Fisheries and Oceans Canada | Pêches et Océans Canada
Science | Sciences

C S A S | ***S C C S***

Canadian Science Advisory Secretariat | **Secrétariat canadien de consultation scientifique**

Research Document 2011/021 | **Document de recherche 2011/021**

Population Genetic Structure of Narwhal (*Monodon monoceros*)

Structure génétique de la population de narvals (*Monodon monoceros*)

Petersen, S.D., D. Tenkula and S.H. Ferguson

Fisheries and Oceans Canada / Pêches et des Océans Canada
501 University Crescent / 501, Université Crescent
Winnipeg, MB R3T 2N6

ISSN 1499-3848 (Printed / Imprimé)
ISSN 1919-5044 (Online / En ligne)

Canada

TABLE OF CONTENTS

Correct citation for this publication:

Petersen, S.D., Tenkula, D. and Ferguson, S.H. 2011. Population Genetic Structure of Narwhal (*Monodon monoceros*). DFO Can. Sci. Advis. Sec. Res. Doc. 2011/021. vi + 20 p.

ABSTRACT

Narwhal (*Monodon monoceros*) is a key element of the eastern Canadian Arctic ecosystem, and is a culturally important species for Inuit. As the Arctic becomes of greater interest for development, and as northern communities grow in size, narwhals will be subject to increasing anthropogenic pressure. There may be additional natural pressures in the form of increased predation by killer whales (*Orcinus orca*) and frequency of ice entrapments. In the face of these pressures, it is important to actively manage anthropogenic impacts. Currently, this is most effectively done through the establishment of harvest regulations directed at specific stocks. Identification of stock structure is needed to facilitate management both nationally and internationally. To identify stock structure, narwhals from the entire species range were genetically profiled using 16 microsatellite markers. Bayesian analysis to determine the number of genetic clusters without sampling location data was not able to resolve population structure at any level. This was in spite of significant genetic differentiation among the Baffin Bay, Northern Hudson Bay, and East Greenland populations (F_{ST} values between 0.011 and 0.028). These results support the population divisions previously established by mitochondrial sequences, contaminant analysis, and satellite telemetry. Within Baffin Bay, partial support was found for the existing stock designations using multivariate analysis to differentiate groups. The analysis suggests that narwhals from Jones Sound and the Somerset Island summering stock are differentiated, although the Somerset Island finding may be influenced by low sample numbers. Increased sampling of most stocks is suggested to increase the understanding of how stocks are related and at what rate migration among stocks may occur.

RÉSUMÉ

Le narval (*Monodon monoceros*) est un élément clé de l'écosystème de l'est de l'Arctique canadien, et il constitue une espèce d'importance culturelle pour les Inuits. Avec le développement de l'Arctique qui revêt de plus en plus d'intérêt, et avec les collectivités nordiques qui prennent de l'expansion, le narval sera soumis à des pressions accrues découlant des activités anthropiques. La prédation accrue exercée par les épaulards (*Orcinus orca*) et l'augmentation de la fréquence d'emprisonnement par les glaces piégeage risquent également d'accentuer les pressions de sources naturelles. Face à de telles pressions, il est important de gérer proactivement les impacts anthropiques. Présentement, les mesures de gestion les plus efficaces passent par l'établissement de règles de récolte visant spécialement les stocks. L'identification de la structure du stock est nécessaire pour faciliter la gestion à l'échelle nationale et internationale. Afin d'identifier la structure du stock, on a établi le profil génétique des narvals pour l'ensemble de l'aire de répartition de l'espèce, au moyen de 16 marqueurs de l'ADN microsatellite. L'analyse bayésienne, visant à déterminer le nombre de groupes génétiques sans avoir recours à des données sur des points d'échantillonnage, n'a pas permis de définir la structure de la population à quelque niveau que ce soit. Ce résultat a été obtenu malgré une forte différenciation génétique entre les populations de la baie de Baffin, du nord de la baie d'Hudson et de l'est du Groenland (valeurs F_{ST} oscillant entre 0,011 et 0,028). Ces résultats appuient les divisions de la population établies antérieurement grâce aux séquences mitochondriales, aux analyses des contaminants et à la télémesure satellitaire. Dans la baie de Baffin, on a trouvé un appui partiel pour l'identification du stock existant en ayant recours à l'analyse multivariable visant à différencier les groupes. L'analyse suggère la différenciation des stocks d'estivage du détroit de Jones et de l'île Somerset, bien qu'il soit possible que les résultats de l'île Somerset aient été biaisés par le faible nombre de spécimens. On suggère d'augmenter l'échantillonnage de la plupart des stocks afin de mieux comprendre les relations entre les stocks et le rythme des migrations entre les stocks.

INTRODUCTION

Narwhals (*Monodon monoceros*) are mid-sized toothed whales that inhabit the Arctic seas north of 60°N. This iconic species is an important resource for Inuit communities' annual food budgets, economic well-being, and cultural identity. In the future, narwhals may be at risk due to increased mortality associated with climate change. In particular, it has been suggested that narwhals will be more vulnerable to ice entrapments (i.e., when animals are not able to reach open water because they are trapped under forming ice) if patterns of sea ice formation become increasingly unpredictable (Laidre and Heide-Jorgensen, 2005; Tynan and DeMaster, 1997). Therefore, it is important to understand the stock structure of this species, and to use that information to effectively manage this important resource. Gaining this understanding is difficult due to the logistical constraints involved in studying an ice-associated marine mammal in the high Arctic. However, several outstanding questions may be resolved by using genetic methods, and then inferring biological and ecological processes that could lead to the structure and patterns in the observed genetic variability.

Genetic structure can be formed when groups of animals, within a species, segregate to some degree and have a higher probability of mating with members of their own group than with members of other groups. However, this loose definition fails to account for the complexity observed in many species, including most cetaceans. These species are characterized by large body size, wide species distribution, and the propensity to make long-distance movements across oceans and ecosystems. This creates a situation where population structure can be very complex (Hoelzel, 1989). Compounding these challenges is the fact that cetaceans are long lived and have overlapping generations, which will counteract the forces of genetic drift that would otherwise differentiate populations (Petersen et al., 2010). Although these characteristics are advantageous in an evolutionary sense, they create significant challenges for research that is aimed at identifying genetic groupings for conservation and management purposes. This is because although groups are demographically separate and reproductively isolated it will take a long time to become genetically differentiated.

In narwhals, several hierarchical levels of population structure have been proposed based on spatial segregation of animals during their annual migration cycle. At the coarsest scale, observational data, surveys, and tagging studies indicate that three populations occur throughout their range (Reeves and Tracey, 1980; Strong, 1988) (Figure 1). These three populations consist of narwhals that move seasonally between: 1) northern Hudson Bay and Hudson Strait; 2) Baffin Bay and the high Arctic; and 3) eastern Greenland and the Greenland Sea. Data from satellite-tagged animals in Canada have supported these groupings (Dietz et al., 2008; Westdal, 2008). Mitochondrial sequence data from the control region, as well control region haplotype frequencies, have also supported the isolation of eastern Greenland narwhals from western Greenland and Canadian narwhals, (Palsbøll et al., 1997) as well as the isolation of northern Hudson Bay from Baffin Bay narwhals (de March et al., 2003).

At a smaller spatial scale, the Baffin Bay population has been divided into a number of stocks based on summering concentrations of narwhals, and has been delineated using data from telemetry, survey, and hunting data (Heide-Jørgensen et al., 2005; Richard, 2010). The term 'stock' refers specifically to resource unit subjected to hunting (DFO 2010) and these may, or may not, correspond to actual genetic units (i.e., populations). The current working hypothesis is that the Baffin Bay population is a meta-population comprised of a number of summering stocks. (Heide-Jørgensen et al., 2005; Heide-Jørgensen et al., 2003). Five stocks have been identified in the Canadian Arctic with up to three more hypothesised to exist in the High Arctic (Parry Channel, Jones Sound, and Smith Sound), and two stocks have been identified in

northwest Greenland (Figure 1 inset). Although these stocks have been delineated using primarily non-genetic data sources, genetic data have been used to support some stock level groups. For example, de March et al. (2003) observed some evidence of differentiation between narwhals from Grise Fiord (Jones Sound) and other Baffin Bay narwhals.

It is also recognised that some stocks may be made up of multiple genetic units. For example, the East Baffin Island summering stock may be subdivided into smaller stocks (Heide-Jørgensen et al., 2005). This has been supported by analyses of contaminants, which suggest that narwhals harvested in Pangnirtung and Qikiqtarjuaq were similar to each other but differentiated from other east Baffin samples (Clyde River) (de March and Stern, 2003). At this scale genetic structure is likely to be the result of seasonal fidelity to specific inlets and fiords, coupled with female-mediated genetic structure as a result of matrifocal pod structure. Matrifocal pod structure has been suggested for narwhals (Hay and Mansfield, 1989), and could contribute to the formation of genetic structure through the accumulation of similar genetic lineages in summer aggregations. Satellite telemetry has suggested that narwhal stocks show a high degree of fidelity to migration routes to and from summering areas (Dietz et al., 2008; Heide-Jørgensen et al., 2002; Heide-Jørgensen et al., 2003). Individual narwhals that have been fitted with satellite transmitters, and for which the instruments functioned long enough to record the spring migration, seem to return to the area where they were initially tagged (Heide-Jørgensen et al., 2003).

Although there is good evidence that the narwhals in each stock behave predictably (i.e., show fidelity to summering grounds, wintering grounds, and migration paths), there is currently uncertainty regarding the ability to differentiate stocks in a manner that will facilitate the management of a mixed-stock harvest. This information is needed, because communities that harvest narwhals often do so during the migration, and may be harvesting from some stocks that are depleted on their summering grounds. Therefore, the objective of this study is to assess whether or not genetic data from nuclear microsatellites can differentiate groupings at the various hierarchical levels outlined above. Specifically, we test the hypotheses: that 1) there are three populations of narwhals within their range, and 2) the stocks that are currently recognized are supported by genetic data. The second hypothesis deals specifically with the Baffin Bay population, which is subdivided into summering stocks for management purposes.

MATERIALS AND METHODS

FIELD

Samples of skin or muscle tissue were obtained from Inuit hunters as part of various whale sampling programs (i.e., contaminants). Tissues were preserved in a solution of 20% DMSO saturated with NaCl for transport at ambient temperatures or frozen. Sample collection has occurred opportunistically over approximately 30 years (1982 to 2009) and has not been focused on any one community or hunt phase (i.e., summering area, floe edge). Additional sample collections have also been made as part of the Fisheries and Oceans Canada (DFO) whale-tagging program. These samples were treated similarly. On arrival at the laboratory, samples were frozen at -20°C until DNA extraction. Samples from Greenland were contributed by the Greenland Institute of Natural Resources, and were obtained through subsistence harvests and scientific sampling programs.

LABORATORY

Genomic DNA was extracted using the BioSprint platform (Qiagen Inc., Valencia, CA, USA), which conducts extractions in a 96-well format, using magnetic particles to isolate nucleic acids. Isolated DNA was then quantified (NanoDrop, Thermo Fisher Scientific Inc., Wilmington, DE, USA) and normalized to 10ng/µl. (Note: in the past phenol:chloroform and Qiagen spin columns [DNeasy Blood & Tissue Kits] have also been employed.)

Sex was determined using primers LGL331 (CAA ATC ATG CAA GGA TAG AC) and LGL335 (AGA CCT GAT TCC AGA CAG TAC CA), which were developed by Shaw et al. (2003). The following concentrations of reagents were included in a total volume of 10µl: 1x PCR buffer (Thermopol: 20mM Tris-HCl, 10mM $(NH_4)_2SO_4$; 10mM KCl; 2mM $MgSO_4$; 0.1% Triton X-100, pH 8.8 @ 25°C) (New England Biolabs Inc., Ipswich, MA, USA); 0.5mM $MgCl_2$; 0.2mM dNTPs; 10%/V BSA; 0.3µM each primer; 0.5U *Taq* polymerase (New England Biolabs Inc., Ipswich, MA, USA); and 20ng of template DNA. Thermocycling conditions consisted of an initial denaturing step of 4min at 94°C; followed by 30 cycles, each consisting of 30s at 94°C, then 30s at 55°C, then 30s at 72°C; and a final extension of 3min at 72°C to complete the amplification. Amplification products were visualized using the QIAxcel system (Qiagen Inc., Valencia, CA, USA). Amplification of two fragments indicated a male (X and Y fragments), while a single amplification product (two X fragments) indicated a female.

Samples were genetically profiled using 16 microsatellite loci developed for other cetaceans (Buchanan et al., 1996; Caldwell et al., 2002; Hoelzel et al., 1998; Krützen et al., 2001; Valsecchi and Amos, 1996). Loci were amplified individually or in duplexes, in reactions with a total volume of 10µl and the following reagent concentrations: 1x GenAmp buffer II (10 mM Tris-HCl pH 8.3, 50 mM KCl) (Life Technologies Corporation, Carlsbad, CA, USA); 1.5mM $MgCl_2$; 0.2mM dNTPs; a variable amount of each primer (Appendix 1); 0.5U Ampli*Taq* Gold polymerase (Life Technologies Corporation, Carlsbad, CA, USA); and 20ng of template DNA. Themocycling conditions varied among reactions (Appendix 1). Single step conditions were as follows: an initial denaturing step of 11min at 94°C; followed by 25 cycles, each consisting of 45s at 94°C; then 45s at annealing temperature; then 45s at 72°C; and a final extension of 15min at 72°C to complete the amplification. One step-up program and one step-down program was also used (Appendix 2). Amplification products were denatured and run with GeneScan 400HD Rox size standard (Life Technologies Corporation, Carlsbad, CA, USA), using an ABI 3130xl Genetic Analyser (Life Technologies Corporation, Carlsbad, CA, USA). Allele scoring was conducted in GeneMarker 1.90 (SoftGenetics, State College, PA, USA).

ANALYSES

Error quantification

Approximately 10% of the samples were extracted and genetically profiled in duplicate. In addition, each plate (~96 samples) carried four to six positive controls from two individuals. Error rates were quantified as the mean number of errors per locus and averaged over all loci used in the analyses.

Diversity and Equilibrium

Departures from Hardy-Weinberg Equilibrium and linkage disequilibrium were assessed using the program FSTAT vers. 2.9.3.2 software (Goudet, 1995). Ten thousand randomizations were used to test for deviations from Hardy-Weinberg equilibrium and for linkage disequilibrium. Calculations to estimate genetic diversity were performed using the FSTAT vers. 2.9.3.2 software (Goudet, 1995) and MStools (Park, 2001). Where necessarily Bonferroni adjustment of alpha (α=0.05) was conducted to account for multiple comparisons.

Population Structure

Non-a priori analyses

Bayesian analysis was used to examine the data for partitions that conform to the expectations of Hardy-Weinberg equilibrium, as well as linkage equilibrium within clusters. No location data are provided to the analyses; therefore, the most likely number of clusters that the data are partitioned into should reflect the number of populations that have been sampled. The program STRUCTURE 2.3.2 (Pritchard et al., 2000) was used to conduct these analyses. STRUCTURE was run using the correlated allele model (Falush et al., 2003) where allele frequencies are correlated among clusters and with the admixture parameter (alpha) inferred from the data. The analysis was initiated with a 100,000 cycle burn-in followed by 500,000 MCMC repetitions. Three independent runs were preformed for each modeled number of clusters (K). The number of clusters investigated was from one to ten.

Spatial a priori analyses

At the largest scale, three populations have been identified, which were used for the analysis to determine if the genetic data support this hypothesis. These three populations are spatially separated; therefore, all samples were included in the analysis. Pairwise F_{ST} (Weir and Cockerham, 1984) was calculated among populations, and tested for significant deviation from zero using 10,000 randomizations. The program FSTAT vers. 2.9.3.2 software (Goudet, 1995) was used for all calculations of differentiation. Pairwise F_{ST} was also calculated among summering stocks (see below) using the same software. Mean relatedness was calculated for each summering stock, because the matrifocal structure of cetaceans has been implicated in the erroneous conclusion of significant population differentiation in some studies (Palsbøll et al., 2002). GENALEX 6.3 (Peakall and Smouse, 2006) was used to calculate population means using the Queller and Goodnight (1989) estimator. Significance was tested by permuting the data among groups (999 permutations) to obtain 95% confidence intervals around the null hypothesis that all individuals are unrelated (r=0) and by bootstrap re-sampling genotypes within groups (999 bootstraps) to obtain 95% confidence intervals around the group mean. When the 95% confidence intervals for the permutations and bootstrapping did not overlap, significance was inferred.

To test if stocks and locations represent distinguishable units, samples were first sorted to determine if they had capture date information associated with them. If a sample lacked the month it was collected, it was dropped from the analysis. Narwhals migrate seasonally, between summering areas in the open water season and wintering areas. Therefore, if the current stock designations represent true genetic units, it will be most pronounced in groups of samples collected during the summer. Other collections, which are usually from animals harvested at the floe edge or from cracks and leads in the sea ice, likely consist of narwhals from multiple stocks. To remove the mixed harvests, all samples were categorized into either summer-collected samples or migration/winter-collected samples. The designation of ‘summer’ was based on harvest data compiled by Romberg and Richard (2005), which indicated that animals harvested

by communities in the Canadian Arctic between July 23 and September 31 were taken from animals on their summering grounds. As the exact date was often missing, samples taken from the months of August and September were grouped by stock and treated as summering stocks. A minimum sample size of ten individuals was used as a threshold to be included in these analyses.

Multivariate analyses of genetic data sets has been gaining acceptance because of their strong mathematical foundation and freedom from the assumptions inherent in traditional genetic population models (Jombart, 2008; Jombart et al., 2009; Parisod and Christin, 2008; Patterson et al., 2006). Inter-class (between-class) eigenanalysis (inter-class PCA) was used to visualize genetic differentiation among groups. Inter-class PCA maximizes the variance among groups, whereas PCA maximizes the total variance (Jombart, 2008). Individuals and groups are visualized along the axis that maximally separates groups. Inter-class PCA provides an ad hoc assessment of how well groups are partitioning the variation. To compliment this analysis, a correspondence analysis (CA), which also seeks the axes which maximize the variance among groups, was also conducted. This analysis produces figures that are simplified compared to the inter-class PCA while still representing the major differences among groups. We conducted these analyses in the statistical software R (www.r-project.org), using the package adegenet (Jombart, 2008) to manage the genetic data and ade4 (Chessel et al., 2008) to conduct the PCA and CA.

RESULTS

We profiled 967 narwhal samples, collected from the Eastern Canadian Arctic and Greenland between 1982 and 2009, at 16 loci. This dataset was reduced to 15 loci because GT39 failed in 50% of the reactions (note: this was due to primer interactions in the multiplex not because of an inherent problem with the marker). Narwhal samples with allelic profiles at less than 10 of the 15 loci were also dropped from these analyses. Per locus error rates ranged from 0 (11 loci) to 0.061 in FCB17. The average error rate for this dataset was 0.008 before loci were dropped from the dataset because of deviations from Hardy-Weinberg equilibrium. The full working dataset, therefore, contained 877 samples. Significant heterozygote deficits were detected at three loci (FCB8, FCB17, and KWM2) in the analysis of the entire dataset, so these loci were also removed. There was no evidence of linkage between loci in the full working dataset, thus the final compliment of microsatellite markers included 12 loci.

At the population level, gene diversity and observed heterozygosity were similar in all three populations (Table 1). The program STRUCTURE was unable to recover three populations from this dataset as was expected. Although this analysis suggested that two populations (K=2) was the model with the most support (Figure 2), further inspection of the assignment of individuals indicated that none were highly assigned to either cluster. This suggests that STRUCTURE was not able to differentiate clusters effectively (STRUCTURE manual). Significant genetic differentiation (pairwise F_{ST}) was detected among all populations (Table 2), although inter-class PCA suggests some overlap in the allele frequencies of all populations (Figure 3).

As previously indicated, this dataset was reduced in size to only include individuals with known date of capture and to retain only locations with a sample of greater than 10 individuals (N = 268; Table 3) for the stock-level analyses. Samples were grouped into stocks based on sampling location. Samples from the winter 2008 entrapment near Pond Inlet (Eclipse Sound; N=235) were not included in these analyses as their source was unknown. Satellite tagged individuals from Admiralty Inlet have entered Eclipse Sound on their migration to wintering

grounds (Dietz et al., 2008), therefore the entrapped animals could be from non-Eclipse Sound stocks. Few groups showed significant differentiation outside of comparisons with the known populations of Northern Hudson Bay and East Greenland (Table 4). F_{ST} values were an order of magnitude larger in these comparisons compared to comparisons among the Baffin Bay stocks. Similar magnitude F_{ST} values but non-significant differences were observed between Inglefield Bredning and Jones Sound (Table 4). Somerset Island was also differentiated from most other stocks (Table 4). Mean relatedness within each stock was not significantly higher than expected for any groups except for the East Greenland sample (Figure 4).

Both PCA and CA analyses suggest that the Jones Sound stock is separate from other stocks and also separate from the Inglefield Bredning samples, which are geographically close. Samples grouped into the Somerset Island stock also separate out in the multivariate analysis (Figure 4). These samples are mostly from Igloolik and are the smallest sample group (N=12), which may have influenced the analyses.

DISCUSSION

This is the first study to use nuclear microsatellite markers and combine samples from most of the species range (Canada and both eastern and western Greenland) to investigate the population structure of narwhals. Genetic structure was investigated at the population and stock level in order to evaluate whether the existing delineations were supported by the present data.

Similar to previous studies that utilized genetic data (de March et al., 2003; Palsbøll et al., 1997), these analyses of microsatellite data generally supported our prediction of three discrete units. This was most evident in the frequency-based analysis (F_{ST}) and suggested in the inter-class PCA analysis (Table 2 and Figure 3, respectively). The F_{ST} values indicate that East Greenland is most differentiated from Northern Hudson Bay, while there is a similar level of differentiation between East Greenland and Baffin Bay and between Northern Hudson Bay and Baffin Bay. This may suggest similar evolutionary histories in terms of population sizes and degree of isolation. It is also interesting to note that the F_{IS} values for all three populations are indicative of excess homozygotes (positive F_{IS}). This could be due to inbreeding, which may be the case in the east Greenland sample, where individuals sampled were more related (Figure 4), but is likely not the case in Baffin Bay. One other possible explanation for the Baffin Bay results is that they are due to the Wahlund effect, wherein a positive F_{IS} is the result of combining multiple populations into one (Wahlund, 1928). Of interest, entrapment samples were analyzed separately and also showed a positive inbreeding coefficient (F_{IS}=0.020).

The inability of the program STRUCTURE to recover this coarsest scale of population subdivision, in spite of relatively high F_{ST} values, is perplexing. F_{ST} measures of differentiation were low (0.011 – 0.028) but at or near the theoretical limit of resolution for the program STRUCTURE (0.02 to 0.03, Latch et al., 2006). One explanation for this pattern is that of ongoing gene flow; however, it is more likely due to a combination of factors, including properties of the data set and characteristics of narwhals. Although the data set is large, some groups likely do not represent true populations and would not be expected to be identified by STRUCTURE. For example, the east Greenland samples show significantly higher relatedness than the null expectation (Figure 4). Furthermore, network analysis of these samples (not shown) suggests that half of the individuals can be clustered into two highly-related groups. It is not likely that this would represent a population at equilibrium and therefore STRUCTURE could not identify it.

Several biological factors could also be influencing the SRUCTURE analysis. Although there has been recent concern about narwhal abundance and population trends, these populations have been historically large in the recent past and the current population estimates are in the thousands of individuals. The Baffin Bay population is estimated at over 60,000 individuals (Richard et al., 2010), and the smaller northern Hudson Bay population and east Greenland population have been estimated in the range of several thousand animals (Heide-Jørgensen et al., 2010; Richard, 2010). These large population sizes would create considerable genetic momentum, such that allele frequencies would not be influenced by genetic drift. Based on the very low mitochondrial diversity, Palsbøll et al. (1997) have suggested that narwhals have recently expanded from a severely bottlenecked population. This would limit the diversity within the species and lead to increased homogeneity among populations. When this pattern is coupled with lifespans near 100 years (Garde et al., 2007) and overlapping generations, both characteristics that increase the effective population size (Petersen et al., 2010), the ability to differentiate populations becomes more challenging. It is likely that the Baffin Bay population is genetically very similar to the ancestral population of all three groups, and the two smaller populations would consist of alleles drawn from the allele frequencies of the larger population. In this situation, STRUCTURE may be unable to determine the 'true' number of clusters because the data could be divided a number of equally likely ways. This problem has lead to the development of spatial explicit models to give STRUCTURE and other programs a 'boost' in these situations (Hubisz et al., 2009). However, we are hampered in our ability to take advantage of these types of analyses because we lack adequate spatial data.

The current stock structure for narwhals is not based on genetic data, and does not claim that stocks are genetic units (Richard, 2010). However, if we use this as a testable hypothesis, we would predict that the currently-identified stocks correspond to genetic units. Under this scenario, stocks could be considered part of the Baffin Bay meta-population as suggested by Heide-Jørgensen et al. (2005) and may be discrete. Given the inability of STRUCTURE to resolve higher-level divisions, it is not surprising that finer-scale divisions were not identified. Genetic differentiation among these groups was only significant in pairwise comparisons with the two known populations outside Baffin Bay (northern Hudson Bay and east Greenland). Inter-class PCA and CA did indicate the differentiation of Jones Sound narwhals as a discrete group (Figure 5). This was not due to high relatedness of individuals within that sample (Figure 4), and may suggest that Jones Sound animals are isolated from other groups, perhaps wintering in the North Water polyna (Mansfield et al., 1975). Hunters from northwest Greenland have observed narwhals in the North Water polyna (Heide-Jørgensen et al., 2005). However, aerial surveys of the polyna have not detected large numbers of narwhals (Finley and Renaud, 1980). Samples from Inglefield Bredning would be expected to cluster with these animals, but do not, and are most similar to other Baffin Bay groups. Narwhals from Inglefield Bredning are thought to be a distinct group and even though they cluster close to other Baffin Bay stocks in this analysis, they separate out in a CA when the analysis is conducted using summer harvest locations as the grouping variable (not shown).

Although genetic data have been useful for discriminating genetic sub-division in many species, it seems difficult to apply these data to solving the current stock issues in narwhals. This may be partially due to the evolutionary history of narwhals, in that their diversity is low and their population sizes are high, but it is also partially due to the current sampling regime. Previous authors who have approached the problem of stock differentiation in narwhals have considered the low and sporadically-collected samples a major limiting factor (de March et al., 2003; Palsbøll et al., 1997). In this study, we have increased the number of samples; however, when partitioned into summer harvests, and further into stock and locations, few groups remain for analysis. In fact areas of great interest (e.g., Igloolik, Iqaluit, Melville Bay) have low sample

sizes usually collected from a single season, which increases the likelihood they are related and not an accurate sample of the groups that would occur in that area. Future collections should focus on obtaining samples of 20 to 50 individuals from locations during the ice-free season. Whether this is possible for many locations is unknown. At the very least, it will likely require a shift to non-lethal sampling of animals and greater participation by local hunters.

Coinciding with an increased sampling effort, finer scale spatial and temporal data are required for each sample. Although whale sampling forms request morphometric and other data, past sample collections are often missing key data such as hunt location (latitude and longitude), hunt date (year, month, and day), and hunt phase (floe edge, ice crack, or open water). Increased effort should be made to obtain these data in order to maximize the sample value for research. Precise spatial data would assist in distinguishing individuals from widely-separated locations that are currently grouped into a single location based on the hunter's community. This is especially important information as hunters often travel great distances to their preferred hunting location. Telemetry data have shown that individuals from certain stocks leave the summering area at a predictable time, and follow similar migrations routes (Dietz et al., 2008; Heide-Jørgensen et al., 2002; Heide-Jørgensen et al., 2003). Precise temporal data, including both date and hunt phase, would allow for adjustments in grouping samples based on the movement of animals in response to ice conditions. Access to finer scale temporal data would help to distinguish samples taken from animals harvested on the summering grounds as opposed to those animals that are harvested while migrating to and from these areas. Finer temporal and spatial resolution may also help to address questions relating to social structure. In this regard, samples collected from entrapments are important because they represent a known group of animals that are likely related.

Overall, these analyses have provided mixed results with regards to the outstanding question of stock structure in narwhals. High F_{ST} values among the three populations confirm their identity and corroborate past research efforts. However, the inability for STRUCTURE to resolve these groups is puzzling and will require further analysis. At the stock level, multivariate analyses indicate that: 1) with more comprehensive sampling, stocks may be identifiable and 2) in the future, individual whales from mixed harvests may be assigned to specific sources. To get to the stage whereby individual harvested whales can be identified with reasonable confidence to a single summering stock, sampling gaps that exist within some geographically widespread stocks will need to be filled. Specifically, more locations in the Somerset Island and east Baffin Island stocks should be included as neither of these stocks are currently understood to any degree. In addition, narwhal samples from Parry Channel, Jones Sound, and Smith Sound would be important to our overall understanding of narwhal population and stock structure.

Overall the genetic data presented here have broadly supported the current population designation but are less clear in their support of the summering stock level designations. This highlights the fact that cetacean populations have been and continue to be a challenge to differentiate genetically as a result of their recent evolutionary history and biological characteristics. The application of novel multivariate methods to explore the differentiation among putative stocks is promising but in order to be fully realized a comprehensive sampling strategy needs to be implemented. The management of narwhals in Canada, and the joint management of stocks with Greenland, will certainly require increased effort to obtain samples – ideally as part of a comprehensive sampling strategy as opposed to the current opportunistic manner.

ACKNOWLEDGEMENTS

Many thanks are owed to the many hunters who have contributed samples to the DFO archive over the years. We are also grateful to the DFO staff for their efforts in collecting and curating samples. Samples from Greenland were generously provided by the Greenland Institute or Natural Resources. Funding for this project has been obtained from DFO, Polar Continental Shelf Project, Nunavut Implementation Fund, Nunavut Wildlife Research Trust, and National Science and Engineering Council.

LITERATURE CITED

Buchanan, F.C., Friesen, M.K., Littlejohn, R.P., Clayton, J.W. 1996. Microsatellites from the beluga whale *Delphinapterus leucas*. Mol. Ecol. 5: 571-575.

Caldwell, M., Gaines, M.S. and Hughes, C.R. 2002. Eight polymorphic microsatellite loci for bottlenose dolphin and other cetacean species. Mol. Ecol. Notes 2: 393-395.

Chessel, D., Dufour, A.-B., Dray, S., Lobry, J.R., Ollier, S., Pavoine, S. and Thioulouse, J. 2008. ade4. Version 1.4-9.

de March, B.G.E. and Stern, G. 2003. Stock separation of narwhal (*Monodon monoceros*) in Canada based on organochlorine contaminants. DFO Can. Sci. Advis. Sec. Res. Doc. 2003/079.

de March, B.G.E., Tenkula, D.A. and Postma, L.D. 2003. Molecular genetics of narwhal (*Monodon monoceros*) from Canada and West Greenland (1982-2001). DFO Can. Sci. Advis. Sec. Res. Doc. 2003/080.

Dietz, R., Heide-Jørgensen, M.P., Richard, P., Orr, J., Laidre, K.L. and Schmidt, H.C. 2008. Movements of narwhals (*Monodon monoceros*) from Admiralty Inlet monitored by satellite telemetry. Polar Biol. 31(11): 1295-1306.

DFO. 2010. Stock definition of Belugas and Narwhals in Nunavut. DFO. Can. Sci. Advis. Sec. Sci. Advis. Rep. 2009/079.

Evanno, G., Regnaut, S. and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611-2620.

Falush, D., Stephens, M. and Pritchard, J.K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567-1587.

Finley, K.J. and Renaud, W.E. 1980. Marine mammals inhabiting the Baffin Bay North Water in winter. Arctic 33(4): 724-738.

Garde, E., Heide-Jørgensen, M.P., Hansen, S.H., Nachman, G. and Forchhammer, M.C. 2007. Age specific growth and remarkable longevity in narwhals (*Monodon monoceros*) from West Greenland as estimated by aspartic acid racemization. J. Mammal. 87(1): 49-58.

Goudet, J. 1995. FSTAT (vers. 1.2): a computer program to calculate F-statistics. J. Hered. 86: 485-486.

Hay, K.A. and Mansfield, A.W. 1989. Narwhal, *Monodon monoceros* Linnaeus, 1758. *In* Handbook of Marine Mammals: volume 4 river dolphins and the larger toothed whales. *Edited by* Ridgway, S.H and Harrison, R. Academic Press. Toronto. p. 145-176.

Heide-Jørgensen, M.P., Dietz, R., Laidre, K.L. and Richard, P. 2002. Autumn movements, home ranges, and winter density of narwhals (*Monodon monoceros*) tagged in Tremblay Sound, Baffin Island. Polar Biol. 25: 331-341.

Heide-Jørgensen, M.P., Dietz, R., Laidre, K.L., Richard, P.R., Orr, J. and Schmidt, H.C. 2003. The migratory behaviour of narwhals (*Monodon monoceros*). Can J Zool 81:1298-1305.

Heide-Jørgensen, M.P., Dietz, R. and Laidre, K.L. 2005. Metapopulation structure and hunt allocation of narwhal in Baffin Bay. NAMMCO/JCNB working group. 37 p.

Heide-Jørgensen, M.P., Laidre, K.L., Burt, M.L., Borchers, D.L., Marques, T.A., Hansen, R.G., Rasmussen, M. and Fossette, S. 2010. Abundance of narwhals (*Monodon monoceros*) on the hunting grounds in Greenland. J. Mammal. 91(5): 1135-1151.

Hoelzel, A.R. 1989. Genetic structure of cetacean populations in sympatry, parapatry, and mixed assemblages: implications for conservation policy. J. Hered. 89(5): 451-458.

Hoelzel, A.R., Dahlheim, M. and Stern, S.J. 1998. Low genetic variation among killer whales (*Orcinus orca*) in the eastern North Pacific and genetic differentiation between foraging specialists. J. Hered. 89: 121-128.

Hubisz, M.J., Falush, D., Stephens, M. and Pritchard, J.K. 2009. Inferring weak population structure with the assistance of sample group information. Mol. Ecol. Res. 9(5): 1322 - 1332.

Jombart, T. 2008. adegenet: a R package for the mulitvariate analysis of genetic markers. Bioinformatics 24: 1403-1405.

Jombart, T., Pontier, D. and Dufour, A.-B. 2009. Genetic markers in the playground of multivariate analysis. Heredity 102: 330-341.

Krützen, M., Valsecchi, E., Connor, R.C. and Sherwin, W.B. 2001. Characterization of microsatellite loci in *Tursiops aduncus*. Mol. Ecol. Notes 1: 170-172.

Laidre, K.L. and Heide-Jorgensen, M.P. 2005. Arctic sea ice trends and narwhal vulnerability. Biol. Conserv. 121(4): 509-517.

Latch, E.K., Dharmarajan, G., Glaubitz, J.C. and Rhodes, O.E. 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. Conserv. Gen. 7:295-302.

Mansfield, A.W., Smith, T.G. and Beck, B. 1975. The narwhal, *Monodon monoceros*, in eastern Canadian waters. J. Fish. Res. Board Can. 32: 1041-1046.

Palsbøll, P.J., Heide-Jørgensen, M.P. and Dietz, R. 1997. Population structure and seasonal movements of narwhals, *Monodon monoceros*, determined from mtDNA analysis. Heredity 78: 284-292.

Palsbøll, P.J., Heide-Jørgensen, M.P. and Bérubé, M. 2002. Analysis of mitochondrial control region nucleotide sequences from Baffin Bay belugas (*Delphinapterus leucas*): detecting pods or sub-populations? NAMMCO Scientific Publications 4: 39-50.

Parisod, C. and Christin, P.-A. 2008. Genome-wide association to fine-scale ecological heterogeneity within a continuous population of *Biscutella laevigata* (brassicaceae). New Phytologist. 178(2): 436–447.

Park, S.D.E. 2001. Trypanotolerance in West African cattle and the population genetic effects of selection. Thesis (Ph.D.). University of Dublin. Dublin.

Patterson, N., Price, A.L. and Reich, D.E. 2006. Population structure and eigenanalysis. PLoS Genet. 2(12): 2074-2093.

Peakall, R. and Smouse, P.E. 2006. GENALEX 6: genetic analysis in excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6: 288-295.

Petersen, S.D., Hainstock, M. and Wilson, P.J. 2010. Population genetics of Hudson Bay marine mammals: current knowledge and future risks. *In* A little less Arctic: changes to top predators in the world's largest nordic inland sea, Hudson Bay. *Edited by* Ferguson, S.H., Lisetto, L. and Mallory, M. Springer. p 237-265.

Pritchard, J.K., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Queller, D.C. and Goodnight, K.F. 1989. Estimating relatedness using genetic markers. Evolution 43(2): 258-275.

Reeves, R.R. and Tracey, S. 1980. *Monodon monoceros*. Mammalian Species 127: 1-7.

Richard, P.R. 2010. Stock definition of belugas and narwhals in Nunavut. Winnipeg, MB: DFO Can. Sci. Advis. Sec. Res. Doc. 2010/022. iv + 14 p.

Richard, P., Laake, J.L., Hobbs, R.C., Heide-Jørgensen, M.P., Asselin, N.C. and Cleator, H. 2010. Baffin Bay narwhal population distribution and numbers: aerial surveys in the Canadian High Arctic, 2002-04. Arctic 63(1): 85-99.

Romberg, S. and Richard, P. 2005. Seasonal distribution and sex ratio of narwhal catches in the Baffin region of Nunavut Territory, Canada. NAMMCO. Report nr NAMMCO/SC/13-JCNB/SWG/2005-JWG/9.

Shaw C.N., Wilson P.J. and White BN. 2003. A reliable molecular method of gender determination for mammals. J Mammal 84(1): 123-128.

Strong, J.T. 1988. Status of the Narwhal, *Monodon monoceros*, in Canada. Can. Feild-Nat. 102(2): 391-398.

Tynan, C.T., DeMaster, D.P. 1997. Observations and predictions for Arctic climate change: potential effects on marine mammals. Arctic 50(4): 308-322.

Valsecchi, E. and Amos, W. 1996. Microsatellite markers for the study of cetacean populations. Mol Ecol 5: 151-156.

Wahlund, S. 1928. Composition of populations and correlation appearances viewed in relation to the studies of inheritance. Hereditas 11:65-106.

Weir, B.S. and Cockerham, C.C. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38(6): 1358-1370.

Westdal, K.H. 2008. Movement and diving of northern Hudson Bay narwhals (*Monodon monoceros*): relevance to stock assessment and hunt co-management. Thesis (M. Env.) University of Manitoba. 103 p.

Table 1. Narwhal genetic diversity summarized for the entire data set and for three populations. Sample size (N), expected unbiased heterozygosity (Hz), observed heterozygosity (obs Hz), mean number of alleles per locus (A_n), mean allelic richness based on a minimum sample of 34 individuals (A_r), and inbreeding coefficient (F_{IS}) for each population

Population	**N**	**Hz**	**obs Hz**	**A_n**	**A_r**	**F_{IS}**
Baffin Bay	785	0.6722	0.6584	11.17	6.89	0.020
East Greenland	46	0.6485	0.6369	6.50	6.19	0.018
Northern Hudson Bay	46	0.6725	0.6539	7.42	7.07	0.028
Global	877	0.6727	0.6570	11.50	6.91	0.021

Table 2. Level of differentiation (pairwise F_{ST}) among populations of narwhal (below diagonal). All comparisons were significantly different. P values are displayed above the diagonal with significant values (P< 0.016) in bold. Bonferroni adjustment of alpha (α=0.05) was conducted to account for multiple comparisons (α=0.016).

	BB	EG	NHB
Baffin Bay		**0.016**	**0.016**
East Greenland	0.016		**0.016**
Northern Hudson Bay	0.011	0.028	

Table 3. Genetic diversity of samples included in the narwhal stock level of analysis. Sample size (N), expected unbiased heterozygosity (Hz), observed heterozygosity (obs Hz), mean number of alleles per locus (A_n), mean allelic richness based on a minimum sample of five individuals (A_r), inbreeding coefficient (F_{IS}), and mean relatedness (R_m) for each stock.

Summering Stocks	**N**	**Hz**	**obs Hz**	**A_n**	**A_r**	**F_{IS}**	**R_m**
Admiralty Inlet	22	0.68	0.69	5.92	3.89	-0.011	-0.006
East Baffin Island	38	0.69	0.70	6.83	3.91	-0.011	-0.015
East Greenland	46	0.65	0.64	6.50	3.64	0.018	0.071
Eclipse Sound	29	0.68	0.64	6.42	3.96	0.068	-0.007
Ingelfield Bredning	31	0.65	0.67	6.08	3.69	-0.027	0.054
Jones Sound	42	0.71	0.70	8.00	4.19	0.026	-0.033
N Hudson Bay	48	0.67	0.65	7.42	3.94	0.030	0.026
Somerset Island	12	0.67	0.70	4.58	3.64	-0.051	0.028
Global	268	0.68	0.67	6.47	3.86	0.014	0.010

Table 4. Level of differentiation (pairwise F_{ST} below the diagonal) observed among summering stocks of narwhal. P values are displayed above the diagonal with significant values in bold (P<0.002). Bonferroni adjustment of alpha (α=0.05) was conducted to account for multiple comparisons (α=0.002).

	AI	EB	EG	ES	IB	JS	NHB	SI
Admiralty Inlet		0.763	**0.002**	0.938	0.221	0.566	0.132	0.063
East Baffin Island	-0.007		**0.002**	0.830	0.080	0.089	**0.002**	0.373
East Greenland	0.020	0.011		**0.002**	**0.002**	**0.002**	**0.002**	0.048
Eclipse Sound	-0.007	-0.002	0.013		0.009	0.020	0.023	0.198
Inglefield Bredning	-0.002	0.002	0.020	0.009		0.068	**0.002**	0.009
Jones Sound	0.001	0.002	0.024	0.009	0.014		0.021	**0.002**
N Hudson Bay	0.007	0.014	0.028	0.007	0.018	0.014		0.014
Somerset Island	0.010	0.005	0.013	0.000	0.026	0.026	0.015	

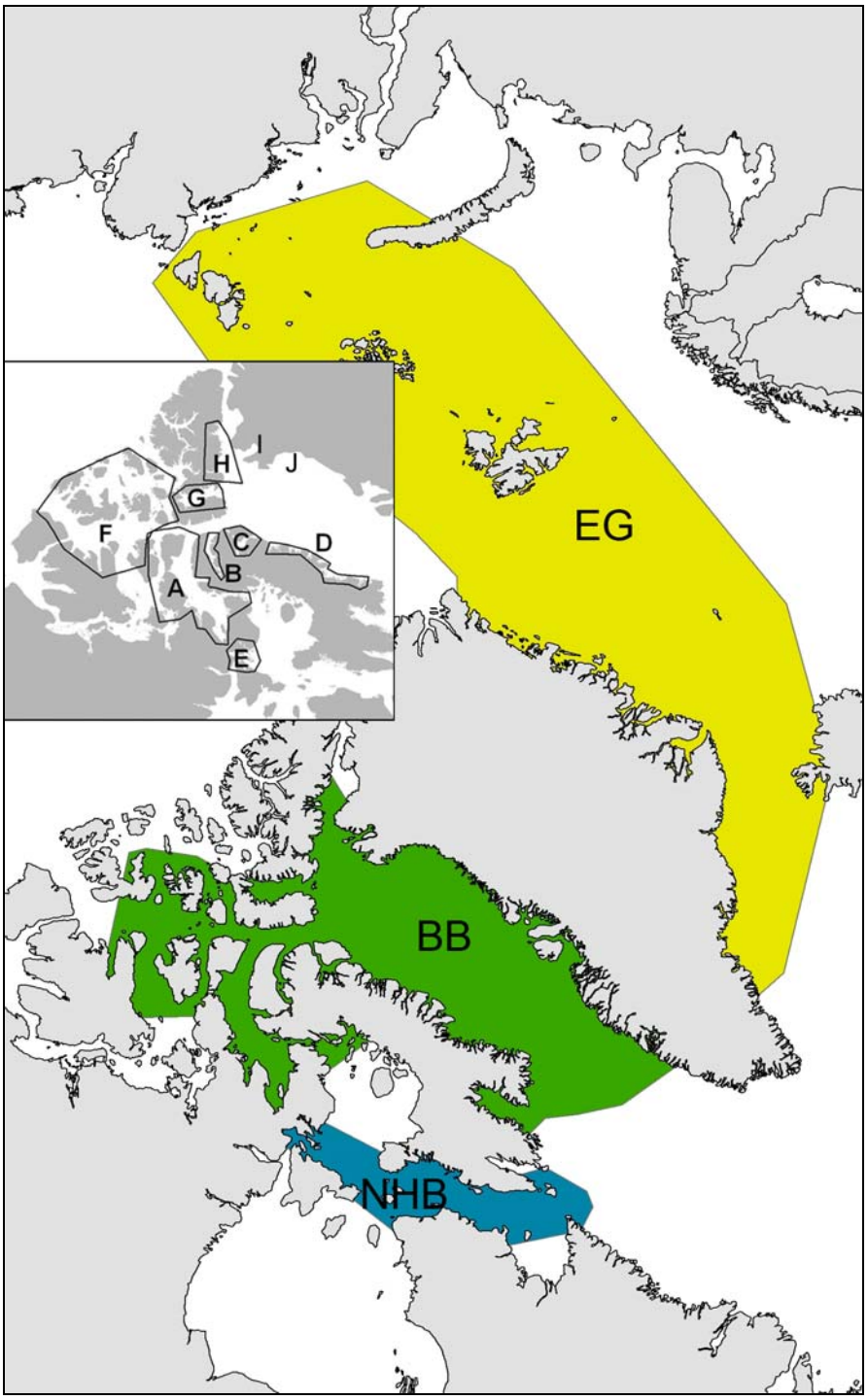

Figure 1. Approximate species range of narwhal in the Arctic with three populations indicated: Baffin Bay (BB), Northern Hudson Bay (NHB), and East Greenland (EG). Inset map indicates summer aggregations of narwhal in Canada and West Greenland: A: Somerset Island; B: Admiralty Inlet; C: Eclipse Sound; D: East Baffin Island; E: Northern Hudson Bay; F: Parry Channel; G: Jones Sound; H: Smith Sound (from DFO, 2010); I: Inglefield Bredning; J: Melville Bay (from Heide-Jørgensen et al. 2010).

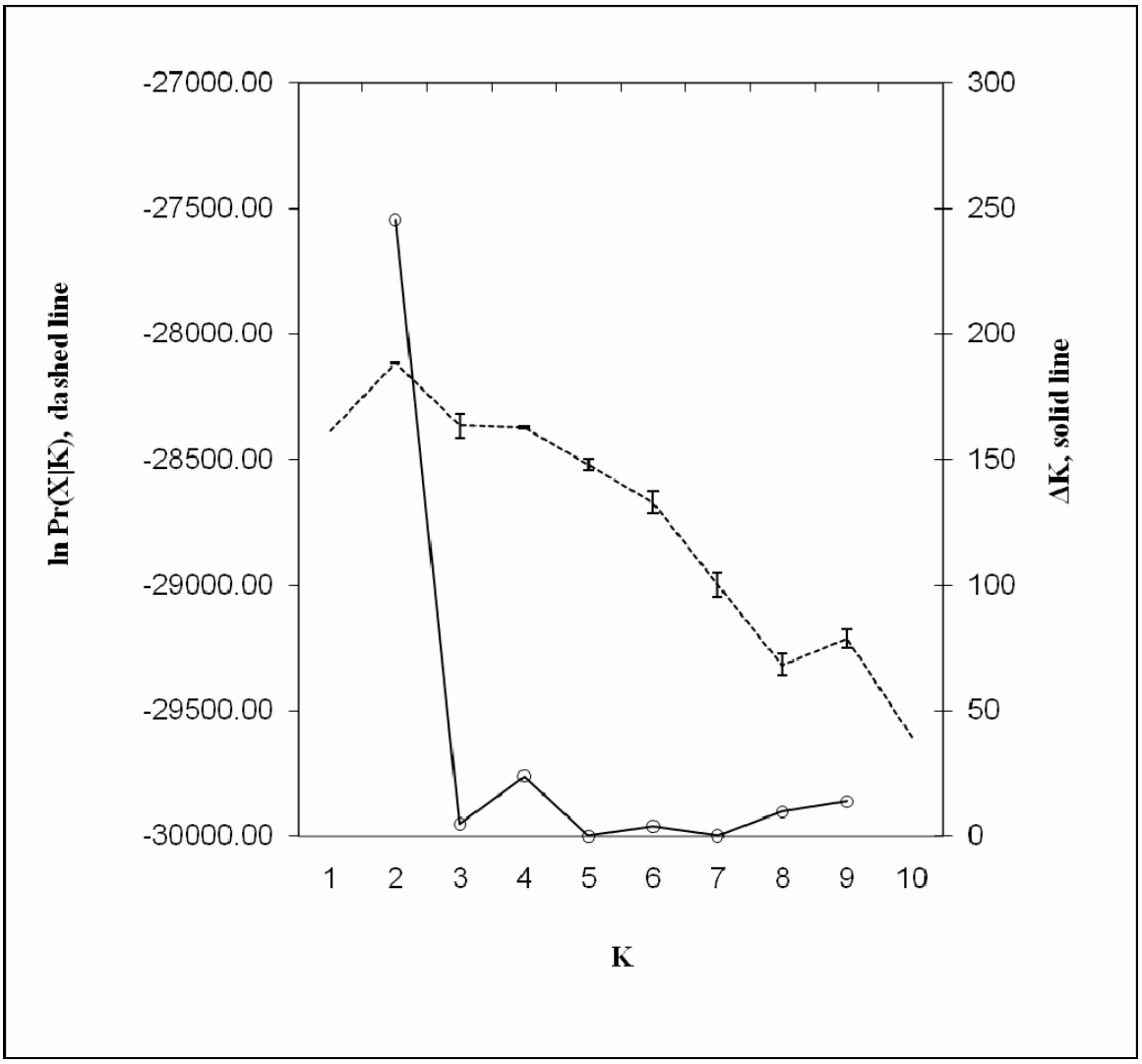

Figure 2. Results from individual based Bayesian analysis using the program STRUCTURE suggesting the presence of two genetic clusters for narwhal. Dashed line indicates log likelihood with error bars representing standard deviation calculated from three independent runs. Solid line represents K determined using the method of Evanno et al. (2005).

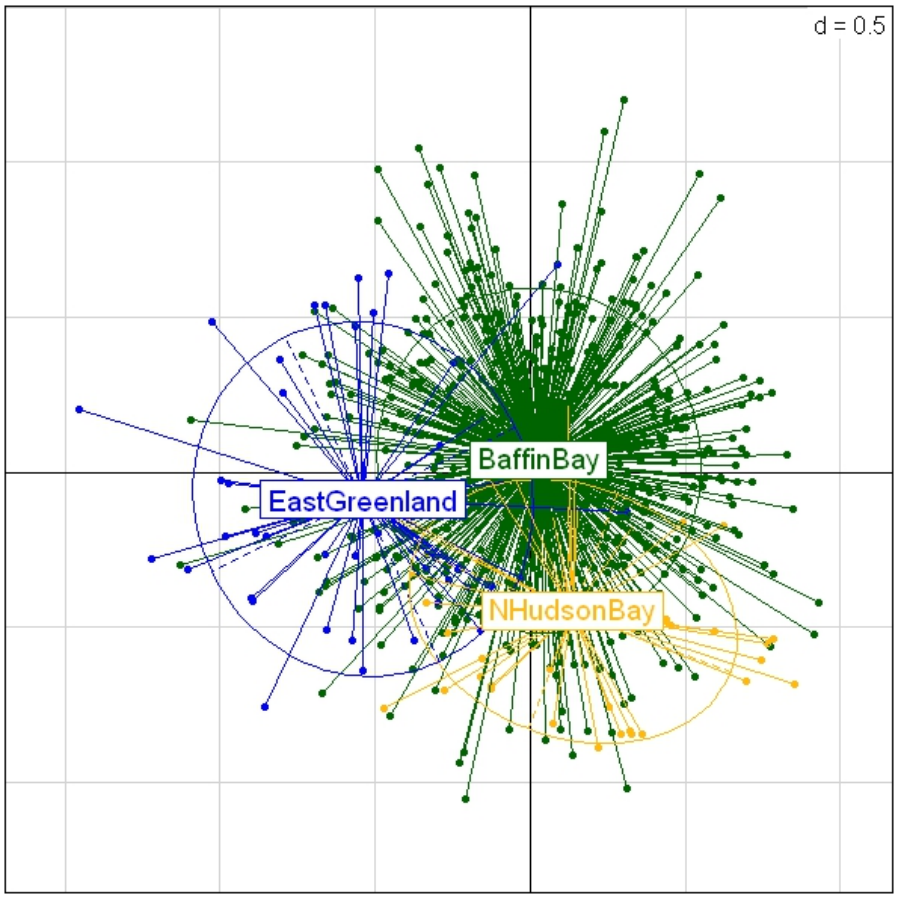

Figure 3. Scatter plot of the first two axis of variation resulting from an interclass PCA conducted on all narwhal samples from the three populations. Individual samples are represented by dots connected to the center of gravity for the population of origin. The scale of both the x (58.5%) and y (41.5%) axes are equal and the gridlines dimension are 0.5 units.

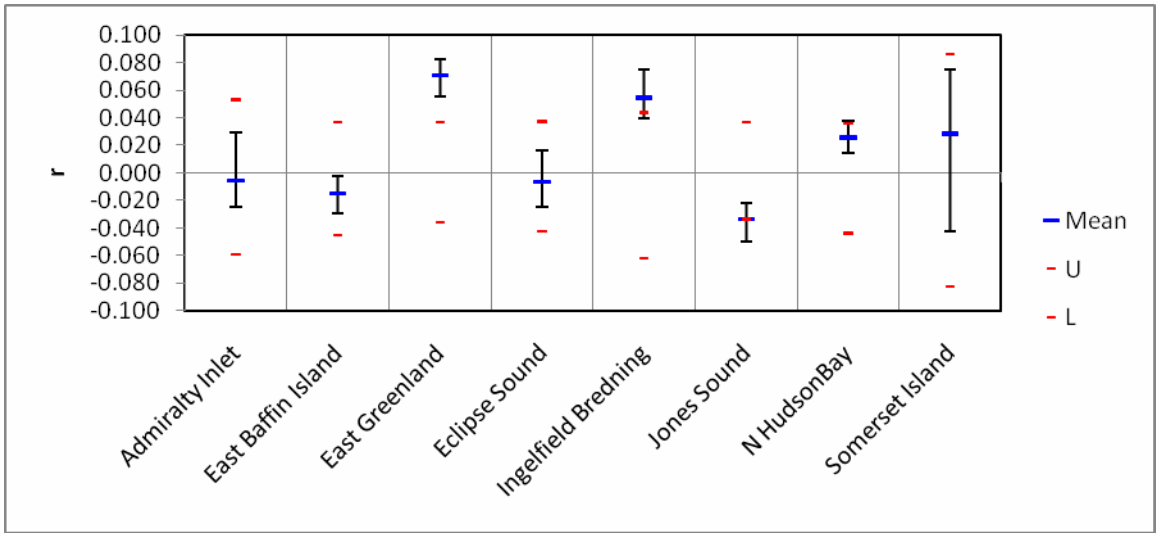

Figure 4. Mean within group relatedness of samples partitioned by summering stock. Confidence intervals (95%) obtained by permuting values around the null hypothesis of no differentiation from a relatedness of 0 (unrelated) are indicated with red bars. Confidence intervals (95%) around population means were obtained by bootstrap re-sampling.

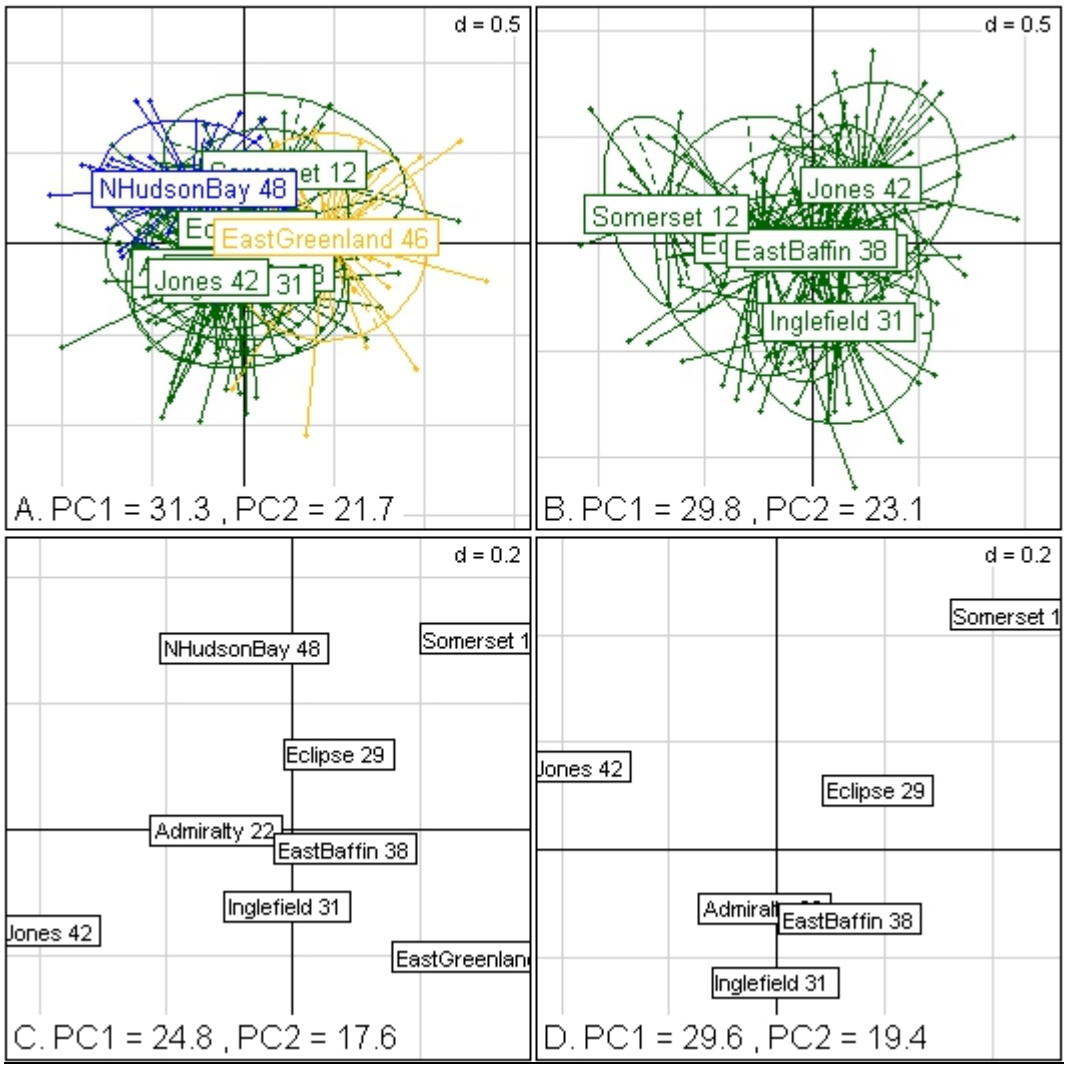

Figure 5. Multivariate analysis of narwhal samples showing considerable overlap in samples from different stocks. Percent variation for the x and y axes (respectively) is given after the figure label for each plot. Top row: Scatter plot of the first two principle components from the inter-class PCA analysis conducted on narwhal samples with stocks used as grouping variable. Dots represent individual samples and ellipses are 95% inertia ellipse for each group. Samples are colour coded to correspond to the population of origin; Baffin Bay (green), northern Hudson Bay (blue), and east Greenland (yellow).Bottom row: Scatter plots of summering stocks situated in the plane described by the first two axes of the correspondence analysis. Plot A and C include the Northern Hudson Bay population and East Greenland population while plots B and D omit them from the analysis. In all figures, the name of the group is followed by the sample size and the box containing this information is over the center of gravity for that group.

Appendix 1.Details regarding PCR multiplexing and loci pooling strategies for narwhal. Refer to text for general reaction recipes and thermocycling conditions.

Panel	Plex	Primers	Source	Program	T_{anneal}
1	A	FCB13 -vic	Buchanan et al. 1996	Step-up	48/53
		MK8 -ned	Krutzen et al. 2001		
	B	FCB4 - 6fam	Buchanan et al. 1996	General	63
		FCB5 -vic	Buchanan et al. 1996		
2	A	FCB1 -ned	Buchanan et al. 1996	Step-up	48/53
		GT39 -vic	Caldwell et al. 2002		
	B	FCB8 -ned *	Buchanan et al. 1996	General	60
		FCB10 -6fam	Buchanan et al. 1996		
3	A	Ev14 -ned	Valsecchi and Amos 1996	Step-up	48/55
		MK9 -6fam	Krutzen et al. 2001		
	B	FCB14 -vic	Buchanan et al. 1996	Step-up	48/55
4	A	FCB3 -ned	Buchanan et al. 1996	Step-down	56/51/48
		KWM12a -6fam	Hoelzel et al. 1998		
	B	KWM2a -vic *	Hoelzel et al. 1998	Step-up	48/55
5	A	FCB17r -vic *	Buchanan et al. 1996	General	55
		Ev37 -6fam	Valsecchi and Amos 1996		

*loci that were dropped from the analyses because of significant deviations from Hardy-Weinberg equilibrium.

Appendix 2. Thermocycling conditions for the general, step-up, and step-down protocols.

General		
Temp (°C)	time	cycle
95	11m	
95	45s	
T_{anneal}	45s	25
72	45s	
72	15m	
4	hold	

Step-down		
Temp (°C)	time	cycle
95	11m	
94	30s	
56	30s	5
72	45s	
94	30s	
51	30s	5
72	45s	
95	30s	
48	30s	35
72	45s	
72	15m	
4	hold	

Step-up		
Temp (°C)	time	cycle
95	11m	
95	45s	
48	45s	10
72	45s	
95	45s	
55	45s	25
72	45s	
72	15m	
4	hold	