Update of the genetic analysis on stock structure of fin whales in the Antarctic based on mitochondrial and microsatellite DNA (SC/65b/J32rev)

Mutsuo Goto, Naohisa Kanda and Luis A. Pastene

Institute of Cetacean Research, 4-5 Toyomi-cho, Chuo-ku, Tokyo 104-0055, Japan

ABSTRACT

This paper is a revised version of SC/F14/J32, which respond to some recommendations from the JARPAII review Workshop. Genetic samples (catches and biopsies) of fin whales obtained by JARPA and JARPAII were analyzed with two genetic markers, mtDNA control region sequencing (479bp-segment) and microsatellite DNA (16 loci), to investigate stock structure of this species in the Antarctic feeding grounds. Genetic samples were available from Areas IIIE (n=6), IV (n=23), V (n=24) and VIW (n=2). No statistical significant difference in mtDNA haplotype frequencies was found between Areas IIIE+IV and Areas V+VIW. Large number of singletons and small sample sizes could have decreased the power of the mtDNA statistical analysis. The microsatellite analysis showed a statistically significant deviation from the Hardy-Weinberg equilibrium in Area V, and the heterogeneity test showed significant differences between Areas IV and V. Results of the genetic analyses therefore suggested the possibility of genetic structuring of fin whales in the JARPAII research area, which should be further explored with the analyses of a large number of samples in the future.

KEYWORDS: ANTARCTIC, GENETICS, FIN WHALE, SCIENTIFIC PERMITS

INTRODUCTION

Little information is currently available on the stock structure of fin whales in the Antarctic. As in the case of the blue whale, earlier mark-recapture analysis showed that most whales return to the same part of the Antarctic year after year (Brown, 1954). Subsequent mark-recapture studies conducted by Brown (1962) suggested that the six whaling areas are probably more valid for blue and humpback whales than for fin whales (see also Mackintosh, 1965). The past information suggested there was certain segregation of fin whales in the feeding ground between certain longitudes in four sectors which lie: South of the Atlantic Ocean, South of the Indian Ocean, South of Western South Pacific Ocean and South of Eastern South Pacific Ocean (Mackintosh, 1965). South of the Indian Ocean correspond approximately to JARPAII Areas IIIE and IV and South of Western South Pacific to JARPA and JARPAII Areas VIW and V. It is important to investigate whether such geographical segregation is supported by genetic differences.

Wada and Numachi (1991) conducted allozyme analysis using North Pacific, Spanish coastal and Antarctic fin whales. They showed significant allele frequency differences between Hemispheres. However they could not detect evidence of more than one stock within the Antarctic or within the North Pacific fin whales.

Only a single study based on JARPA biopsy samples and mtDNA has been conducted to examine genetic differences between the whales from IIIE+IV and VW (Pastene *et al*., 2005). Although the authors found no evidence of the genetic differences between IIIE+IV and VW, the sample size was too small (8 and 15, respectively) to make a firm conclusion on stock structure of this species in the Antarctic.

Additional genetic samples were obtained during the JARPAII surveys. The analysis of those samples is important for testing the hypothesis of segregation in the feeding grounds of the IWC management areas.

In this study we conducted mtDNA and microsatellite analyses using additional biopsy and research samples obtained by JARPA and JARPA II to investigate further the stock structure of fin whales in Areas IIIE, IV, V and VIW.

This paper is a revised version of SC/F14/J32, which respond to the recommendations of the JARPAII

Review Workshop. The Workshop recommended that an updated paper be provided that:

(1) omits the analysis using *post-hoc* groupings unless the authors can provide a robust statistical justification for this approach;
(2) discusses in more detail the fact that three microsatellite loci showed significant departures from Hardy-Weinberg (HW) expectations after accounting for multiple testing before the data are used in other analyses;
(3) provides F_{IS} values for each locus in table 7 of SC/F14/J32, which would indicate whether the HW departures represent an excess or deficiency of heterozygotes; (that information is essential to understand the possible causes for the deviations);
(4) shows how the samples were arranged into management Areas in fig.1; and
(5) recognises that conclusions about possible differences between Areas IV and V are provisional until the nature, cause, and influence of these HW deviations are resolved.

MATERIALS AND METHOD

Samples

Genetic samples were available from fin whales caught by JARPAII surveys between 2005/06 and 2010/11 and from biopsies obtained from the sighting surveys of the JARPA and JARPAII, on an opportunistic basis. Table 1 and Fig. 1 show the number and geographical position of the genetic samples used in this study, by year and Area.

Laboratory procedures

Total DNA was extracted from 0.05g of skin tissue using the protocol of Sambrook *et al*. (1989). Extracted DNA was stored in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

MtDNA

The first half of the control region of the mitochondrial genome was amplified by using the polymerase chain reaction (PCR). For the amplification of approximately 500bp of the mtDNA control region, primers MT4 (Árnason *et al*., 1993) and Dlp 5R (*5′*-CCA TCG AGA TGT CTT ATT TAA GGG GAA C-*3′*), were used. Reactions were carried out in 25 uL volumes containing 10-100ng of DNA, 2.5 pmole of each primer, 0.5 units of Ex Taq DNA polymerase (Takara), 2mM of each dNTP, and 10x reaction buffer. After an initial denaturation step at 95° C for 5 minutes, a PCR amplification cycle of 30 seconds at 94°C, followed by 30 seconds at 50°C and 30 seconds at 72°C are repeated 30 times. The amplification is completed with a final extension step of 10 minutes at 72°C. Subsequent cycle sequencing reactions are performed with 100ng of products generated in the above PCR amplifications using the BigDye® Terminator Cycle Sequencing Kit (Applied Biosystems). The oligo-nucleotides used to prime the cycle sequencing reaction were the same as employed in the initial PCR amplification listed above. A total of 25 cycles for 10 seconds at 96°C, 20 seconds at 56°C and four minutes at 60°C are performed. The nucleotide sequence of each cycle sequencing reaction is determined using Applied Biosystems 3500 Genetic Analyzer (Life Technology) under standard conditions. Both strand samples are sequenced in their entirety for all samples.

Microsatellite

Genetic variation at microsatellite loci were analyzed using 16 sets of primers as follow: EV1, EV14, EV21, EV94, EV104 (Valsecchi and Amos, 1996), GT011 (Bérubé *et al*., 1998), GT23, GT211, GT271, GT310, GT575 (Bérubé *et al*., 2000), GATA28, GATA53, GATA98, GATA417, GGAA520 (Palsbøll *et al*., 1997), and DlrFCB17 (Buchanan *et al*., 1996). Primer sequences and PCR cycling profiles generally followed those of the original authors. PCR amplifications were performed in 15ml reaction mixtures containing 10-100ng of DNA, 5 pmole of each primer, 0.625 units of Ex Taq DNA polymerase (Takara Shuzo), and 2mM of each dNTP, and 10x reaction buffer containing 20mM $MgCl_2$ (Takara Shuzo). Amplified products with internal size standard (GENESCAN400HD, Applied Biosystems Japan) were run on a 6% polyacrylamide denaturating gel (Long Ranger) using BaseStation100 DNA fragment analyzer (Bio-Rad). Although alleles were visualized using Cartographer software specifically designed for the BaseStation, allelic sizes were determined manually in relation to the internal size standard and fin whale's DNA of known size that were rerun on each gel.

Statistical analysis

MtDNA

Sequences were aligned initially using Sequence Navigator (Applied Biosystems, Inc). The aligned

sequences were then corrected by eye for minor inconsistencies.

Haplotype and nucleotide diversity were calculated following Nei (1987). Conventional hypothesis testing procedure was conducted based on the randomized chi-square test of independence (Roff and Bentzen, 1989) to assess genetic differences among Areas. *Fst* was calculated using the AMOVA (Excoffier *et al*., 1992). Furthermore heterogeneity tests were conducted as described in Hudson *et al.* (1992), using the *Hst* and the *Kst** statistics. The level of statistical significance was estimated from 10,000 Monte Carlo simulations as the proportion of simulations in which a similar or more extreme value of *Hst* or *Kst** was observed.

Microsatellites

The number of alleles per locus and expected heterozygosity per locus was calculated using FSTAT 2.9.3 (Goudet, 1995). Statistical tests for the deviations from expected Hardy-Weinberg genotypic proportions, as well as their F_{IS} values (Weir and Cockerham, 1984) were calculated, using GENEPOP 4.0 (Rousset, 2008).

Probability test (or Fisher's exact test) implemented in GENEPOP 4.0 (Rousset, 2008) was used to conduct the heterogeneity tests. When simultaneous multiple tests were conducted, correction for the multiple tests was performed (Rice, 1989). Statistical significance was also determined using the chi-square value obtained from summing the negative logarithm of p-values over the 16 microsatellite loci (Sokal & Rohlf 1995). A p-value smaller than 0.05 was used as a criterion to reject the null hypothesis of panmixia.

RESULTS AND DISCUSSION

MtDNA

Diversity

A segment of 479bp of the mtDNA control region was sequenced for all samples. A total of 32 polymorphic sites identified a total of 45 unique sequences (haplotypes) in the total sample of 55 fin whales (Table 2). Except for two transversion sites, all substitutions were transitions.

Level of mtDNA diversity based on nucleotide and nucleon diversity is shown in Table 3. Both indices showed a high and similar degree of mtDNA diversity through the Areas. Nucleotide diversity estimate for the total sample was 0.0121 (Table 3).

Heterogeneity test

No significant differences were found in the statistical comparison between Areas IIIE+IV and Areas V+VIW (Table 4). Values of *Fst, Hst* and *Kst** were negative. It is possible that the large number of haplotypes represented by a single individual (singletons) summed to the small sample size had decreased the power of the statistical analysis.

Microsatellite

Diversity

Table 5 shows the number of alleles, expected heterozygosity, test results for expected Hardy-Weinberg genotypic proportions and inbreeding indices (F_{IS}) at 16 microsatellite loci by the total sample and by the different areas. Both average number of the alleles and average heterozygosity over the 16 loci were the highest in Area V, so that there were alleles unique to this area. Three loci (GAT98, EV21, TAA31) after correction for the simultaneous multiple tests and overall value showed significant deviation from the expected Hardy-Weinberg genotypic proportions. F_{IS} values of these three loci were all positive, indicating homozygote excess. Although these F_{IS} values were relatively small figures, the homozygote excess suggested existence of individuals from multiple stocks in our samples. It is important to notice that homozygote excess was not caused by the null allele. Evidence of the deviation from the expected Hardy-Weinberg genotypic proportions were detected only in the Area V with stronger indication of the homozygote excess when the tests were conducted for each of the areas separately. It should be noticed that there was a single individual in the area V sample having one or two alleles unique to only this individual at eight of the 16 loci. Although this fact could have caused the observed significant deviation from the expected Hardy-Weinberg genotypic proportions in the area, removal of this particular individual unchanged the test result (data not shown).

Heterogeneity test

Pair-wise comparisons among the areas showed the statistically significant differences between whales

from Areas IV and V, but not between IIIE and IV, and between IIIE and V, most likely due to the small sample size of Area IIIE. (Table 6).

Stock structure

One of the genetic markers (microsatellite) detected significant genetic heterogeneity between Areas IV (Indian Ocean) and V (Pacific Ocean). Significant departure from the Hardy-Weiberg equilibrium was detected in Area V. In contrast to the past genetic studies (Wada and Numachi, 1991; Pastene *et al*., 2005), our genetic study raised the possibility of genetic structuring in the JARPAII research area. One of the interpretations for the results based on microsatellite is that at least two stocks distribute in the research area, one in Area IV and the other in Area V, which possible mixing of different stocks in Area V.

Our study implies that effective management of fin whales in Antarctic should be conducted Area by Area although it is premature yet to conclude on their stock structure in the research area. In order to better understand the stock structure of fin whales, further genetic analyses should be conducted in the future based on larger sample sizes.

ACKNOWLEDGEMENTS

We would like to thank the researchers and crew members of the JARPA and JARPAII surveys for their efforts in collecting the samples. Our gratitude goes to H. Oikawa and S. Azumi for their laboratory assistance.

REFERENCES

Arnason, U., A. Gullberg, and B. Widegren. 1993. Cetacean mitochondrial DNA control region: sequences of all extant baleen whales and two sperm whale species. *J. Mol. Biol Evol*. 70: 960-970.

Bérubé, M., Aguilar, A., Dendanto, D., Larsen, F., Notarbartolo-di-Sciara, G., Sears, R., Sigurjónsson, J., Urban-Ramirez, J., and Palsbøll, P.J. 1998. Population genetic structure of North Atlantic, Mediterranean Sea and Sea of Cortez fin whales, *Balaenoptera physalus* (Linnaeus, 1758); analysis of mitochondrial and nuclear loci. *Mol. Ecol*. 7:585-599.

Bérubé, M., Jørgensen, H., Mcewing, R., and Palsbøll, P.J. 2000. Polymorphic di-nucleotide microsatellite loci isolated from the humpback whale, *Megaptera novaeanglliae. Mol. Ecol.* 9:2181-2183.

Brown, S.G. 1954. Dispersal in blue and fin whales. *Discovery Reports*. Vol. XXVI:355-384.

Brown, S.G. 1962. The movement of fin and blue whales within the Antarctic zone. *Discovery Reports*. Vol XXXIII:1-54.

Buchanan, F.C., Friesen, M.K., Littlejohn, R.P., and Clayton, J.A. 1996. Microsatellites from beluga whale *Delphinapterus leucas. Mol. Ecol*. 5:571-575.

Excoffier, L., Smouse, P.E. and Quatro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479-91.

Goudet, J. 1995. FSTAT, version 1.2: a computer program to calculate F-statistics. *J. Hered.* 86:485-486.

Hudson, R.R., Boos, D.D. and Kaplan, N.L. (1992). A statistical test for detecting geographic subdivision. *Mol Biol Evo* 9:138-151.

Nei, M. 1987. *Molecular Evolutionary Genetics. Columbia University Press, New York. x+512pp.*

Palsbøll, P.J., Bérubé, M., Larsen, A.H., and Jørgensen, H. 1997. Primers for the amplification of triand tetramer microsatellite loci in baleen whales. *Mol. Ecol.* 6:893-895.

Pastene, L.A., Goto, M., Kanda, N. and Nishiwaki, S. 2005. Genetic analyses on stock identification in the Antarctic humpback and fin whales based on samples collected under the JARPA. Paper JA/J05/JR16 submitted to the JARPA review meeting. January 2005. (unpublished) 13pp.

Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution* 43:223-225.

Roff, D.A. and Bentzen, P. 1989. The statistical analysis of mtDNA polymorphisms: chi-square and the problem of small samples. *Mol. Biol. Evol.* 6:539-45.

Rousset, F. 2008. Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molec. Ecol. Resources* 8:103-106.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. *Molecular Cloning: A Laboratory Manual. Second Edition*. Cold Spring Harbor Laboratory, New York.

Sokal, R.R., and Rohlf, F.J. 1995. *Biometry: the Principles of Statistics in Biological Research*. Freeman and Company, New York.

Valsecchi, E., and Amos, W. 1996. Microsatellite markers for the study of cetacean populations. *Mol. Ecol.* 5:151-156.

Wada, S. and Numachi, K. 1991. Allozyme analyses of genetic differentiation among the populations and

species of the *Balaenoptera*. *Rep. int. Whal. Commn* (special issue 13):125-54.
Weir, B.S., and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.

Table 1 Number of samples used in this studies by year on Antarctic fin whales collected during JARPA and JARPAII sueveys.

Areas	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	Total
IIIE					4							2		6
IV				4				19						23
V	3		6		6				6		1		2	24
VIW						2								2

Table 2. Variable sites defining 45 mtDNA haplotypes in the Antarctic fin whale. The column on the left are haplotype ID. The numbers above listed on the right side are the nucleotide positions of the polymorphic sites starting from the 5' end of the mtDNA control region. Haplotypes '2' through '45' are listed with reference to haplotype '1'. A dot indicates an identical nucleotide at the position relative to haplotype '1'. The left side of the table are the frequencies of the 45 haplotypes in the four Antarctic Areas.

Haplotype	Area IIIE	Area IV	Area V	Area VIW	10 11 1667788900 4082978748	20 1112222222 3790015555 4910144689	30 2222223333 6888890026 6347811245	34 90 80
1	1	0	0	0	GCAGTTTTTA	GCCCTTAACG	CTTTTATCAT	CA
2	1	1	0	0	.TG..C....	....C.....	......C...	..
3	1	0	0	0	.TG..C....	....C....A	......CT..	..
4	1	0	0	0	.TG.......	....C.....	........G.	..
5	1	0	0	0	..G.......	........T.	C	..
6	1	0	0	0	...A.C....	....	......C...	..
7	0	1	0	0	..G.......	.....CG...	.C....C...	..
8	0	2	0	0	.....C....	....C.....	........G.	..
9	0	1	0	0	.TG..C.CC.	....C.....	...C...T..	..
10	0	1	0	0	....	....	......C...	..
11	0	1	1	0	..G..C....	....	......C...	..
12	0	1	0	0	..G.......	.........A	.....G.T..	..
13	0	1	0	0	..G.....C.	....	......C...	..
14	0	2	2	0	.....C....	....C.....	....	..
15	0	1	0	0	..G..C....	....C.....	......C...	..
16	0	1	0	0	..G..C....	.TATC.....	.C..C.....	..
17	0	1	0	0	....	....C.G...	......CT..	..
18	0	1	0	0	ATGT.....G	....C.....	....	.G
19	0	1	0	0	.TG..C....	....C.....	..C...CT..	..
20	0	1	0	0	..G....C..	A...C.....	....	..
21	0	1	0	0	.TG.......	....	......C...	.G
22	0	1	0	0	.TG.......	....	......C...	..
23	0	1	0	0	..G.......	........T.	....	..
24	0	1	1	0	.TG.......	....	T.....C...	.G
25	0	1	0	0	.TG.......	....C.....	......CT..	.G
26	0	1	1	0	..G.......	...TC.....	....	.G
27	0	0	1	0	..G.......	....C..G..	....	.G
28	0	0	1	0	..G..C....	....C.G...	....	..
29	0	0	1	0	.TG.....C.	....C.....	......C...	.G
30	0	0	1	0	.TG.......	....	..CC..CT..	..
31	0	0	2	0	..G.......	....CCG...	......C...	..
32	0	0	1	0	..G.......	....	..CC......	..
33	0	0	1	0	.TG..C....	....C.....	......CT..	..
34	0	0	1	0	.TG..C....	....	......CT..	..
35	0	0	1	0	..G...C...	....C.....	..C...C...	..
36	0	0	1	0	..G..C....	.........A	......C...	..
37	0	0	1	0	..G.......	....	......C...	..
38	0	0	2	0	.TG..C....	....C.....	......CT..	.G
39	0	0	1	0	..G.......	....	.C........	T.
40	0	0	1	0	..G.......	....	.....G.T..	..
41	0	0	1	0	....	....	....	T.
42	0	0	1	0	.TG.......	....	..C...C..C	.G
43	0	0	1	0	.....C....	....C.....	....	T.
44	0	0	0	1	....C.....	....C.....	....	..
45	0	0	0	1	.TG......G	....C...T.	....	.G
	6	23	24	2	2223222222	2222222222	2222222222	22

Table 3. Estimates of the nucleon and nucleotide diversities in the Antarctic fin whale.

	H	π	SE
IIIE+IV	0.99261	0.01076	0.00086
V+VIW	0.99077	0.01052	0.00049

Table 4. Results of the heterogeneity tests using randomized chi-square, *Fst*, *Hst* and *Kst** in pair-wise comparison.

	χ^2	*Fst*	*Hst*	*Kst**
P	0.413	0.524	0.574	0.935
value		−0.001	−0.0005	−0.0058

Table 5. The number of alleles per locus, expected heterozygosity, test results for expected Hardy-Weinberg genotypic proportions (p-values) and inbreeding index (F_{IS}) at 16 microsatellite loci by total and by different areas.

Genetic index	Total		IIIE		IV		V		VIW	
No. alleles per locus	14.4		5.5		9.9		13.1		2.8	
Heterozygosity	0.816		0.799		0.789		0.834		0.875	
Hardy-Weinberg / Fis										
EV37	0.854	0.012	1.000	-0.044	0.948	-0.048	0.043	0.088	-	
EV1	0.777	-0.032	1.000	0.000	0.845	-0.063	0.969	-0.010	1.000	0.000
GT310	0.650	-0.073	1.000	-0.081	0.721	-0.086	0.783	-0.055	1.000	-0.333
GAT28	0.993	0.005	1.000	-0.053	0.883	-0.076	0.622	0.068	-	
GT575	0.801	-0.066	1.000	0.172	0.897	-0.184	0.568	-0.006	-	
EV94	0.197	0.053	1.000	0.000	0.307	-0.002	0.068	0.115	-	
GT23	0.102	0.045	0.164	0.059	0.528	0.021	0.336	0.053	1.000	-0.333
GAT98	0.000	0.110	1.000	-0.032	0.244	-0.010	0.004	0.132	0.334	1.000
EV104	0.608	-0.043	0.619	0.000	0.134	-0.020	0.573	-0.045	-	
GAT417	0.041	0.004	0.622	0.086	0.445	-0.083	0.626	0.058	-	
GT211	0.486	0.066	0.350	0.314	0.320	0.029	0.435	0.058	-	
EV21	0.009	0.085	0.302	0.172	0.185	-0.063	0.016	0.184	-	
FB14	0.032	0.186	1.000	-0.391	1.000	0.018	0.059	0.273	0.334	1.000
EV14	0.334	0.002	1.000	-0.333	0.763	0.006	0.046	0.049	-	
GT195	0.030	0.178	1.000	-0.177	0.051	0.255	0.053	0.173	-	
TAA31	0.000	0.059	1.000	-0.143	0.809	-0.104	0.000	0.138	1.000	-0.333
All	High. Sign.		1.000		0.753		High. Sign.		0.975	

Table 6. Results of the heterogeneity test comparing allele frequencies of 16 microsatellite loci of fin whales, among Antarctic management Areas.

	P value
IIIE vs IV	0.6776
IIIE vs V	0.9997
IV vs V	0.0136

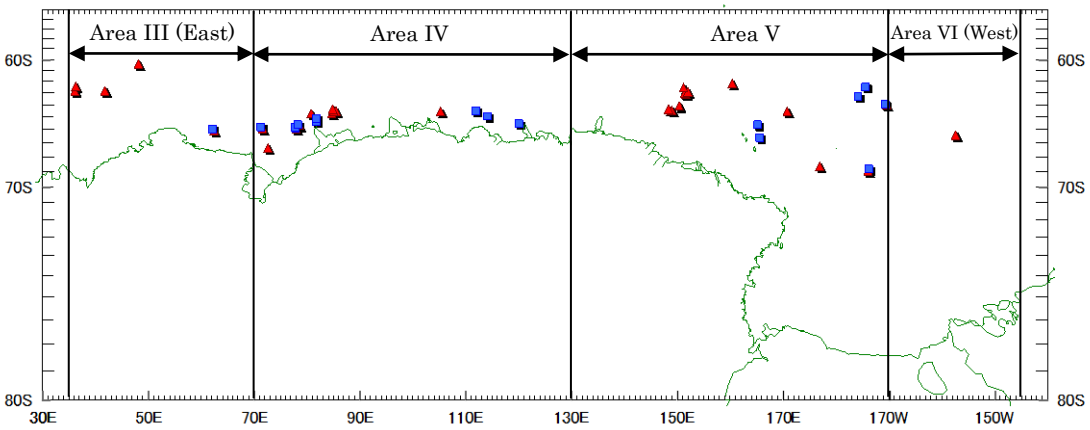

Figure 1. Geographic sampling localities of fin whales by catching (■) and biopsy (▲) during JARPA and JARPAII.