
G. O’Corry-Crowe, W. Lucey, F. I. Archer, and B. Mahoney

**Introduction**

Individuals on the frontier of a species’ range typically define the geographic limits of suitable conditions for the maintenance of viable populations, but they can also represent the vanguard of a range or niche expansion or the relics from past retreats. As such, groupings of individuals at the range edge can tell us much about a species’ niche preference limitations, its demographic history, and its adaptability and potential for range expansion or shifts. It is for these reasons that range limits are receiving growing attention in the context of determining how species will face the increasing challenges of environmental and climate change (Gaston, 2009; Sexton et al., 2009, 2011).

The geographic range limits of species are often characterized by low population density (Jarema et al., 2009) where populations may be subject to the genetic and demographic consequences of small population size, including negative aspects of inbreeding (Hedrick and Kalinowski, 2000) and the Allee effect (Couchamp et al., 1999). In social species, these effects can be either compounded or alleviated by the size, genetic composition (e.g., kin- vs. nonkin-based), and breeding patterns of founding and immigrant groups. Furthermore, the extent of gene flow with larger central populations or other edge populations can influence the effective population size of marginal populations and the allelic variation upon which selection can act (Sexton et al., 2009, 2011).

Thus, determining the genetic origins of marginal populations, their genetic composition and breeding patterns, including levels of relatedness and inbreeding, and the extent of continuing genetic exchange with central populations can provide insight into the factors that influence the establishment and persistence of small groups or populations on the geographic frontier of a species’ range. Such insight is key to predicting how species will respond to climate and other environmental change and has immediate application to assessing the viability of core populations, especially those experiencing declines and habitat loss.

Beluga whales, *Delphinapterus leucas*, have a nonuniform distribution across the Arctic and Subarctic (O’Corry-Crowe, 2008) shaped by...
glacial history, contemporary resource dispersion, physiological and physical constraints, and the behavioral tendency of this species to form social aggregations and return to traditional locations each year (Suydam et al., 2001; Lydersen et al., 2001; Laidre et al., 2008; Colbeck et al., 2013). Within the Gulf of Alaska, a small population of beluga whales is resident in the waters of Cook Inlet (Fig. 1; Rugh et al., 2000; Allen and Angliss, 2012). Satellite telemetry and sightings reveal that the population’s current range is restricted to the mid and upper reaches of the inlet (Rugh et al., 2000; Hobbs et al., 2005). While dispersal cannot be completely excluded, this population of whales has been effectively geographically and genetically isolated from populations to the west and north for a long enough time that genetic divergence via drift is evident at neutral markers (O’Corry-Crowe et al., 1997, 2010).

The Cook Inlet population experienced dramatic declines in the 1990’s, with an estimated reduction from 653 whales in 1994 to 347 whales by 1998 (Hobbs et al., 2000). This decline was accompanied by a northward contraction of population range within the inlet (Rugh et al., 2000, 2010). Subsequent research revealed that the population is not recovering but continues to decline at a much reduced rate of 0.6–1.6% per year (Lowry et al., 2006; Allen and Angliss, 2012; Hobbs et al., 2015). Its restricted range, low abundance, and apparent failure to recover following more than a decade of increased protection raise concerns over the viability of the Cook Inlet population (Allen and Angliss, 2012). Its geographic isolation and location on the warm edge of the species’ range as well as its proximity to an increasingly urbanized region of Alaska add...
concerns over the population’s ability to respond to climate change and other challenges.

Over the past four decades periodic sightings of small numbers of beluga whales (typically ≤ 2) have been reported in the Gulf of Alaska outside of Cook Inlet (Harrison and Hall, 1978; Laird et al., 2000; Calkins1). These records have been attributed to short-term extralimital movements by Cook Inlet animals (Laird et al., 2000). One location, however, stands in stark contrast to the sporadic pattern of sightings in the rest of the gulf. Beluga whales, about 5–12 individuals, have been regularly reported in Yakutat Bay, some 700 km to the east of Cook Inlet and 800 km southeast of the conterminous range of the depleted Cook Inlet population (Fig. 1; Hubbard et al., 2003; O’Corry-Crowe et al., 2000). One term extralimital movements by Cook Inlet animals (Laidre et al., 2000). Further, traditional ecological knowledge (TEK) from the Yakutat Tlingit Tribe and local fishermen document consistent sightings of small groups (≤20) of beluga whales in the bay dating back to 1938 (Lucey et al., 2015).

Yakutat Bay is a deep glacial fiord fed by several tidewater glaciers at its head. A field study initiated in 2005 found that beluga whales selected habitats associated with actively calving glaciers that may reflect a unique ecology for this species in the North Pacific (O’Corry-Crowe et al., 2005). Furthermore, sightings over the past decade have occurred in all seasons and include adults, juveniles, and young of the year (Lucey et al., 2015).

The evidence to date may indicate that beluga whales regularly move between Cook Inlet and Yakutat Bay, and that the current whales are recent visitors or colonizers from Cook Inlet, possibly even a single family group. Such scenarios would potentially extend the known range and habitat of the endangered Cook Inlet population. Alternatively, these findings raise the possibility of the persistence of a very small reproductive group or population of whales in a previously unknown location and unique habitat beyond the known southern limit of the species’ range.

Determining the genetic origins of the Yakutat whales, their breeding patterns, and level of kinship and inbreeding, and the extent of continuing gene flow with the likely source population, Cook Inlet, will provide critical insight into: 1) the relationship between Yakutat Bay and Cook Inlet, and 2) the viability of populations of beluga whales at the warm edge of the species’ range. This is essential to modeling the future prospects of the larger Cook Inlet population already struggling to recover from recent declines. We conducted such a study and present our findings here. Using remote tissue biopsy methods, multi-locus genotyping, and a suite of moment and likelihood-based statistical analyses, including a new group exclusion-assignment test (GELATo) we introduce here, we addressed the following:

1) Is there evidence of long-term use of Yakutat Bay by individual whales?
2) Do the Yakutat whales have a lower neutral genetic diversity than similar sized sample sets from the Cook Inlet population?
3) Are Yakutat whales, on average, more related or inbred than whales sampled in Cook Inlet?
4) As a group, are the Yakutat whales likely to originate from Cook Inlet and or any other reference population?
5) Are individual Yakutat whales likely to be immigrants from Cook Inlet or any of the other populations of beluga whales in Alaskan waters?

**Methods**

Skin plug samples were collected from free-swimming belugas in Yakutat Bay, Alaska, via remote biopsy between 2002 and 2008. Tissue was preserved in 20% DMSO-NaCl solution and archived at -20°C. Total DNA was extracted from each tissue sample by established protocols and screened for variation within 410bp of the mtDNA control region and 8 independent microsatellite loci according to previously published methods (O’Corry-Crowe et al., 1997, 2010). The gender of each sample was also determined by PCR-based methods (Fain and LeMay, 1995), and replicate genotyping, sequencing, and gender determination was conducted to confirm genetic scores.

Permutation-based estimation of observed heterozygosity (Hs) and cumulative probabilities of identity (P Id), as well as calculation of the total number of alleles across all loci were conducted using the program Doh (Brzustowski, 2002). To compare estimates of such genetic parameters between Yakutat and similar-sized sample sets from another population we used a customized script written in Access, Go-Random (Beaman and O’Corry-Crowe2) to generate multiple subsets of samples of equivalent magnitude chosen at random from the population’s sample set. MCMC exact tests (500,000 iterations) were used to test for deviations from Hardy-Weinberg expectations and levels of genetic differentiation at both microsatellite and mtDNA loci using Genepop 4.1 (Rousset, 2008). F statistics were also performed in Genepop.

Considering the small sample size from Yakutat (see below), we used POWSIM 4.1 (Ryman and Palm, 1993) to estimate critical probabilities of identity (P Id) for the estimated number of free-swimming belugas in Yakutat Bay.

4Mention of trade names or commercial firms does not imply endorsement by the National Marine Fisheries Service, NOAA.
Bootstrap resampling (10,000 reps) was used to test significance in mean \( r \) among groupings and mean \( F \) within individuals. Data from 78 whales sampled from Cook Inlet were used in some of these analyses.

We used assignment methods and genetic data from five reference populations of beluga whale in Alaska waters, including the Cook Inlet population, (Cook Inlet, Bristol Bay, Norton Sound, Chukchi Sea, and Beaufort Sea, \( n=494 \)) (O’Corry-Crowe et al.\(^6\)) to determine the likely origins of the Yakutat whales. Individuals were assigned to populations based on estimated likelihoods of their genotype and mtDNA haplotype arising in each of the sampled populations under assumptions of random assortment of alleles and independence of loci (Paetkau et al., 1995; Banks and Eichert, 2000). Analyses of diploid and haploid data were conducted separately as well as combined using the WHICHRUN 4.1 program (Banks and Eichert, 2000).

We used the Bayesian model-based clustering method, STRUCTURE 2.3.4 (Pritchard et al., 2000, 2010), to estimate the likely number of genetic clusters or populations, \( K \), in the Gulf of Alaska dataset (Yakutat and Cook Inlet) and to assign individuals to cluster of origin. Multiple MCMC runs (50,000 burn-in followed by 10\(^6\) iterations) of the data for a range of values of \( k \) were conducted for both admixture and no admixture ancestry models. To accommodate possible limited information content (i.e., due to small \( n \), low divergence) on structure in our dataset for this algorithm, we also applied the LOCPRIOR option that uses information on sampling location as an unbiased informative prior (Hubisz et al., 2009). We also used the non-MCMC, iterative method, FLOCK 3.1 (Duchesne et al., 2013) to estimate the number of populations, \( K \), and allocate individuals to those populations (Duchesne and Turgeon, 2012). To assess the method’s ability to accommodate limited information content on structure in the data, we conducted a series of trials (each with 20 runs X 50 iterations) on two genetically differentiated populations (\( K=2 \), Bristol Bay and Cook Inlet) varying sample size with each trial.

We also used GENECLASS2 (Piry et al., 2004) to assign or exclude reference populations as the origin of individuals or groups of individuals. Employing the Bayesian and frequency-based criteria of Rannala and Mountain (1997), Baudouin and Lebrun (2000), and Paetkau et al. (1995) log likelihood ratio scores (\( \text{Score}_{i,T} \)) were used to assign likely population of origin (highest rank score) and exclude other populations (rank score falls below a predetermined threshold, e.g. \( \text{Score}_{i,T} <0.5 \)). Furthermore, MCMC methods were used in a semi-Bayesian exclusion procedure where observed probabilities of test samples were compared to those of simulated samples (\( n=1,000 \) to 10,000) from each population (Cornuet et al., 1999, Baudouin et al., 2004).

In determining whether a test group of individuals comes from one, any, or none of a series of reference populations, however, the approach used in GENECLASS2 has a series of shortcomings. By only considering relative likelihoods among a series of reference populations, the logL methods do not estimate genotypic (or haplotypic) likelihoods of test groups relative to those of other similar sized groupings within each population, while the MCMC methods tend towards excessive exclusion of reference populations (see Discussion). We therefore developed a test that estimates these relative likelihoods and improves exclusion accuracy (Group ExcLusion-Assignment Test – GELAto). Specifically, we addressed the following question: “assuming the whales sampled in Yakutat Bay were from a single group or population, were any of the reference populations sampled the likely source, and if so which was the most likely?”

GELATO is a permutation-based test that compares the distributions of $F_{st}$ values for a population of unknown origin to potential source populations with null $F_{st}$ distributions generated from comparing similarly sized random samples from the source populations back to themselves. For a given “unknown” population, $U$, with a sample size of $N_u$, the likelihood of its membership to a given source population, $K$, is calculated as follows. A random sample of $N_k$ individuals is selected from $K$ without replacement ($K_u$). These individuals are used to calculate an $F_{st}$ value from the null distribution ($F_{st-null}$) by comparing them to the remaining individuals in $K$ ($K_k$). An “observed” $F_{st}$ ($F_{st-obs}$) is then calculated between all individuals from $U$ and $K_k$. This process is repeated 1,000 times to generate a distribution of $F_{st-null}$ and $F_{st-obs}$. The null distribution is then modeled as a Normal distribution with the mean and standard deviation of $F_{st-null}$.

The likelihood that $U$ came from $K$ is the product of the likelihood (sum of log-likelihoods) of all values in $F_{st-obs}$ given the fitted null distribution. The probability that population $U$ came from $K$ is its likelihood of $F_{st-obs}$ given the null for $K$ divided by the product of likelihoods to each potential source population. In the results below, we express the likelihoods as deviations of the log-likelihood from the maximum observed log-likelihood. As can be seen, this method can also be used as an exclusion test, where low likelihoods (negative log-likelihoods) for all source populations would indicate that the test group was either of mixed origin, a nonrandom sample from one of the source populations, or from an unsampled population.7

Results
A total of 10 skin biopsy samples were collected from free swimming beluga whales in Yakutat Bay, Alaska, between 2002 and 2008. PCR-based tests found similar proportions of males and females in the sample set (Table 1).

**Genetic Identity Analysis**

The genetic identity analysis revealed that of the 45 pairwise comparisons involving the 10 samples across 8 microsatellite loci, two dyads differed at one allele in one locus (locus D3 in both cases), while all others differed by multiple alleles at two or more loci. Reanalysis revealed that both single allele differences were the result of genotyping error. An independent assessment of genotyping error rates in a large dataset from Bristol Bay (n=547) found that single allelic differences at one of eight loci were likely the result of a genotyping error, while differences at two or more loci reflected real differences among individuals (O’Corry-Crowe et al.8). This has also been found in other species for as few as six loci (Kalinowski et al., 2006). Our findings indicate that the two exact matches represent two cases of genetic recapture. Finally, the complete removal of locus D3 resulted in estimated pairwise $P_{ID}$ orders of magnitude lower (2.72 x 10-8 and 1.01 x 10-5) than probable whale abundance in Yakutat further indicating that two whales were resampled during the course of the study. The first case involved a male whale, DL-Yak-2-04, initially sampled on 01 May 2004 and resampled a year later on 19 May 2005. The second case involved a female whale, DL-Yak-1-04, biopsied on the same day as the initial sampling of DL-Yak-2-04 and subsequently resampled 3 years and 3 months later on 04 Aug. 2007. These findings establish for the first time the occurrence of individual whales in Yakutat Bay across multiple years.

**Genetic Diversity and Differentiation**

Only one mtDNA haplotype, Haplotype 2, was documented in all the Yakutat animals. Haplotype 2 is present within Cook Inlet at moderate frequencies (14%) but is not the most common lineage in this population. This haplotype has also been found in Norton Sound (0.52%), the Chukchi Sea (47%), and the Beaufort Sea (5.6%) (O’Corry-Crowe et al.8). Excluding recaptures (see below), average heterozygosity per nuclear locus for Yakutat ranged from 0.25–0.875, and no loci were found to deviate from Hardy-Weinberg expectations ($p > 0.136$).

The Yakutat whales were found to have lower mean levels of heterozygosity ($H_{obs} = 0.580$ vs. 0.715), lower numbers of alleles (total allele count = 22 vs. 36.3), and higher estimated cumulative probabilities of identity-by-descent ($P_{ID} = 3.24 x 10^{-5}$ vs. 2.92 x 10^{-8}) than sets of samples (n=50) of similar magnitude drawn at random from the Cook Inlet population (Fig. 2). Even when the Cook Inlet resampling (n=25 sets) was restricted to individuals with the same mtDNA haplotype, the diversity indices were still lower and $P_{ID}$ higher for the single-matrilinie group in Yakutat Bay compared to Cook Inlet (Fig. 2).
The Yakutat sample set was significantly differentiated from Cook Inlet for both microsatellite loci ($F_{st} = 0.106; P = 0.065 - 0.0001$) and mtDNA ($F_{st} = 0.536; P < 0.0001$). The POWSIM analysis revealed that the eight microsatellite loci used had sufficient power to provide a 99% probability of detecting an $F_{st}$ as low as 0.05 for the nuclear data when analyzing skewed sample sizes of 8 and 78 for pop1 and pop2 (i.e., Yakutat Bay and Cook Inlet sample sizes) and yielded a type 1 error close to 0.05 ($\alpha = 0.03 - 0.08$).

**Inbreeding and Relatedness**

All three estimators of relatedness yielded similar results although ML methods tended toward higher $r$ values for pairs of highly homozygous individuals with few differences. When the entire Gulf of Alaska (i.e., Cook Inlet and Yakutat Bay) was treated as a single population, the highest estimated relatedness for all estimators was found among pairs of Yakutat whales with many in excess of $r = 0.3$ (Fig. 3a). Average relatedness among beluga whales within Yakutat was significantly higher than among whales within Cook Inlet (e.g., observed difference in mean $r_{QeullerGt} = 0.410$, which lies well outside the 99% confidence intervals of the bootstrap analysis, Fig. 3b). Furthermore, average relatedness within Yakutat was significantly higher than average relatedness between Cook Inlet and Yakutat whales (e.g., observed difference in mean $r_{QeullerGt} = 0.418$, which lies well outside the 99% confidence intervals of the bootstrap analysis, Fig. 3c).

The Yakutat whales were also found to have higher inbreeding coefficients, on average, than whales in Cook Inlet. Significance tests differed among the two estimators with the observed difference falling outside the 95% confidence intervals of the bootstrap analysis for the Lynch and Ritland estimator and inside for the Ritland estimator. Restricting the estimated allelic frequencies to include only Yakutat whales, average dyadic relatedness coefficients were lower but close relationships ($r > 0.4$) were still evident.
Assignment and Exclusion Tests

Using the assignment criteria in GENECLASS2, all eight Yakutat whales had the highest likelihood of arising within the Cook Inlet population. However, most of the Yakutat whales’ assignment likelihoods to Cook Inlet fell within the lower range of likelihoods observed for Cook Inlet animals assigned to their own population. The remaining Yakutat whales had lower likelihoods (Fig. 4). This pattern was also observed with WHICHRUN, including tests that used microsatellite and mtDNA data combined. These findings indicate that while the likelihood of encountering each of the Yakutat whales’ multilocus genotype in a given reference population is highest for the Cook Inlet population, most of the Yakutat genotypes have relatively low likelihoods of arising in the Cook Inlet population.

The STRUCTURE analyses where information on sample location was not used to assist clustering outcomes did not provide a clear signal of structure in the Gulf of Alaska dataset. When structure was apparent (i.e., \( k \geq 2 \)), the Yakutat whales always clustered together. When prior information on sample origin was included, the most likely number of genetic groups was \( K = 2 \) (i.e., \( \Pr(K | X) = 1 \)). These two groupings corresponded unambiguously to Yakutat and Cook Inlet for both admix and noadmix ancestry models with all individuals from Yakutat assigned to one cluster and all individuals from Cook Inlet assigned to the other (Fig. 5).

The FLOCK analysis identified \( K = 2 \) genetic groups (PL \( \geq 6 \)), allocated the Cook Inlet samples across both groups, and allocated all the Yakutat whales to one of these groups. The trials involving Bristol Bay and Cook Inlet found that with moderate to large sample sizes (\( n \geq 20 \)) FLOCK performed well (i.e., estimated \( K = 2 \), allocated >80% of individuals to their sampled population). By contrast, when the sample size from even one location was lowered to values on the order of \( n=10 \), FLOCK rarely yielded pla-
teau sequences that were long enough (i.e., PL ≥ 6 for a run number of 50) to estimate K with confidence (i.e., yielded an “undecided” verdict). Furthermore, a substantial number of individuals from either population were mis-assigned.

The analytical group assignment tests in GENECLASS2 all assigned the Yakutat whales to Cook Inlet. The assignment criterion, -log10(L), was lowest for Cook Inlet and the score for the Yakutat group of individuals, ScoreYak = 100%. By contrast, the MCMC resampling procedure in GENECLASS2 gave the opposite result: Yakutat as a group was excluded from Cook Inlet. Further analysis, however, revealed that this test consistently “excluded” subsets of individuals from their own reference population when tested as a group (see Discussion).

The group exclusion-assignment test, GELATo, revealed that while the Yakutat whales have the strongest affinity with Cook Inlet, as a group they have a very low likelihood (negative LogL) of coming from Cook Inlet (Fig. 6). The test excluded the Yakutat whales from all baseline populations including Cook Inlet for both the microsatellite (LogL = -68.63 to -268.6) and the mtDNA (LogL = -17.29 to -106.1) data. The Yakutat sample set did have a relatively high, but still negative, likelihood (lnL = -17.3) of being sampled in the Chukchi Sea based on mtDNA due to the relatively high occurrence of Hap#2 (47%) in that population (O’Corry-Crowe et al., 1997, 2010).

**Discussion**

The genetic analysis of beluga whales sampled in Yakutat Bay between 2002 and 2008 revealed that the whales in this glacial fjord system, recently estimated to number less than 20 individuals (Lucey et al., 2015), are unlikely to represent recent arrivals or seasonal visitors from the Cook Inlet population to the northwest. The Yakutat whales do not appear to experience substantial contemporary genetic exchange with the larger Cook Inlet population. Further, the genetic “re-sights” of the same individuals across years, when taken with recent sightings of beluga whales in all seasons, including juveniles and young-of-year calves (Hubbard et al., 1999; O’Corry-Crowe et al., 2002) and TEK of beluga whales in this region spanning 80 years (Lucey et al., 2015) indicate that the Yakutat beluga whales likely represent a small, resident reproductive group of whales on the southern edge of the species’ range.

High average relatedness among the Yakutat whales combined with higher inbreeding coefficients and lower heterozygosity and allelic (and haplotypic) diversity relative to the Cook Inlet population indicate that this group may be exhibiting negative genetic consequences of low abundance and genetic isolation.

The use of a broad suite of complementary methods of genetic analysis, including a power analysis and a new group-based assignment/exclusion method, revealed that genetic data from a relatively small number of samples was sufficient to address a number of key questions. Individual likelihood-based assignment methods identified Cook Inlet as the most likely origin of Yakutat whales when compared to other baseline populations. It does not necessarily follow, however, that the Yakutat whales were primarily resident in, or had recently immigrated from Cook Inlet. Rather, the low individual likelihoods of many Yakutat genotypes occurring in Cook Inlet (relative to Cook Inlet whales) suggest that Yakutat Bay whales have Cook Inlet ancestry but may not have been born in Cook Inlet. Estimates of relatedness and inbreeding revealed that the Yakutat sample set was not a random sample from the entire Gulf of Alaska and that preferential interbreeding among the Yakutat whales has led to high levels of relatedness among group members, and slightly higher average inbreeding within individual whales.

A number of widely different approaches found that the Yakutat whales, as a group, were genetically distinct from Cook Inlet. F statistics and homogeneity tests revealed substantial and statistically significant differentiation between Yakutat and Cook Inlet, STRUCTURE identified Yakutat as its own distinct genetic cluster, and GELATo excluded Yakutat as unlikely to occur in Cook Inlet or any other population studied. The close affinity of Yakutat to the Chukchi Sea in
the GELATo mtDNA analysis differed markedly from that for the multilocus microsatellite analysis. This is likely due to the demographic history of the maternally inherited mtDNA that has resulted in the wide distribution of Hap#2 across much of the species' range rather than contemporary dispersal between Yakutat and the Chukchi Sea, and cautions against relying on a single marker for assignment- or allocation-based tests.

The group assignment criteria of Paetkau et al. (1995, 2004), Rannala and Mountain (1997), and Baudouin and Lebrun (2000) as implemented in the GENECLASS2 (Piry et al., 2004) program all estimated the highest likelihoods of the Yakutat whales coming from Cook Inlet, each giving a rank score of 100% to this population. In all three methods the assignment criterion is simply the highest estimated likelihood (rank score) while the exclusion criterion is an arbitrary rank score value below which it is considered that population-of-origin is unlikely (Piry et al., 2004).

These criteria, however, do not provide any insight into whether the group of whales tested is in fact likely or unlikely to occur in each of the reference populations. To illustrate this, we tested subsets of whales similar in magnitude to that of the Yakutat test group (i.e., n=8) from the genetically isolated Cook Inlet population against the four remaining reference populations in Alaska and Northwest Canada. Earlier studies had found negligible levels of contemporary gene flow between Cook Inlet and these four populations and that all individuals tested here had the highest likelihoods of arising within Cook Inlet (O’Corry-Crowe et al.). And yet, in all individual and group assignment tests in GENECLASS2, the Cook Inlet whales were assigned to one of the other populations, some with rank scores > 80% (not shown). Similarly, using an exclusion criterion based on a rank score can only exclude the lowest scores (as long as they fall below the threshold, e.g., 0.05%) but cannot exclude all populations. This is a problem if the test group is from none of the reference populations.

The method we present (GELATo) overcomes both of these limitations. The assignment and exclusion criteria are based on the likelihood of the test group relative to randomly chosen groups of individuals (of similar magnitude) from each population. The tests are conducted on each reference population independently and rank scores are not relied upon. Strong overlap between the population distribution (based on resampled replicates of the original reference dataset) and test group distribution (based on observed differences between the test group and resampled reference population) indicates a high likelihood of the group coming from that population, while non-overlapping distributions excludes the group.

A further shortcoming of the exclusion methods in GENECLASS2 is that while the rank-score methods tended to under-exclude populations, the MCMC methods (Rannala and Mountain, 1997; Cornuet et al., 1999; Paetkau et al., 2004) tended to do the reverse, excessively exclude populations that were in fact the real origin of the test group. We determined this by running a series of tests in GELATo and GENECLASS2 where 100 “groups” of individuals (n=8) were randomly subsampled from their own population (e.g., Cook Inlet, Chukchi Sea). In all cases the former approach correctly assigned these groups to their population whereas the MCMC methods almost always excluded (probability < 0.05) groups numbering three or more individuals.

Our method, as with all other group methods, assumes random associations of individuals and random breeding among group members, which may not hold for small groups of beluga whales. Beluga whales are highly social, characterized by a wide variety of grouping patterns. If such grouping behavior is kin-based, it is possible that a small group of closely related

Figure 5.—Inferred membership in K=2 genetic clusters from a STRUCTURE analysis of all Gulf of Alaska beluga whale samples. The summary plot is for the analysis that allowed admixture and used the LOCPRIOR option. Each individual is represented by a single vertical line broken into K colored segments, with lengths proportional to membership in each of the K inferred clusters.
individuals recently dispersed from Cook Inlet would accentuate measured genetic differences between the source population and the splinter group, and we would thus erroneously exclude Cook Inlet.

Similar erroneous exclusions are conceivable if mating patterns differ greatly from random mating. We found no evidence of non-random mating within the Yakutat whales and a recent genetic analysis of beluga whale grouping patterns revealed that most grouping types other than cow-calf pairs are not comprised of close kin (O’Corry-Crowe et al.6). We conclude, therefore, that it is unlikely that colonizing groups are comprised solely of close kin or that breeding in these groups is highly skewed. So, while Cook Inlet is the likely original source of whales in Yakutat Bay, our test found that the group of whales we sampled in Yakutat is unlikely to have recently emigrated from Cook Inlet. Finally, while the new group-based assignment/exclusion method we present here has shown great promise, development is ongoing, including validation with simulated data and the exploration of other measures of genetic discreetness.

Figure 6.—The likelihood of the Yakutat group of beluga whales coming from the 5 population stocks of this species in Alaska and NW Canadian waters including Cook Inlet, using the program GELAto: A) mtDNA, and B) microsatellite loci. The null distribution of genetic distance ($F_{st\text{-null}}$) within each reference population centered on $F_{st} = 0$ (light grey bars), is compared to the observed distribution of genetic distance ($F_{st\text{-obs}}$) between the Yakutat sample set and the reference population, centered on $F_{st} \geq 0$ (dark grey bars). The distributions are generated from 1,000 re-samplings of the reference data set and associated estimation of genetic distance. The log likelihood of each $F_{st\text{-obs}}$ is also reported. A high degree of overlap among the null and observed distributions indicates a high likelihood the sampled group came from that population stock.
Perspectives

This study, when viewed with growing evidence from ecological and traditional knowledge studies, established that a small breeding group of beluga whales is resident in Yakutat Bay at the warm edge of the species’ range. This group has likely persisted in this region for at least several decades spanning multiple generations, and occupies a habitat unique to beluga whales in the North Pacific. Recurring immigration from Cook Inlet appears limited and dispersal to Cook Inlet from Yakutat Bay may be rare.

That these whales represent a viable functioning population is less clear. The number of breeding adults may be less than 10 individuals. While the adaptive potential of this group can influence its chances of survival in the long term, it is its small current size and the attendant demographic and genetic effects on fitness and survival that will dictate its fate in the short term.

Nevertheless, that such a small number of whales have survived on the southern edge of the species’ range is encouraging. Their persistence may reflect an innate resilience of this species in marginal habitats or locations that may increase their prospects when dealing with predicted environmental and climate change. Genetic and behavioral aspects of the social nature of this species may enhance the chances of small pioneering or relic groups of beluga whales surviving on the edge of the species’ range.

The management of small groupings on the species’ range limit presents both unique challenges and opportunities. On the one hand, the persistence of peripheral groups may be short lived, their direct role in the maintenance or recovery of central populations over ecological timescales may be limited and their contributions to the evolutionary potential of species on evolutionary timescales may be minimal. It can be debated, therefore, whether they qualify for inclusion within Distinct Population Segments (Waples, 1991; USFWS and NMFS, 1996) under the U.S. Endangered Species Act.

On the other hand, groups, no matter how small, at the geographic limit of a species’ range define the boundary of management units where the objective is to maintain all populations as functioning elements of their ecosystem, as in Population Stocks under the Marine Mammal Protection Act (Wade and Angliss, 1997). They also provide rare opportunities to investigate how species respond to environmental perturbations, including climate change, and cope with the challenges of small population size and isolation.

In the case of the belugas of Yakutat Bay, their survival thus far has likely been aided by their obscurity and the relative pristineness of their environment. Further, the Tlingit of the Lost Coast have no tradition of hunting the whale and view the beluga as an integral part of an ecosystem that should be preserved. We recommend that further research be conducted on the population dynamics, ecology, genetics, and movement patterns of the Yakutat Bay whales. This study was a collaborative effort with the communities of this region and it is only through continued community-based research will further insight be gained and the chances of survival of this group be maximized.

Acknowledgments

Funding for this research was provided by research grants from the U.S. Marine Mammal Commission and the National Marine Fisheries Service. Biopsy sampling was conducted under NOAA permit no. 782-1719-03 issued to John Bengston, Director, National Marine Mammal Laboratory, Seattle, Washington. The establishment of a field camp and associated activities in Disenchantment and Yakutat Bays were conducted under a Scientific Research and Collecting Permit #WRST-2005-SCI-003 from the U.S. Department of the Interior, National Park Service. We thank NOAA’s NMFS Southwest Fisheries Science Center, La Jolla, California for providing lab facilities and Carolina Bonin and Amy Frey for assistance with lab work. Thanks to Rudy, Jeremiah, and Jonathon Pavlik, our field guides in Yakutat, for their assistance in so many ways. Thanks to John Jansen, Bert Adams Jr., Carolina Bonin, Eileen Henniger, Amy Frey, Aviva Nessler, Sam Williams, Ray Sensmier, and Bert Adams Sr., for their assistance in the field and to Kaja Brix and Brad Smith of the NMFS Office of Protected Resources, Alaska, and Rod Hobbs of the National Marine Mammal Laboratory for all their support. Thanks to Nelson Beaman for the Go-Random Access script. A special word of gratitude to Barbara Taylor, Phil Clapham, and Brad Smith, their comments greatly improved this paper. HBOI contribution #1973.


