

Genetic study on close relatedness of common minke whale *Balaenoptera acutorostrata* in the Central and Northeast Atlantic.

STEFANIA BENÓNÍSDÓTTIR^{1,2}, HANS J. SKAUG², KEVIN GLOVER³, BJARKI ÞÓR ELVARSSON^{1,4}, GÍSLI A. VIKINGSSON¹ AND CHRISTOPHE PAMPOULIE¹

¹Marine Research Institute, Skúlagötu 4, 121 Reykjavík, Iceland: chrisp@hafro.is

²University of Bergen, Faculty of mathematics and Natural Sciences, Norway.

³Institute of Marine Research, Bergen, Norway.

⁴Institute of Science, University of Iceland, Dunhagi 5, 107 Reykjavík, Iceland.

ABSTRACT

One of the main problems to the genetic structure of highly vagile pelagic species such as the cetacean lies in their almost continuous habitat, the lack of information on the location of their breeding ground and their potential connectivity. Most of the available data have often been collected at feeding grounds which might be composed of a mixture of several distinct populations. Thus, traditional population genetic studies of common minke whale in North Atlantic have not been able to considerably improve the knowledge about stock structure of this species. Alternative statistical approaches such as the close-kin relatedness approach can however, shed more light into their migration pattern and relationship between individuals from different feeding location. A common alternative statistic used for such approach is the LOD score, which is easily calculated from a pair of DNA-profiles. In the present paper, we present analyses based on a newly developed algorithm using LOD score. A total of 817 DNA-profiles at 16 microsatellite loci were used to look for relatedness among individuals within the North Atlantic. We found 124 significant related individuals, including 40 among mother-foetuses, 12 among foetuses of different females and among foetuses and other whales than their mother. No trio matches between mother-foetus pairs and an alleged father were found. A total of 23 relatedness pairwise comparisons were significant between Iceland and Norway, spanning from few years to few generation suggesting that within/among generation, gene flow among separated by the current IWC area boundaries is quite high, therefore confirming previous genetic analyses (SC/F13/SP17).

INTRODUCTION

Traditionally, the management of the North Atlantic common minke whales has been based on four geographical subdivisions divided by the International Whaling Commission (Donovan, 1991), namely the Canadian East coast stock, the West Greenland stock, the Central stock (Iceland) and the Northeastern stock (Norway). These management regions have been primarily established through studies based on catch statistics, biological characteristics and tagging. Lately, genetic studies tend to confirm the established subdivisions. Genetic variation among potential (sub)populations of minke whale have been investigated using isozyme (Daníelsdóttir *et al.*, 1992; 1995) and human α -globin 3'HVR (Árnason and Spillaert, 1991). Both types of markers revealed a significant genetic differentiation among samples collected at West Greenland, Icelandic and Norwegian feeding grounds; hence suggesting the potential existence of at least three different populations. A recent study was carried out on a sample of 306 individuals using microsatellite loci and the D-loop in the mtDNA, to determine the population structure of the common minke whale in Greenland, Central, NE Atlantic and North Sea (Andersen *et al.*, 2003). This study confirmed the genetic distinctiveness of four subpopulations: 1) West Greenland, 2) Central North Atlantic-East Greenland-Jan Mayen area, 3) NE Atlantic (Svalbard, The Barents Sea and the North Sea), and 4) North Sea. The authors suggested that the common minke whales within different ecological regions represented genetically distinct subpopulations. The most recent intensive genetic study performed on common minke whale however tend to contradict previous studies as a suite of 16 microsatellite loci and the mtDNA control region failed to reveal any genetic pattern related to geographical areas of management or any detection of different genetic groups in the data (Pampoulie *et al.*, 2008). A more recent study however detected cryptic populations across the North Atlantic (Anderwalt *et al.*, 2011). One of the main challenges to study the possible genetic structure of highly vagile pelagic species, such as the

cetacean, lies in their almost continuous habitat, the lack of information on the location of their breeding ground and their potential connectivity. Most of the available data have often been collected at feeding grounds which might be composed of a mixture of several distinct populations, and therefore indirect methods of gene flow offer an alternative approach to assess genetic structure and gene flow among potential cetacean sub-populations. As such, the use of microsatellite loci to study individual relatedness within and among populations remains an alternative approach to investigate potential reproductive success differences among group of individual and/or migration patterns of individual common minke whale (Skaug and Oien, 2005). A common alternative statistic used for a given hypothesis about relatedness is the LOD score, which is easily calculated from a pair of DNA-profiles (Skaug *et al.*, 2010). In the present paper, we analysed 817 DNA-profiles at 16 microsatellite loci and look for relatedness among individuals within the North Atlantic.

MATERIALS AND METHODS

All the genetic protocols used during this study were previously described in Pampoulie *et al.* (2008) and followed the IWC guidelines (see Tiedemann *et al.* 2012. SC/64/SD4 for an update).

Relatedness analyses were performed using the genetic profiles previously analyzed in SC/F13/SP17. In addition, 50 mothers-foetus pairs were collected from 2003-2007 and the foetus DNA profiles were added to the original dataset presented in SC/F13/SP17. Out of these 50 foetuses, five contained missing alleles at some loci and/or some non-common alleles with the mother (mutation) and were not considered for further analyzes. In addition, Norwegian minke whales collected from 2004-2006 and encompassed in the Norwegian minke whale DNA register (Glover *et al.*, 2012), with validated genotyping error rates (Haaland *et al.*, 2011) were also added in the dataset (with permission of the Norwegian DNA-registry managers), a total of 272 individuals. Because the techniques highly depends on matching of alleles, all individuals with missing alleles at certain loci were deleted from the dataset prior to the analyzes.

Detection of relatives was done by computing pairwise LOD scores for 817 individuals in the sample, for each relatedness of interest. If D_i and D_j were the genetic profiles of individuals i and j then the LOD score for their relatedness was calculated according to Skaug *et al.* (2010):

$$LOD_{i,j} = \log\left(\frac{P(D_i, D_j | H_1 : \text{related})}{P(D_i, D_j | H_0 : \text{unrelated})}\right)$$

In this paper, the relatedness of interest was parent-offspring relations. The parent-offspring LOD score is a very strict measure of relatedness and unless individuals i and j have at least one matching allele at every single locus, their parent-offspring LOD score will be infinitely negative. A high LOD score indicates relatedness but that entails an issue of what should be considered to be high enough. That issue was accounted for by developing a method to estimate a single p -value with each LOD score. The p -values can then be attained via simulation (Benónísdóttir, 2012). A total number of 265 unrelated individuals were simulated with the same population allele frequencies as the one estimated from the available dataset and then their pairwise LOD scores were computed. This procedure was replicated 80 times. The p -values were computed by comparing the original LOD scores with the simulated ones but $p_{i,j}$ can be described as the probability of attaining as extreme or more extreme LOD score than $LOD_{i,j}$ just by chance. All computations and simulations were done by using the open source program R (R Core Team 2012), but the codes written are available upon request to the authors of this paper.

Relatedness was tested for every possible pair in the dataset. The high number of pairwise comparisons raised another statistical issue, the problem of multiple testing, which was addressed by the FDR procedure, e.g. the false discovery rate. The FDR procedure has been suggested to be more appropriate than the Bonferroni correction for example, as it takes the number of erroneous false discoveries of relatedness into account instead of only the question of whether any error was made (Benjamini and Hochberg, 1995).

RESULTS

Out of the 333,336 ($n(n-1)/2$) pairwise-relatedness comparisons possible, 670 exhibited a high LOD score larger than zero (suggesting relatedness). After FDR corrections for multiple tests, only 124 remained significant (Table 1). The correction procedure (FDR) was set-up to capture most of the mother-foetus pairs, e.g. the q values was set-up at 0.52 and only 5 mother-foetus pairs out of 45 were not detected (mainly because of too many common allele types) between the mother and the foetus, resulting in a low LOD score).

Among the 124 significant pairwise-relatedness comparisons, 40 were among mother-foetus pairs and 12 involved foetuses of different females and foetus and other females than their mother. No trio matches mother-foetus pairs – potential alleged father were found.

Table 1. Number of significant relatedness pairwise comparisons among minke whales from different geographical regions, after FDR correction procedure at $q = 0.52$. Mother-foetus pairs were removed from these numbers ($n = 40$) as well as relationship among foetuses and among foetus and other whales than their mother ($n = 12$). Note that numbers of individuals related can be assessed by multiplying the number of pairs per 2, as none of the individuals were observed twice in the comparisons.

	ICELAND	NORWAY	NORTH SEA	GREENLAND	SPITSBERGEN	BARENTS SEA
ICELAND	11					
NORWAY	23	14				
NORTH SEA	1	0	0			
GREENLAND	0	4	0	4		
SPITBERGEN	7	6	0	0	1	
BARENTS SEA	1	1	0	0	0	0

DISCUSSION

In the face of inconclusive genetic pattern using classical genetic statistics, other means of using genetic markers can be address to give better insight into the biology of evasive species such as baleen whales. The LOD score, implemented to study the relatedness among individuals, has been proposed as an alternative approach to genetic to investigate potential migration patterns among and/or within families of marine organisms (Skaug *et al.*, 2010). Here, by applying a newly developed algorithm (Benónísdóttir, 2012), we investigated the relatedness of individual minke whales across the North Atlantic, as well as potential matches between mother-foetuses pairs and alleged fathers.

The first interesting result was that, despite the large sample size of Icelandic and Norwegian dataset, we could not detect any trio-matches of a mother-foetus pair and an alleged father. Interestingly, some relatedness was found among foetuses of different females, suggesting that they might have shared a common alleged father that was not contained in the dataset.

The largest amount of significant related-pairs comparisons were observed among Norway and Iceland, where a total of 46 individuals were found to be related. Additional biological information might give more insight into the type of relationships detected, as only few of them could be looked into at present. Few examples of relationships can be examined further, e.g. for example:

- A female of 17 years old captured in Iceland in 2004 was related to a female captured in Norway in 2006

- A female of 20 years old captured in Iceland in 2007 was related to a male captured in Norway in 2004
- A male of 23 years old captured in 2006 in Iceland was related to a male of 19 years old captured in 2007 at the same location (potential brothers)
- A male of 22 years old captured in 2007 in Iceland was related to a male captured in 1984 at the same location (potential match father-son)
- A female of 14 years old captured in 2007 in Iceland was related to a male captured in the same location in 1982 (potential grand father-grand daughter)

Older relationships among Icelandic and Norwegian minke whales were detected such as the male captured in 2004 in Norwegian waters which was related to a female captured in 1983 in Iceland.

Although a relatively low number of individuals were found to be related when comparisons involved samples collected at Greenland, Spitsbergen, North Sea and the Barents Sea geographical regions, the detected related pairs might suggest that relatedness matches might be more important, as sample sizes were fairly low in these areas.

Importantly, these results also show that within and among generations, gene flow among groups separated by the current IWC area boundaries is quite high, therefore confirming previous genetic analyses (SC/F13/SP17). Relatedness analyzes seem to be promising for future investigation of relationships among these groups, but would crucially need to be complemented by other biological information which will help to assess the type of relationships observed.

REFERENCES LIST

- Andersen L.W., Born E.W., Dietz R., Haug T., Oien N., and Bendixen C., 2003. Genetic population structure of minke whales *Balaenoptera acutorostrata* from Greenland, the North East Atlantic and the North Sea probably reflects different ecological regions. *Marine Ecology Progress Series*, 247: 263-280.
- Anderwald P., Daniélsdóttir A.K., Haug T., Larsen F., Lesage V., Reid R.J., Víkingsson G.A., and Hoelzel A.R., 2011. Possible cryptic stock structure for minke whales in the North Atlantic: implications for conservation and management. *Biological Conservation*, 144: 2479-2489.
- Árnason A., and Spillaert R., 1991. A study of variability in minke whale (*Balaenoptera acutorostrata*) in the North Atlantic using a human hypervariable region probe alpha-globin 3'HRV. *Report of the International Whaling Commission*, (special issue 41): 439-443.
- Benjamini Y., and Hochberg Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B, (Methodological)*: 289-300.
- Benónisdóttir S., 2012. Statistical approach to relatedness analysis in large collections of genetic profiles. An application to a DNA-registry of fin whales. Ms Thesis in Statistics – Data analyse, University of Bergen, Faculty of Mathematics and Natural Sciences, Norway.
- Daniélsdóttir A.K., Duke E.J., and Árnason A., 1992. Genetics variation at enzyme loci in North Atlantic minke whale, *Balaenoptera acutorostrata*. *Biochemical Genetics*, 30: 189-202.
- Daniélsdóttir A.K., Halldórsson S.D., Guðlaugsdóttir S., and Árnason A., 1995. Genetic variation in northeastern Atlantic minke whales (*Balaenoptera acutorostrata*). pp. 105-118. In: A.S. Blix, L.Walløe, Ø. Ulltang (eds) Whales, seals and man.
- Donovan G.P., 1991. A review of IWC stock boundaries. *Report of the International Whaling Commission*, (special issue 13): 39-68.
- Glover K.A., Haug T., Øien N., Walløe L., Lindblom L., Seliussen B.B., and Skaug H.J., 2012. The Norwegian minke whale DNA register. *Fish and Fisheries*, 13: 313-332.
- Haaland Ø.A., Glover K.A., Seliussen B.B., and Skaug H.J., 2011. Genotyping errors in a calibrated DNA-register: implications for identification of individuals. *BMC Genetics*, 12: 36.
- Pampoulie C., Daniélsdóttir A.K., and Víkingsson G.A., 2008. Genetic structure of the North Atlantic minke whale (*Balaenoptera acutorostrata*) at feeding grounds: a microsatellite loci and mtDNA analysis. Paper SC/60/PF110 presented to the IWC Scientific Committee, June 2008, Santiago, Chile.
- R Core Team, 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Skaug H.J., and Oien N., 2005. Genetic tagging of male North Atlantic minke whales through comparison of maternal and foetal DNA-profiles. *Journal of Cetacean Research and Management*, 7: 113-117.
- Skaug H.J., Bérubé M., and Palsbøll P.J., 2010. Detecting dyads of related individuals in large collections of DNA-profiles by controlling the false discovery rate. *Molecular Ecology Resources*, 10: 693-700.
- Tiedemann R., Cipriano F., Morin P. A., Hoelzel A. R., Palsbøll P., Waples R., Natoli A., Bachmann L., Postma L., Double M., Pampoulie C., Skaug H., Baker C. S., and Jackson J., 2012. Updated guidelines for DNA data quality control and error rate estimation, for genetic studies relevant to IWC management advice. (First presented at IWC-SC60; updated by RT, FC, PM, RW, and JJ April/May 2012), Paper SC/64/SD4 presented at the IWC, Panama City, Panama.