Fisheries and Oceans Canada | Pêches et Océans Canada
Ecosystems and Oceans Science | Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Research Document 2023/069

Ontario and Prairies Region and Arctic Region

Stock definition and genetic composition of Cumberland Sound beluga whales (*Delphinapterus leucas*)

Cortney A. Watt[1*], Luca Montana[2*], Justine Hudson[1], and Geneviève J. Parent[2]

[1]Freshwater Institute
Fisheries and Oceans Canada
501 University Crescent
Winnipeg, Manitoba, R3T 2N6

[2]Maurice Lamontagne Institute,
850 route de la Mer
Mont-Joli, QC, G5H 3Z4

September 2023

Canada

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca

ISSN 1919-5044
ISBN 978-0-660-49974-1 Cat. No. Fs70-5/2023-069E-PDF

Correct citation for this publication:

Watt, C.A., Montana, L., Hudson, J., and Parent, G.J. 2023. Stock definition and genetic composition of Cumberland Sound beluga whales (*Delphinapterus leucas*). DFO Can. Sci. Advis. Sec. Res. Doc. 2023/069. iv + 38p.

Aussi disponible en français :

*Watt, C.A., Montana, L., Hudson, J., et Parent, G.J. 2023. Définition du stock et de la composition génétique des bélugas de la baie Cumberland (*Delphinapterus leucas*). Secr. can. des avis sci. du MPO. Doc. de rech. 2023/069. iv + 40 p.*

TABLE OF CONTENTS

ABSTRACT

Less than 1,500 beluga whales (*Delphinapterus leucas*) are estimated to be in Cumberland Sound. These whales, considered as a single population of beluga, have been listed as Threatened under the *Species at Risk Act* and recently assessed as Endangered by the Committee on the Status of Endangered Wildlife in Canada. This beluga whale group has been defined based on satellite tracking data, contaminants, morphometrics, and previous genetic analyses. However, questions regarding whether multiple populations of beluga whales visit Cumberland Sound, their genetic stock discrimination, and the timing of migration and distribution of whales from different groups within Cumberland Sound remain. A recent reexamination of a long haplotype of the mitochondrial DNA (mtDNA) control region showed greater discrimination among beluga whale populations in Eastern Canada, including a small subsample of whales harvested from Cumberland Sound. In this study, we reexamined the genetic distinctiveness of beluga whales hunted in Cumberland Sound (N = 208) compared with other Eastern Arctic whales (N = 657), analyzing all samples collected from this area with the long haplotype of mtDNA. We also genotyped a subsample of whales harvested in Cumberland Sound (N = 27) and Western Hudson Bay (N = 121) using 12,370 nuclear DNA (nDNA) single nucleotide polymorphisms (SNPs) to investigate their distinctiveness. Our results using mtDNA confirmed that approximately 35% of beluga whales harvested in July and August from Cumberland Sound had haplotypes private to this region. The rest of the harvested whales had haplotypes shared with other populations from the Hudson Bay-Strait Complex. Nuclear DNA results also suggested the presence of two populations in Cumberland Sound during summer with approximately 74% of the whales belonging to the CSB population. The degree of differentiation between the CSB and WHB populations was low (F_{ST}=0.014), but this is expected given the recent colonization of the Hudson Bay-Strait Complex. Our results support that there are two populations of beluga whales that summer in Cumberland Sound. Based on all current information, managing beluga whales inhabiting Cumberland Sound in the summer as a single stock comprised of two genetic populations is the most precautionary approach.

INTRODUCTION

BACKGROUND AND OBJECTIVES

The definition of ‘stock’ in regard to marine mammals has been discussed and debated extensively over the last 30 years (Stewart 2008). The International Whaling Commission (IWC) has defined a management stock as a human construct defined in the context of management. It refers to “animals that happen to be present in a defined region and defined season where management is taking place or is contemplated” (IWC 2002). In Canada, the term ‘stock’ is generally used in a management framework to refer to a group of animals that are subject to harvest (Richard 2010).

There are many definitions for “population” that vary widely across the biological sciences and the conservation and management disciplines. In an evolutionary context, one of the suggested definitions for population refers to a group of interbreeding individuals that exist together in time and space (for a review see Waples and Gaggiotti 2006). Genetic and genomic data are often used to describe populations in an evolutionary context and these results are widely used by conservation and management end-users. Populations are also seen as relevant units to maximise conservation of genetic diversity (Funk et al. 2012).

In this assessment, a stock refers to a management unit defined geographically and temporally, as suggested by the IWC (2002). A stock may include more than one population if they overlap spatially and temporally during the management or harvest season. Conversely, a population will be defined as a group of interbreeding individuals that exist together during summer. Our definition of population will be informed using two sources of single nucleotide polymorphisms (SNPs), namely mitochondrial DNA (mtDNA) and nuclear DNA (nDNA). Sequences from mtDNA, referred to as haplotypes, provide information about maternal relatives since this genome is inherited exclusively from the mother in most species, including beluga whales. The geographic distribution of maternal lineages in summering areas was, in the past, the only information used to define beluga whale populations in the Hudson Bay-Strait Complex (e.g., Turgeon et al. 2012, Parent et al. 2023). In this assessment, we also present results from nDNA which provide information about the genomic composition inherited from both parents. Consequently, it may be possible to infer if an individual is the offspring of beluga whales from two different populations.

The target stock or population of this study are Cumberland Sound beluga whales (*Delphinapterus leucas*). These beluga whales have been hunted commercially from the late 1800s until the mid-1900s (Mitchell and Reeves 1981). It is estimated that there were more than 5,000 whales prior to 1923; however, due to commercial whaling, less than 1,000 whales were left by the 1980s (Mitchell and Reeves 1981). In 1990, the Southeast Baffin Island-Cumberland Sound beluga whales were first designated as Endangered by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC; COSEWIC 2020) and in 1991 a quota system was put in place which limited the number of harvested beluga whales in Cumberland Sound to 35 individuals per year (DFO 2002). In 2002, the quota was increased to 41 beluga whales on the condition that females with calves and calves were not hunted, and information on struck and loss rates and number of animals harvested were collected by the community of Pangnirtung (DFO 2002). The most recent aerial survey in 2017 estimated an abundance of 1,381 (95% confidence interval [CI] = 1,270-1,502) beluga whales in Cumberland Sound (Watt et al. 2021). Based on the survey and consecutive abundance modelling, the probability of decline in beluga whales in Cumberland Sound over the next 10 years with the current harvest rate of 41 whales annually is approximately 96% (DFO 2019). At a harvest rate of 20 whales annually, the

probability of decline is 50% (DFO 2019). A harvest reduction to reduce this risk has not yet been implemented, in part because Inuit knowledge states there are two groups of beluga whales that visit Cumberland Sound, which is not reflected in the risk assessment.

This study aims to 1) review current evidence for delineation of Cumberland Sound beluga whales, and 2) provide new results from genetic and genomic analyses contrasting Cumberland Sound beluga whales with other groups of beluga whales from the Hudson Bay-Strait Complex. A full review of Inuit knowledge on Cumberland Sound beluga whales is not presented herein and is outside the scope of this assessment. Instead, we aim to provide an overview of the scientific information used to define Cumberland Sound beluga whales historically, and a new scientific assessment of genetic information.

How were Cumberland Sound beluga whales identified historically?

Initially, Cumberland Sound beluga whales were included as Southeast Baffin Island-Cumberland Sound beluga whales, which also included whales harvested by hunters from Kimmirut and Iqaluit (Richard and Orr 1986) (Figure 1). In 1990 these whales were designated as Endangered by COSEWIC (COSEWIC 2004). Subsequently, satellite tracking data, contaminants, morphometrics, and genetic analyses suggested that beluga whales hunted in Cumberland Sound were distinct from those hunted in Kimmirut and Iqaluit (Sergeant and Brodie 1969, de March et al. 2002, Innes et al. 2002, Richard and Stewart 2008, Turgeon et al. 2012, Watt et al. 2016). Therefore, in 2004, the Southeast Baffin Island-Cumberland Sound beluga whales were re-delineated and the Southeast Baffin Island animals, which included beluga whales from Kimmirut and Iqaluit, were considered part of Western Hudson Bay (WHB). Cumberland Sound beluga whales were listed as “Threatened” under the *Species at Risk Act* in 2017 and assessed as Endangered by COSEWIC in 2020 (COSEWIC 2020).

Satellite tracking data

Between 1998 and 2008, 19 beluga whales from Cumberland Sound were equipped with satellite tags and tracked over a period of 2 to 270 days (Richard and Stewart 2008, Watt et al. 2016) (Table 1). Tracking data showed that beluga whales ranged from Nettilling Fiord to Clearwater Fiord in July and August (Richard and Stewart 2008) (Figure 1). In September, most tagged whales were found in Clearwater Fiord and around the entrance of Kangilo Fiord; however, beluga whales were also located on the southwestern side of Cumberland Sound (Richard and Stewart 2008, Watt et al. 2016). In late fall and early winter, whales were located in the southeastern side of Cumberland Sound, along the Cumberland Peninsula, where a persistent polynya has been observed, which may allow beluga whales to overwinter in the Sound and avoid entrapments during the winter (Richard and Stewart 2008, Watt et al. 2016). Beluga whales likely mate in March-June on the wintering grounds (Heide-Jørgensen and Teilmann 1994). Although the number of tagged whales is small, telemetry data suggests that Cumberland Sound beluga whales spend the entire year in Cumberland Sound and are resident to the area (Richard and Stewart 2008).

Contaminants

Organochlorines (OCs), including polychlorinated biphenyls and persistent contaminants, including polybrominated diphenyl ether (PBDE) flame retardants can accumulate in high-trophic level marine mammals, such as beluga whales (Innes et al. 2002, McKinney et al. 2006, Smythe et al. 2018). Since variations in the type of OCs and OC concentrations vary due to prey composition in different foraging locations, OCs can be used to differentiate whale groups (Aguilar 1987). For example, blubber contaminants from whales harvested in Cumberland Sound differed from whales harvested in Kimmirut (Innes et al. 2002). Beluga whale SPBDE concentrations were also lower in males harvested in Cumberland Sound than males harvested

in Western Hudson Bay (McKinney et al. 2006). Additionally, decreasing trends in perfluorocarboxylic acid (S_5PFCA) concentrations were observed in belugas harvested from the Eastern Beaufort Sea, Eastern Hudson Bay, and Western Hudson Bay regions, while an increase in concentrations was observed in Cumberland Sound beluga whales, potentially indicating a higher source of exposure unique to Cumberland Sound (Smythe et al. 2018). Differences in OC concentrations and ratios were also observed between whales harvested in Cumberland Sound (by the community of Pangnirtung), Kimmirut, and Iqaluit, supporting the hypothesis that Cumberland Sound beluga whales are distinct from Kimmirut and Iqaluit whales (de March et al. 2004). Differences in OCs among the three southeast Baffin Island beluga whale groups may be due to variation in foraging strategies, with beluga whales in Cumberland Sound potentially feeding at lower trophic levels, on benthic fish or fish that eat benthic organisms, and have relatively lower OC loads (de March et al. 2004).

Morphometrics, hormones, and diet

Beluga whales harvested in Cumberland Sound are significantly larger than those harvested in Western Hudson Bay with an asymptotic body length of 370.9 cm compared to 354.4 cm, respectively (Ferguson et al. 2020). Cumberland Sound beluga whales also have significantly higher cortisol (a stress-related hormone) levels than the Eastern High-Arctic Baffin Bay or Western Hudson Bay beluga whales, and a significantly different $\delta^{13}C$ signature (indicative of where animals forage; Newsome et al. 2010) compared to Western Hudson Bay beluga whales (Kucheravy et al. 2022). In addition, Cumberland Sound beluga whales have a different temporal $\delta^{13}C$ trend than whales harvested in Eastern High-Arctic Baffin Bay and Western Hudson Bay (Matthews and Ferguson 2014). Although a similar $\delta^{15}N$ signature between Cumberland Sound and Western Hudson Bay beluga whales suggest they forage at a similar trophic level, Cumberland Sound beluga whales have a unique fatty acid signature (Kucheravy et al. 2022).

Genetics

Previous genetic studies using mtDNA indicate that Cumberland Sound beluga whales are genetically distinct from other beluga whales (Brown Gladden et al. 1997, de March et al. 2002, Turgeon et al. 2012). Brown Gladden et al. (1997) first determined that beluga whales harvested in Cumberland Sound differed from those harvested in Kimmirut and Iqaluit. The authors found a temporal distinction within beluga whales harvested from Cumberland Sound, with beluga whales hunted in the 1980s differing from those hunted in the 1990s, which was thought to be caused by a ban on hunting in Clearwater Fiord, an important calving area for this population (Brown Gladden et al. 1997). de March et al. (2002) also found that beluga whales harvested in Cumberland Sound were distinct from whales harvested from Kimmirut, while beluga whales harvested in Iqaluit had similar genetic characteristics to beluga whales from both Cumberland Sound and Kimmirut. In concordance, Turgeon et al. (2012) found that beluga whales from Hudson Bay and nearby areas formed three genetically distinct summering populations, including Eastern Hudson Bay, Western Hudson Bay, and Cumberland Sound. Cumberland Sound beluga whales were identified as being distinct from beluga whales harvested in Kimmirut and Iqaluit (Turgeon et al. 2012). Together these three studies suggest that Cumberland Sound beluga whales are a genetically distinct population, and since mtDNA is maternally inherited, they suggest that genetic differences between populations are likely sustained by maternal site fidelity (Turgeon et al. 2012).

A new genetics and genomics assessment for Cumberland Sound beluga

A new study confirmed the mixed genetic composition of beluga whales in Cumberland Sound using mtDNA (Parent et al. 2023), supporting Inuit knowledge from Pangnirtung that two groups

of beluga whales inhabit the area. This genetic assessment included 2,861 samples from the Hudson Bay-Strait Complex and was done using a longer mtDNA control region sequence (615 nucleotides) than previous studies (Brown Gladden et al. 1997, de March et al. 2002, Turgeon et al. 2012). With this larger sample size and longer haplotype sequence, the study had greater power to differentiate genetic populations in the Hudson Bay-Strait Complex. In Cumberland Sound, beluga whales had either private haplotypes or haplotypes common to other populations. Moreover, self-assignment rates for beluga harvested during summer in Cumberland Sound were the lowest among the four reference groups described by Parent et al. (2023), either indicating a marginal differentiation with belugas from Western Hudson Bay, or that belugas from other populations inhabit the Sound as well, thus, prompting further analyses.

Parent et al. (2023) identified five genetic beluga whale populations, including Eastern Hudson Bay (EHB), James Bay (JAM), WHB, Cumberland Sound (CSB), and a newly described population in the Belcher Islands (BEL). To conserve genetic diversity, the authors suggested these five populations should be considered as independent evolutionary units. In Cumberland Sound, 34.5% of whales harvested had private haplotypes (specific to CSB) whereas the remaining beluga whales had haplotypes in common with WHB and BEL beluga whales (Parent et al. 2023). The small sample size for whales harvested in Cumberland Sound used in Parent et al. (2023) prompted a more thorough investigation into genetic discrimination of whales harvested within Cumberland Sound, which is presented here.

METHODS

SAMPLES

Cumberland Sound beluga whale skin samples have been collected by hunters as part of the Marine Mammal Sampling Program in Pangnirtung since the 1980s. The exact location where each animal was harvested is sometimes reported, and in some cases the location name is cited. In these cases, location names were converted to approximate latitude and longitude locations. In many cases the hunt location was not provided or indicated as ‘Pangnirtung’, in which case no latitude or longitude were assigned. Occasionally tissues were preserved in a saturated salt dimethyl sulphoxide (DMSO) solution, while many were frozen initially upon collection and later preserved in a DMSO solution once shipped to the Freshwater Institute in Winnipeg, Manitoba.

We also used summer samples from the other three reference groups (BEL-EHB, JAM, WHB) which were identified in Parent et al. (2023) using mtDNA. Samples were obtained from harvested, biopsied, or satellite tagged beluga whales between 1982 and 2021. Again, the exact location where each animal was harvested was usually unknown and was attributed to the area in which the harvest event occurred (see Parent et al. 2023 for more information). For these samples, most tissues were preserved in a saturated salt solution containing 20% DMSO and 0.5 mol/L ethylene diamine tetraacetic acid (Seutin et al. 1991). Some samples were only frozen while others were frozen first and preserved later using the DMSO solution.

DNA extraction and characterization

DNeasy Blood and Tissue kit (QIAGEN, Valencia, USA) was used to extract DNA from skin samples. We extracted DNA from 846 samples for mtDNA haplotyping and from 182 samples for nDNA genotyping (N_{CS} = 36, N_{WHB} = 146) (Table 2). For nDNA, DNA extracts were visually inspected on agarose gel to ensure high quality samples and their concentration was estimated on a Synergy LX (BioTek, Santa Clara, USA) fluorescence plate reader using picogreen as the fluorescent marker.

Mitochondrial DNA

Haplotyping

Partial sequences (615 bp) of the mtDNA control region were amplified and sequenced for the 846 samples according to the protocol detailed in Parent et al. (2023) using PCR conditions from Postma et al. (2012), with the Multiplex PCR Kit (QIAGEN, Valencia, USA). Sequencing was performed on an ABI 3130 sequence (Applied Biosystems Inc., Foster City, USA) at Maurice Lamontagne Institute, following the procedure in Postma et al. (2012).

For each specimen, consensus sequences using the forward and the reverse sequencing outputs were produced and manually edited using Geneious Prime 2020.1 (Biommatters, Ltd, Auckland, New Zealand) and Sequencher® 5.4.6 (Gene Codes Corporation, Ann Arbor, MI USA). Consensus sequences were then aligned in R (R Core team 2022) using the *muscle* algorithm available with the package Biostrings 2.64.1 (Pagès et al. 2023, penalties for gap opening: 10,000, extension: 400). The starting nucleotide of the 615-nt sequence corresponds to position 38 of the complete mtDNA control region (Lillie et al. 1996). Indels were not detected.

Haplotypes were defined following the procedure described in Bonnet et al. (in prep.[1]). Briefly, single nucleotide polymorphisms (SNPs) were identified using adegenet 2.1.7 (Jombart 2008, Jombart and Ahmed 2011). Control region mtDNA haplotypes based on a minimal sequence of 570 nucleotides are listed in a haplotype library. This library was used to assign a haplotype to each beluga whale.

Lineage analyses

Reference samples from beluga whales harvested in July and August (N = 846), while they reside in their summer grounds, were used to ascertain if mtDNA control region haplotypes can be used to distinguish genetic units within Cumberland Sound summering beluga whales.

A statistical parsimony network of haplotypes was generated using sequences in PopART (Leigh and Bryant 2015). We performed a principal component analysis (PCA) using haplotype frequencies with the function *dudi.pca* of the ade4 1.7-19 R package (Dray and Dufour 2007). Finally, summary statistics on private haplotypes (i.e., those haplotypes specific to a population) were also estimated in R (R Core Team 2022).

Nuclear DNA

Library synthesis and sequencing

DNA (20 ng) from each of 182 beluga whale samples from Cumberland Sound (N = 36; Table 2) and Western Hudson Bay (N = 146; Table 2) were sent to the Plateforme d'analyse génomique (IBIS, Université Laval) for the preparation of ddRADseq libraries using *PstI* and *MspI* restriction enzymes. Libraries were sequenced on Illumina NovaSeq 6000 S4 PE 150 at Génome Québec (Montréal, Canada) with 10% PhiX.

Quality control steps

Overall quality of the reads and presence of adapters were assessed using FastQC 0.11.9 (Andrews 2010) and multiQC 1.10 (Ewels et al. 2016). Illumine adapters as well as three base pairs (bp) from read 2 (R2) were removed from raw sequence files using Trimmomatic 0.39

[1] Bonnet, C., Montana, L., St-Pierre, A.P., Sauvé C., Hammill, M.O., and Parent, G. In prep. Genetic monitoring program for beluga (*Delphinapterus leucas*) harvested by the Nunavik regions. DFO Can. Sci. Advis. Sec. Tech. Report. In preparation.

(Bolger et al. 2014). The three bp removed correspond to the *MspI* restriction site, where we detected a decrease in sequence quality. Reads were then visualized again with FastQC and multiQC to ensure removal of Illumina adaptors and examine read quality anew. Reads were then processed with the Stacks 2.55 pipeline (Catchen et al. 2013, Rochette et al. 2019), and the module *process_radtags* was used for demultiplexing and quality filtering: reads were truncated to 135 bp, and *PstI* restriction site quality at read 1 (R1) was assessed. Demultiplexed reads were aligned to a new beluga genome assembly (project accession number: PRJNA925093; Bringloe et al. in prep.[2]) with the Maximal Exact Match (MEM) algorithm in BWA-MEM (Li and Durbin 2010, Li 2013) using default parameters. To call SNPs with reference alignments, aligned reads were then sorted using SAMtools 1.12 (Danecek et al. 2021). A sample with an alignment rate below 96% was discarded from further analysis (Table A1). Aligned pair-end reads were finally assembled with *gstacks* modules. Samples with a mean coverage below 5X were discarded from further analyses, resulting in a sample size of 158 beluga whales (N_{CSB} = 29; N_{WHB} = 129; Table A1).

Filtration of single nucleotide polymorphisms

The *population* module of Stacks was used to perform the first SNPs filtration step on the 158 samples: SNPs were removed if they were not found in at least 75% of individuals, or with a minor allele frequency (MAF) of ≤ 1% (Table A1). The resulting SNP panel was exported in VCF format. The second filtration step aimed to discard those loci with low read depth (or coverage). If read depth of a locus is insufficient, alleles may not be detected and thus result in false homozygotes (O'Leary et al. 2018, Rochette et al. 2019). We inspected if some loci were more homozygous than expected by comparing observed vs expected heterozygosity (Figure A1A). Using VCFtools 0.1.16 (Danecek et al. 2011), we removed loci with median read depth ≤ 12X because most loci with a higher than expected homozygosity had median read depth between 5X and 12X (Figure A1B), as well as those with median depth ≥ 28X, which represents the 99th percentile of the median read depth distribution (Table A1). The third filtration step aimed to discard individuals with more than 30% missing loci and loci with more than 10% missing data using VCFtools 0.1.16 (Table A1; Danecek et al. 2011). Next, loci with an observed heterozygosity ≥ 60% were identified after importing VCF files in R (R Core Team, 2022) with vcfR 1.13.0 (Knaus and Grunwald 2017), and discarded using VCFtools (Table A1; Danecek et al. 2011). Loci were screened to identify potential sequencing plate effects or sex-linked loci. One locus with clear sequencing plate effect was removed from further analysis (Table A1). For most samples (168 out of 182 original samples) sex was identified with a qPCR-based method (Bonnet et al. in prep.[1]). Sex-linked SNPs were identified through a redundancy analysis using the function *rda* of the R package *vegan* 2.6-2 (Oksanen et al. 2022) and discarded (Table A1). Relatedness between samples was estimated using the method of Manichaikul et al. (2010) available in VCFtools. High relatedness (Φ > 0.25) can be caused either because two samples belong to the same individual (Manichaikul et al. 2010), or because contamination between samples occurred. For samples with metadata indicating they belonged to the same specimen, one of the duplicates was discarded (N = 4). If contamination occurred, both samples were discarded (N = 5). This filtering step left 148 specimens for analyses (Table 2; Table A1). We then kept one SNP per locus by selecting that with the highest MAF (Table A1). Finally, we re-estimated the MAF and the number of missing loci within the finalized reduced dataset. We eliminated loci with more than 5% missing data and with MAF ≤ 5% or ≤10%. A MAF ≤ 5% is a common threshold in marine mammals studies whereas a MAF ≤ 10% is a conservative

[2] Bringloe, T., and Parent, G. In prep. Contrasting new and available reference genomes to highlight uncertainties in assemblies and areas for future improvement: an example with monodontid species. In preparation.

threshold that should avoid the inclusion of alleles associated to sequencing errors in the analyses, a practice discussed in Díaz-Arce and Rodríguez-Ezpeleta (2019).

Remaining SNPs were screened for outlier loci (possibly under selection) using the PCA-based methods implemented in pcadapt 4.3.3 using Mahalanobis distance (Luu et al. 2017, Privé et al. 2020). The number of PCs used to screen for outlier loci was chosen from a visual observation of the scree plots using Cattell's rule (K = 4; Cattell 1966) as suggested by pcadapt authors (bcm-uga.github.io/pcadapt/articles/pcadapt.html). SNPs with a q-value < 0.05 were identified as outliers. Results for the dataset without outlier loci and MAF > 10% are presented from this point onward ($N_{loci\ MAF\ 0.1}$ = 12,370), except indicated otherwise. Note that all analyses were run for datasets with MAF > 5% and > 10%, with and without outliers, and we observed concordant results regardless of datasets used.

Population structure analyses

Exploratory PCAs were performed at the individual level over all datasets for the first two PCs with the *glPca* function of the package *adegenet* 2.1.7 in R (Jombart 2008, Jombart and Ahmed 2011). We then inferred the number of genetic groups present in the dataset using the maximum-likelihood approach in ADMIXTURE 1.3.0 (Alexander et al. 2009; default parameters) and estimated the membership-probability to each of these groups. Standard errors (SEs) for membership-probability point estimates were estimated using the moving block bootstrap procedure (1000 bootstrap replicates; Alexander et al. 2009). We used all samples from Cumberland Sound (N = 27; Table 3) and randomly selected 35 samples from Western Hudson Bay to ensure similar sample size from both putative populations, and thus maximize the likelihood the ADMIXTURE analysis identifies populations ($N_{ADMIXTURE}$ = 62). We implemented the cross-validation approach, testing one to four populations (K), to determine the number of populations estimated using nDNA SNPs (Alexander et al. 2009). Finally, we estimated pairwise F_{ST} (the fixation index, which measures allele frequency divergence among populations; Holsinger and Weir 2009) between Cumberland Sound and Western Hudson Bay beluga whales with the function *gl.fst.pop* in dartR 2.0.4 (Gruber et al. 2018) following the equation of Weir and Cockerham (1984). Confidence intervals (95%) were estimated by running 999 bootstraps.

RESULTS

MITOCHONDRIAL DNA

This study provides the haplotypes of 207 beluga whales harvested in Cumberland Sound between 1982 and 2021 (Figure 1). From these samples, 189 were harvested in July and August (Table 2) and the other 18 samples were collected either in spring or fall. We also included 657 beluga whales from BEL-EHB (N = 261), JAM (N = 78), WHB (N = 318) and haplotyped all samples at the 38 SNPs within the mtDNA control region.

Fifteen haplotypes (21, 23, 24, 28, 29, 40, 57, 73, 82, 128, 139, 140, 141, 142, and 143) private to CSB were again observed in the haplotype network (Figure 2). Private haplotypes were found in 35% of samples harvested in Cumberland Sound.

The PCA showed that the haplotype composition of CSB was very different from BEL-EHB and JAM, and differed slightly from WHB (Figure 3; Table A2). JAM separated from WHB and CSB along the first axis, whereas BEL-EHB separated from JAM, WHB, and CSB along the second axis of the PCA. The difference between WHB and CSB was greater between WHB July samples and CSB August samples. The difference in proportions of haplotype 24 (HL024) contributed most to the difference between WHB and CSB (Figure 3).

The sample size in Clearwater Fiord and Kangilo was insufficient to determine if the proportion of private haplotypes differs among surveyed regions within Cumberland Sound ($N_{Clearwater\ Fiord}$ = 12; $N_{Kangilo}$ = 4; $N_{North\ Stratum}$ = 21; $N_{West\ Stratum}$ = 44). In the West Stratum 18 beluga whales out of 44 had private haplotypes, while in the North Stratum 10 beluga whales out of 21 had private haplotypes (Figure 4).

Fourteen beluga whales satellite tagged in Cumberland Sound at the mouth of Clearwater Fiord had samples available for mtDNA analysis (Table 1). Of the seven whales tagged in 2006-2008, five beluga whales had private haplotypes (2x HL028, 2x HL040, 1xHL057). Out of the seven beluga whales tagged in 1998-1999, one had CSB private haplotypes (HL024).

NUCLEAR DNA

A mean of 5,619,009 reads per individual were obtained for the 148 Cumberland Sound and Western Hudson Bay beluga whales (min = 1,459,944; max = 16,633,981; Table 2). Loci were thoroughly filtered to avoid potential bias in interpretation of neutral population structure. Each beluga was genotyped for 12,370 loci (Table A1), excluding outlier loci (min = 9,252; max = 12,365). The mean sequencing depth for these loci from the 148 individuals was 17.67X (min = 11.41X; max = 35X).

The exploratory PCA showed two genetic clusters of beluga whales present in Cumberland Sound (Figure 5). Most belugas harvested in Cumberland Sound (N = 20, CSB cluster) had PC1 values larger than 3.2 while all but one individual harvested in Western Hudson Bay had PC1 values lower than 1.2 (WHB cluster). Within the CSB cluster, four samples separated from the other 16 samples along the PC1 axis (the two circles and two triangles on PC1 > 13). These four samples did not have higher proportions of missing loci (Figure A2), or were not more homozygous (Figure A3) than samples found in the main CSB cluster. Using all specimens available per harvest location, we estimated the F_{ST} between beluga whales harvested in Cumberland Sound (Figure 5; N = 27) and Western Hudson Bay (Figure 5; N = 121), and obtained F_{ST} = 0.0083 (0.0078-0.0088, P = 0.00). Seven beluga whales harvested in Cumberland Sound grouped with the WHB cluster. Of these seven samples, the one at the edge of the WHB cluster had a greater proportion of missing data or lower observed heterozygosity compared to the other six samples (Figure A2; Figure A3). One beluga whale harvested in Western Hudson Bay grouped with the CSB cluster.

The ADMIXTURE analysis and cross validation results identified one or two genetic groups in the subset dataset of 62 beluga harvested either in Cumberland Sound (N = 27) or Western Hudson Bay (N = 35; Figure A4; Figure 6). The best model for estimating the number of genetic groups (K) was of one genetic group using the cross validation results (Figure A4). However, low genetic differentiation may impede the detection of the “real” K (Cullingham et al. 2020). Thus, we present the model results with K = 2 due to the PCA analysis results and biologically relevant hypotheses based on previous genetic analyses (Turgeon et al. 2012) and traditional knowledge (Kilabuk 1998). For K = 2, one genetic group was more abundant in whale samples harvested in Cumberland Sound, while the other genetic group was more abundant in whale samples harvested in Western Hudson Bay (Figure 6; Table A3; see also Figure A5 for ADMIXTURE results for analyses using the MAF > 5% dataset). For beluga whales harvested in Cumberland Sound, six whales had membership probabilities identifying them completely to the CSB genetic group. Four of these beluga whales had a PC1 value greater than 12 in the PCA (Figure 5). In the ADMIXTURE analysis, 20 beluga whales harvested in Cumberland Sound had a membership probability ≥ 50% to the CSB genetic group (Figure 6; Table 3). We used this arbitrary threshold as a conservative way to identify individuals most likely having a higher proportion of their genome associated to CSB. These results are congruent with the PCA

results. For beluga harvested in Western Hudson Bay, all but one beluga whale had membership probability greater than 75% to the WHB genetic group (Table A3). On average, the SE associated to membership-probability point estimates was SE = 0.043 (min = 0.00; max = 0.089 (Table A3). Since PCA and ADMIXTURE results concordantly indicated the presence of potential migrants, we estimated a second F_{ST} using the smaller dataset used for the ADMIXTURE analysis (N = 62). Additionally, putative migrants (i.e. animals harvested in Cumberland Sound or western Hudson Bay with membership probability < 25% to their local population; N = 8) were removed from the dataset, leaving a final sample size of 54 whales (N_{CSB} = 20; N_{WHB} = 34), which generated a F_{ST} = 0.014 (0.013-0.015; P = 0.00). We selected the < 25% threshold since WHB membership probability varied between 0 and 25% for the WHB population (except for the one migrant). Note that the F_{ST} value estimated without migrants was almost twice as large as the F_{ST} estimated with putative migrants. For beluga whales harvested in Cumberland Sound, we had the geographic coordinates for 14 individuals (North Stratum N = 7, West Stratum N = 6, Kangilo N = 1). We estimated that 71% (or 5 out of 7) and 83% (or 5 out of 6) beluga whales harvested in the North and West Strata, respectively, were from the CSB population.

Note that we avoided presenting membership coefficient results in terms of admixing between the two groups. ADMIXTURE bar plots may be overinterpreted (Lawson at al. 2018). Different explanations may cause similar patterns in these bar plots. In this study, the low genetic differentiation detected between the two populations precludes any interpretation of the values in the ancestry coefficient associated to each population. We used a conservative approach and determine that any individuals having a proportion of CSB population ≥ 40% was a CSB individual. We used this arbitrary threshold as a conservative way to identify individuals most likely having a high proportion of their genome associated to CSB. Those classification results were congruent with the PCA results, which provided some confidence in the utilisation of this threshold.

DISCUSSION

Cumberland Sound beluga whale management has been challenging over the last ~30 years, partly because of a disconnection between science and Inuit knowledge about whether more than one group of beluga whales summer in Cumberland Sound (Kilabuk 1998). In our study, the combination of analyses from mtDNA and nDNA support the presence of two genetically distinct populations of beluga whales inhabiting Cumberland Sound during the summer, namely WHB and CSB. Based on mtDNA, a large proportion of genetic diversity is unique to CSB, as described in previous studies (Turgeon et al. 2012, Postma 2017, Parent et al. 2023). Based on nDNA, this study showed that most beluga whales harvested in Cumberland Sound are from the CSB genetic group. We also showed that a smaller proportion of beluga whales harvested in Cumberland Sound (26%, or 7 out of 27) was similar to the WHB cluster or genetic group. Our results are in agreement with Inuit knowledge from the Pangnirtung community which has indicated the presence of two distinct beluga whale groups in Cumberland Sound.

CUMBERLAND SOUND AND WESTERN HUDSON BAY POPULATIONS WERE DISTINCT USING NUCLEAR DNA

A reminder of the use of terminology in this document is of the utmost importance because of the new type of genetic information provided in this study. Previously, the definition for beluga whale populations was based solely on mtDNA, which identified or used the information about maternal lineages specific to summering areas (e.g., Hammill et al. 2021, Parent et al. 2023). In an evolutionary context, however, populations are usually defined as a group of interbreeding individuals that exist together in time and space (Waples and Gaggiotti 2006). This study is the

first to characterize interbreeding groups of belugas using nDNA SNPs. Our results using nDNA SNPs highlight that beluga whales harvested in Cumberland Sound were mostly from the CSB genetic group, which is distinct from the WHB genetic group. For this study and subsequent analyses of this genetic group, the term population will refer to beluga whales from the distinct CSB or WHB genetic groups identified with the PCA or ADMIXTURE analyses.

Analyses of nDNA markers showed that CSB and WHB populations had distinct genetic profiles. A previous study used microsatellite loci from the nuclear genome to infer genetic structure among whales harvested in the Hudson Bay-Strait Complex, but could not replicate the structure identified with mtDNA (Brown Gladden et al. 1997, de March and Postma 2003, Turgeon et al. 2012). The dissimilarity between the results presented here and those originating from microsatellites underlines the greater resolution that nDNA SNPs offer for the study of genetic differences among populations. The estimated genetic differentiation between these populations was low regardless of whether putative migrants were included. Low genetic differentiation is expected given the recent isolation and divergence of these populations, as inferred for beluga whales in Alaska and northwestern Canada (O'Corry-Crowe et al. 1997). It is unknown when this isolation began but ice retreat started less than 10,000 years ago in the Hudson Strait (Dyke 2004). This left little evolutionary time for CSB and WHB populations to isolate and diverge, possibly explaining the low genetic differentiation. Members of the CSB population interbreeding during late winter/early spring only at the entrance of Cumberland Sound may have accelerated the divergence from WHB. Future work using demographic models could help elucidate the timing of divergence and interbreeding since colonization of these areas.

Both the PCA and ADMIXTURE results identified substructure within the CSB population. With the PCA, four individuals from the CSB population stood out as separating further from WHB whales. Those four individuals and two more were identified with 100% membership probabilities to the CSB population in the ADMIXTURE analysis. Uncertainty in assigning some individuals with high confidence to a specific population could be the result of admixing with the WHB population (but see Lawson et al. (2018) warning on over-interpretation of admixture membership-probabilities). However, larger sample sizes and other statistical approaches (e.g., assignment rates with simulated genotypes) would be necessary to provide reliable information about putative interbreeding between both populations (Lawson et al. 2018).

The PCA and ADMIXTURE also highlighted the presence of individuals with genotypes similar to those of WHB beluga whales (26%) among the Cumberland Sound summer harvest, as well as one CSB whale harvested in Western Hudson Bay. The presence of whales with a genetic profile distinct from that of the summering population might have originated by dispersal of young adults. Among mammals, female philopatry and male dispersal has been described as the general pattern (Wolff 1997), and beluga whales are no exception (O'Corry-Crowe et al. 2018). Despite the limited sample size, here we show that among putative migrants, 7 out of 8 beluga whales were males. Yet, to this day, evidence for male biased dispersal to new summering areas has not been presented (de March and Postma 2003, Turgeon et al. 2012, Colbeck et al. 2013), but could have impacts on the genetic connectivity and survival of threatened populations (Lowe and Allendorf 2010). Beluga whales summering in the Hudson Bay-Strait Complex migrate from their summering waters (with the possible exception of the JAM population) to spend the winter season in the Hudson Strait, Labrador Sea, southwest Davis Strait, Ungava Bay, and southeast Cumberland Sound among shifting pack-ice (Finley et al. 1982, Richard et al. 1990, Lewis et al. 2009, Luque and Ferguson 2010, Watt et al. 2016), but the location of each independent population is poorly understood. Satellite tracking data suggests that CSB beluga whales overwinter in a persistent polynya located in the southeast side of the Sound (Richard and Stewart 2008, Watt et al. 2016), but it is not known if some CSB

whales move to adjacent wintering waters and breed with beluga whales belonging to other populations, or if individuals from populations other than CSB overwinter in the southeast side of Cumberland Sound and thus have the possibility to interbreed. These results and the inferences above highlight that our knowledge of beluga migration, dispersal, and population gene flow (either through dispersal/emigration to different summer grounds or breeding) is still limited (O'Corry-Crowe et al. 1997, 2018, Turgeon et al. 2012, Colbeck et al. 2013).

MITOCHONDRIAL DNA UNDERESTIMATED THE PROPORTION OF CSB POPULATION IN CUMBERLAND SOUND DURING SUMMER

Our results confirm that the CSB population has the largest collection of private haplotypes (N = 15) shared among a large proportion (35%) of beluga whales compared to populations summering in Hudson Bay (this study, Parent et al. 2023). The beluga whale social system is hypothesized to center around mothers and their offspring, with females remaining with their maternal pod and males dispersing to join other male pods (O'Corry-Crowe et al. 2018). Beluga whales appear to develop migratory culture via social learning of migration routes and destinations, through females who are less likely to disperse to other areas, which facilitates population genetic and evolutionary divergence over time (Colbeck et al. 2013, O'Corry-Crowe et al. 2018). Annual migrations of beluga whales to summering groups may be facilitated by social culture, while common wintering groups may facilitate outbreeding (O-Corry-Crowe et al. 2018). A combination of genetic and satellite telemetry studies have revealed that genetically distinct subpopulations of beluga whales follow migratory corridors that have been maintained over multiple generations, and that some corridors overlap in space if not in time (Citta et al 2016, O'Corry-Crowe et al. 2018). A loss of unique genetic lineages may result in a loss of knowledge regarding migration routes and summering areas.

Information provided from mtDNA about maternal lineages private to CSB was underestimating the proportion of CSB population present in Cumberland Sound during summer. The nDNA analyses showed that 74% of beluga whales harvested in Cumberland Sound were from the CSB population (N = 20 out of 27 with PCA and ADMIXTURE). Such a result suggests the importance of maternal lineages in the population structure of beluga whales; however, it also highlights that mtDNA classification is performing rather poorly at discriminating individuals from the CSB population. While no other tools are available, it would be more prudent to assume most animals harvested during the summer in Cumberland Sound are from CSB than using mtDNA classification results. This study highlights the urgency of developing a new tool to monitor harvesting from the CSB population.

A subset of nDNA SNPs could be targeted to discriminate reliably the CSB and WHB populations summering in Cumberland Sound. The subset of nDNA SNPs should, however, be selected from a larger dataset including more populations of beluga whales from the Hudson Bay-Strait Complex. Those nDNA SNPs could then be used to discriminate more than two populations reliably and putative migrants across regions. Each sample harvested could be genotyped using massive parallel sequencing and an AmpliSeq™ approach (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Population assignment using these SNPs could then be performed and contribute to a genetic monitoring program, such as the one conducted for mtDNA and Nunavik beluga whale harvests (Bonnet et al. in prep.[1]). Such information would be highly valuable to monitor precisely the impact of harvesting on beluga populations.

MIGRATION IN AND OUT OF CUMBERLAND SOUND

Our analyses of nDNA and mtDNA markers supported the presence of two genetic populations in Cumberland Sound during summer. Of particular importance for management is to determine

if beluga whales from the two populations can be discriminated from one another. Inuit knowledge suggests the two types display different behaviour patterns and have morphological differences (Kilabuk 1998). Although we cannot verify this currently from these analyses, we did evaluate where and when beluga whales from the two populations were harvested and investigated two hunt samples indicated by hunters to come from another population based on traditional knowledge and neither had haplotypes private to Cumberland Sound.

We showed, using nDNA SNPs, that both CSB and WHB populations are harvested in the Northern and Western Strata. Consequently, CSB and WHB populations overlap in space within Cumberland Sound. Seven out of 27 beluga whales harvested in Cumberland Sound were similar to the WHB population (mostly males), but it is unknown if and when migration occurred, and how frequent it is. Past studies have shown sexual segregation among adult whales, with males and females inhabiting different regions at certain times of year (Richard et al. 2001, Krasnova et al. 2012). Our study also suggests that the observed difference in adult movements by different sexes may be associated with distinct populations. Limited satellite telemetry data support that beluga whales with CSB private and non-private haplotypes remained in Cumberland Sound year-round, but the number of tags that lasted into the winter was only seven (Table 1). Whether or not they are effective migrants, i.e., WHB beluga whales reproducing with CSB beluga whales, is currently unknown and would necessitate larger sample sizes, as highlighted earlier.

This assessment compared whales hunted in Cumberland Sound to those harvested in Western Hudson Bay and did not evaluate samples from whales harvested from the Eastern High Arctic-Baffin Bay beluga population that is morphologically similar to Cumberland Sound whales (Ferguson et al. 2020). Based on survey data it is assumed the Eastern High Arctic – Baffin Bay beluga whales overwinter in the North Water Polynya (Finley and Renaud 1980, Richard et al. 1998), and west Greenland (Heide-Jørgensen et al. 1993). Based on satellite tagging data the majority of whales remain near the North Water and approximately 8,000 beluga migrate farther south to west Greenland (Richard et al. 2001, Heide-Jørgensen et al. 2017), although this population was considerably larger prior to overharvesting (Innes and Stewart 2002). During the late winter mating season both west Greenland and Hudson Bay beluga populations are similar distances from the Cumberland Sound population; however, during summer it is more likely, based on relative population abundance, that whales from the Western Hudson Bay population may move into Cumberland Sound. It would be useful to include samples from the Eastern High Arctic-Baffin Bay beluga population in future genetic assessments.

DESCRIPTION OF THE CUMBERLAND SOUND BELUGA WHALE STOCK

Our study showed that most beluga whales summering in Cumberland Sound form a distinct population, of which a large proportion has private maternal lineages. We also showed that there are some migrants, most likely from WHB, inhabiting Cumberland Sound. Therefore, based on the results presented here, two beluga whale populations inhabit Cumberland Sound during summer. Consequently, Cumberland Sound beluga whales should be defined as a stock based on mtDNA and nDNA evidence. New tools to monitor the harvest from the CSB population should be employed to monitor the harvest and could potentially inform on how to reduce harvesting impacts on the CSB population in the future. The evidence of two populations inhabiting Cumberland Sound during summer is in agreement with the Pangnirtung community who have maintained that multiple groups of beluga whales summer in this area. There were no spatial or temporal characteristics to enable identification of individuals from each of the two populations in Cumberland Sound; however, it would be worth investigating any possible association between genetic, physical, and behavioural traits in future studies.

Using nDNA we estimated that 74% (20 out of 27 samples) of beluga whales from Cumberland Sound were from the CSB population. Theoretically, this may indicate the survey estimate for CSB whales is an over-estimate. We could not identify spatial or temporal variability in proportions of CSB and WHB in different surveyed regions of Cumberland Sound due to the sample size analyzed. Such variation would be important to characterise as there are different densities of beluga whales in the different surveyed regions across Cumberland Sound (Watt et al. 2021), which could increase or decrease the proportion of animals associated to the CSB population.

Our results show that hunters in Pangnirtung harvesting beluga whales in Cumberland Sound are mostly targeting the CSB population. The estimated proportion of beluga whales from the WHB population (26%) was low and mostly comprised of males. It is unknown whether those whales from WHB were seasonal or permanent migrants to Cumberland Sound. It is also unknown if migration is sporadic or recurrent. Based on current information, managing beluga whales inhabiting Cumberland Sound in the summer as a single stock comprised of two genetic populations is the most precautionary approach.

ACKNOWLEDGEMENTS

We are grateful to the hunters from Nunavik and Pangnirtung for collecting tissue samples from their beluga whales. We also thank all members of the Cumberland Sound Beluga Working Group, particularly the Pangnirtung Hunters and Trappers Association, for supporting and advocating for the Marine Mammal Sampling Program. Our gratitude goes also to Thomas Wood and Denise Tenkula for assistance with tissue subsampling, Audrey Bourret for assistance with genomic analyses, and Claudie Bonnet for managing the beluga whale dataset. Finally, we thank the CSAS review committee whose comments improved the final version of this research document.

REFERENCES CITED

Aguilar, A. 1987. Using organochlorine pollutants to discriminate marine mammal populations: a review and critique of the methods. Mar. Mamm. Sci. 3(3): 242–262.

Alexander, D.H., Novembre, J., and Lange, K. 2009. Fast model-based estimation of ancestry in unrelated individuals. Gen. Res. 19: 1655–1664.

Andrews, S. 2010. FastQC: A quality control tool for high throughput sequence data. Cambridge, England.

Bolger, A.M., Lohse, M., and Usadel, B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30(15): 2114–2120.

Brown Gladden, J.G., Ferguson, M.M., and Clayton, J.W. 1997. Matriarchal genetic population structure of North American beluga whales *Delphinapterus leucas* (Cetacea: Monodontidae). Mol. Ecol. 6(11): 1033–1046.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., and Cresko, W.A. 2013. Stacks: An analysis tool set for population genomics. Mol. Ecol. 22(11): 3124–3140.

Cattell, R.B. 1966. The scree test for the number of factors. Multivariate Behav. Res. 1(2): 245–76.

Citta, J.J., Quakenbush, L.T., Frost, K.J., Lowry, L.F., Hobbs, R.C., and Aderman, H. 2016. Movements of beluga whales (*Delphinapterus leucas*) in Bristol Bay, Alaska. Mar. Mamm. Sci. 32(4): 1271–1298.

Colbeck, G.J., Duchesne, P., Postma, L.D., Lesage, V., Hammill, M.O., and Tugeon, J. 2013. Groups of related belugas (*Delphinapterus leucas*) travel together during their seasonal migrations in and around Hudson Bay. Proc. R. Soc. Biol. Sci. Ser. B 280(1752): 20122552.

COSEWIC. 2004. COSEWIC assessment and status report on the beluga whale *Delphinapterus leucas*. Comm. Status Endanger. Wildl. Canada. Ottawa. ix + 70 pp.

COSEWIC. 2020. COSEWIC assessment and status report on the Beluga Whale *Delphinapterus leucas*, Eastern High Arctic - Baffin Bay population, Cumberland Sound population, Ungava Bay population, Western Hudson Bay population, Eastern Hudson Bay population and James Bay population in Canada. Comm. Status Endanger. Wildl. Canada. Ottawa. xxxv + 84 pp.

Cullingham, C.I., Miller, J.M., Peery, R.M., Dupuis, J.R., Malenfant, R.M., Gorrell, J.C., and Janes, J.K. 2020. Confidently identifying the correct *K* value using the Δ*K* method: When does K = 2?. Mol. Ecol. 29(5): 862–869.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., and Durbin, R. 2011. The variant call format and VCFtools. Bioinformatics 27(15): 2156–2158.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H. 2021. Twelve years of SAMtools and BCFtools. GigaScience 10(2): 1–4.

de March, B.G.E., and Postma, L.D. 2003. Molecular genetic stock discrimination of belugas (*Delphinapterus leucas*) hunted in eastern Hudson Bay, Northern Quebec, Hudson Strait, and Sanikiluaq (Belcher Islands), Canada, and comparisons to adjacent populations. Arctic 56(2): 111–124.

de March, B.G.E., Maiers, L.D., and Friesen, M.K. 2002. An overview of genetic relationships of Canadian and adjacent populations of belugas (*Delphinapterus leucas*) with emphasis on Baffin Bay and Canadian eastern Arctic populations. NAMMCO Sci. Publ. 4: 17–38.

de March, B.G.E., Stern, G.A., and Innes, S. 2004. The combined use of organochlorine contaminant profiles and molecular genetics for stock discrimination of white whales (*Delphinapterus leucas*) hunted in three communities on southeast Baffin Island. J. Cetac. Res. Manag. 6(3): 241–250.

DFO. 2002. Stock Status Report E5-32.

DFO. 2019. Sustainable harvest advice for Cumberland Sound beluga based on the 2017 aerial survey and modelled abundance estimates. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2019/024.

Díaz-Arce, N. and Rodríguez-Ezpeleta, N. 2019. Selecting RAD-Seq data analysis parameters for population genetics: the more the better? Front Genet. 10: 533.

Dray, S., and Dufour, A. 2007. The ade4 Package: Implementing the Duality Diagram for Ecologists. J. Stat. Soft. 22(4): 1–20.

Dyke, A.S. 2004. An outline of North American deglaciation with emphasis on central and northern Canada. Dev. Quaternary Sci. 2: 373–424.

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. 2016. MultiQC: Summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32(19): 3047–3048.

Ferguson, S.H., Willing, C., Kelley, T.C., Boguski, D.A., Yurkowski, D.J., and Watt, C.A. 2020 Reproductive parameters for female beluga whales (*Delphinapterus leucas*) of Baffin Bay and Hudson Bay, Canada. Arctic 73(4): 405–420.

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

Finley, K.J., Miller, G.W., Allard, M., Davis, R.A., and Evans, C.R. 1982. The belugas (*Delphinapterus leucas*) of Northern Quebec: distribution, abundance, stock identity, catch history and management. Can. Tech. Rep. Fish. Aquat. Sci. 1123:1–32.

Funk, W.C., McKay, J.K., Hohenlohe, P.A., and Allendorf, F.W. 2012. Harnessing genomics for delineating conservation units. Trends Ecol. Evol. 27(9): 489–496.

Gruber, B., Unmack, P.J., Berry, O.F., and Georges, A. 2018. dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Mol. Ecol. Resour. 18(3): 691–699.

Hammill, M.O., Mosnier, A. and Bordeleau, X. 2021. An update of impacts of harvesting on the abundance of Nunavik beluga. DFO Can. Sci. Advis. Sec. Res. Doc. 2021/003. iv + 21 p.

Heide-Jørgensen, M.P., and Teilmann, J. 1994. Growth, reproduction, age structure, and feeding habits of white whales (*Delphinapterus leucas*) in West Greenland waters. Med. om Grǿnland Biosci. 39: 195–212.

Heide-Jørgensen, M.P., Lassen, H.,Teilmann, J., and Davis, R.A. 1993. An index of the relative abundance of wintering belugas, *Delphinapterus leucas*, and narwhals, *Monodon monoceros*, off West Greenland. Can. J. Fish. Aquat. Sci. 50(11): 2323–2335.

Heide-Jørgensen, M.P., Hansen, R.G., Fossette, S., Nielsen, N.H., Borchers, D.L., Stern, H., and Witting, L. 2017. Rebuilding beluga stocks in West Greenland. Anim. Conserv. 20(3): 282–293.

Holsinger, K.E., and Weir, B.S., 2009. Genetics in geographically structured populations: defining, estimating and interpreting FST. Nat. Rev. Gen. 10: 639–650.

Innes, S., and Stewart, R.E. 2002. Population size and yield of Baffin Bay beluga (*Delphinapterus leucas*) stocks. NAMMCO Sci. Pub. 4: 225–238.

Innes, S., Muir, D.C., Stewart, R.E., Heide-Jørgensen, M.P., and Dietz, R. 2002. Stock identity of beluga (*Delphinapterus leucas*) in Eastern Canada and West Greenland based on organochlorine contaminants in their blubber. NAMMCO Sci. Publ. 4: 51–68.

International Whaling Commission (IWC). 2002. Annex I - Report of the Working Group on Stock Definition. In: Report of the Scientific Committee. J. Cetacean Res. Manage. 4: 47–52.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24(11): 1403–1405.

Jombart, T., and Ahmed, I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27(21): 3070–3071.

Kilabuk, P. 1998. A study of Inuit knowledge on the Southeast Baffin Beluga. Report prepared for The Southeast Baffin Beluga Management Committee. Nunavut Wildlife Management Board, Iqaluit, Canada. 74pp.

Knaus, B.J., and Grünwald, N.J. 2017. vcfr: a package to manipulate and visualize variant call format data in R. Mol. Ecol. Resour. 17(1): 44–53.

Krasnova, V.V., Chernetsky, A.D., Kirillova, O.I., and Bel'kovich, V.M. 2012. The dynamics of the abundance, age, and sex structure of the Solovetsky reproductive gathering of the beluga whale *Delphinapterus leucas* (Onega Bay, White Sea). Russ. J. Marine Biol. 38: 218–225.

Kucheravy, C.E., Trana, M.R., Watt, C.A., Roth, J.D., Tomy, G.T., Anderson, W.G., and Ferguson, S.H. 2022. Blubber cortisol in four Canadian beluga whale populations is unrelated to diet. Mar. Ecol. Prog. Ser. 698: 171–189.

Lawson, D.J., van Dorp, L., and Falush, D. 2018. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. Nat. Commun. 9(3258).

Leigh, J.W., and Bryant, D. 2015. PopART: Full-feature software for haplotype network construction. Methods Ecol. Evol. 6(9): 1110–1116.

Lewis, A.E., Hammill, M.O., Power, M., Doidge, D.W., and Lesage, V. 2009. Movement and aggregation of eastern Hudson Bay Beluga whales (*Delphinapterus leucas*): a comparison of patterns found through satellite telemetry and Nunavik Traditional Ecological Knowledge. Arctic 62(1): 13–24

Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.

Li, H., and Durbin, R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26(5): 589–595.

Lillie, W.R., Brown Gladden, J.G., and Tretiak, D.N. 1996. Amplification and sequencing of control region mitochondrial DNA from the beluga whale, *Delphinapterus leucas*. Can. Tech. Rep. Fish. Aquat. Sci. 2080: iv + 8pp.

Lowe, W.H., and Allendorf, F.W. 2010. What can genetics tell us about population connectivity? Mol. Ecol. 19(15): 3038–3051.

Luque, S.P., and Ferguson, S.H. 2010. Age structure, growth, mortality, and density of belugas (*Delphinapterus leucas*) in the Canadian Arctic: responses to environment? Polar Biol. 33: 163–178.

Luu, K., Bazin, E., and Blum, M.G.B. 2017. pcadapt: an R package to perform genome scans for selection based on principal component analysis. Mol. Ecol. Resourc. 17(1): 67–77.

Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.M. 2010. Robust relationship inference in genome-wide association studies. Bioinformatics 26(22): 2867–2873.

Matthews, C.J.D., and Ferguson, S.H. 2014. Validation of dentin deposition rates in beluga whales by interspecies cross dating of temporal $\delta^{13}C$ trends in teeth. NAMMCO Sci. Publ. 10.

McKinney, M.A., De Guise, S., Martineau, D., Béland, P., Lebeuf, M., and Letcher, R.J. 2006. Organohalogen contaminants and metabolites in beluga whale (*Delphinapterus leucas*) liver from two Canadian populations. Environ. Toxicol. Chem. 25(5): 1246–1257.

Mitchell, E., and Reeves, R.R. 1981. Catch history and cumulative catch estimates of initial population size of cetaceans in the Eastern Canadian Arctic. Rep. Int. Whal. Comm. 31.

Newsome, S.D., Clementz, M.T., and Koch, P.L. 2010. Using stable isotope biogeochemistry to study marine mammal ecology. Mar. Mamm. Sci. 26(3): 509–572.

O'Corry-Crowe, G.M., Suydam, R.S., Rosenberg, A., Frost, K.J., and Dizon, A.E. 1997. Phylogeography, population structure and dispersal patterns of the beluga whale *Delphinapterus leucas* in the western Nearctic revealed by mitochondrial DNA. Mol. Ecol. 6(10): 955–970.

O'Corry-Crowe, G., Suydam, R., Quakenbush, L., Potgieter, B., Harwood, L., Litovka, D, Ferrer, T., Citta, J., Burkanov, V., Frost, K., and Mahoney, B. 2018. Migratory culture, population structure and stock identity in North Pacific beluga whales (*Delphinapterus leucas*). PLOS one. 13(3).

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Solymos, P., Stevens, M., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H., Fitz, J.R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M., Lahti, L., McGlinn, D., Ouellette, M., Ribeiro Cunha, E., Smith, T., Stier, A., Ter Braak, C., and Weedon, J. 2022. *vegan*: Community Ecology Package. R package version 2.6-2.

O'Leary, S.J., Puritz, J.B., Willis, S.C., Hollenbeck, C.M. and Portnoy, D.S. 2018. These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists. Mol. Ecol. 27(16): 3193–3206.

Pagès, H., Aboyoun, P., Gentleman, R., and DebRoy, S. 2023. Biostrings: efficient manipulation of biological strings. R package version 2.62.0,

Parent, G.J., Mosnier, A., Montana, L., Cortial, G., St-Pierre, A.P., Bordeleau, X., Lesage, V., Watt, C., Postma, L., and Hammill, M.O. 2023. Re-examining populations of beluga in the Hudson Bay-Strait Complex and assessing the impact on harvests in Nunavik and Sanikiluaq management units. DFO Can. Sci. Advis. Sec. Res. Doc. 2023/004. iv + 31 p.

Postma, L.D. 2017. Genetic diversity, population structure and phylogeography among belugas (*Delphinapterus leucas*) in Canadian waters: broad to fine-scale approaches to inform conservation and management strategies. Thesis (Ph.D) University of Manitoba, Winnipeg, MB. 314 p.

Postma, L.D., Petersen, S.D., Turgeon, J., Hammill, M.O., Lesage, V., and Doniol-Valcroze, T. 2012. Beluga whales in James Bay: a separate entity from eastern Hudson Bay belugas? DFO Can. Sci. Advis. Sec. Res. Doc. 2012/074. iii + 23 p.

Privé, F., Luu, K., Vilhjálmsson, B.J., and Blum, M.G.B. 2020. Performing highly efficient genome scans for local adaptation with R Package pcadapt Version 4. Mol. Biol. Evol. 37(7): 2153–2154.

R Core Team. 2022. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

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

Richard, P.R., and Orr, J.R. 1986. A review of the status and harvest of white whales (*Delphinapterus leucas*) in the Cumberland Sound area, Baffin Island. Can. Tech. Rep. Fish. Aquat. Sci. 1477: iv + 25 p.

Richard, P., and Stewart, D.B. 2008. Information Relevant to the Identification of Critical Habitat for Cumberland Sound Belugas (*Delphinapterus leucas*). DFO Can. Sci. Advis. Secr. Res. Doc. 2008/085: iii + 24 p.

Richard, P.R., Orr, J.R., and Barber, D.G. 1990. The distribution and abundance of belugas, *Delphinapterus leucas*, in eastern Canadian subarctic waters: a review and update. In: Smith, E.G., St. Aubin, J., Geraci, J.R. (eds) Advances in research on the beluga whale, *Delphiapterus leucas*. Can. Bull. Fish. Aquat. Sci. 224: 23–38

Richard, P.R., Heide-Jørgensen, M.P., and St. Aubin, D. 1998. Fall movements of belugas (*Delphinapterus leucas*) with satellite-linked trasnmitters in Lancaster Sound, Jones Sound, and Northern Baffin Bay. Arctic 51(1): 5–16.

Richard, P.R., Martin, A.R., and Orr, J.R. 2001. Summer and autumn movements of belugas of the eastern Beaufort Sea stock. Arctic. 54(3): 223–236.

Rochette, N.C., Rivera-Colón, A.G., and Catchen, J.M. 2019. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. Mol. Ecol. 28(21): 4737–4754.

Sergeant, D.E., and Brodie, P.F. 1969. Body size in white whales, *Delphinapterus leucas*. J. Fish. Res. Board Canada. 26(10): 2561–2580.

Seutin, G., White, B.N., and Boag, P.T. 1991. Preservation of avian blood and tissue samples for DNA analyses. Can. J. Zool. 69(1): 82-90.

Smythe, T.A., Loseto, L.L., Bignert, A., Rosenberg, B., Budakowski, W., Halldorson, T., Pleskach, K., and Tomy, G.T. 2018. Temporal trends of brominated and fluorinated contaminants in Canadian Arctic beluga (*Delphinapterus leucas*). Arct. Sci. 404(3): 388–404.

Stewart, R.E.A. 2008. Redefining walrus stocks in Canada. Arctic. 61(3): 292-308.

Turgeon, J., Duchesne, P., Colbeck, G.J., Postma, L.D., and Hammill, M.O. 2012. Spatiotemporal segregation among summer stocks of beluga (*Delphinapterus leucas*) despite nuclear gene flow: Implication for the endangered belugas in Eastern Hudson Bay (Canada). Conserv. Genet. 13: 419–433.

Waples, R.S., and Gaggiotti, O. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol. Ecol. 15(6): 1419–1439.

Watt, C.A., Orr, J., and Ferguson, S.H. 2016. A shift in foraging behaviour of beluga whales *Delphinapterus leucas* from the threatened Cumberland Sound population may reflect a changing Arctic food web. Endanger. Species Res. 31: 259–270.

Watt, C.A., Marcoux, M., Ferguson, S.H., Hammill, M.O., and Matthews, C.J.D. 2021. Population dynamics of the threatened Cumberland Sound beluga (*Delphinapterus leucas*) population. Arct. Sci. 7(2): 545–566.

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

Wolff, J. O. 1997. Population regulation in mammals: an evolutionary perspective. J. Anim. Ecol. 66(1): 1–13.

TABLES AND FIGURES

*Table 1. Beluga whales tagged in Cumberland Sound from 1998-2007 and their dominant mtDNA haplotype. *Indicates private haplotypes. Shading indicates whales which were tracked into the winter season and all remained within Cumberland Sound. Note that beluga whales tagged in September were not used for analyses in this study.*

Year	Tag	Start	End	Duration	Sex	mtDNA
1998	17000	Aug 19	Aug 31	12	M	HL055
1998	17001	Aug 20	Oct 27	68	M	NA
1998	20682	Aug 25	Nov 8	75	M	NA
1998	20683	Aug 30	Sept 4	5	M	HL024*
1998	20684	Aug 30	Oct 25	56	F	HL003
1998	20685	Aug 30	Nov 3	65	F	NA
1999	20162	Sept 3	Jan 16	135	M	NA
1999	20682	Sept 3	Nov 23	81	F	HL003
1999	20683	Sept 8	Sept 10	2	F	HL022
1999	20684	Sept 3	Dec 7	95	M	HL050
1999	20685	Sept 5	Dec 11	97	M	HL003
1999	7926	Sept 8	Nov 18	71	M	NA
2006	57594	July 17	March 3	229	F	HL003
2007	37023	July 12	Nov 3	114	F	HL040*
2007	57602	July 12	Nov 15	126	F	HL057*
2008	39296	Sept 5	May 14	251	F	HL028*
2008	39308	Sept 6	June 3	270	F	HL040*
2008	39323	Sept 9	April 28	231	M	HL028*
2008	40623	Sept 9	Sept 16	7	M	HL003

Table 2. Sample sizes of the five beluga whale populations used for genetic and genomic analyses of mtDNA and nDNA presented in this study. Note that the mtDNA haplotype was not available for all samples used for nDNA analyses, hence the total sample size (N) is greater than the sample size for mtDNA analyses (N_{mtDNA}).

Population	**N**	**Years**	**N_{mtDNA}**	**$Years_{mtDNA}$**	**N_{nDNA}**	**$Years_{nDNA}$**
Cumberland Sound	191	1982 - 2021	189	1982 - 2021	36*	2002 - 2007
Eastern Hudson Bay	183	1994 - 2019	183	1994 – 2019	NA	NA
Belcher Islands	78	1994 - 2020	78	1994 - 2020	NA	NA
James Bay	78	2003 - 2021	78	2003 - 2021	NA	NA
Western Hudson Bay	350	1985 - 2015	318	1985 - 2015	146*	1992 - 2015

* Final samples size for statistical analyses after quality control and SNPs filtration: N_{CSB} = 27, N_{WHB} = 121 (N = 148).

Table 3. Beluga whales harvested in Cumberland Sound from 2002-2007 that were assessed for nDNA and their ADMIXTURE probability to the CSB population. The asterisk in the nDNA column indicates individuals with mtDNA haplotypes that were private to the CSB population.

Sample	Date	Latitude	Longitude	Sex	Age	ADMIXTURE probability ± SE	nDNA
ARPG-02-1034	22/07/02	NA	NA	M	30	0.63 ± 0.08	CSB
ARPG-02-1037	NA/07/02	66.32	-67.10	M	32	0.53 ± 0.03	CSB
ARPG-02-1039	10/07/02	NA	NA	M	NA	1.00 ± 0.00	CSB
ARPG-02-1040	21/07/02	NA	NA	F	18	0.61 ± 0.03	CSB
ARPG-02-1054	06/07/02	65.97	-67.02	F	17	0.60 ± 0.03	CSB
ARPG-02-1072	25/07/02	NA	NA	F	NA	0.57 ± 0.04	CSB*
ARPG-02-1077	06/07/02	65.97	-67.02	M	38	0.58 ± 0.05	CSB*
ARPG-02-1181	22/07/02	65.22	-65.75	F	23	1.00 ± 0.00	CSB
ARPG-02-1198	22/07/02	NA	NA	F	NA	0.60 ± 0.03	CSB
PGDL-02-02	NA/07/02	65.24	-66.66	F	NA	1.00 ± 0.00	CSB*
PGDL-02-03	NA/07/02	65.24	-66.66	M	NA	1.00 ± 0.00	CSB*
PGDL-02-05	NA/07/02	65.24	-66.66	M	NA	0.65 ± 0.03	CSB
ARPG-05-1211	12/07/05	NA	NA	F	29	0.14 ± 0.05	WHB*
ARPG-05-1220	03/08/05	NA	NA	M	24	0.20 ± 0.04	WHB
ARPG-05-1225	03/08/05	NA	NA	M	22	0.08 ± 0.06	WHB
ARPG-05-1230	NA/NA/05	66.15	-65.70	M	NA	0.21 ± 0.07	WHB
ARPG-05-1232	04/08/05	65.22	65.75	M	31	0.10 ± 0.05	WHB
ARPG-05-1241	09/07/05	NA	NA	M	6	0.12 ± 0.07	WHB
ARPG-05-1255	03/08/05	NA	NA	M	17	1.00 ± 0.09	CSB*
ARPG-06-1265	01/07/06	NA	NA	M	41	0.63 ± 0.04	CSB
ARPG-06-1280	01/07/06	65.97	67.02	M	16	0.58 ± 0.04	CSB*
ARPG-06-1281	30/06/06	65.97	67.02	M	14	0.23 ± 0.04	WHB
ARPG-07-01	12/07/07	66.27	-67.11	F	NA	1.00 ± 0.05	CSB*
ARPG-07-02	12/07/07	66.27	-67.11	F	NA	0.60 ± 0.05	CSB*
ARPG-07-1314	04/07/07	NA	NA	F	26	0.68 ± 0.05	CSB*
ARPG-07-1317	04/07/07	NA	NA	F	13	0.59 ± 0.04	CSB
ARPG-07-1328	04/07/07	66.32	-67.63	M	29	0.63 ± 0.04	CSB*

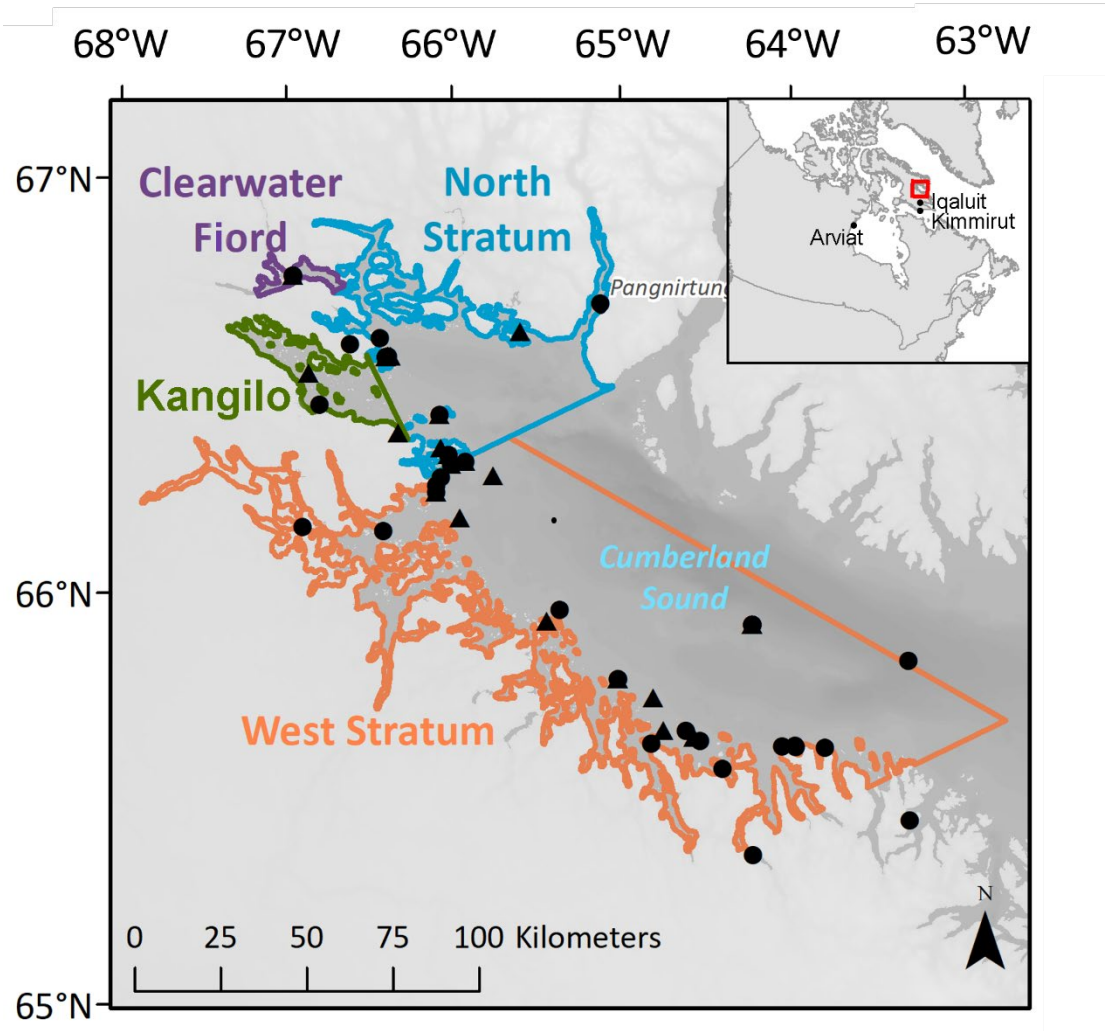

Figure 1. Geographic location of 98 out of 207 beluga whales harvested in Cumberland Sound that were haplotyped for the control region of mitochondrial DNA. Circles indicate haplotypes shared with one of the four other reference populations, whereas the triangles indicate private haplotypes to Cumberland Sound beluga whales.

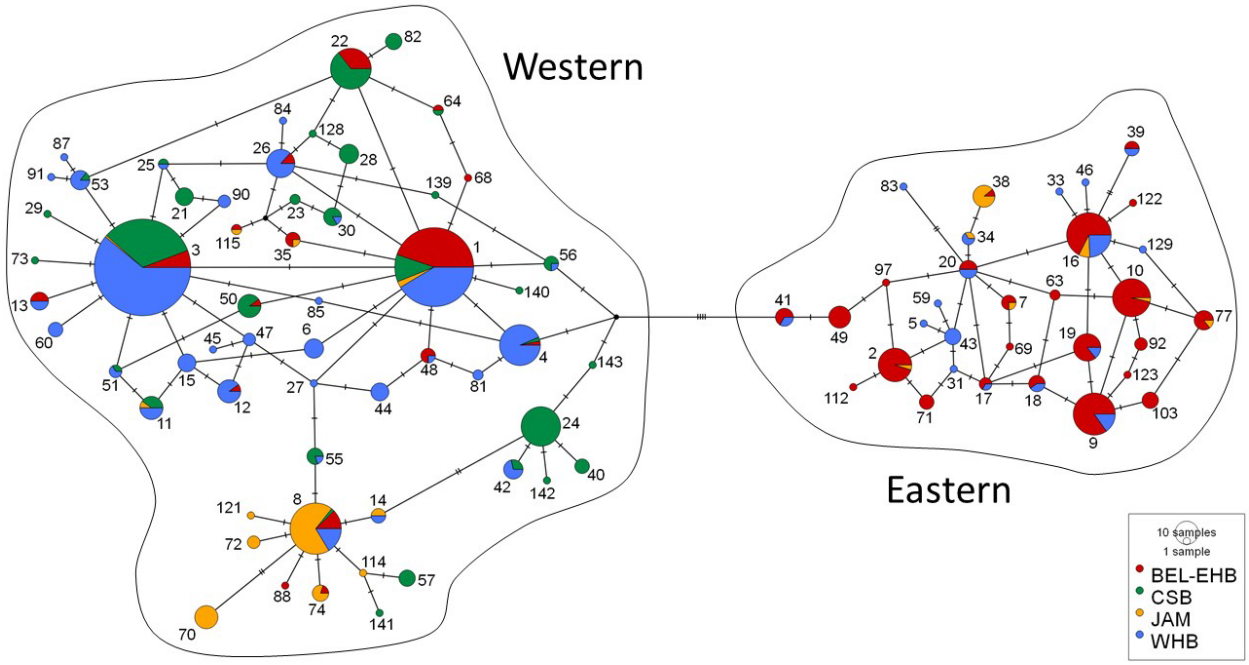

Figure 2. Haplotype network for the mitochondrial DNA control region from beluga whales from four reference groups (BEL-EHB, CSB, JAM, and WHB) in the Hudson Bay-Strait Complex. A statistical parsimony (TSC) network using PopArt is presented. Small perpendicular bars along lines between two haplotypes indicate the number of mutations between haplotypes. Black circles lacking haplotype numbers indicate missing haplotypes in the evolution of the network. The two haplogroups identified in previous studies (Eastern and Western) are indicated (Postma 2017, Parent et al. 2023).

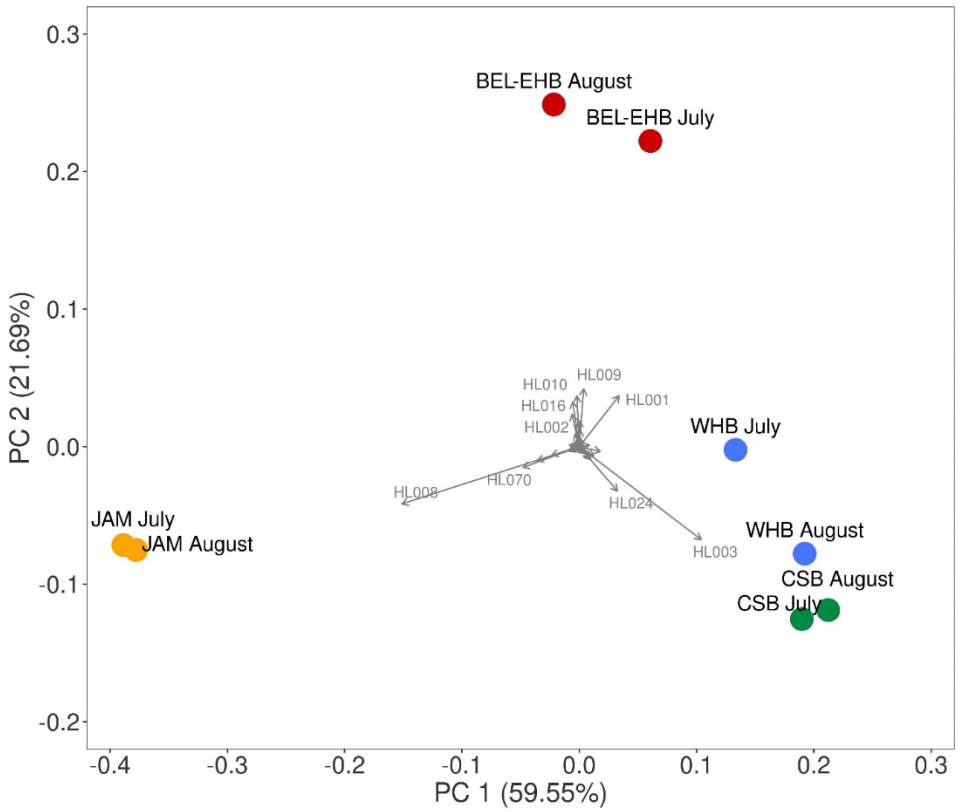

Figure 3. Genetic distinctiveness of Cumberland Sound beluga whales (CSB) with a biplot from the principal component analyses (PCA) using haplotype frequencies of mitochondrial DNA control regions for beluga whales from four reference groups (CSB, Western Hudson Bay (WHB), James Bay (JAM), and Belchers-Eastern Hudson Bay (BEL-EHB)). The length and direction of arrows indicate the effect of the haplotypes on the distance between reference groups. PCA figure made using ggplot2, ggrepel and factoextra.

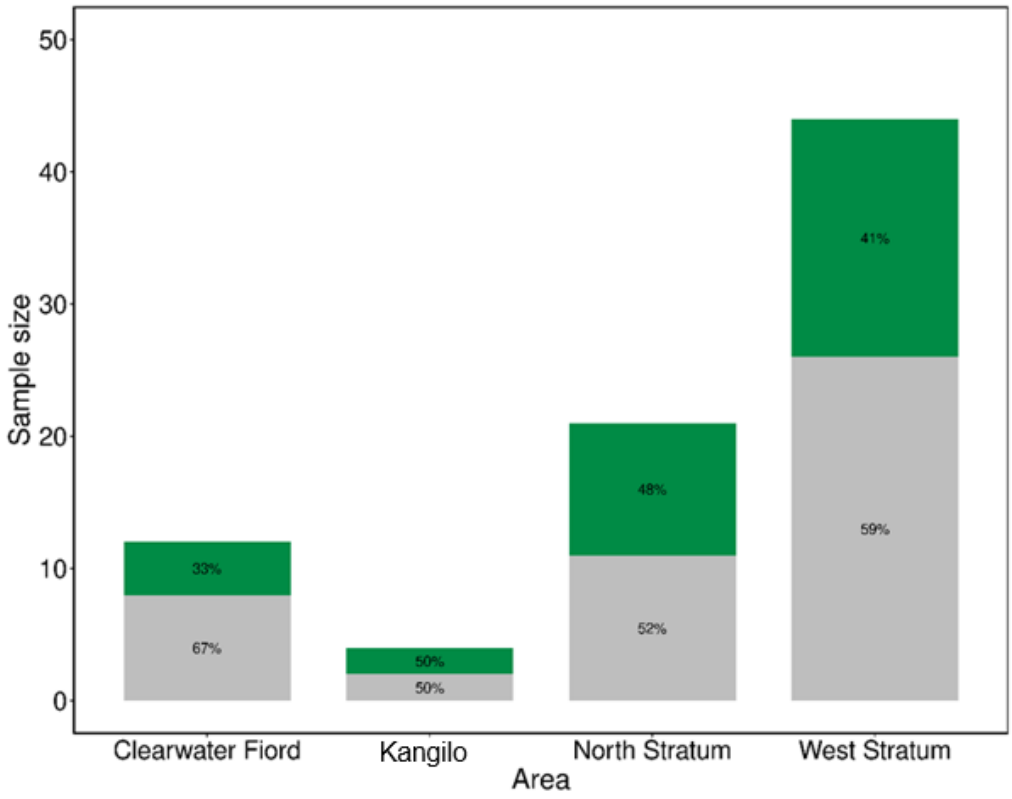

Figure 4. Proportions of shared (grey) and private (green) mitochondrial DNA haplotypes of the control region during summer (July, August) across four survey areas in Cumberland Sound (see Figure 1). Note the difference in sample sizes between areas ($N_{Clearwarter\ Fiord}$ = 12; $N_{Kangilo}$ = 4; $N_{North\ Stratum}$ = 21; $N_{West\ Stratum}$ = 44).

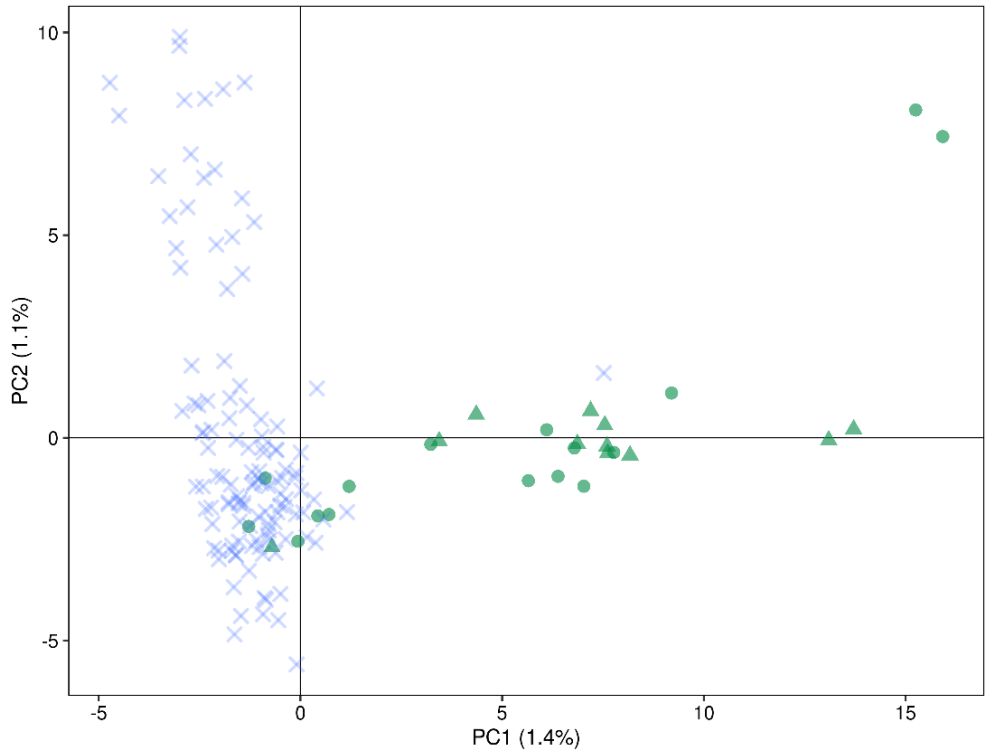

Figure 5. Principal Component Analysis (PCA) using 12,370 nuclear DNA loci and showing genetic distinctiveness between Cumberland Sound (CSB, green) and Western Hudson Bay (WHB, blue) beluga whales. Within CSB, mitochondrial information is also provided: triangles indicate private haplotypes to CSB, while circles represent shared haplotypes with other reference groups.

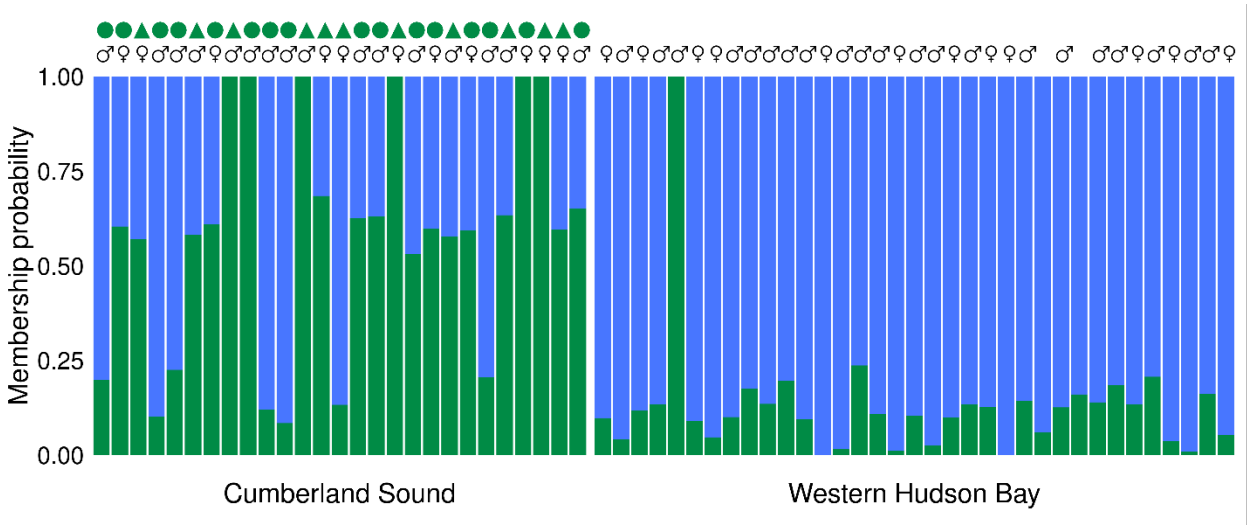

Figure 6. ADMIXTURE analysis (nDNA) results for two groups. Sex for each individual (vertical bars) are indicated on top (except two which could not be determined with qPCR, see Bonnet et al. (in prep.[1]) for details). Private CSB (triangle) and shared (circles) mtDNA haplotypes are indicated.

APPENDIX A. SUPPLEMENTARY INFORMATION

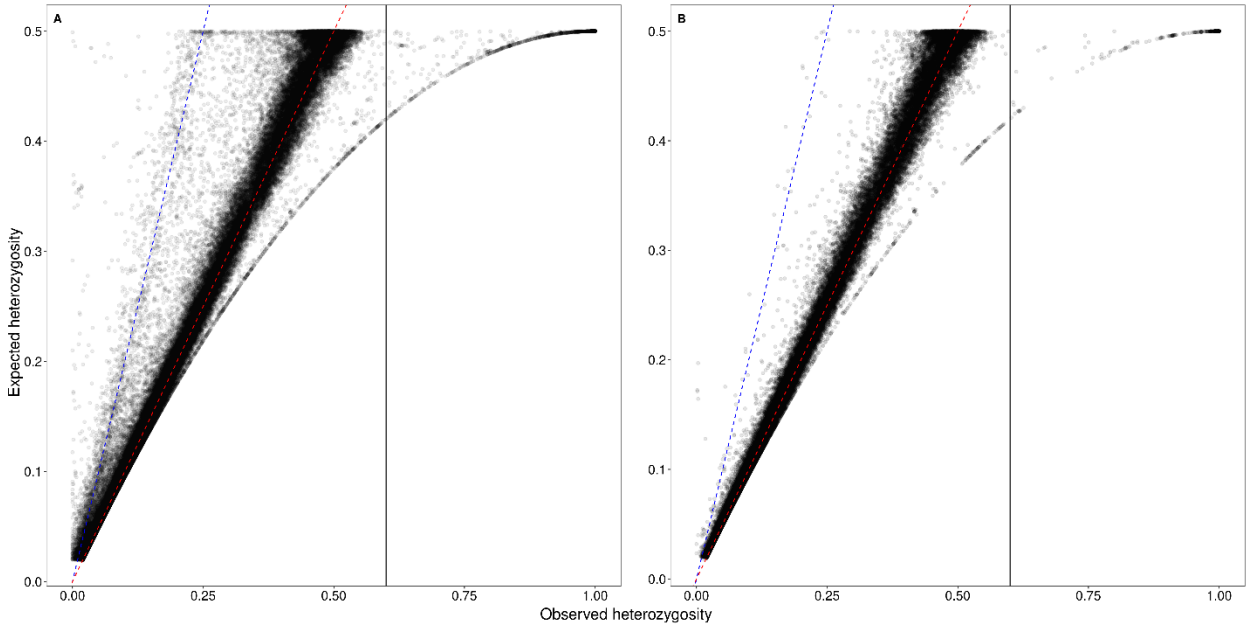

Figure A1. Observed versus expected heterozygosity for nuclear DNA (nDNA) loci before (A = 136,884 loci) and after (B = 96,232 loci) removing loci with low (< 12) and high (> 29) median read depth. The black vertical line at observed heterozygosity = 0.6 represents the threshold used in one of the filtering step to remove highly heterozygous loci. The blue and red dashed lines have slopes = 2 and 1, respectively (both α = 0). Loci close to the red dashed line have similar observed and expected heterozygosity whereas loci close to the blue dashed line have half of expected heterozygosity.

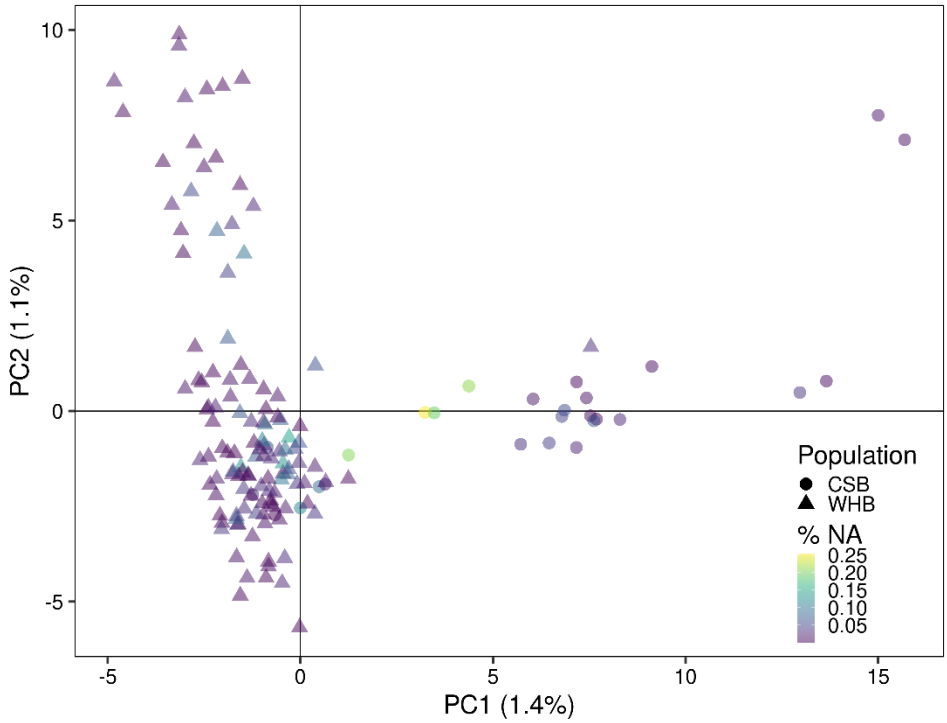

Figure A2. Proportion of nuclear DNA (nDNA) loci with missing data for each beluga whale used in the Principal Component Analysis (PCA; Figure 5).

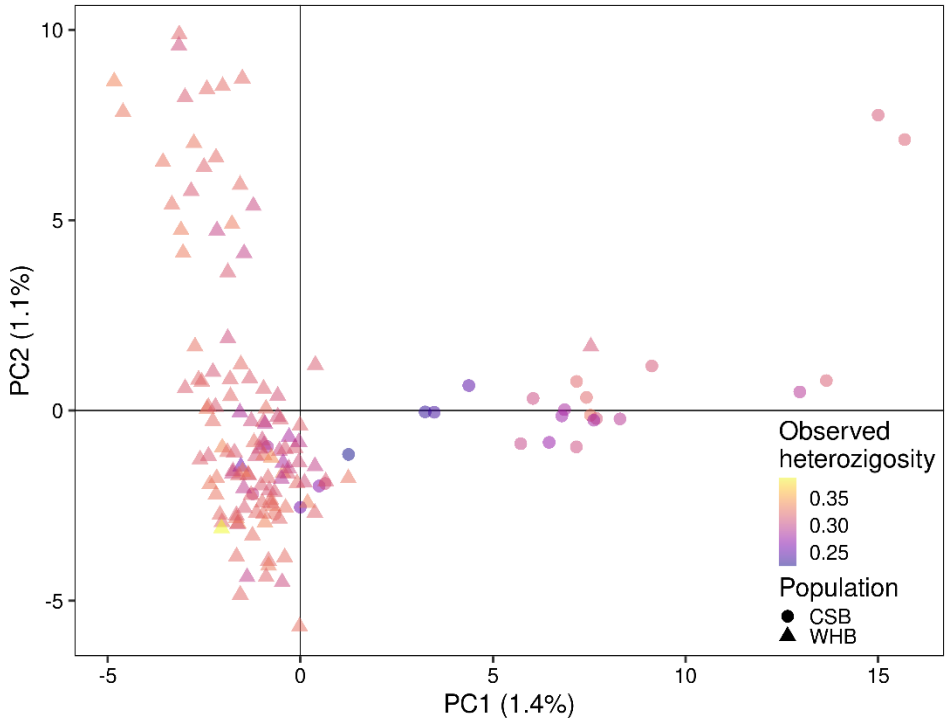

Figure A3. Heterozygosity estimated from nuclear DNA (nDNA) loci for each beluga whale from the Principal Component Analysis (PCA; Figure 5).

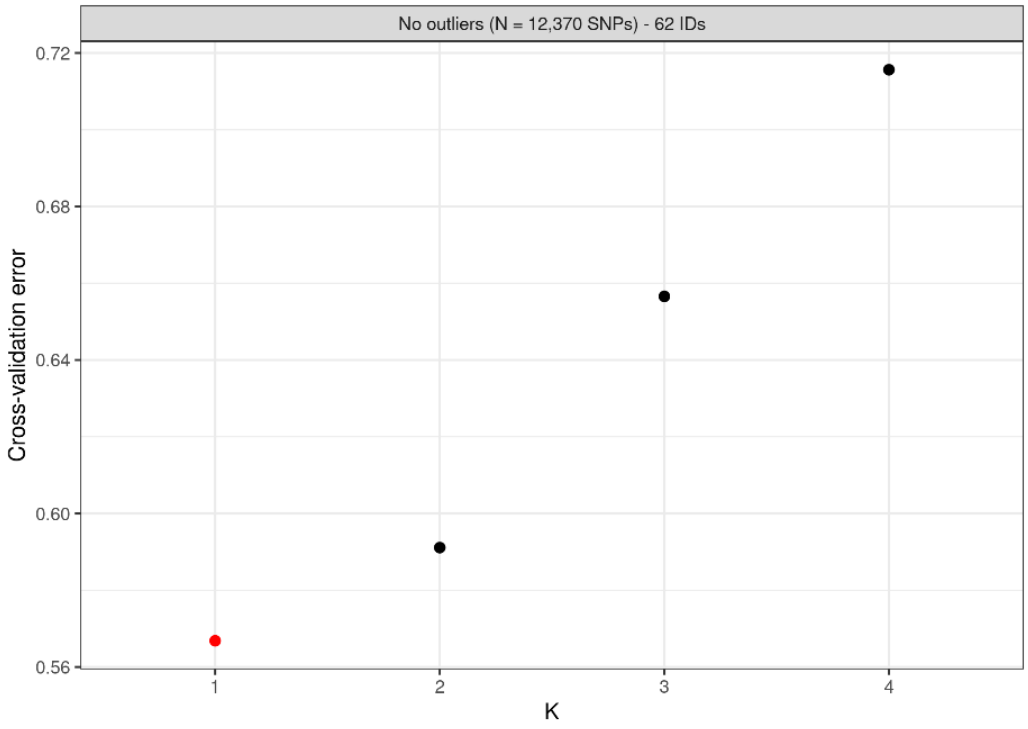

Figure A4. Cross validation approach used to infer the number of populations (K) detected by ADMIXTURE (Alexander et al. 2009) with 27 belugas harvested in Cumberland Sound and 35 belugas harvested in Western Hudson Bay. The red point indicates the lowest CV value.

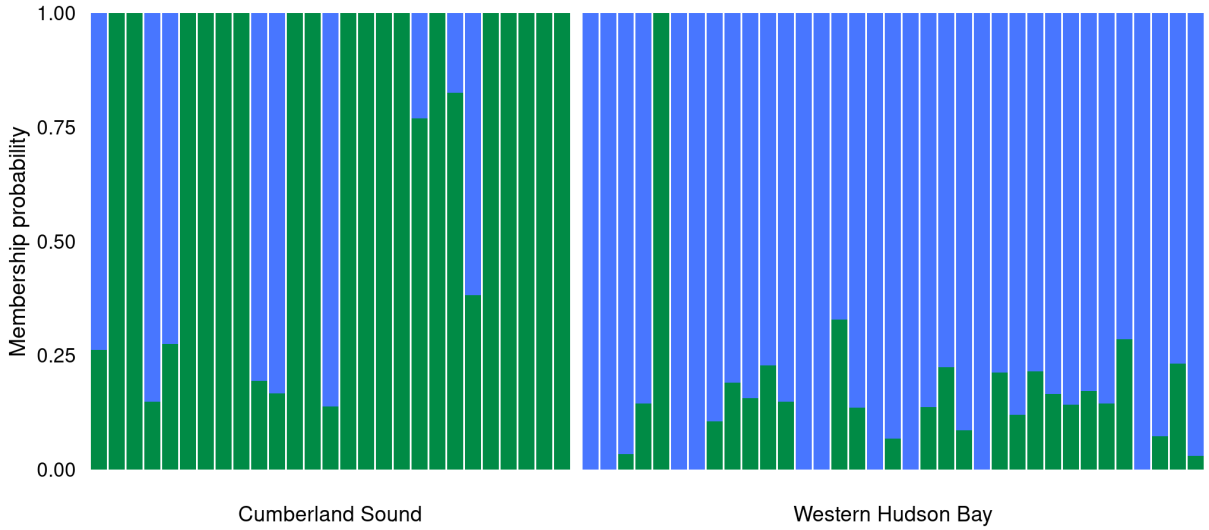

Figure A5. ADMIXTURE analysis (nuclear DNA) results for K = 2 using a minor allele frequency (MAF) threshold of 5%, and using 20,957 loci (excluding 383 aberrant SNPs).

Table A1. Counts of loci and samples/individuals after filtering steps. Each row includes filter steps above.

Filtration steps	**SNPs**	**ddRADloci**	**N**
Initial samples	-	-	182
Exported from STACKS (reads mapping ≥ 96%)	-	-	181
Minimum 5X coverage	-	-	158
Minor Allele Frequency (MAF) > 1% or SNP detected in at least 25% of individual	136,807	88,565	158
Read depth (median read depth > 12X)	96,111	63,421	158
Individual with missing loci < 30% and loci with less than 10% missing data	86,673	58,341	157
Observed heterozygosity lower than 60%	86,343	58,229	157
Sequencing plates effect	86,342	58,228	157
Sex-linked SNPs	86,334	58,226	157
Relatedness (Φ < 0.25)	86,334	58,226	148
One SNP per locus	58,226	58,226	148
MAF > 5% and less than 5% missing data	21,350	21,350	148
MAF > 10% and less than 5% missing data	12,381	12,381	148

Table A2. Haplotype frequencies for the control region of mitochondrial DNA in July and August of beluga populations from Cumberland Sound and the Hudson Bay-Strait Complex (used to generate Figure 3).

Haplotype	CSB		BEL-EHB		JAM		WHB	
	July	August	July	August	July	August	July	August
HL001	0.054	0.125	0.279	0.096	0.048	0.035	0.205	0.132
HL002	0.000	0.000	0.102	0.044	0.048	0.000	0.000	0.000
HL003	0.302	0.300	0.054	0.018	0.000	0.018	0.265	0.357
HL004	0.007	0.000	0.007	0.000	0.000	0.000	0.012	0.119
HL005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004
HL006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.030
HL007	0.000	0.000	0.014	0.009	0.048	0.000	0.000	0.000
HL008	0.007	0.000	0.007	0.044	0.429	0.439	0.048	0.017
HL009	0.000	0.000	0.068	0.158	0.000	0.000	0.048	0.004
HL010	0.000	0.000	0.082	0.123	0.000	0.018	0.000	0.000
HL011	0.027	0.000	0.000	0.000	0.000	0.018	0.000	0.021
HL012	0.000	0.000	0.007	0.000	0.000	0.000	0.036	0.026
HL013	0.000	0.000	0.000	0.026	0.000	0.000	0.024	0.004
HL014	0.000	0.000	0.000	0.000	0.000	0.035	0.024	0.000
HL015	0.000	0.000	0.000	0.000	0.000	0.000	0.036	0.013
HL016	0.000	0.000	0.061	0.140	0.000	0.053	0.000	0.038
HL017	0.000	0.000	0.014	0.000	0.000	0.000	0.000	0.004
HL018	0.000	0.000	0.014	0.009	0.000	0.000	0.012	0.004
HL019	0.000	0.000	0.014	0.088	0.000	0.000	0.000	0.009
HL020	0.000	0.000	0.020	0.000	0.000	0.000	0.000	0.013
HL021	0.034	0.025	0.000	0.000	0.000	0.000	0.000	0.000
HL022	0.121	0.050	0.075	0.000	0.000	0.000	0.000	0.000
HL023	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HL024	0.141	0.200	0.000	0.000	0.000	0.000	0.000	0.000

Haplotype	CSB		BEL-EHB		JAM		WHB	
	July	August	July	August	July	August	July	August
HL025	0.000	0.025	0.000	0.000	0.000	0.000	0.000	0.004
HL026	0.000	0.000	0.014	0.000	0.000	0.000	0.108	0.017
HL027	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004
HL028	0.027	0.075	0.000	0.000	0.000	0.000	0.000	0.000
HL029	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HL030	0.013	0.075	0.000	0.000	0.000	0.000	0.000	0.004
HL031	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.000
HL033	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.000
HL034	0.000	0.000	0.000	0.000	0.000	0.018	0.000	0.009
HL035	0.000	0.000	0.020	0.000	0.048	0.000	0.000	0.000
HL038	0.000	0.000	0.007	0.000	0.095	0.105	0.000	0.000
HL039	0.000	0.000	0.000	0.018	0.000	0.000	0.000	0.009
HL040	0.027	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HL041	0.000	0.000	0.007	0.026	0.000	0.000	0.012	0.004
HL042	0.013	0.000	0.000	0.000	0.000	0.000	0.012	0.017
HL043	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.017
HL044	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.026
HL045	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004
HL046	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004
HL047	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.013
HL048	0.000	0.000	0.020	0.000	0.000	0.000	0.012	0.000
HL049	0.000	0.000	0.041	0.026	0.000	0.000	0.000	0.000
HL050	0.047	0.050	0.007	0.000	0.000	0.000	0.000	0.000
HL051	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.009
HL053	0.007	0.000	0.000	0.000	0.000	0.000	0.048	0.009

Haplotype	CSB		BEL-EHB		JAM		WHB	
	July	August	July	August	July	August	July	August
HL055	0.020	0.025	0.000	0.000	0.000	0.000	0.000	0.004
HL056	0.020	0.000	0.000	0.000	0.000	0.000	0.000	0.004
HL057	0.027	0.025	0.000	0.000	0.000	0.000	0.000	0.000
HL059	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.000
HL060	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.013
HL063	0.000	0.000	0.000	0.018	0.000	0.000	0.000	0.000
HL064	0.007	0.000	0.007	0.000	0.000	0.000	0.000	0.000
HL068	0.000	0.000	0.007	0.000	0.000	0.000	0.000	0.000
HL069	0.000	0.000	0.007	0.000	0.000	0.000	0.000	0.000
HL070	0.000	0.000	0.000	0.000	0.143	0.123	0.000	0.000
HL071	0.000	0.000	0.027	0.000	0.000	0.000	0.000	0.000
HL072	0.000	0.000	0.000	0.000	0.000	0.053	0.000	0.000
HL073	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HL074	0.000	0.000	0.007	0.000	0.095	0.035	0.000	0.000
HL077	0.000	0.000	0.000	0.053	0.000	0.018	0.000	0.000
HL081	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.009
HL082	0.034	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HL083	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.000
HL084	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004
HL085	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004
HL087	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004
HL088	0.000	0.000	0.000	0.009	0.000	0.000	0.000	0.000
HL090	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.013
HL091	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.000
HL092	0.000	0.000	0.000	0.026	0.000	0.000	0.000	0.000

Haplotype	CSB		BEL-EHB		JAM		WHB	
	July	August	July	August	July	August	July	August
HL097	0.000	0.000	0.000	0.009	0.000	0.000	0.000	0.000
HL103	0.000	0.000	0.007	0.035	0.000	0.000	0.000	0.000
HL112	0.000	0.000	0.007	0.000	0.000	0.000	0.000	0.000
HL114	0.000	0.000	0.000	0.000	0.000	0.018	0.000	0.000
HL115	0.000	0.000	0.000	0.009	0.000	0.018	0.000	0.000
HL121	0.000	0.000	0.000	0.000	0.048	0.000	0.000	0.000
HL122	0.000	0.000	0.000	0.009	0.000	0.000	0.000	0.000
HL123	0.000	0.000	0.000	0.009	0.000	0.000	0.000	0.000
HL128	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HL129	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.000
HL139	0.000	0.025	0.000	0.000	0.000	0.000	0.000	0.000
HL140	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HL141	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HL142	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HL143	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table A3. Individual membership-probability point estimates to populations (Cumberland Sound: CSB, Western Hudson Bay: WHB) and associated standard errors (SE) estimated with ADMIXTURE.

ID	Harvest region	CSB	WHB	SE
ARPG_05_1220	CSB	0.199	0.801	0.039
ARPG_02_1054	CSB	0.604	0.396	0.033
ARPG_02_1072	CSB	0.570	0.430	0.040
ARPG_05_1232	CSB	0.102	0.898	0.053
ARPG_06_1281	CSB	0.225	0.775	0.041
ARPG_06_1280	CSB	0.582	0.418	0.035
ARPG_02_1040	CSB	0.609	0.391	0.027
ARPG_05_1255	CSB	1.000	0.000	0.087
ARPG_02_1039	CSB	1.000	0.000	0.000
ARPG_05_1241	CSB	0.119	0.881	0.072
ARPG_05_1225	CSB	0.084	0.916	0.052
PGDL_02_03	CSB	1.000	0.000	0.000
ARPG_07_1314	CSB	0.684	0.316	0.056
ARPG_05_1211	CSB	0.133	0.867	0.045
ARPG_02_1034	CSB	0.625	0.375	0.077
ARPG_06_1265	CSB	0.630	0.370	0.044
PGDL_02_02	CSB	1.000	0.000	0.000
ARPG_02_1037	CSB	0.531	0.469	0.027
ARPG_02_1198	CSB	0.599	0.401	0.030
ARPG_02_1077	CSB	0.577	0.423	0.048
ARPG_07_1317	CSB	0.593	0.407	0.045
ARPG_05_1230	CSB	0.206	0.794	0.074
ARPG_07_1328	CSB	0.633	0.367	0.066
ARPG_02_1181	CSB	1.000	0.000	0.000
ARPG_07_01	CSB	1.000	0.000	0.054

ID	Harvest region	CSB	WHB	SE
ARPG_07_02	CSB	0.597	0.403	0.052
PGDL_02_05	CSB	0.651	0.349	0.029
ARAR_03_1061	WHB	0.098	0.902	0.050
ARAR_03_1075	WHB	0.042	0.958	0.036
ARAR_03_1041	WHB	0.118	0.882	0.045
ARNR_04_02	WHB	0.133	0.867	0.040
ARAR_99_1009	WHB	1.000	0.000	0.079
ARAR_99_1036	WHB	0.089	0.911	0.045
ARHU97_019	WHB	0.048	0.952	0.044
ARCI_99_1023	WHB	0.100	0.900	0.040
ARNR_03_04	WHB	0.176	0.824	0.045
ARAR_03_1063	WHB	0.135	0.865	0.045
ARAR_03_1064	WHB	0.196	0.804	0.036
ARAR_03_1047	WHB	0.095	0.905	0.052
ARAR_03_1048	WHB	0.000	1.000	0.035
ARAR_03_1057	WHB	0.016	0.984	0.039
ARNR_04_04	WHB	0.237	0.763	0.036
ARCI_99_1003	WHB	0.109	0.891	0.046
ARRB_02_1179	WHB	0.011	0.989	0.044
FMMM_CH_009	WHB	0.104	0.896	0.042
B96_254_COH	WHB	0.026	0.974	0.040
B96_258_COH	WHB	0.099	0.901	0.038
ARRB_01_1101	WHB	0.135	0.865	0.053
ARRB_01_1103	WHB	0.127	0.873	0.061
ARRB_01_1146	WHB	0.000	1.000	0.031
ARCH_99_1004	WHB	0.143	0.857	0.041

ID	Harvest region	CSB	WHB	SE
ARCH_99_1007	WHB	0.059	0.941	0.042
ARCH_99_1008	WHB	0.126	0.874	0.041
ARCH_99_1012	WHB	0.160	0.840	0.041
ARCH_00_1057	WHB	0.140	0.860	0.046
ARCH_00_1064	WHB	0.185	0.815	0.038
ARCI_99_1006	WHB	0.133	0.867	0.037
ARCI_99_1013	WHB	0.207	0.793	0.038
ARCI_99_1020	WHB	0.036	0.964	0.035
ARCHL_15_128164	WHB	0.009	0.991	0.038
ARCHL_15_128157	WHB	0.161	0.839	0.045
ARCHL_15_128160	WHB	0.054	0.946	0.044