Sea Duck Joint Venture
Annual Project Summary for Endorsed Projects
FY05 – (October 1, 2004 to September 30, 2005)

Project Title: (SDJV #46) Population structure and annual survival estimation of female Black Scoters using genetic tagging

Principal Investigators: Sandra Talbot, Jason Schamber, Paul Flint, U.S. Geological Survey, 1011 E. Tudor Rd., Anchorage, Alaska, 99503; Matthew Perry, USGS, Patuxent Wildlife Research Center, 11410 American Holly Drive, Laurel, MD 20708-4015; Fred J. Broerman, USFWS, Yukon Delta NWR, PO Box 346, Bethel, Alaska 99559; sandy_talbot@usgs.gov

Project Description: Identify individuals through genetic techniques and estimate annual female survival using a mark-recapture approach. Use the same markers to examine population delineation in these species.

Objectives or Hypothesis:

1. Obtain an estimate of annual female survival based on collection and genetic analysis of contour feathers from nesting Black Scoter females.

2. Characterize population structure of Black Scoters at Aropuk Lake.

This project fills the information gap and addresses high-priority needs identified by the Sea Duck Joint Venture Strategic Plan 2001 – 2006, by using population genetic analyses and genetic tagging in a mark-recapture study to define populations, estimate population size and trends, and understand population dynamics of a Black Scoter population.

Results and Discussion:

Locus Selection

We obtained blood samples from females trapped at the nest during 2001 – 2003, and feather samples from 113 nests from Aropuk Lake across three years (2002 – 2004). We screened 70 microsatellite loci and typed 29 individuals from Aropuk Lake at 12 variable loci, to create a baseline dataset (data not shown). Following informal exploration of data from these loci, we selected 8 (Table 1) for further consideration, based on published guidelines for choosing markers for genetic tagging studies (Paetkau 2000, 2003). These 8 loci were initially selected because they demonstrated the highest mean expected heterozygosity (H_E), high resolution, with few or no 1 base-pair (bp) insertions, no initial indication of null alleles or allelic dropout, and relatively high variability (Talbot et al. 2004).

We used these 8 loci to collect data from DNA extracted from feather samples collected from nest bowls during 2002 – 2004. Seven of the 8 loci selected were sufficiently robust to yield high quality results from the feather samples. One locus, Sfiμ11 (Libants et al. 1999) was reliable if extractions were from blood tissue, but less reliable when extracted from feather

tissue. We therefore dropped this locus from further analysis in the genetic-tagging portion of this study. In the place of that locus, we added another (Smo4, Paulus & Tiedemann 2003), which demonstrated very high levels of variation (Table 1), no 1 – bp insertions, and no indication of allelic dropout in amplifications of DNA extracted from feathers.

After testing the loci for linkage disequilibrium and conformation to Hardy-Weinberg equilibrium (HWE) expectations, we used the statistical program, GIMLET 1.3.2 (Valieré 2002), to generate $P_{(ID\mathit{obs})}$ and $P_{(ID\mathit{sib})}$ values yielded by the multilocus genotype to determine the probability of detecting individuals within this population. $P_{(ID\mathit{obs})}$ is the probability at which another individual with the same genotype would be observed, given the sample frequency of the alleles observed at those loci, within the population. $P_{(ID\mathit{sib})}$ estimates the probability of observing identical multilocus genotypes between two individuals sampled from a population comprised of first-order relatives (e.g., between siblings or parent-offspring). General guidelines for genetic tagging studies suggest using a suite of markers that achieve a reasonably low $P_{(ID)}$ bounded between 0.01 and 0.0001; $P_{(ID\mathit{sib})}$ provides a conservative upper bound on this estimate (Waits et al. 2001).

Table 1. Measures of genetic variability and probability of identity [$P_{(ID)}$] values for 8 microsatellite loci characterized from 29 Black Scoter females breeding at Aropuk Lake, Yukon-Kuskokwim Delta, Alaska. Loci are ranked according to $P_{(ID)}$ value. $P_{(shadow)}$ is the probability that two individuals within the Aropuk Lake breeding population will share the same 8-locus genotype. H_E = expected heterozygosity; *A* = number of alleles observed.

Locus	H_E	*A*	$P_{(ID\mathit{obs})}$	$P_{(ID\mathit{sib})}$	Locus Source
Smo4	0.94	20	0.0039	0.0292	Paulus & Tiedemann 2003
Aph07	0.89	12	0.0274	0.3185	Maak et al. 2000
Bcaμ11	0.79	9	0.0827	0.3809	Buchholtz et al. 1998
Aph04	0.68	4	0.1778	0.4593	Maak et al. 2003
CRG	0.65	3	0.2068	0.4821	A. Baker, unpublished
Aph18	0.63	4	0.2157	0.4944	Maak et al. 2003
Smo7	0.50	2	0.3837	0.6042	Paulus & Tiedemann 2003
Smo11	0.39	4	0.4064	0.6598	Paulus & Tiedemann 2003
Mean	**0.68**	**7.3**	-	-	
Multilocus product			2.121×10^{-8}	1.990×10^{-3}	
$P_{(shadow)}$			1/47,147,572	1/503	

We observed no deviations from HWE, and no significant linkage disequilibrium, among the 8 loci within the Aropuk population. The observation of significant deviation from HWE in Aph18 observed in preliminary analyses (Talbot et al. 2004) was not observed after the completion and proofing of the dataset, so was retained for this study.

Probability of identity values indicate that an 8-locus genotype comprised of the loci listed in Table 1 are sufficient to distinguish a single individual among approximately 47,147,570 individuals drawn from the Aropuk Lake breeding population [multilocus $P_{(ID\mathit{obs})} = 2.121 \times 10^{-8}$], and among approximately 500 single-order relatives [multilocus $P_{(ID\mathit{sib})} = 1.990 \times 10^{-3}$] (Table 1).

Since the Aropuk Lake population does not likely exceed 300 individuals (Schamber & Flint, unpublished data), a multilocus genotype comprised of these 8 loci clearly has adequate statistical power for use in individual identification, even if the population were comprised mostly of first-order relatives. This marker set adheres to published guidelines recommending genetic tagging markers achieve a $P_{(ID)}$ bounded between 0.01 and 0.0001(Waits et al. 2001).

Renesting

Multilocus genotypes were obtained from DNA extracted from feathers from 36 nests in 2002, 33 nests in 2003, and 44 nests in 2004. We observed 14 instances in which feathers from more than one individual were recovered from a single nest (data not shown). In these cases, a single individual was used to represent that nest for the purposes of the renesting study. Subsequent laboratory analyses are planned to obtain complete genotypes for the second individual for each of these 14 nests.

In 9 instances, individual Black Scoters were identified nesting in at least one subsequent year, and in one instance, an individual renested within the same year (Table 2). We were also able to attribute one instance of within-season mortality to a nesting female (Table 2). These data were used to obtain an estimate of survival among female Black Scoters nesting at Aropuk Lake.

Table 2. Individual Black Scoters observed in subsequent years, based on genetic tagging studies. A total of 9 between-year observations were made among 113 total nests (see text).

Year/ Sample ID	Year/ Nest Sample ID	Year/ NestSample ID	Year/ NestSample ID	Comments
2001/287	2002/JLC10	2003/KRS019		First observed in 2001; reobserved (nested) in 2002 and 2003
2001/313	2002/CWP017			First observed in 2001; reobserved (nested) in 2002
2001/314			2004/EKJ013	First observed in 2001; reobserved (nested) in 2004
2002/310	2002/PLF002	2003/KRS011	2004/JTP018	First observed in 2002; reobserved (nested) in 2003 and 2004
	2002/JLC07		2004/TRJ030	First observed in 2002; reobserved (nested) in 2004
	2002.JLC08	2003/TRJ017	2004/JLS041	First observed in 2002; reobserved (nested) in 2003 and 2004
	2002/JLS12	2003/JLS31		First observed in 2002; reobserved (nested) in 2003
		2003/HDW007	2004/JTP012	First observed in 2003; reobserved (nested) in 2004
		2003/KRS015	2004/A04F1	First observed in 2003; reobserved (nested) in 2004
		2003/JLS022	2003/TRJ012	Likely renest within same year
04A04UNK			2004/A04F2	04A04UNK was from the wing of a dead bird

Estimate of Survival

We generated a capture-recapture matrix (Figure 1) using the individuals determined via multilocus genotypes derived from contour feathers. Sample sizes were small with only 9 individuals being documented in more than 1 year and only 3 individuals being documented in more than 2 years. We performed a Cormack Jolly-Seber live captures analysis in Program

MARK and restricted model structure such that annual survival rate was constant and recapture rate varied across time ($\Phi.p_t$). While this model was estimable, these data appear to be ill-conditioned. That is, recapture rate for the last year is estimated as 1.0, which is clearly incorrect.

1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	1	1	0
1	1	0	0
1	0	0	1
0	1	0	0
0	1	0	0
0	1	0	0
0	1	0	0
0	1	0	0
0	1	1	1
0	1	0	1
0	1	1	1
0	1	1	0
0	0	1	1
0	0	1	1

Figure 1. Capture-recapture matrix using individuals determined based on multilocus-genotypes derived from DNA extracted from contour feathers.

Because of the inherent negative association between survival and recapture rate in the model likelihood, this results in underestimation of the survival rate. In this case, survival is estimated as 0.46 (SE 0.08). Therefore, we re-examined the same model ($\Phi.p_{t)}$, but constrained the final recapture rate to the mean of the previous years. The model estimate for survival increased to 0.63 (SE 0.11). Thus it appears that ill-conditioning of these data results in underestimation of the survival rate by at least 17 percentage points. Fully accepting that sample sizes were very small, we realize that the point estimates are questionable. However, we are estimating apparent survival which is the product of survival and fidelity. Given the reproductive output of Black Scoters (Flint & Schamber, unpublished data), annual survival would be expected to exceed 90%. Therefore, we interpret the relatively low estimates of survival to indicate that permanent emigration off of our study area is likely. Levels of emigration can be inferred indirectly, using estimates of gene flow that occurs among populations over many generations. This is possible, given adequate sampling of populations and sufficient levels of population differentiation. However, based on preliminary comparative analysis suggesting low levels of population differentiation (see below), we suggest in addition that future studies examine emigration directly.

Population Comparisons

Genetic Diversity. We determined levels of genetic diversity across years (2001 – 2004) within the Aropuk population, based on 8 microsatellite loci extracted from both blood and feather samples (Table 3). We also compared levels of genetic diversity using 9 microsatellite loci (with Sfiμ11 added to the suite of markers listed in Table 1) in the Aropuk population (based on blood samples only) with values obtained from samples representing geographically distinct populations aggregated during different seasonal periods: Nelson Lagoon on the Alaska Peninsula (collected between January and April 2004), Prince William Sound in the Gulf of Alaska (collected between May and June 2004), and coastal British Columbia (collected in December 2003).

Levels of genetic diversity were concordant across years within Aropuk Lake, and across populations or aggregates (Table 3). Among populations and aggregates, average expected heterozygosities (H_E), based on 9 microsatellite loci, ranged from 0.617 in wintering populations off the coast of British Columbia, to 0.657 in Nelson Lagoon. Average allelic richness (rg), which corrects A (average number of alleles) for disparity in sample size, was also concordant across populations and ranged from 3.99 to 4.05 (Table 3). We observed no signature of inbreeding (*F*) within any of the populations (Table 3). Values for H_O and rg among years for the Aropuk Lake breeding population were concordant (Table 3)

Estimates of average relatedness within populations were assessed using Queller and Goodnight's R-values (Queller and Goodnight 1989) and obtained using the program Identix (Belkhir et al. 2002). Average relatedness (R) was concordant and close to zero within all populations, with the exception of the aggregation in Prince William Sound. That population demonstrates a substantially negative average R (Table 3). Substantially negative R-values suggest the population assayed is comprised of individuals that are less related than expected in a randomly breeding population (Queller and Goodnight 1989). This population is represented by

very low samples sizes and the observed deviation in this population may be due to inadequate sampling, or admixture. Average relatedness within year-classes for the Aropuk breeding population was concordant across years and close to zero (Table 3).

Table 3. Genetic diversity at 9 bi-parentally-inherited microsatellite loci, for Aropuk Lake (pooled across years; blood samples only), Prince William Sound, British Columbia, and Nelson Lagoon, and for 8 bi-parentally-inherited microsatellite loci (only Aropuk Lake, 2001 – 2004, blood and feather samples).

Population	N[1]	H_E[2]	H_O[3]	A[4]	rg[5]	*F*[6]	R[7]	$R_{(var)}$[8]
Aropuk Lake (pooled)	29	0.641	0.630	6.44	3.93	0.017 *ns*	-0.008	0.052
(2001 only)	21	0.643	0.601	5.75	5.75	0.065 *ns*	-0.052	0.057
(2002 only)	34	0.667	0.670	7.75	6.87	-0.004 *ns*	-0.031	0.055
(2003 only)	26	0.652	0.599	6.75	6.41	-0.011 *ns*	-0.040	0.050
(2004 only)	38	0.621	0.628	7.00	6.49	0.031 *ns*	-0.024	0.050
Prince William Sound	6	0.650	0.630	4.00	4.00	0.031 *ns*	-0.191	0.063
British Columbia	34	0.617	0.586	6.44	3.99	0.050 *ns*	-0.030	0.051
Nelson Lagoon	17	0.657	0.654	5.56	4.05	0.005 *ns*	-0.060	0.044

[1]number of individuals sampled
[2]unbiased expected heterozygosity (Nei 1987; eq. 7.39, pg. 164)
[3]observed heterozygosity
[4]average number of alleles at 9 microsatellite loci
[5]allelic richness (El Mousadik and Petit 1996)
[6]inbreeding coefficient (Wright 1951). *ns* = not significant (significance of *F* was tested using methods of Li and Horovitz (1953)
[7]Relatedness values according to Queller and Goodnight (1989)
[8]Variance of R, tested using Identix (Belkhir et al. 2002)

Genetic Differentiation. Significance of spatial variation among populations (the breeding population at Aropuk, and winter and spring aggregates, as well as among years for the Aropuk population) was assessed using F-statistics (Weir and Cockerham 1984). These measures can be viewed simply as variance components that describe the apportionment of allelic variance among individuals within (F_{IS}) and among (F_{ST}) populations. Values of F_{ST} are summary statistics ranging essentially from 0 to 1 that describe the extent of spatial variation among populations or population groups. A value of 1 at a specific locus would imply that all populations are fixed for different alleles (i.e., the total variance at that locus is segregating among populations). A value of 0 implies all populations share the same alleles in equal frequency (panmixia). Overall (multilocus) microsatellite estimates of F_{ST} variance, θ, were obtained using FSTAT (Ver. 2.9.3, Goudet 2001). Estimates of interpopulational variance (θ) were derived using the program ARLEQUIN 2.0 (Schneider et al. 2000). Significance of θ values were based on random permutation tests (n = 1,000), whereby alleles were randomly permuted between populations. A significant value of θ implies that a significant portion of the total genomic variation across loci is partitioned among populations. We also tested for significance of heterogeneity of microsatellite alleles between populations, as described in Raymond and Rousset (1995), using ARLEQUIN. For significance testing, all α-values were set at 0.05 and, where appropriate, adjusted using Bonferroni procedures (Rice 1995).

Overall interpopulational variance among breeding populations and seasonal aggregates was not significantly different from zero (θ = -0.002, p > 0.05). Similarly, θ across years within Aropuk was not significantly different from zero (θ = 0.001 p > 0.05). Lack of population differentiation for both among-population and between-years is further supported by G-statistics analyses of Raymond and Rousset (1995) (data not shown). For all population pairwise comparisons, and among years for Aropuk breeding populations, we also failed to reject the null hypothesis of panmixia (p > 0.05, data not shown for each pairwise comparison).

Lack of differentiation at the inter-annual level among Black Scoters nesting at Aropuk differs from that reported for nesting Steller's Eider on the North Slope (Deering 2002, but see Pearce & Talbot 2004). The apparent absence of population differentiation at various spatial and temporal scales among different aggregations of Black Scoters in western North America (breeding populations and winter and spring aggregations) suggests the species, like King and Steller's eiders (Pearce et al. 2004, 2005), may be comprised of populations that are not demographically distinct at moderate geographic distances within its North American range. Lack of observed interpopulational structuring can result from high levels of dispersal of juveniles from natal sites as well as of adults from breeding sites. Although instances of breeding site fidelity was demonstrated in 9 instances for females within the Aropuk breeding population, fidelity at this level may be insufficient to allow differentiation among populations. Analysis of other breeding populations is necessary to determine whether breeding aggregations contribute to significant differentiation within the species. In addition, analysis of mitochondrial DNA in concert with the biparentally-inherited microsatellite data will help clarify whether the observed panmixia among aggregates is largely due to male-mediated gene flow.

Project Status: The work described above support both objectives of obtaining an estimate of female survival using genetic tagging for individual identification and description of the population genetics characteristics of the Black Scoter population breeding at Aropuk Lake. Although we have been able to compare wintering and breeding populations for the western portion of the species' range, we were unable to meet the objective of comparing eastern and western populations; this will be accomplished upon analysis of samples collected from New Brunswick, Canada or elsewhere on the eastern coast of North America when they become available.

Literature Cited

Belkhir K, Castric V, Bonhomme F (2002) IDENTIX, a software to test for relatedness in a population using permutation methods. *Molecular Ecology Notes*, **2**, 611-614

Buchholtz W, Pearce JM, Pierson BJ, Scribner KT (1998) Dinucleotide repeat polymorphisms in waterfowl (family Anatidae): characterization of a sex-linked (z-specific) and 14 bi-parentally inherited loci. *Animal Genetics*, **29**, 323-325.

Deering MM (2002) Steller's Eider nesting in Barrow, Alaska: Philopatry, individual nest site selection, and association with nesting associated with pomerine jaegers. PhD dissertation, University of California Davis

El Mousadik A, RJ Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic of Morocco. *Theoretical and Applied Genetics,* **92**, 832-839.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html.

Li CC, Horovitz DG (1953) Some methods for estimating the inbreeding coefficient. *American Journal of Human Genetics,* **5**,107-117.

Libants S, Oswald K, Olle E, Scribner K (1999) Isolation of microsatellite loci from the spectacled eider: GenBank Accession Number AF180501. Available (September 28, 2004) at: http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=nucleotide&val=12734691

Maak S, Neumann K, von Langerken G, Gattermann R (2000) First seven microsatellites developed for the Peking Duck (*Anas platyrhynchos*). *Animal Genetics*, **31**, 228-241.

Maak S, Wimmers K, Weigend S, Neumann K (2003) Isolation and characterization of microsatellites in the Peking Duck (*Anas platyrhynchos*) and their application in other waterfowl studies. *Molecular Ecology Notes*, **3**, 224-227.

Nei M (1987) *Molecular Evolutionary Genetics*. New York: Columbia University Press.

Paetkau D (2000) Genotyping errors in DNA-based inventories: a protocol that controls them, complete with empirical validation. Pages 12-13 *In*: Managing for Bears in Forested Environments. Proceedings of the Columbia Mountains Institute of Applied Ecology Workshop, October 17-19, 2000, Revelstoke, British Columbia, Canada.

Paetkau D (2003) An empirical exploration of data quality in DNA-based population inventories. *Molecular Ecology*, **12**, 1375-1387.

Paulus KB, Tiedemann R (2003) Ten polymorphic autosomal microsatellite loci for the Eider duck *Somateria mollissima* and their cross-species applicability among waterfowl species (Anatidae). *Molecular Ecology Notes*, **3**, 250-252.

Pearce JM, Talbot SL, Petersen MR, Rearick JR (2005) Limited genetic differentiation among breeding, molting, and wintering groups of the threatened Steller's eider: the role of historic and contemporary factors. *Conservation Genetics*, in press.

Pearce JM, Talbot SL (2004) Enhancing genetic studies of Steller's eiders. A final report to the Alaska SeaLife Center and the Eider Recovery Team, Grant No. R3604-01. Unpublished report, Alaska Science Center, U. S. Geological Survey, Anchorage, AK. September 30, 2004. 46pp.

Pearce JM, Talbot SL, Pierson BJ, Petersen MR, Scribner KT, Dickson DL, Mosbech A (2004) Lack of spatial genetic structure among nesting and wintering king eiders. *Condor*, **106**, 229-240.

Raymond M, Rousset F (1995) GENEPOP (Ver 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**,248-249.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**,223-225.

Schneider S, Roessli D, Excoffier L (2000) Arlequin ver. 2.000: a software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland. Available: http://anthropologie.unige.ch/arlequin.

Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers. *Evolution* **43**:258-275.

Valieré N (2002) GIMLET: a computer program for analyzing genetic individual identification data. *Molecular Ecology Notes*, **2**, 377-379.

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*, **10**, 249-256.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**,1358-1370

Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323-354