# Temporal patterns of population structure of humpback whales in West coast of Africa (B stock)

Inês CARVALHO[1,2,3], Jacqueline LOO[2,3,4], Tim COLLINS[3,5], Cristina POMILLA[2,3], Jaco BARENDSE[6], Peter B. BEST[6], Rebecca HERSCH[2], Matt S. LESLIE[2,3], Meredith THORNTON[6] & Howard C. ROSENBAUM[2,3]

[1] *Faculdade de Ciências do Mar e Ambiente – Universidade do Algarve, Campus Gambelas, 8000-139 Faro, Portugal.*

[2] *Sackler Institute for Comparative Genomics, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA.*

[3] *Ocean Giants Program, Wildlife Conservation Society, 2300 Southern Blvd., Bronx, NY 10460-1099, USA.*

[4] *Department of Biology, New York University, 100 Washington Square, New York, NY 10012, USA*

[5] *Oman Whale and Dolphin Research Group*

[6] *Mammal Research Institute, University of Pretoria, c/o South African Museum, PO Box 61, Cape Town, 8000 South Africa*

*Contact Email:* carvalho.inesc@gmail.com

**ABSTRACT**

In the eastern South Atlantic Ocean (Region B) humpback whales are distributed along the west coast of South Africa and winter in the Gulf of Guinea. The most recent data available suggest that Breeding stock B is possibly sub-structured, with B1 considered a breeding ground and B2 a summer feeding ground and a winter migration corridor. However questions remain over the population structure of B stock. Here, we present an assessment of temporal population structure in humpback whales on the west coast of Africa using maternally (mitochondrial DNA control region) and bi-parentally (10 microsatellites) inherited markers. We amplified, sexed, genotyped and sequenced a total of 2018 samples from B1 (Gabon, Angola, São Tomé) and B2 (West South Africa). The results showed significantly differentiation based on haplotype frequencies ($F_{ST}$) and molecular distances ($\Phi_{ST}$) between B1 and B2; similar results were obtained with the microsatellite data, however very low gene flow was detected the two regions. For the temporal analysis, significant results were obtained only for haplotype frequency statistics ($F_{ST}$), where west South Africa seasons were significantly different from seasons in Gabon. When the samples were stratified by sex, significant differentiation at the haplotype level were found for both sexes and nucleotide level only for females. The direct detection of movements by genetically identified individuals, females and males, suggests that interchange occurs between regions. However, all movements to date are from a northbound to southbound direction. The results presented here indicate that there is some spatial and temporal population substructure in humpback whales in B stock.

Keywords: Humpback whales, Atlantic Ocean, Breeding ground, population structure, B Stock

## INTRODUCTION

Southern Hemisphere humpback whales (*Megaptera novaeangliae*) undertake annual migrations from summer high latitude feeding grounds to winter breeding grounds in tropical and sub-tropical waters (Mackintosh, 1942). In the Southern Hemisphere, humpback whale distribution on wintering grounds occurs within seven Wintering Regions (A-G) based on low latitude distributions (IWC, 1997).

In the eastern South Atlantic Ocean (Region B) whales are distributed from the west coast of South Africa (Ølsen, 1914; Matthews, 1938) to wintering areas in Gulf of Guinea (Budker and Collignon, 1952; Van Waerebeek *et al*. 2009; Walsh *et al*, 2000). Recent studies have been conducted in Gabon and Angola confirming that these areas are used as breeding grounds by these animals (Rosenbaum *et al.*, 2009; Best *et al*, 1995; Best *et al,* 1999). Humpback whales have been reported in other areas off the West coast of Africa, including the Bight of Benin, Togo, Nigeria, Ghana (Van Waerebeek 2001, Van Waerebeek 2009, Rosenbaum & Mate, 2006), the coasts of Equatorial Guinea and Congo (Best *et al.* 1999, Rosenbaum & Collins 2006), offshore islands of São Tomé (Picanço *et al.* 2009, Carvalho *et al.* submitted), Bioko and Pagalu (Aguilar 1985). The most recent data available suggest that Breeding stock B is possibly sub-structured with B1 as a breeding ground and B2 as a summer feeding ground and a winter migration corridor (Rosenbaum *et al.* 2009; Barendse *et al.* 2010; Pomilla *et al.* 2006; Carvalho *et al* 2009; SC/58/Rep.5). Some degree of movements of animals has been detected between Gabon (B1) and west South Africa (B2): nine individuals (five females and four males) were sighted in Gabon and then were re-sighted off the west coast of South Africa (Carvalho *et al*. 2009), in the same way satellite telemetry studies show animals moving between these two areas (Rosenbaum & Mate, 2006). However, significant differences have been found in mitochondrial DNA haplotype frequencies and nuclear data between these two regions (Rosenbaum *et al.* 2009; Carvalho *et al.* 2009), with very low gene flow rates detected (Rosenbaum *et al.* 2009). On the west South African coast, recent observations indicate a regular and extended presence of whales as late as March, direct observation of feeding during sring and summer months and 19$^{th}$ century catches during the summer season off Namibia suggest that this area besides functioning as a migration corridor, may also serve as an important supplementary feeding ground for some animals of the population (Barendse *et al*., 2010; Best *et al*. 1995: Findlay, 2000).

A number of questions regarding the population structure of B stock remain unclear. One possible explanation for some of these differences could be a temporal structuring of this stock. Evidence of temporal and geographic sub-structure has been found in other wintering grounds (Smultea, 1994; Medrano-Gonzalez et al. 1995, Brown *et al.*, 1995, Dawbin, 1997). There are also significant differences in the time period that the whales arrive at and depart from a breeding area based on their reproductive status (Dawbin, 1997). Differences in group size and composition during a season (Smultea, 1994) and temporal distribution of matrilines at breeding grounds have also been demonstrated (Medrano-Gonzalez et al., 1995).

In this report we update our previous report (Carvalho *et al*., 2009), including here new samples from west South Africa coast.

## MATERIALS AND METHODS

### Sample collection, DNA extraction and sex determination

Samples were collected from humpback whales at five sampling sites on the west coast of Africa (Fig. 1, table I). Skin tissues were mostly obtained using the biopsy dart procedure (Lambertsen 1987), but included sloughed skin and stranded specimens. Samples were preserved in 95% Ethanol or salt saturated 20% Dimethyl Sulfoxide solution (DMSO) and later stored at –20ºC until processed. Total genomic DNA was extracted from the skin tissues using proteinase K digestion followed by a standard Phenol/Chloroform extraction method (Sambrook et al. 1989) or using DNAeasy tissue kit (Qiagen) following manufacturer's protocol. Sex determination was carried out by PCR amplifications, followed by TaqI 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 (Bérubé and Palsbøll, 1996).

### Mitochondrial DNA sequencing

A segment of 550 bp fragment within the mtDNA control region (Kocher *et al*., 1989; Baker *et al*., 1993) was amplified by the Polymerase Chain Reaction (PCR) with primers Dlp 1.5 and Dlp 5 described by Baker *et al.* (1993). Reactions of 25µL total volume containing 50mM KCL, 10mM Tris-HCL pH 8.8, 2.5 mM MgCl2, 200µM of each dNTP, 1.0 µM of each primer, and 0.05 U/µl Taq polymerase were conducted either on a Perkin-Elmer thermocycler or an Eppendorf Gradient Mastercycler following standard PCR procedures (94˚C for 4 min for initial denaturing, followed by 30 cycles of 94˚C for 45 s, 54˚C for 45 s, and 72˚C for 45 s). Amplified PCR products were cycle sequenced with dye-labeled terminators. Reactions were run on an ABI3700 or ABI3730 DNA analyzer (Applied Biosystems, Inc). PCR amplifications included negative control reactions to check for exogenous contamination.

### Microsatellite genotyping

Ten microsatellite loci, which have 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 & Amos, 1996). The 5'-end of the forward primer from each locus was labeled with a fluorescent tag (HEX, 6-FAM, and TET, Qiagen-Operon; NED, Applied Biosystems, Inc). PCRs were carried out in a 20µl 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 (ROX, Applied Biosystems, Inc) on an ABI 3700 or ABI 3730 DNA analyzer (Applied Biosystems, Inc). Microsatellite alleles were identified by their sizes in base pairs using the software GENEMAPPER 4.0 (Applied Biosystems, Inc).

### Sample size

To limit genotyping errors, some specific guidelines were used during laboratory work and scoring procedures. First, given the elevated number of samples, 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 GENOTYPER 2.1 and GENEMAPPER 4.1, and allele sizing was successively checked by hand. Samples that yielded ambiguous allele peaks were repeated a second time. The presence of null alleles was be monitored comparing the genotypes of known mother-offspring pairs. For detecting errors in our database, we used the program DROPOUT 1.3 (McKelvey & Schwartz, 2005), that identifies both loci and samples that likely contain errors affecting capture-

mark-recapture; this program uses a "*bimodal test*" that enumerates the number of loci different between each pair of samples, and provides information to determine the source of the errors, and uses a "*difference in capture history test*" to determine those loci producing the most errors; this test allows one to determine that a data set is error-free.

Duplicate samples within each population were detected from microsatellite genotype identity using MS_TOOLKIT (Park, 2001) and DROPOUT 1.3, and were excluded from the analysis. The probability of different individuals in each population sharing the same genotype by chance (Probability of Identity, PI) was estimated using the Excel add-in GENEALEX 5.1 (Peakall and Smouse, 2001).

**Statistical analysis of genetic variation**

DNA sequence variation patterns were characterized into mtDNA haplotype definitions for this species as previously recorded in Rosenbaum *et al*. (2002). From the 520 bp mtDNA Control Region fragment, a 477 bp consensus region that contains the majority of variable nucleotide positions in the mtDNA control region of humpback whales was examined for all samples (Baker *et al*., 1993). Sequences were aligned and assembled using *Sequencher 4.5* (Gene Codes, Inc). Sequences for this portion of the mtDNA Control Region were maintained for each individual in *MacClade 4.08*. Matching of sequences to a haplotype was done using *Collapse 1.2*.

The diversity of humpback whale mtDNA sequences was estimated at both the haplotype and nucleotide level (Nei, 1987) using *Arlequin 3.01*. (Schneider *et al*., 2000). At the haplotype level, diversity and its standard error were calculated without reference to the genetic distance (i.e., number of nucleotide substitutions) between two mtDNA sequences. At the nucleotide level, diversity (Nei, 1987) and its standard error for both sampling and stochastic processes were calculated from the pairwise differences between the mtDNA sequences.

Microsatellite variation was measured using the mean number of alleles per locus ($K$), the observed heterozygosity ($H_o$), and the heterozygosity expected ($H_e$) under Hardy–Weinberg assumptions (Nei, 1987) using the Excel add-in MS_TOOLKIT package (Park, 2001). Deviations from HWE and LD were tested by the Markov chain method implemented in GENEPOP (Guo and Thompson, 1992; Raymond and Rousset, 1995).

**Statistical analysis of population structure**

For mtDNA sequences, genetic variation was estimated as haplotype diversity ($h$) and nucleotide diversity ($\pi$) (Nei, 1987) over all samples, for each sampling site using ARLEQUIN 3.0.1 (Excoffier *et al.,* 1992, Schneider *et al.,* 2000). Nucleotide diversity and its standard error were calculated from pairwise differences between mtDNA sequences and haplotype diversity and its standard error were calculated without regard to the genetic distance. The diversity and geographic variation of haplotypes were quantified using the Analysis of Molecular Variance procedure (AMOVA; Excoffier *et al.,* 1992) as implemented in the software ARLEQUIN 3.01. AMOVA was performed incorporating both $\Phi_{ST}$ and $F_{ST}$. The $\Phi_{ST}$ takes into account the relationships between haplotypes based on molecular distance (Excoffier *et al.,* 1992), whereas the $F_{ST}$ considers only the difference in overall haplotype frequencies (Weir and Cockerham, 1984). The statistical significance of these values was tested by 10000 permutations.

For microsatellites, AMOVA was performed using ARLEQUIN 3.01. The distances between microsatellite genotypes is estimated either by counting the number of different alleles between two genotypes, the equivalent of estimating weighted $F_{ST}$ over all loci (Weir and Cockerham, 1984), 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).

For tests of spatial structure, we tested the B1 (samples from Gabon, São Tomé, Angola) and B2 (samples from west coast of South Africa) division. For the temporal analysis the data were partitioned according to the seasons for each area, taking into account the data available for these regions and the sampling months that we have for these analyses we only used Gabon and South African data. Based on distributional data, whales sampled off the coast of Gabon were divided into "Early" season (July and August) and "Late" season (September and October). For west South Africa the season was divided as "Regular" season (July- October) and "Late" season (November-February). In this way the regular season in South Africa corresponded to the entire season in Gabon. Data were further divided into strata to evaluate the effect of sex (males, females).

**Measures of dispersal**

To estimate recent, asymmetric of gene flow rates we used BayesAss program (Wilson and Rannala, 2003). This program estimates recent migration rates between populations from multilocus genotype data using a Bayesian inference framework. This method requires fewer assumptions than estimators of long-term gene flow. The analysis was run three times to check for consistency of results.

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 consisting in searching for genotype matches between different areas (Waits *et al*, 2001). We used the Excel add-in MS_TOOLKIT package (Park, 2001) and DROPOUT 1.3 program (McKelvey & Schwartz, 2005).

## RESULTS

**Sample size and sex ratio**

Table I illustrates the sample sizes for each sampling site and for the divisions described above, and the number of known males and females. The 2018 tissue samples were inferred to represent 1640 different whales. Average probability of identity (PI) for each population was small enough to exclude within-site re-samples with confidence. In Gabon the sex ratio is biased towards males (2.1M:1F), while in South Africa the sex ratio is close to parity (1M:1.1F).

**Genetic variation**

A consensus region of 477 bp of the mitochondrial DNA control region was assembled in which 134 maternal haplotypes were detected from 64-81 polymorphic sites. Haplotype diversity and nucleotide diversity estimates are in Table II.

All 10 microsatellite loci were highly polymorphic, with the number of alleles per locus ranging from 4 (Ev1) to 22 (GATA417), with a mean of 11.60 for B2 and 13 for B1. No significant differences were found between the observed heterozygosity ($H_0$) and the heterozygosity expected ($H_e$) under Hardy-Weinberg assumptions across loci (Table III).

**Population Structure**

For B1-B2 analysis, significant differentiation occurred between areas at haplotype level ($F_{ST}$= 0.00741, $p$=0.0000) and nucleotide level ($\Phi_{ST}$= 0.00468, $p$=0.00693;). Similar results were obtained with microsatellites, when we tested in the AMOVA, both $F_{ST}$ (0.00183, p=0.0000) and $R_{ST}$ (0.00275, p=0.00723)

For the temporal analysis, significant differentiation was founded between seasons at haplotype level ($F_{ST}$= 0.00294, $p$=0.0000), but not at nucleotide level. In the pairwise comparisons based on haplotype frequencies differentiation was founded in all seasons between the two regions. (Table IV), and between "Late" season in west South Africa (WSA) and both seasons in Gabon related with molecular distances (Table IV).

When the samples were stratified by sex, we founded significant differentiation at the haplotype level ($F_{ST}$=0,00451, $p$=0.0000) and nucleotide level ($\Phi_{ST}$= 0,00619, $p$=0.027) for females; for males only haplotype frequencies ($F_{ST}$=0,00291, $p$=0.0001) showed differentiation. In pairwise comparisons Females from "Late" season in west South Africa were significantly different from Gabon females (independent of season) and females from "Regular" season in South Africa were different from females in "Late" season in Gabon (Table V). For males (Table VI), the differences were only between males from the two regions.

**Measure of dispersal**

The migration rate from B1 to B2 was much lower ($m$=0.046; confidence interval=0.015-0.099) than the migration rate in the opposite direction ($m$= 0.317; confidence interval= 0.304-0.327). Multiple runs showed consistent results, and the 95% CI suggested that the data contained sufficient information for reliable migration rate estimates.

The genetic capture-recapture approach recovered a total of 11 matches (Table VII) between Gabon and west

South Africa coast. Two of these had already been described by Pomilla *et al*. (2006). All the movements were from a northbound to southbound direction when considering month of year (and not year itself). Four of them were re-sampled in the same year and the other seven were re-sampled in different years (maximum interval of eight years between the capture and re-capture). Seven matches involve females (one of them with a small calf) and four males. Two individuals were sampled in Gabon and then were re-sampled twice (each of them) in South Africa. Two of the individuals sampled in Gabon were re-sampled in South Africa in a situation related with feeding behavior (feeding or defecation). Two individuals were sampled in west South Africa and then were re-sampled later (five year's interval) in Gabon.

## DISCUSSION

As in our previous report (Carvalho *et al.* 2009), the tests for population differentiation of B1 and B2 were significant, based on mtDNA haplotype frequencies and molecular distance, and microsatellite data. Previous analyses of mtDNA and nuclear DNA of B stock have shown significantly differentiation between B1 and B2 sub-stocks (Rosenbaum *et al*. 2009; Pomilla, 2005, Pomilla *et al.* 2006), however in these studies the differences were only found at haplotypic level (Rosenbaum *et al.* 2009) and for microsatellite data when Angola and Gabon were grouped together (Pomilla *et al.* 2006). These differences with previous data could be explained by the use of a much larger dataset from Gabon and South Africa in analyses presented here, which provide greater resolution.

Different estimations of gene flow conducted between these two areas show that the expected exchange is very low, when compared with estimations in other regions (sometimes in oceans apart), in the order of 1 $N_em$ (migrants per generation) for nuclear DNA (Pomilla et 2006) and 0.386 $N_em$ for mithocondrial DNA (Rosenbaum *et al.* 2009) when the direction is between B1 to B2 and slightly higher in the opposite direction (B2 to B1). We got consistent results with the previous ones when calculated the recent migration rates, and again with higher values when the direction is from B2 to B1. However connectivity between these two areas was demonstrated through detection of movements by genetically identified individuals between Gabon and South Africa, both for females and males. All of the detected movements were from a northbound to southbound 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 (July/Aug in Gabon and October-November-December in South Africa) reflects an overall pattern of movement of some animals from B1 to B2. However this could be just an artifact of the sample effort in west South Africa, since less effort was completed when the animals were migrating in a northerly direction (prior to the Regular season in West South Africa).

These results seem in certain way to be confounding. However they could be explained by spatial or temporal patterns in this stock. Based on catch histories, Findlay (2000) suggested a degree of spatial segregation of this stock, with some whales using alternate migration routes and thus avoiding capture, (e.g. some animals migrate further offshore and others inshore). Recent telemetry data indicate the same pattern; two whales tagged in Gabon traveled offshore at the Walvis Ridge and passed west South Africa far offshore *en route* to Bouvet Island, arriving in polar latitudes at around 0º longitude (Rosenbaum & Mate, 2006), suggesting that humpbacks whales wintering in the Gulf of Guinea region use at least two different migration routes. We found some degree of temporal heterogeneity for these two regions, significant differentiation at haplotype level, the differences were found among seasons between areas. However in pairwise comparisons among the seasons we found that animals that are in Gabon (independent of the season) were significantly different from the ones that are sampled in summer (December to February) in west South Africa, for both statistics. These results were slightly different from the results from the last report (Carvalho *et al.* 2009) and could be explained by the incorporation of new South African samples, which were almost entirely collected during November (Late Season). This heterogeneity means that west South Africa coast is not a feeding ground destination for the majority of whale's migration from Gabon. Recent data from humpback whales movements and seasonality in west South Africa coast, showed mid-spring (October) to be a turning point in their behaviour, after which a significant reduction in swimming speed, an increase in non-directional movement and a distribution further from shore were observed (Barendse *et al.*, 2010). Our results are consistent with this behaviour change, since our late season (November-February) in west South Africa is after this alteration and are the season that shows more genetic differentiation at all levels with Gabon.

When we analyzed the data stratified by sex, significant results were obtained among seasons between areas, for females and for males only at haplotype level. Based on haplotype frequencies and molecular distances, females that occurred in summer months in west South Africa were different from females from Gabon,

independent of season. In Gabon the sex ratio is highly biased toward males, especially in the Late season (2.4M:1F). This could be explained by sampling bias (because big competitive groups are easier to detect), group structure segregation (because not all females migrate to breeding areas each year) (Chittleborough, 1965: Dawbin, 1997; Brown et al 1995) or by some spatial structure among groups within the Gulf of Guinea breeding area, where some groups of animals (like mother-calves or pregnant females) prefer some areas more than others (Smultea, 1994; Morete *et al.*, 2007). In fact differences occur between percentages of mother-calves pairs observed in Gabon (4.6%) (Rosenbaum & Collins, 2006) and off shore in São Tomé and Principe archipelago (65.15%) (Carvalho *et al.* submitted). Conversely, the sex ratio on the west coast of South Africa (1M:1.2F – Regular season and 1M:1.3F – Late season) is much more like that of a feeding area, where normally it is biased toward females (Mackintosh,1942; Brown *et al*.,1995) A lower or even a lack of differentiation for male was expected since the sex with higher dispersal (males in the case of humpback whales) will have a lower between-subpopulations $F_{ST}$ value than the sex that is dispersing less (Prugnolle & Meeus, 2002). For species in which females are philopatric and males disperse, genetic differentiation between populations is expected to be higher when estimated using mtDNA (or another maternal marker) (Prugnolle & Meeus, 2002); comparisons with microsatellite data may be more useful, but contain bi-parental contributions to genetic variance.

The results presented here indicate that there is some temporal and spatial population substructure in whales in B stock and seems to suggest that a portion of whales feeding off west South Africa are breeding and calving in a area which has not yet been surveyed. Additional samples from other regions between these two areas and in other parts of Gulf of Guinea, along other areas in west coast of Africa and further west (e.g. Bamy *et al.* 2010) could give additional information regarding with B1-B2 stock structure.

## ACKNOWLEDGEMENTS

We are grateful to the staff of the WCS/AMNH New York Staff and WCS Country Programs of Madagascar and Gabon, in particular George Amato, Rob DeSalle, Eleanor Sterling, Helen Crowley Matthew Hatchwell and Lee White, and all the summer interns, students, volunteers and research assistants who have contributed to field and laboratory work, in particular Carla Freitas, Nic Mbadinga and Vanessa Rasoamampianina. Guy-Philippe Sounguet of Aventure Sans Frontier, and Solange Ngouessono of the Department of Wildlife, Ministry of the Environment, provided valuable support in Gabon. Generous logistic support was supplied in Gabon by Rombout Swanborn and the personnel of Operation Loango. Peter Best would like to acknowledge the contributions of Meredith Thornton, Shaun Dillon and Simon Elwen to sample collection, and the Earthwatch Institute and Mazda Wildlife Fund for field support. Logistical assistance was generously provided by the South African Navy and the Military Academy, Saldanha. This material is based upon work supported by the National Research Foundation (South Africa) under Grant number 2053539. This work has the support of Portuguese Foundation for Science and Technology (Fundação para a Ciência e Tecnologia - PhD grant to Inês Carvalho - SFRH / BD / 18049 / 2004).

## LITERATURE CITED

Aguilar A (1985) Aboriginal whaling off Pagalu (Equatorial Guinea) *Reports to the* 2 *International Whaling Commission* **35**: 385-386.
Bamy, I., Van Waerebeek, K., Bah, S.S., Dia, M., Kaba, B., Keita, N. and Konate, S. (2010). Species occurrence of cetaceans in Guinea, including humpback whales with southern hemisphere seasonality. *Marine Biodiversity Records*, vol 3; e48.
Baker, C.S., Perry, A., Bannister, J.L., Weinrich, M.T., Abernethy, R.B., Calambokidis, J., Lien, J.,Lambertsen, R.H., Urbán, J., Vásquez, O., Clapham, P.J., Alling, A., O´Brien, S.J and Palumbi, S.R. 1993. Abundant mitochondrial DNA variation and world-wide population structure in humpback whales. *Proceedings Natural Academic Science* 90: 8239-8243.
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. *African Journal of Marine Science*, 32:1-22.
Berubé, M. & Palsbøll, P. 1996. Identification of sex in Cetaceans by multiplexing with three ZFX and ZFY specific primers. *Molecular Ecology*, 5, 283-287.
Best, P. B., D. Reeb, M. Morais, and A. Baird. 1999. A preliminary investigation of humpback whales off northern Angola. Paper SC/51/CAWS33 presented to the IWC Scientific Committee, 1999 (unpublished).
Best, P. B., K. Sekiguchi, and K. P. Findlay. 1995. A suspended migration of humpback whales Megaptera novaeangliae on the west coast of South Africa. *Marine Ecology Progress Series* 118, 1-12.
Brown, M., Corkeron, P., Hale, P., Schultz. K. and Bryden, M. 1995. Evidence for a sex-segregated migration in the humpback whale (*Megaptera novaeangliae*). *Proceedings of Royal Society of London, Serie B*, 259, 229-234.
Budker P, Collignon J (1952) Trois campagnes balenieres au Gabon: 1949-1950-1951. *Bulletin de l'Institut d'etudes centrafricaines*, **3**, 75-100.
Carvalho, I., Brito, C., dos Santos, M. and Rosenbaum H.C (*submitted*). The waters of São Tomé: a calving ground for West African humpback whales?
Carvalho, I; Loo, J.; Pomilla, C.; Leslie, M.; Collins, T.; Barendse, J.; Best, P. & Rosenbaum, H. (2009). Temporal patterns of population structure of humpback whales in west coast of Africa (B1-B2 sub-stocks) based on mitochondrial DNA and microsatellite variation. Paper SC_61_SH6 presented to the IWC the Scientific Committee, (unpublished).
Chapman, D.G. 1974. Status of Antarctic rorqual stocks. Pages 218-238. In: W.E. Schevill (ed.) *The Whale Problem*. Harvard University Press, Cambridge. 419pp .
Dawbin, W. 1997. Temporal segregation of humpback whales during migration in southern hemisphere waters. *Memoirs of the Queensland Museum*, 42, 105-138.
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.
Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479-491.
Findlay K.P, (2000) A Review of humpback whale caches by modern whaling operations in the Southern Hemisphere. *Memoirs of the Queensland Museum* **47,** 411-420.
Guo S, Thompson E (1992) Performing the exact test of Hardy Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361-372.
IWC. 1997. Report of the scientific committee. Report of the International Whaling Commission 47.
IWC. 2000. Report of the scientific committee. Report of the International Whaling Commission 52.
IWC. 2007. Report of the Workshop on the Comprehensive Assessment of Southern Hemisphere Humpback Whales. Paper SC/58/Rep5.
Lambertsen, R. H. 1987. A biopsy system for large whales and its use for cytogenetics. *Journal of Mammalogy* 68, 443-445.
Mackintosh, N. A. 1942. The southern stocks of whalebone whales. *Discovery Reports* 22, 197-300.
Matthews LH (1938) The humpback whale, *Megaptera nodosa*. *Discovery Reports*, **17**, 7-92.
McKelvey, K.S. & Schwartz, M.K. (2005). DROPOUT: a program to identify problem loci and samples for noninvasive genetic samples in a capture-mark-recapture framework. *Molecular Ecology 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. *Canadian Journal of Zoology*, 73, 1735-1743.
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
Morete, M.; Bisi, T. & Rosso, S. 2007. Temporal pattern of humpback whale (*Megaptera novaeangliae*) group structure around Abrolhos Archipelago breeding region, Bahia, Brazil. *Journal Marine Biological Association of United Kingdom*, 87, 87-92.
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.
Ølsen, O. (1914). Whales and whaling in South Africa. *Bergens Museum Årbok Naturvitenskap* 5:1-56. [Translated from Norwegian].
Palsbøll PJ, Bérubé M, A.H. L, Jørgensen H (1997b) Primers for the amplification of tri- and tetramer microsatellite loci in baleen whales. *Molecular Ecology*, **6**, 893-895.
Palsbøll PJ, Vader A, Bakke I, Raafat El-Gewely M (1992) Determination of gender in cetaceans by thepolymerase chain reaction. *Canadian Journal of Zoology*, **70**, 2166-2170.
Palsbøll, P. J., P. J. Clapham, D. K. Mattila, F. Larsen, R. Sears, H. R. Siegismund, J. Sigurjonsson, O. Vasquez, and P. Arctander. 1995. Distribution of mtDNA haplotypes in North Atlantic humpback whales: the influence of behavior on population structure. *Marine Ecology Progress Series* 116:1-10.
Park S (2001) MS_Toolkit: Excel add-in tool package for microsatellite data. Genetics Department, Trinity College, Dublin.
Peakall, R and P Smouse. 2001. GenAlEx V5: genetic analysis in Excel. Population genetic software for teaching and research,. Canberra, Australian National University
Picanço C, Carvalho I, Brito C (2009) Occurrence and distribution of cetaceans in São Tomé 1 and Príncipe tropical archipelago and their relation to environmental variables. *Journal* 2 *of Marine Biological Association of United Kingdom* **89**:1071-1076
Pomilla C., Best P.B., Findlay K., Kotze P.J.H., Engel M.H., Barendse J., and Rosenbaum H.C. 2004. Population structure of Southern Hemisphere humpback whales from wintering Regions A, B, and C based on nuclear microsatellite variation, pp. 12, Paper SC/56/SH4 presented to the IWC the Scientific Committee, (unpublished).
Pomilla C., Best P.B., Findlay K.P., Collins, T., Engel. M., Minton, G., Ersts, P., Barendse, J., Kotze, P.G.H., Razafindrakoto, Y., Ngouessono, S., Meyer, M., Thornton, M. and 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. SC/A06/HW38 presented to the IWC Scientific Committee Workshop on the Comprehensive Assessment of Southern Hemisphere humpback whales. April 2006, Hobart, Australia (unpublished).
Prugnolle, F. & Meeus, T. 2002. Inderring sex-biased dispersal from population genetic tools: a review. *Heredity*, 88, 161-165.
Raymond M and F Rousset. 1995. GENEPOP: population genetic software for exact tests and ecumenicism. *Journal of Heredity* **86**: 248-249.
Rosenbaum, H. C., P. Ersts, Y. Razafindrakoto, G. Sounguet, C. Pomilla, S. Ngouessono, V. Rasoamampianina, and L. White. 2002. Population characteristics, distribution, and relative abundance of humpback whales off the coasts of Madagascar and Gabon: an update on recent and planned research. Pp.13. Paper SC/54/H20 presented to the IWC Scientific Committee, (unpublished).
Rosenbaum, H.C. and 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. pp. 425-433. In: *Natural History of the Gamba Complex*. Alonso, A. and Campbell, P. (eds). Smithsonian Press, Washington D.C.
Rosenbaum, H.C. and Mate, B. (submitted) From north of the equator to the Antarctic: unique and unexpected movements for humpback whales off the coast of West Africa and throughout the eastern South Atlantic Ocean.
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.
Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: a Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, New York.
Schlötterer C, Amos B, Tautz D (1991) Conservation of polymorphic simple sequence loci in cetacean species., *Nature*, **354**, 63-65
Schlötterer, C., B. Amos, and D. Tautz. 1991. Conservation of polymorphic simple sequence loci in cetacean species. *Nature* 354:63-65.
Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN (Version 2.0): A software for population genetic data analysis. University of Geneva, Geneva, Switzerland.
Schneider, S., D. Roessli, and L. Excoffier. 2000. ARLEQUIN (Version 2.000): A software for population genetics data analysis. University of Geneva, Geneva, Switzerland.
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. *Canadian journal of Zoology*, 72, 805-811.
Takahata, N., and S R. Palumbi. 1985. Extranuclear differentiation and gene flow in the finite island model. *Genetic* 109:441-457.
Valsecchi E, Amos W (1996) Microsatellite markers for the study of cetacean populations. *Molecular Ecology*, **5**,151-156.
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 Scientific Committee of the IWC, (unpublished).
Van Waerebeek K, Ofori-Danson P.K. and Debrah J. (2009). The cetaceans of Ghana: a validated faunal checklist. *West African Journal of Applied Ecology* 15:61-90.
Waits L, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*, **10**, 249-256.
Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, **38**:1358-1370.
Wilson, G. and Rannala, B. (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163:1177-2291.

**Figure 1**- B1 and B2 areas in the South Atlantic Ocean.

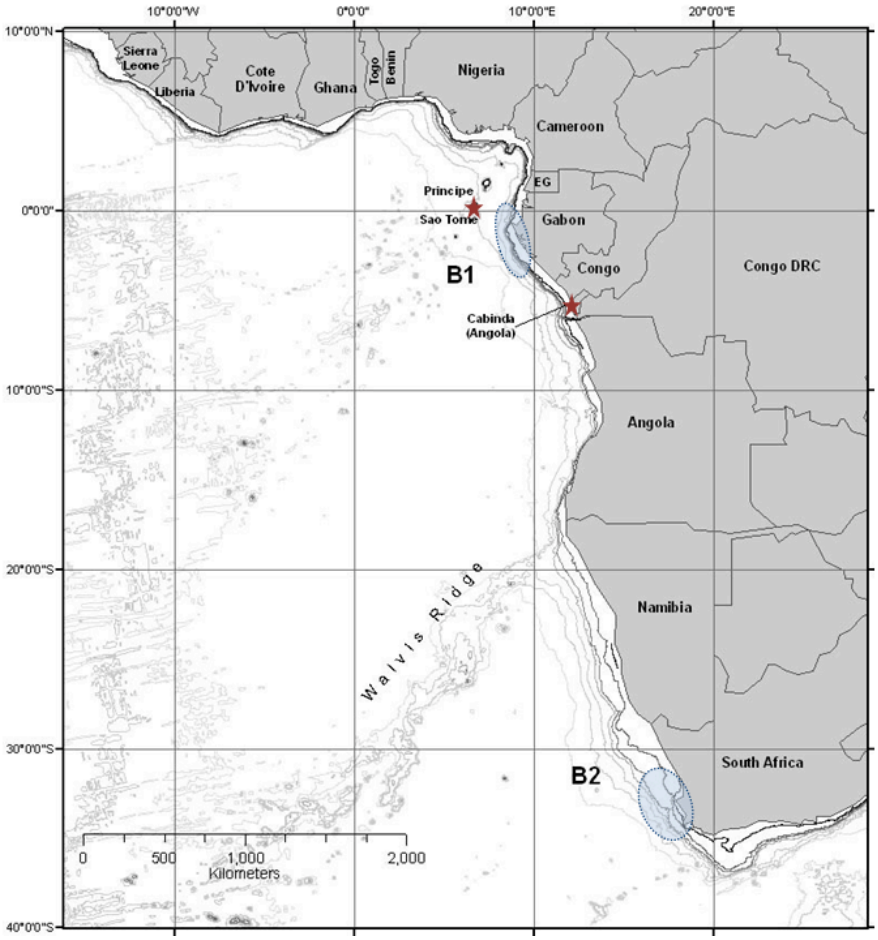


**Table I** – Years of sampling and number of samples collected in west coast of Africa. Number of known males (M) and female (F) individuals is also shown.

| Sampling Regions | Years | Total Samples | M | F |
|---|---|---|---|---|
| **B1** | | | | |
| São Tomé (ST) | 2004-2005 | 5 | 1 | 4 |
| Angola – Cabinda (ANG) | 1998 | 13 | 8 | 4 |
| Gabon (GA) | 2000-2006 | 1696 | 992 | 460 |
| **B2** | | | | |
| West South Africa (WSA) | 1990-2009 | 304 | 122 | 139 |

**Table II** – Mitochondrial DNA control region variability. Sampling site, number of individuals (N), number of haplotypes, polymorphic sites, haplotype (*h*) and nucleotide ($\pi$) diversities, as well as their standard deviations (SD).

| Sampling site | N | Haplotypes | Polymorphic sites | *h* ± SD | $\pi$ ± SD |
|---|---|---|---|---|---|
| **B1** | 1336 | 129 | 81 | 0,980±0,000 | 0,0213±0,011 |
| **B2** | 251 | 71 | 64 | 0,9712±0,004 | 0,021±0,011 |

**Table III** – Genetic variability in humpback whale genotyped at ten loci. *N* = mean number of genotyped individuals per locus; *K* = mean number of alleles per locus; $H_o$ and $H_e$ =observed and expected heterozygosities (SD- standard deviation). *PI* = probability of identity.

| Sampling site | *N* | *K* | $H_o$ ± SD | $H_e$ ± SD | PI |
|---|---|---|---|---|---|
| **B1** | 1495 | 13.00 | 0,743±0,0036 | 0,745±0,049 | $1.9\times10^{-12}$ |
| **B2** | 263 | 11.60 | 0,751±0,008 | 0,740±0,051 | $3.2\times10^{-12}$ |

**Table IV** – Pairwise comparisons between **Seasons** in Gabon and West South Africa, using molecular distances ($\phi_{ST}$) (above the diagonal in italic) and haplotype frequencies ($F_{ST}$) (below the diagonal). Significant values are highlighted in bold ($p$ <0.05).

| | Early Gabon | Late Gabon | Late WSA | Regular WSA |
|---|---|---|---|---|
| **Early Gabon** | | *-0.00076* | ***0.00524*** | *0.00111* |
| **Late Gabon** | 0.00002 | | ***0.00510*** | *0.00013* |
| **Late WSA** | **0.00750** | **0.00958** | | *0.00380* |
| **Regular WSA** | **0.00350** | **0.00485** | -0.00124 | |

**Table V** – Number of **females** in each season (N). Pairwise comparisons, for females, between different seasons in Gabon and West South Africa, using molecular distances ($\phi_{ST}$) (above the diagonal in italic) and haplotype frequencies ($F_{ST}$) (below the diagonal). Significant values are highlighted in bold ($p$ <0.05).

| | N | Early Gabon | Late Gabon | Late WSA | Regular WSA |
|---|---|---|---|---|---|
| **Early Gabon** | 287 | | *0.00150* | *0.00719* | *0.00762* |
| **Late Gabon** | 122 | 0.00132 | | ***0.01156*** | ***0.02124*** |
| **Late WSA** | 115 | **0.00631** | **0.01100** | | *-0.00222* |
| **Regular WSA** | 47 | 0.00358 | **0.00957** | -0.00167 | |

**Table VI** – Number of **males** in each season (N). Pairwise comparisons, for males, between different seasons in Gabon and West South Africa, haplotype frequencies ($F_{ST}$) (below the diagonal). Significant values are highlighted in bold ($p$ <0.05).

| | N | Early Gabon | Late Gabon | Late WSA | Regular WSA |
|---|---|---|---|---|---|
| **Early Gabon** | 585 | | | | |
| **Late Gabon** | 294 | -0.0003 | | | |
| **Late WSA** | 90 | **0.00893** | **0.01085** | | |
| **Regular WSA** | 40 | **0.00712** | **0.00750** | 0.0078 | |

**Table VII** – Matches founded between Gabon (GA) and west coast of South Africa (WSA). Sex, date of collection in each site and the situation when the sample was collected.

| Matches | Sex | GA date | Situation | WSA date | Situation |
|---|---|---|---|---|---|
| 1 | F | **08/09/2001** | Pair | **16/12/2001** | 2 Adults/ Defecation |
| | | | | **17/11/2009** | |
| 2 | F | **09/09/2002** | M-C pair | **10/01/2003** | M-C pair |
| | | | | **25/10/2004** | 2 adults |
| 3 | M | **07/08/2001** | ? | **21/11/2005** | Male with a M-C pair |
| 4 | F | **16/08/2002** | Pair | **07/11/2006** | 2 adults |
| 5 | M | **04/09/2003** | Competitive | **18/01/2003** | 2 adults |
| 6 | F | **26/09/2004** | Competitive | **08/11/2004** | 2 Adults/ Defecation |
| 7 | M | **14/08/2005** | Competitive | **17/10/2005** | 3 adults |
| 8 | F | **20/08/2006** | Competitive | **12/10/2001** | 2 ind in mixed group of whales |
| 9 | M | **04/09/2006** | Competitive | **17/12/2001** | 20 ind./milling and feeding |