Distribution of mtDNA haplotypes in North-Atlantic humpback whales

Palsboll, PJ; Clapham, PJ; Mattila, DK; Larsen, F; Sears, R; Siegismund, HR; Sigurjonsson, J; Vasquez, O; Arctander, P

Published in:
Marine Ecology Progress Series

DOI:
10.3354/meps116001

Vol. 116: 1–10, 1995 | MARINE ECOLOGY PROGRESS SERIES Mar. Ecol. Prog. Ser. | Published January 12

Distribution of mtDNA haplotypes in North Atlantic humpback whales: the influence of behaviour on population structure

Per J. Palsbøll[1], Phillip J. Clapham[2], David K. Mattila[2], Finn Larsen[3], Richard Sears[4], Hans R. Siegismund[5], Johann Sigurjónsson[6], Oswaldo Vasquez[7], Peter Arctander[1]

[1] Department of Population Biology, Zoological Institute, University of Copenhagen, Universitetsparken 15, DK-2100 København Ø, Denmark

[2] Cetacean Research Program, Center for Coastal Studies, Provincetown, Massachusetts 02657, USA

[3] Greenland Fisheries Research Institute, Tagensvej 135, 1., DK-2200 København N, Denmark

[4] Mingan Island Cetacean Study, Inc., 285 Green Street, St-Lambert, Quebec, Canada J4P 1T3

[5] Arboretum, The Royal Veterinary and Agricultural University, Kirkegårdsvej 3A, DK-2970 Hørsholm, Denmark

[6] Marine Research Institute, Skúlagata 4, PO Box 1390, 121 Reykjavik, Iceland

[7] Centro de Investigacciones de Biologia Marina, Universidad Autonoma de Santo Domingo, Jonas Salk 55, Santo Domingo, Dominican Republic

ABSTRACT: Samples from 136 humpback whales *Megaptera novaeangliae*, representing 5 feeding aggregations in the North Atlantic and 1 in the Antarctic, were analyzed with respect to the sequence variation in the mitochondrial (mt) control region. A total of 288 base pairs was sequenced by direct sequencing of asymmetrically amplified DNA. Thirty-one different haplotypes were identified. The nucleotide diversity for the total sample was estimated to be 2.6%, which is high relative to other North Atlantic cetaceans. The degree of genetic differentiation in various subsets of the samples was estimated and tested for statistical significance by Monte Carlo simulations. Significant degrees of heterogeneity were found between the Antarctic and all North Atlantic areas, as well as between Iceland and the western North Atlantic samples. A genealogical tree was estimated for the 31 haplotypes and rooted with the homologous sequence from a fin whale *Balaenoptera physalus*. The branching pattern in the genealogical tree suggests that the North Atlantic Ocean has been populated by 2 independent influxes of humpback whales. The combined results from the homogeneity tests and the genealogical tree indicate that behaviour (in this case maternally directed site fidelity to a foraging area) can influence the population structure of marine cetaceans on an evolutionary time scale.

KEY WORDS: Philopatry · Population genetics · Mitochondrial control region · PCR

INTRODUCTION

While the dispersal and segregation of populations of terrestrial mammals are frequently influenced by geographic features or climatic characteristics, few such obvious barriers exist in the oceans. Consequently, we might expect population structure among the pelagic cetaceans to reflect this lack of constraint. Indeed, behavioural studies have shown that the humpback whale *Megaptera novaeangliae* is one of the widest-ranging mammals in the world, making annual migrations of up to 8000 km from high-latitude feeding grounds to breeding and calving areas in tropical waters (Dawbin 1966, Stone et al. 1990). However, while all studied populations of this species undertake such migrations, it is apparent that, during the summer feeding season, the whales exhibit more restricted movement and geographic segregation (Clapham & Mayo 1987, Katona & Beard 1990, Clapham et al. 1993a).

In the North Atlantic, the identification of individual whales over the past 15 yr has suggested the existence of a number of relatively discrete feeding aggrega-

tions, fidelity to which is determined matrilineally (Katona & Whitehead 1981, Katona & Beard 1990). These areas include the Gulf of Maine (Clapham et al. 1993a), Newfoundland/Labrador (Perkins & Whitehead 1977), the Gulf of St. Lawrence (Katona & Beard 1990), West Greenland (Perkins et al. 1984) and Iceland (Sigurjonsson & Gunnlaugsson 1990). Humpbacks are also known to feed in the waters of arctic and subarctic Norway (Christensen et al. 1992), although little is known of the current status of this population. As more individuals have been identified, an increasing number of humpbacks have been recorded in more than one feeding area, although the proportion remains very low relative to the total number of resightings (Katona & Beard 1990). Within specific feeding areas, distribution is largely determined by the availability of prey (Payne et al. 1990).

With the exception of Norwegian humpbacks, whose migratory destination is presently unknown, whales from the various North Atlantic feeding aggregations all migrate to the West Indies each winter (Martin et al. 1984, Mattila et al. 1989, Katona & Beard 1990). Furthermore, observations of whales from different high-latitude areas in the same competitive groups (Clapham et al. 1993b), and the spatial mixing that is evident throughout the West Indies breeding range, have strongly suggested that the North Atlantic population of this species constitutes a single panmictic unit.

The investigations noted above have recently been complemented by genetic studies based upon molecular analyses of skin samples collected as biopsies from free-ranging whales. Baker et al. (1993) examined worldwide variation in the mitochondrial (mt) control region and showed that humpback whales in the North Pacific, North Atlantic and Southern Ocean constitute separate populations with very limited inter-oceanic exchange. Within the North Pacific, highly significant differences in mt haplotype frequencies have been found among whales from regions that behavioural studies have suggested represent different panmictic entities (Baker et al. 1990).

A natural extension of the molecular studies conducted to date is to test whether the observed segregation of individuals into different feeding aggregations within the same ocean is maintained over periods that are long enough to be of evolutionary significance. Given that the 2 factors which appear to determine present population structure (maternally directed foraging area philopatry and spatial distribution of prey) presumably are, from an evolutionary perspective, relatively plastic and short-lived phenomena, we might expect little or no reflection of these patterns in the genetic structure of the population.

We report here the results of an analysis of the distribution of substitutions in the nucleotide sequence of the mt control region (also called the D-loop) of 136 humpback whales from 5 feeding aggregations in the North Atlantic and 1 in the Antarctic. We find that, contrary to our expectations, the observed matrilineal fidelity to specific areas appears to have a significant effect on the overall North Atlantic population structure on an evolutionary time scale.

The mt genome is maternally inherited, non-recombining, and evolves at a rate 5 to 10 times higher than single-copy nuclear sequences (Brown et al. 1979, Brown 1985). The control region is the only major non-coding region and most rapidly evolving part of the mt genome (Upholt & Dawid 1977, Cann et al. 1984), making it particularly well suited for the study of intraspecific evolution. The combination of these characteristics makes the mt control region an ideal genetic marker for testing the hypothesis of long-term segregation of maternally directed feeding aggregations in an otherwise panmictic population. Although Hoelzel et al. (1991) and Baker et al. (1993) found that the cetacean mt control region evolved at a lower rate than that of other mammals, the present study shows a high degree of genetic variation within the North Atlantic humpback whale population, at levels similar to that observed in human populations (Vigilant et al. 1991). This variation has allowed us to detect statistically significant segregation, and to estimate a reliable genealogy for major clades of haplotypes.

METHODS

Sample collection and conservation. The samples included in this study were collected either as skin biopsies from free-ranging whales (Lambertsen 1987, Palsbøll et al. 1991) or as sloughed skin (Clapham et al. 1993c). Sampling was conducted on 4 feeding grounds: the Gulf of Maine, the Gulf of St. Lawrence, West Greenland and Iceland. A number of other whales sampled in Samana Bay, Dominican Republic (a breeding area), were later found from photo-identification to have been previously observed in the Newfoundland/Labrador area, or in the Gulf of Maine, thus providing additional samples for the study of genetic variation among high-latitude feeding aggregations. Eleven samples collected off the Antarctic Peninsula were also included to provide an outgroup as a frame of reference for interpreting the North Atlantic samples.

All sampled whales were photographically identified by variations in natural markings, specifically those of the ventral fluke pattern and the dorsal fin (Katona & Whitehead 1981). All photographs were compared to

ensure the exclusion of multiple samples from single individuals.

Sighting histories and knowledge concerning degree of consanguinity between individuals were available for many of the Gulf of Maine whales from a long-term study of identified individuals in that region (Clapham et al. 1993a). These data were used to confirm and interpret the results of the mtDNA analyses.

Samples were conserved in saturated sodium chloride with 25% dimethylsulfoxide at −20°C (Amos & Hoelzel 1991).

DNA extraction, amplification, sequencing and sex determination. DNA was extracted from the tissue samples following standard protocols with cell lysis in 1.0% SDS (sodium lauryl sulfate), 0.15 M sodium chloride, 10 mM Tris-HCl (pH 8.0), 1 mM EDTA (ethylenediamine tetraacetic acid) and digested with proteinase K (100 µg ml^{-1}) at 65°C for a minimum of 3 h followed by phenol/chloroform extractions and precipitation with ethanol (Maniatis et al. 1982).

The nucleotide sequence of the 5' end of the mt control region was determined by direct sequencing of asymmetrically amplified DNA by the polymerase chain reaction (PCR) (Gyllensten & Erlich 1988, Saiki et al. 1988). Symmetric double-stranded amplification of the whole control region was performed with primers MT3 and MT4 (Table 1) in 20 µl reactions; [0.067 M Tris-HCl (pH 8.8), 0.002 M $MgCl_2$, 0.0166 M NH_3SO_4, 0.01 M β-mercaptoethanol, 0.2 µM per nucleotide], 1 µM for each primer, 0.4 units AmpliTaq™ DNA polymerase and 10 ng extracted genomic DNA. The temperature profile was: 1 min at 94°C, 1 min at 57°C and 3.5 min at 72°C for 28 cycles. For asymmetric amplification the primers MT4 and Mn312 were used (Table 1). Amplifications were conducted under the same conditions as above except that the total reaction volume was 50 µl, the annealing temperature 55°C and the extension time 1.5 min for 30 cycles. Blank extractions and blank PCR reactions were included as controls to detect possible contamination. Both strands were sequenced for approximately half of the samples. Sequencing was conducted according to the manufacturer's instructions (US Biochemicals, Inc.). MT4, Mn312, Mn196 and Mn152 were used as sequencing primers (Table 1). The sex of each sampled whale was determined by a PCR-based method (Palsbøll et al. 1992).

Table 1. Primer sequences. F and R, respectively, denote a forward- or reverse-oriented primer, with reference to the light strand

Primer code	Nucleotide sequence	Source
MT4-F	CCT CCC TAA GAC TCA AGG AAG	Arnason et al. (1993)
MT3-R	CAT CTA GAC ATT TTC AGT G	Arnason et al. (1993)
Mn312-R	CGT GAT CTA ATG GAG CGG CCA	This study
Mn152-R	GCA CGA ATT ACA TAA TGG TAT	This study
Mn196-F	ACT GAT AGC ACC TTC CAT GAG T	This study

Data analysis. Nucleotide diversity was estimated as defined by Nei & Li (1979). The extent of genetic differentiation between various partitionings of samples was estimated using either the sequence (K_{ST}) or haplotype (H_{ST}) statistic proposed by Hudson and coworkers (Hudson 1992, Hudson et al. 1992, Roff & Bentzen 1992). K_{ST} is identical to γ_{ST} defined by Nei (1982), except for the weighting of the average number of differences between the sampled sequences within each locality. For K_{ST} (the statistic used in this study) the actual number of sampled sequences is used, whereas for γ_{ST} each sample location is usually given the same weighting in the pairwise comparisons. The choice of statistic used for testing the level of significance in the pairwise comparisons was that suggested by Hudson et al. (1992). When the degree of variation was high or one of the 2 samples was small, the K_{ST} statistic was chosen.

Samples in pairwise comparisons were considered to be significantly heterogeneous when the probability of obtaining the observed, or a more extreme, value of either H_{ST} or K_{ST} was less than 0.05 in 1000 Monte Carlo simulations (Hudson et al. 1992).

Homogeneity tests were performed by pairwise comparisons of samples in a hierarchical manner as follows: males/females within each area; residents/non-residents (for the southern Gulf of Maine only — see below); year/year; feeding aggregation/feeding aggregation; and hemisphere/hemisphere.

Homogeneous samples were combined in the subsequent comparisons. Other partitionings of the specimens were tested for homogeneity, and these are described in more detail in 'Results'.

The Gulf of Maine samples were divided into 'resident' and 'non-resident' individuals based upon the frequency with which an individual had been observed in the southern Gulf of Maine study area (Massachusetts Bay and the Great South Channel; see Clapham et al. 1993a). For the sampled individuals the overall mean number of days sighted per year was 4.3 (range = 0 to 21.3 d yr^{-1}). Individuals with a mean of less than 4.3 d yr^{-1} were categorized as non-residents, and those with a mean above 4.3 as residents. The extent of genetic differentiation between residents and non-residents was estimated and tested for significance.

A rooted genealogical tree was estimated, with the homologous sequence from a North Atlantic fin whale *Balaenoptera physalus* as an

outgroup (Arnason et al. 1991). Using the PHYLIP 3.5c computer package (Felsenstein 1993) a tree was estimated based on the maximum likelihood method. A total of 10 trees with random input order were estimated based on a transition/transversion ratio of 25:1 (the ratio that yielded the highest log-likelihood score). Using the Neighbor-joining method (Saitou & Nei 1987), a consensus tree was estimated from 400 bootstrap samples generated from the original data set. Distance matrices were calculated using the algorithm defined by Kimura (1980) and input order randomized. The estimated topologies obtained from the maximum likelihood and Neighbor-joining methods were compared using the Kishino-Hasegawa-Templeton test (Kishino & Hasegawa 1989) which is included in the PHYLIP 3.5c DNAML program.

Table 2. Number of samples from each area and years in which they were collected. Sampled individuals from Samana Bay, Dominican Republic, were later identified by the pigmentation pattern on the ventral side of the fluke to humpback whales previously observed in Newfoundland/Labrador or the Gulf of Maine

Area	1988	1989	1990	1991	Samana Bay	Total
Gulf of Maine			23		7	30
Gulf of St. Lawrence			9	16		25
Newfoundland					13	13
West Greenland	16	17		11		44
Iceland				13		13
Antarctic Peninsula		11				11
Total	16	28	32	40	20	136

RESULTS

Samples and haplotypes

The number of samples and the areas and years in which they were collected are summarized in Table 2.

The first 288 base pairs of the 5' end of the mt control region were successfully sequenced in 136 individuals; the quality of these sequences has been demonstrated in a previous publication (Clapham et al. 1993c). A total of 36 polymorphic sites defined 31 haplotypes. With the exception of 1 deletion and 2 transversions, all substitutions were transitions. Fig. 1 shows the sequence of haplotype no. 9 (Table 3; see also Figs. 2 & 3) aligned with a fin whale sequence (Arnason et al. 1991), and Table 3 summarizes the haplotype and sex of all specimens.

In the 11 specimens from the Antarctic Peninsula, 7 haplotypes were detected, 3 of which were found only in single specimens. Two Antarctic whales shared a haplotype with 1 North Atlantic whale from the Gulf of St. Lawrence (haplotype no. 10 in Table 3). Because of the unexpected nature of this discovery, all 3 specimens were amplified and sequenced a second time, with the same result. The 125 specimens from the North Atlantic divided into 25 haplotypes, of which 12 were represented by single individuals.

Comparisons with data previously published in a study of the worldwide variation in the mt control region of humpback whales (Baker et al. 1993) were not possible since the sequences concerned were not available from any of the common sequence databases. The portion of the control region analyzed appears to be different from that sequenced in this study, starting about 150 base pairs into the mt control region.

```
Base no.                1
M. novaeangliae         AAAAAATGTA TATTGTACAA TAACCATAGG GCCACAGTAC TATGTCCGTA TTGAAAATAG CTCATTCCAT
B. physalus             .....--... .......... ......C... A......... .......... .........A ..TGCCTT..

                                  101
TACATACCAT TATGTAATTC GTGCATGTAT GTATTACCAC ATAACCAACT GATAGCATCT T-CCATGGGT AT-------- ----------
..G...TT.. .......C.. .......C.. ...C.T.... ....TT..-- --.....G... .T......... ..GAACAGAT ATACATGCTA

                                  201
-GTATAATCG TGCATTCAAT TATTTTCACT ACCAGCAGTT AAAGCTCGTA TTAAATTTTA TTAATTTTAC ATATTACATA ATATGTATTA
T........T. .......... .........C .......... C......... .......... .......... .......... ..........

                                                  301
ATAATACAAT AGTACATGTC CTTATAAATT CCCAGGTTTA TTTAAATCAA ATGATCCC
...G..... ..CG.....T .....GC..C .....A.C.. .......... .....T..
```

Fig. 1. *Megaptera novaeangliae*. The sequence of the first 288 nucleotides of the mt control region in haplotype no. 9 in Table 3 aligned with the published homologous sequence in the fin whale *Balaenoptera physalus* (Arnason et al. 1991). Dots indicate an identical nucleotide; hyphens were included to optimise alignment

Table 3. *Megaptera novaeangliae*. Haplotype and gender for the analyzed specimens. The first column represents the haplotype number corresponding to the numbers in Figs. 2 & 3. The 4 middle columns show the polymorphic sites in the 31 haplotypes identified in the 136 analyzed specimens. The numbers above list the nucleotide position of the polymorphic site starting from the 5' end of the mitochondrial control region. Haplotypes 2 through 31 are listed with reference to haplotype number 1. -: a deletion; .: an identical nucleotide at the position relative to haplotype number 1; A: adenine; C: cytosine; G: guanine; T: thymine. The final columns list the ID numbers of the analyzed samples with respect to haplotype. The first 2 letters denote the sampling location, the first 2 digits the sampling year and the last 4 digits are a serial identification number. AP: Antarctic Peninsula; GL: Gulf of St. Lawrence; GM: Gulf of Maine; IL: Iceland; SB: Samana Bay; WG: West Greenland. ID numbers which are underlined are males; those not underlined are females. The following ID numbers of samples collected in Samana Bay were later matched to individuals previously observed in the Gulf of Maine: SB90004, SB90028, SB90036, SB90039, SB90063, SB90080, SB91019 or Newfoundland: SB90009, SB90013, SB90015, SB90019, SB90058, SB90073, SB90078 SB91003, SB91005, SB91008, SB91022, SB91039, SB91043

	11 1111111111 1122222222 222222 112268801 2233345666 7844566666 777788 6387978940 8978994458 0629546789 026767	Sample ID nos
1	-TCCGCTTTC TCAGTCATTA TTGTCAGTTT ATATTC	GL91100 WG89016 WG91011
2	....T.....	GL90011 GL90028 GL91011 GL91029 GL91101 SB90004 SB91019 GM90054 IL91004 SB90009 SB91039 SB91043 WG89001 WG89008 WG91002
3	....G... ....T.....	GL91022
4	....T..... ....C.	GM90015 GM90019 GM90033 GM90051 WG88004 WG88018
5	....T.... ....T..... ....C.	GL90019
6	........C. ....T..... ....CT	GL91013 SB91003
7	........C. ....T.A... ....C.	SB90039 SB91005
8	........C. ....T..... ....C.	GL91103 IL91005 IL91012 IL91013 WG89003 WG89006 WG89009 WG89019
9	A..T....C. ..G..... ....T..... ....C.	IL91003
10	........C. C......... ....T..... ....C.	AP89003 AP89006 GL91105
11	......C.C. C......... .C..T..... ....C.	GL91065
12	......C.C. C......... ....T..... ....C.	GL91064 GM90008 IL91010 WG88025 WG89005 WG89013 WG91003
13	........C. C...C..C.. ....T..... ..G.C.	GL91027
14	........C. C......C.. ....T..... ..GAC.	WG88014
15	........CT C......C.. ....T..... ..G.C.	WG88001
16	........C. C......C.G ....T..... ..G.C.	GM90056 GM90057
17	........C. C......C.. ....T..... ..G.C.	GL90007 GL90015 GL91028 GL91104 GM90013 GM90021 GM90024 GM90028 GM90034 GM90035 GM90041 GM90055 GM90059 IL91001 IL91002 IL91006 IL91008 IL91009 IL91016 SB90058 SB90073 SB91008 WG88015 WG88016 WG89004 WG89007 WG89018 WG91009 WG91012
18	...T....C. C......C.. ....T..... ..G.C.	IL91015
19	....AT.CA. CT......C. C..CT....C ..G.CT	AP89007 AP89010
20	A....T.CA. CT.A..... C..CT....C ..G.CT	AP89004 AP89005
21	....T.CA. CT.A..... C..CT..... ..G.CT	AP89009
22	.....T..C. CT.A..... C..CT...C. ..G.CT	AP89011 AP89012
23	.....TC.C. CT...T..C. CC..T..C.C .CG.CT	GL90030 SB90080 GM90006 GM90026 SB90015 SB90078 SB91022 WG88002 WG88013 WG88021 WG88022 WG88023 WG89011 WG89012 WG91001 WG91004 WG91005 WG91006 WG91007 WG91008
24	.....TC.C. CT......C. CC..TG.C.C .CG.CT	GM90058
25	..T..TC.C. CT...T..C. CC..TG.C.C .CG.CT	SB90019
26	.....TC.C. CT...T..C. CC..TG.C.C .CG.CT	GL90004 SB90063 GM90011 GM90012 SB90013 WG88012 WG88017 WG89002 WG89017
27	...T.TC.C. CT...T..C. CC..T..C.C .CG.CT	GL91102
28	.....T..C. CT...T..C. CC..T..C.C .CG.CT	GL90029 GL90031 GL91051 SB90028 SB90036 WG88003 WG88020 WG89010
29	.....TC.C. CT...T..C. C...T..C.C .CG.CT	GL91030 GM90050
30	AC...T.CC. CT...T..C. C...T....C ..G.CT	AP89002
31	.....T..C. C.G..T.... C.A.T...CC GC..CT	AP89001

Homogeneity tests

No heterogeneity was detected within the 6 areas between sexes, residents/non-residents, and years. Consequently, males and females in each area, as well as the residents/non-residents in the Gulf of Maine, were pooled in the subsequent homogeneity tests. Within the western North Atlantic (all North Atlantic feeding grounds except Iceland) no heterogeneity was found and consequently the samples from this area were pooled.

Significant degrees of heterogeneity were found between the Antarctic Peninsula and all North Atlantic areas, as well as between Iceland and all other North Atlantic regions (Table 4); the latter (Gulf of Maine, Gulf of St. Lawrence, Newfoundland and West Greenland) are hereinafter collectively referred to as the 'western North Atlantic'.

The 31 specimens from the Gulf of Maine included whales from 3 known maternal lineages. The analysis revealed that each lineage possessed a different haplotype. Although samples from all feeding grounds probably include related individuals we could only test any effects of such an inclusion for the Gulf of Maine. We removed all but 1 animal of each known maternal lineage in the Gulf of Maine sample and included the reduced sample in the pairwise comparisons with other feeding grounds. No additional heterogeneity was detected. There was a slight but non-significant increase in the nucleotide diversity in the reduced sample relative to the non-reduced Gulf of Maine sample.

Nucleotide diversity

The estimated nucleotide diversity for the total sample was 2.6%. The estimates for the feeding aggregations within the western North Atlantic and the Antarctic Peninsula varied between 2.4 and 2.9% with no significant inter-area differences (Table 5). The nucleotide diversity in the Icelandic sample (0.86%) was 3 times lower than in any of the western North Atlantic feeding aggregations.

Table 4. K_{ST} values (indicating the extent of genetic differentiation) in pairwise comparisons and the significance level. $^{**}p < 0.01$; $^{***}p < 0.001$

	Antarctic Peninsula	Western North Atlantic	Iceland
Antarctic Peninsula			
Western North Atlantic	0.047***		
Iceland	0.33***	0.040**	

Table 5. *Megaptera novaeangliae*. Estimates of the nucleotide diversity and standard error within each humpback whale feeding aggregation

Feeding aggregation	Nucleotide diversity	SE
Antarctic Peninsula	0.024	0.0038
Gulf of Maine	0.025	0.0022
Gulf of St. Lawrence	0.025	0.0027
Newfoundland	0.029	0.0031
West Greenland	0.026	0.0010
Iceland	0.0086	0.0015

Estimated genealogy

None of the estimated genealogies based on either the maximum likelihood or the Neighbor-joining method were statistically different. The tree with the highest log-likelihood value is shown in Fig. 2 and the majority rule consensus tree in Fig. 3 with the bootstrap values over 50%. All estimated trees have the same basic topology, as indicated by the bootstrap values.

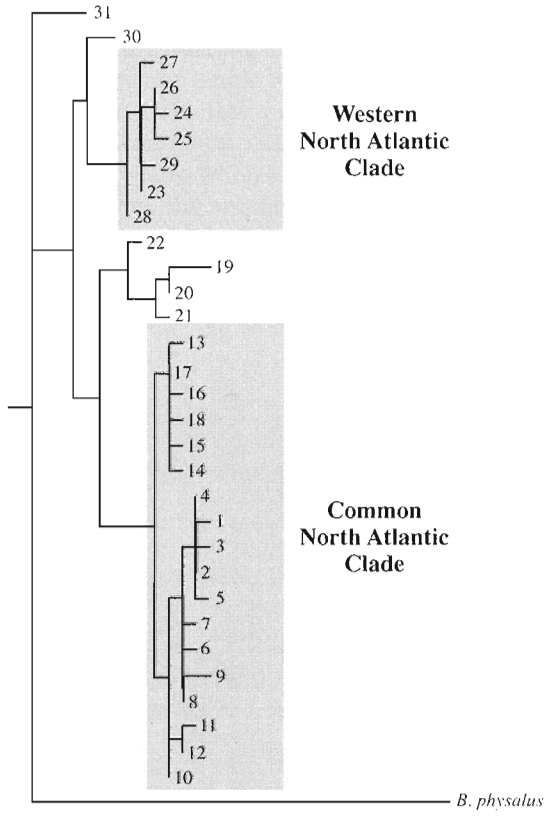

Fig. 2. *Megaptera novaeangliae*. Genealogy with the highest log-likelihood value. Numbers labeling the tip of the nodes correspond to the haplotype number in the far left column in Table 3. *B. physalus*: fin whale *Balaenoptera physalus*

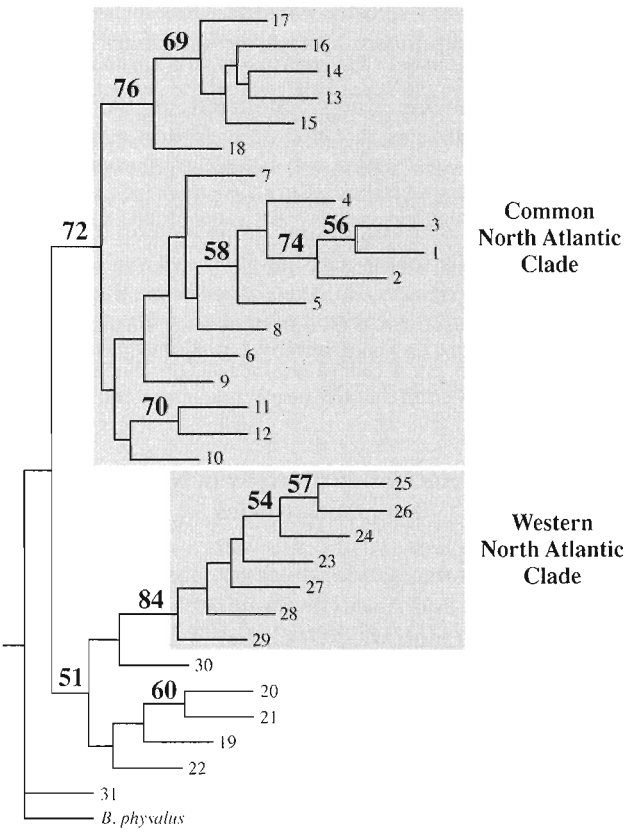

Fig. 3. *Megaptera novaeangliae*. The majority rule consensus genealogy estimated by the Neighbor-Joining method on 400 datasets created by bootstrapping. Large boldface numbers are bootstrap values. Numbers labeling the tip of the nodes correspond to the haplotype number in the far left column in Table 3. *B. physalus*: fin whale *Balaenoptera physalus*

The genealogy divides the 25 North Atlantic haplotypes into 2 major clades: we have termed these the 'Common North Atlantic' and the 'Western North Atlantic' (Fig. 2). The Western North Atlantic clade contains specimens from only the western part of the North Atlantic (Gulf of Maine, Gulf of Lawrence, Newfoundland and West Greenland).

The Common North Atlantic clade encompasses 72% (18 of 25) of the North Atlantic haplotypes accounting for 66% (83 of 125) of all North Atlantic specimens. The nucleotide diversity in the Common North Atlantic clade was estimated to be 0.010 (± 0.00040 SE); this is 3 times higher than the nucleotide diversity within the Western North Atlantic clade (0.0033 ± 0.00043). Although direct comparison is impossible it seems that the Common North Atlantic clade is equivalent to clade 'CD', and the Western North Atlantic clade to clade 'IJ' in the study by Baker et al. (1993).

DISCUSSION

Maternally directed philopatry maintains long-term segregation between feeding aggregations

Baker et al. (1993) compared 283 nucleotides of the mt control region in 90 specimens of which 34 were from the North Atlantic. They found no heterogeneity between the 3 feeding grounds Iceland (n=3), Newfoundland (n=12) and the Gulf of Maine (n=16). The present study is based upon a larger sample representing more North Atlantic feeding aggregations. The finding that a group of matrilineally related humpbacks (in what appears to be a panmictic population; Mattila et al. 1989, Clapham et al. 1993b) migrates only to feeding grounds within the western North Atlantic illustrates that behaviour influences population structure on an evolutionary time scale. There are no known geographical or environmental barriers, nor any physiological mechanisms, which would prevent western North Atlantic clade individuals from migrating to the feeding grounds off Iceland. Humpback whales have been observed to undertake migrations of almost 8000 km (Stone et al. 1990). Despite this propensity for long-distance movements, the sole mechanism of maternally directed foraging area philopatry appears to maintain segregation between the western North Atlantic and Icelandic feeding aggregations.

Lack of heterogeneity within the western North Atlantic

The genetic homogeneity within the 112 specimens representing 4 western North Atlantic feeding aggregations is in contrast to the results of studies based upon the identification of individual humpback whales, which have convincingly demonstrated an annual and maternally directed fidelity to specific feeding areas within the western North Atlantic (Katona & Beard 1990, Clapham et al. 1993a). Our results corresponds to those of Baker and coworkers (Baker et al. 1993), who found no heterogeneity within the western North Atlantic in their worldwide study which included 28 specimens from 2 feeding grounds, the Gulf of Maine and Newfoundland. This apparent genetic homogeneity could have several causes, the most immediate being a high degree of mixing within the western North Atlantic. However, this seems unlikely for 2 reasons. Although resightings of a number of individuals on more than one feeding ground demonstrate that movements between areas do indeed occur, such events represent a small fraction of all resightings, a fact which argues

for limited mixing (Stevick et al. 1993). Whether sightings of individuals in more than one feeding ground could represent whales *en route* to a final migratory destination, or the existence of very large feeding ranges is presently unknown. Our attempt to divide the Gulf of Maine specimens into residents and non-residents resulted in no heterogeneity, which would have been expected if what we define as 'non-residents' were primarily whales in transit to another feeding ground, although the sample sizes clearly are very small.

A more probable cause for the lack of heterogeneity in the western North Atlantic is the difference in the time scale of events between the life histories of individual whales and the accumulation of nucleotide differences in the mt genome. The greater part of the feeding grounds used by humpback whales in the western North Atlantic became accessible only after the last glaciation some 10 000 yr ago (Johnsen et al. 1992). Even if the population had been structured into discrete matrilineally directed feeding aggregations since the retreat of the ice, a much greater time scale would probably be required for evolutionary processes (such as accumulation of a significant number of divergent base substitutions and stochastic lineage extinction) to create similar structuring at the genetic level.

Gene flow between oceanic populations

The estimated genealogy placed all North Atlantic samples in 1 of 2 well-supported but divergent clades. The Common North Atlantic clade includes haplotypes which represent specimens from all North Atantic feeding aggregations, while the western North Atlantic clade is comprised of haplotypes found solely among from the western North Atlantic humpbacks. The co-existence of whales descending from 2 such divergent clades in the North Atlantic (each of which is more closely related to Antarctic haplotypes than the other North Atlantic clade) indicates that the North Atlantic Ocean has been populated by significant influxes of humpback whales on 2 occasions.

In the worldwide study conducted by Baker and co-workers (Baker et al. 1993), 6 migration events between oceans were postulated to explain the estimated genealogy. In the current study we found 1 Gulf of St. Lawrence and 2 Antarctic specimens with an identical haplotype. Thus this study demonstrates such migration between oceans, in this case from the North Atlantic to the Southern Ocean, since the haplotype of the 3 specimens is estimated to be within the Common North Atlantic clade.

High level of nucleotide diversity in North Atlantic humpback whales

The overall nucleotide diversity observed in the North Atlantic and Antarctic humpback whales found in this study is high relative to that of other cetacean species in the North Atlantic (Table 6) and is similar to that observed in some human populations (Vigilant et al. 1991). The estimate is similar to that found in previous studies (Baker et al. 1993).

The high level of diversity (relative to other cetaceans) can be explained relatively easily by the above hypothesis of 2 independent influxes of humpbacks into the North Atlantic. Thus the North Atlantic population of humpback whales is currently composed of whales originating from 2 different populations. Within each of the 2 clades, the Common North Atlantic and the Western North Atlantic, the estimate of nucleotide diversity is at a level similar to that found in other North Atlantic populations of baleen whales, except the heavily exploited northern right whale *Eubalaena glacialis*. The Antarctic Peninsula feeding aggregation is believed to contain humpback whales from more than one breeding ground (Dawbin 1966) and is therefore expected to have a high diversity, as indeed appears to be the case.

Table 6. Estimates of nucleotide diversity in other North Atlantic cetaceans. RFLP: restriction fragment length polymorphisms

Species	No. of analyzed specimens	Gene and method of analysis	Nucleotide diversity
Minke whale *Balaenoptera acutorostrata*	111	First 303 nucleotides of mt control region	0.0072[a]
Fin whale *Balaenoptera physalus*	103	First 288 nucleotides of mt control region	0.012[b]
Northern right whale *Eubalaena glacialis*	126	RFLP analysis of total mt genome	0.0026[c]
Humpback whale *Megaptera novaeangliae*	136	First 288 nucleotides of mt control region	0.026[d]
Narwhal *Monodon monoceros*	206	First 287 nucleotides of mt control region	0.0016[e]

[a] Heiberg et al. unpubl. data; [b] Bérubé et al. unpubl. data; [c] Schaeff et al. pers. comm.; [d] this study; [e] Palsbøll et al. unpubl. data

Acknowledgements. We thank our colleagues at the Department of Population Biology, the Center for Coastal Studies, the Greenland Fisheries Research Institute, the Mingan Island Cetacean Study, Inc., Haffrannstofnuninn and CIBIMA for support and assistance in numerous ways. Terje Härkonen and Mads-Peter Heide-Jørgensen are thanked for the Antarctic samples, Allied Whale for fast and efficient matching of important photos, as well as Drs Volker Loeschcke and Jotun Hein and 3 anonymous reviewers for comments on the manuscript. For technical assistance with the molecular analyses we thank Christina Færch-Jensen, Martine Bérubé and Hanne Jørgensen and, with the photos, Lisa Baraff. The work was supported by the involved institutions, as well as the University of Copenhagen, the Commission for Scientific Research in Greenland, the Greenland Home Rule, Åge V. Jensens Charity Foundation and World Wildlife Fund Denmark.

LITERATURE CITED

Amos, W., Hoelzel, A. R. (1991). Long-term preservation of whale skin for DNA analysis. In: Hoelzel, A. R. (ed.) Genetic ecology of whales and dolphins. Rep. int. Whal. Commn (spec. iss. 13): 99–104

Arnason, U., Gullberg, A., Widegren, B. (1991). The complete sequence of the mitochondrial DNA of the fin whale, *Balaenoptera physalus*. J. molec. Evol. 33: 556–568

Arnason, U., Gullberg, A., Widegren, B. (1993). Cetacean mitochondrial DNA control region: sequences of all extant baleen whales and two sperm whale species. Molec. Biol. Evol. 10: 960–970

Baker, C. S., Palumbi, S. R., Lambertsen, R. H., Weinrich, M. T., Calambokidis, J., O'Brien, S. J. (1990). Influence of seasonal migration on geographic distribution of mitochondrial DNA haplotypes in humpback whales. Nature 344: 238–240

Baker, C. S., Perry, A., Bannister, J. L., Weinrich, M. T., Abernethy, R. B., Calambodikis, J., Lien, J., Lambertsen, R. H., Urban Ramirez, J., Vasquez, O., Clapham, P. J., Alling, A., O'Brien, S. J., Palumbi, S. R. (1993). Abundant mitochondrial DNA variation and world-wide population structure in humpback whales. Proc. natl Acad. Sci. USA 90: 8239–8243

Brown, W. M. (1985). The mitochondrial genome of animals. In: MacIntyre, R. J. (ed.) Molecular evolutionary genetics. Plenum Press, New York, p. 95–130

Brown, W. M., George, M. Jr, Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. Proc. natl Acad. Sci. USA 76: 1967–1971

Cann, R. L., Brown, W. M., Wilson, A. C. (1984). Polymorphic sites and the mechanism of evolution in the human mitochondrial DNA. Genetics 106: 479–499

Christensen, I., Haug, T., Øien, N. (1992). Seasonal distribution, exploitation and present abundance of stocks of large baleen whales (Mysticeti) and sperm whales (*Physeter macrocephalus*) in Norwegian and adjacent waters. ICES J. mar. Sci. 49: 341–355

Clapham, P. J., Baraff, L. S., Carlson, C. A., Christian, M. A., Mattila, D. K., Mayo, C. A., Murphy, M. A., Pittman, S. (1993a). Seasonal occurrence and annual return of humpback whales in the southern Gulf of Maine. Can. J. Zool. 71: 440–443

Clapham, P. J., Mattila, D. K., Palsbøll, P. J. (1993b). High-latitude-area composition of humpback whale groups in Samana Bay: further evidence for panmixis in the North Atlantic population. Can. J. Zool. 71: 1065–1066

Clapham, P. J., Palsbøll, P. J., Mattila, D. K. (1993c). High-energy behaviors in humpback whales as a source of sloughed skin for molecular analysis. Mar. Mammal Sci. 9: 182–194

Clapham, P. J., Mayo, C. A. (1987). Reproduction and recruitment of individually identified humpback whales, *Megaptera novaeangliae*, observed in Massachusetts Bay, 1979–1985. Can. J. Zool. 65: 2853–2863

Dawbin, W. H. (1966). The seasonal migratory cycle of the humpback whale. In: Norris, K. S. (ed.) Whales, dolphins and porpoises. Univ. of California Press, Berkeley, p. 145–170

Felsenstein, J. (1993). PHYLIP 3.5c. Manual and program available from Joe Felsenstein. Dept of Genetics SK-50, University of Washington, Seattle

Gyllensten, U. B., Erlich, H. A. (1988). Generation of single-stranded DNA by the polymerase reaction and its application to direct sequencing of the *HLA-DQA* locus. Proc. natl Acad. Sci. USA 85: 7652–7656

Hoelzel, A. R., Hancock, J. M., Dover, G. A. (1991). Evolution of the cetacean mitochondrial control region. Molec. Biol. Evol. 8: 475–493

Hudson, R. R. (1992). Reply to Roff and Bentzen. Molec. Biol. Evol. 9: 969

Hudson, R. R., Boos, D. D., Kaplan, N. L. (1992). A statistical test for detecting geographic subdivision. Molec. Biol. Evol. 9: 138–151

Johnsen, S. J., Clausen, H. B., Dansgaard, W., Fuhrer, K., Gundestrup, N., Hammer, C. V., Iversen, P., Jouzel, J., Stauffer, B., Steffensen, J. P. (1992). Irregular glacial interstidials recorded in a new Greenland ice core. Nature 359: 311–313

Katona, S. K., Beard, J. A. (1990). Population size, migrations, and feeding aggregations of the humpback whale, *Megaptera novaeangliae*, in the western North Atlantic Ocean. In: Hammond, P. S., Mizpoch, S. A., Donovan, G. P. (eds.) Individual recognition of cetaceans: Use of photo-identification and other techniques to estimate population parameters. Rep. int. Whal. Commn (spec. iss. 12): 253–282

Katona, S. K., Whitehead, H. P. (1981). Identifying humpback whales using their natural markings. Polar Record 20: 439–444

Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. molec. Evol. 16: 111–120

Kishino, H., Hasegawa, M. (1989). Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from sequence data, and the branching order in Homonoidae. J. molec. Evol. 129: 170–179

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

Maniatis, T., Fritsch, E. F., Sambrock, J. (1982). Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor

Martin, A. R., Katona, S. K., Mattila, D., Hembree, D., Waters, T. D. (1984). Migration of humpback whales between the Caribbean and Iceland. J. Mammal. 65: 330–333

Mattila, D. K., Clapham, P. J., Katona, S. K., Stone, G. S. (1989). Population composition of humpback whales, *Megaptera novaeangliae*, on Silver Bank, 1984. Can. J. Zool. 67: 281–285

Nei, M. (1982). Evolution of human races at the gene level. In: Bonne-Tamir, B., Cohen, T., Goodman, R. M. (eds.) Human genetics, part A: the unfolding genome. Alan R. Liss, New York, p. 167–181

Nei, M., Li, W.-H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. natl Acad. Sci. USA 76: 5269–5273

Palsbøll, P. J., Larsen, F., Sigurd Hansen, E. (1991). Sampling of skin biopsies from free-ranging large cetaceans in West Greenland: development of new biopsy tips and bolt designs. Rep. int. Whal. Commn (spec. iss.) 13: 71–79

Palsbøll, P. J., Vader, A., Bakke, I., El-Gewely, M. R. (1992). Gender determination in cetaceans by the polymerase chain reaction. Can. J. Zool. 70: 2166–2170

Payne, P. M., Wiley, D. N., Young, S. B., Pittman, S., Clapham, P. J., Jossi, J. W. (1990). Recent fluctuations in the abundance of baleen whales in the southern Gulf of Maine in relation to changes in selected prey. Fish. Bull. U.S. 88: 687–696

Perkins, J., Balcomb, K. C., Nichols, G., Deavilla, M. (1984). Abundance and distribution of humpback whales, *Megaptera novaeangliae*, in West Greenland waters. Can. J. Fish. Aquat. Sci. 41: 533–536

Perkins, J., Whitehead, H. (1977). Observations on three species of baleen whales off northern Newfoundland and adjacent waters. J. Fish. Res. Bd Can. 34: 1436–1440

Roff, D. E., Bentzen, P. (1992). Detecting geographic subdivision: a comment on a paper by Hudson et al. Molec. Biol. Evol. 9: 968

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487–491

Saitou, N., Nei, M. (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. Molec. Biol. Evol. 4: 406–425

Sigurjonsson, J., Gunnlaugsson, T. (1990). Recent trends in abundance of blue, *Balaenoptera musculus*, and humpback whales, *Megaptera novaeangliae*, off west and southwest Iceland based on systematic sightings records with a note on the occurrence of other cetacean species. Rep. int. Whal. Commn 40: 537–551

Stevick, P. T., Allen, J. M., Katona, S. K. (1993). Dispersal patterns of humpback whales, *Megaptera novaeangliae*, in the western North Atlantic. 10th Conference on the Biology of Marine Mammals, Galvaston, Texas, 11–15 November. Marine Mammal Program, Texas A&M University, Galvaston, TX, p. 103

Stone, G. S., Florez-Gonzalez, L., Katona, S. (1990). Whale migration record. Nature 346: 705

Upholt, W. B., Dawid, I. B. (1977). Mapping of mitochondrial DNA of individual sheep and goats: rapid evolution in the D loop region. Cell 11: 571–583

Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K., Wilson, A. C. (1991). African populations and the evolution of human mitochondrial DNA. Science 253: 1503–1507

This article was submitted to the editor

Manuscript first received: June 9, 1994
Revised version accepted: August 19, 1994