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 2006/051 | **Document de recherche 2006/051**

Molecular genetic support of a single population of bowhead whales (*Balaena mysticetus*) in Eastern Canadian Arctic and Western Greenland waters

La génétique moléculaire en support à l'hypothèse d'une seule population de baleine boréale (*Balaena mysticetus*) dans les eaux de l'est de l'Arctique canadien et de l'ouest du Groenland

L.D. Postma[1], L.P. Dueck[1], M.P. Heide-Jørgensen[2], and S.E. Cosens[1]

[1]Fisheries and Oceans Canada, Central and Arctic Region
501 University Crescent, Winnipeg, Manitoba, R3T 2N6, Canada

[2]Greenland Institute of Natural Resources
Box 570, DK-3900 Nuuk, Greenland

ISSN 1499-3848 (Printed / Imprimé)

Canada

ABSTRACT

Molecular genetic relationships among bowhead whales (*Balaena mysticetus*) were examined and tested for population sub-structuring of samples collected in the waters of the Eastern Canadian Arctic and Western Greenland. An analysis of 15 nuclear DNA microsatellite loci was completed for 286 individual bowheads sampled at Pelly Bay, Igloolik, Repulse Bay and Pangnirtung in Nunavut, Canada and from Disko Bay in Western Greenland. An additional sample of whales from the Beaufort Sea representing the putative Bering-Chukchi-Beaufort (B-C-B) Sea stock/population was also included in the analysis. A Bayesian clustering (assignment) procedure was used to interpret the genetic profiles obtained from the samples in order to identify the inferred population structure detected from the observed genotypes. The analysis consistently revealed a lack of identifiable structure for these samples and the clustering analysis supports the results obtained from satellite tracking and aerial survey studies that indicate a single population of bowheads in the Eastern Canadian Arctic and Western Greenland. However, the small sample of whales from the Beaufort Sea was not clearly distinguished from the other samples in the analysis. Additional collaborative work is currently ongoing to increase the number of samples from the B-C-B population for comparison to the Eastern Canadian samples and to increase the number of loci examined in order to increase the power of the analysis.

RÉSUMÉ

Une comparaison de la génétique moléculaire des baleines boréales (*Balaena mysticetus*) a été entreprise pour vérifier la sous-structuration de la population à l'aide d'échantillons recueillis dans les eaux de l'est de l'Arctique canadien et de l'ouest du Groenland. L'analyse de 15 loci de microsatellites d'ADN nucléaire a été réalisée chez 286 baleines boréales échantillonnées dans les régions de Pelly Bay, Igloolik, Repulse Bay et Pangnirtung, au Nunavut (Canada), et de Disko Bay dans l'ouest du Groenland. Un échantillon additionnel de baleines de la mer de Beaufort, représentant le stock/population présumé des mers de Béring, des Tchouktches et de Beaufort (B-T-B) faisait aussi partie de l'analyse. Une méthode d'agrégation (affectation) bayésienne a été utilisée pour interpréter les profils génétiques tirés des échantillons, en vue de définir la structure présumée de la population, déduite des génotypes observés. L'analyse a révélé de façon constante un manque de structure identifiable au sein de ces échantillons et l'analyse typologique appuie les résultats obtenus grâce au suivi par satellite et aux relevés aériens qui montrent une seule population de baleines boréales dans l'est de l'Arctique canadien et dans l'ouest du Groenland. Toutefois, il n'a pas été possible de faire une distinction claire entre le petit échantillon de baleines de la mer de Beaufort et les autres échantillons de l'analyse. D'autres travaux coopératifs sont en cours pour élargir le nombre d'échantillons de la population B-T-B afin de les comparer aux échantillons de l'est canadien, et pour augmenter le nombre de loci examinés en vue d'accroître la valeur de l'analyse.

INTRODUCTION

The bowhead whale (*Balaena mysticetus*) is the largest of three Arctic species of whale inhabiting Canadian waters. Its very size, including a blubber layer which can measure 43 to 50 cm (Montague 1993), made the bowhead a primary target of the European whaling industry in the 18th, 19th and early 20th centuries (Reeves *et al.,* 1983; Ross, 1993). This intensive and unmanaged commercial hunting resulted in a reduction of numbers of bowheads from a minimum stock size of 452 in Hudson Bay to approximately 100 animals, and from a minimum of 11,759 animals in Davis Strait to approximately 1000 (Woodby and Botkin, 1993). These numbers are indicative of the numbers of animals remaining in the stock at the end of the peak harvest decade. As commercial hunting did continue for many years until collapse, the numbers of animals were reduced even further from these estimates of residual stock size.

The present day distribution of bowheads in the eastern Canadian Arctic spans the areas of Fury and Hecla Strait, Foxe Basin, northern Hudson Bay, Hudson Strait, the southeastern coast of Baffin Island, Baffin Bay, Lancaster Sound, Prince Regent Inlet, and the Gulf of Boothia, (Fig. 1) with fairly well documented areas of summering and wintering aggregations (Reeves *et al*., 1983; Reeves, 1991). Aerial surveys for bowhead whales were completed during 2002 – 2004 and covered the areas of southern Gulf of Boothia, west Foxe Basin, northwestern Hudson Bay (2002), Admiralty Inlet, east coast of Baffin Island (2003), Eclipse Sound, Admiralty Inlet and Barrow Strait (2004). From these surveys, the estimate for a combined number of bowheads in the eastern Canadian Arctic was found to be approximately 5000 individuals (Cosens *et al*., 2005).

Based largely on the absence of commercial hunting and these recent survey estimates, the Committee On the Status of Endangered Wildlife in Canada (COSEWIC) revised the designation of bowheads in the eastern Canadian Arctic from “Endangered” to “Threatened” in May, 2005 (COSEWIC, 2005). At this time, COSEWIC also split the Eastern Arctic bowhead whales into two populations, the Hudson Bay-Foxe Basin population (HB-FB) and the Baffin Bay-Davis Strait population (BB-DS). The two-stock hypothesis for bowheads has previously been suggested as a conservative approach for management purposes (Reeves and Mitchell, 1990), however it was recognized that the data in the past was insufficient for evaluating stock affinities. The International Whaling Commission (IWC) also recognizes two stocks of bowheads in the Northwest Atlantic (IWC 1978, 1992). Again, this designation is mostly based on the separation of summer distributions of bowheads where exchange of animals was thought to be unlikely.

More recent studies on the distributions and movements of bowhead whales in these areas directly challenge this two population model (Heide-Jørgensen and Finley, 2003; Cosens and Blouw, 2003; Heide-Jørgensen *et al*., 2003; Heide-Jørgensen *et al*., 2006). Bowheads photographed in Northern Foxe Basin during mid-summer revealed that the area is mostly utilized by females with young-of-the-year calves and juveniles (Cosens and Blouw, 2003). This suggests that this area is perhaps a nursery area and these animals are an age and sex-segregated portion of the population. Satellite tracking studies on bowheads have shown that they are capable of traveling long distances in relatively short periods of time (Heide-Jørgensen *et al*., 2003) and that whales tagged in the spring in Disko Bay, West Greenland eventually moved to the Hudson Strait wintering ground in mid-November (Heide-Jørgensen *et al*., 2006). Heide-Jørgensen *et al.* (2006)

suggest that movement of these Baffin Bay whales to a wintering ground that is thought to be used by bowheads from Foxe Basin and Hudson Bay brings into question the stock discreteness of these two putative populations. Other recent satellite tracking results have shown that HB-FB whales move into Prince Regent Inlet (an area identified as part of the BB-DS stock range) from Northern Foxe Basin (Dueck *et al*., 2006)

Satellite tracking studies do provide direct evidence on the degree of spatial and temporal heterogeneity of movement patterns across a species' range and the impact of environmental parameters on movement patterns (Bossart and Prowell, 1998). These are key elements for the detection of population structure within a species (Ihssen, 1981; Hartl, 1988). However, the role of gene flow is also an important aspect of defining populations (Pianka, 1988; Dawson and Belkhir, 2001) and barriers to gene flow may not always be obvious from direct observation (e.g. differences in habitat preference within the same geographic range, assortative mating, gender-specific migration rates) (Evanno *et al*., 2005; Dawson and Belkhir, 2001; Tiedemann *et al*., 2000). Reliable estimates of genetic population differentiation are also important considerations for conservation biology where the degree of genetic isolation of populations affects their evolutionary potential (Balloux and Lugon-Moulin, 2002).

In this study, molecular markers were used to assess genetic population structure among bowhead whales sampled in the eastern Canadian Arctic and western Greenland. A Bayesian approach was used to try and delineate clusters of individuals based on their genotypes at multiple loci. This method has many advantages over the more classical, frequentist statistical methods for analyzing genetics data (Shoemaker *et al*., 1999; Pearse and Crandall, 2004). It allows for a more direct approach to evaluating population models and uses the incorporation of prior information. It may also provide more straightforward interpretations of results that are better able to differentiate among alternative explanations for a given genetic signal.

MATERIALS AND METHODS

Sample collection

Biopsy samples of bowhead whale skin were obtained during post mortem examinations of beached and hunted animals and during biopsy sample programs targeting free-ranging whales (Table 1). The majority of samples were obtained during biopsy sampling programs of free-ranging bowhead whales in Foxe Basin (Igloolik), Repulse Bay, Pelly Bay and Cumberland Sound (Pangnirtung). Samples from Foxe Basin and Cumberland Sound were collected from June through August, while those in Repulse Bay were collected in August and September. Based on the assumption that northern Hudson Bay bowhead whales might be distinct from Foxe Basin whales, we avoided collection of samples from Repulse Bay earlier than August to preclude the possibility that whales sampled in Repulse Bay were actually Foxe Basin whales migrating through Repulse Bay. Samples were also obtained during an ongoing program for satellite tracking of bowhead whales from Western Greenland (Heide-Jørgensen *et al.,* 2003, 2006). A skin biopsy sample was taken for genetic analyses when whales were first approached for attachment of the satellite transmitter. These samples were collected in May in northwest Disko Bay.

All biopsy sampling was conducted from a two-person kayak, boat, or from an ice platform. The majority of whales sampled were initially approached by boat and either pursued and fired at from the boat, or alternatively, a kayak was launched from the floe edge or boat and used to approach the whales to within firing range of the biopsy system. Sampling from the floe edge was conducted opportunistically when bowhead whales were moving along or moving toward and diving beneath the floe edge.

Biopsy tips were cleaned and sterilized using a two stage process involving immersion and cleaning in hydrogen peroxide to dissolve and remove previous genetic material, and then in Betadyne antiseptic solution. Skin samples were transferred from the biopsy tip into vials containing a salt-saturated 20% dimethylsulfoxide (DMSO) solution (Seutin *et al*., 1991) within 1 to 15 minutes of extraction from the whale. These samples were then kept cool until genetic analyses were initiated.

In addition, a total of n=9 samples was obtained from free-ranging bowhead whales in the Mackenzie Delta area (Shingle Point and King Point) in 1990 and 1992. An additional sample was collected after a bowhead whale hunt in Shingle Point in 1996, bringing the sample total for this area to n=10.

DNA analysis

For earlier samples, total cellular DNA was extracted from bowhead whale skin using the methods described in Maiers *et al*. (1996) with some modifications. The bowhead whale skin has a very tough, rubbery texture after preservation and it required several weeks of incubation at 37°C and repeated additions of proteinase K (20 mg/mL) to digest the tissue to the point where it was suitable for extraction. Once this process was complete, in most samples sufficient quantities of DNA was recovered for analyses. More recent samples (2000 to present) were extracted using commercial DNA tissue extraction kits (DNeasy, Qiagen).

The sex of each of the animals sampled was determined using a PCR-based method for the identification of sex in cetaceans (Bérubé and Palsbøll 1996 or Shaw *et al*, 2003). This method amplifies ZFX-and ZFY-specific regions of nuclear DNA that results in a product that corresponds to a portion of the X chromosome and a product specific to the Y chromosome (if present). Separation and visualization of these products on an agarose gel allows for the reliable assignment of a sex.

Microsatellite analysis

A total of 15 microsatellite loci were analyzed using primers from a variety of sources (Table 2). The analysis was performed using Applied Biosystems' fluorescence-based technology on a 3100 genetic analyzer. The PCR and primer conditions were as described in the reference papers for each locus with some modifications to the annealing temperatures (Table 2), and were generally analyzed a single locus at a time. Allele sizes for genotypes were determined by co-running a size standard (ROX-HD350, Applied Biosystems) and using Genotyper software (Applied Biosystems). Designations were checked visually with the lanes aligned by scan. Any errors in allele sizing were corrected using a comparison to a set of reference samples that were analyzed with every run of samples.

Descriptive statistics

Geographic sample groups that were statistically compared for the purposes of this document were: 1. Mackenzie Delta (Shingle Point and King Point); 2. Gulf of Boothia (Pelly Bay); 3. Hudson Bay-Foxe Basin (Igloolik and Repulse Bay); 4. Davis Strait (Pangnirtung); and 5. Baffin Bay (Western Greenland) (Figure 2).

The numbers of alleles, observed heterozygosity, and expected heterozygosity were generated using an in-house descriptive statistics program written with Visual Basic (Brigitte de March, pers. comm..) and using POPGENE ver 1.31 (http://www.ualberta.ca/~fyeh/).

Homogeneity of allele distributions for all pairs of sample groups (the null hypothesis being "the allelic distribution is identical across populations") was tested using an unbiased estimate of the *P*-value of the probability test or Fisher exact test, when possible (Raymond and Rousset, 1995). Each sample group was also tested at each locus for departure from Hardy-Weinberg equilibrium (HWE) using the U-test (Rousset and Raymond, 1995) with the hypothesis of heterozygote deficiency. These tests were performed using GENEPOP ver 3.4 (Raymond and Rousset, 1995).

Population structure

The microsatellite dataset was checked for genotyping errors due to null alleles (nonamplified alleles), short allele dominance (large allele dropout), scoring errors, and typographic errors using the software program MICRO-CHECKER (van Oosterhout *et al*., 2004). Though adjustment of allele and genotype frequencies to correct for null alleles is also possible with this software, loci that were found to be consistently out of Hardy-Weinberg equilibrium were eliminated from the final analyses. The final proof-read dataset was converted to appropriate formats for further analysis using CONVERT (Glaubitz, 2004) which is designed to manipulate codominant, diploid genotypic data and also provides summaries of allele frequencies with the identification of private alleles.

Assignment tests were used to identify genetic structure and to assign individuals to their likely population of origin using STRUCTURE ver. 2.0 (Pritchard *et al*., 2000). This program uses a Bayesian clustering method that takes a sample of genotypes and uses the assumption of Hardy-Weinberg and linkage equilibrium within sub-populations to determine (i) the most likely number of inferred populations, *K*, that best fits the data given the observed genotypes, (ii) the proportion of each predefined sample group contributed by each inferred population, and (iii) the individual assignments for each sample indicating the proportion of each individual animal's genotype contributed by each inferred population. With no prior information on population sampling design, the results of this analysis provide an estimate of the number of subpopulations, each of which contains a set of individual genotypes that are in Hardy-Weinberg equilibrium. In the input file, the individual samples are identified to a predefined group according to the geographical area from which the samples were obtained (Fig. 2) and this information can later be correlated to the clustering results.

The results generated from analyses using STRUCTURE were based on simulations from one to five (K = 1-5) inferred populations to allow for the possibility that each of the geographic sample groups may be a separate population. A variety of burn-in periods and number of iterations of the Markov chain Monte Carlo (MCMC) simulations (which is an iterative method used for approximating the posterior probabilities of

population genetic parameters for larger datasets and for finding likelihood maixima) were used. These trials allow for the best selection of run length to minimize the effect of the starting configuration and to get accurate parameter estimates. In this document, the burn-in period was 1,000,000 iterations with 500,000 iterations of the MCMC simulation. Other prior information used was the assumption of an admixture model (that individuals may have mixed ancestry – it is a flexible model likely most reflective of real populations) with a uniform prior for α (the degree of admixture) and an initial α = 1.0. The correlated allele frequency model (Falush *et al*., 2003) was used as the allele frequencies in the different groups are likely to be similar due to the very high likelihood of migration and shared ancestry. The prior mean was set at 0.01 which corresponds to very low levels of subdivision (which is expected for these bowhead samples given the satellite tracking results). Lambda, the parameter of the distribution of allelic frequencies, was set to one as recommended by Pritchard *et al*. (2000). This is thought to be best for situations where allele frequencies are not skewed and when other parameters are being estimated.

As different runs of STRUCTURE can produce different likelihood values, each data set for the individual population simulations was analyzed for 20 runs in order to assess the amount of variation of the likelihood for each *K*.

RESULTS

Descriptive statistics

Nuclear DNA microsatellite analysis was performed at 15 loci using primers from several sources (Table 2). The numbers of alleles detected at each locus ranged from 4 – 18 and the microsatellite diversity of individual loci was lowest for EV37 (0.000 – 0.335) and highest for RW34 (0.758 – 0.890). The lowest numbers of microsatellite alleles were found in Repulse Bay (68 alleles), Pelly Bay (72 alleles) and the W. Arctic (74 alleles) samples and the highest number in the Igloolik samples (118 alleles) (Table 3). Repulse Bay, Pelly Bay and W. Arctic (0.600; 0.608; 0.607) diversities were lower than those from Igloolik, Pangnirtung and W. Greenland (0.657; 0.643; 0.658).

A test for goodness of fit to Hardy Weinberg Equilibrium revealed significant deviations from the HWE at 16 of 90 locus X location tests (marked as **bold** on Table 3). No one locus consistently deviated from HWE, however, Igloolik had the most loci not in HWE (5 out of 15 loci), followed by Pangnirtung (4 out of 15 loci). Inbreeding coefficients (F_{is}) were positive for all except one locus in the Igloolik samples (14 out of 15 loci) and for most loci in samples from Pelly Bay (9 out of 15 loci) and W. Greenland (9 out of 15 loci).

Population structure

The use of MICROCHECKER revealed the presence of null alleles at loci TV11 and GATA28. The data for these loci were eliminated from further analyses.

Running the simulations with various lengths of burn-in time and MCMC iterations did not change significantly the results. However, different runs did produce different likelihood values (data not shown). For 20 runs at *K* = 1, these values ranged from -8692.7 to -8683.0. However, most of the values for this inferred cluster were -8683.4 +/- 0000.4. The 20 runs for *K* = 2 to 5 had lower likelihood values (data not shown). Values for α

Individual assignments of samples to population clusters are shown in Figure 3. These graphs are representations of the estimated membership coefficients for each individual, in each cluster. Each individual in the data set is represented by a single vertical line, which is partitioned into *K* colored segments that represent that individual's estimated membership fraction in each of the *K* inferred clusters. The samples are grouped by sampling location and are also indicated on the bar graph. The graph for *K* = 1 is obvious – since there is only one inferred cluster, all individuals are assigned 100% to that cluster. At *K* = 2, varying proportions of membership to the clusters is found for each individual sample and range from 0.1/0.9 to 0.8/0.2. This trend continued and analysis for each value of *K* = 1 - 5 revealed the individual samples to be fairly admixed.

Table 4 summarizes the overall proportion of membership of the samples from each location for the *K* = 5 graph. Each location had the highest proportion of membership to different clusters (with the exception of S.E. Baffin Island and W. Greenland), however, the sample assignment for all locations is roughly symmetric (~1/K in each cluster). This result was the same for all other simulations of different numbers of inferred populations (data not shown).

DISCUSSION AND CONCLUSIONS

The Bayesian analysis used in STRUCTURE is a model-based clustering method for using multilocus genotype data to infer population structure and assign individuals to populations (Pritchard *et al*., 2000). The main modeling assumptions are Hardy-Weinberg equilibrium within populations and complete linkage equilibrium between loci within populations. Thus, the quality of data is critical to the success of this analysis. In this study, every attempt was made to ensure that alleles in the dataset were entered correctly and that the loci used did not display evidence of scoring errors, null alleles, or linkage disequilibrium.

The model choice criterion in STRUCTURE to detect the true value of *K* is an estimate of the posterior probability of the data for a given *K*, Pr(X/*K*) (Pritchard *et a*l., 2000). This value, called 'Ln P(D)' in the output of STRUCTURE is obtained by computing the log likelihood of the data at each step of the MCMC, then an average is calculated and then half their variance subtracted from the mean. The true number of populations (*K*) is often identified using the maximal value of Ln P(D) (or L(*K*)). Using this criterion to interpret the results in this study, the value of *K* suggested by this data is *K* = 1.

Pritchard and Wen (2004) also point out that when there is no population structure, the results will show that the proportion of sample assigned to each population is roughly symmetric and most individuals admixed. Conversely, if some individuals are strongly assigned to one population or another, and if the proportions assigned to each group are asymmetric, then this is a strong indication that there is real population structure. The results in this study indicate a lack of population structure (Fig. 3 and Table 4). If the number of runs is limited to fewer simulations for *K* (e.g. *K* = 1 to 3), the results are the same and the proportions of assignments become even more symmetric. These results would suggest that the bowheads in Hudson Bay-Foxe Basin and Baffin Bay-Davis Strait are from a single population.

It is important to note some of the potential limitations of these analyses. First, the use of prior information for the analyses introduces a subjective quality to the method (Huelsenbeck *et al*., 2002). The results should be examined with respect to the sensitivity of the results to the priors used and the reliability of the MCMC approximations. This can be evaluated by performing many runs of the simulations using different settings of the priors and different (especially longer) lengths of MCMC iterations.

The ability of STRUCTURE to detect clusters of individuals at different levels of *K* when dispersal among clusters is more intense is not clear (Evanno et al., 2005). In other words, this type of analysis is more effective when levels of genetic differentiation are relatively strong. It is possible, however, to compensate for weak genetic signals by increasing the number of loci used for the genotyping (Rosenberg *et al*., 2002). In general, it is recommended that at least 12 – 15 highly variable loci should be genotyped in at least 15 – 20 individuals per hypothesized population. The addition of several more loci of information to the dataset in this document would strengthen the analysis, however, the combination of number of loci and numbers of samples make the current results meaningful. As the amount of data added to a Bayesian analysis increases, the influence of the prior beliefs on the posterior distribution decreases (Huelsenbeck *et al*., 2002).

The composition of the samples may also be important for this method. Ideally, the individuals sampled should belong to the same generation (or to the same cohort for organisms with overlapping generations) as allele frequencies vary not only over space, but also over time (Balloux and Lugon-Moulin, 2002). Given the long life span of bowhead, the data set in this study likely contains individuals from several generations. Analysis of only samples from adult animals would address this issue, however this would drastically reduce the sample sizes. This introduces another problem, as the power of this approach directly depends on sample size (Balloux and Lugon-Moulin, 2002).

Ultimately, Pritchard *et al*. (2000) stress that care should be taken in the interpretation of the inferred value of *K*. However, their methods can produce highly accurate clustering and sensible choices for *K*, especially when used with other biologically meaningful information. Given the context of information revealed from satellite tracking studies and aerial surveys of bowheads in the Eastern Canadian Arctic and Western Greenland, the molecular genetic results presented in this study support the model of a single stock of whales in the Hudson Bay-Foxe Basin and Baffin Bay-Davis Strait areas.

ACKNOWLEDGEMENTS

The authors would like to thank Brad Parker, the Repulse Bay HTO, and the Pelly Bay HTO for the collection of some of the samples used in this study. Logistic support was provided by the communities of Igloolik, Repulse Bay, Pangnirtung and Pond Inlet during the sampling of free-ranging animals. Thank-you to Margaret Friesen, Denise Tenkula, Rob Bajno and Melissa McPherson for providing laboratory assistance. Funding for this project has been provided by Fisheries and Oceans Canada; in the Eastern Arctic by the Nunavut Implementation Fund (NIF) and the Species at Risk Recovery Fund (SARCEP); and in the Western Arctic by the Fisheries Joint Management Committee (FJMC).

REFERENCES

Balloux, F. and N. Lugon-Moulin. 2002. The estimation of population differentiation with microsatellite markers. Molecular Ecology 11:155-165.

Bérubé, M. and P. Palsbøll. 1996. Identification of sex in cetaceans by multiplexing with three ZFX and ZFY specific primers. Molecular Ecology 5: 283-287.

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

Cosens, S.E. and A. Blouw. 2003. Size- and age-class segregation of bowhead whales summering in Northern Foxe Basin: a photogrammetric analysis. Marine Mammal Science 19:284-296.

Cosens, S.E., H. Cleator and P. Richard. 2005. Results of aerial surveys of bowhead whales (*Balaena mysticetus*) in the Eastern Canadian Arctic in 2002, 2003 and 2004. DFO Canadian Science Advisory Research Document 2005/006.

COSEWIC 2005. COSEWIC assessment and update status report on the bowhead whale *Balaena mysticetus* in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa. Viii + 51 pp. (www.sararegistry.gc.ca/status/status_e.cfm).

Davis, R.A. and W.R. Koski. 1980. Recent observations of the bowhead whale in the eastern Canadian high Arctic. Report of the International Whaling Commission 30: 439-444.

Dueck, L.P., M.P. Hiede-Jørgensen, M.V. Jensen, and L.D. Postma. 2006. Update on investigations of bowhead whale (*Balaena mysticetus*) movements in the eastern Arctic, 2003-2005, based on satellite-linked telemetry. DFO Can. Sci. Advis. Sec. Res. Doc. 2006/050

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

Falush, D., M. Stephens, and J.K. Pritchard. 2005. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567-1587.

Finley, K.J. 1990. Isabella Bay, Baffin Island: an important historical and present-day concentration area for the endangered bowhead whale (*Balaena mysticetus*) of the eastern Canadian Arctic. Arctic 43: 137-152.

Guo, S.W. and E.A. Thompson. 1992. Performing exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48: 361-371.

Haig, S. 1998. Molecular contributions to conservation. Ecology 79: 413-425.

Hartl, D.L. 1988. A Primer of Population Genetics. Sinauer Associates, Inc., Sunderland, Massachusetts, 305 pp.

Hay, K. 1997. Inuit bowhead knowledge study interim report: Northwest Hudson Bay and High Arctic. Nunavut Wildlife Management Board, Iqaluit, NT.

Heide-Jørgensen, M.P., K.L. Laidre, Ø. Wiig, M.V. Jensen, L. Dueck, L.D. Maiers, H.C. Schmidt and R.C. Hobbs. 2003. From Greenland to Canada in ten days: tracks of bowhead whales, Balaena mysticetus, across Baffin Bay. Arctic 56: 21-31.

Heide-Jørgensen, M.P., K.L. Laidre, M.V. Jensen, L. Dueck and L.D. Postma. 2006. Dissolving stock discreteness with satellite tracking: bowhead whales in Baffin Bay. Marine Mammal Science 22: 34-45.

Huelsenbeck, J.P., B. Larget, R.E. Miller and F. Ronquist. 2002. Potential applications and pitfalls of Bayesian Inference of Phylogeny. Systematic Biology 51: 673-688.

Maiers, L.D., M.K. Friesen, A.V. Wiens and J.W. Clayton. 1996. Use of DNA microsatellites in beluga whale (*Delphinapterus leucas*) population genetics. Canadian Technical Report of Fisheries and Aquatic Sciences 2115, v + 17pp.

Milligan, B.G., J. Leebens-Mack and A.E. Strand. 1994. Conservation genetics: beyond the maintenance of marker diversity. Molecular Ecology 3: 423-435.

Mitchell, E.D. and R.R. Reeves. 1982. Factors affecting abundance of bowhead whales *Balaena mysticetus* in the eastern Arctic of North America, 1915-1980. Biological Conservation 22: 59-78.

Montague, J.J. 1993. Introduction, pp. 1-21. *In* J.J. Burns, J.J. Montague and C.J. Cowles [eds.], Allen Press Inc., Lawrence, KS.

Moore, S. and R.R. Reeves. 1993. Distribution and movement, pp. 313-386. *In* J.J. Burns, J.J. Montague and C.J. Cowles [eds.] The Bowhead Whale, Allen Press Inc., Lawrence, KS.

Palsbøll, P., M. Bérubé, A.H. Larsen and H. Jørgensen. 1997. Primers for the amplification of tri- and tetramer microsatellite loci in baleen whales. Molecular Ecology 6: 893-895.

Pearse, D.E. and K.A. Crandall. 2004. Beyond Fst: analysis of population genetic data for conservation. Conservation Genetics 5: 585-602.

Pianka, E.R. 1988. Evolutionary Ecology. Harper and Row, Publishers, New York, 397 pp.

Pritchard, J.K. and W. Wen. 2004. Documentation for STRUCTURE software: Version 2. Department of Human Genetics, University of Chicago.

Raymond M. and F. Rousset. 1995. GENEPOP (version 3.1b): population genetics software for exact tests and ecumenicism. Journal of Heredity 86: 248-249.

Reeves, R., E. Mitchell, A. Mansfield and M. McLaughlin. 1983. Distribution and migration of the bowhead whale, *Balaena mysticetus*, in the eastern North American Arctic. Arctic 36: 5-84.

Reeves, R.R. and E. Mitchell. 1990. Bowhead whales in Hudson Bay, Hudson Strait, and Foxe Basin: a review. Naturalist Canadien (Rev. Ecol. Syst.) 117: 25-43.

Reeves, R.R. 1991. The bowhead whale. Underwater World, Communications Directorate, Department of Fisheries and Oceans, Ottawa, Canada.

Reeves, R.R. and S.E. Cosens. 2003. Historical population characteristics of bowhead whales (*Balaena mysticetus*) in Hudson Bay. Arctic 56: 283-292.

Rosenberg, N.A., T. Burke, K. Elo, M.W. Feldman, P.J. Freidlin, M.A.M. Groenen, J. Hillel, A. Maki-Tanila, M. Tixier-Boichard, A. Vignal, K. Wimmers, and S. Weigend. Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. Genetics 159: 699-713.

Ross, W.G. 1993. Commercial whaling in the North Atlantic sector, pp. 511-561. *In* J.J. Burns, J.J. Montague and C.J. Cowles [eds.] The Bowhead Whale, Allen Press Inc., Lawrence, KS.

.

Rousset, F. and M. Raymond. 1995. Testing heterozygote excess and deficiency. Genetics 140: 1413-1419

Royce, W.F. 1972. Introduction to the Fishery Sciences. Academic Press, New York, 351 pp.

Seutin G., B.N. White and P.T. Boag. 1991. Preservation of avian blood and tissue samples for DNA analyses. Canadian Journal of Zoology 69: 82-90.

Shaw, C.N., P.J. Wilson and B.N. White. 2003. A reliable molecular method of gender determination for mammals. Journal of Mammalogy 84: 123-128.

Shoemaker, J.S., I.S. Painter and B.S. Weir. 1999. Bayesian statistics in genetics: a guide for the uninitiated. Trends in Genetics 15: 354-358.

Tiedemann, R., O. Hardy, X. Vekemans, and M.C. Milinkovitch. 2000. Higher impact of female than male migration on population structure in large mammals. Molecular Ecology 9: 1159-1163.

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

Waldick, R.C., M.W. Brown and B.N. White. 1999. Characterization and isolation of microsatellite loci from the endangered North Atlantic right whale. Molecular Ecology 8: 1763-1765.

Waser, P.M., and Strobeck, C. 1998. Genetic signatures of interpopulation dispersal. Trends Ecol. Evol. 13: 43-44.

Woodby, D.A. and D.B. Botkin. 1993. Stock sizes prior to commercial whaling, pp. 387-407. *In* J.J. Burns, J.J. Montague and C.J. Cowles [eds.] The Bowhead Whale, Allen Press Inc., Lawrence, KS.

Table 1. Bowhead sample collection information.

Major Sample Group	Sample Location	Minor Sample Group	Year(s)	n samples with Haplotypes	n samples with Microsatellites	n samples with both	Number of Females	Number of Males	Season collected*
1	Beaufort Sea (Mackenzie Delta)	1	1990	2	2	2	1	1	unkn
		2	1992	7	7	7	3	4	unkn
			1996	1	1	1	0	1	unkn^
2	Pelly Bay	3	2000	1	1	1	0	1	unkn
			2001	2	2	2	1	1	unkn
		4	2002	5	5	5	2	3	Sept
3	Igloolik	5	1994	1	1	1	1	0	unkn^
			1995	13	13	13	9	4	4Jul-6Jul
		6	1996	20	20	20	9	11	3Jul-9Jul
			1997	1	1	1	1	0	unkn^
		7	2001	42	42	42	19	23	30Jun-6Jul
		8	2002	65	65	65	32 (one unknown)	32	1Jul-15Jul
		9	2003	31	33	31	24	9	unkn
4	Repulse Bay	10	1997	4	5	4	4	1	Aug, Sept
		11	1998	4	0	0	1	3	Sept
		12	2000	4	4	4	3	1	Sept
		13	2001	4	4	4	0	4	Sept
5	Pangnirtung	14	1997	25	25	25	8	17	unkn
		15	2002	10	10	10	7 (one unknown)	2	unkn
6	West Greenland	16	2000	7	7	7	2	5	
		17	2001	13	13	13	7	6	28Apr-8May
		18	2002	10	10	10	6	4	4May-13May
		19	2003	9	11	9	11	0	4May-18May
Totals:	286 bowhead			281	282	277	153	134	

* Samples were collected as a biopsy of a free-ranging animal using a crossbow or during satellite tag attachment, unless indicated otherwise.

^ Sample collected from harvested animal.

Table 2. Details of the 15 microsatellite loci based on all individuals (n=286) analyzed in this study.

Microsatellite Locus[1]	Annealing Temperature	Reference	n Alleles	Range of Sizes (base pairs)
EV1*Pm*	48°C / 53°C	Valsecchi & Amos, 1996	13	137 - 195
EV37*Mn*	48°C / 53°C	Valsecchi & Amos, 1996	5	181 - 195
EV76*Mn*	48°C / 53°C	Valsecchi & Amos, 1996	4	152 - 162
EV104*Mn*	48°C / 53°C	Valsecchi & Amos, 1996	9	147 - 165
TexVet11	64°C /59°C / 54°C	Rooney *et al*., 1999	7	242 - 256
TexVet16	62°C / 57°C / 52°C	Rooney *et al*., 1999	6	184 - 196
TexVet17	56°C / 51°C / 46°C	Rooney *et al*., 1999	11	192 - 214
rw18	48°C / 53°C	Waldick *et al*., 1999	5	187 - 195
rw31	48°C / 53°C	Waldick *et al*., 1999	6	114 - 132
rw34	50°C / 55°C	Waldick *et al*., 1999	18	84 - 128
rw48	50°C / 55°C	Waldick *et al*., 1999	10	129 - 149
DlrFCB4	48°C / 53°C	Buchanan *et al.*, 1996	18	150 - 206
DlrFCB11	48°C / 53°C	Buchanan *et al*., 1996	6	120 - 130
GATA028	48°C / 53°C	Palsboll *et al*., 1997	9	118 - 186 (tetramer)
GATA098	48°C / 53°C	Palsboll *et al.*, 1997	6	86 - 110 (tetramer)

[1] The 15 loci are designated as listed in the reference (usually according to species and/or by the initials of the person who developed the primers; or, in Palsboll *et al*., as the repeat unit and locus identifier.).

Table 3. Microsatellite information for genetic analysis of bowhead samples. N = number of individual samples scored; A = number of alleles; H_o = observed heterozygosity; H_e = expected heterozygosity; F_{is} = inbreeding coefficient and test for goodness of fit to Hardy Weinberg Equilibrium (deviations from HWE are **bolded**).

Sample Group	Variable	EV104	TV16	GATA28	EV1	EV37	EV76	FCB4	RW18	RW31	RW48	TV11	TV17	RW34	FCB11	GATA98	overall
Western Arctic	N	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	A	6	2	4	3	3	3	8	3	4	7	3	8	12	4	4	74
	H_o	0.7778	0.6000	0.6667	0.6000	0.6000	0.1000	0.8571	0.5000	0.4444	1.0000	0.3333	0.9000	1.0000	0.5000	0.5714	0.6301
	H_e	0.8086	0.4200	0.5139	0.6150	0.5800	0.3350	0.8061	0.4600	0.4444	0.8150	0.4861	0.8200	0.8900	0.5750	0.5306	0.6067
	F_{is}	0.0620	-0.3970	-0.0830	0.0260	-0.0240	**0.4070**	-0.0100	-0.0950	-0.0190	-0.1200	0.3290	-0.0200	-0.0400	0.1730	-0.0370	
Pelly Bay	N	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
	A	6	5	7	4	4	2	6	3	4	5	3	7	8	5	3	72
	H_o	0.5000	0.5000	0.8571	0.8750	0.5714	0.2500	0.6250	0.2857	0.5000	0.7500	0.2857	0.8571	0.8750	0.8571	0.4000	0.5993
	H_e	0.7578	0.6875	0.8265	0.7109	0.6633	0.2188	0.5781	0.2551	0.5625	0.7656	0.4388	0.8163	0.8438	0.6531	0.3400	0.6079
	F_{is}	**0.3210**	**0.3320**	0.0230	-0.1430	0.1070	-0.0820	-0.0140	-0.0070	0.2860	0.0890	0.3360	0.0830	0.0200	-0.1100	-0.0160	
Igloolik	N	175	175	175	175	175	175	175	175	175	175	175	175	175	175	175	175
	A	9	6	9	8	4	4	14	5	5	9	7	10	16	6	6	118
	H_o	0.8047	0.6587	0.8286	0.6982	0.4821	0.1951	0.7029	0.5829	0.5930	0.8229	0.4368	0.7914	0.7412	0.5679	0.5814	0.6325
	H_e	0.8181	0.6036	0.8636	0.7193	0.5753	0.1912	0.6975	0.5919	0.5984	0.7760	0.5391	0.7977	0.7808	0.6887	0.6070	0.6565
	F_{is}	0.0200	**0.2660**	**0.0650**	0.0090	0.0480	0.0120	0.0060	0.0130	0.0030	-0.0360	**0.1600**	**0.0850**	0.0010	**0.0750**	0.0120	
Repulse Bay	N	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
	A	6	4	6	8	3	1	6	3	4	4	2	5	8	3	5	68
	H_o	1.0000	0.7273	0.6154	0.8462	0.6667	0.0000	0.5833	0.6667	0.7692	0.7273	0.4545	0.5455	0.6923	0.5455	0.7143	0.6369
	H_e	0.7574	0.6157	0.7574	0.7722	0.5312	0.0000	0.6528	0.5313	0.5680	0.7273	0.3512	0.6983	0.8077	0.4835	0.7449	0.5999
	F_{is}	-0.1640	-0.1250	0.1680	-0.0440	-0.1310	**n/a**	0.0630	-0.1980	-0.2030	0.1330	-0.2590	0.1810	0.0830	-0.0950	0.1920	
Pangnirtung	N	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35
	A	9	3	8	12	3	3	8	4	4	8	6	7	12	4	6	97
	H_o	0.8824	0.4000	0.6286	0.9412	0.6286	0.2500	0.4412	0.5882	0.4118	0.7714	0.4706	0.7879	0.6000	0.7188	0.5926	0.6075
	H_e	0.8157	0.5159	0.8082	0.7872	0.5629	0.2222	0.6631	0.5887	0.4035	0.7318	0.6306	0.7842	0.7584	0.6982	0.6797	0.6434
	F_{is}	-0.0610	**0.2480**	**0.1300**	-0.0630	-0.0870	-0.0560	**0.1420**	-0.0100	-0.0150	-0.0160	**0.1350**	-0.0110	0.0500	-0.0520	0.1540	
W. Greenland	N	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41
	A	8	3	8	10	5	3	8	5	4	7	4	9	13	4	6	97
	H_o	0.9000	0.4634	0.5854	0.9000	0.5122	0.3415	0.6250	0.5641	0.5128	0.7750	0.3902	0.7692	0.9000	0.6970	0.6571	0.6395
	H_e	0.8087	0.5446	0.8096	0.7850	0.5181	0.2965	0.7331	0.6160	0.4895	0.7656	0.5535	0.8008	0.8269	0.6515	0.6747	0.6583
	F_{is}	-0.0390	0.1170	**0.2230**	-0.0480	0.0030	-0.0880	**0.1040**	0.0930	-0.0180	-0.0120	**0.3420**	0.0170	-0.0510	0.0840	0.0390	

Table 4. Proportion of membership of each pre-defined population in each of the 5 clusters. Bolded values indicate the cluster to which the highest proportion of each population was assigned.

Given population	Inferred Clusters 1	2	3	4	5	Number of Individuals
W. Arctic	**0.291**	0.087	0.202	0.222	0.198	10
Pelly Bay	0.205	0.150	**0.265**	0.179	0.202	8
HB-FB	0.187	**0.230**	0.199	0.196	0.189	191
S.E. Baffin Island	0.222	0.132	0.195	0.216	**0.235**	35
W. Greenland	0.210	0.164	0.194	0.203	**0.228**	41

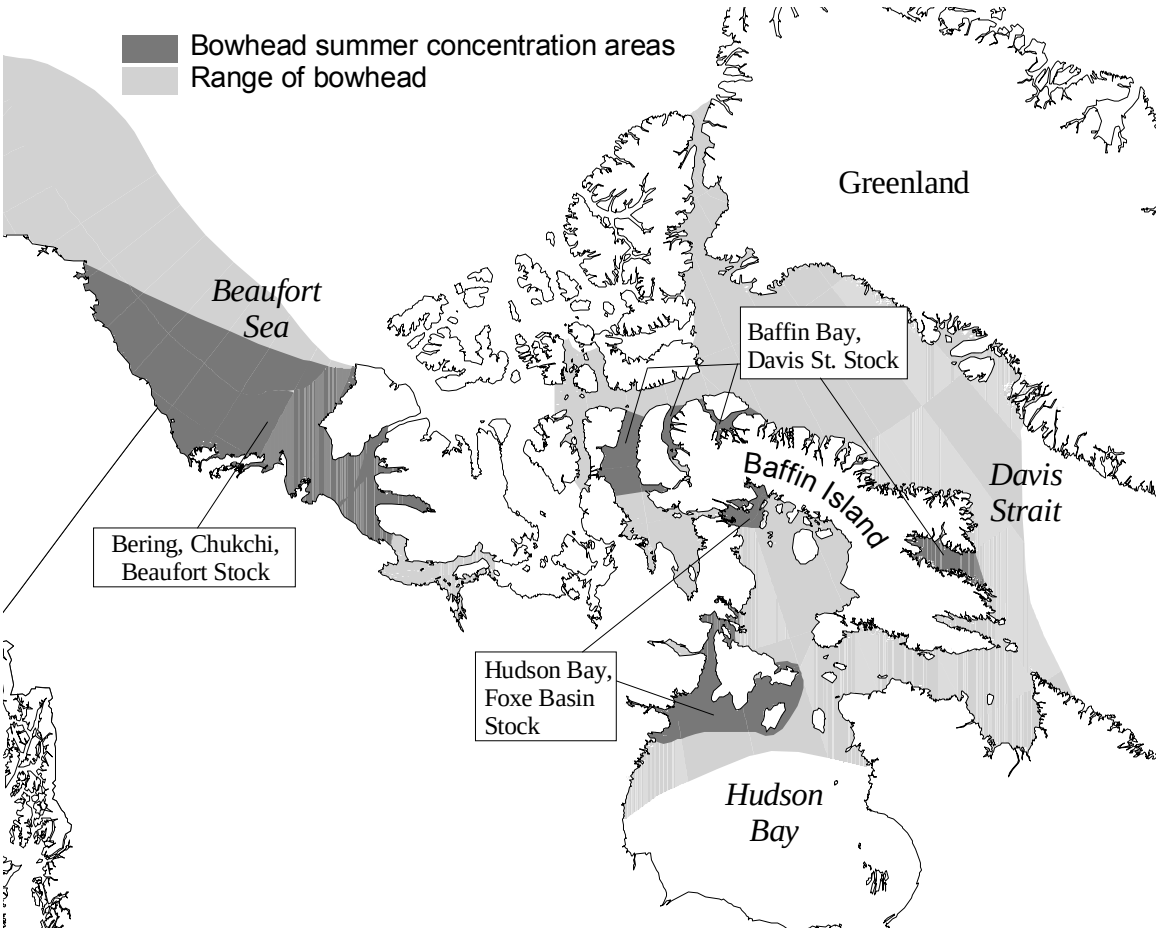


Figure 1. Distribution and summer concentrations of bowhead whales in Canadian and western Greenland waters.

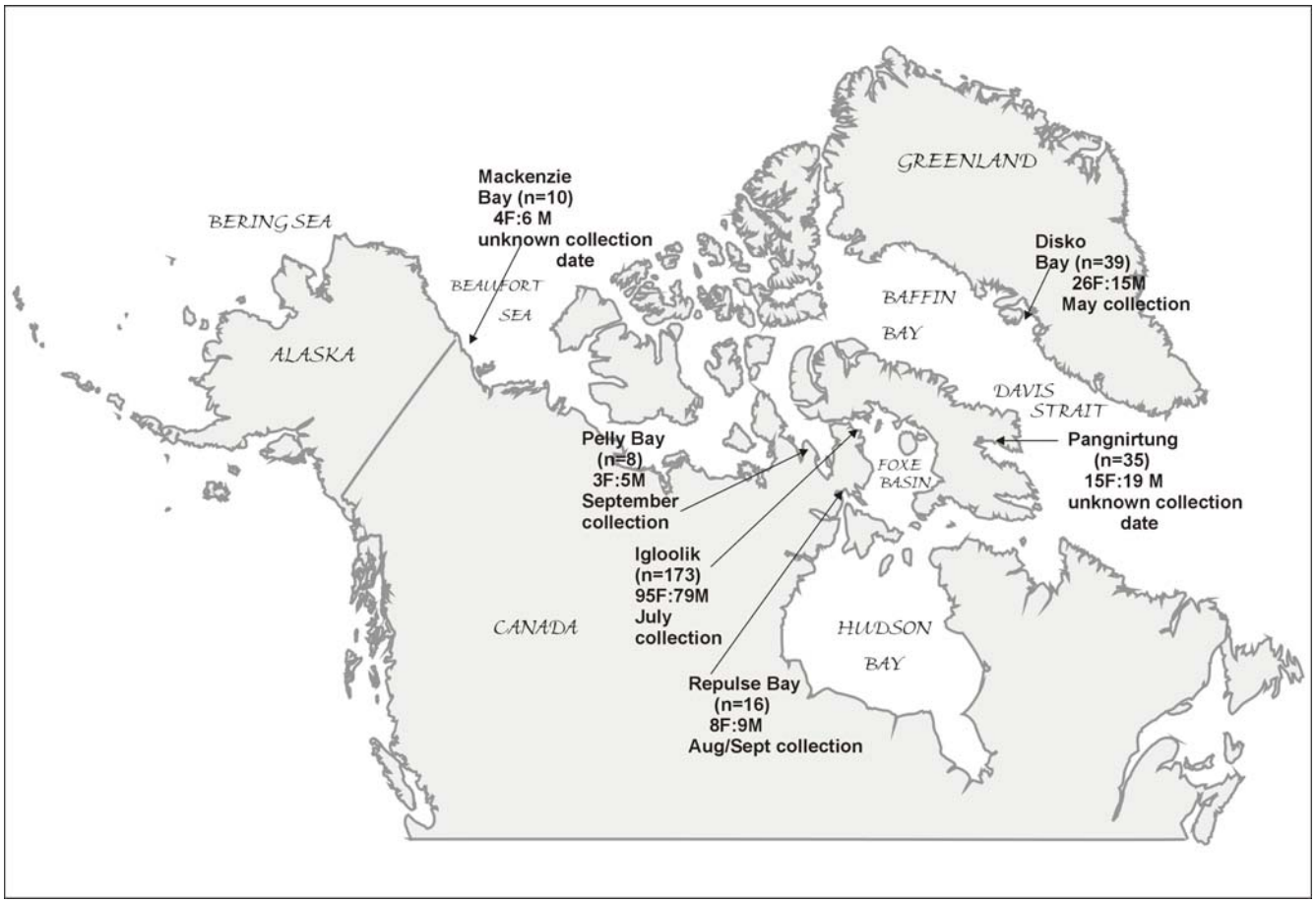
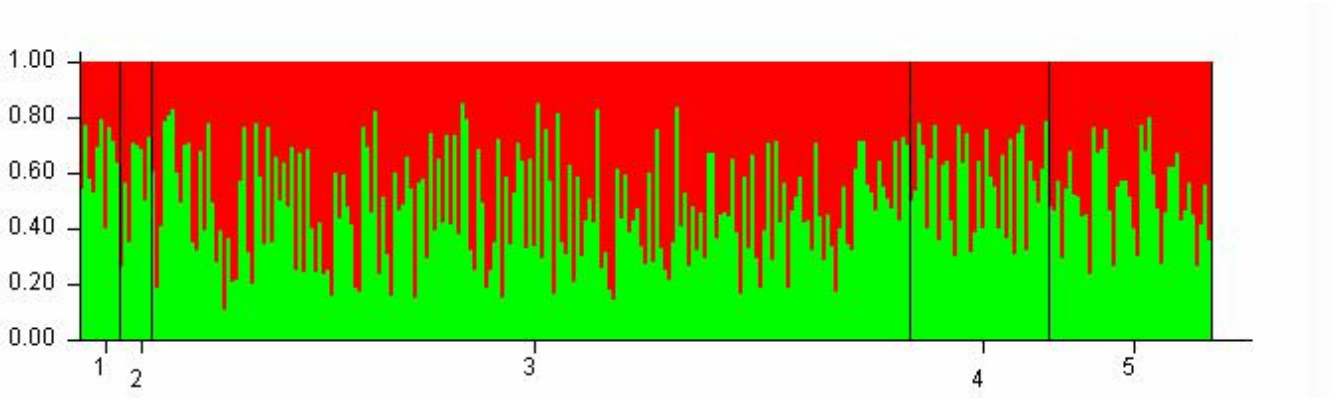


Figure 2. Sampling locations and sample summary for bowheads (n=281) analyzed for molecular genetic markers

Total individuals n = 285
N loci = 12
Burn-in period = 1000000
MCMC reps = 500000

1 = Mackenzie Delta n=9
2 = Pelly Bay n=8
3 = Hudson Bay n=186
4 = Pangnirtung n=35
5 = W. Greenland n=41

K=2



K=3

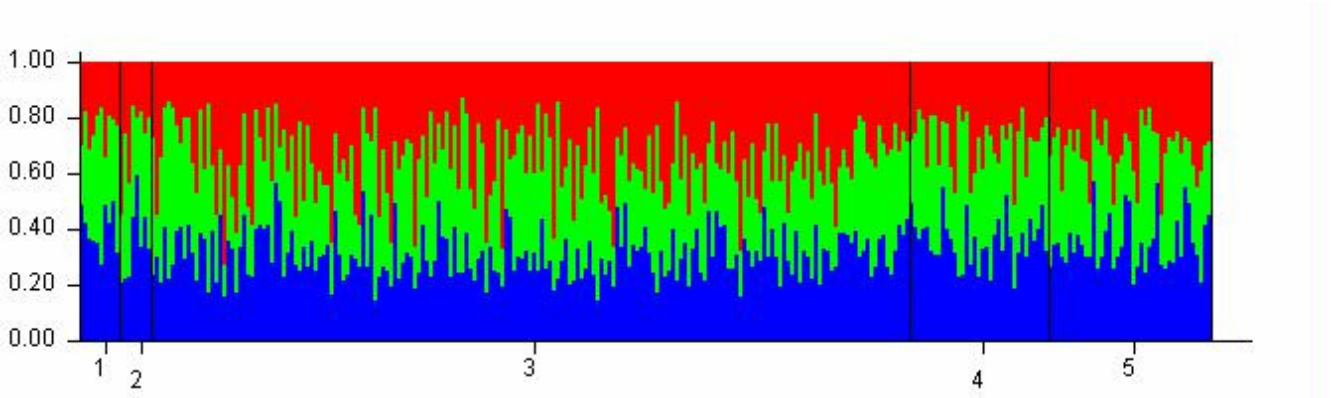
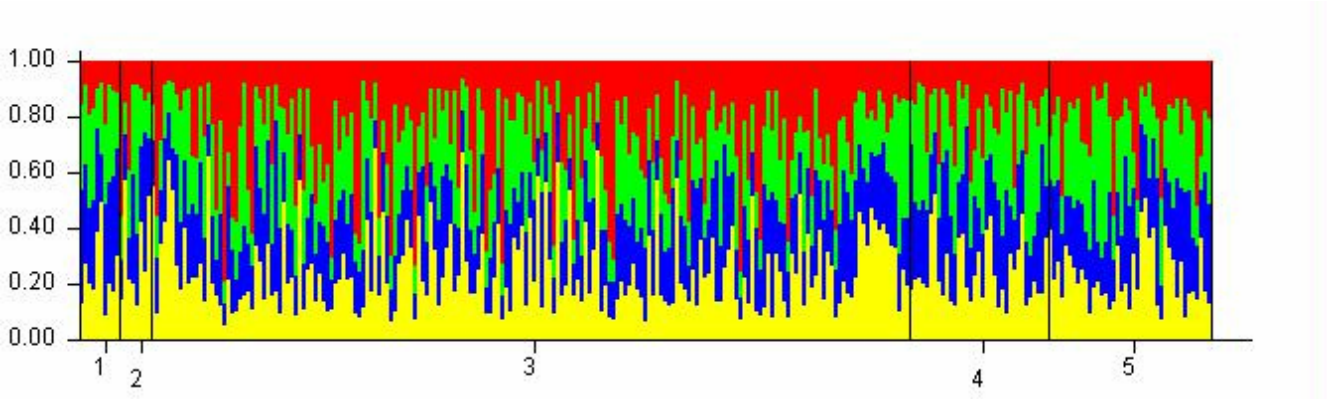


Figure 3. Individual bowhead sample assignments to populations K = 1 to 5 using Structure ver. 2.1.

Total individuals n = 285
N loci = 12
Burn-in period = 1000000
MCMC reps = 500000

1 = Mackenzie Delta n=9
2 = Pelly Bay n=8
3 = Hudson Bay n=186
4 = Pangnirtung n=35
5 = W. Greenland n=41

K=4



K=5

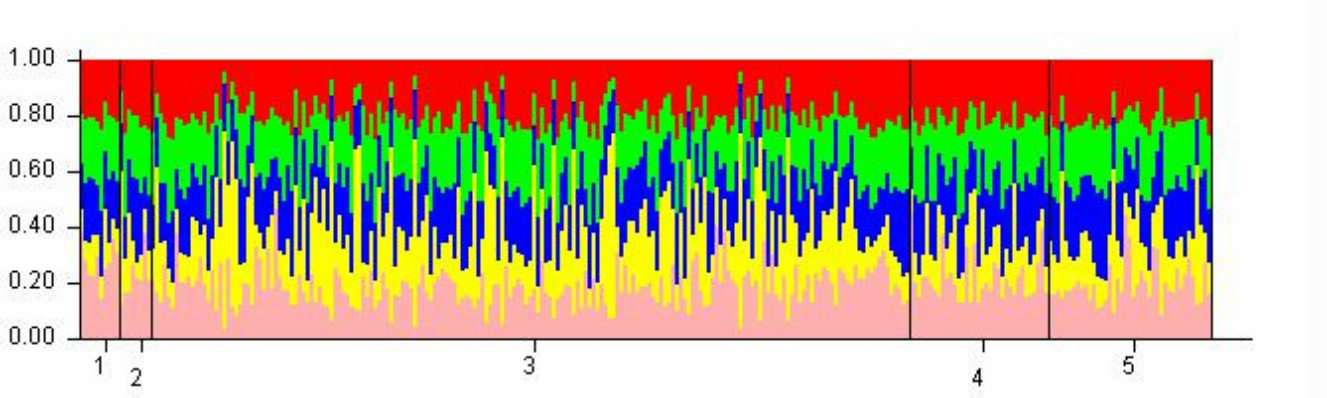


Figure 3 (con't). Individual bowhead sample assignments to populations K = 1 to 5 using Structure ver. 2.1.