Pairwise relatedness patterns of beluga whales (*Delphinapterus leucas*) in the St. Lawrence River population

By
copyright Katelyn Cherwonick

Supervised By
Dr. Timothy Frasier

A Thesis Submitted to
Saint Mary's University, Halifax, Nova Scotia
in Partial Fulfillment of the Requirements for the
Degree of Bachelor of Science, Honours Certificate in Biology.

April, 2013, Halifax, Nova Scotia

Approved:	Dr. Timothy Frasier Supervisor
Approved:	Dr. Hugh Broders Reader
Approved:	Dr. Genlou Sun Reader
Date:	April, 2013

Table of Contents

List of Figures and Tables

Pairwise relatedness patterns of beluga whales (*Delphinapterus leucas*) in the St. Lawrence River population

Katelyn Cherwonick

Abstract

The St. Lawrence River beluga whale (*Delphinapterus leucas*) population is geographically and reproductively isolated from other beluga populations. This small population is also endangered, therefore learning the population dynamics and social structure of the St. Lawrence belugas is of great interest for not only biological reasons, but also for establishing conservation strategies. The main goal of this study was to uncover the genetic structure and final-scale relatedness patterns of the St. Lawrence River beluga population, and in particular to establish if the population follows a matrilineal social structure, similar to other cetaceans such as sperm whales and orcas. These queries are addressed by genotyping DNA of beluga whales in the St. Lawrence at 22 microsatellite loci. These genotypes were used to estimate pairwise relatedness of all individuals. Network analyses were then conducted on these data to assess clustering patterns within the population, as well as to assess the patterns of genetic connectivity among individuals. The population shows significant signs of clustering, with females being more clustered than males, and males having higher connectivity than females. These findings suggest that the St. Lawrence River beluga population does follow a matrilineal social structure since females cluster with other females based on close genetic relatedness, and males act as conduits of gene flow between these maternal clusters. Future studies will combine these network data with field observation to identify how genetic clusters reflect habitat use patterns and differential exposure to anthropogenic disturbances in the area.

April, 2013

Acknowledgements

I would like to acknowledge the help and guidance of the following:

Dr. Timothy Frasier
Dr. Brenna McLeod
Lynne Burns
Krista Arseneault
Laura Johnson

Introduction

The social structure of animal populations can be diverse and complex, particularly in long-lived vertebrates that display group living such as primates and elephants (Goldzien, 1987; Archie *et al*., 2006). Many toothed whale species (odontocetes) are highly social and live in societies with complex social structures (Mann *et al*., 2000). At a large scale the dynamics of group-living species are thought to be driven by two major factors; ecological conditions that make group living beneficial; and differential relatedness among individuals within the group, which results in different strategies for maximizing inclusive fitness (Emlen, 1982). Thus, understanding the social structure of complex animal societies requires long-term ecological data for understanding the ecological conditions driving group living, as well as high-resolution genetic data for understanding relatedness patterns and kin selection. Detailed field data on the St. Lawrence beluga population have been collected for over 20 years. Described below are the recent advancements made in the genetic aspect of this work, and the subsequent insights these have provided to our understanding of the social structure of the St. Lawrence beluga population.

Group Living

Group-living is characteristic for many species, despite the inevitable costs associated with it. Costs associated with living in a group are that it can be energetically demanding, lead to an increased stress response (mainly experienced by a subordinate in relation to its social status of a dominate within the same group), result in resource competition, reproductive competition, as well as aggressive interactions between members of the group (Creel, 2001; Craig *et al*., 2002; Majolo *et al*., 2008; Schradin *et al*., 2010). However, conditions imposed by the ecological

environment can make group living beneficial, such as a lack of food resources or the constraint of having poor options for independent breeding (Emlen, 1982a).

Group-living can only evolve when its innate benefits are greater than the costs (Emlen, 1982b). Several hypotheses on the benefits of group living include the suggestion hunting in a group increases the possibility of gaining larger prey and the idea that group-living decreases the risk of predation (Mech, 1981; Inman and Krebs, 1987). The benefits of group-living and the mechanisms of ecological constraints interact with each other to form tolerable living conditions (Keonig *et al*., 1992). For example, striped mice display group-living during their breeding season, even though this is an unstable social group dynamic, because during this time period when the population is at a high density, dispersal is unfavorable due to the lack of quality territory available (Schradin *et al*., 2010).

Inclusive Fitness and the Evolution of the Family

Inclusive fitness refers to how successfully an organism passes copies of its genes to future generations (Creel, 1990; Queller, 1992). More specifically, it is a combination of the fitness an individual attains through its own reproductive success (direct fitness) combined with the fitness of that individual through the reproductive success of its genetic relatives (indirect fitness). The rationale is that relatives share genes, and therefore an individual can indirectly pass on alleles through the reproductive success of relatives (Hamilton, 1964). Thus, individuals should show differential willingness to help, or sustain fitness costs from, other individuals in proportion to how related they are. This process can result in complex social structures in species, such as families, where ecological conditions favour group living because variation in

relatedness of individuals within the group will lead to variation in strategies of, and conflicts among, individuals as they each try to maximize their own inclusive fitness.

The evolution of the family builds off factors that drive group-living, but instead of random congregations of individuals within a population, families are composed of genetic relatives. Families seem to form when early-in-life breeding opportunities are limited and low fitness would be experienced by early dispersal, thus forcing offspring to remain with their parents after maturity (Emlen, 1994; 1995). Hamilton's theory of kin selection is directly related to this concept in that the greater the frequency of shared alleles, the more willing individuals should be to help each other (i.e. to incur costs at the expense of a relative's benefits) (Hamilton, 1964).

As with group-living, there are potential costs and benefits associated with delaying dispersal from parents, instead of dispersing upon maturity. Thus, dispersal of offspring is one aspect of behaviour that has been studied in relation to inclusive fitness (Emlen, 1994; Emlen, 1995; Pamillo and Crozier, 1996; Clutton-Brock, 2009). Families are rare because as an individual, not reproducing offspring upon maturity (not dispersing) can have negative consequences on the fitness of that individual (Emlen, 1995). However, if the inclusive fitness of individuals that delay dispersal (and therefore their own reproductive success) is greater than that of offspring who disperse upon maturity, family living is expected to develop (Emlen, 1994).

While breeding opportunities are fewer within the family group (avoiding inbreeding depression), there are many advantages derived from a family structure: helping rear closely related individuals (cooperative breeding), forming coalitions with relatives to better acquire breeding vacancies, and the possibility of inheriting parental territory (Emlen, 1994). The latter is

especially true for dynasties who rule territory with high quality resources in which a consistent genetic lineage inhabits the same area over many successive generations (Emlen, 1995). Such behaviours, when individuals are closely related, can create variation in breeding success and survival which influences the success of a social group (West *et al*., 2007; Clutton-Brock, 2009). For example, in African elephants there exist stable sub-units of close maternal relatives which combine to form larger groups of maternal relatives (Archie *et al*., 2006). Periodically these groups fuse with other units that consist of more distant maternal relatives (Durand *et al*., 2007). It was discovered through studies on genetic relatedness that these group changes appear to be directed by the female matriarch of each group because social groups are representative of individuals who share the same haplotype (Archie *et al*., 2006).

The Beluga Whale

Beluga whales (*Delphinapterus leucas*) are toothed whales that are relatively small compared to other whale species; generally ranging in length and weight from 3 to 5.5 m and 1,350 to 1,500 kg, respectively (Sergeant and Brodie, 1969; Béland, 1996). Belugas are characterized by their lack of a dorsal fin and white skin, although neonates are grey to brown in colour. Whitening of the skin occurs after six years as belugas reach adulthood and the colour change takes place over a number of years until they are completely white around age 13 (Brodie, 1971).

Belugas have a discontinuous circumpolar distribution in Arctic and sub-Arctic waters. The range of each population changes seasonally in response to the movement of the ice edge, with migrations between higher-latitude summering grounds and lower-latitude wintering grounds (Brodie, 1971; Stewart and Stewart, 1989). Summering grounds exist around the mouths

of rivers and in estuaries, which are areas of high productivity and therefore represent important feeding areas. As temperatures drop and the water begins to freeze, the belugas migrate to deeper off-shore locations that are referred to as wintering grounds. Five main wintering areas have been identified in North America, which may be composed of multiple summering groups (Donovan, 1992).

Based on observations of peak calving times in mid-summer, and an estimated gestation time of 14.5 months, conception likely occurs while beluga whales are in their wintering grounds or in early spring during their migration to their summering grounds (Brodie, 1971). However, poor weather conditions during the Arctic winter, and the beluga's preference to remain under offshore ice cover, make studying belugas in the winter difficult to impossible. Thus, data on most beluga populations is limited to information obtained while in their summering grounds, resulting in a large gap in our understanding of what happens during the winter, which includes mating (Brodie, 1971; Sergeant, 1973). Although mating cannot be directly observed, patterns of gene flow and reproductive success can be inferred through genetic analyses. Thus, for this species genetic analyses are particularly useful and can shed light on important aspects of their biology, such as reproduction and gene flow, which are otherwise unavailable.

Genetic analyses involving mitochondrial DNA (mtDNA) have been helpful in understanding the population structure of many whale species by inferring maternal lineages from the maternal inheritance of mtDNA (Baker *et al*., 1993; Palsbøll *et al*., 1995; Witteveen *et al*., 2004; Hoelzel, 2009). For example, in humpback whales mtDNA sequencing showed that all whales within an ocean converge on one breeding ground in the Caribbean during the winter, but during the summer months whales of different matrilines use different feeding areas (Witteveen *et al*., 2004; Hoelzel, 2009). This maternally-based site fidelity results in strong differentiation of

mitochondrial haplotypes among summer feeding areas, despite the complex mixing of these whales on one common breeding ground (Baker *et al*., 1994).

Studies involving mtDNA have also been conducted on beluga whales (Brennin *et al*., 1997; Brown Gladden, 1997; O'Corry-Crowe *et al.*, 1997). These analyses have shown that beluga whale haplotypes differ between eastern populations and western populations, indicating that matrilines return to the same summer habitats after spending time in wintering locations with multiple beluga populations. The differentiation of beluga haplotypes in the various summering grounds is evidence of site fidelity to these locations year after year (Brennin *et al*., 1997; Caron and Smith, 1990).

Of all the beluga whale populations around North America, the population inhabiting the St. Lawrence River is the only one that resides south of the Arctic Circle and has an estimated population size of 1,000 individuals (Béland, 1996; Stewart and Stewart, 1989). The summering ground for this population is the Saguenay River and the estuary of the St. Lawrence River. Once the estuary freezes over in the winter belugas are forced out to the icy waters of the Gulf of St. Lawrence, which appears to be the location of their wintering grounds. These data on habitat use, combined with genetic data, have shown that the haplotypes common to belugas in the St. Lawrence River beluga population is reproductively isolated from all other beluga populations (Brennin *et al*., 1997; Brown Gladden *et al.*, 1999; O'Corry-Crowe *et al*., 1997).

Genetic Methods of Analysis

Earlier studies on beluga populations and other social animals focused on the use of mtDNA since it is more abundant and is more resistant to degradation due to its robust circular structure than nuclear DNA (nDNA) (Brown *et al*., 1979; Butler and Levin, 1993). The pattern

of inheritance for mtDNA primarily follows the matriline since mitochondria originate from the egg of the mother. This haploid nature of mtDNA allows for easier sequencing; however it also means that mtDNA data must be interpreted carefully because it does not represent gene flow of the whole population (Butler and Levin, 1993; Godinho *et al*., 2008). One advantage nDNA has over mtDNA is biparental inheritance, which means that nDNA represents patterns of gene flow to and from both sexes (Butler and Levin, 1998). Hence, many studies that had limited resolution based on analysis of mtDNA were improved by incorporating nDNA microsatellites to further assess population structure and dynamics (Moore, 1995; Godinho *et al*., 2008).

A major advancement in the use of nDNA in population studies was the discovery of microsatellites or STRs (short tandem repeats) which have a relatively uniform distribution throughout an organism's genome (Queller *et al*., 1993). Microsatellites have repeat units of 1-6 base pairs in length and entire microsatellite regions are typically small enough to be amplified using PCR (polymerase chain reaction). They are also abundant in the genome and are highly polymorphic; making them ideal as genetic markers for population studies (Queller *et al*., 1993; Bennett, 2000). The ability to add fluorophores of different colours to primers, and therefore label the resulting PCR products, allows for several loci to be amplified and analyzed in one reaction called a multiplex reaction (Church and Kieffer-Higgins, 1988; Bennett, 2000). These multiplex reactions increase the efficiency of the genotyping process and allow for studies to be carried out that look at the genetic relationships among groups, within groups, and between individuals (Mann *et al*., 2000).

The analysis of nDNA microsatellites is now widespread (Maiers *et al*., 1996; Brown Gladden *et al*., 1999; Burg *et al*., 1999; Mann *et al*., 2000; Witteveen *et al*., 2004; Frère *et al*., 2010; Martien *et al*., 2012). The proportion of shared alleles at certain loci can be used to

estimate genetic relatedness patterns between individuals, which allows for inferences to be made about aspects of population structure (Mann *et al*., 2000). More specifically, genetic relatedness among social groups can be estimated by comparing allele frequencies from microsatellites between individuals to infer information about the structure, mating systems and stability of animal populations (Blouin *et al*., 1996).

Microsatellites were used to compare nuclear data with previously studied mtDNA data on the distribution of select North American beluga whales (Brennin *et al*., 1997; O'Corry-Crowe *et al.*, 1997; Brown Gladden *et al.*, 1999). Brown Gladden *et al*. (1999) found the nDNA results matched previous mtDNA studies in that there was a non-uniform distribution of microsatellite alleles between populations. From this study it was concluded that belugas from different summering sites mate with individuals from other summering sites that winter in the same location. However, belugas from different wintering grounds do not mate, as there was still differentiation among whales from different wintering locations. Conversely, if mating of different wintering groups did occur, then a relatively uniform distribution between microsatellite alleles would have been observed. These results were reflected in a study by Turgeon *et al*. (2012), which found that belugas in eastern and western Hudson Bay showed differentiation of mitochondrial haplotypes, and represent distinct summering groups that follow maternally-inherited migration patterns. However, microsatellite analyses showed a lack of differentiation at the nuclear markers suggesting interbreeding between belugas from Hudson Bay does occur on their wintering grounds (Turgeon *et al*., 2012).

Some studies have used results from relatedness data to conduct network analyses as a means to further ascertain the complex social relationships of a particular population (Lusseau, 2003; Martien *et al*., 2012). This is a relatively new application of network analysis, which is a

long standing mathematical technique (Newman, 2006; Wey *et al*., 2008). In animal population network analysis an individual is considered a "node", and the connectivity of individuals and groups is determined by the number and orientation of edges attached to each node. Through network analysis, clustering patterns of individuals can be interpreted as a group that has a high level of connectivity; furthermore, it can be possible to recognize specific nodes as "key players" which are individuals that have a high level of connectivity (Newman, 2006; Wey *et al*., 2008).

Study Objective

The goal of this study was to investigate the social structure and finer-scale relatedness patterns of the St. Lawrence River beluga population by genotyping microsatellites to conduct network analysis. However, instead of looking at structuring differences between summering and wintering sites, as was conducted in previous beluga whale studies on relatedness, I researched the clustering patterns within one summering site (Brown Gladden *et al.*, 1999).

Based on patterns observed in studies of large social animals such as elephants and other whales (Mann *et al*., 2000; Archie *et al*., 2006), I hypothesized that the St. Lawrence River beluga population follows a matrilineal social structure. Derived from field observations of beluga whales in the St. Lawrence, it was predicted that the population would display intrapopulation genetic clustering, with females being more clustered than males, representing groups of closely related females. Consistent with matrilineal-based groups, it was also predicted and that males would be more connected than females, since males should serve as a conduit of gene flow between these clusters for reproduction with females from different matrilines.

Materials and Methods

Sample Collection and Preparation

Beluga whale skin samples were obtained in collaboration with the Group for Research and Education on Marine Mammals (GREMM) using an air propelled dart that takes a small biopsy from free-swimming whales (Barrett-Lennard *et al*., 1996). This method has been used for collecting samples from many cetacean populations and research indicates that it does not have any short- or long-term impacts on the whales, other than an initial startle response (Noren and Mocklin, 2011). Tissue samples were also collected from dead individuals during necropsies. All samples were stored in a 20% dimethyl sulfoxide solution (DMSO) which contains; 20% DMSO, 0.25 M Ethylenediamineteraacetic acid (EDTA) (pH 8.0) and is saturated with sodium chloride (NaCl) (Seutin *et al*., 1991).

In preparation for DNA extraction, 35-45 mg of whale tissue from each sample were weighed, then diced on a KimWipe using a scalpel blade. The sample and 100 µl of lysis buffer were then added to a mortar that had been cooled with liquid nitrogen (LN). The purpose of lysis buffer is to break open the cell membrane encasing the DNA and contains; 10 mM EDTA (pH 8.0), 2% sodium dodecyl sulfate (SDS), 40 mM Dithiothreitol (DTT), 10 mM Tris base (pH 8.0) and 0.1 M NaCl (Budowle *et al*., 2000). The tissue was ground into a powder using a pestle that had also been cooled with LN. The ground skin material was scraped from the mortar with a LN cooled scoopula into a 1.5 ml tube and another 300 µl of lysis buffer was added. For each batch of extractions, a positive control of beef muscle tissue was prepared in the same manner as the beluga whale samples and a negative control – containing only reagents and no DNA – was also prepared to identify if any of the extraction reagents were contaminated. Samples were incubated at room temperature for 5-6 days, mixing two to three times per day by upending.

On the last day of incubation in the lysis buffer, 0.5 units per milligram of tissue (which for a 40 mg sample is 20 units) of the enzyme proteinase K (pro-K) was added to each sample tube, and mixed by flicking. After sitting overnight at room temperature samples received another 20 units of pro-K, were mixed, placed in 65°C water bath for 1 hour, and then placed in the incubator at 37°C for 1 hour. All samples had another 20 units of pro-K added, were mixed, and left overnight at room temperature.

DNA Extraction

Phenol:chloroform methods were used to extract DNA from the beluga tissue samples. In a fume hood an equal volume (400 µl) of phenol:chloroform was added to each sample and the samples were upended for 5 minutes. After being spun in the centrifuge for 4 minutes at 12,000 x g the aqueous layer was removed and aliquoted into to a new 1.5 ml tube. These steps were repeated by adding another equal volume (400 µl) of phenol:chloroform to each sample. Then the samples were upended for 5 minutes and spun in the centrifuge for 4 minutes at 12,000 x g. After this second aqueous layer was transferred to a new 1.5 ml tube an equal volume (400 µl) of chloroform was added to the tube, upended for 5 minutes, and spun in the centrifuge for 4 minutes at 12,000 x g. The aqueous layer was removed and aliquoted into a new 1.5 ml tube then 80 µl of 10 M ammonium acetate and 800 µl of 95% ice-cold ethanol were added to each sample, and were mixed well. Samples were left if the freezer overnight to facilitate DNA precipitation.

After removing samples from the freezer, they were spun at 12,000 x g for 10 minutes, and the ethanol was decanted. Then 100 µl of 70% ethanol was added to each sample and the tubes were spun again at 12,000 x g for 10 minutes. After the 70% ethanol was decanted and a KimWipe was used to remove excess ethanol, tubes were left open for 10-20 minutes to allow

any remaining ethanol to evaporate. Lastly a volume, enough to immerse the DNA pellet, of $TE_{0.1}$ pre-warmed to 65°C was added to each tube and mixed well to re-dissolve the pellet. $TE_{0.1}$ is a storage solution for DNA that contains 10 mM Tris (pH 8.0) and 0.1 mM EDTA (pH 8.0).

Assessing DNA Quantity and Quality

DNA concentrations were initially estimated based on spectrophotometery using a NanoDrop 2000 (Thermo Scientific). The NanoDrop was blanked using $TE_{0.1}$ and four calf thymus standards at know concentrations of 50 ng/µl, 10 ng/µl, 5 ng/µl, and 1 ng/µl were used for calibration. Based on the amount of $TE_{0.1}$ that each sample was diluted in, the concentration readings from the spectrophotometer were used to estimate the yield of DNA in each sample as well as create 5 ng/µl dilutions to be used for further assessments of DNA quality and quantity in an agarose gel.

Agarose gel electrophoresis was used as a second method of estimating the quantity of DNA in each sample, as well as to assess the amount of DNA degradation in each. Twenty nanograms of DNA were loaded into 2% agarose gels stained with the intercalating dye ethidium bromide. With every gel a Low Mass DNA Ladder (Invitrogen) was added to the first well, as a standard to which DNA samples could be compared to assess DNA quality and quantity (**Figure 1**). Based on the intensity (fluorescence) of each sample, the DNA concentration and yields are re-estimated and a functional concentration was calculated which was then used to make new 5 ng/µl dilutions to be used for further DNA analyses.

Sexing Beluga Samples

Two regions of the nDNA were targeted and amplified to assign a sex to each individual, using the primer pairs described in Gilson *et al*. (1998). One pair (P1-5EZ and P2-3EZ) amplified a 445 base pair (bp) portion of a zinc finger transcription factor gene present on both sex chromosomes (**Table 1**). The second region, a 224 bp region of the *SRY* gene located on the Y-chromosome of males, was amplified using the forward and reverse primers Y53-3C and Y53-3D respectively (Table 1). PCR amplifications were conducted in 20 µl reaction volumes containing; 10 ng of DNA, 1X PCR buffer (20 mM Tris-HCl (pH 8.4), 50 mM KCl) (Invitrogen), 1.5 mM $MgCl_2$ (Invitrogen), 0.2 mM of each dNTP, 0.4 mg/ml BSA, 0.3 µl *Taq* polymerase (Invitrogen) and 0.3 µM of each primer. Cycling conditions consisted of an initial denaturing step for 5 minutes at 95°C to better separate the long fragments of double stranded DNA, 30 cycles of 95°C for 30 seconds, 55°C for one minute, 72°C for one minute, and a final extension step of 64°C for 45. Every round of samples run through these PCR conditions also contained a positive control in which 10 ng calf thymus DNA were added and a negative control in which 2 µl of reagent water were added in place of DNA.

PCR products were run on a 2% agarose gel stained with 0.5 µl/mL ethidium bromide and visualized under UV light. For each sample 4 µl of PCR product were combined with 2 µl of Orange G loading dye. A sample that has two bands indicates the sample is a male, while a single band is indicative of a female (**Figure 2**).

Microsatellite Genotyping

Each sample was amplified at 22 microsatellite loci using protocols that had previously been developed in our laboratory (**Table 2**). The forward primers are labelled with one of four

fluorophores; 6-FAM, VIC, NED or PET (Applied Biosystems) which visualizes the alleles at loci as either blue, green, yellow or red, respectively. PCR amplifications of microsatellite loci were conducted in 10 µl reaction volumes containing; 10ng of DNA, 1X PCR buffer (Invitrogen), 1.5 mM $MgCl_2$ (Invitrogen), 0.2 mM of each dNTP, 0.16 mg/ml BSA, 0.05 u/µl *Taq* polymerase (Invitrogen) and concentration specific to each primer (Table 2). Cycling conditions were as follows: 5 minutes at 95°C, 30 cycles of 95°C for 30 seconds, annealing temperature (Ta, Table 2), 72°C for one minute, and a final extension step of 64°C for 45 minutes. Every round of samples run through these PCR conditions also contained a negative control in which 2 µl of reagent water were added in place of DNA.

In preparation for genotyping, the PCR product of the target regions were first diluted by combining 2 µl of PCR product with 18 µl of water, then 2 µl of each dilution was suspended in 10 µl HiDi formamide that also contained a size standard called GS600 which is labelled with LIZ, the orange fluorescent tag (Applied Biosystems). PCR products were size-separated and visualized via capillary electrophoresis on an ABI 3500xl genetic analyser.

Microsatellite Scoring, Relatedness and Network Analysis

Microsatellite alleles at each locus were scored using the GeneMarker 2.2 computer software (SoftGenetics) (**Figure 3**). Allele calls were edited manually at the 22 loci to ensure correct scoring. Any scores that were difficult to call were re-assessed by the principle investigator (Tim Frasier) to either come to a consensus on the appropriate call, or to conclude not to score a specific locus for an individual. Unsuccessful loci and samples were removed from the dataset for further analysis since samples with a lack of loci information have the potential to sway the data and provide results that are not representative of the population.

Network analysis was initiated by first creating pairwise matrices of the relatedness of genotypes, by using the relatedness estimator described in Li *et al*. (1993), and the software program GeNetwork (Frasier, 2011). Once relatedness between genotypes was estimated, the statistical program R was used to perform individual-based network analyses on the genetic data using the *igraph* package (Csardi and Nepusz, 2006; R Development Core Team, 2012). There are many methods for estimating the existence of group structures within a network; however the fast greedy community cluster analysis was the most appropriate fit for the data.

Network analyses were then conducted independently for males and females to test for patterns of clustering within each of the sexes. There are several means to quantify how "connected" individuals are within a network, essentially how many connections go through each individual. One metric of this is called "eigenvector centrality", which measures the strength of connectivity of each individual relative to other individuals within the same network (Croft *et al*., 2008). Individuals with high eigenvector centrality are those which are connected to many other individuals, and conversely, individuals with low eigenvector centrality are those which have few connections to other individuals within the same network. The eigenvector centrality values between males and females in the St. Lawrence beluga network were plotted and compared.

The metric of "modularity" was calculated for the total beluga network, based on the fastgreedy community cluster analysis. Modularity measures how well divided clusters of a network are relative to the total groups within a network, and reflects the concentration of individuals within identified clusters compared to the random distribution of all individuals in the network (Croft *et al*., 2008). A low modularity value would indicate connections between individuals are relatively equal, whereas a high modularity value would represent a network that

has many close connections between individuals within clusters, but sparse connections between individuals belonging to different clusters. GeNetwork was then used to calculate the significance of clustering patterns in the beluga network by comparing the modularity of the observed network to the expected modularity values of 100 randomly generated, non-clustered populations, consisting of the same number of individuals. Modularity values were also established for each sex separately, and the modularity of these observed male and female networks were tested for significance following the same methods of analysis carried out for the total beluga network.

The resulting networks were then combined with the long term field photo-identification data obtained in collaboration with GREMM, such as whether the sample was taken from a live individual or during a necropsy, as well as the gender of the samples, to identify association patterns among individual beluga whales. Networks were visualized using the software program Cytoscape (Smoot *et al*., 2011).

Results

Most of the samples showed a bright band with high molecular weight when assessed for DNA quality and quantity (**Figure 1**). Samples that did not appear in the initial assessment of DNA on the 2% agarose gel still contained enough DNA for analysis, since the amplification of the target loci in the multiplex reactions were successful. Of the 222 samples used for analysis, 214 amplified successfully using the primers for sexing, 126 of which were males and 88 of which were identified as females (**Figure 2**).

Previous members of our laboratory extracted and amplified 177 beluga samples, following the same protocol as was followed for the 68 samples I extracted, for a total of 245 samples that required microsatellite analysis. Out of the 22 loci amplified in the multiplex reactions, 16 were ultimately used for the subsequent analyses. Six loci (FCB4, TexVet5, TexVet19, Mk6, GATA018 and Ev1Pm) were excluded from the relatedness and network analysis due to a lack of samples with successful allele scores. Additionally, any samples missing data from five or more loci were also not included in subsequent analyses. This resulted in 222 samples amplified at a minimum of 11 loci being used to represent the St. Lawrence River beluga whale population in the network analysis and calculation of population metrics.

Network analyses based on the pairwise relatedness showed evidence of clustering, with clustering analysis detecting four clusters in the beluga population (**Figure 4**). The modularity of the total beluga network was estimated at 0.164 (**Figure 5**). Although distinct clusters are not immediately detectable with the eye, the degree of clustering within these data is statistically significant, as indicated by the modularity of the observed network being significantly higher than those from the simulated data sets ($P < 0.01$).

Network analyses were then conducted independently for males and females to test for patterns of clustering within each of the sexes (**Figure 6, Figure 7**). The modularity of the male network was 0.156, which was not significantly more clustered than the simulated networks, with 10 of the simulated networks having as high, or higher, modularity than the observed data ($P < 0.1$) (**Figure 8**). The modularity of the female network was 0.206, which was significantly higher than expected based on the simulated networks ($P < 0.01$) (**Figure 9**). Three main clusters were found in the female network, which are colour-coded for better visualization in **Figure 10**. The eigenvector centrality values for males have higher values than females, and this difference is statistically significant (Wilcoxon rank-sum test, $P = 0.00036$) (**Figure 11**).

Discussion

Network analysis based on the pairwise relatedness shows evidence of clustering. The high degree of clustering within these data indicates that there is significant sub-structuring within the population and that the clusters represent groups of closely related individuals. From the separated male and female information the results show that only females have significant clustering patterns. This indicates that the clustering pattern of the entire St. Lawrence Beluga population is driven by patterns of relatedness in females, which suggests a matrilineal-based social structure.

The difference in eigenvector centrality values for males and females clearly show that males have significantly higher values than females, which indicates that males are "more connected" than females. Males possess the main flow of connections between the different clusters, with few connections directly linking females from different clusters. The higher connectivity that males have in comparison to females is representative of their role in facilitating gene flow between the identified clusters for reproduction with females from different matrilines.

The resulting data from this study shows there is intrapopulation clustering in the St. Lawrence River beluga network, with the cluster patterns reflecting patterns of relatedness within the population. The data show that females are more clustered than males. This indicates that females are maternally related, since there is a higher proportion of shared alleles at the scored microsatellite loci between females of closer genetic relatedness than between unrelated individuals. If belugas in the St. Lawrence follow a matrilineal social structure, with individuals primarily associating with matrilineal-based groups, it is expected that each female would have a cluster of maternal relatives within the population, with male movements facilitating gene flow

between these clusters for reproduction with females from different matrilines. This is the pattern that was found in the genetic networks.

There are several reasons that could explain why beluga whales group together based on genetic relatedness. Ecological constraints of being a small population and having low reproduction rates may have made group living more beneficial by clustering based on pairwise relatedness as a means to better each individual's inclusive fitness (Clutton-Brock, 2009; Emlen, 1982a, 1994; Keonig *et al*., 1992). Furthermore, site fidelity and learned behaviour of the neonates from the mother could be at play (Brodie, 1971; Brennin et al., 1997). Since the population consistently returns to the same summering location of the St. Lawrence River, clusters may also be returning to specific areas within the estuary as well. Beluga whales also lactate over a two year period; however the neonate may spend three or more years with its mother (Brodie, 1969). Therefore, the relatively long lactation time beluga calves are subjected to would allow the neonates to learn their mother's migratory route, leading to site fidelity, and eventually genetic differentiation of clusters within the population if females show preferential use of areas within the estuary (Brodie, 1971).

This research is part of an on-going project that will continue to add more samples annually to the existing dataset. The subsequent addition of future samples will create a more comprehensive understanding of the St. Lawrence River beluga population dynamics and social structure by increasing the accuracy of information on relatedness of the individuals within the population. Comparing known association patterns identified through photo-identification with these network data will give a good perception to understanding the dynamics of the St. Lawrence River beluga population.

Further analysis could be done which explore the female social structure more closely to investigate if there are key players within each matrilineal cluster. Key players in a network are individuals that have a higher connectivity compared to others in the population (Newman, 2006; Wey *et al*., 2008). This type of analysis could help reveal if the matrilineal social groups contain dominant females, which would be representative of a matriarchal society, such is the case for wild African elephants (Archie *et al*., 2006).

A continuation of this research will combine the network data from this study with field observations as a means to identify how the genetic clusters reflect habitat-use patterns and differential exposure to anthropogenic disturbances in the St. Lawrence River, such as toxins, vessel traffic, and habitat degradation (Bailey and Zinger, 1995; Environment Canada, 1992; 1997; Richman and Dreier, 2001; Martineau *et al*., 2002). This is an important application since this population is endangered; it is listed as near threatened on the international union for the conservation of nature (IUCN) red list, as threatened by the committee of the status of endangered wildlife in Canada (COSEWIC) and are listed on schedule 1 of the species at risk act (SARA). Once data on differential exposure to contaminant loads are available, having access to the genetic network data from this research will provide a means to correlate the observed differential use of identified clusters of belugas to the exposure rates of pollution and other anthropogenic disturbances, so that conservation initiatives targeting the most highly affected areas can be created.

Another future goal is to genetically link dead individuals, who cannot be identified based on natural markings, back to specific clusters of previously sampled individuals, since not all deceased individuals were also sampled when they were alive. This will provide a way to link life history data with information obtained during the necropsy, such as exposure to toxic

chemicals, which may correspond to preferential use of the St. Lawrence estuary of the cluster from which the individual belonged. For creating conservation strategies it will be informative to link dead individuals to previously established clusters within the St. Lawrence River beluga population network if these clusters of individuals show differential habitat use patterns and therefore have differential exposure rates to anthropogenic disturbances. Separating the sample types into categories of live individuals and dead individuals is a step towards establishing the foundation for this avenue of research (**Figure 12**).

Table 1. Primers used to sex the beluga whale DNA samples by amplifying the sex determining regions in the genome of mammals.

Primer	Reaction Concentration (μM)	Ta (°C)	Reference
P1-5EZ	0.3	55	Aasen and Medrano, 1990
P2-3EZ	0.3	55	Aasen and Medrano, 1990
Y53-3C	0.3	55	Gilson *et al*., 1998
Y53-3D	0.3	55	Gilson *et al*., 1998

Table 2. Primers used to amplify microsatellites in the genome of beluga whales with the use of fluorescent labels and ten multiplex reactions. There are 22 primer pairs in total and all amplify dinucleotide microsatellite repeats, except for GATA417 and GATA028 which amplify tetranucleotide microsatellite repeats.

Primer	Fluorescent Label	Multiplex Reaction	Reaction Concentration (μM)	Ta (°C)	Reference
IGF1	VIC	#1	0.2	58	Barendse *et al*., 1994
Ev14Pm	6-FAM	#1	0.3	58	Valsecchi and Amos, 1996
GATA417	PET	#1	0.3	58	Palsbøll *et al*., 1997
RW31	VIC	#2	0.18	52	Waldick *et al*., 1999
SW19	PET	#2	0.4	52	Richard *et al*., 1996
FCB14	VIC	#3	0.35	55	Buchanan *et al*., 1996
RW34	6-FAM	#3	0.35	55	Waldick *et al*., 1999
RW48	VIC	#4	0.35	52	Waldick *et al*., 1999
FCB3	VIC	#5	0.3	58	Buchanan *et al*., 1996
Ev37Mn	VIC	#6	0.3	54	Valsecchi and Amos, 1996
FCB17	6-FAM	#6	0.85	54	Buchanan *et al*., 1996
FCB5	NED	#6	0.45	54	Buchanan *et al*., 1996
FCB10	6-FAM	#6	0.6	54	Buchanan *et al*., 1996
Ev94Mn	6-FAM	#7	0.8	56	Valsecchi and Amos, 1996
FCB6	NED	#7	1.2	56	Buchanan *et al*., 1996
FCB1	VIC	#7	0.45	56	Buchanan *et al*., 1996
FCB4	PET	#8	0.2	58	Buchanan *et al*., 1996
TexVet5	NED	#9	0.3	52	Rooney *et al*., 1999
GATA028	NED	#9	0.4	52	Palsbøll *et al*., 1997
MK6	VIC	#9	0.3	52	Krützen *et al*., 2001
TexVet19	6-FAM	#9	0.3	52	Rooney *et al*., 1999
Ev1Pm	NED	#10	0.3	52	Valsecchi and Amos, 1996

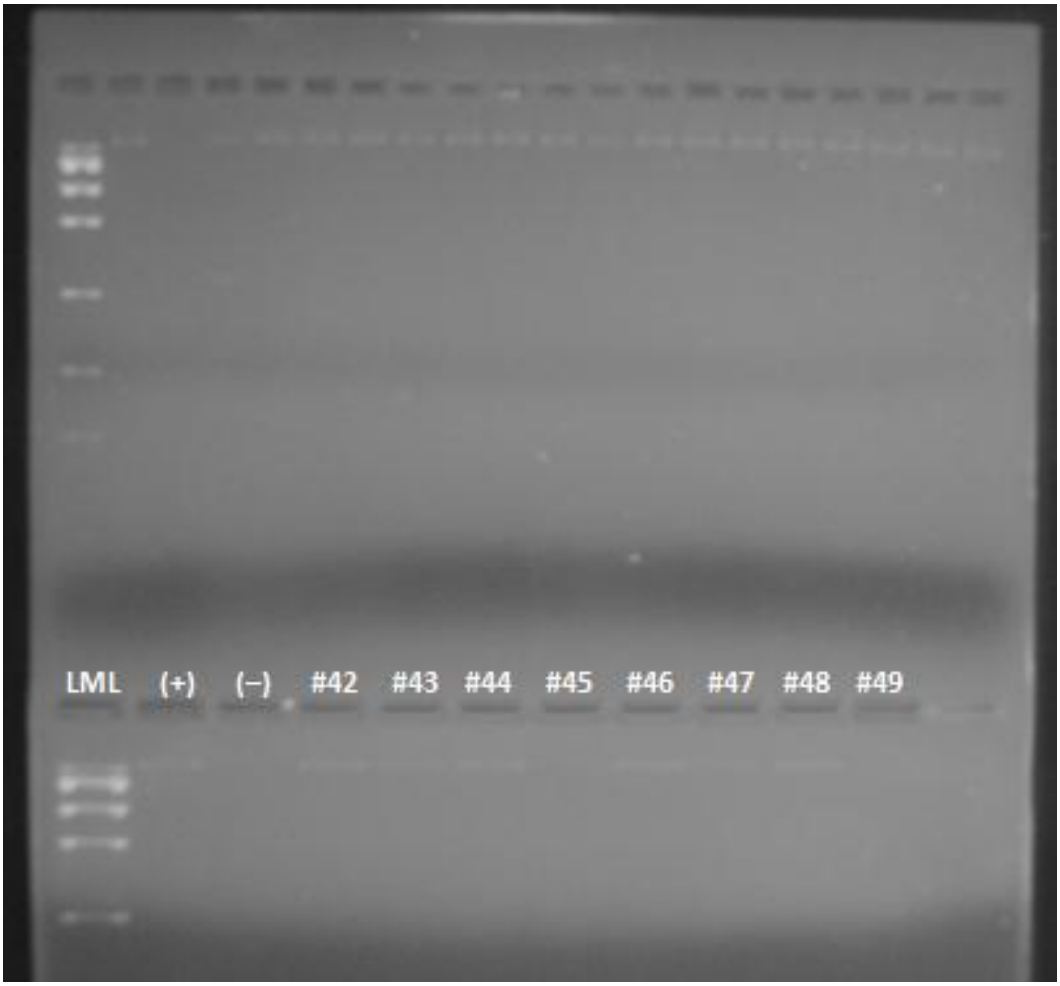

Figure 1. Example of a 2% agarose gel stained with ethidium bromide loaded with 20ng DNA dyed with Orange-G for assessment of DNA quality and quantity. (+) indicates a positive male calf thymus standard. Samples #45 and #49 show DNA of low quantity while all other labelled samples show DNA with high molecular weight bands.

Figure 2. Example of a 2% agarose gel stained with ethidium bromide for visualization of Sexing PCR product for determination of sex. A single band indicates a female while two bands indicate a male. (+) indicates a positive male calf thymus standard. Samples #3 and #4 indicate scorable females, while samples #6 and #8 indicate scorable males.

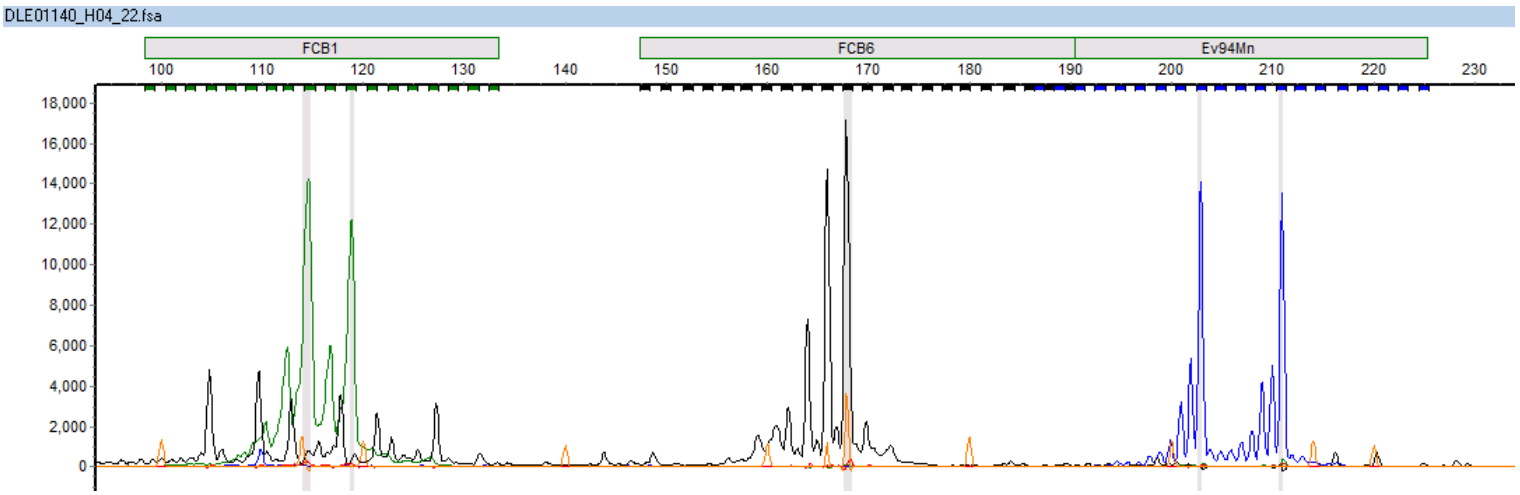

Figure 3. An example of an electropherogram of the amplified loci FCB1, FCB6 and Ev34Mn in Multiplex #1. The peaks for FCB1 and Ev34Mn show this individual is heterozygous at those loci. The peaks for FCB6 show this individual is homozygous at that locus.

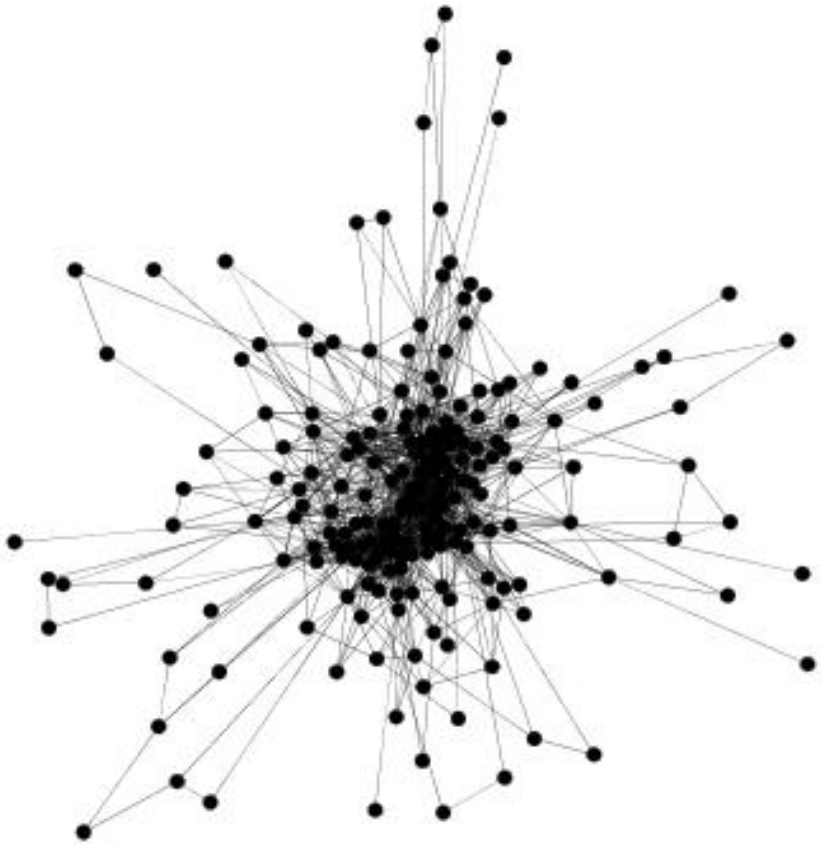

Figure 4. Network of the St. Lawrence beluga population (222 samples) based on pairwise relatedness values. Pairwise sharing values less than 0.4 were filtered out for better visualization. Each node represents one individual; the edges represent the number of connections between the individuals. The edge length between nodes is proportional to the genetic distance. The clustering in this network is significant ($P < 0.01$).

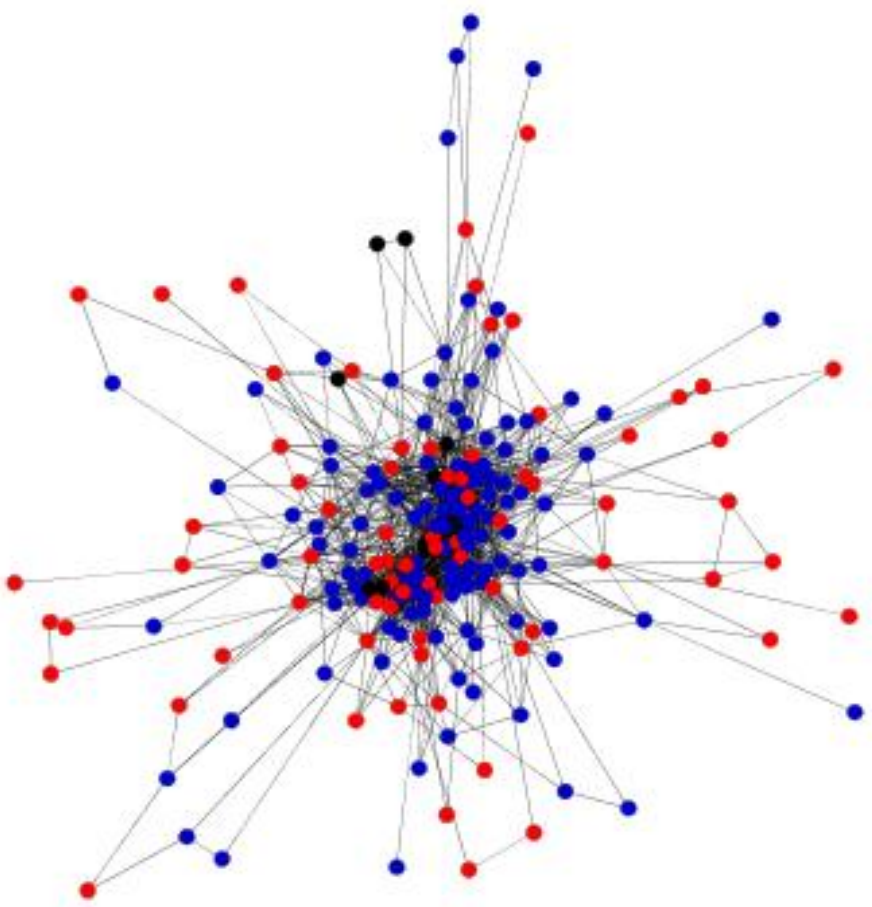

Figure 5. Network of the St. Lawrence beluga population (222 samples) based on pairwise relatedness values. Pairwise sharing values less than 0.4 were filtered out for better visualization. Males are shown in blue (126 samples), females in red (88 samples), and individuals of unknown gender in black (8 samples).

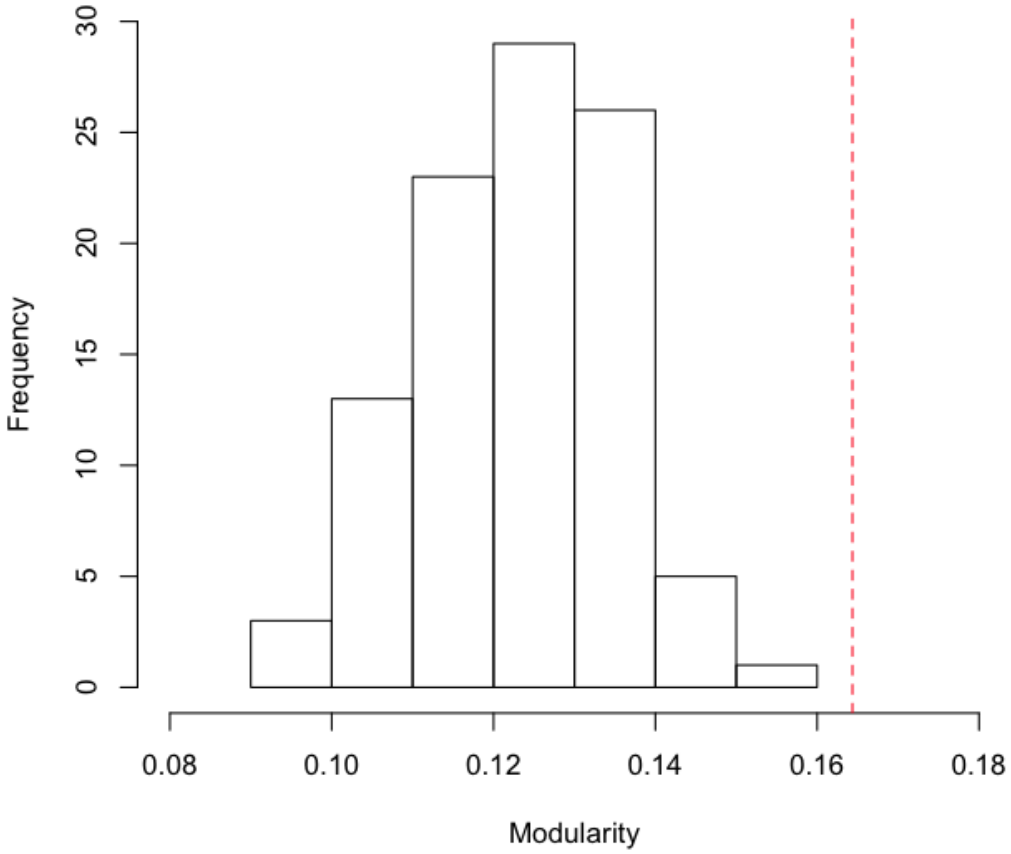

Figure 6. The modularity value for the St. Lawrence beluga population is estimated at 0.164 as represented by the red dotted line. This value falls outside of the expected modularity values of 100 randomly generated networks of 222 individuals. This measure of clustering is significant ($P < 0.01$).

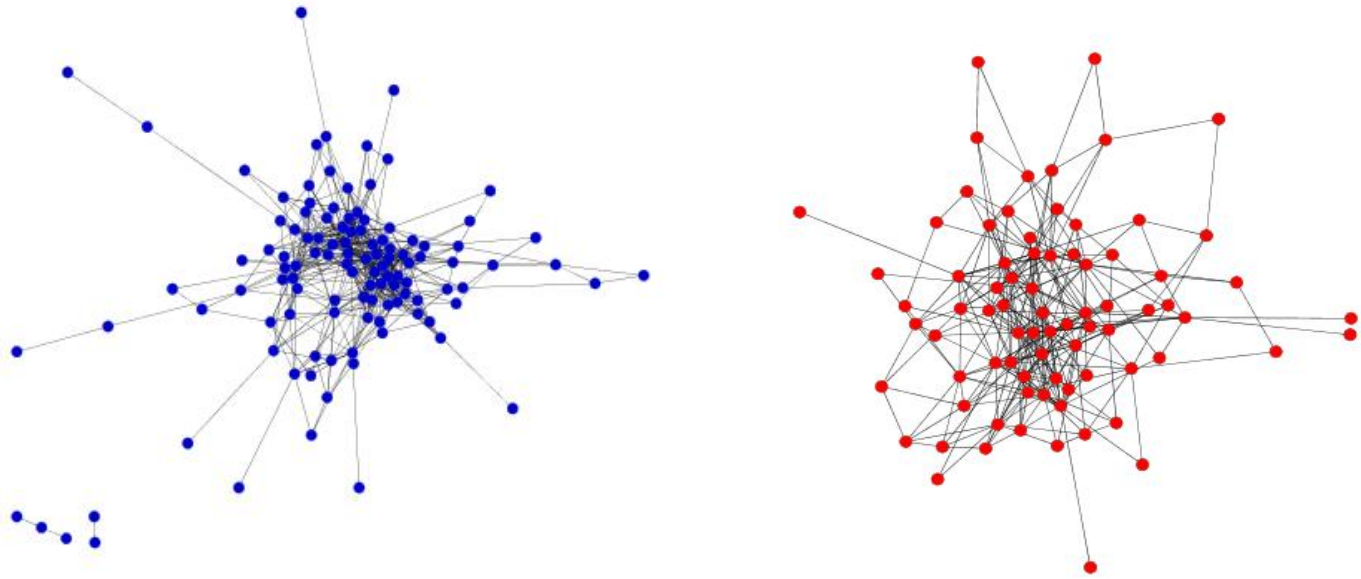

Figure 7. The networks of each sex in the St. Lawrence River beluga population independently. Pairwise sharing values less than 0.4 were filtered out for better visualization. The network for males is on the left, females on the right. Clustering patterns exist for both sexes, but this pattern is only significant for females (males, $P > 0.1$; females, $P < 0.01$).

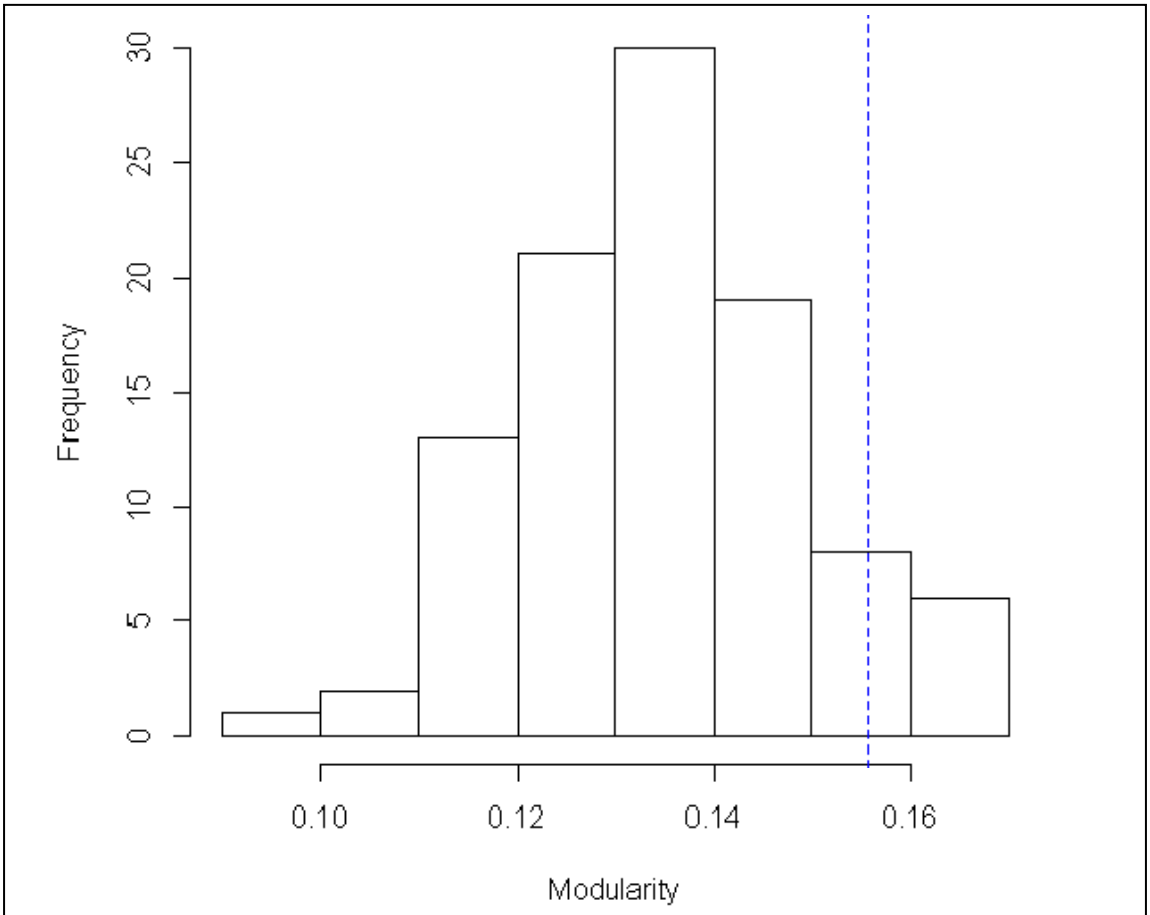

Figure 8. The modularity value for the St. Lawrence River male beluga population is estimated at 0.156 as represented by the blue dotted line. This value falls within the expected modularity values of 100 randomly generated networks of 126 individuals. This measure of clustering is not statistically significant ($P > 0.1$).

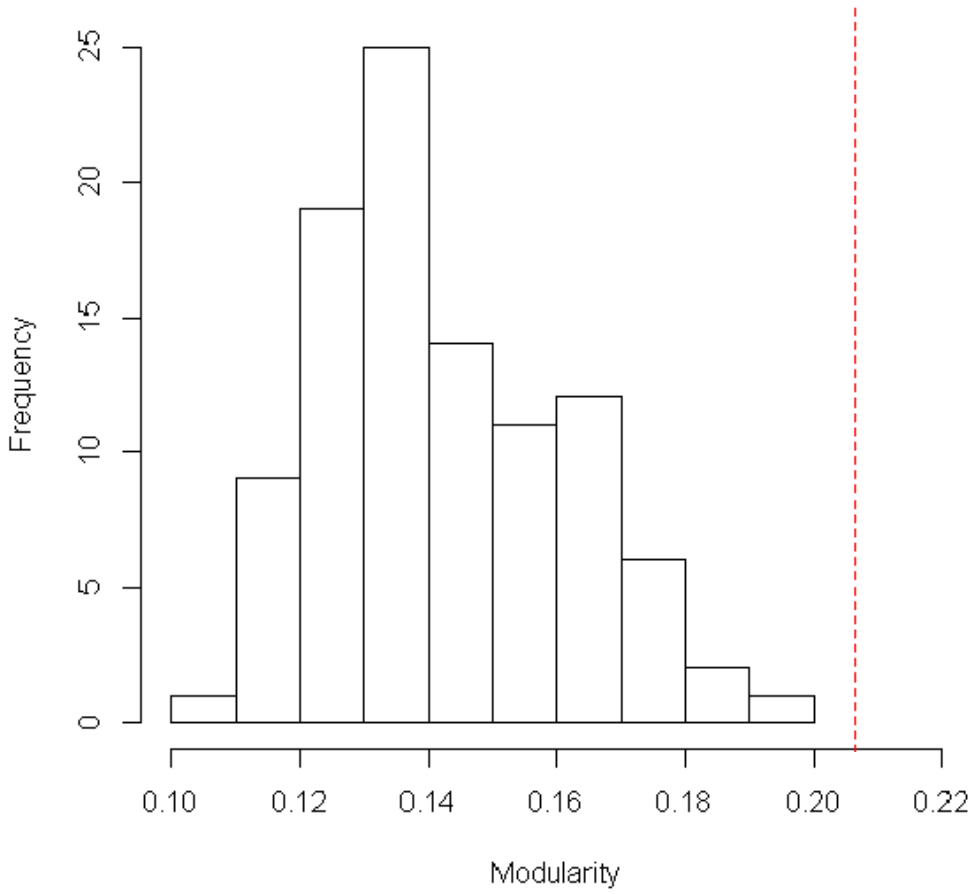

Figure 9. The modularity value for the St. Lawrence River female beluga population is estimated at 0.206 as represented by the red dotted line. This value falls outside of the expected modularity values of 100 randomly generated networks of 88 individuals. This measure of clustering is significant ($P < 0.01$).

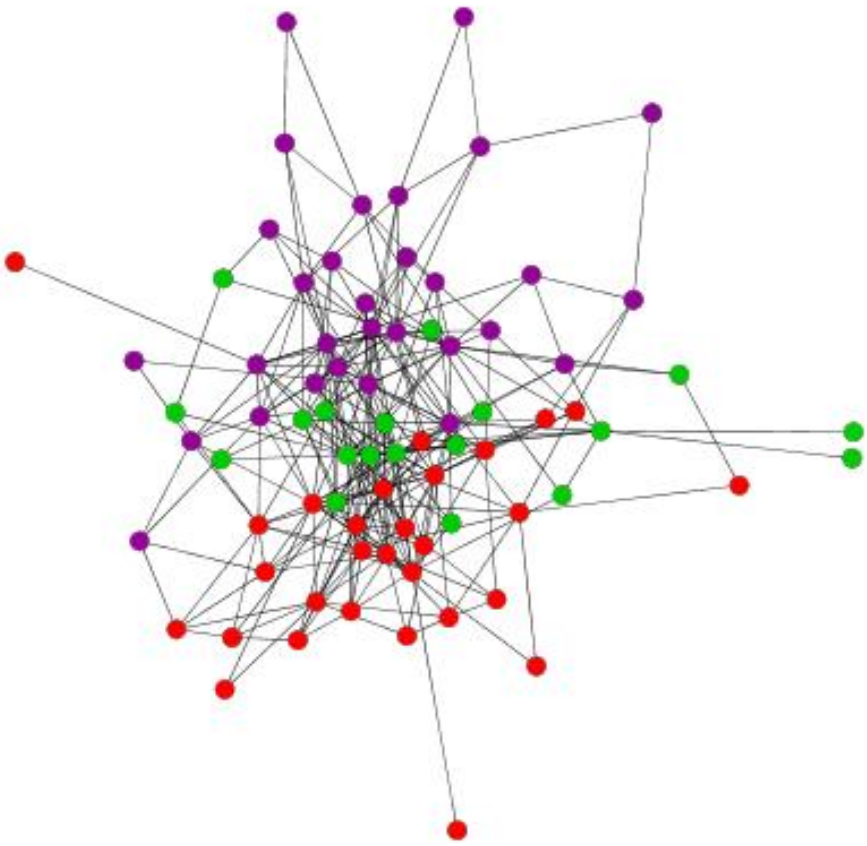

Figure 10. The network of the St. Lawrence River female beluga population independently. Three main clusters exist based on the fast greedy community cluster analysis. These groups have been colour-coded for enhanced visualization.

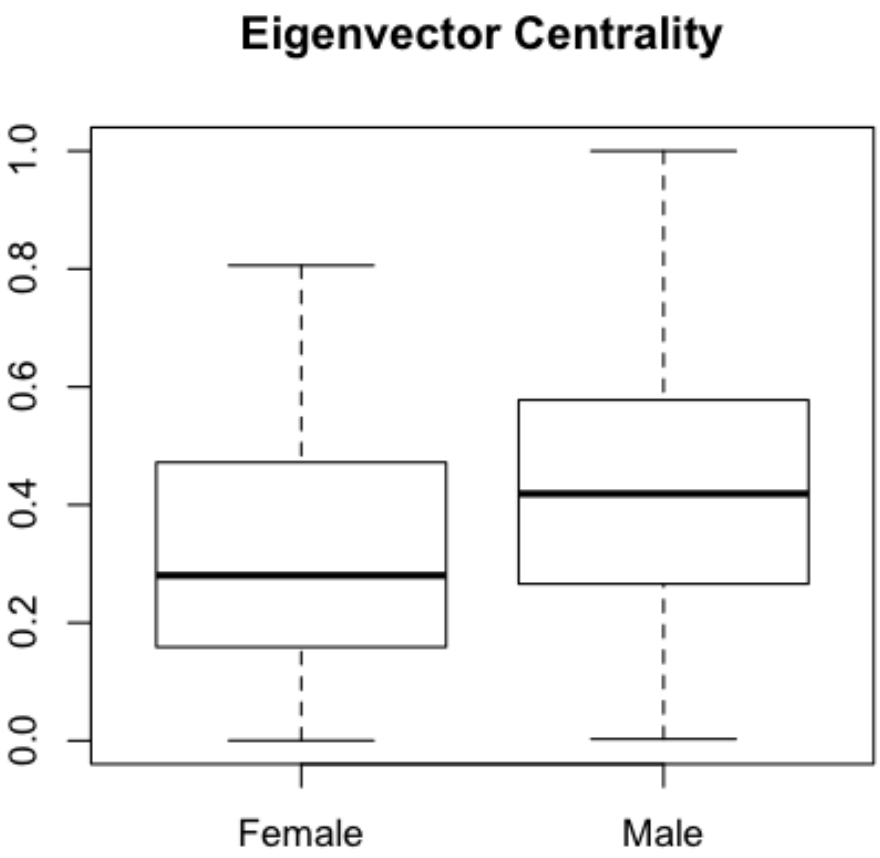

Figure 11. Box plot of the eigenvector centrality values for males and females in the St. Lawrence River beluga population network. The data clearly show that males have higher values than females and are therefore more connected than females. The difference for this measure for connectivity is statistically significant (P = 0.0004).

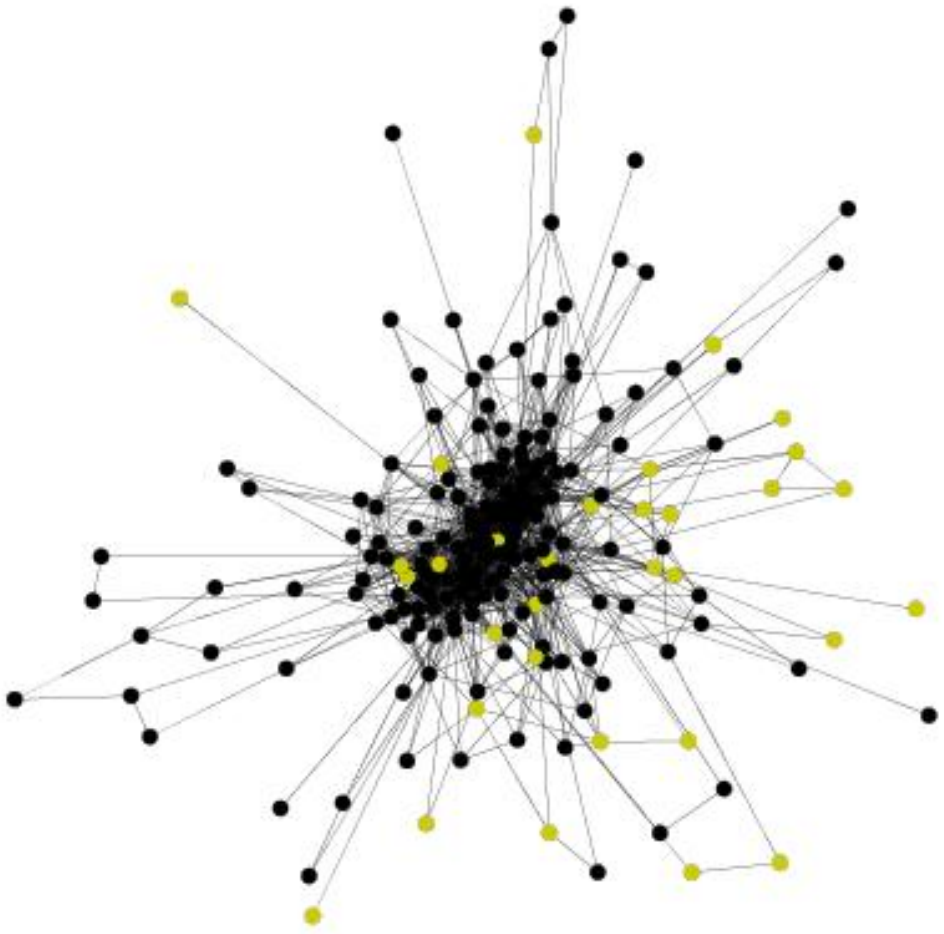

Figure 12. Network of the St. Lawrence beluga population (222 samples) based on pairwise relatedness values. Pairwise sharing values less than 0.4 were filtered out for better visualization. Samples taken from live individuals are represented in black and samples collected from dead individuals are represented in yellow.

References

Aasen E and Medrano JF (1990) Amplification of the ZFY and ZFX genes for sex identification in humans, cattle, sheep and goats. *Biotechnology* **8**:1279-1281.

Archie EA, Moss CJ and Alberts SC (2006) The ties that bind: genetic relatedness predicts the fission and fusion of social groups in wild African elephants. *Proceedings of the Royal Society B: Biological Sciences* **273**:513-522.

Bailey R and Zinger N (1995) *St. Lawrence Beluga Recovery Plan*. World Wildlife Fund, Toronto. Department of Fisheries and Oceans, Mont-Joli, Quebec.

Baker CS, Perry A, Bannister JL, Wienrich MT, Abernethy RB, Calambokidis J, Lein J, Lambertsen RH, Urban Ramirez J, Vasquez O, Clapham PJ, Alling A, O'Brien SJ and Palumbi SR (1993) Abundant mitochondrial DNA variation and world-wide population structure in humpback whales. *Proceedings of the National Academy of Sciences of the USA* **90**:8239-8243.

Barendse W, Armitage SM, Kossareklom LM, Shalom A, Kirkpatrick BW, Ryan AM, Clayton D, Li L, Neibergs HL, Zhang N, Grosse WM, Weiss J, Crighton P, McCarthy F, Ron M, Teale AJ, Fries R, McGraw RA, Moore SS, Georges M, Soller M, Womack JE and Hetzel DJS (1994) A Genetic linkage map of the bovine genome. *Nature Genetics* **6**:227-235.

Barrett-Lennard L, Smith TG and Ellis GM (1996) A cetacean biopsy system using lightweight pneumatic darts, and its effect on the behavior of killer whales. *Marine Mammal Science* 12:14-27.

Béland P (1996) *Beluga: a farewell to whales*. Lyons and Burford Publishing, New York.

Béland P, DeGuise S, Girard C, Lagacé A, Martineau D, Michaud R, Muir DCG, Norstrom RJ, Pelletier E, Ray S and Shugart LR (1993) Toxic Compounds and Health and Reproductive Effects in St. Lawrence Beluga Whales. *Journal of Great Lakes Research* **19**:766-775.

Bennett P (2000) Microsatellites. *Molecular Pathology* **53**:177-183.

Bergmüller R, Heg D and Taborsky M (2005) Helpers in a cooperatively breeding cichlid stay and pay or disperse and breed, depending on ecological constraints. *Proceedings of the Royal Society of London, Series B, Biological Sciences* **272**:325-331.

Bérubé M, Jørgensen H, McEwing R and Palsbøll PJ (2000) Polymorphic di-nucleotide microsatellite loci isolated from the humpback whale, *Megaptera novaeangliae*. *Molecular Ecology* **9**:2181-2183.

Blouin M, Parsons M, Lacaille V, Lotz S (1996) Use of microsatellite loci to classify individuals by relatedness. *Molecular Ecology* **5**:393-401.

Brennin R, Murray BW, Friesen MK, Maiers LD, Clayton JW and White BN (1997) Population genetic structure of beluga whales (*Delphinapterus leucas*): mitochondrial DNA sequence variation within and among North American populations. *Canadian Journal of Zoology* **75**:795-802.

Brodie PF (1969) Duration of lactation in cetacean: an indicator of required learning?. *American Midland Naturalist* **82**:312-314.

Brodie PF (1971) A reconsideration of aspects of growth, reproduction, and behaviour of the white whale (*Delphinapterus leucas*), with reference to the Cumberland South, Baffin Island population. *Journal of the Fisheries Research Board of Canada* **28**:1309-1318.

Brown WM, George M Jr and Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the USA* **76**:1967-1971.

Brown Gladden JG, Ferguson MM and Clayton JW (1997) Matriarchal genetic population structure of North American beluga whales *Delphinapterus leucas* (Cetacea: Monodontidae). *Molecular Ecology* **6**:1033-1046.

Brown Gladden JG, Ferguson MM, Friesen MK and Clayton JW (1999) Population structure of North American beluga whales (*Delphinapterus leucas*) based on nuclear DNA microsatellite variation and contrasted with the population structure revealed by mitochondrial DNA variation. *Molecular Ecology* **8**:347-363.

Bruford M, Ciofi C, Funk S (1998) Characteristics of Microsatellites. In: *Molecular Tools for Screening Biodiversity* (eds Karp A, Isaac P, Ingram D) pp 202-205. Chapman & Hall, London.

Buchanan FC, Friesen MK, Littlejohn RP and Clayton JW (1996) Microsatellites from the beluga whale *Delphinapterus leucas*. *Molecular Ecology* **5**:571-575.

Budowle B, Smith J, Moretti T and DiZinno J (2000) *DNA Typing Protocols: Molecular Biology and Forensic Analysis*. Eaton Publishing, Massachusetts.

Burg TM, Trites AW and Smith MJ (1999) Mitochondrial and microsatellite DNA analyses of harbour seal population structure in the northeast Pacific Ocean. *Canadian Journal of Zoology* **77**:930-943.

Butler JM and Levin BC (1998) Forensic application of mitochondrial DNA. *Trends in Biotechnology* **16**:158-162.

Caron LMJ and Smith TG (1990) Philopatry and Site Tenacity of Belugas, *Delphinapterus leucas*, Hunted by the Inuit at the Nastapoka Estuary, Eastern Hudson Bay. In: *Advances in Research on the Beluga Whale, Delphinapterus leucas* (eds Smith TG, St. Aubin DJ and Geraci GR) pp 69-79. Canadian Bulletin of Fisheries and Aquatic Science 224. Department of Fisheries and Oceans, Ottawa.

Church GM and Kieffer-Higgins S (1988) Multiplex DNA Sequencing. *Science* **240**:185-188.

Clutton-Brock TH (2009) Structure and function in mammalian societies. *Philosophical Transaction of the Royal Society B* **364**:3229-3242.

Craig AS, Herman LM and Pack AA (2002) Male mate choice and male-male competition coexist in the humpback whale (*Megaptera novaeangliae*). *Canadian Journal of Zoology* **80**:745-755.

Creel S (1990) How to Measure Inclusive Fitness. *Proceedings: Biological Sciences* **241**:229-231.

Creel S (2001) Social dominance and stress hormones. *Trends in Ecology and Evolution* **16**:491-495.

Croft DP, James R, Krause J (2008) *Exploring Animal Social Networks*. Princeton University Press, New Jersey.

Csardi G, Nepusz T (2006) The igraph software package for complex network research. InterJournal, Complex Systems 1695.

Environmental Health Perspectives **103**:73-77.

Donovan GP (1992) Report of the sub-committee on small cetaceans. *Report of the International Whaling Commission* **42**:178-234.

Durand E, Blum MGB and François O (2007) Prediction of group patterns in social mammals based on a coalescent model. *Journal of Theoretical Biology* **249**:262-270.

Emlen ST (1982a) The evolution of helping I: An ecological constraints model. *American Naturalist* **119**:29-39.

Emlen ST (1982b) The evolution of helping II: The role of behavioral conflict. *American Naturalist* **119**:40-53.

Emlen ST (1994) Benefits, constraints and the evolution of the family. *Trends in Ecology and Evolution* **9(8)**:282-285.

Emlen ST (1995) An evolutionary theory of the family. *Proceedings of the National Academy of Science of the USA* **92**:8092-8099.

Environment Canada (1992) Remedial action plan for the St. Lawrence River (Cornwall) Area of Concern Stage 1 Report. *Environment Canada*, Ontario Region, Toronto.

Environment Canada (1997) Remedial action plan for the St. Lawrence River (Cornwall) Area of Concern Stage 2 Report. *Environment Canada*, Ontario Region, Toronto

Frasier TM (2011) *GeNetwork: A program to facilitate individual-based network analysis on genetic data*. Saint Mary's University, Halifax.

Frasier TR, Koroscil SM, White BN and Darling JD (2011) Assessment of population sub-structure in relation to summer feeding ground use in the eastern North Pacific gray whale. *Endangered Species Research* **14**:39-48.

Frère CH, Krützen M, Mann J, Watson-Capps JJ, Tsai YJ, Patterson EM, Connor R, Bejder L and Sherwin WB (2010) Home range overlap, matrilineal and biparental kinship drive female associations in bottlenose dolphins. *Animal Behaviour* 80:481-486.

Gilson A, Syvanen M, Levine K and Banks J (1998) Deer gender determination by polymerase chain reaction: validation study and application to tissues, bloodstains, and hair forensic samples in California. *California Fish and Game* **84**:159-169.

Godinho R, Crespo EG and Ferrand N (2008) The limits of mtDNA Phylogeography: complex patterns of population history in a highly structured Iberian lizard are only revealed by the use of nuclear markers. *Molecular Ecology* **17**:4670-4683.

Goldzien AW (1987) Tamarins and marmosets: communal care of offspring. In *Primate*

Societies (eds Smutz BB, Wrangham RW, Cheney DL, Strusaker TT and Seyfarth RM) pp 34-43. Chicago University Press, Chicago.

Hamilton WD (1964) The genetical evolution of social behaviour. *Journal of Theoretical Biology* **7**:1-52.

Hoelzel AR (2009) Evolution of population genetic structure in marine mammal species. In *Population genetics for animal conservation* (eds Bertorelle G, Bruford MW, Hauffe HC, Rizzoli A and Vernesi C) pp 295-318. Cambridge University Press, New York.

Inman AJ and Krebs JR (1987) Predation and group living. *Trends in Ecology and Evolution* **2**:31-32.

Koenig WD, Pitelka FA, Carmen WJ, Mumme RL and Stanback MT (1992) The evolution of delayed dispersal in cooperative breeders. *Quarterly Reviews in Biology* **67**:111-150.

Krützen M, Valsecchi E, Connor RC, Sherwin WB (2001) Characterization of microsatellite loci in *Tursiops aduncus*. *Molecular Ecology Notes* **1**:170-172.

Li CC, Weeks DE, Chakravarti A (1993) Similarity of DNA fingerprints due to chance of relatedness. *Human Heredity* **43**:45-52.

Lusseau D (2003) The emergent properties of a dolphin social network. *Proceedings of the Royal Society B: Biological Sciences* **270**:186-188.

Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. *Genetics* **152**:1753-1766.

Maiers LD, Friesen MK, Wiens AV and Clayton JW (1996) *Use of DNA Microsatellites in Beluga Whale (Delphinapterus leucas) Population Genetics*. 77th Canadian Technical Report of Fisheries and Aquatic Sciences 2115. Fisheries and Oceans, Manitoba.

Majolo B, Vizioli AD and Schino G (2008) Costs and benefits of group living in primates: group size effect on behaviour and demography. *Animal Behaviour* **76**:1235-1247.

Mann J, Connor RC, Tyack PL and Whitehead H (2000) *Cetacean Societies: Field Studies of Dolphins and Whales*. University of Chicago Press, Chicago.

Martien KK, Baird RW, Hedrick NM, Gorgone AM, Thieleking JL, McSweeny DJ, Robertson

KM and Webster DL (2012) Population structure of island-associated dolphins: Evidence from mitochondrial and microsatellite markers for common bottlenose dolphins (*Tursiops truncates*) around the main Hawaiian Islands. *Marine Mammal Science* **28(3)**:E208-E232.

Martineau D, DeGuise S, Fournier M, Shugart L, Girard C, Lagacé A and Béland P (1994) Pathology and toxicology of beluga whales from the St. Lawrence Estuary, Quebec, Canada. Past, present and future. *The Science of the Total Environment* **154**:201-215.

Martineau D, Lemberger K, Dallaire A, Labelle P, Lipscomb TP, Michel P, Mikaelian I (2002) Cancer in wildlife, a case study: beluga from the St. Lawrence estuary, Quebec, Canada. *Environmental Health Perspectives* **110**:285-292.

Mech LD (1981) *The Wolf: The Ecology and Behaviour of an Endangered Species*. University of Minnesota Press, Minneapolis.

Moore WS (1995) Inferring Phylogenies from mtDNA Variation: Mitochondrial-Gene Trees Versus Nuclear-Gene Trees. *Evolution* **49(4)**:718-726.

Newman MEJ (2006) Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the USA* **103**:8577-8582.

Noren DP and Mocklin JA (2011) Review of cetacean biopsy techniques: Factors contributing to successful sample collection and physiological and behavioural impacts. *Marine Mammal Science* **28**:154-199.

O'Corry-Crowe GM, Lydersen C, Heide-Jørgenson MP, Hansen L, Mukhametov LM, Dove O and Kovacs KM (2010) Population genetic structure and evolutionary history of North Atlantic beluga whales (*Delphinapterus leucas*) from West Greenland, Svalbard and the White Sea. *Polar Biology* **33**:1179-1194.

O'Corry-Crowe GM, Suydam RS, Rosenberg A, Frost KJ and Dizon AE (1997) Phylogeography, population structure and dispersal patterns of the beluga whale *Delphinapterus leucas* in western Nearctic revealed by mitochondrial DNA. *Molecular Ecology* **6**:955-970.

Packer CR, Pusey AE and Eberly LE (2001) Egalitarianism in female African lions. *Science* **293**:690-693.

Pamillo P and Crozier RH (1996) Reproductive Skew Simplified. *Oikos* **75**:533-535.

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

Palsbøll PJ, Clapham PJ, Mattila DK, Larsen F, Sears R, Siegismund HR, Sigurjonsson J, Vasquez O and Arctander P (1995) Distribution of mtDNA haplotypes in North Atlantic humpback whales: the influence of behaviour on population structure. *Marine Ecology Progress Series* **116**:1-10.

Queller DC (1992) Quantitative Genetics, Inclusive Fitness, and Group Selection. *The American Naturalist* **139**:540-558.

Queller DC, Strassmann JE, Hughes CR (1993) Microsatellites and Kinship. *Trends in Ecology and Evolution* **8**:285-288.

R Development Core Team (2012) R: a language and environment for statistical computing and graphics. R foundation for statistical computing, Vienna, Austria. URL http://www.R-project.org.

Richard KR, Whitehead H and Wright JM (1996) Polymorphic microsatellites from sperm whales are their use in the genetic identification of individuals from naturally sloughed pieces of skin. *Molecular Ecology* **5**:313-315.

Richman LA and Dreier SI (2001) Sediment contamination in the St. Lawrence River along the Cornwall, Ontario waterfront. *Journal of Great Lakes Research* **27**:60-83.

Rooney AP, Merritt DB and Derr JN (1999) Microsatellite Diversity in Captive Bottlenose Dolphins (Tursiops truncates). *Journal of Heredity* **90**:228-231.

Schradin C, König B and Pillay N (2010) Reproductive competition favours solitary living while ecological constraints impose group-living in African striped mice. *The Journal of Animal Ecology* **79**:515-521.

Sergeant DE (1973) Biology of white whales (*Delphinapterus leucas*) in western Hudson Bay. *Journal of Fisheries Research Board of Canada* **30**:1065-1090.

Sergeant DE and Brodie PF (1969) Body Size in White Whales *Delphinapterus leucas*. *Journal of Fisheries Research Board of Canada* **26**:2561-2580.

Seutin G, White B and Boag P (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology* **69**:82–90.

Sherman PW, Lacey EA, Reeve HW and Keller L (1995) The eusociality continuum. *Behavioural Ecology* **6**:102-108.

Smoot M, Ono K, Ruscheinski J, Wang PL, Ideker T (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**: 431–432.

Stewart BE and Stewart REA (1989) Delphinapterus leucas: Mammalian Species. *American Society of Mammologists* **336**:1-8.

Turgeon J, Duchesne P, Colbeck GJ, Postma LD and Hammill MO (2012) Spatiotemporal segregation among summer stocks of beluga (*Delphinapterus leucas*) despite nuclear gene flow: implication for the endangered belugas in eastern Hudson Bay (Canada). *Conservation Genetics* **13**:419-433.

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

Van de Castreele T, Galbusera P, Matthysen E (2001) A comparison of microsatellite-based pairwise relatedness estimators. *Molecular Ecology* **10**:1539-1549.

Waldick RC, Brown MW and White BN (1999) Characterization and isolation of microsatellite loci from the endangered North Atlantic right whale. *Molecular Ecology* **8**:1763-1765.

West SA, Griffin AS and Gardiner A (2007) Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology* **20**:415-432.

Wey T, Blumstein DT, Shen W and Jordan F (2008) Social network analysis of animal behaviour: a promising tool for the study to sociality. *Animal Behaviour* **75**:333-344.

Witteveen BH, Straley JM, von Ziegesar O, Steel D and Baker CS (2004) Abundance and mtDNA differentiation of humpback whales (Megaptera novaeangliae) in the Shumagin Islands, Alaska. *Canadian Journal of Zoology* **82**:1352-1359.