Early' season (E_GAB), including samples collected in July and August, and 'Late' season (L_GAB), with samples collected in September and October. For WSA, the whales pass by during the northward (July/August) and southward (October/November) migrations, while some whales are present in the area until late summer (February; Barendse et al. 2010). Here, the season was divided into 'Regular' season (R_WSA), including samples collected between July and October, and 'Late' season (L_WSA), with samples collected between November and February (late Spring and Summer). In this way, the regular season in South Africa corresponded to the entire season in Gabon. Note that the term 'season' is used hereafter to refer to these seasonal groupings, unless stated otherwise. The dataset was further stratified to evaluate the effect of sex (males, females).

The degree of geographic (TSA vs WSA) and temporal (time of the season in each region) population subdivision for mitochondrial DNA was assessed using Φ_{ST} and F_{ST}. Φ_{ST} takes into account the relationships between haplotype based on molecular distance (Excoffier et al. 1992), whereas F_{ST} considers only the difference in overall haplotype frequencies (Weir and Cockerham 1984). The significance of the observed Φ or F statistic was tested using a null distribution generated from 10,000 nonparametric random permutations of the data. Chi square statistics have been shown to have greater power than sequence-based statistics for detecting population structure (Hudson et al. 1992). To test for a random distribution of individuals between the two populations, we conducted an Exact Test of population differentiation. The significance of this test is an indication of the non-random association of individuals among populations for both global and pairwise comparisons. All of these tests were performed in ARLEQUIN. No correction for multiple tests was applied to significant levels of pairwise comparisons (Perneger 1998; Narum 2006).

For microsatellites, three genetic metrics were used. The distances between microsatellite genotypes were estimated either by counting the number of different alleles between two genotypes, the equivalent of estimating weighted F_{ST} over all loci, or by counting the sum of the squared number of repeat differences between two haplotypes, the equivalent of estimating R_{ST} over all loci (Slatkin 1995; Michalakis and Excoffier 1996), both metrics were performed using ARLEQUIN. Additionally, D_{est} (Jost 2008) was calculated using SWAP 0.9 software (http://www.uni-marburg.de/fb17/fachgebiete/naturschutz/naturschutzbiologie/downloads?language_sync=1), which evaluates population differentiation by quantifying genetic diversity within populations based on the effective number of alleles (using a rarefaction method).

A Bayesian model-based clustering approach was implemented in order to identify the number of populations (K) present in the dataset using the software STRUCTURE 2.3.4 (Pritchard et al. 2000; Hubisz et al. 2009), which applies a Markov Chain Monte Carlo (MCMC) algorithm. An admixture model, as well as a correlated allele frequency model, was used with prior information on the origin of samples (TSA and WSA). The maximum number of populations (K) was assumed to vary between 1 and 5. For each potential value of K, 6 independent runs were performed. Burn-in length and length simulation were set at 500,000 and 5,000,000 steps, respectively. The most likely number of clusters (K) was chosen as that showing the maximum estimated mean log-likelihood of the data (LnP(D)) (Pritchard et al. 2000) and confirmed by calculating ΔK, the second-order rate of change of LnP(D) with respect to K after Evanno et al. (2005) using STRUCTURE HARVESTER (Earl and VonHoldt 2011).

Measures of migration rates, dispersal and connectivity

Effective migration rates and divergence time estimates between the two regions were obtained using the mitochondrial DNA data set and a maximum likelihood framework based on coalescence theory, implemented in the software MDIV (Nielsen and Wakeley 2001). MDIV simultaneously estimates the migration rate per gene per generation between populations, scaled by effective population size ($M = 2N_{ef}m$), the divergence time scaled by the effective population size ($T = t/2N_{ef}$) and the parameter theta ($\theta = 4N_{ef}$). These parameters were obtained using a finite sites model (HKY) to allow for the possibility of multiple mutations per site. We ran 5×10^6 cycles with a burn-in of 5×10^5 to minimize dependence on initial conditions.

The maximum value for the theta parameter was 0. The choice of $\theta = 0$ provides the model with a flat prior hypothesis, which has the least influence on parameter estimation. Maximum values for T and M were set at 3 and 45, respectively. Five converged runs with different random seeds were used to identify M, T, and θ values corresponding to the maximum likelihood.

Assignment/exclusion of individuals using predefined subpopulations and possible first generation migrants was assessed using GENECLASS 2.0 (Piry et al. 2004), which does not assume that all potential source populations have been sampled. A Bayesian-based method for computation and a resampling algorithm for probability computation with 10,000 simulated individuals were chosen (Rannala and Mountain 1997). This method tests the null hypothesis that an individual was born into the population where it was sampled by estimating the probability of drawing that individual's multilocus genotype from the allele distributions observed in a series of study populations. When compared with other assignment methods, the Bayesian method described by Rannala and Mountain (1997) allows recent immigrants to be identified on the basis of their multilocus genotype and has considerable power for detecting immigrants, even when the overall level of genetic differentiation among populations is low (Cornuet et al. 1999). Statistical thresholds for the confidence of excluding a population as the origin of an individual whale were established with Monte Carlo resampling methods with $P < 0.01$ for excluding populations.

Evidence of sex-biased dispersal was tested using the program FSTAT 2.9.3 (Goudet 2001) and included examination of differences in mean relatedness between same-sex pairs F_{IS}, sex-specific F_{ST}, and sex-specific mean-corrected assignment indices (AIc). A randomized approach with 10,000 permutations was employed to test whether the difference in dispersal between the sexes was significant. In the case of sex-biased dispersal, the sex with greater dispersal tendencies is predicted to show a lower mean AIc than the more philopatric sex (Favre et al. 1997; Mossman and Waser 1999). This genetic test has been shown to perform well even with species that have no extreme sex-biased dispersal tendencies (Mossman and Waser 1999).

Movements of specific individuals between different areas are suggestive of current interchange. To document such movements, we used a genetic capture–recapture approach based on the attainment of unique individual genetic profiles and searched for genotype matches between different areas (Waits et al. 2001). GENALEX was used for this analysis and applied only to those samples for which we have information on 9 or 10 loci.

Results

Sample size and sex ratio

Location and sample sizes for each sampling site and for the divisions described above are illustrated in Fig. 1. The 2,018 analysed tissue samples were determined to represent 1,640 different whales: 1,421 individuals in region TSA and 219 in region WSA. The average probability of identity $P_{(ID)}$ for each population was small enough to exclude duplicate samples with high confidence (Table 1). The overall resampling rate (proportion of known individuals re-sampled in the same area in different years) was significantly lower ($\chi^2 = 29.86$, $P < 0.000$, Yates correction) for TSA (6.61 %) than for WSA (17.56 %).

Sex was determined for 1,600 individuals, resulting in 938 males and 449 females for TSA and 103 males and 110 females for WSA. For TSA, the observed overall proportion of 2.1:1 males to females differed significantly ($\chi^2 = 171.7$, $\alpha = 0.000$, $df = 1$) from a 1:1 sex ratio; for WSA (1:1.1 males to females) there was no significant deviation from parity.

Genetic diversity

A consensus region of 477 bp of the mitochondrial DNA control region was assembled and 135 maternal haplotypes were detected based on a total of 84 polymorphic sites (Table 1). Two new haplotypes (all WSA samples) were found that had not been previously described (Rosenbaum et al. 2009). Sequences representing these new haplotypes have been deposited in GenBank (Accession numbers

Table 1 Genetic diversity in humpback whales based on mitochondrial DNA and 10 microsatellite loci for two regions: TSA—comprising samples from Gabon, São Tomé and Angola and WSA—samples from the west coast of South Africa

Region	Mitochondrial DNA					Microsatellites		
	H	P	PS	h	π	H_o	H_e	P(ID)
TSA	129	42	81	0.980 (0.000)	0.021 (0.011)	0.743 (0.004)	0.745 (0.049)	1.9×10^{-12}
WSA	71	3	64	0.971 (0.004)	0.021 (0.011)	0.751 (0.008)	0.740 (0.051)	3.2×10^{-12}

H number of haplotypes, P number of private haplotypes, PS polymorphic sites, h haplotype diversity, π nucleotide diversity, H_o observed heterozygosity, H_e expected heterozygosity, PI probability of identity. Values in parentheses are standard deviation

Table 2 Allelic richness and private allelic richness (values in parentheses) estimated from a rarefaction algorithm calculated over the 10 microsatellite loci, for the spatial and temporal divisions

Locus	Spatial		Temporal			
	TSA	WSA	E_GAB	L_GAB	R_WSA	L_WSA
28	12.60 (2.87)	10.57 (0.83)	12.28 (0.92)	11.28 (0.94)	10.95 (0.05)	10.50 (0)
53	12.09 (1.43)	11.39 (0.73)	11.66 (0.24)	12.21 (0.51)	10.89 (0.10)	10.79 (0.08)
199–200	7.34 (0.57)	7.42 (0.65)	7.09 (0.12)	7.25 (0.45)	6.96 (0.09)	7.21 (0.31)
417	16.96 (2.64)	17.23 (2.91)	16.68 (1.48)	15.64 (0.74)	14.93 (0.07)	16.30 (1.31)
417–418	10.88 (0.77)	11.53 (1.42)	10.23 (0.19)	11.12 (0.59)	9.96 (0.06)	11.03 (0.80)
464–465	6.65 (0.76)	6.00 (0)	6.28 (0.30)	6.71 (0.63)	6 (0)	5.98 (0)
Ev1	4.00 (0)	4.00 (0)	4.00 (0)	4.00 (0)	4 (0)	3.93 (0)
Ev37	17.35 (1.66)	16.84 (1.15)	16.94 (0.56)	16.89 (0.38)	17.89 (1.06)	16.06 (0.01)
Ev94	9.60 (0.27)	9.63 (0.30)	9.49 (0.06)	9.50 (0.06)	8.97 (0)	9.58 (0.12)
Ev96	12.19 (0.80)	11.47 (0.08)	11.82 (0.04)	12.24 (0.56)	11.96 (0.07)	10.96 (0)
All loci	10.97 (1.17)	10.61 (0.82)	10.65 (0.39)	10.68 (0.49)	10.25 (0.15)	10.23 (0.26)

KF768654-KF768655); 42 haplotypes were only present in the TSA region and 3 were only present in the WSA region. Haplotype (h) and nucleotide diversity (π) estimates obtained were similar for TSA and WSA regions (Table 1).

All 10 microsatellite loci were highly polymorphic, with the number of alleles per locus ranging from 4 (Ev1) to 22 (GATA417). No significant differences were found between the observed heterozygosity (H_o) and the heterozygosity expected (H_e) under Hardy–Weinberg assumptions across loci. When Fisher's method was applied to combine tests across populations, all pairwise combinations of loci were in linkage equilibrium. The overall allelic richness is similar to each region (TSA and WSA) and to each season (E_GAB; L_GAB; R_WSA; L_WSA); however, by locus, there were differences in allelic richness and several private alleles were present not only in each region, but also in each season too (Table 2).

Population structure

For the spatial (TSA vs WSA) analysis, the among-groups component of the Analysis of Molecular Variance analysis was significant for mtDNA data when both the haplotype frequencies as well as molecular distances were considered ($F_{ST} = 0.007$, $P = 0.001$; $\Phi_{ST} = 0.005$, $P = 0.007$). For microsatellite data all estimators showed differentiation ($F_{ST} = 0.002$, $P = 0.001$; $R_{ST} = 0.003$, $P = 0.007$; $D_{est} = 0.016$, SD $= 0.002$). The Exact test of population differentiation was consistent with these results ($P < 0.05$).

Clustering analysis conducted on the whole dataset indicates that the most likely number of distinct genetic entities is $K = 3$, based on ΔK (Fig. 2) and $K = 1$ based on LnP(D) [LnP(D): $-64{,}947.55$ for $K = 1$; LnP(D): $-65{,}039.13$ for $K = 2$; LnP(D): $-65{,}172.8$ for $K = 3$)
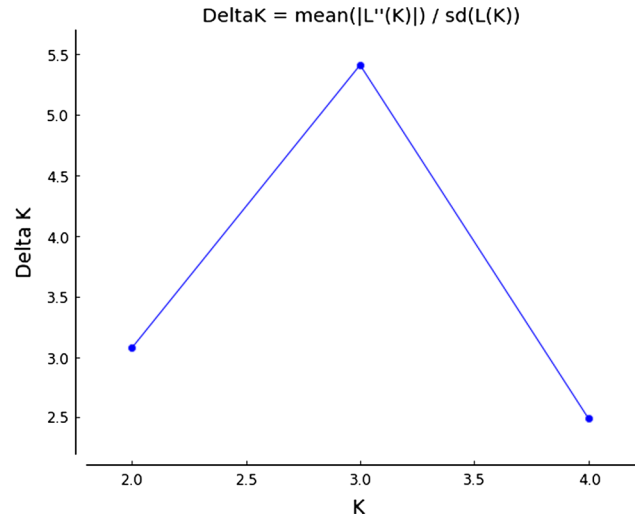

Fig. 2 Variation of ΔK for $K = 2$ to $K = 5$, following Evanno et al. (2005)

Table 3 Pairwise comparisons of genetic differentiation of humpback whales between seasons in Gabon and WSA for mitochondrial DNA and microsatellites

Seasons	mtDNA		Microsatellites	
	F_{ST}	Φ_{ST}	F_{ST}	R_{ST}
E_GAB vs L_GAB	0.000	0.000	0.000	0.000
E_GAB vs R_WSA	**0.004***	0.001	0.000	0.000
E_GAB vs L_WSA	**0.008***	**0.005**	**0.003**	**0.003**
L-GAB vs R_WSA	**0.005***	0.000	0.001	0.002
L_GAB vs L_WSA	**0.010***	**0.005**	**0.003**	**0.003**
R_WSA vs L_WSA	0.000	0.004	0.001	0.000

Significant values are highlighted in bold ($P < 0.05$). (*) indicate significant values for the Exact test

Table 4 Pairwise comparisons of genetic divergence of humpback whales between Seasons in Gabon and WSA for mitochondrial DNA (F_{ST} and Φ_{ST}) and microsatellites (only F_{ST}) by sex

Seasons	mtDNA				Microsatellites	
	F_{ST}		Φ_{ST}		F_{ST}	
	Females	Males	Females	Males[a]	Females	Males
E_GAB vs L_GAB	0.001	0.000	0.002	–	0.001	0.000
E_GAB vs R_WSA	0.004	**0.007***	0.008	–	0.002	0.002
E_GAB vs L_WSA	**0.006***	**0.009***	0.007	–	0.000	**0.002**
L_GAB vs R_WSA	**0.010***	**0.008***	**0.021**	–	**0.002**	0.002
L_GAB vs L_WSA	**0.011***	**0.011***	**0.012**	–	**0.001**	**0.003**
R_WSA vs L_WSA	0.000	0.008*	0.000	–	0.000	0.002

Significant values are highlighted in bold ($P < 0.05$). (*) indicate significant values for the Exact test

[a] As none of the Φ_{ST} values for males were significant, we chose to not show their values

Table 5 Sex-biased dispersal test results based on 10 microsatellite loci of humpback whales from TSA and WSA regions

	AIc	vAIc	F_{IS}	F_{ST}
Males	−0.048	14.641	0.008	0.002
Females	0.089	12.888	−0.003	0.001
P value	0.459	0.101	0.654	0.543

Differences in sex-specific assignment index (AIc), variance of corrected assignment index (vAIc), mean relatedness (F_{IS}) and F_{ST} were tested for significance using 10,000 permutations

(STRUCTURE results—Figure S1). Although, there was evidence of admixture between the two groups based on the assignment of individuals to clusters, when $K = 2$, the proportion of samples assigned to each population was not completely symmetric; 35.5 % of animals showed strong assignment ($Q > 0.8$) to Cluster 1 (TSA) and 31.5 % strongly assigned to Cluster 2 (WSA). Interestingly, of the animals from WSA that were strongly assigned to Cluster 2, 78 % were animals sighted in Late Season.

In the temporal analysis, weak but significant population subdivision was found between seasons but only between regions and only at haplotype frequency level ($F_{ST} = 0.003$, $P = 0.000$) (Table 3). In the pairwise comparisons, F_{ST} showed differentiation between different seasons in the two regions while Φ_{ST} only showed differentiation between Late season in WSA and both seasons in Gabon (Table 3). The Exact test was significant between the two regions. A lower level of differentiation was found for microsatellite data using both F_{ST} and R_{ST} statistics ($F_{ST} = 0.001$, $P = 0.000$; and $R_{ST} = 0.001$, $P = 0.033$). In the pairwise comparisons, significant genetic differentiation was only found between Late season in WSA and both seasons of Gabon, supporting results obtained with Φ_{ST} (Table 3).

When the samples were stratified by sex, the pairwise comparisons using the mitochondrial DNA dataset showed that Late season females in WSA were significantly different from females in Gabon (independent of the season), and Regular season females in South Africa were different from Late season females in Gabon. Significant differences were also found between males from the two regions (Table 4). For the microsatellite dataset, F_{ST} results were significant for both females and males from Late season in WSA, and were significantly different from females and males sampled in Gabon (Table 4).

Measures of migration rates, dispersal and connectivity

Likelihood estimates of population migration rates per generation were high, with $M = 39.776$ (SD = 0.45747), $T = 0.0345$ (SD = 0.003) and the parameter theta $\theta = 15.657$ (SD = 0.789). As expected, estimated scaled divergence times exhibit the opposite behaviour to migration rate.

The power to assign any individual to one of the populations was relatively high. When the assignment criterion was to assign the individual to the population with the highest probability value, the method was able to assign 76.8 % of the samples to the population from which they had been sampled. If the assignment criterion was set by exclusion at $P \leq 0.01$, 42 migrants were identified, 22 in TSA (1.47 %) and 20 in WSA (7.6 %).

Tests for sex-biased dispersal did not yield significant differences between sexes for any of the indicators. Furthermore, while three indicators (AIc, vAIc, FIS) tended to match the expectations for higher dispersal in males than females, the other (F_{ST}) showed the opposite trend (Table 5).

The genetic capture–recapture approach recovered a total of 11 matches involving nine individuals (Table 6)

Table 6 Matches found between Gabon (GA) and WSA

Individuals, sex, date of collection in each site and the group composition (GC) when the sample was collected are provided. (Group description: M–C—mother–calf pair; Pair—two adults; Competitive—at least three adults with one nuclear animal, a principal escort and challengers)

[a] Matches described by Pomilla et al. (2006)

Individual	Sex	GA date	GC	WSA date	GC
#1	F	08/09/2001[a]	Pair	16/12/2001[a]	Pair—defecation
#1		–		17/11/2009	?
#2	F	09/09/2002[a]	M–C	10/01/2003[a]	M–C
#2		–		25/10/2004	Pair
#3	M	07/08/2001	?	21/11/2005	Male with a M–C
#4	F	16/08/2002	Pair	07/11/2006	Pair
#5	M	04/09/2003	Competitive	18/01/2003	Pair
#6	F	26/09/2004	Competitive	08/11/2004	Pair—defecation
#7	M	14/08/2005	Competitive	17/10/2005	3 Adults
#8	F	20/08/2006	Competitive	12/10/2001	Pair—mixed group of whales
#9	M	04/09/2006	Competitive	17/12/2001	20 individuals—milling/feeding

between Gabon and the WSA. Two of these have already been described by Pomilla et al. (2006). All of these whales were sampled later in the year off WSA, between 12 October and 18 January, then off Gabon. Four of them were re-sampled in the same year (with a minimum of 43 days and a maximum of 136 days between sightings) and the other seven were re-sampled in different years (a maximum interval of 8 years between capture and re-capture). Seven matches involved females (one of them with a small calf) and four males. Two individuals were sampled in Gabon and then re-sampled twice (each of them) in South Africa. Three of the individuals sampled in Gabon were re-sampled while feeding in South Africa (including observed defecation). Two individuals were sampled in WSA and then re-sampled in Gabon (5-year interval).

Discussion

Using a combination of maternally and biparentally inherited markers, as well as a large sample size, this study presents evidence for the existence of sub-structure in Breeding Stock B, which may be due to possible spatial and/or temporal segregation.

Explanation for complex population structure

The present study provided new insights into the complex population differentiation of Breeding Stock B as a whole, and re-affirmed previous genetic analyses (Rosenbaum et al. 2009; Pomilla 2005; Pomilla et al. 2006) that indicated significant differentiation between these two areas (TSA and WSA). However, the differences found in previous studies were only at F_{ST} level for mtDNA and for microsatellite data differentiation was found only when Angola (Cabinda) and Gabon were grouped together.

The results from STRUCTURE seem to support this subtle differentiation but were not conclusive. The best K found was three using the ΔK method of Evanno et al. (2005). This method showed that the real number of groups is best detected by the modal value of ΔK, a quantity based on the second-order rate of change with respect to K of the likelihood function. It also efficiently detects the uppermost level of population structure of the data. The subgroups created by the best individual assignment produced by STRUCTURE permits identification of sublevels of structuring (Evanno et al. 2005). Using the LnP(D) method, no structure was found ($K = 1$). However, it has been demonstrated that this method does not provide a correct estimation of the number of clusters, K (Evanno et al. 2005). The new version of STRUCTURE (Hubisz et al. 2009) incorporates a model in which sampling location is incorporated into the Bayesian prior for the analysis, thus increasing power to detect population structure in data sets where divergence is low or the number of loci is moderate (<20). Nonetheless, even this new model does not perform well at either determining the correct value of K or estimating admixture proportions for admixed populations with $F_{ST} < 0.2$ (Hubisz et al. 2009), as was the case in this study.

The present study, based on a more extensive sample than prior assessments and combining maternally and biparentally inherited markers, identified subtle yet statistically significant spatial and temporal sub-structuring for humpback whales sampled along the west coast of Africa. In pairwise comparisons among seasons and sites, animals sampled in Gabon (independently of the season) were significantly different from the ones that were sampled in the late season (spring = November) and summer (December to February) off WSA, for both statistics and markers. This differentiation was true for both sexes. This is of further interest given that many of the matched individuals between Gabon and WSA were sampled at the latter locality during the late season.

Together, the spatial and temporal heterogeneity found suggests that the WSA coast is not wholly a migratory corridor and/or feeding ground utilized by all whales migrating from Gabon. This is supported by additional evidence: (1) Based on catch histories, Findlay (2000) suggested a degree of spatial segregation of this stock, with some whales habitually using alternate migration routes and thus avoiding capture (e.g. some animals migrated further offshore outside of the range of shore-based whalers). The existence of a concentration of pelagic catches in the South Atlantic 750–1,400 km offshore at the same latitude and the same season as the WSA feeding ground confirms that some whales must take a more direct route to their Southern Ocean feeding grounds from further north (Barendse et al. 2011); (2) Recent telemetry data indicated a similar pattern; whales tagged in Gabon travelled south to the central Angolan coast and then offshore along the Walvis Ridge, passing WSA far offshore before reaching 40°S, where they diverged and continued south as far as Bouvet Island (54°26′S–3°24′E; Rosenbaum et al. accepted); (3) A recent model averaged abundance estimate for humpback whales sighted in WSA was 510 (SE 143; 95 % CI 230–790; Barendse et al. 2011), which is far fewer animals compared with 1,404 individuals genetically identified in Gabon alone.

The telemetry data suggest that some humpback whales wintering in the greater Gulf of Guinea region use at least two different migration routes for different summer feeding grounds (Rosenbaum et al. accepted). One of these feeding grounds is around the border between Areas II and III (at 0° of longitude; IWC 1998) in the vicinity of Bouvet Island, and migration to and from this feeding ground is mostly offshore. This could explain the different catch histories shown for the whaling ground off Gabon and others to the south (e.g. Findlay 2000).

The other known feeding area is off the west coast of South Africa (Best et al. 1995; Barendse et al. 2010) to which animals possibly take a coastal migration route between breeding and feeding grounds: whether this is the final or only temporary destination (with Antarctic as final destination) is not known. Genetic differences observed between samples from Gabon and WSA could therefore be an evolutionary consequence of long-term-directed site fidelity to these specific feeding grounds. The different levels of resampling of individuals found in the two areas are consistent with this hypothesis; the overall resampling rate of photographically identified individuals in the WSA area is almost three times higher than in Gabon, indicating high site fidelity, with several animals returning to the WSA region in multiple years (Barendse et al. 2010, 2011). Variations in site fidelity by humpback whales to breeding and feeding grounds have been shown elsewhere, with high annual return rates to feeding grounds (up to 90 %;

Clapham et al. 1993; Acevedo et al. 2006) whereas fidelity to breeding grounds appear to be more fluid, with whales roaming widely during the winter (Mattila et al. 1994). In the North Atlantic and North Pacific, humpback whales feed in a number of relatively discrete sub-populations in summer, but mix at a common breeding ground in winter (e.g. Palsbøll et al. 2004, 2007).

Given the known exchange of individuals between Gabon and WSA, and the paucity of samples from elsewhere in the Gulf of Guinea, from BSB migratory routes and feeding areas in general, the differentiation found between the sampled areas likely represents only part of a more complex picture. Evidence of temporal and geographic sub-structure has been found in other wintering grounds (Smultea 1994; Medrano-González et al. 1995; Brown et al. 1995; Dawbin 1997). It is likely that Gabon is not the final destination in the greater Gulf of Guinea for some individuals. Humpback whales have been reported present in several other areas, including Guinea (Bamy et al. 2010), Togo, Nigeria, Ghana (Van Waerebeek et al. 2001, 2009), the coasts of Equatorial Guinea and Congo (Best et al. 1999; Rosenbaum and Collins 2006) and the islands of Bioko, Pagalu (Aguilar 1985) and São Tomé (Carvalho et al. 2011). Some of the whales satellite tagged in Gabon in 2002 moved northwards, deeper into the Gulf of Guinea, to destinations that included São Tomé and Príncipe, Cameroon, Bioko Island (Equatorial Guinea) Nigeria and Ghana (4.7°N) using inshore routes over the continental shelf as well as offshore movements across deep waters (Rosenbaum et al. accepted). Recent data also show the regular presence of humpback whales in Angola during the austral winter and spring (Weir 2011), with singing activity relatively constant for 4 months from early July to early November and with distinct spatial patterns (inshore and offshore; Cerchio et al. accepted). Spatial segregation on breeding grounds is common on wintering areas; in the south-western Indian Ocean (stock C) substructure within several longitudinally separated breeding regions apparently exists, suggesting demographic independence with recent exchange of individuals between regions (Ersts et al. 2011) and probably a high degree of shared ancestry among whales in sub-populations (Rosenbaum et al. 2009). On the other hand, according to contemporary statements from early twentieth century whalers, trends in pre-WWI seasonal catches (or availability data) and incidental sightings recorded from oil platforms in 1995, the seasonal occurrence of humpback whales of Angola is distinctly bimodal, with waves of humpbacks going north from May to July and south from August to November (Barendse et al. 2011). This pattern of availability would not be completely consistent with spatial segregation of whales within the greater Gulf of Guinea by latitude, at least south of Gabon.

Numerous studies of dispersal and population genetics have shown evidence of geographic population structure and dispersal limitation (Palumbi and Warner 2003; Palumbi 2004). Behaviour appears to be an important factor driving geographic structure in marine populations and it has been suggested that philopatry has an important role in driving population structure in marine turtles, sharks and whales (e.g. Lee et al. 2007; Palumbi and Baker 1994; Pardini et al. 2001). Due to the remarkable mobility of humpback whales and the apparent absence of geographic barriers within ocean basins, the formation of significant genetic divisions between stocks indicates strong fidelity to migratory destinations, as a result of a calf's early maternally inherited experience (Baker et al. 1994), especially in its first year of life (Baker et al. 1998). The life history of these whales suggests a likely mechanism for a 'cultural' transmission of migratory destinations (Baker et al. 1990), and is supported by analysis of maternally inherited mtDNA.

The stock structure of Breeding Stock B may be more complex due to the apparent subdivision of groups of animals with different behavioural and migratory patterns. A multi sub-stock structure model that assumes two breeding units, with the presence of multiple migratory streams (inshore vs offshore) to different feeding areas and/or the existence of some spatial and/or temporal segregation on breeding grounds in the greater Gulf of Guinea would provide some explanation for the genetic differentiation found in this study. Many populations are composed of a mixture of individuals that reproduce at different times within a particular season and location. The timing of reproduction is highly heritable and is thus influenced by genetic as well as (seasonal) environmental factors (Hendry and Day 2005). Under these conditions, gene flow is expected to be limited between early and late reproducers even within populations having a unimodal temporal distribution of reproductive activity (Hendry et al. 2004). A temporal, instead of spatial, restriction on gene flow would hence create a pattern of isolation by time (IBT; Hendry and Day 2005); this could be a plausible explanation for the population structure found in BSB stock if the animals arrived in the greater Gulf of Guinea at slightly different times due to the fact that they had migrated from different feeding areas. However, habitually early or late breeding individuals have not been detected in any baleen whale population to date.

Connectivity, novel biogeography: a re-occurring pattern?

Direct connectivity between the two areas was demonstrated through detection of movements by genetically identified individuals between Gabon (TSA) and WSA, both for females and males. All of the detected movements were from a northerly to southerly direction (when considering month of recapture). Therefore, when recaptures did not occur in the same year, the presumption is that the monthly nature of recaptures (August/September in Gabon and October–November–December–January in South Africa) reflects an overall pattern of movement of some animals from TSA to WSA, consistent with photo-identification results (Barendse et al. 2011). Some of the animals from Gabon feed on the WSA coast (sighted while engaged in feeding behaviours) and the migration rate between the two areas is high. Potentially similar latitudinal gradations have been observed among other populations of humpback whales. In the south east Pacific, the main summer feeding ground for the G Stock of humpback whales that breeds off Colombia, Costa Rica, Peru and Ecuador (Flórez-González 1991; Scheidat et al. 2000; Félix and Haase 2005), apparently extends along the western coast of the Antarctic Peninsula (Mackintosh 1942; Stone et al. 1990; Stevick et al. 2004). However, the cold inshore waters of western South America, including in the southern Patagonian fjords and the Strait of Magellan in Southern Chile, have recently been described as an alternative summer feeding ground for humpback whales of this stock, with a high site fidelity and residency to the area (Gibbons et al. 2003). However, no genetic comparison between these areas has yet been made.

Oceanographic features may serve as migratory cues or oceanographic boundaries. Upwelling areas, ocean currents and fronts, sea surface temperature and depth have been recently identified as important factors influencing the distribution and abundance of cetaceans, and may influence dispersal and population structure (e.g. Davis et al. 2002; Mendez et al. 2010, 2011; Ainley et al. 2012). For humpback whales, breeding areas are at approximately 20° latitude in both hemispheres with average sea surface temperatures of between 24 and 28 °C (Rasmussen et al. 2007). Breeding areas in the greater Gulf of Guinea and the eastern Pacific (breeding stock G) occur near the equator where SST $\geq$ 24 °C (Rasmussen et al. 2007). These two regions are considerably further to the north of other southern hemisphere breeding grounds, implying that humpback whales in these areas need to migrate farther and suggest that the large marine upwelling ecosystems of each region (Benguela Current and Humboldt/Peru Current, respectively) influence the location of breeding grounds for these populations.

Conclusions

Spatial differences and fidelity/philopatry to certain areas have been identified as important mechanisms driving genetically differentiated sympatric populations (Hoelzel 1998). Temporal influences may also play a role and should be considered as segregating factors. In this case, it is postulated that differentiated breeding stocks mix in temporary

assemblages, and the pattern of mixing can vary on both spatial and temporal scales.

The results presented here do not fully resolve the questions about Breeding Stock B structure, principally because sampling in the region has been extremely limited geographically, with 99 % of the samples coming from two areas 3,500 km apart and collected in largely different seasons. Nevertheless, some population segregation is clearly evident, largely for whales sampled off Gabon in winter and those in WSA in spring/summer, which suggests that some of the whales feeding off WSA may be breeding and calving in an area which has not yet been identified. The implementation of further satellite tagging efforts could be useful for assessing movements throughout the region and could clarify patterns of connectivity between regions and help identify the locations of unknown breeding or feeding areas. Additional genetic samples from other areas in the region (inshore and offshore) especially from other parts of the Gulf of Guinea, Angola, Namibia, and further north and west (e.g. the Bight of Benin) could give additional information on this stock's population structure. Increasing the number of markers (microsatellites) analysed could improve the resolution of analyses and thus the detection of the population structure in this region. Additionally, using other types of markers like SNPs (Genomic Single Nucleotide Polymorphism) can be used to infer population structure, estimate the spatial scale of gene flow and examine population connectivity, aspects which were typically unidentified using traditional approaches (Allendorf et al. 2010). Improving our knowledge of the patterns of connectivity and population structure is important for the conservation and management of these humpback whale populations.

References

Acevedo JA, Aguayo-Lobo A, Pastene LA (2006) Filopatría de la ballena jorobada (Megaptera novaeangliae Borowski, 1781), al área de alimentación del Estrecho de Magallanes. Rev Biol Mar Oceanog 41:11–19

Aguilar A (1985) Aboriginal whaling off Pagalu (Equatorial Guinea). Paper SC/36/PS13 presented to the International Whaling Commission Scientific Committee

Ainley DG, Jongsomjit D, Ballard G, Thiele D, Fraser WR, Tynan CT (2012) Modeling the relationship of Antarctic minke whales to major ocean boundaries. Polar Biol 35:281–290

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nat Rev Genet 10: 697–709

André C, Larsson LC, Laikre L, Bekkevold D, Brigham J, Carvalho GR, Dahlgren TG, Hutchinson WF, Mariani S, Mudde K, Ruzante DE, Ryman N (2011) Detecting population structure in a high gene-flow species, Atlantic herring (*Clupea harengus*): direct, simultaneous evaluation of neutral vs putatively selected loci. Heredity 106:270–280

Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge

Baker CS, Palumbi SR, Lambertson RH, Weinrich MT, Calambokidis J, O'Brien SJ (1990) Influence of seasonal migration on geographic distribution of mitochondrial DNA haplotypes in humpback whales. Nature 344:238–240

Baker CS, Perry A, Bannister JL, Weinrich MT, Abernethy RB, Calambokidis J, Lien J, Lambertsen RH, Urbán J, Vásquez O, Clapham PJ, Alling A, O'Brien SJ, Palumbi SR (1993) Abundant mitochondrial DNA variation and world-wide population structure in humpback whales. Proc Natl Acad Sci 90:8239–8243

Baker CS, Slade RW, Bannister JL, Abernethy RB, Weinrich MT, Lien J, Urbán Ramirez J, Corkeron P, Calambokidis J, Vasquez O, Palumbi SR (1994) Hierarchical structure of mitochondrial DNA gene flow among humpback whales *Megaptera novaeangliae*, worldwide. Mol Ecol 3:313–327

Baker CS, Flórez-González L, Abernethy B, Rosenbaum HC, Slade RW, Capella J, Bannister JL (1998) Mitochondrial DNA variation and maternal gene flow among humpback whales of the southern hemisphere. Mar Mammal Sci 14:721–737

Bamy I, Van Waerebeek K, Bah SS, Dia M, Kaba B, Keita N, Konate S (2010) Species occurrence of cetaceans in Guinea, including humpback whales with southern hemisphere seasonality. Mar Biodivers Rec 3:e48

Barendse J, Best P, Thornton M, Pomilla C, Carvalho I, Rosenbaum HC (2010) Migration redefined? Seasonality, movements, and group composition of humpback whales *Megaptera novaeangliae* off the west coast of South Africa. Afr J Mar Sci 32:1–22

Barendse J, Best PB, Thornton M, Elwen SH, Rosenbaum HC, Carvalho I, Pomilla C, Collins T, Meyer M (2011) Transit station or destination? Attendance patterns, regional movement, and population estimate of humpback whales *Megaptera novaeangliae* off west South Africa based on photographic and genotypic matching. Afr J Mar Sci 33:353–373

Barendse J, Best PB, Carvalho I, Pomilla C (2013) Mother knows best: occurrence and associations of resighted humpback whales suggest maternally derived fidelity to a Southern Hemisphere coastal feeding ground. PLoS ONE 8:e81238

Berubé M, Palsbøll P (1996) Identification of sex in Cetaceans by multiplexing with three ZFX and ZFY specific primers. Mol Ecol 5:283–287

Best PB, Sekiguchi K, Findlay KP (1995) A suspended migration of humpback whales *Megaptera novaeangliae* on the west coast of South Africa. Mar Ecol Prog Ser 118:1–12

Best PB, Reeb D, Morais M, Baird A (1999) A preliminary investigation of humpback whales off northern Angola. Paper SC/51/CAWS33 presented to the International Whaling Commission Scientific Committee

Bowen BW, Karl SA (2007) Population genetics and phylogeography of sea turtles. Mol Ecol 16:4886–4907

Brown M, Corkeron P, Hale P, Schultz K, Bryden M (1995) Evidence for a sex-segregated migration in the humpback whale (*Megaptera novaeangliae*). Proc R Soc Lond Ser B 259: 229–234

Budker P, Collignon J (1952) Trois campagnes baleneires au Gabon: 1949–1950–1951. Bulletin l'Institut d'etudes centrafricaines 3:75–100

Carvalho I, Brito C, dos Santos M, Rosenbaum HC (2011) The waters of São Tomé: a calving ground for West African humpback whales? Afr J Mar Sci 33:91–97

Cerchio S, Strindberg S, Collins T, Bennett C, Rosenbaum HC (accepted) Seismic surveys negatively affect humpback whale singing activity off northern Angola. PLoS ONE

Chittleborough RG (1965) Dynamics of two populations of humpback whales, *Megaptera novaeangliae* (Borowski). Aust J Mar Fresh Res 16:33–128

Chow S, Okamoto H, Miyabe N, Hiramatsu K, Barut N (2000) Genetic divergence between Atlantic and Indo-Pacific stocks of bigeye tuna (*Thunnus obesus*) and admixture around South Africa. Mol Ecol 9:221–227

Clapham PJ, Baraff LS, Carlson CA, Christian MA, Mattila DK, Mayo CA, Murphy MA, Pittman S (1993) Seasonal occurrence and annual return of humpback whales, *Megaptera novaeangliae*, in the southern Gulf of Maine. Can J Zool 71: 440–443

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics 153:1989–2000

Davis RW, Ortega-Ortiz JG, Ribic CA, Evans WE, Biggs DC, Ressler PH, Cady RB, Leben RR, Mullin KD, Würsig B (2002) Cetacean habitat in the northern oceanic Gulf of Mexico. Deep-Sea Res (1 Oceanogra Res Pap) 49:121–142

Dawbin W (1966) The seasonal migratory cycle of humpback whales. In: Norris KS (ed) Whales, dolphins, and porpoises. University of California Press, Berkeley and Los Angeles, pp 145–171

Dawbin W (1997) Temporal segregation of humpback whales during migration in southern hemisphere waters. Mem Queensl Mus 42:105–138

Earl D, vonHoldt B (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4(2):359–361

Ersts PJ, Pomilla C, Kiszka J, Cerchio S, Rosenbaum HC, Vély M, Razafindrakoto Y, Loo JA, Leslie M, Avolio M (2011) Observations of individual humpback whales utilising multiple migratory destinations in the southwestern Indian Ocean. Afri J Mar Sci 33:333–338

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Excoffier L, Lischer H (2010) 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

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Favre L, Balloux F, Goudet J, Perrin N (1997) Female-biased dispersal in the monogamous mammal *Crocidura russula*: evidence from field data and microsatellites patterns. Proc R Soc Lond Ser B 264:127–132

Félix F, Haase B (2005) Distribution of humpback whales along the coast of Ecuador and management implications. J Cetacean Res Manag 7:21–31

Findlay KP (2000) A review of humpback whale caches by modern whaling operations in the Southern Hemisphere. Mem Queensl Mus 47:411–420

Flórez-González L (1991) Humpback whales *Megaptera novaeangliae* in the Gorgona Island, Colombian Pacific breeding waters: population and pod characteristics. Mem Queensl Mus 30:291–295

Fontaine MC, Baird SJE, Piry S, Ray N, Tolley KA, Duke S, Birkun A, Ferreira M, Jauniaux T, Llavona A, ztúrk BO, ztúrk AAO, Ridoux V, Rogan E, Sequeira M, Siebert U, Vikingsson GA, Bouquegneau J, Michaux JR (2007) Rise of oceanographic barriers in continuous populations of a cetacean: the genetic structure of harbour porpoises in Old World waters. BMC Biol 5:30–46

Gibbons J, Capella J, Valladares C (2003) Rediscovery of a humpback whale, *Megaptera novaeangliae*, summering ground in the Strait of Magellan, Chile. J Cetacean Res Manag 5:203–208

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). http://www.unil.ch/izea/softwares/fstat.html. Accessed 26 June 2008

Guo S, Thompson E (1992) Performing the Exact test of Hardy Weinberg proportion for multiple alleles. Biometrics 48:361–372

Hendry AP, Day T (2005) Population structure attributable to reproductive time: isolation by time and adaptation by time. Mol Ecol 14:901–916

Hendry AP, Morbey YE, Berg OK, Wenburg JK (2004) Adaptive variation in senescence: reproductive lifespan in a wild salmon population. Proc R Soc Lond Ser B 271:259–266

Hoelzel AR (1998) Genetic structure of cetacean populations in sympatry, parapatry and mixed assemblages; implications for conservation policy. J Hered 89:451–458

Hubisz M, Falush D, Stephens M, Pritchard J (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332

Hudson RR, Boos DD, Kaplan NL (1992) A statistical test for detecting geographic subdivision. Mol Biol Evol 9:138–151

IWC (1998) Report of the scientific committee—annex G. Report of the sub committee on the comprehensive assessment of southern hemisphere humpback whales. Rep Int Whaling Comm 48:170–182

Jost L (2008) G_{ST} and its relatives do not measure differentiation. Mol Ecol 17:4015–4026

Kalinowski ST (2005) HP-rare: a computer program for performing rarefaction on measures of allelic diversity. Mol Ecol Notes 5:187–189

Knutsen H, Jorde PE, André C, Stenseth NC (2003) Fine-scaled geographic population structuring in a highly mobile marine species: the Atlantic cod. Mol Ecol 12:385–394

Laidre KL, Heagerty PJ, Heide-Jørgensen MP, Witting L, Simon M (2009) Sexual segregation of common minke whales (*Balaenoptera acutorostrata*) in Greenland, and the influence of sea temperature on the sex ratio of catches. ICES J Mar Sci 66:2253–2266

Lambertsen RH (1987) A biopsy system for large whales and its use for cytogenetics. J Mammal 68:443–445

Lee PL, Luschi P, Hays GC (2007) Detecting female precise natal philopatry in green turtles using assignment methods. Mol Ecol 16:61–74

Loseto LL, Richard P, Stern GA, Orr J, Ferguson SH (2006) Sexual segregation of Beaufort beluga whales during the open-water season. Can J Zool 84:1743–1751

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

Mackintosh NA (1942) The southern stocks of whalebone whales. Discov Rep 22:197–300

Matthews LH (1938) The humpback whale, *Megaptera nodosa*. Discov Rep 17:7–92

Mattila DK, Clapham PJ, Vasquez O, Bowman R (1994) Occurrence, population composition and habitat use of humpback whales in Samana Bay, Dominican Republic. Can J Zoolog 72:1898–1907

McKelvey KS, Schwartz MK (2005) DROPOUT: a program to identify problem loci and samples for noninvasive genetic samples in a capture-mark-recapture framework. Mol Ecol Notes 5:716–718

Medrano-González L, Aguayo-Lobo A, Urbán-Ramirez J, Baker S (1995) Diversity and distribution of mitochondrial DNA lineages among humpback whales, *Megaptera novaeangliae*, in the Mexican Pacific Ocean. Can J Zool 73:1735–1743

Mendez M, Rosenbaum HC, Subramaniam A, Yackulic C, Bordino P (2010) Isolation by environmental distance in mobile marine species: molecular ecology of franciscana dolphins at their southern range. Mol Ecol 19:2212–2228

Mendez M, Subramaniam A, Collins T, Minton G, Baldwin R, Berggren P, Sarnblad A, Amir OA, Peddemors VM, Karczmarski L, Guissamulo A, Rosenbaum HC (2011) Molecular ecology meets remote sensing: environmental drivers to population structure of humpback dolphins in the Western Indian Ocean. Heredity 107:349–361

Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distance between alleles with special reference for microsatellite loci. Genetics 142:1061–1064

Mossman CA, Waser PM (1999) Genetic detection of sex-biased dispersal. Mol Ecol 8:1063–1067

Narum SR (2006) Beyond Bonferroni: less conservative analyses for conservation genetics. Conserv Genets 7:783–787

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov Chain Monte Carlo approach. Genetics 158:885–896

Palsbøll PJ, Vader A, Bakke I, Raafat El-Gewely M (1992) Determination of gender in cetaceans by the polymerase chain reaction. Can J Zool 70:2166–2170

Palsbøll PJ, Bérubé M, Larsen A, Jørgensen H (1997) Primers for the amplification of tri- and tetramer microsatellite loci in baleen whales. Mol Ecol 6:893–895

Palsbøll PJ, Bérubé M, Aguilar A, Notarbartolo-Di-Sciara G, Nielsen R (2004) Discerning between recurrent gene flow and recent divergence under a finite-site mutation model applied to North Atlantic and Mediterranean Sea fin whale (*Balaenoptera physalus*) populations. Evolution 58:670–675

Palsbøll PJ, Bérubé M, Allendorf FW (2007) Identification of management units using population genetic data. Trends Ecol Evol 22:11–16

Palumbi SR (1994) Genetic divergence, reproductive isolation, and marine speciation. Annu Rev Ecol Syst 25:547–572

Palumbi SR (2004) Marine reserves and ocean neighborhoods: the spatial scale of marine populations and their management. Annu Rev Environm Resourc 29:31–68

Palumbi SR, Baker CS (1994) Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. Mol Biol Evol 11:426–435

Palumbi SR, Warner RR (2003) Why gobies are like hobbits. Science 299:51–52

Pardini AT, Jones CS, Noble LR, Kreiser B, Malcolm H, Bruce BD, Stevens JD, Cliff G, Scholl MC, Francis M, Duffy CAJ, Martin AP (2001) Sex-biased dispersal of great white sharks—in some respects, these sharks behave more like whales and dolphins than other fish. Nature 412:139–140

Park S (2001) The excel microsatellite toolkit: Trypanotolerance in west African cattle and the population genetic effects of selection. Ph.D. dissertation, University of Dublin. Available at http://oscar.gen.tcd.ie/~sdepark/ms-toolkit/

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel—population genetic software for teaching and research—an update. Bioinformatics 28:2537–2539

Perneger TV (1998) What's wrong with Bonferroni adjustments. BMJ 316:1236–1238

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GeneClass2: a software for genetic assignment and first generation migrants detection. J Hered 95:536–539

Pomilla C (2005) Genetic structure of humpback whale (Megaptera novaeangliae) populations on Southern Hemisphere wintering grounds. Ph.D. Dissertation, New York University

Pomilla C, Best PB, Findlay KP, Collins T, Engel M, Minton G, Ersts P, Barendse J, Kotze PGH, Razafindrakoto Y, Ngouessono S, Meyer M, Thornton M Rosenbaum H (2006) Population structure and sex-biased gene flow in humpback whales from Wintering Regions A, B, C, and X based on nuclear microsatellite variation. Paper SC/A06/HW38 presented to the International Whaling Commission Scientific Committee

Pritchard JK, Stephens M, Donelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proc Natl Acad Sci 98:9197–9201

Rasmussen K, Palacios DM, Calambokidis J, Saborío MT, Dalla Rosa L, Secchi ER, Steiger GH, Allen JM, Stone GS (2007) Southern Hemisphere humpback whales wintering off Central America: insights from water temperature into the longest mammalian migration. Biol Lett 3:302–305

Raymond M, Rousset F (1995) GENEPOP: population genetic software for Exact tests and ecumenicism. J Hered 86:248–249

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225

Rice DW (1998) Marine mammals of the world: systematics and distribution. Special Publication Number 4, Society for Marine Mammalogy. Allen Press, USA

Rosenbaum HC, Collins T (2006) The ecology, population characteristics and conservation efforts for humpback whales (*Megaptera novaeangliae*) on their wintering grounds in the coastal waters of Gabon. In: Alonso A, Campbell P (eds) Natural history of the gamba complex. Smithsonian Press, Washington DC, pp 425–433

Rosenbaum HC, Ersts P, Razafindrakoto Y, Sounguet G, Pomilla C, Ngouessono S, Rasoamampianina V, White L (2002) Population characteristics, distribution, and relative abundance of humpback whales off the coasts of Madagascar and Gabon: an update on recent and planned research. Paper SC/54/H20 presented to the International Whaling Commission Scientific Committee

Rosenbaum HC, Pomilla CC, Mendez MC, Leslie M, Best P, Findlay K, Minton G, Ersts P, Collins T, Engel M, Bonatto S, Kotze D, Meÿer M, Barendse J, Thornton M, Razafindrakoto Y, Ngouessono S, Vely M, Kiszka J (2009) Population structure of humpback whales from their breeding grounds in the South Atlantic and Indian Oceans. PLoS ONE 4:e7318

Rosenbaum HC, Maxwell S, Kershaw F, Mate B (accepted) Quantifying long-range movements and potential overlap with anthropogenic activity for humpback whales in the South Atlantic Ocean. Conserv Biol

Rozas J, Sánchez-DelBarrio JC, Messegyer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496–2497

Ruckstuhl KE, Neuhaus P (2005) Sexual segregation in vertebrates: ecology of the two sexes. Cambridge University Press, Cambridge

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York

Scheidat M, Castro C, Denkinger J, González J, Adelung D (2000) A breeding area for humpback whales (*Megaptera novaeangliae*) off Ecuador. J Cetacean Res Manag 2:165–172

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

Shaw PW, Pierce GJ, Boyle PR (1999) Subtle population structuring within a highly vagile marine invertebrate, the veined squid Loligo forbesi, demonstrated with microsatellite DNA markers. Mol Ecol 8:407–417

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462

Smultea M (1994) Segregation by humpback whale (*Megaptera novaeangliae*) cows with a calf in coastal habitat near the island of Hawaii. Can J Zool 72:805–811

Stevick P, Aguayo A, Allen J, Avila IC, Capella J, Castro C, Chater K, Engel MH, Félix F, Flórez-González L, Freitas A, Haase B, Llano M, Lodi L, Muñoz E, Olavarría C, Secchi E, Scheidat M, Siciliano S (2004) A note on the migrations of individually identified humpback whales between the Antarctic Peninsula and South America. J Cetacean Res Manag 6:109–113

Stone GS, Flórez-González L, Katona S (1990) Whale migration record. Nature 346:705

Valsecchi E, Amos W (1996) Microsatellite markers for the study of cetacean populations. Mol Ecol 5:151–156

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Van Waerebeek K, Tchibozo S, Montcho J, Nobime G, Sohouhoue P, Dossou C (2001) The Bight of Benin, a North Atlantic breeding ground of the Southern Hemisphere humpback whale population, likely related to Gabon and Angola substocks. Paper SC/53/IA21 presented to the International Whaling Commission Scientific Committee

Van Waerebeek K, Ofori-Danson PK, Debrah J (2009) The cetaceans of Ghana: a validated faunal checklist. West Afr J Appl Ecol 15:61–90

Waits L, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. Mol Ecol 10:249–256

Walsh PD, Fay JM, Gulick S, Sounguet GP (2000) Humpback whale activity near Cap Lopez, Gabon. J Cetacean Res Manag 2:63–68

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J Hered 89:438–450

Ward RD, Woodwark M, Skibinski DOF (1994) A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. J Fish Biol 44:213–232

Weir C (2011) Distribution and seasonality of cetaceans in tropical waters between Angola and the Gulf of Guinea. Afr J Mar Sci 33:1–15

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370

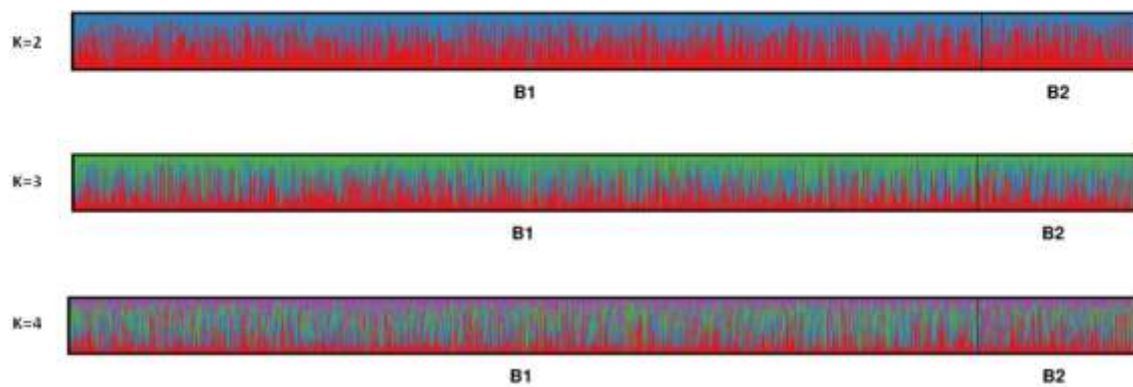Whitehead H (2003) Sperm whales: social evolution in the oceans. University of Chicago Press, Chicago

Supplementary material

Figure S1 (Supplementary material) - STRUCTURE clustering results for K = 2, 3 and 4. Group labels are B1 = TSA, combined samples from Gabon, São Tomé, and Angola; B2 = WSA, samples from the West coast of South Africa. Each individual is represented by a vertical column partitioned into colour segments that represents its estimated admixture fraction in each cluster
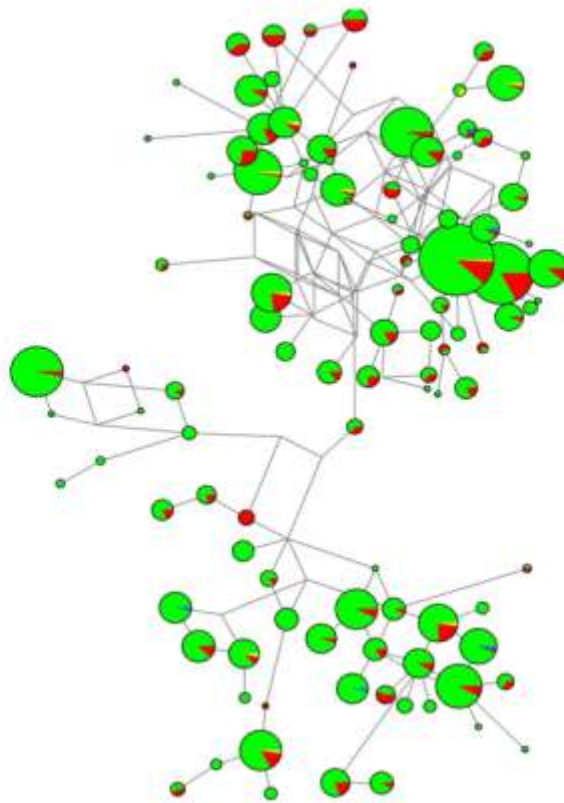
Figure S2 (Supplementary material) - Median-joining network of control region mtDNA haplotypes of humpback whales, implemented in NETWORK 4.6. (Bandelt et al. 1999). Circle size is proportional to the number of individuals exhibiting the corresponding haplotype. Each location within each haplotype is coloured according to the legend: TSA region: *green circle* Gabon; *blue circle* São Tomé; *yellow circle* Angola – Cabinda; WSA region: *red circle* West South Africa.