ICES CM 2007/L:21

Population genetic studies in Northeastern Atlantic minke whales

M. Berube, H. Skaug, L. Andersen, T. Haug, N. Øien

Introduction

Minke whales are the most abundance species of baleen whales in the North Atlantic. As part of current management of minke whales in Norwegian and adjacent waters, a DNA-register have been established. The register ensures that samples are taken of each animal caught under the Norwegian catch quota, and that a DNA-profile is established and stored in a database from each individual whale. Previous studies have indicated that genetic population sub-structure exists within the North Atlantic, but sample sizes were limited. We present an analysis based on the sex, mtDNA control region sequences and 10 microsatellite loci from the 4500 individuals that currently constitute the DNA-register. Information about population structure is an essential input to the management procedure applied for North Atlantic minke whales.

Material and methods

The establishment of the Norwegian minke whale DNA-register ensures that samples (muscle tissues) are taken of each animal caught under the Norwegian catch quota, and that a DNA-profile is established and stored in a database from each individual whale (Olaisen, 1997). The DNA-profile consists of 10 microsatellites, mtDNA and a sex-marker (Dupuy and Olaisen, 1999). In addition, for each animal the register contains information about the time and geographical location of capture, as well as some biological parameters (length etc.). For the period 1997-2005, the DNA-register contains DNA-profiles for approximately 4,500 individuals. Two different genetic laboratories have been involved in the analysis: VitaTech in Canada (years 1997-2002) and Marine Research Institute (MRI) at Iceland (years 2003-2005).

We have run three kinds of analyses;

1. GENEPOP to estimate the probability of homogeneity in allele frequencies between pairs of sample partitions (Rousset and Raymond, 1995),
2. GENECLASS which estimates the probability of each multi-locus genotype in each sample partition (Piry et al., 2004), and
3. MDIV, which estimates the population divergence time, gene flow and genetic diversity (as theta) from nucleotide sequences (Nielsen and Wakeley, 2001).

Data partitioning

We only used mitochondrial DNA sequences in the analysis (3). The way that the Norwegian minke whale DNA-register registers and abbreviates the mtDNA sequences is not appropriate for an MDiv analysis. Instead, we have used mtDNA sequences from 127 individuals, taken from Berube and Palsbøll (unpublished) and from Bakke et al. (1996). In addition we added 27 sequences from the Gulf of St. Lawrence (unpublished) to use the most distant North Atlantic "population" in the MDiv assessment.

For the remainder of the data analyses, we have used the DNA register microsatellite data. We have kept the VitaTech and MRI data sets separate as there are some disagreements in terms of what alleles are called. Splitting the dataset in two has the unfortunate effect of reducing the power of the analysis.

In each case, we partitioned samples into sex, year and month, e.g. one sample would be females, May month 2001. This was done because it is known that females arrive earlier and go further north than the males. This would mean that if whales came from different breeding populations then you could have females from one population with males from another in any one area at a given time. Of course the onset of migration may not happen exactly the same time every year why females in one year and area may be from a different breeding population than in another year in the same area(s). Of course you may have (e.g. males) in different areas within one month and thus males should be further divided into say a southern and northern partition but that requires mapping all the data.

Results

In the GenePop analysis (1), no pair-wise comparisons aimed at testing homogeneity in allele frequencies yielded any p-values of less than 0.05 after Bonferroni sequential corrections.

For the GeneClass analysis, the VitaTech data was 3284 individuals and 44 sample partitions. In no case was an individual "rejected" (p-value < 0.001) from its "own" sample partition. Next, individuals all individuals were tested against other sample partitions. Of the total of 141212 (3284 times 43) tests conducted, 332 yielded p-values less than 0.001, that is some 170 more than the 140 or so expected. The IMR data showed the same trend, 1793 individuals and 22 sample partitions. No sample was rejected from its own partition, and of the 37653 remaining tests 157 resulted in p-values < 0.001 (the expected number would be ca. 37). All the GeneClass assessment yielded slightly more statistically rejections than expected with an α value of 0.001, but whether this is a sign of some low levels of "population structure" or simple due chance is hard to say without a more thorough analysis of the estimated likelihood values. However, there was no readily and obvious trend among the "significant" p-values. However, since half of them at least are likely due to chance then it would be hard to discern any trend.

The MDiv analysis was based upon 27 Gulf of St. Lawrence and 127 Northeastern Atlantic mtDNA sequences. Multiple runs (based upon a finite-site mutation model) indicated relatively high levels of gene flow at Nm around 30 or more, but a very poorly defined likelihood surface. The estimates of population divergence time was at or very close to zero (some runs significantly more likely than zero but still very low). How low exactly requires assumptions about the mutation rate which is tricky, but *very recent* is definitively the case. The mt control region nucleotide sequences in the register should help this kind of analysis a lot, but (as stated above) the current form of the data makes this difficult. The MDiv analyses leads us to conclude that the current population is not in mutation-drift-migration equilibrium, but a more formal assessment would require a comparison of estimates obtained from e.g., MIGRATE or LAMARCK.

Discussion

The results of this analysis are preliminary. It seems that the North Atlantic minke whale populations are either (1) essentially one more or less panmictic population (in all practical terms), or (2) that the current populations are of very recent origin (from the same ancestral population) but not yet in the (usually assumed) mutation-drift-migration equilibrium.

In order to resolve this, one likely needs to get a handle on the tempo-spatial distribution of close relatives, and likely in a "small" population. West Greenland might be a good candidate as this population presumably is small and if a reasonable large proportion of individuals were sampled and genotyped, the observed number of offspring (identified using the genetic data) with the expected distribution in a "closed" population obtained from demographic simulations. If the number of observed parent-offspring dyads is less that the 2.5 percentile of the expected distribution then this would suggest migration to and from West Greenland.

References

Bakke I, Johansen S, Bakke Ø, El-Gewely MR. 1996. Lack of population subdivision among minke whales (Balaenoptera acutorostrata) from Icelandic and Norwegian waters based on mitochondrial DNA sequences. Mar Biol 125:1–9

Dupuy, B. and B. Olaisen. 1999. Typing procedure for the Norwegian minke whale DNA register. Institute of Forensic Medicine, University of Oslo.

Nielsen, R. and J. W. Wakeley. 2001. Distinguishing Migration from Isolation: an MCMC Approach. *Genetics* 158: 885-896.

Olaisen, B. 1997. Proposed specifications for a Norwegian DNA database register for minke whales. Unpublished report to the Scientific Committee of the International Whaling Commission (Available upon request from the IWC secretariat) SC/49/NA1.

Piry S, Alapetite A, Cornuet, J.-M., Paetkau D, Baudouin, L., Estoup, A. (2004) GeneClass2: A Software for Genetic Assignment and First-Generation Migrant Detection. *Journal of Heredity* **95**:536-539.

Rousset, F. and Raymond, M. (1995). GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *Journal of Heredity*, 86: 248-249.