Microsatellite Analysis of North American Wolverines

Christopher J. Kyle
Department of Biological Sciences, University of Alberta
Edmonton, AB, T6G 2E9, Canada
ckyle@ualberta.ca

Curtis Strobeck
Department of Biological Sciences, University of Alberta
Edmonton, AB, T6G 2E9, Canada

ABSTRACT

Elucidating the population genetic structure of a species gives us insight into the levels of gene flow and connectivity between geographic regions. These are important considerations when trying to manage a threatened or Blue-listed species, such as wolverines, as we are then provided with information on how far and where animals may be dispersing and where population boundaries occur. Population genetic structure can be estimated using fast evolving, highly variable molecular markers such as microsatellites. These molecular markers have a high mutation rate, can be analyzed from a very small amount of DNA using polymerase chain reaction, and provide reproducible results. Using microsatellites, we have addressed questions of genetic variation within populations with estimates of population heterozygosity, mean number of alleles, and probability of identity. Gene flow among populations was also estimated via an assignment test and genetic distance measures. In this preliminary study, 4 geographic regions (2 from the Northwest Territories and 2 from British Columbia) were sampled and analyzed at 7 microsatellite loci. As expected, more gene flow occurs between proximal regions than those separated by large geographic distances. Our preliminary results also suggest that there is a directional trend of gene flow from the Northwest Territories and Revelstoke to northern British Columbia, whereas the Revelstoke region seems to be more genetically isolated from the other regions analyzed here.

Key words: gene flow, genetic structure, *Gulo gulo*, microsatellites, population, wolverine.

The population genetic structure of North American wolverines (*Gulo gulo*) is unknown to date. With increasing concern about the status of this forest carnivore and how it might be managed effectively, this information is needed to determine what constitutes a population of wolverines and what levels of effective migration occur between these geographic areas. Elucidating the population genetic structure of a species can also provide estimates of a species' vagility, which can then be correlated to ecological estimates of dispersal. Furthermore, successful reintroduction programs, ideally, reintroduce animals that come from a similar genetic background to those that once existed in that area. Hence, determining the population genetic structure may help to establish which populations are best suited for reintroduction into a depauperate area.

Molecular markers, such as microsatellites, can provide estimates of genetic distance between populations which reflect the effective migration between populations, and not simply movement between them. Direct observations of individuals and migration rates between separate geographic areas are often very difficult and costly, if not impossible for some species when attempting to estimate gene flow. Furthermore, direct observations of migration can be misleading if migrants do not contribute to the gene pool of the population they have moved into. When the population genetic structure is related to the levels of gene flow between geographic areas it enables us to determine how, and if, individuals are moving between populations. From this we can also determine population boundaries for a particular species (if they exist). Populations can be identified by the genetic distinctiveness of individuals taken from different geographic areas which may not always be evident from direct observations.

Measures of migration and population boundaries are important factors in managing a species. First, we can use them to elucidate whether a management unit is a distinct population or simply part of a larger panmictic population. Second, we can determine if a harvested population is being replaced solely from within by the level of fecundity in that population or if replacement from other populations plays a role in its numbers. Levels of genetic variation within populations can also help identify which populations (if any) act as source populations for smaller and more isolated regions. This becomes an important factor when a species' habitat becomes fragmented, separating source populations from sink populations. When this occurs, concerns for inbreeding may arise.

Gene flow, and hence the population genetic structure of a species, are determined by the life history characteristics of that species. Those animals with larger home ranges and elevated levels of vagility are likely to have less population

L. M. Darling, editor. 2000. Proceedings of a Conference on the Biology and Management of Species and Habitats at Risk, Kamloops, B.C., 15 - 19 Feb.,1999. Volume Two.
B.C. Ministry of Environment, Lands and Parks, Victoria, B.C. and University College of the Cariboo, Kamloops, B.C. 520pp.

genetic structure that those with smaller home ranges and limited vagility. Other life history traits such as fecundity, longevity, and density are also important in determining population genetic structure.

Wolverines have relatively large home ranges and can disperse across large distances in a short time, especially as juveniles (Gardner 1985). Furthermore, *Gulo gulo* is thought be relatively diffuse throughout its distribution (Banci 1994), especially in the extreme southern regions of this animal's distribution. These characteristics of high vagility, large home ranges, and relatively high longevity might be expected to lower the levels of population subdivision in this species. Yet, low population numbers in the southern regions of this species' distribution and high levels of juvenile mortality as they disperse may lead to very isolated populations. If genetic variation does exist among populations, high resolution (fast evolving) neutral genetic markers will be needed to detect it. For this reason we have chosen to use hypervariable polymorphic microsatellites as in other mammalian studies of species with high vagility, such as the polar bear (*Ursus maritimus*). Clear genetic differentiation was found in these animals using microsatellites despite the known long range movements of the species and that little variation was detected using other methods (Paetkau et al. 1995).

STUDY AREAS AND METHODS

Seven polymorphic loci were used to examine the population genetic structure of *Gulo gulo* from 4 geographic regions including: the Kugluktuk (N = 45) and Bay Chimo (N = 40) regions in the Northwest Territories (NWT) and the Revelstoke (N = 48) and Williston Lake (northern B.C.; N = 37) regions in British Columbia (Fig. 1). Samples were provided as frozen tissue, ear plugs, and hair roots. DNA was extracted using a QIAamp® Tissue Extraction Kit. The 7 loci used to amplify the microsatellites were developed by Davis and Strobeck (1998) in badgers (BA-1 and BA-4) and wolverines (GG-3, GG-4, GG-7) and by Duffy et al. (1998) in wolverines (Ggu 101 and Ggu 234). DNA at all loci was amplified using the same protocol as in Davis and Strobeck (1998). The DNA fragments were then visualized using an ABI Prism™ 377 DNA sequencer. DNA fragments were analyzed using the programs GeneScan™ Analysis 2.02 and Genotyper® 2.0.

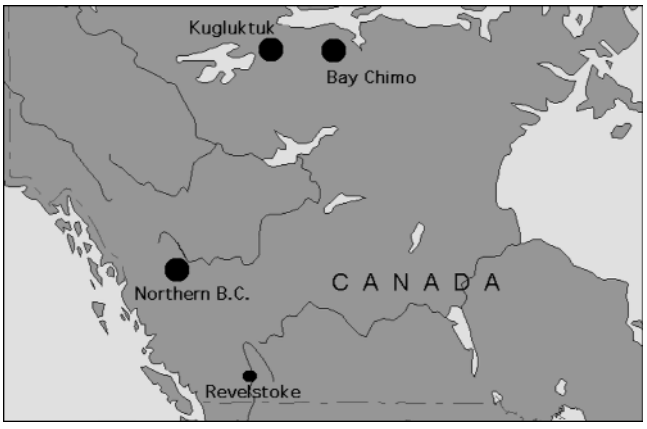

Figure 1. Map of sampled regions. Dot size reflects relative size of sampled regions.

Data Analysis

To assess the relative genetic variation of each of the 4 geographic areas sampled the allele frequencies for each loci were first determined. From these frequencies the mean number of alleles, average heterozygosity, and overall unbiased probabilities of identity (Paetkau et al. 1998) were calculated. The genetic distances among the 4 geographic areas were estimated using 2 measures: 1) both Nei's standard genetic distance, D_S, which is calculated from genotype frequencies (Nei 1972); and 2) the genotype likelihood ratio, D_{LR}, calculated from genotype probabilities (Paetkau et al. 1997). An unrooted tree of the D_{LR} values was then created using PHYLIP 3.572 (Felsenstein 1995). The assignment test, which determines the probability of a genotype occurring in the region from which it was sampled and the probabilities of it occurring in the regions it is being compared with (Paetkau et al. 1995), was run for all 4 geographic areas. The programs used to calculate the assignments and the genetic distance measure (D_{LR}) were designed by J. Brzustowski and are available at:
http://www.biology.ualberta.ca/jbrzusto/alpha/Doh.html.

The assignment test (Paetkau et al. 1995), also found on the website by J. Brzustowski, was run for all populations. This program determines both the probability of a genotype occurring in the region from which it was sampled and the probability of it occurring in the regions it is being compared with. It then assigns individuals to the population for which that individuals genotype has the highest probability of occurring. The assignment test was run several times using the following options: replacement of gene frequency values of 0 with 0.01, resampling of the populations from within each gene pool assuming Hardy-Weinburg equilibrium (HWE), and resampling of the populations by combining the gene pools. The option of resampling the data from each gene pool lends a significance value to the assignment of individuals to particular populations, and may help identify potential migrants between populations. The randomizations of the data indicate the percentage (of the number of iterations used) of when more individuals were assigned to a population than the original assignment. Therefore, if a small percentage is obtained from the randomization then it is unlikely that all of the misassignments of individuals are due simply to

Table 1. Sample size, mean number of alleles, mean heterozygosity, and overall unbiased probability of identity for each of the 4 geographic sampling regions.

Population	N	Mean no. of alleles	Mean heterozygosity	Prob. of identity[a]
Revelstoke	48	4.27	57.92	1/42,407
Northern B.C.	37	4.57	61.79	1/254,373
Bay Chimo	40	4.71	62.14	1/157,299
Kugluktuk	45	4.71	63.99	1/302,836

[a] $\sum_i p_i^4 + \sum_i \sum_{j>i} (2p_i p_j)^2$

chance; some of the misassigned individuals are most likely migrants from the population to which they were misassigned into the one from which they were sampled. When the option of combining the gene pools is used, the value obtained suggests how heterogeneous the populations are. If most of the randomizations of the populations sets have fewer self-assignments from a population to itself, then this is evidence for population structure and not simply a reflection of regional sampling efforts.

We chose the following options: 1) replacement of allele frequency with a value of 0.01 instead of 0; 2) resampling (10,000 times) of the populations from within each gene pool assuming HWE; and 3) resampling (10,000 times) of the populations by combining the gene pools. The second option lends a significance to the individual assignments. The third option indicates that misassignments are due to "true" migration between the 2 regions compared, or simply that the probability of the genotype occurring in either region is low. A high value from the resampling results for the individuals correctly assigned implies that any individuals misassigned are truly migrants between those two regions. Conversely, a low value indicates that the origin of the misassigned individuals has no significance.

RESULTS

Four geographic areas were sampled in this preliminary study (Fig. 1). The relative measures of genetic variation among the sampled regions are summarized in Table 1. For all 3 measures (mean number of alleles, mean heterozygosity, and overall unbiased probability of identity), the Revelstoke region displayed the least variation, whereas the Kugluktuk region displayed the most. The mean number of alleles (ranging from 4.57 to 4.71), and the mean heterozygosity (ranging from 57.92 to 63.99%) did not vary greatly amongst these 4 regions. The unbiased probabilities of identity estimates, however, did have a greater range with values from Revelstoke having a probability of 1/42,407 to Kugluktuk having a value of 1/302,836.

Two estimates of genetic distance were calculated, D_{LR} and D_S (Table 2). The estimates show that Revelstoke is genetically more similar to the northern B.C. area (D_{LR} = 0.544, D_S = 0.0565) than the Bay Chimo or Kugluktuk areas from which it is roughly generically equidistant (D_{LR} = 0.922, D_S = 0.0806 and D_{LR} = 1.09, D_S = 0.0922, respectively). The northern B.C. area is roughly genetically equidistant from all 3 other areas with values of D_{LR} ranging from 0.534 to 0.582 and D_S ranging from 0.0422 to 0.0661 The 2 NWT areas are the mostly closely associated areas with a D_{LR} value of 0.240. An unrooted tree of the D_{LR} values (Fig. 2) reveals the relationship of the aforementioned genetic distances represented by the length of the tree branches.

The randomizations of the assignment test were run in 2 ways: 1) by resampling the regions from each gene pool assuming HWE, and 2) by combining all the gene pools assuming HWE (Table 3). The Revelstoke area had the fewest individuals assigned to other regions with 37/48 individuals assigned to itself and 5, 4, and 2 assigned to northern B.C., Bay Chimo, and Kugluktuk, respectively. The northern B.C. region had 14/38 individuals assigned to itself with 10 assigned to Revelstoke, 7 assigned to Bay Chimo, and 6 assigned

Table 2. Genetic distance estimates.

D_{LR}

	Revelstoke	Northern B.C.	Bay Chimo	Kugluktuk
Revelstoke	0			
Northern B.C.	0.5440033	0		
Bay Chimo	0.9218481	0.58198035	0	
Kugluktuk	1.0903602	0.53401995	0.23970568	0

Nei's D_S

	Revelstoke	Northern B.C.	Bay Chimo	Kugluktuk
Revelstoke	0			
Northern B.C.	0.05646155	0		
Bay Chimo	0.08063514	0.06607150	0	
Kugluktuk	0.09221709	0.04246151	0.04215205	0

Table 3. Summary of assignment test results and randomizations of genotype frequencies.

Individual assignments
From: To:

	Revelstoke	Northern B.C.	Bay Chimo	Kugluktuk
Revelstoke	37	5	4	2
Northern B.C.	10	14	7	6
Bay Chimo	2	7	20	11
Kugluktuk	4	3	14	24

Randomizations from within each gene pool

	Revelstoke	Northern B.C.	Bay Chimo	Kugluktuk
Revelstoke	3,154	7,045	6,854	7,503
Northern B.C.	1,050	9,970	1,284	2,350
Bay Chimo	9,488	986	9,252	1,228
Kugluktuk	5,528	9,061	387	8,671

Randomizations from the combination of all gene pools

	Revelstoke	Northern B.C.	Bay Chimo	Kugluktuk
Revelstoke	0	9,972	9,992	10,000
Northern B.C.	945	4,510	8,603	9,346
Bay Chimo	9,997	9,087	27	4,161
Kugluktuk	9,985	9,997	2,175	6

to Kugluktuk. Bay Chimo had 20/40 individuals assigned to itself, 2 to Revelstoke, 7 to northern B.C., and 11 to Kugluktuk. The Kugluktuk region had 24 animals assigned to itself, 4 to Revelstoke, 3 to northern B.C., and 14 to Bay Chimo.

CONCLUSIONS

The relative measures of genetic variation (mean number of alleles, mean heterozygosity, and unbiased probability of identity) among the geographic areas sampled suggest that less variation is found from northern to southern regions

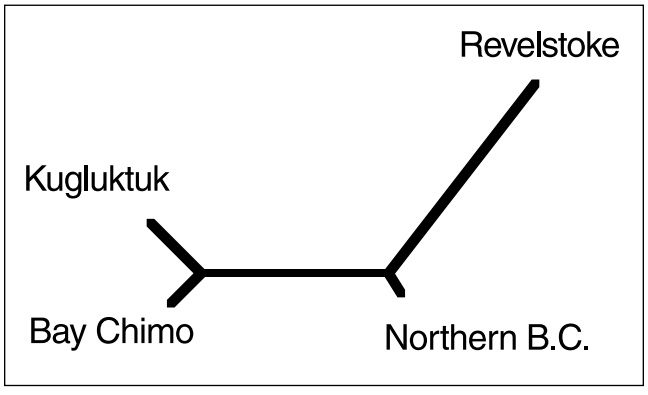

Figure. 2. Unrooted neighbour joining tree of genotype likelihood ratio values.

Table 4. Straight line distances (km) between centres of sampled regions and their approximate areas.

	Revelstoke	Northern B.C.	Bay Chimo	Kugluktuk
Revelstoke (7,000 km^2)	0			
Northern B.C. (10,000 km^2)	550	0		
Bay Chimo (15,000 km^2)	1,550	1,200	0	
Kugluktuk (15,000 km^2)	1,450	1,000	350	0

(Table 1). The relative size of each sampled area, however, could be influencing these results (Table 4). The Revelstoke area has the least variation, but it is also the smallest region sampled (approx 7,000 km^2), which may account for the lower mean number of alleles, mean heterozygosity, and unbiased probability of identity estimates. More likely, however, is that gene flow towards Revelstoke from the other areas sampled is somehow impeded, which accounts for its decreased genetic variation.

The 2 genetic distance measures, D_{LR} and D_S, present similar results (Table 2). Both measures suggest that the northern B.C. area is roughly genetically equidistant from both the Revelstoke area (D_{LR} = 0.544, D_S = 0.0565) and the NWT regions of Bay Chimo (D_{LR} = 0.582, D_S = 0.0661) and Kugluktuk (D_{LR} = 0.534, D_S = 0.0425). These results also suggest that the 2 NWT regions are relatively close genetically (D_{LR} = 0.240, D_S = 0.0421). The more proximal regions (Fig. 2) are more closely related genetically. The Revelstoke region, however, is disproportionately genetically distant from the northern B.C. region as compared with the NWT regions when taking their geographic locations into account (Tables 2 and 4). Once again, this result suggests that some factor impedes gene flow between the Revelstoke area and the other regions sampled, and there are fewer impediments among the remaining 3 areas.

The randomization of the genotype frequencies in the assignment test (Table 3), with all gene pools combined, suggests that there is a significant level (95%) of population genetic structure between all of these areas except the northern B.C. region. This region revealed a slightly less significant level of structure (90% confidence). The test also estimated that 37/48 animals were correctly assigned to Revelstoke from Revelstoke, whereas only 14/37 in northern B.C., 20/40 in Bay Chimo, and 24/45 in Kugluktuk were correctly assigned (Table 3). These results, when compared with one another, suggest that some impediment to gene flow into Revelstoke from the other regions exists. Note that these assignments do not necessarily mean that individuals captured in one region and assigned to another are actually migrants from one population to the other. What it does

demonstrate is that the probability of a particular genotype occurring in an area with similar genotypes is higher.

To estimate the levels of migration among the sampled groups the assignment test was run using the option of resampling the genotypes from their original gene pools (Table 3). The results for Revelstoke region suggest that there is little migration into the Revelstoke region; the number of misassigned individuals into this population is not well supported by the resampling of the genotype frequencies, whereas movement from Revelstoke to northern B.C. is well supported by the randomizations. The northern B.C. region seems to be receiving the most gene flow into it from both Revelstoke and the NWT regions. Ten individuals sampled from northern B.C. are more likely to have come from the Revelstoke region, while 7 and 6 individuals are more likely to have come from Bay Chimo and Kugluktuk, respectively. The resampling of these genotypes supports (approx 90% significance for individuals from Revelstoke and Bay Chimo) the initial assignment of individuals to each region, and that these individuals are actually migrants into the northern B.C. region. The Bay Chimo and Kugluktuk regions reflect the highest levels of gene flow between 2 groups with 14/45 individuals going from Bay Chimo to Kugluktuk and 11/40 going from Kugluktuk to Bay Chimo. This is expected from their geographic proximity. The gene flow to these NWT regions from Revelstoke is not significant. The gene flow from northern B.C. is a bit higher, however, with 7 individuals assigned (with 90% significance) to northern B.C. that were sampled in Bay Chimo.

The assimilation of the 3 types of data analyzed in this study (genetic variation, genetic distance, and gene flow via the assignment test) suggests that there is some impediment to gene flow into the Revelstoke region. The northern B.C. region, having less genetic structure and receiving the most gene flow into it from both northern and southern regions, seems to be acting as a sink population. However, only 4 regions have been analyzed in this preliminary study. The 2 NWT regions, although displaying some population genetic structure, do have extensive gene flow between them. This structure was not expected because of the ability of these animals to disperse and the more contiguous landscape in the northern regions of NWT. Further investigation with more loci and more sampled regions will enable us to refine and resolve the suggested levels of gene flow found in this species across its North American distribution.

ACKNOWLEDGEMENTS

We thank the following people for providing samples for this study: J. Krebs and D. Lewis from the Columbia Basin Fish and Wildlife Compensation Program; D. Reid and E. Lofroth from the Northern Wolverine Project; and R. Mulders and A. Bourque from the Department of Resources, Wildlife and Economic Development, NWT. This research was funded by Natural Sciences and Engineering Research Council and Parks Canada grants to C. Strobeck.

LITERATURE CITED

Banci, V. A. 1994. Wolverine. Pp. 99–127 *in* L. F. Ruggiero, K. B. Aubry, S. W. Buskirk, L. J. Lyon, and W. J. Zielinski, eds. The scientific basis for conserving forest carnivores, American marten, fisher, lynx and wolverine in the western United States. U.S. Dep. Agric., For. Serv. Gen. Tech. Rep. RM-254.

Davis, C. S., and C. Strobeck. 1998. Isolation, variability, and cross-species amplification of polymorphic microsatellite loci in the family Mustelidae. Mol. Ecol. 7:1771–1788.

Duffy, A. J., A. Landa, M. O'Connell, C. Stratton, and J. M. Wright. 1998. Four polymorphic microsatellites in wolverine, *Gulo gulo*. Anim. Genet. 29:63–72.

Felsenstein, J. 1995. PHYLIP (Phylogeny Inference Package) version 3.572. Univ. Wash., Dep. Genetics, Seattle, WA.

Gardner, C. L. 1985. The ecology of wolverines in southcentral Alaska. Thesis, Univ. Alaska-Fairbanks, Fairbanks, AK. 82pp.

Nei, M. 1972. Genetic distances between populations. Am. Nat. 106:283–292.

Paetkau, D., W. Calvert, I. Stirling, and C. Strobeck. 1995. Microsatellite analysis of population structure in Canadian polar bears. Mol. Ecol. 4:347–354.

_____, L. P. Waits, P. L. Clarkson, L. Craighead, and C. Strobeck. 1997. An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. Genetics 147:1943–1957.

_____, _____, _____, _____, E. Vyse, R. Ward, and C. Strobeck. 1998. Variation in genetic diversity across the range of North American brown bears. Conserv. Biol. 12:418–429.