GENETIC CHARACTERIZATION OF THE WESTERN NORTH ATLANTIC RIGHT WHALE: YEAR III

FINAL REPORT
for the period
September 1997 to September 1998

Contract Number 40EANF700069

December 1998

Submitted by:

Moira Brown[1], Sobia Malik[2,3], Ruth Waldick[2], Cindy Tomlinson[2], and Brad White[2]

[1]Center for Coastal Studies, PO Box 1036, Provincetown, MA 02657 and East Coast Ecosystems, Freeport, NS B0V 1B0

[2]McMaster University, 1280 Main Street West, Hamilton, ON L5S 4K1

[3]Centre for Forensic Sciences, 25 Grosvenor St. Toronto, ON M7A 2G8

I. Executive Summary

A total of 289 individual right whales (*Eubalaena glacialis*) have had their mitochondrial DNA (mtDNA) haplotyped directly or their haplotypes inferred from their matrilineal pedigree. Four males first seen as adults prior to 1985 are the only animals identified with haplotype E and therefore this matriline may already have effectively disappeared. The gender of 315 of 388 known whales has been determined by a combination of sighting records and direct molecular sexing. Two hundred animals have been genotyped at nine microsatellite loci giving a probability of two animals sharing the same profile of less than one in 200,000. Two animals apparently did share the same profile but this discrepancy was resolved by reviewing the biopsy records. A few other discrepancies have been identified using multiple samples of the same animal and mother calf pairs. However there is a very high congruence between photo-identification and DNA profiles. A total of 20 paternities have been assigned to 25 calves. The allele frequencies at the microsatellite loci are very different to those in *E. australis* and are consistent with total reproductive isolation. In addition, none of the 5 mitochondrial haplotypes were found in *E. australis*. The sequence differences suggest a separation of the two species three to nine million years ago. Genetic diversity is much lower in *E. glacialis* than in *E. australis* for both mtDNA and microsatellite loci and is consistent with a small effective population size for about 400 years.

Between 1980 and 1996, 60% of calves were brought to the Bay of Fundy nursery. Differences in frequencies of mitochondrial haplotypes were observed in mothers that always used the Bay of Fundy nursery (*Bof all*) compared to those mothers who never used the Bay of Fundy (*Bof none*) as a nursery. There is a suggestion that this site fidelity is based on cultural transmission from mother to daughter. Haplotype C shows a gender bias (13 females: 26 males) and is the one found primarily in *Bof none* mothers. The under representation of females with genotype C appears to be due to the fact that they are rarely seen in areas where biopsy sampling has occurred. The reproductive performance of *Bof all* and *Bof none* mothers appears different but this may be due to the sighting and sampling bias. This reinforces the need for more sighting and biopsy effort in the southeast US calving area where all mothers appear. Due to the sighting and sampling bias against *Bof none* mothers, it may be advisable to test the use of this as a sighting index similar to that of the offshore index.

We are analyzing the matrilineal pedigrees to assess the rate of genome loss from generation to generation. There are 82 females (of 156 in the catalogue) whose mothers are unknown and can be considered the potential founders. Some of the genetic diversity in these potential founders has already been lost due to presumed dead females either not reproducing or producing only one calf. Paternity analyses suggest that the variance in male reproductive success is lower than that for females suggesting a more rapid loss of genomic representation from females than males. We are now able to connect some of the pedigrees by identifying fathers from their DNA profiles. This, in conjunction with identifying newly sighted non-calf whales as unidentified calves from earlier years, will improve our estimates of rates of loss of genetic diversity and the effective population size.

II. Deliverables

1. Methods

a) DNA Extraction

Samples collected prior to May 1998 were extracted using standard phenol/chloroform methods. Some of these samples presented amplification problems and contaminants, which can be carried over and were suspected to be the problem. A new method of cleaning DNA (Qiagen) which involves binding DNA to silica was assessed. Samples which had previously not amplified with mitochondrial or microsatellite primers showed good amplification products after Qiagen cleaning. Samples which, in the past, were unable to be genotyped were put through this process. All new tissue samples will be extracted with Quiagen to avoid further amplification problems.

b) Gender

Molecular sexing of individuals was done by polymerase chain reaction (pcr) amplification of the ZFY/ZFX loci on the sex chromosomes, followed by *TaqI* restriction enzyme digestion. Electrophoresis through an agarose gel results in bands visualized by ethidium bromide staining showing 2 bands in males and one band in females.

c) Matriline Haplotyping

The mitochondrial control region was amplified through pcr, and run on non-denaturing polyacrylamide gels to assess single stranded conformation polymorphism (SSCP) patterns. In addition, the first 550bp in the mitochondrial control region were sequenced for new 43 individuals (in addition to the 12 sequenced for 500bp in 1996/1997) to confirm SSCP haplotypes.

d) Microsatellite genotyping

Using standard cloning and screening methodologies, 34 microsatellite regions identified from the genome of the North Atlantic right whale and primer sets were established for 18 loci. Genotyping of individuals was conducted using 25-50 ng of DNA per PCR and the products were size separated using gel electrophoresis (polyacrylamide). Eleven of the 34 loci identified in the right whale were found to be polymorphic with an average of 3 alleles per locus. Genetic profiles have been established for 200 individual right whales at 9 microsatellite loci. Unique profiles were established for 199 of the 200 samples. Presently, genetic profiles are being developed for pairs of samples collected from the same individual on separate biopsy-sampling days. The source of discrepancies in genetic profiles for samples from the same animal will be identified and used to categorize the rate of sample classification. Mismatching profiles have been identified among duplicate samples and between mother-calf pairs. The source of three mismatches have been identified. Both photo-identification and laboratory sample misclassifications have been identified. The protocol for identifying the source of errors is: to recheck genotyping records and lab codes within the lab, identify the sampling date (Field tag), re-evaluate biopsy-sampling records from the collection date, re-evaluate the photographs of the sampled whale. This protocol for identifying sample misidentification was implemented as a standard procedure at the Right Whale Consortium meetings in October, 21, 1998.

e) Major Histocompatibility Complex Typing

The amplification reaction using primers from the preliminary study of DQB in the right whale (Murray 1997) was optimized. The resulting products are being analyzed via SSCP and existing right whale clones of specific alleles will enable genotypes to be determined. Any individuals with new alleles will be cloned and sequenced.

2. Data
a) Mitochondrial haplotypes and gender

The total number of whales sampled, sexed and mtDNA haplotyped to date are presented in Tables 1-5. One sample did not amplify with mitochondrial primers but the haplotype could be inferred through the matrilineal pedigree. A further two samples presented problems with sexing. All three of these had been Qiagen cleaned.

Of the 76 whales which have been sexed both in the field and in the lab (Table 2), only one apparent mismatch occurred which was resolved after accepting a field misidentification. The five individuals with multiple biopsy samples which differ in haplotype will have their microsatellite profiles assessed to see if one of the samples matches another whale in our database.

Of the 388 identified whales in the catalogue, 239 have been genotyped directly with a further 50 having genotypes inferred from matrilineal pedigrees (Table 3). We have directly sequenced 58 individuals to confirm that the method of SSCP is detecting all genotypes. There are ten variable positions over 550bp which define five matrilines, with 2 to 6bp differing among the 5 matrilineal genotypes.

The four genotype 'E' individuals are all males first sighted before 1985. In the years since then, no sampled mother or calf has been found with this genotype. This matriline may therefore be extinct as only females pass on the mitochondrial genome.

b) Site Fidelity and Genetic Sub-structure

Between 1980 and 1996, 60% of calves were brought to the Bay of Fundy nursery (BoF) grounds. We found that some mothers bring all their calves to the BoF nursery (*BoF all*), some mothers are never seen in the BoF with a calf (*BoF none*), while another group of females bring some of their calves to the BoF, but not all (*BoF some*).

From an analysis of females who had had two or more calves, it was shown that there is significant site fidelity to a particular nursery (Malik *et. al.* subm. a.). Mothers used a particular nursery significantly more often than would be expected given the proportion of calves brought to the different nurseries for the population as a whole. Multigenerational pedigrees were used to analyze the possibility of cultural transmission of nursery use from mothers to daughters. Eight *BoF all* pedigrees included 12 second generation mothers. Ten of these 12 have used the BoF nursery grounds indicating a trend for the cultural transmission of nursery use (Malik *et. al.* subm. a.). There were no multigenerational pedigrees available for analysis in *the BoF none* group. We will be examining whether there is any evidence of sub-structuring at microsatellite loci.

c) Implications of Genetic Sub-structuring

Mothers and their calves were examined to assess the effect of this philopatric behavior on the genetic sub-structuring of the population. The frequencies of the genotypes differed significantly between the *BoF all* group, and the *BoF none* group (Fig. 1). Genotype 'A' is relatively uncommon in *BoF none* individuals while genotype 'C' is relatively more common. When looking at the frequency of genotypes distributed among the sexes, it is evident that the ratio of genotype 'C' differs from the 50:50 sex ratio of the population as a whole (Table 4). There are half as many females as males with this genotype. As genotype C is relatively most common in the *BoF none* group, we can infer that we are missing a number of *BoF none* females. There appears to have been a sampling bias against individuals from the BoF none group as almost 40% of the living *BoF none* population has not been sampled, as opposed to 15% or 18% for individuals in the *BoF some* or *all* groups (Table 5). This has resulted in close to 50% of the known *BoF none* pedigrees not being sampled (Table 6). This reinforces the need for more biopsy sampling in the south east calving area where all mothers appear. This lack of available samples from mothers and/or calves not appearing in the BoF may be indicative of a similar lack of sighting information, which could have implications for our knowledge of the demographics of the group not using the BoF. There was a significant difference in the average number of calves per female with *BoF none* females producing an average of 1.77 calves (s.d. 0.95) and the *BoF all + some* females producing 2.53 calves (s.d. 1.48) (Fig 2., t test p<0.05). It is possible that there are fewer *none* females with 4 or more calves because of limited sighting effort in the southeast U.S. until recently.

The average inter-birth interval of the three groups at first appears to be higher for the *BoF none* group (Table 7). However, two females had inter-birth intervals of 13 years, after having been presumed dead (Fig. 3). When these 2 females are excluded, there is no obvious difference between the inter-birth intervals of the 3 groups, but the lack of sighting information for the two resurrected females points to a possible sighting bias against mothers not appearing in the BoF.

Looking at the distribution of maximum number of years not sighted for the mothers of the 3 groups there are obvious differences (Fig. 4). The *BoF all* mothers predominate among those seen every year or two, while the *BoF none* females are hardly ever seen one or two years apart and a large proportion of them are presumed to be dead. Given the example of two 13 year inter-birth intervals, and the proportion of females resurrected for the *BoF none* group, it may be reasonable to expect that some of the females still presumed dead may in fact be alive. This will only be known, however, if sighting effort in the southeast remains strong.

d) Comparisons of North and South Atlantic Genetic Variability
i) Mitochondrial haplotypes

Five mtDNA control region haplotypes were found in 180 North Atlantic individuals compared to 10 haplotypes in 16 Sounth Atlantic individuals (Malik *et. al.* subm. b.). None of the five *E. glacialis* haplotypes were found in *E. australis* confirming the reproductive isolation of the two species. Nucleotide diversity for North Atlantic haplotypes was 0.6% while that for South Atlantic haplotypes was 2.0%. The average haplotypic diversity was 0.87 in the North Atlantic

population and 0.96 in the South Atlantic samples. Genetic divergence between the North and South Atlantic populations was estimated to have occurred 3.0-9.0 mya. The long bottleneck experienced by the North Atlantic right whale has apparently resulted in a reduction in both haplotypic and nucleotide diversity when compared to the South Atlantic species.

ii) Microsatellites

The levels of polymorphism and allelic diversity were compared in the North and South Atlantic right whales at 18 of the 34 *Eubalaena*-derived microsatellite loci and 6 additional, cross-species loci (Tables 8 and 9). Forty-six percent (11/24) of the microsatellite regions were polymorphic in *E. glacialis* compared to 91% (22/24) in *E. australis*. The allelic diversity (A) was also significantly lower in *E. glacialis* than in *E. australis* (A=3 and 7, respectively). The exclusion probability (PE) for parentage analysis was determined to be significantly higher in the South Atlantic right whale than in for the North Atlantic right whale population. The PE for the North and South Atlantic right whale populations is 62% vs. 98% when neither parent is known and 97% vs. 99.9% when one parent is known, respectively.

e) Completion of Paternity Analysis

Paternity assignments are necessary to identify the mating system in the right whale. When DNA is available for both mother and calf, the probability of excluding non-parental males is 97% (based on genetic profiles using the nine polymorphic microsatellite loci). The low genetic variability among males, however, is preventing us from discriminating among father-son pairs and pairs of male siblings. Biopsy samples available to January, 1997 provide DNA for approximately 50% of the mother-calf pairs (Tables 10 and 11). Paternity analysis conducted for 31 calves born between 1991 and 1996 using the pool of an estimated 61-65% if the living adult males indicates that the variance in male reproductive success is low in this population . Putative fathers were identified for 25 of the 31 calves from the available pool of biopsy-sampled males. Using Maximum-likelihood calculations based on allele frequencies at microsatellite loci, these 25 calves were attributed to 20 different males. None of the males assigned multiple paternities (N=3) matched as fathers to multiple calves within a single cohort year. Once additional microsatellite loci are typed (n=5) and functional gene loci are added to the genetic profiles, the primary limitation to completing paternity analysis will be limited to the proportion of adult males that have been biopsy-sampled (N=61-65%) and the proportion of female-calf pairs for which DNA is available (Table 11).

f). Population History

The North Atlantic right whale has low genetic diversity at minisatellite and microsatellite loci compared to the South Atlantic population and most other large cetacean species (Table 9). The magnitude of the difference in genetic diversity is too great to be accounted for by a recent bottleneck (<200 years ago). The low allelic diversity within these microsatellite regions is reflected in the high degree of allele sharing observed among individual right whales. Most of the low genetic variability observed in the North Atlantic right whale appears to be a consequence of a historically small effective population size. Preliminary screening of a functional gene in the MHC and at a microsatellite loci associated with a functional gene (Insulin-like growth factor-1) suggest that levels of genetic variability may be higher within loci that are subject to natural selection.

g). Major Histocompatibility Complex (MHC) Analyses

A preliminary study of DQB variation in right whales found 9 alleles among 16 individuals, but with up to-4 alleles per individual at least two loci are involved and the locus designation was not possible (Murray 1997). We are using the more complete pedigrees, including those where both parents are known or where there are multiple sets of half-sibs, to determine the locus affiliation of these alleles. Clones will be used as standards for genotyping. Once the locus affiliation is determined the rest of the population will be screened to a) assess any association between MHC haplotype and reproductive and skin lesions and b) improve the discriminating power for paternity analyses. Microsatellite primers associated with MHC in cattle and/or sheep are being tested for their possible application in the right whale. Three primer sets have been tested but optimization procedures are ongoing. In addition, DRB primers used with beluga, cattle and/or humans will be assessed.

h). Incorporation into NEA database

All new sexes and mtDNA haplotypes have been communicated to the database curators at NEA. They have also been brought up to date with the index used for classifying females and their offspring as *BoF all, none or some*, and the possible implications of the use of this category in demographic modelling using sighting information. The relatedness and parentage data will be submitted to the NEA database on an annual basis.

i) Additional Field Work/Analyses Required

From the distribution of haplotypes it was apparent that a number of mothers, which do not use the BoF as a nursery, and their calves, have not been sampled. The best location for sampling of these individuals would be the southeast calving grounds. If all mothers and calves appearing there could be sampled it is expected that the number of 'new' individuals would be reduced, as a greater majority of them could be identified from their genetic profiles to known calving events, especially where the calf was not seen in the fall when it's callosity patterns become established. This would also aid in paternity analyses as it would increase the percentage of males in the population that could be included in the analyses. Paternity analyses will be conducted for the more recent calves, and the number of loci used in these analyses will be expanded. These loci include a few microsatellites, which have not yet been examined as well as the DQB locus and possibly, the DRB and/or the three MHC microsatellites. The inclusion of these loci will improve our ability to discriminate among potential fathers.

Table 1. Number of individuals in the database which have been sampled, sexed or haplotyped.

	Totals in Database	Sampled	Sexed	Haplotyped
Identified	388	245	191	239*
Unidentified	25	25	15	23
Totals	413	271	206	262

*In addition, five individuals had discrepant haplotypes where two or more biopsy samples per individual had differing haplotypes.

Table 2. Total numbers of whales sexed in the field and/or by molecular DNA techniques.

Method of Sexing	No. Males	No. Females	No. Unknown	Totals
Field	50	74		124
Molecular	83	32		115
Field and Molecular	26	50		76
Unsexed			73	73
Totals	159	156	73	388

Table 3. Numbers of individuals for which their mtDNA haplotype was inferred or determined by molecular genotyping.

Genotype	Inferred	Genotyped	Total
A	21	66	87
B	5	26	31
C	4	36	40
D	20	107	127
E	0	4	4
Total	50	239	289

Table 4. Distribution of mtDNA genotypes among the sexes

Genotypes	Male	Female	Unknown	Totals
A	43	38	6	87
B	15	14	2	31
C	26	13	1	440
D	55	65	7	127
E	4	0	0	4
Totals	143	130	16	289

Table 5. Numbers of living individuals categorized by nursery use, and the numbers within each which have been biopsy darted.

Nursery Use	Total Living	Darted	% Not Darted
BoF all	85	70	17.6
BoF some	52	44	15.4
BoF none	34	21	38.2

Table 6. The total number of pedigrees for each of the nursery use categories and the number of pedigrees where at least one individual has been biopsy darted.

Nursery Use	Number of Pedigrees	Number Darted	% Not Darted
BoF all	25	22	12.0
BoF some	22	21	4.6
BoF none	26	14	46.2
	73	58	

Table 7. The average inter-birth intervals for mothers of different nursery use categories. *BoF none* (a) includes and (b) excludes two females who were resurrected with an inter-birth interval of 13 years.

	BoFall	*BoFsome*	*BoFnone*(a)	*BoFnone*(b)
average	4.34	4.09	5.20	3.90
st.dev.	0.91	0.90	3.43	1.01
n	16	22	14	12

Table 8. Size and frequency of microsatellite alleles at nine polymorphic loci in *Eubalaena glacialis* (n=180-200). Sizes of alleles are expressed as the number of base pairs with the allele frequencies shown beneath.

Locus	Allele size frequency								
	119	135	137						
	0.005	0.645	0.350						
1Pm*									
	199	201	203						
	0.410	0.529	0.071						
37Mn**									
	151	153	155	157	159	161	163	165	167
	0.365	0.0182	0.0651	0.0573	0.396	0.140	0.0391	0.219	0.0078
IgF***									
	191	193	195	197	201				
	0.045	0.0185	0.688	0.214	0.034				
RW18									
	163	165							
	0.298	0.702							
RW26									
	122	128	130	132	136				
	0.166	0.0449	0.668	0.0900	0.0309				
RW31									
	122	128	130	132	134	136	138	140	142
	0.475	0.0027	0.0516	0.0217	0.0109	0.128	0.188	0.0299	0.038
RW34									
	102	108	110	112	114	124			
	0.0108	0.0649	0.130	0.743	0.0486	0.0027			
RW48									
	96	98	100	102					
	0.442	0.293	0.050	0.215					
RW417									

Table 9 Allelic diversity, locus-specific heterozygosity (H), and an evaluation of the informativeness of microsatellite loci, expressed as the exclusion probabilities and polymorphism information content (PIC) values for *E. glacialis*.

Locus	Number of alleles	Observed Heterozygosity	Expected Hetero
1Pm	3	0.437	0.463
37Mn	4	0.576	0.559
IgF	10	0.786	.0766
RW18	5	0.471	0.479
RW26	2	0.451	0.419
RW31	5	0.511	0.516
RW34	11	0.668	0.717
RW48	6	0.432	0.425
RW417	5	0.718	0.672

Table 10. Yearly summary of the number of newly sighted whales and the proportion of missed calvings.

Year	# of New sightings (Excluding cows and young-of-the-year)	Cumulative number of observed newborns that were never photo-identified
1985-1990	69	8
1991-1996	20	26
TOTAL	**89**	**35**

Table 11. Summary of the availability of DNA samples from mother-calf pairs and adult males that can be used for parentage analyses (to Jan. 1997).

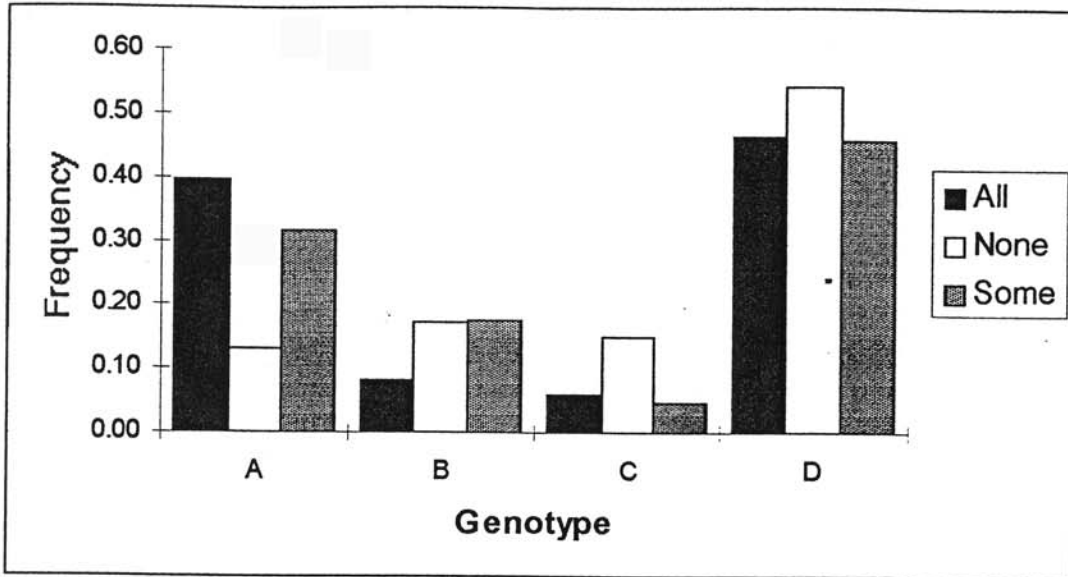| Year	# of known mother-calf pairs	# of mother-calf pairs available for genetic testing	Proportion of unsampled, non-BoF calves	Proportion of unsampled moms that are BoF-nones
1985-1990	70	48	55	77
1991	17	8	56	23
1992	12	5	43	29
1993	6	3	33	33
1994	9	2	57	14
1995	7	3	100	75
1996	21	8	46	31
TOTAL	**141**	**77**		

Fig. 1. Genotype frequency distributions among mothers who brought all, none or some of their calves to the Bay of Fundy and their calves ($p = 0.0012$ for all vs. none).
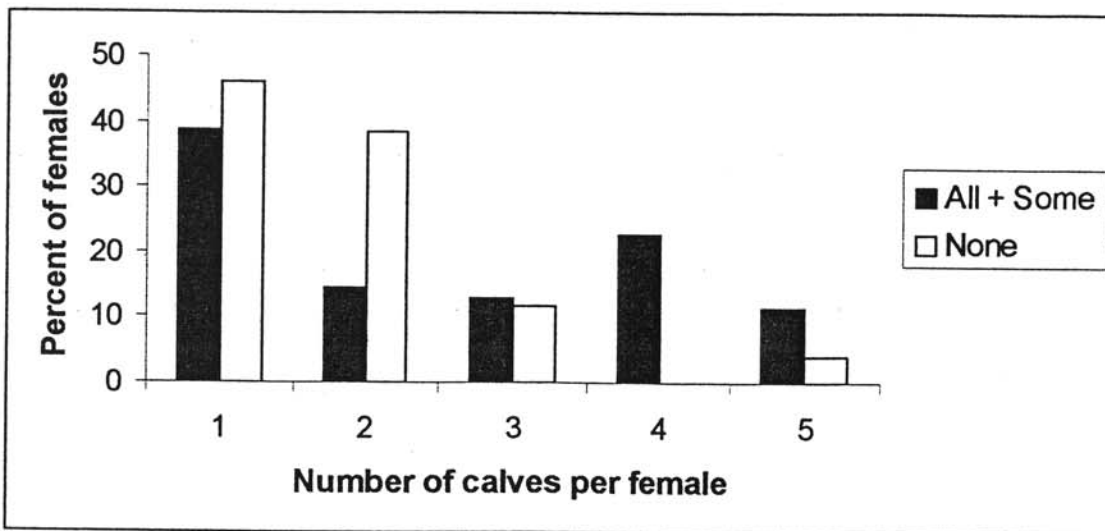

Fig. 2. The number of calves per female for mothers never bringing a calf to the BoF and mothers who sometimes or always bring their calves to the BoF.
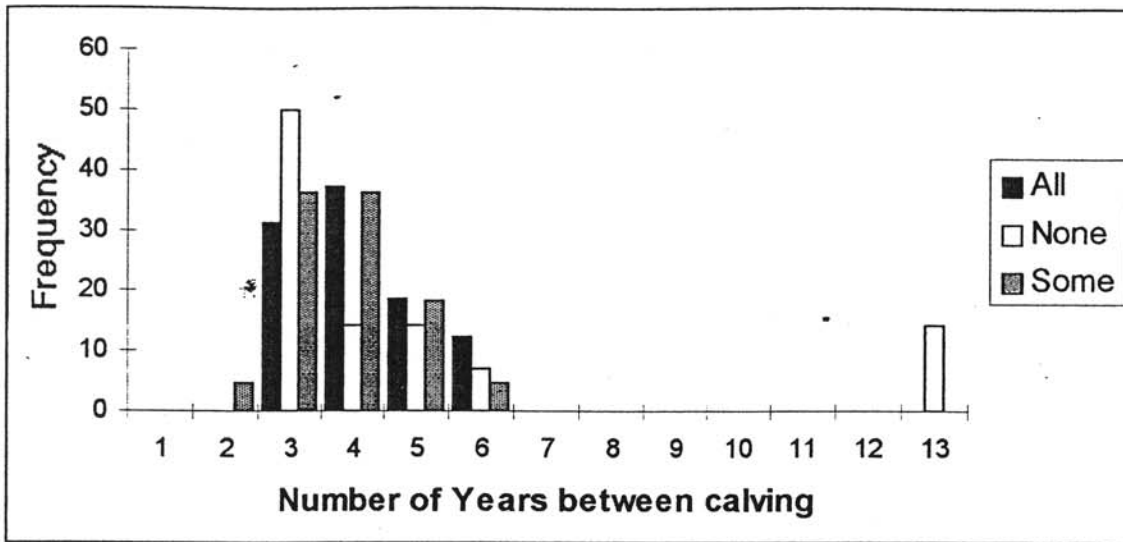
Fig. 3. Interbirth intervals for mothers who brought all, none or some of their calves to the BoF.
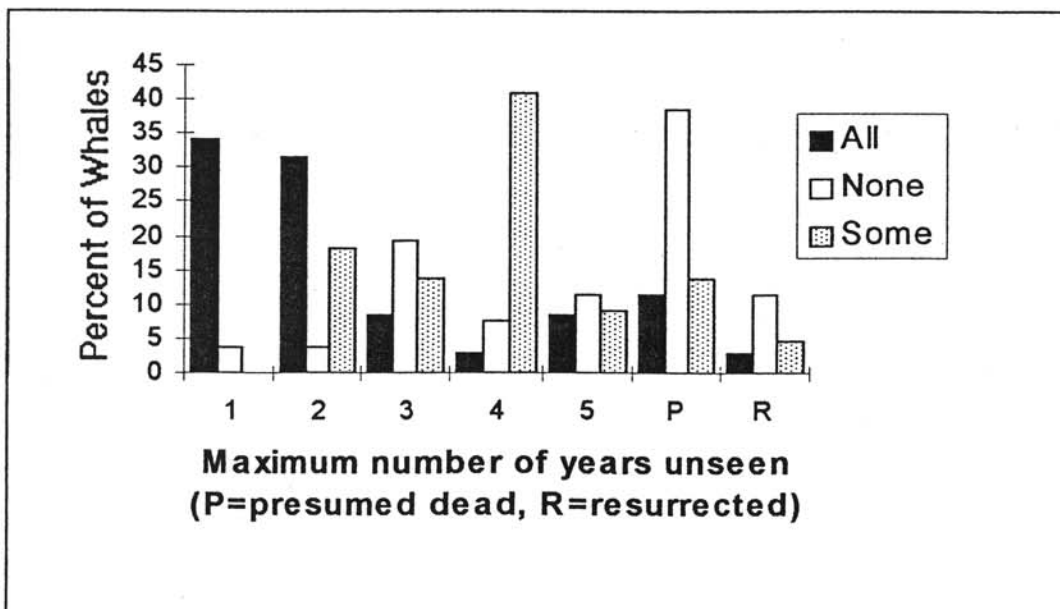

Fig. 4. Maximum number of years between sightings of females for the different nursery use categories.

References

Malik, S., Brown, M.W., Kraus, S.D. Knowlton, A., Hamilton, P. and White, B.N. Assessment of genetic structuring and habitat philopatry in the North Atlantic right whale (*Eubalaena glacialis*). Submitted (a) to Canadian Journal of Zoology.

Malik, S., Brown, M.W., Kraus, S.D. and White, B.N. Analysis of mitochondrial DNA diversity within and between north and south Atlantic right whales. Submitted (b) to Marine Mammal Science.

Murray, W.B. 1997. Major Histocompatibility complex class II sequence variation in cetaceans: DQB and DRB variation in beluga (*Delphinapterus leucas*) and DQB variation in North Atlantic right whales (*Eubalaena glacialis*). PhD thesis, McMaster University, Hamilton, Ontario, Canada.