SC/60/PFI11

This working paper is not to be cited without prior consent with the authors

Lack of genetic divergence among samples of the North Atlantic Fin Whale collected at feeding grounds: congruence among microsatellite loci and mtDNA in the new Icelandic dataset.

C. PAMPOULIE[1], A.K. DANÍELSDÓTTIR[1], M. BÉRUBÉ[3], P.J. PALSBØLL[3], A. ÁRNASON[1], Th. GUNNLAUGSSON[1], D. ÓLAFSDÓTTIR[1], N. ØIEN[4], L. WITTING[5], and G.A. VÍKINGSSON[1].

[1] *Marine Research Institute, Skúlagata 4, IS-101 Reykjavík, Iceland*

[2] *MATIS ohf., Food Research, Innovation and Safety, Skúlagata 4, IS-101 Reykjavík, Iceland*

[3] *University of California Berkeley, Ecosystem Sciences Division –ESPM, Hilgard Hall #3110, Berkeley, CA 94720-3110, USA*

[4] *Institute of Marine Research, P.O. Box 1870 Nordnes, N-5817 Bergen, Norway*

[5] *Greenland Institute of Natural Resources, c/o Box 570, Kivioq 2, 3900 Nuuk, Greenland*

ABSTRACT

The main objective of this study was to confirm the level of genetic differentiation observed among samples of fin whale collected at feeding grounds using a new dataset composed of 500 selected individuals genotyped for 15 microsatellite loci and the mtDNA was sequenced. New methods of amplification have been used (especially for the microsatellite loci) to successfully genotyped these selected individuals. Both types of genetic markers employed revealed a lack of genetic differentiation among samples collected and therefore suggest that extensive gene flow might prevail among fin whale population and/or that a recent population expansion of the species occurred. The signal of population expansion detected with the mtDNA might suggest that population of fin whale are not at equilibrium and that further analyses based on this information should be carried out (calculation of potential time of expansion and divergence). However, the comparison of these results to the allozyme data that have been recently submitted revealed that the observed F_{ST} values at the allozyme loci are likely to be due to diversifying selection at different breeding sites that accentuates differences not readily apparent at neutral loci. At neutral markers such as microsatellite loci, in the absence of gene flow, genetic differentiation among populations due to drift and/or mutation will rise given enough time, while selection on non-neutral markers (allozyme) in the absence of gene flow will accentuate differences among populations in a shorter time period, depending on the strength of the selective processes.

KEYWORDS: BALAENOPTERA PHYSALUS, NORTH ATLANTIC OCEAN, GENETICS, MICROSATELLITE LOCI, MTDNA.

INTRODUCTION

The fin whale *Balaenoptera physalus* is the second largest living marine mammal and it has consequently been the target of commercial and aboriginal hunting which severely depleted its abundance across most of their distribution range. From 1868 to 1985, some 79,000 fin whales were killed for commercial purposes in the North Atlantic Ocean (Sigurjónsson, 1995). Systematic sightings surveys conducted regularly since 1987 (Sigurjónsson *et al.*, 1991; Víkingsson *et al.*, in press) have shown a substantial increase in abundance of the species in the central North Atlantic. According to the most recent estimate agreed by the Scientific Committies of the IWC and NAMMCO there are around 24,000 fin whales in the Central North Atlantic (EGI stock see below) (Víkingsson *et al.*, in press). In recent scientific assessments of the status of North Atlantic fin whales, lack of knowledge on stock structure has been identified as the main source of uncertainty (IWC, 2007; NAMMCO, 2006). This is merely due to the fact that the numbers and location of breeding areas have not yet been established and that it is not known whether individuals located at different feeding areas represent different populations.

The North Atlantic fin whales are believed to undertake long, annual migrations between high-latitude summer feeding areas and low-latitude winter breeding areas like most of the baleen whales. Little is known about the genetic composition and biological characteristics of the group of individuals located at feeding grounds. So far, the geographical delineation of fin whale stocks in the North Atlantic Ocean is not well established, although it is believed that several distinct stocks of the species inhabit the North Atlantic Ocean. Rørvik and Jonsgård (1981) divided fin whales into six independent groups, namely: (1) off North Norway and in the Arctic Eastern North Atlantic; (2) off East Greenland and West Iceland; (3) off West Norway and the Faroe Islands; (4) off The Scottish islands, Spain and Portugal; (5) off West Greenland and (6) off Nova Scotia, Newfoundland and Labrador. The International Whaling Commission (IWC) divides the North Atlantic fin whales into seven management stocks, splitting above group No. (6) into two (Donovan, 1991). This division was originally based on various methods including catch data, occurrence and length distributions. Later on, morphological studies (Jover, 1987, 1991), mark-recapture data (Gunnlaugsson and Sigurjónsson, 1989; Sigurjónsson *et al.*, 1991) and analysis of genetic markers (Árnason and Jónsdóttir, 1988; Daníelsdóttir *et al.*, 1991) supported the discrimination of these different areas.

Population genetic studies have been carried out on the North Atlantic fin whale stocks (Daníelsdóttir *et al.*, 1992; Daníelsdóttir, 1994; Bérubé *et al.*, 1998; Palsbøll *et al.*, 2004; Daníelsdóttir *et al.*, 2005, 2006). Both genetic studies and studies on morphological characters have revealed significant divergence among some fin whales stocks in the North Atlantic which are also supported by tagging experiments and other non-genetic evidence (Árnason and Jónsdóttir, 1988; Árnason *et al.*, 1992; Bérubé *et al.*, 1998; Daníelsdóttir *et al.*, 1991, 1992; Daníelsdóttir, 1994; Jover, 1987, 1991; Sigurjónsson and Gunnlaugsson, 1985; Gunnlaugsson and Sigurjónsson, 1989; Sigurjónsson *et al.*, 1991; Víkingsson and Gunnlaugsson, 2006). However, Árnason (1981) suggested that the occurrence of genetically distinct fin whale stocks is unlikely within a limited area and that no solid evidence has been reported in favor of the existence of several fin whale stocks in the Northeastern Atlantic. In a previous study of fin whale from the Icelandic feeding ground, Árnason and Sigurðsson (1983) nevertheless showed, using red blood cell carbonic anhydrase-like esterases, that the fin whale off Iceland could be considered as a discrete stock.

The main objective of this new study was to gather a better genetic dataset to assess the potential genetic structure of fin whale at feeding grounds in the North Atlantic Ocean. The plan was to select individuals from the previously gathered database and to genotyped them

for 15 microsatellite loci and the mtDNA. A total of 475 individuals have been genotyped for more than 13 microsatellite loci while 466 have been successfully amplified for the mtDNA. Here we present the results of genetic analyses performed on these samples collected at five different geographical locations.

MATERIALS AND METHODS

The dataset presented here is on conventional genetic analysis of a new dataset (n=475, see Table 1) of 15 microsatellite loci (EV001, EV037, GATA028, GATA053, GATA098, GATA417, GT011, GT023, GT195, GT211, GT271, GT310, GT575, TAA023 and GGAA520; see Table 2) and the D-loop of the mtDNA of 466 fin whales sampled in five North Atlantic locations; i.e. off West Iceland, Canada, Norway, Spain and West Greenland. The 13 Canadian samples originate from the years 1992 and 1993 from the Newfoundland Grand Banks provided by Dwaine Pittman and Jon Lien, Memorial University of Newfoundland, Canada. The Icelandic samples, a total of 365 animals caught west and southwest off Iceland came from the 1983-2006 commercial catch and the catch under scientific permit. The Icelandic samples were divided into early, mid and late seasons within each year. Samples from the period 1st to 30th of June were grouped into early samples, 1st to 31st of July were grouped into mid and 1st of August to 31st of September were grouped into late. Biopsy of 15 fin whales were taken from free swimming animals using biopsy technique off North Norway in the summer of 1991 by the Marine Research Institute, Reykjavík, Iceland and 39 biopsy samples from Norwegian waters (off the Norwegian coast up to the Spitsbergen area) were collected by the Institute of Marine Research, Bergen, Norway in the years 1999-2003. The 43 Spanish samples were collected at the Spanish whaling station in 1985 from whales taken northwest of the Iberian Peninsula. The 15 samples from West Greenland waters come from the years 2000 and 2003-2005 (Table 1).

Sequencing analyzes for microsatellite loci.

15 microsatellite markers and 1 sex marker (ZFYX0582) were used in 4 multiplex PCR reaction. Multiplex 1 contained GATA098, EV1, ZFYX0582, GT310, EV37 and GT023. Annealing temperature was 54°C and 32 PCR cycles were run. Multiplex 2 contained GATA417, GATA028, GT211 and GT575. Annealing temperature was 56°C and 35 PCR cycles were run. Multiplex 3 contained GT195, GATA053 and GT011. Annealing temperature was 58°C and 35 PCR cycles were run. Multiplex 4 contained TAA023 and GGAA520. Annealing temperature was 54°C and 32 PCR cycles were run. Marker GT271 was run in a separate PCR reaction. Annealing temperature was 64°C and 35 PCR cycles were run. The PCR product was added to multiplex 4 before loading it to the machine.

PCR reactions were performed in a total reaction volume of 10 µl. Each reaction consisted of 2 µl of DNA template isolated with 15% Chelex 100 Resin (BioRad, cat.143-2832) (Walsh *et al*., 1991) and ProteinaseK, 0.2 µl Teg DNA polymerase (3U/µl, comparable with Taq DNA polymerase), 1.0 µl of 10x buffer, 0.8 µl dNTP (10 mM), 0,07-0,40 µl reverse- and forward primers (100 µM) and dH_2O up to 10 µl total reaction volume. The amplification cycle consisted of a 4 min denaturation at 94°C followed by 32-35 cycles of 94°C denaturing for 50 s, 54-64°C annealing for 50 s and 72°C extension for 90 s. Cycling was concluded with a 7 min extension at 72°C. Thermal cycling was performed in a Peltier Thermal Cycler 225 (MJ Research). Thermal cycler was configured with a heated lid.

Amplified DNA fragments were separated by an ABI 3730 DNA Analyser (Applied Biosystems) and were sized according to the Applied Biosystem GeneScan™ - 500LIZ™ size standard. Alleles were scored manually with the GeneMapper™ Analysis Software version 4.0 (Applied Biosystem).

Table 1. Sampling location, year of sampling, number of individuals scored (n), observed heterozygosity (H_O), expected heterozygosity (H_E), mean number of alleles (MnA) and allelic richness loci (A_R based on 9 individuals) at 15 microsatellite from 475 samples of fin whales from the North Atlantic. Bold values: F_{IS} values deviating from HWE after corrections for multiple tests (Rice, 1989).

Location	Year	Seasons	15 microsatellite loci data				
			n	H_O	H_E	MnA	A_R
Iceland	1983	Early	28	0.766	0.782	8.80	6.11
		Mid	78	0.804	0.798	10.13	6.19
		Late	18	0.760	0.768	7.73	6.17
Iceland	1985	Early	78	0.799	0.795	10.27	6.19
		Mid	70	0.791	0.795	9.47	6.16
		Late	10	0.827	0.767	6.53	6.09
Iceland	1987	Mid	9	0.726	0.749	6.47	6.17
Iceland	1989	Early	28	0.769	0.785	8.67	6.06
		Mid	39	0.793	0.782	9.07	6.07
Iceland	2006	Late	7	0.791	0.735	5.87	NA
Norway	1991,1999-2003	1991,1999-2003	39	**0.786**	0.795	9.20	6.31
Spain	1985	1985	43	0.744	0.793	9.07	6.05
West Greenland	2000, 2003-2005	2000, 2003-2005	15	0.806	0.779	8.20	6.43
Newfoundland Grand Banks Canada	1992-1993	1992-1993	13	**0.718**	0.731	6.67	5.72

Sequencing analyzes for mitochondrial DNA control region (mtDNA).

Primers MT4-PCR-F (5' CCTCCCTAAGACTCAAGGAAG 3) and MW-PCR-r 5' (GGTCCTGAAGTAAGAACCAGATG 3´) were used to amplify mtDNA fragment. PCR was performed in a total reaction volume of 20 µl. Each reaction consisted of 2 µl of DNA template isolated with 15% Chelex 100 Resin (BioRad, cat.143-2832) (Walsh *et al*., 1991) and ProteinaseK, 0.4 µl Teg DNA polymerase (3U/µl, comparable with Taq DNA polymerase), 2.0 µl of 10x buffer, 0.4 µl dNTP (10 mM), 0.05 µl of each reverse-and forward primers (100 µM) and 15.1 µl dH_2O.

The amplification cycle consisted of a 3 min denaturation at 94°C followed by 35 cycles of 94 °C denaturing for 50 s, 56°C annealing for 50 s and 72°C extension for 90 s. Cycling was concluded with a 7 min extension at 72°C. Thermal cycling was performed in a Peltier Thermal Cycler 225 (MJ Research). Thermal cycler was configured with a heated lid. PCR fragment was purified by using ExoSAP-IT method according to producer (cat. 78201, usb). ExoSap method was performed in a Peltier Thermal Cycler 225 (MJ Research), 37°C for 25 min and then 80°C fo 15 min. Purified PCR product was sequenced with forward or reverse using primers MN312-seq-r (5´-GATCTAATGGAGCGGCCATAAG-´3) and BP15851-seq-F (5´-CATCACACTCCACCATCAG-´3). PCR was performed in a total reaction volume of 10 µl. Each reaction consisted of 5 µl purified PCR product, 0.5 µl Big Dye, 1.5 µl of 5x buffer, 1 µl of either reverse-or forward primers (3.5 µM) and 2.0 µl dH_2O. The amplification cycle consisted of 35 cycles of 96°C denaturing for 30 s, 50°C annealing for 15 s and 60°C extension for 4 min. Thermal cycling was performed in a Peltier Thermal

Cycler 225 (MJ Research). Thermal cycler was configured with a heated lid. For Big Dye terminator removal of the cycle sequencing reaction CleanSeq® from Agencourt, Bioscience Corporation (cat no: 000136) was used according to protocol 000383v027 from producer. Sequencing was performed in ABI 3730 DNA Analyser (Applied Biosystems).
During the laboratory process, careful attention was given to the labelling of individuals to avoid confusion and duplicates. Later on, possible duplicates were manually checked within the genotype matrix and kinship and relatedness analyses.

Table 2. Characteristics of the 15 microsatellite DNA markers scored for 475 fin whale individuals.

Locus	Primer sequence (5'→3')	Reference	Size range	No. of alleles
EV01	CCC TGC TCC CCA TTC TC ATA AAC TCT AAT ACA CTT CCT CCA AC	Valsecchi and Amos (1996)	143-179	17
EV037	AGC TTG ATT TGG AAG TCA TGA TAG TZG AGC CGT GAT AAA GTG C	Valsecchi and Amos (1996)	179-215	17
GATA028	AAA GAC TGA GAT CTA TAG TTA CGC TGA TAG ATT AGT CTA GG	Palsboll *et al.* (1997)	191-243	15
GATA053	ATT GGC AGT GGC AGG AGA CCC GAC ACA GAG ATG TAG AAG GAG	Palsbøll *et al.* (1997)	242-278	11
GATA098	TGT ACC CTG GAT GGA TAG ATT TCA CCT TAT TTT GTC TGT CTG	Palsbøll *et al.* (1997)	88-120	9
GATA417	CTG AGA TAG CAG TTA CAT GGG TCT GCT CAG GAA ATT TTC AAG	Palsbøll *et al.* (1997)	209-297	24
GT011	CAT TTT GGG TTG GAT CAT TC GTG GAG ACC AGG GAT ATT GC	Bérubé *et al.* (1998)	117-133	8
GT023	CAT TTC CTA CCC ACC TGT CAT GTT CCC AGG CTC TGCACT CTG	Bérubé *et al.* (2000)	109-143	13
GT195	TGA GAA AGA TGA CTA TGA CTC TGA AGT AAC AGT TAA TAT ACC	Bérubé *et al.* (2000)	161-181	11
GT211	CAT CTG TGC TTC CAC AAG CCC GGC ACA AGT CAG TAA GGT AGG	Bérubé *et al.* (2000)	106-126	11
GT271	GCT CAC ACT GGT AAT CTG TGG CCC TAG GAA GGA TAG ACA TAG	Bérubé *et al.* (2000)	104-128	11
GT310	TAA CTT GTG GAA GAT GCC AAC GAA TAC TCC CAG TAG TTT CTC	Bérubé *et al.* (2000)	106-130	11
GT575	TAT AAG TGA ATA CAA AGA CCC ACC ATC AAC TGG AAG TCT TTC	Bérubé *et al.* (2000)	146-160	8
TAA023	CTC GCA CAG AAA TGA AGA CCC AGA GCC TGA ACC AGA ACA AGG	Palsbøll *et al.* (1997)	86-104	6
GGAA520	TAG CAG AYC TGA GTT ATT TCC TAG CAT TTT AGTCTT GGG TGG	Palsbøll *et al.* (1997)	199-235	11

Statistical analyses

Microsatellite loci.
Wright's single-locus *F*-statistics (Wright, 1969) were calculated from allele frequencies at all loci examined for each population according to Weir and Cockerham's method (1984) using FSTAT vers. 2.9.3.2 (Goudet, 1995), and significance of pairwise F_{ST} values was assessed by permuting alleles among samples. The Mean F_{ST} (Weir and Cockerham, 1984) over all loci and samples were calculated and the 95% CI of the pairwise estimates of F_{ST} were bootstrapped (1,000) over all loci in the same software. Significance levels were adjusted with Bonferroni correction for multiple tests (Rice, 1989). These tests were primarily samples-based (variation among location-year) and then geographically based. FSTAT was also used to test for fit to Hardy-Weinberg proportions (HWE); to test for linkage disequilibrium between loci, to compare allelic richness (A_R) and mean number of alleles (MnA). In all cases 15,000 permutations were used for significance testing. As a final means of assessing genetic

substructure, we calculated assignment index (AI) values for males and females in our study populations using FSTAT 2.9.3 (Goudet, 1995). Assignment indices determine the probability that an individual genotype should occur in the population from which it was sampled (Favre *et al*., 1997). When dispersal is frequent (i.e. the individuals sampled represent a mixture of natal and immigrant animals), the mean AI value for a population should be small and the variance around the mean should be large. Accordingly, in populations with sex-biased dispersal, mean AI should be smaller and variance should be larger for the sex in which dispersal is more common.

Mitochondrial DNA control region (mtDNA).

For each sample, gene (H) and nucleotide diversity (π) were calculated using Arlequin vers. 3.0 (Excoffier *et al*., 2005). Shared haplotypes among populations were identified using the same software. Genetic differentiation among populations was assessed taking into account nucleotide differences between haplotypes (Φ_{ST}, Weir and Cockerham, 1984), after correction by the Tamura-Nei formula. Significance was assessed by a permutation procedure (5,000 permutations). Corrected genetic distances accounting for intra-population variability (PiXY-(PiXY + PiY)/2) were also calculated using Arlequin. DnaSP 4.20 (Rozas *et al*., 2003) was used to perform tests of neutrality (Tajima's D) and to analyse mismatch distributions (Slatkin and Hudson, 1991) for the inference of demographic history from mtDNA data. These tests were primarily samples-based (variation among location-year) and then geographically based. Relationships among mitochondrial haplotypes were depicted by a median-joining tree in the software MEGA4 (Tamura *et al*., 2007) and in Network4 (Bandelt *et al*., 1999). We selected the Median Joining approach (MJ) implemented by the software Network4 (Bandelt *et al*., 1999), for being one of the most efficient network building methods available to date (Cassens *et al*., 2003). The homoplasy parameter (ε) was set to zero. Two weighting schemes were applied in order to account for differences in substitution rates: (1) equal weight for all classes of changes, and (2) weight of 10 for transitions and 30 for transversions and gaps, as suggested by the authors.

Estimates of the number of migrants per generation (*Nm*) were calculated in GENETIX (Belkhir *et al*., 2002) e.g. $N\text{m} = (1-F_{ST})/4*F_{ST}$ for microsatellite loci and $N\text{m} = (1-\Phi_{ST})/2*\Phi_{ST}$ for the mtDNA (Wright, 1969).

RESULTS

Microsatellite loci.

All microsatellite loci were highly polymorphic. The number of alleles per locus across samples ranged from 6 (TAA023) to 24 (GATA417). H_o averaged over loci and all samples ranged from 0.718 to 0.827, while H_e ranged from 0.731 to 0.800. Genotypic proportions were significantly out of HWE in three out of 210 exact tests (this is less than expected by chance alone). Genotypic frequencies in all but two samples (Norway and Canada) were in HWE. The genetic variation among and within the samples showed a non-significant F_{ST} of 0.0012 (95% CI: -0.0004-0.0028) and a non-significant F_{IS} of 0.0200 (95% CI: 0.0058-0.0347). After Bonferroni correction, pairwise differentiation between samples did not yield any significant values (Appendix 1). When pairwise differentiation were based on geographical locations, the comparison among Iceland and Canada was significant after Bonferroni correction (Appendix 3).

In Addition, F_{ST} as well as relatedness and variance values were larger for females than males, but in neither case were they significant (Table 3)

Table 3. Test results for sex-biased dispersal in the fin whale, based on all samples and 15 microsatellite loci.

	F_{IS}	F_{ST}	Rel.	Mean assignment values	Variance
Females	0.019	**0.0034**	**0.0066**	-0.027	**15.18**
Males	0.019	0.0009	0.0018	0.037	14.04
P-values	0.42	0.12	0.12	0.57	0.65

Mitochondrial DNA control region (mtDNA).
A total of 28 polymorphic sites were observed among which transitions were highly dominant (Table 4).

Table 4. Maximum composite likelihood estimate of the pattern of the rate of nucleotide substitution.

	A	T	C	G
A	-	*0.05*	*0.02*	**8.84**
T	*0.04*	-	**22.28**	*0.02*
C	*0.04*	**47.19**	-	*0.02*
G	**21.42**	*0.05*	*0.02*	-

NOTE: Each entry shows the probability of substitution from one base (row) to another base (column) instantaneously. Only entries within a row should be compared. Rates of different transitional substitutions are shown in bold and those of transversionsal substitutions are shown in italics. The nucleotide frequencies are 0.335 (A), 0.358 (T/U), 0.169 (C), and 0.138 (G). The transition/transversion rate ratios are k_1 = 484.029 (purines) and k_2 = 1000 (pyrimidines). **The overall transition/transversion bias is R = 241.139**, where $R = [A*G*k_1 + T*C*k_2]/[(A+G)*(T+C)]$. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 285 positions in the final dataset. All calculations were conducted in MEGA4 (Tamura *et al*., 2007).

The 28 sites defined a total of 63 different haplotypes where haplotype 1 is shared and common among all sampling areas. Estimates of nucleotide diversity ranged from 0.0097 for IC87 to 0.0150 for IC06, and the haplotypic diversity ranged from 0.8154 for SP to 0.9524 for IC06 (Table 5).

Table 5. Number of samples (N), gene diversity (H), number of haplotypes (nH), and nucleotide diversity (π) by population with standard deviation (s.d.).

	NOR	CAN	GREEN	SP85	IC83	IC85	IC87	IC89	IC06
N	40	14	15	40	123	155	7	66	7
nH	21	9	10	19	37	44	5	26	6
H	0.9321	0.9359	0.9238	0.8154	0.9288	0.9098	0.8571	0.8942	0.9524
s.d.	0.0261	0.0507	0.0530	0.0611	0.0144	0.0175	0.1371	0.0282	0.0960
π	0.0137	0.0140	0.0146	0.0112	0.0122	0.0141	0.0097	0.0131	0.0150
s.d.	0.0078	0.0084	0.0086	0.0066	0.0069	0.0078	0.0067	0.0074	0.0097

Before correction for multiple tests, pairwise differentiation based on nucleotide differences among haplotypes (Φ_{ST}) between populations yielded five significant values among which

three were due to the Canada sample (Table 6). After Bonferroni corrections for multiple tests, none of the pairwise comparisons were significant. Neither the corrected genetic distances accounting for intra-population variability (PiXY-(PiXY + PiY)/2) between populations nor the exact test of sample differentiation based on haplotype frequencies did yield significant values after Bonferroni corrections (Appendix 2). The test based on geographical location did not yield any significant pairwise comparisons (Appendix 3).

Table 6. Pairwise Φ_{ST} comparisons among North Atlantic fin whales samples from five locations (above diagonal) based on mtDNA. The number of migrants *Nm* was calculated according to Wright (1969), $Nm = (1-\Phi_{ST})/2*\Phi_{ST}$ and is indicated below the diagonal. Bold values indicated comparisons which were significant prior to Bonferroni correction.

	NOR	CAN	GREEN	SP85	IC83	IC85	IC87	IC89	IC06
NOR	0	**0.108**	0.004	0.025	0.011	**0.022**	0.022	**0.026**	-0.036
CAN	4.12	0	0.015	**0.058**	**0.055**	0.033	0.031	0.012	0.046
GREEN	124.5	32.83	0	0.016	-0.015	-0.019	0.003	-0.009	-0.054
SP85	19.6	8.89	30.75	0	0.009	0.015	-0.041	0.005	-0.036
IC83	44.95	8.6	-	55.06	0	-0.001	0.01	-0.002	-0.03
IC85	22.23	14.65	-	32.83	-	0	0.006	0.002	-0.037
IC87	22.23	15.63	166.17	-	49.5	82.83	0	0.008	-0.04
IC89	18.73	41.17	-	99.5	-	249.5	62	0	-0.018
IC06	-	10.37	-	-	-	-	-	-	0

The relationship of the different haplotypes detected was assessed by a neighbour joining tree (Figure 1) constructed with the software MEGA 4 (Tamura *et al*., 2007).

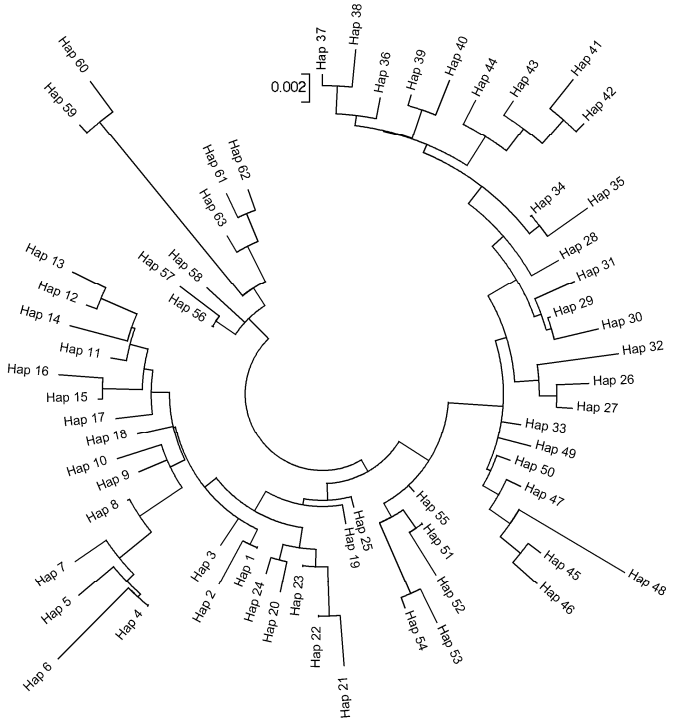

Figure 1. Neighbour joining tree showing relationship among the 63 fin whale mtDNA haplotypes detected. The tree was constructed with the software MEGA 4 (Tamura *et al*., 2007).

The Median Joining network did not reveal the presence of different groups based on haplotype distribution (Figure 2). The application of differential weights did not alter significantly the phyletic relationships between haplotypes. The weighted network was identical to the unweighted one.

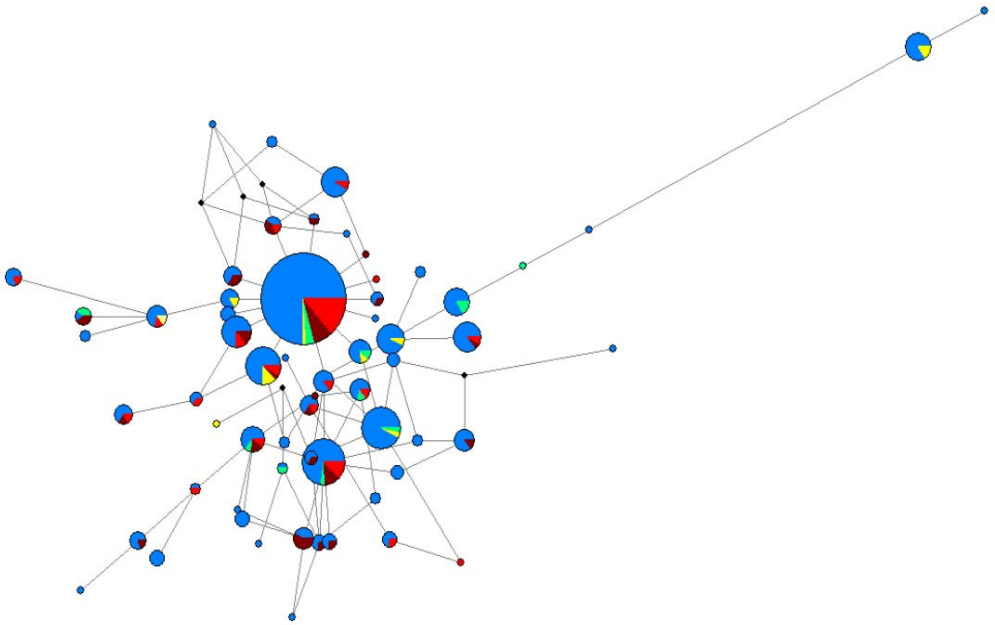

Figure 2. Median joining tree showing relationship among the 63 fin whale mtDNA haplotypes detected. The tree was constructed using the software Network4. Filling pattern represent the provenance of the samples (Blue=Iceland, yellow=Canada, Green=Greenland, Red=Spain, Brown=Norway). Connector length is proportional to the number of substitutions. Small closed circles represent potential intermediate haplotypes that were not sampled.

In Addition, the mismatch distributions analysis (Figure 3) was consistent with exponential population expansion as suggested by Bérubé *et al*. (1998), suggesting that population of fin whale are not at equilibrium.

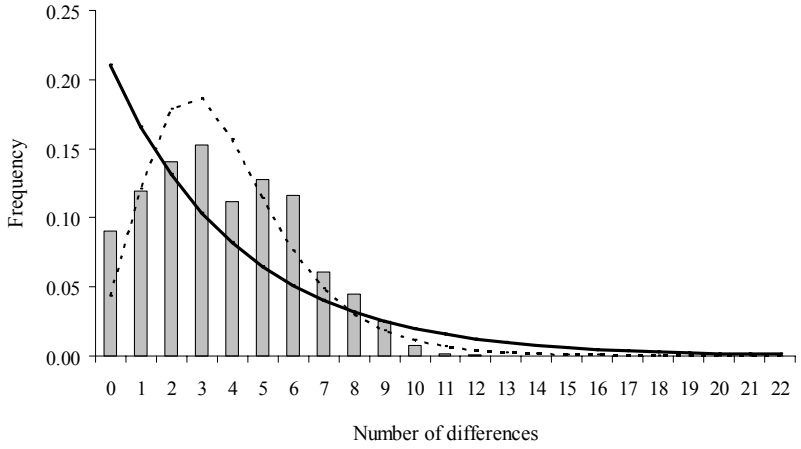

Figure 3. Observed mitochondrial DNA haplotype mismatch distribution (grey bars) and theoretical curves for an historical population expansion (dashed line) and for constant population size (plain line).

Discussion

The main objective of this new study was to gather a better genetic dataset to assess the potential genetic structure of fin whale at feeding grounds in the North Atlantic Ocean. The plan was to select individuals from the previously gathered database and to genotyped them for 15 microsatellite loci and the mtDNA with a maximum of success. A total of 475 individuals have been genotyped for more than 13 microsatellite loci while 466 have been successfully amplified for the mtDNA.

The observed genetic diversity as well as the overall level of genetic differentiation was similar to the one observed during previous studies (Daníelsdóttir *et al*., 2005, 2006) for both type of genetic markers. Additionally, both mtDNA and microsatellite loci gave congruent results and revealed a lack of significant genetic differentiation among samples collected at different feeding grounds of the North Atlantic Ocean. After Bonferroni corrections for multiple tests, none of the pairwise F_{ST} comparisons were significant, although it can be noted that F_{ST} values have a tendency to be 10-fold higher for comparisons involving the Canada sample collected in 1971.

These results based on both type of genetic markers could be explained by two different but non exclusive hypotheses: 1) the level of genetic connectivity among populations is rather high, 2) the pattern of genetic structure observed here is the result of a recent divergence event among populations. Indeed, the genetic pattern of present-day organisms has been shown to be related to historical events such as estuarine formation or postglacial expansion of the studied species which may, to some extent, have played a key role in the origin of marine population divergence. Therefore one of the most challenging tasks for geneticists has been to disentangle historical from contemporary genetic signatures (Hilbish, 1996). For example, the Pleistocene glaciations have left a distinct genetic imprint on the structure of many freshwater and terrestrial species, which can be interpreted as a postglacial range expansion from one or multiple ice-free refugia in the south. It has even been recently suggested that the isolation of marine fish populations in several refuges during the Pleistocene glaciations was reflected in their genetic structure across the North Atlantic (Bigg *et al*., in press; Pampoulie *et al*., in press). Most areas sampled during this study only became accessible to fin whale with the retreat of the last glaciation some 18,000 years ago, and may thus have been recolonised recently. This scenario will have resulted in a non-equilibrium phase and insufficient time may have passed for the level of genetic divergence to correspond to current gene flow. The mismatch distribution analysis performed on the tmDNA tends to confirm the latter as the pattern was consistent with exponential population expansion.

However, careful interpretation of the data should be performed, knowing that unbalanced sampling size might affect the estimate of genetic divergence among samples collected. Whatever the reason might be, the genetic markers employed during this study didn't reveal genetic pattern among samples collected at feeding grounds.

However, the observed genetic structure of species will crucially depend on the type of genetic markers used and their characteristics (mutation rate and neutral/non-neutral markers for example). Although the present study suggests limited genetic variation among samples collected at feeding grounds, which might reveal the presence of one single population or several populations not genetically distinguishable, the comparisons of these results with the allozyme data might give a better insight in the structure of this species. In fact, Daníelsdóttir *et al*. (submitted) clearly distinguished samples collected at Icelandic and Spain (1985)

feeding grounds as well as samples collected at Icelandic (1988), Norwegian (1991) and Canadian (1971) feeding grounds. Using locus-by-locus AMOVA, they reveal that some allozyme loci were outliers, possibly reflecting selection on these loci. Although possible selection effects have not been tested statistically on the allozyme data, this comparison might suggest that the observed F_{ST} values at the allozyme loci are likely to be due to diversifying selection that accentuates differences not readily apparent at neutral loci.

Acknowledgements
We would like to thank Dwaine Pittman, Jon Lien and all partners who kindly provided samples for this study.

References list

Árnason Ú (1981) Fin whales in the NE Atlantic, relationship between abundance and distribution. *Holarctic Ecology* **4**, 245-251.

Árnason A, Sigurðsson JH (1983) An electrophoretic study of protein and enzyme markers of the blood in three species of whales: *Balaenoptera physalus, Balaenoptera borealis, Physeter macrocephalus.* Paper SC/34/08 presented to the IWC Scientific Committee, 1983 (unpublished), 14pp.

Árnason A, Jónsdóttir S (1988) An electrophoretic study of cardiac esterases and proteins of fin whales (*Balaenoptera physalus*) from Icelandic and Spanish water, and sei whales caught off Iceland. Paper SC/39/Ba 6 presented to the IWC Scientific Committee, June 1988 (unpublished), 14pp.

Árnason A, Daníelsdóttir AK, Spilliaert R, Sigurðsson JH, Jónsdóttir S, Pálsdóttir A, Duke EJ, Joyce P, Groves V, Trowsdale J (1992) A brief review of protein and DNA marker studies in relation to the stock identity of fin whales (*Balaenoptera physalus*) from Iceland and Spain. *Report of the International Whaling Commission* 42, 701-705.

Bandelt H-J, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**: 37-48.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2002). *Genetix, a WindowsTM Based Software for Population Genetic Analyses* Version 4.05. Laboratoire Génome, Populations, Interaction CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Bérubé M, Aguilar A, Dendanto D, Larsen F, Notarbartolo di Sciara G, Sears R, Sigurjónsson J, Urban-R J and Palsbøll PJ (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. *Molecular Ecology* **7**, 585-599.

Bérubé M, Jorgensen H, McEwing R, Palsbøll PJ (2000) Polymorphic di-nucleotide microsatellite loci isolated from the humpback whale, *Megaptera novaeangliae*. *Molecular Ecology* **9**, 2181-2183.

Bigg GR, Cunningham CW, Ottersen G, Pogson GH, Wadley MR, Williamson P (in press) Ice-age survival of Atlantic cod: agreement between palaeoecology models and genetics. *Proceedings of the Royal Society of London B*.

Cassens I, van Waerebeek K, Best PB, Crespo EA, Reyes J, Milinkovitch MC (2003) The phylogeography of dusky dolphins (*Lagenorhynchus obscurus*): a critical examination of network methods and rooting procedures. *Molecular Ecology* **12**, 1781-1792.

Daníelsdóttir AK, Duke EJ, Joyce P, Árnason A (1991) Preliminary studies on the genetic variation at enzyme loci in fin whales *(Balaenoptera physalus)* and sei whales *(Balaenoptera borealis)* from the North Atlantic. *Report of the International Whaling Commission* (special issue 13), 115-124.

Daníelsdóttir AK, Sigurjónsson J, Mitchell E, Árnason A (1992). Report on a pilot study of genetic variation in skin samples of North Atlantic fin whales (*Balaenoptera physalus*). Paper SC/44/NAB16 presented to the IWC Scientific Committee, June 1992, 8pp.

Daníelsdóttir AK (1994) Genetic variation among different species and populations of baleen whales from the North Atlantic Ocean. Ph.D. Thesis, University College Dublin, Ireland. 308pp.

Daníelsdóttir AK, Stefánsson MÖ, Thorgilsson B, Jörundsdóttir ThD, Ragnarsdóttir A, Árnason A, Gunnlaugsson Th, Víkingsson GA, Ólafsdóttir D, Bérubé M, Palsbøll PJ, Øien N, Witting L, Pampoulie C (2005) Genetic analysis of North Atlantic fin whales (*Balaenoptera physalus*): Is there more than one breeding unit at the feeding ground west off Iceland? Paper presented at the *International Whaling Commission (IWC) Scientific Meeting*, 30th May – 8th June 2005, Ulsan, Republic of Korea. IWC

SC/57/PFI4, 12pp.
Daníelsdóttir AK, Bérubé M, Palsbøll PJ, Stefánsson MÖ, Thorgilsson B, Jörundsdóttir ThD, Ragnarsdóttir A, Árnason A, Gunnlaugsson Th, Ólafsdóttir D, Øien N, Witting L, Pampoulie C, Víkingsson GA (2006). Genetic stock delineation of fin whales. Paper presented at the joint North Atlantic Marine Mammal Commission (NAMMCO) and International Whaling Commission (IWC) Scientific workshop on the catch history, stock structure and abundance of North Atlantic fin whales, 23rd March – 26th March 2006, Reykjavík, Iceland. SC/58/FW/3 – SC/M06/FW5. 25pp.
Daníelsdóttir AK, Pampoulie C, Árnason A, Víkingsson GA, Duke EJ (submitted) What does re-analysis of North Atlantic fin whale allozymes tell us about their genetic composition at different feeding grounds?
Donovan GP (1991) A review of IWC stock boundaries. *Report of the International Whaling Commission* (special issue 13), 39-68.
Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47-50.
Favre L, Balloux F, Goudet J, Perrin N (1997) Female-biased dispersal in the monogamous mammal *Crocidura russula*: evidence from field data and microsatellite patterns. *Proceedings of the Royal Society of London B* **264**, 127-132.
Goudet J (1995) FSTAT (vers. 1.2): a computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.
Gunnlaugsson Th, Sigurjónsson J (1989) Analysis of the North-Atlantic fin whale marking data from 1979-1988 with special reference of Iceland. *Report of the International Whaling Commission* **39**, 383-388.
Hilbish TJ (1996) Population genetics of marine species: the interaction of natural selection and historically differentiated populations. *Journal of the Experimental Marine Biology and Ecology* **200**, 67-83.
International Whaling Commission (IWC) (2007) Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure (RMP). *Journal of Cetacean Research Management* **9**, 88-128.
Jover L (1987) Biometrical discrimination between Icelandic and Spanish stocks of fin whales (*Balaenoptera physalus*). Paper SC/38/Ba3 presented to the IWC, 1989 (unpublished), 12pp.
Jover L (1991) Reanalysis of morphometric differences between Icelandic and Spanish stocks of fin whales *(Balaenoptera physalus)*. Paper SC/F91/F44 presented to the IWC Scientific Committee, February 1991 (unpublished), 9pp.
North Atlantic Marine Mammal Commission (NAMMCO) (2006) Report of the NAMMCO Scientific Committee working group on fin whales. Oslo, 20th – 22nd October 2005. NAMMCO Annual Report 2006. *Report of the Scientific Committee, Annex 2*, 252-278.
Palsbøll PJ, Bérubé M, Larsen AH, Jorgensen H (1997) Primers for the amplification of tri- and tetramer microsatellite loci in baleen whales. *Molecular Ecology* **6**, 893-895.
Palsbøll PJ, Bérubé M, Aguilar A, Notarbartolo-Di-Sciara G, Rasmus N (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.
Pampoulie C, Stefánsson MÖ, Danilowicz BS, Daníelsdóttir AK (in press) Recolonisation history and large scale dispersal in the open sea: the case study of the North Atlantic Cod (*Gadus morhua* L.). *Biological Journal of the Linnean Society.*
Rice WR (1989) Analyzing table of statistical tests. *Evolution* **43**, 223-225.
Rørvik CJ, Jonsgård Å, (1981) Review of balaenopterids in the North Atlantic Ocean. *F.A.O.*

Fish. Ser. [Mammals in the Seas] **3**: 269-286.
Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496-2497.
Sigurjónsson J (1995) On the life history and autoecology of North Atlantic Rorquals. In. whales, seals, fish and man (eds Blix AS, Walløe L, Ulltang Ø), pp425-441. Elsevier, Amsterdam.
Sigurjónsson J, Gunnlaugsson Th, Ensor P, Newcomer M, Víkingsson G (1991) North Atlantic Sightings Survey 1989 (NASS-89): Shipboard surveys in Icelandic and adjacent waters July-August 1989. *Report of the International Whaling Commission* **41**, 559-572.
Sigurjónsson J, Gunnlaugsson Th (1985) Further mark-recapture analysis of fin whales caught off Iceland with a note on stock identity and movements of the East-Greenland/Icelandic population. *Report of the International Whaling Commission* **5**, 257-362.
Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**, 555-562.
Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596-1599.
Valsecchi E, Amos W (1996) Microsaetllite markers for the study of cetacean populations. *Molecular Ecology* **5**, 151-156.
Víkingsson GA, Gunnlaugsson Th, (2006) Stock structure of fin whales (*Balaenoptera physalus*) in the North Atlantic – Indications from non-genetic data. Paper presented at the joint North Atlantic Marine Mammal Commission (NAMMCO) and International Whaling Commission (IWC) Scientific workshop on the catch history, stock structure and abundance of North Atlantic fin whales, 23rd March – 26th March 2006, Reykjavík, Iceland. SC/58/FW/7 – SC/M06/FW7.
Víkingsson GA, Pike DG, Desportes G, Øien N, Gunnlaugsson Th, Bloch D (2008) Distribution and abundance of fin whales (*Balaenoptera physalus*) in the Northeast and Central Atlantic as inferred from the North Atlantic Sightings Surveys 1987-2001. NAMMCO Scientific Publication vol (in press).
Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* **10**, 506-513.
Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.
Wright S (1969) Evolution and the genetics of populations. Vol. 2. The theory of gene frequencies. University of Chicago Press, Chicago.

Appendix 1 *Microsatellite loci analysis*. Above diagonal: Pairwise F_{ST} comparisons among North Atlantic fin whales samples from five locations. Bold values indicated *P*-values which were significant prior to Bonferroni correction. None of the comparisons were significant after Bonferroni correction (Indicative adjusted nominal level (5%) for multiple comparisons is: $P = 0.000549$). Below diagonal: The number of migrants *Nm* was calculated according to Wright (1969), $Nm = (1-F_{ST})/4*F_{ST}$.

	CAN	SP85	GREEN	IC06L	IC83E	IC85E	IC83L	IC85L	IC83M	IC85M	IC87M	IC89E	IC89M	NOR
CAN		0.0169	**0.0197**	0.0027	**0.0180**	**0.0144**	**0.0200**	0.0152	**0.0201**	**0.0182**	0.0208	**0.0132**	**0.0258**	**0.0204**
SP85	14.52		0.0005	-0.0083	0.0055	0.0018	-0.0022	-0.0012	-0.0003	-0.0004	-0.0040	-0.0041	0.0028	0.0017
GREEN	12.42	552.26		-0.0013	-0.0012	-0.0009	0.0049	0.0020	0.0001	0.0004	-0.0061	0.0013	-0.0023	-0.0064
IC06L	93.15	∞	∞		-0.0099	-0.0102	-0.0134	0.0013	-0.0072	-0.0089	-0.0296	-0.0099	-0.0016	-0.0037
IC83E	13.65	45.1	∞	∞		-0.0020	0.0005	0.0051	0.0016	0.0008	-0.0084	0.0039	0.0018	0.0042
IC85E	17.17	138.05	∞	∞	∞		-0.0017	0.0045	-0.0011	-0.0005	-0.0094	-0.0016	0.0019	0.0036
IC83L	12.24	∞	51.21	∞	493.97	∞		0.0104	-0.0011	-0.0003	-0.0143	-0.0020	0.0021	0.0083
IC85L	16.24	∞	125.08	191.34	48.81	54.77	23.7		0.0059	0.0034	0.0019	0.0006	0.0006	0.0023
IC83M	12.22	∞	2987.04	∞	160.18	∞	∞	42.36		-0.0010	-0.0049	-0.0009	0.0005	0.0026
IC85M	13.48	∞	685.82	∞	319.38	∞	∞	73.08	∞		-0.0047	0.0000	0.0015	0.0021
IC87M	11.8	∞	∞	∞	∞	∞	∞	132.27	∞	∞		-0.0106	-0.0071	-0.0011
IC89E	18.67	∞	186.84	∞	63.69	∞	∞	411.97	∞	6708.04	∞		0.0029	0.0020
IC89M	9.43	89.87	∞	∞	141.96	131.24	117.16	388.33	534.69	167.3	∞	86.62		-0.0017
NOR	12	150.82	∞	∞	59.71	69.46	29.74	108.67	97.63	118.7	∞	124.99	∞	

Appendix 2 *Mitochondrial (mtDNA) data analyses*. Above diagonal: Exact test *P*-values of samples differentiation based on haplotype frequencies. Below diagonal: Corrected genetic distances accounting for intra-population variability (PiXY-(PiXY + PiY)/2). Bold values indicated *P*-values which were significant prior to correction for multiple tests. None of the comparisons were significant after Bonferroni correction.

	NOR	CAN	GREEN	SP85	1983	1985	1987	1989	2006
NOR	0	**0.0122**	0.4856	0.3874	0.0518	**0.0031**	0.9151	**0.0069**	0.3069
CAN	**0.4744**	0	0.1423	**0.0337**	0.4319	0.3436	0.2460	0.3746	0.2781
GREEN	0.0117	0.0630	0	0.1783	0.2827	0.5455	0.7076	0.0725	0.7529
SP85	0.0895	0.1872	0.0343	0	0.2787	0.6210	0.6202	0.7176	0.2896
1983	0.0373	**0.1858**	-0.0728	0.0321	0	**0.0362**	0.5172	**0.0301**	0.3782
1985	**0.0918**	0.1376	-0.0787	0.0664	-0.0012	0	0.3655	**0.0376**	0.1489
1987	0.1522	0.1575	0.0667	-0.0976	0.0824	0.1132	0	0.4954	1.0000
1989	**0.1009**	0.0354	-0.0448	0.0194	-0.0084	-0.0124	0.0914	0	0.1496
2006	-0.1597	0.1868	-0.2191	-0.1810	-0.1568	-0.1602	-0.1361	-0.1034	0

Appendix 3

Microsatellite analysis. Above diagonal: Pairwise F_{ST} comparisons among North Atlantic fin whale geographical samples. Below diagonal: F_{ST} P values. Bold values indicated *P*-values which were significant after Bonferroni correction (Indicative adjusted nominal level (5%) for multiple comparisons is: $P = 0.005$).

	NORWAY	CANADA	GREENLAND	SPAIN	ICELAND
NORWAY	0	0.020	-0.006	0.002	0.003
CANADA	0.006	0	0.020	0.017	0.018
GREENLAND	0.721	0.069	0	0.000	0.000
SPAIN	0.118	0.061	0.351	0	0.001
ICELAND	0.079	**0.001**	0.430	0.306	0

Mitochondrial (mtDNA) data analyses. Above diagonal: Pairwise Φ_{ST} comparisons among North Atlantic fin whale samples from five geographical locations (above diagonal) based on mtDNA. Below diagonal: Corrected genetic distances accounting for intra-population variability (PiXY-(PiXY + PiY)/2). None of the comparisons were significant after Bonferroni correction.

	NORWAY	CANADA	GREENLAND	SPAIN	ICELAND
NORWAY	0	0.027	-0.005	0.014	0.004
CANADA	0.026	0	0.020	0.055	0.010
GREENLAND	-0.005	0.019	0	0.011	-0.005
SPAIN	0.012	0.045	0.007	0	0.009
ICELAND	0.004	0.008	-0.005	0.010	0