SC/F16/JR50

A note on the genetic diversity and phylogeny of western North Pacific and southern right whales based on mitochondrial and microsatellite DNA

Luis A. Pastene[1], Mutsuo Goto[1] Mioko Taguchi[2], Shigetoshi Nishiwaki[1] and Toshihide Kitakado[3]

[1] *Institute of Cetacean Research, 4-5 Toyomi-cho, Chuo-ku, Tokyo 104-0055, Japan*
[2]*Okitsunaka-cho 1347-3-302, Shimizu-ku, Shizuoka-shi, Shizuoka 424-0204, Japan*
[3] *Tokyo University of Marine Science and Technology, 1-5-7 Konan, Minato-ku, Tokyo 108-0075, Japan*

ABSTRACT
Genetic variation at 14 microsatellite DNA loci and mitochondrial DNA (mtDNA) control region sequences (275bp) was examined in right whales from the western North Pacific and Antarctic Area IV. Genetic analyses were based on biopsy samples collected during the surveys of the JARPNII in 2011 and 2012 (n=15), and JARPAII in 1993/94-2009/10 (n=67). The overall heterozygosity was 0.630 and 0.650 for North Pacific and southern right whales, respectively, while the nucleotide diversity/haplotype diversity were 0.0222/0.9048 and 0.0234/0.7743, respectively. Statistical tests found no evidence of deviation from the expected Hardy-Weinberg genotypic proportion in each of the oceanic basins. The Kimura's two parameter net interpopulational distance was 0.0358 (mtDNA) while the Nei's genetic distance (Da) was 0.7582 (microsatellite DNA), between North Pacific and southern right whale. A phylogenetic tree separated clearly mtDNA haplotypes of the North Pacific, North Atlantic and southern right whales.

INTRODUCTION

Rosenbaum *et al*. (2000) provided mtDNA phylogenetic evidences to separate North Atlantic, North Pacific and Southern Hemisphere right whales into three distinct lineages. Consequently, whales from these three oceanic basins have been considered different species *Eubalaena glacialis*, *E. japonica* and *E. australis*, respectively.

There are several genetic studies on southern right whales, which indicate that this species is divided into several genetic stocks: Argentina, South Africa, Western Australia, and the New Zealand sub-Antarctic (Baker, *et al*., 1999; Patenaude *et al*., 2007; Carrol *et al*., 2011). Apart from the study by Carrol *et al*. (2011) that used both mtDNA and microsatellite DNA, these studies have been based largely on mtDNA.

None of the genetic studies on southern right whales have examined genetic samples from whales in the Antarctic. Kanda *et al*. (2014a) examined the population genetic structure of southern right whales from Antarctic Area IV based on mtDNA and microsatellite DNA at 14 loci, and suggested that some whales could migrate long distance between breeding grounds and this feeding ground in the Antarctic.

In contrast, little is known about the population genetic structure of North Pacific right whales. A number of biopsy samples were obtained by JARPNII surveys in the western North Pacific, and these were examined genetically in this study to investigate level of genetic diversity in this endangered species. This study also compares the levels of genetic diversity between North Pacific and southern right whales examined by Kanda *et al*. (2014a). Finally this study expands the phylogenetic analysis of right whales worldwide using all available mtDNA sequences.

MATERIALS AND METHODS

Samples

Biopsy samples were obtained during the dedicated sighting surveys under JARPNII in the western North Pacific in 2011 and 2012, and JARPAII in Antarctic Area IV between 1993/94 and 2009/10 austral

summer seasons. A total of 15 and 67 biopsy samples were available for North Pacific and southern right whales, respectively. Figure 1 shows the distribution of the samples.

DNA extraction
The IWC guidelines for DNA data quality (IWC, 2009) was followed as much as possible (see Kanda *et al*., 2014b). Total DNA was extracted from 0.05g skin biopsy 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).

Microsatellite analysis
Genetic variation at microsatellite DNA was analyzed using 14 loci, none of which was designed specifically from right whales: EV1, EV14, EV21, EV37, EV94 (Valsecchi and Amos, 1996), GT23, GT211, GT310 (Bérubé *et al*., 2000), GATA28 (Palsbøll *et al*., 1997), DlrFCB17 (Buchanan *et al*., 1996), TR2F3, TR3G2, TR2G5 and TR3F2 (Frasier *et al*., 2006). Primer sequences and PCR cycling profiles generally followed those of the original authors.

PCR amplifications were performed in 15μl 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 right whale's DNA of known size that were rerun on each gel.

Data analysis

Level of polymorphisms
The number of alleles per locus, expected heterozygosity per locus and inbreeding coefficient per locus were calculated using FSTAT 2.9.3 (Goudet, 1995). Statistical tests for the deviations from expected Hardy-Weinberg genotypic proportions were conducted using GENEPOP 4.0 (Rousset, 2008).

Genetic differentiation
Conventional hypothesis testing procedure was conducted using heterogeneity test in microsatellite allele frequencies among samples. Probability test (or Fisher's exact test) implemented in GENEPOP 4.0 (Rousset, 2008) was used to conduct the heterogeneity tests. Statistical significance was determined using the chi-square value obtained from summing the negative logarithm of *p*-values over the 14 microsatellite loci (Sokal and Rohlf, 1995). The False Discovery Rate (FDR) approach (Benjamini and Yekutieli, 2001) was used for adjustment of *p*-value in case of multiple comparisons. F_{ST} value was calculated using FSTAT 2.9.3 (Goudet, 1995).

Individual matching
Matching exercises between the individuals were conducted using a computer program CERVUS (Marshall *et al*., 1998).

Mitochondrial DNA
Sequencing analysis of the 275bp control region of mtDNA was conducted using the primers light-strand MT4 (Árnason *et al*., 1993) and heavy-strand P2 (5'-GAAGAGGGATCCCTGCCAAGCGG-*3'*; Hori *et al*., unpublished). PCR products were purified by MicroSpin S-400HR columns (Pharmacia Biotech). Cycle sequencing was performed with the same primers, using BigDye terminator cycle sequence Kit (Applied Biosystems, Inc). The cycle sequencing products were purified by AutoSeq G-50 spin Columns (Pharmacia Biotech). The labeled sequencing fragments were resolved by electrophoresis through a 5% denaturing polyacrylamide matrix on an ABI 377™ or ABI3100 Automated DNA Sequencer (Applied Biosystems, Inc), following the protocols of the manufacturer. For each sample both strands were sequenced.

Data analysis

Level of polymorphisms
The number of haplotypes and haplotype diversity were calculated following Nei (1987). The nucleotide diversity (Nei, 1987: equation 10.5) and its standard error for population sampling and stochastic processes were calculated from the pair-wise differences between the mtDNA sequences using the Kimura's 2- parameter adjustment (Kimura, 1980).

Phylogenetic analysis
The genealogy of the mtDNA haplotypes was estimated using the Neighbor-Joining method (Saitou and Nei 1987) as implemented in the program PHYLIP (Felsenstein 1993). Genetic distances among haplotypes were estimated using the program DNADIST of PHYLIP, based on Kimura's 2-parameter model (Kimura 1980). A transition-transversion ratio of 5:1was used. The genealogy was rooted using the homologous sequence from bowhead whale. To estimate support for each node a total of 1,000 bootstrap simulations were conducted and the majority-rule consensus genealogy estimated.

RESULTS AND DISCUSION

Microsatellites

Level of polymorphisms
All 14 loci analyzed were polymorphic. For total samples (North Pacific and Antarctic) the total number of alleles per locus ranged from two at the TR2F3 to 19 at the EV1 with an average of 8.36 (Table 1). Expected heterozygosity at each of the loci ranged from 0.22 at EV21 to 0.92 at DlrFCB17 with an average of 0.68. As expected, significant deviation from the expected Hardy-Weinberg genotypic proportions was found. Consistently the magnitude of F_{IS} was large in several loci as expected if the samples were collected from different stock.

For North Pacific right whales, the total number of alleles per locus ranged from two at the EV21 to 7 at the GATA28 with an average of just 4.15 (Table 2). Expected heterozygosity at each of the loci ranged from 0.30 at TR3F2 and EV94 to 0.81 at EV37 with an average of 0.60. No significant deviation from the expected Hardy-Weinberg genotypic proportions was observed although the magnitude of F_{IS} was large in several loci.

For southern right whales, the total number of alleles per locus ranged from two at the TR2F3 and TR2G5 to 14 at the EV1 with an average of 6.93 (Table 3). Expected heterozygosity at each of the loci ranged from 0.09 at EV21 to 0.89 at DlrFCB with an average of 0.65. No significant deviation from the expected Hardy-Weinberg genotypic proportions was observed although the magnitude of F_{IS} was large in some loci.

Expected heterozygosities estimated for North Pacific and southern right whales were similar although the average number of alleles per locus was smaller in the North Pacific. Values of heterozygosity and average number of alleles were smaller than those estimated by Carrol *et al.* (2011) for whales in several localities of the Southern Hemisphere examined with 13 loci. The ranges estimated by those authors were 0.80-0.82 and 6.31-12.15, respectively. These differences could be related to sample size (substantially larger in that study), localities (several localities in that study) and loci used (different sets between the two studies).

Genetic differentiation
As expected F_{ST} value between North Pacific and southern right whales was high (0.176 over all the loci) and highly significant (Table 4). Nei's genetic distance (Da) was estimated at 0.7582.

Movement (mark-recapture)
There was no case of duplicate sampling in the western North Pacific. In the case of the Antarctic there were two cases of re-sampling involving different years. These duplicates were confirmed by genotype profiles as well by photo-identification (see details in Kanda *et al*., 2014a). These finding in the Antarctic suggest fidelity to feeding areas in Area IV.

mtDNA

Levels of polymorphisms

Nucleotide diversity was similar between North Pacific and southern right whales, and these were relatively high and similar to estimates for several regions of the Southern Hemisphere. Haplotype diversity was somewhat higher in the North Pacific (Table 5). The values of haplotype and nucleotide diversity estimated for several localities of the Southern Hemisphere ranged from 0.595-1.000 and 0.014-0.033, respectively (Patenaude *et al*., 2007). The estimates by Carrol *et al*. (2011) were 0.67-0.95 and 0.017-0.028, respectively.

Genetic differentiation

The Kimura's two parameter net interpopulational distance between North Pacific and southern right whales was estimated at 0.036. This value is much lower than those among recognized large whale species and lower than common minke whales from different ocean basins, which are currently recognized as sub-species.

Phylogenetic analysis

The phylogenetic analysis of haplotypes was based on an enlarged samples size with regard the phylogenetic analysis conducted by Rosenbaum *et al*. (2000). Haplotypes from North Pacific, North Atlantic and southern right whales were placed in different clusters. Within the southern right whale cluster, two sub-clusters were observed, which were supported by high bootstrap value. However these sub-clusters were not geographically specific. Results confirmed that North Pacific, North Atlantic and Southern Hemisphere right whales are different Evolutionary Significant Units.

ACKNOWLEDGEMENTS

We gratefully acknowledge the researchers and crew members that participated in JARPN/JARPNII and JARPAII surveys for the collection of biopsy samples. We thank H. Oikawa and S. Azumi (ICR) who collaborated in the process of DNA extraction. We also thank N. Kanda (JANUS-Japan) for preparing the microsatellite data set used in this study.

REFERENCES

Árnason, Ú., Gullberg, A. and Widegten, B. 1993. Cetacean mitochondrial DNA control region: Sequences of all extant baleen whales and two sperm whale species. *Mol. Biol. Evol.*, **10**, 960-970.

Baker, C.S., Patenaude, N.J., Bannister, J.L., Robins, J. and Kato, H. 1999. Distribution and diversity of mtDNA lineages among Southern right whales (*Eubalaena australis*) from Australia and New Zealand. *Marine Biology* 134: 1-7.

Benjamini, Y. and Yekutieli, D. 2001. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*. 29, No. 4, 1165–1188

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.

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.

Carroll, E.L., Patenaude, N.J., Alexander, A.M., Steel, D., Harcourt, R., Childerhouse, S., Smith, S., Bannister, J.L., Constantine, R. and Baker, C.S. 2011. Population structure and individual movement of southern right whales around New Zealand and Australia. *Mar. Ecol. Prog. Ser*. 432: 257-268.

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) 3.5c. University of Washington, Seattle, USA.

Frasier, T.R, Rastogi, T., Brown, M.W., Hamilton, P.K., Kraus, S.D., and White, B.N. 2006. Characterization of tetranucleotide microsatellite loci and development and validation of multiplex reactions for the study of right whale species (genus *Eubalaena*). *Mol Ecol. Note* 6: 1025-1029.

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

International Whaling Commission. 2009. Report of the Scientific Committee. Annex I. Report of the working group on stock definition. Appendix 2. Guidelines for DNA data quality control for genetic studies relevant to IWC management advice. *J. Cetacean Res. Manage. (Suppl.)* 11:252-256.

Kanda, N., Goto, M., Nishiwaki, S. and Pastene, L.A. 2014a. Long-distance longitudinal migration of southern right whales suspected from mtDNA and microsatellite DNA analysis on JARPA and JARPAII biopsy samples. Paper SC/F14/J33rev presented to the JARPA II Review Workshop, Tokyo, February 2014 (unpublished). 6pp.

Kanda, N., Goto, M., Oikawa, H. and Pastene, L.A. 2014b. Update of note on sampling and laboratory procedure protocols of the genetic work at the Institute of Cetacean Research (SC/F14/J27rev). Paper SC/65b/DNA1 presented to the IWC Scientific Committee, May 2014 (unpublished). 6pp.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16: 111-120.

Marshall, T.C., Slate, J., Kruuk, L., and Pemberton, J.M. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* 7: 639-655.

Nei, M. 1987. *Molecular Evolutionary Genetics.* Columbia University Press, New York.

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

Patenaude, N.J., Portway, V.A., Schaeff, C.M., Bannister, J.L., Beat, P.B., Payne, R.S., Rowntree, V.J., Rivarola, M. and Baker, C.S. 2007. Mitochondrial DNA Diversity and Population Structure among Southern Right Whales (*Eubalaena australis*). *J. Hered.* 98: 147-157

Rosenbaum, H.C., Brownell JR,R.L., Brown,M.W., Schaeff,C., Portway,V., White,B.N., Malik,S., Pastene,L.A., Pateneude,N.J., Baker,C.S., Goto,M., Best,P.B., Clapham,P.J., Hamilton,P., Moore,M., Payne,R., Rowntree,V., Tynan,C.T., Bannister,J.L., DeSalle,R. 2000: World-wide genetic differentiation of *Eubalaena*: questioning the number of right whale species. *Molecular Ecology* 9:1793-1802.

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

Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4(4):406-25.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. *Molecular cloning: A laboratory manual. 2nd Ed.*, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

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.

Table 1. The number of alleles (A), expected heterozygosity (H_E), test results for expected Hardy-Weinberg genotypic proportion (HW) and inbreeding coefficient (F_{IS}) in right whales from the North Pacific and Antarctic, at 14 microsatellite loci.

Microsatellite loci	A	F_{IS}	H_E	HW
EV1	19	0,09	0,91	**<0.001**
GT310	5	0,05	0,68	**0,003**
GT23	9	0,04	0,80	0,598
EV94	6	-0,01	0,41	0,975
EV14	10	0,03	0,78	0,425
GT211	11	0,19	0,85	**<0.001**
EV37	10	0,07	0,84	0,082
GATA28	10	0,04	0,80	0,535
EV21	3	0,21	0,22	0,179
DlrFCB17	18	0,10	0,92	**<0.001**
TR2F3	2	0,03	0,49	1,000
TR3G2	7	-0,01	0,80	0,175
TR2G5	3	-0,04	0,51	0,649
TR3F2	4	0,23	0,53	0,065
Overall	8,36	0,07	0,68	**High. Sign.**

Table 2. The number of alleles (A), expected heterozygosity (H_E), test results for expected Hardy-Weinberg genotypic proportion (HW) and inbreeding coefficient (F_{IS}) in North Pacific right whale, at 13 microsatellite loci.

Microsatellite loci	A	F_{IS}	H_E	HW
EV1	5	0,14	0,78	0,695
GT310	3	0,18	0,65	0,018
GT23	5	0,00	0,73	0,760
EV94	3	-0,11	0,30	1,000
EV14	4	0,06	0,57	0,253
GT211	3	-0,03	0,45	1,000
EV37	5	0,01	0,81	0,070
GATA28	7	-0,01	0,79	0,488
EV21	2	-0,31	0,51	0,328
DlrFCB	6	0,18	0,74	0,128
TR3G2	5	-0,15	0,70	0,882
TR2G5	3	0,21	0,51	0,356
TR3F2	3	-0,11	0,30	1,000
Overall	4,15	0,02	0,60	0,384

Table 3. The number of alleles (A), expected heterozygosity (H_E), test results for expected Hardy-Weinberg genotypic proportion (HW) and inbreeding coefficient (F_{IS}) in Antarctic right whale, at 14 microsatellite loci.

Microsatellite loci	A	F_{IS}	H_E	HW
EV1	14	0,02	0,88	0,390
GT310	5	-0,04	0,65	0,094
GT23	8	0,03	0,80	0,797
EV94	5	-0,01	0,43	1,000
EV14	9	-0,08	0,75	0,934
GT211	10	0,10	0,82	0,145
EV37	10	0,04	0,81	0,691
GATA28	8	0,02	0,78	0,661
EV21	3	0,31	0,09	0,115
DlrFCB	12	0,01	0,89	0,438
TR2F3	2	0,03	0,49	1,000
TR3G2	6	-0,05	0,77	0,545
TR2G5	2	-0,10	0,50	0,468
TR3F2	3	0,12	0,47	0,620
Overall	6,93	0,01	0,65	0,768

Table 4. Results of the microsatellite DNA heterogeneity test between North Pacific and Antarctic right whales. Bold indicate significant differences after FDR correction.

Microsatellite loci	*P* -value	F_{ST}
EV1	**<0.001**	0,166
GT310	**<0.001**	0,134
GT23	**0,001**	0,056
EV94	**0,006**	0,047
EV14	**<0.001**	0,216
GT211	**<0.001**	0,294
EV37	**<0.001**	0,131
GATA28	**<0.001**	0,066
EV21	**<0.001**	0,522
DlrFCB17	**<0.001**	0,179
TR3G2	**<0.001**	0,155
TR2G5	0,058	0,018
TR3F2	**<0.001**	0,379
Overall	**High. Sign.**	0,176

Table 5. Results of mtDNA diversity for North Pacific and Antarctic right whale.

Location	Number of haplotypes	Nucleotide diversity (SE)	Haplotype diversity
North Pacific (n=15)	8	0.02223 (0.00279)	0.90476
Antarctic (n=67)	10	0.02343 (0.00093)	0.77431

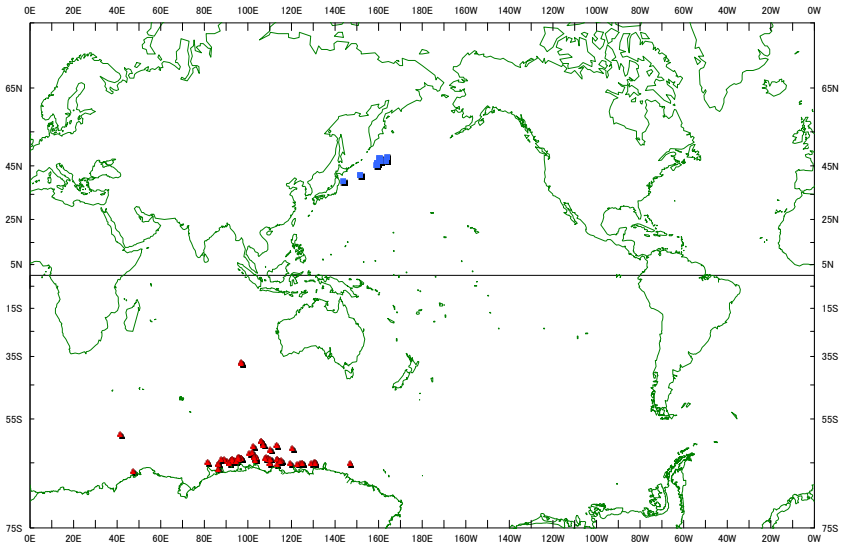

Figure 1. Geographical distribution of right whales examined in this genetic study. Blue: biopsy samples from JARPNII; Red: biopsy samples from JARPAII.

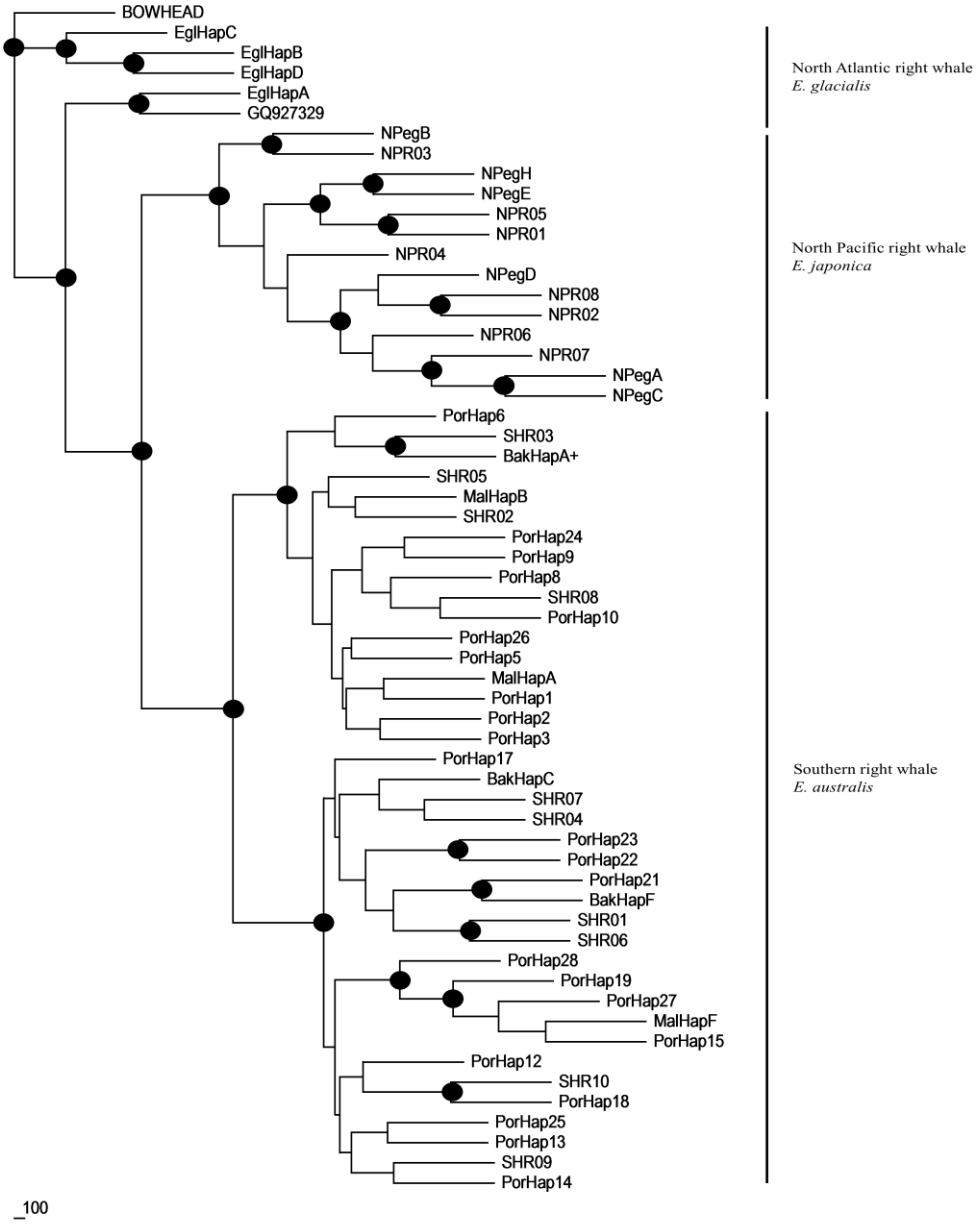

Figure 2. Phylogenetic relationship among right whales mtDNA haplotypes from several ocean basins.