The Journal of Wildlife Management; DOI: 10.1002/jwmg.736

Research Article

A Long-Term Population Monitoring Approach for a Wide-Ranging Carnivore: Noninvasive Genetic Sampling of Gray Wolf Rendezvous Sites in Idaho, USA

CARISA R. STANSBURY,[1] *Department of Fish and Wildlife Sciences, University of Idaho, 875 Perimeter Drive MS 1136, Moscow, ID 83844, USA*

DAVID E. AUSBAND, *Montana Cooperative Wildlife Research Unit, Natural Sciences Room 205, University of Montana, Missoula, MT 59812, USA*

PETER ZAGER, *Idaho Department of Fish and Game, 3316 16th Street, Lewiston, ID 83501, USA*

CURT M. MACK, *Wildlife Management Division, Nez Perce Tribe, 260 Phinney Drive, Lapwai, ID 83540, USA*

CRAIG R. MILLER, *Department of Mathematics, University of Idaho, P.O. Box 441103, Moscow, ID 83844, USA; and Department of Biological Sciences, University of Idaho, P.O. Box 441103, Moscow, ID 83844, USA*

MATTHEW W. PENNELL, *Department of Biological Sciences, University of Idaho, P.O. Box 441103, Moscow, ID 83844, USA; and Institute for Bioinformatics and Evolutionary Studies (IBEST), University of Idaho, PO Box 443051, Moscow, ID 83844, USA*

LISETTE P. WAITS, *Department of Fish and Wildlife Sciences, University of Idaho, 875 Perimeter Drive MS 1136, Moscow, ID 83844, USA*

ABSTRACT Various monitoring methods have been developed for large carnivores, but not all are practical or sufficiently accurate for long-term monitoring over large spatial scales. From 2009 to 2010, we used a predictive habitat model to locate gray wolf rendezvous sites in 4 study areas in Idaho, USA and conducted noninvasive genetic sampling (NGS) of scat and hair found at the sites. We evaluated species and individual identification PCR success rates across the study areas, and estimated population size with a single-session population estimator using 2 different recapture-coding methods. We then compared NGS population estimates to estimates generated concurrently from telemetry data. We collected 1,937 scat and 166 hair samples and identified 193 unique individuals over 2 years. For fecal DNA samples, species identification success rates were consistently high (>92%) across areas. Individual identification success rates ranged from 78% to 80% in the drier study areas and dropped to 50% in the wettest study area. The degree of agreement between NGS- and telemetry-derived population estimates varied by recapture-coding method with considerable variability in 95% confidence intervals. Population estimates derived from NGS methods were most influenced by the average number of detections per individual. We demonstrate how changes in field effort and recapture-coding method can affect population estimates in a widely used single-session population estimation model. Our study highlights the need to further develop reliable population estimation tools for single-session NGS data, especially those with large differences in capture frequencies among individuals stemming from severe capture heterogeneity (i.e., overdispersion). © 2014 The Wildlife Society.

KEY WORDS *Canis lupus*, fecal DNA, genetic, gray wolf, monitoring, noninvasive genetic sampling, population estimation, rendezvous site, single-session models.

Various monitoring methods have been developed for large carnivores including telemetry, capture-mark-recapture, harvest data, and sign survey (Wilson and Delahay 2001, Barea-Azcón et al. 2007). Noninvasive genetic sampling (NGS) techniques, which use DNA extracted from animal sign such as hair, scat, saliva, urine, or regurgitates (Waits and Paetkau 2005), have become an effective method for studying wildlife populations and is the preferred monitoring method for some species and populations (e.g., Rudnick et al. 2005, De Barba et al. 2010, Borthakur et al. 2011). The effectiveness of NGS for monitoring recolonizing gray wolf (*Canis lupus*) populations was first evaluated in Europe (Lucchini et al. 2002, Valière et al. 2003). More recently, researchers have implemented NGS approaches to monitor the reintroduced wolf population in Idaho, USA (Ausband et al. 2010, Stenglein et al. 2010*a*). With the 2011 delisting of wolves in the northern Rocky Mountains (U.S. Fish and Wildlife Service 2011), the responsibility for gray wolf monitoring has shifted from the federal government to state and tribal governments, and federal funds for monitoring will decrease annually and will no longer be available at the end of the required 5-year post-delisting monitoring period in 2016. Thus, more cost effective survey and monitoring techniques are urgently needed to meet post-recovery management goals (Kunkel et al. 2005) and maintain an

Received: 12 April 2013; Accepted: 5 April 2014

[1]*E-mail: carisarichele@gmail.com*

adequate, sustainable population size capable of withstanding an annual harvest.

Gray wolves use rendezvous sites, or resting sites, during the summer where pups can grow and learn basic social and physical skills until they are ready to travel with the pack. These sites are valuable for NGS monitoring because they provide an opportunity to sample all members of a pack within a small spatial area. Ausband et al. (2010) developed a predictive habitat model to identify highly probable rendezvous sites from habitat characteristics. Stenglein et al. (2010*a*) combined this predictive habitat model with NGS of scat and hair at these sites and demonstrated agreement between NGS- and telemetry-derived population estimates indicating that NGS could be an efficient and cost-effective monitoring alternative to telemetry-based techniques in areas of both low and high wolf density. However, this approach was applied only to a limited number of sites ($n = 2$) with very similar ecological conditions, so the effectiveness of these genetic monitoring methods needs to be evaluated across a large and more heterogeneous spatial area.

Our goal was to assess the efficacy of the Stenglein et al. (2010*a*) method to monitor gray wolves throughout Idaho, USA. We applied this approach in 2 existing study areas (East and West) and 2 new study areas (Central and North) in Idaho with different climate and habitat characteristics. We assessed the effectiveness of the approach to 1) detect wolf presence using genetic techniques, 2) provide a minimum count of wolves, and 3) estimate wolf abundance. We monitored all study areas concurrently using radio-telemetry techniques, which enabled us to compare independently derived telemetry- and NGS-based population estimates.

STUDY AREA

We conducted NGS surveys from mid-June to mid-August in 2009 and from mid-June to mid-July in 2010. In 2009, we sampled the Central, East, and West study areas that comprised 9,976 km^2 and 5 Game Management Units (GMUs) in Idaho, USA (Fig. 1). In 2010, we re-sampled the Central study area only and sampled the North study area comprising 5,933 km^2 to expand the spatial coverage of our study. The wolves in the Central study area were harvested between the sampling years of 2009 and 2010 for the first time since their reintroduction into Idaho. Therefore, we differentiated between the 2009 (Central09) and 2010 (Central10) datasets for the Central study area. The Central and North sites were characterized by average summer daily temperatures ranging from 8°C to 29°C and average annual precipitation of 71 cm and 56 cm, respectively. Vegetation was generally dominated by red cedar (*Thuja plicata*) and western hemlock (*Tsuga heterophylla*) forests. The East and West areas were characterized by average summer daily temperatures ranging from 4°C to 31°C and average annual precipitation of 42 cm and 62 cm, respectively. Vegetation was generally dominated by ponderosa pine (*Pinus ponderosa*), lodgepole pine (*P. contorta*), and spruce (*Picea englemannii*) forests, and sagebrush (*Artemesia tridentata*) steppe (Western Regional Climate Center 2011).

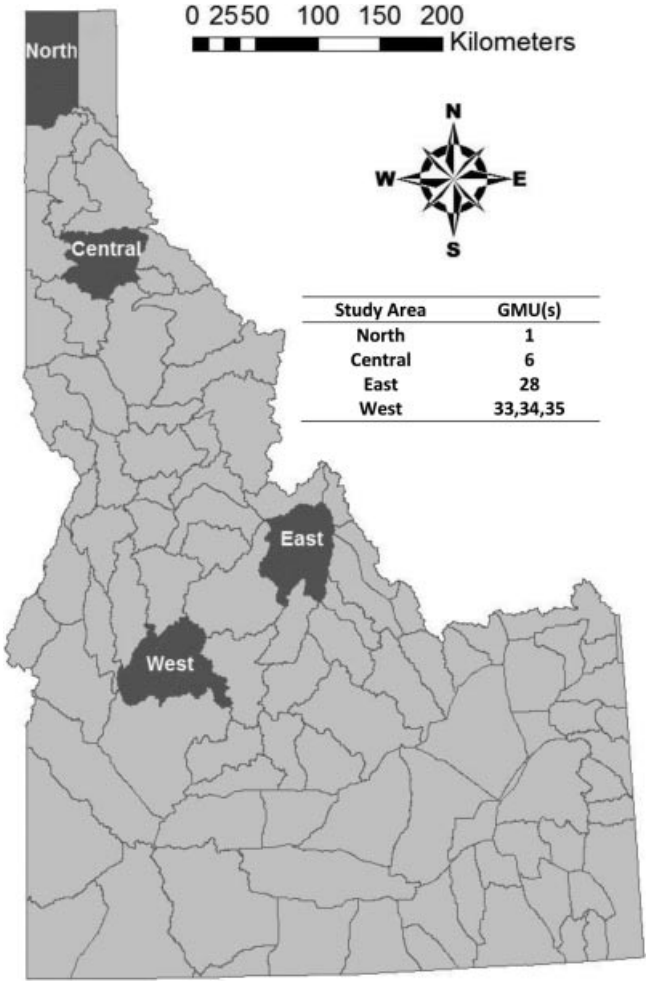

Figure 1. Study areas and corresponding Game Management Units (GMUs) in Idaho, USA.

METHODS

Field Methods

In combination with Idaho Fish and Game wolf monitoring personnel, we conducted annual monitoring of wolves, which included 1) trapping and radio-collaring $\geq$1 individual(s) from packs known to be within study areas, 2) collecting tissue samples during radio-collaring to provide a known genetic reference sample, and 3) tracking individuals through ground searches and aerial observations at rendezvous sites to determine pack size. The number of observed wolves during summer months (Jun–Aug) provided a minimum count per study area. We added 9% to our minimum count, to account for lone wolves, which provided a telemetry-derived population estimate (Stenglein et al. 2010*a*). We were not always able to obtain precise counts when conducting independent telemetry surveys and for these cases we provide a range for the minimum count and population estimate. This research was carried out under an Institutional Animal Care and Use Committee protocol obtained through The University of Montana (008-09MMMCWRU-031009).

Detailed field sampling methods are described in Stenglein et al. (2010*a*), and only key aspects or protocol changes are

summarized here. We applied a predictive habitat model developed for the state of Idaho that uses green leaf biomass, profile curvature, and roughness to identify highly probable ($\geq$70% suitability) rendezvous sites (Ausband et al. 2010). The model predicts rendezvous site habitat suitability for each 15-m^2 pixel across Idaho. We surveyed all probable rendezvous sites with $\geq$70% suitability at dawn and dusk when wolves are most active (Harrington and Mech 1982). Upon arrival at a site, we conducted a howl survey where we howled and waited for a response. If we did not hear a response, we surveyed the probable rendezvous site for 30–45 minutes for wolf sign. We categorized probable rendezvous sites without extensive wolf sign as inactive sites. If wolves reciprocated the howl or we observed wolves or found extensive sign (i.e., $\geq$10 scats) of wolves (hereafter active sites), we searched extensively to locate the activity center (i.e., area with concentrated wolf sign). Upon locating an activity center, we sampled for scat and hair within a 500-m radius around the activity center. We surveyed and sampled each site 1 time. After we detected and sampled an active site, we excluded other probable rendezvous sites within a 6.4-km radius from surveys to avoid duplicate sampling of packs and to increase efficiency of the approach (Ausband et al. 2010). We also collected adult scats (diameter $\geq$ 2.54 cm; Ausband et al. 2010, Stenglein et al. 2010*a*) opportunistically as we traveled between probable rendezvous sites (hereafter incidentals). We collected scat samples by removing a small portion from the side of the scat (Stenglein et al. 2010*b*) and stored samples in a 2-ml vial containing DET buffer (Frantzen et al. 1998). We used sterilized forceps to collect hair found snagged on trees or distinct tufts on the ground.

In 2008, we collected all samples at a site (Stenglein et al. 2010*a*, 2011), but in 2009 we did not collect hair from daybeds to minimize survey time (approx. 5 hr). In 2010, we changed our field protocol to sample active sites for 3 hours instead of 5 hours (Stenglein et al. 2011) and included 1-hour searches for hair in daybeds. At active sites, 6 technicians sampled in teams of 2. One pair of technicians searched for and sampled scat in the activity center (i.e., areas most concentrated with sign). A second pair of technicians searched for and sampled hair in the activity center for 1 hour then switched to searching for and sampling scat in the activity center. The third pair of technicians searched for and sampled both scat and hair on trails leading from the activity center throughout the 500-m radius sampling area in an effort to target additional adult wolves traveling to and from or resting outside of the activity center.

Genetic Methods

We analyzed samples at the Laboratory for Ecological, Evolutionary and Conservation Genetics (LEECG) in a facility dedicated to low quality and quantity DNA at the University of Idaho, USA. We extracted DNA from scat and hair samples using Qiagen kits (Qiagen, Inc., Valencia, CA) as described in Stenglein et al. (2010*a*) and included a negative control to test for contamination. We initially screened all samples in a species-identification test using a 3-primer mitochondrial DNA (mtDNA) control region polymerase chain reaction (PCR; H16145, SIDL, H3R; Murphy et al. 2000, Dalén et al. 2004, Onorato et al. 2006) to remove non-target species (e.g., coyote [*C. latrans*]) and low-quality samples (Lucchini et al. 2002). For individual identification, we genotyped all 2009 samples identified as wolf or dog in a 7-μl, 9-locus PCR multiplex as described in Stenglein et al. (2010*a*). For the 2010 samples, we re-optimized the 7-μl multiplex in an effort to improve PCR success rates at the following concentrations for primer pairs: 0.04 μM C09.173, FH2088 and FH2137; 0.05 μM FH2054 and FH2611; 0.09 μM Cxx.119, FH2001 and FH3725; 0.1 μM FH2670. We used an Applied Biosystems 3130xl capillary machine (Applied Biosystems, Inc., Foster City, CA) to separate PCR products by size and verified peaks individually by eye with GENEMAPPER 3.7 (Applied Biosystems, Inc.). We genotyped 66 tissue samples from wolves captured in Idaho, 17 dogs (*C. lupus familiaris*), and 40 coyotes to serve as reference samples. For genotyping, we initially amplified all samples twice; we required successful amplification of alleles at $\geq$5 loci for the sample to continue for an additional 1–3 replications, whereas we discarded samples that amplified at <5 loci. For each locus, we required $\geq$2 independent PCR amplifications for consensus of a heterozygote and $\geq$3 independent PCR amplifications for consensus of a homozygote. We included a negative control in all PCRs to test for contamination. Given the potentially large number of first-degree relatives in our dataset, we required $\geq$7 loci to consider a sample genotyped at consensus and $\geq$8 loci ($P_{(ID)}$sibs ranging from 0.000354 to 0.00118 across study areas; Waits et al. 2001) to confirm multiple detections of the same individual. We compared all consensus genotypes and all unique genotypes of previously identified individuals using GENALEX (Peakall and Smouse 2006) to match samples and distinguish unique genotypes. We analyzed samples matching at all but 1 locus in a second 9-locus multiplex (AHT103, AHT109, AHT121, AHTk200, C05.377, C37.172, Cxx.250, FH2004, and FH2010; Ostrander et al. 1993, Holmes et al. 1994, Breen et al. 2001) as described in Stenglein et al. (2011) to confirm a match or mismatch. To avoid overestimation and account for undetected genotyping errors, we lumped samples mismatching by allelic dropout (ADO) at only 1 locus (e.g., 102, 102 vs. 102, 106) as a single individual (Adams and Waits 2007). We used RELIOTYPE (Miller et al. 2002) to test the accuracy of unique genotypes detected in only 1 sample (i.e., single captures) by ensuring the genotype attained a 95% accuracy threshold. Last, we ran all unique genotypes in STRUCTURE v2.3.3 (Pritchard et al. 2000) with reference samples of known wolves, domestic dogs, and coyotes at $K = 3$ under the general admixture model, with a burn-in of 100,000, and 500,000 Markov Chain Monte Carlo (MCMC) repetitions and 10 iterations to estimate individual ancestry and remove samples highly probable as dogs or coyotes from the dataset.

Analysis

To evaluate the survey for detecting wolves through NGS, we calculated 1) the proportion of fecal samples that were

positively identified as wolf using the mtDNA species-identification test, 2) the proportion of probable rendezvous sites confirmed to have wolf sign, 3) the proportion of samples that resulted in consensus genotypes and unique genotypes, and 4) the average number of samples collected per individual.

For each study area, we assessed DNA quality by calculating species- and individual-identification PCR success rates and average error rates due to ADO and false alleles (FA). We calculated species-identification success rates as the number of successful PCRs of gray wolf, or other sympatric carnivore species identifiable by the test, divided by the total number of samples. We calculated individual-identification success rates as the number of samples producing a consensus genotype divided by the number of samples identified as wolf. We calculated ADO and FA error rates for the first 2 PCRs on 100 randomly chosen scat samples per study area that reached a consensus genotype (Broquet and Petit 2004); we did not include hair when calculating genotyping error rates because of low sample sizes. We used Fisher's exact test to evaluate differences in species- and individual-identification success rates. Also, we evaluated Hardy-Weinberg equilibrium and genetic diversity statistics (see Supplementary Material, available online at www.onlinelibrary.wiley.com).

We used the number of unique genotypes as the NGS minimum count of wolves in each study area. We calculated probability of capture ($\hat{P}$) by dividing the number of known, radio-collared wolves identified through NGS by the total number of known, radio-collared wolves in the study areas. We estimated population size using the R package CAPWIRE (Miller et al. 2005, Pennell et al. 2012), a maximum likelihood, single-session population estimator designed for NGS data. We estimated abundance using the two innate rates model (TIRM) because we expected capture heterogeneity in our datasets due to age, sex, and social status (Piggott and Taylor 2003). CAPWIRE assumes 1) geographic and demographic closure, 2) correctly identified genetic tags, and 3) observations are independent events (Miller et al. 2005).

We applied 2 recapture-coding methods to estimate population size in CAPWIRE. First, we used capture rules (CAP) developed by Stenglein et al. (2010*a*) with detections condensed into recaptures as follows: an individual may be recaptured within a rendezvous site between data types (i.e., scat and hair), between rendezvous sites, or outside of a rendezvous site as an incidental, and among incidentals from different incidental locations. Second, we used all of the data (ALL) for a given study area where we marked each sample exhibiting the same genotype as a recapture following the first detection regardless of its spatial location.

RESULTS

We collected 1,217 samples in 2009 and 886 samples in 2010 (Table 1). Incidental scats, those not collected at rendezvous sites, accounted for 7% of the total samples. We did not collect any incidental hair samples. Of the total hair samples collected, we identified 82% and 88% as wolf in 2009 and 2010, respectively. In both years, we confirmed 82% of all scats as wolf through species identification. For 2009 and 2010 respectively, wolf scats comprised 85% and 79% of all scats collected within predicted inactive or active rendezvous sites and 55% and 46% of scats collected as incidentals. We confirmed wolf presence with species identification at or near 18% of predicted rendezvous sites (Table 1). Coyotes were the most common non-target species, accounting for 8% of all samples. Coyote scats comprised 9% of scat samples collected within predicted rendezvous sites and 44% of scats collected as incidentals. One sample exhibited a high probability of being a domestic dog ($q = 0.8$ in STRUCTURE) and we removed the sample from analyses. Microsatellite analysis produced 1,168 consensus genotypes, which resulted in identification of 193 unique individuals. We detected individuals 1–30 times (Fig. 2), and the average number of samples per individual ranged from 3 to 10 across study areas (Table 1). We identified all 193 individuals with scat and 37 of those with hair; we did not identify new individuals with hair alone.

Success Rates and Error Rates

Species-identification success rates for scat were high (>92%) for all study areas in both years (Table 2), and we detected no differences among areas ($P = 0.065$). Species identification success rates for hair varied considerably in 2009 (53–95%), likely because of small sample sizes in the West and Central study areas, with the only difference occurring between the East and West study areas ($P \leq 0.001$). In 2010, species-identification success rates for hair were >90% for both study areas and we found no difference ($P = 0.48$) among areas.

In 2009, we found no difference ($P = 0.059$) in individual-identification success rates for scat in the East and West study areas (Table 2). However, we detected a difference between both the East and West sites and the Central study area (the wettest site) where the individual-identification success rate dropped by roughly 29% ($P \leq 0.001$). In 2010, individual-identification success rates for scats were lower in the Central site (66%) than the North area (72%), but the difference was not significant ($P = 0.083$). We did not statistically compare individual-identification success rates for hair samples among study areas because of low sample sizes (Table 1). The average number of PCRs required to complete a consensus genotype ranged 3.2–3.8 for scat and 2.9–4.8 for hair for all areas in both years (Table 2).

We calculated genotyping error rates for scats only. The ADO rates were higher (4.2–9.3%) and more variable than FA rates (0.5–0.9%; Table 2). We observed the highest ADO rate in the Central area in both years. In 2010, we detected a significant difference between ADO rates in the Central and North areas ($P = 0.005$). We found no differences in FA error rates among study areas in 2009 ($P = 0.31$) or in 2010 ($P = 0.57$).

Minimum Count and Population Estimates

We identified 139 individuals in 2009 and 67 individuals in 2010 through genetic analyses (Table 1). We detected 13

Table 1. Number of gray wolf rendezvous sites surveyed, number of samples collected by data type, percent of surveyed sites with genetically confirmed wolf sign (scat, hair), average number of samples per individual, and the minimum number of wolves from genetic analysis (NGS) and counted through radio-telemetry (RT) by year and study area in Idaho. Incidentals (Inc.) were scat samples we collected opportunistically as we traveled between probable rendezvous sites.

Year	Area	Sites surveyed	No. samples			% Sites with confirmed wolf sign	Samples/individual	Min. no. wolves	
			Scat	Hair	Inc.			NGS	RT
2009	West	166	470	15	21	22	7.0	50	62
	East	151	367	44	28	20	5.0	55	46
	Central	118	232	3	37	19	3.0	34	33
	Total	435	1,069	62	86	21	5.0	139	141
2010	Central	117	312	40	44	16	5.2	37	31
	North	131	403	64	23	11	10	30	31
	Total	248	715	104	67	14	7.7	67	62

individuals in both years (i.e., redetection of the same individuals) in the Central study area; thus, our study resulted in an overall minimum count of 193 individuals across the 4 study areas. The NGS and telemetry minimum counts were similar across study areas; neither method was consistently higher or lower than the other (Table 1). The average number of detections per individual varied considerably across years, study areas, and recapture-coding methods (Table 3). In 2009, using NGS, we detected 15 of 22 (68%) radio-collared wolves, representing 10 of 16 packs (63%) known to be in the study areas. In 2010, we detected 10 of 13 (77%) radio-collared wolves, representing 10 of 13 known packs (77%).

The NGS-derived population estimates (Fig. 3) using the 2 recapture-coding methods produced different estimates with considerable variability in 95% confidence intervals. Under CAP, all NGS-derived estimates were higher than telemetry-derived estimates. The confidence intervals were highly variable (width: 23–140), and overlap with the telemetry-derived estimate occurred in only 2 datasets where the average number of detections per individual exceeded 1.7 (West and North). Under ALL, NGS-derived estimates were below the telemetry-derived estimate for 2 datasets (West and North), and 4 of the 5 estimates were closer to the NGS minimum count than the telemetry-derived estimate. The confidence intervals were generally small (1–16) and

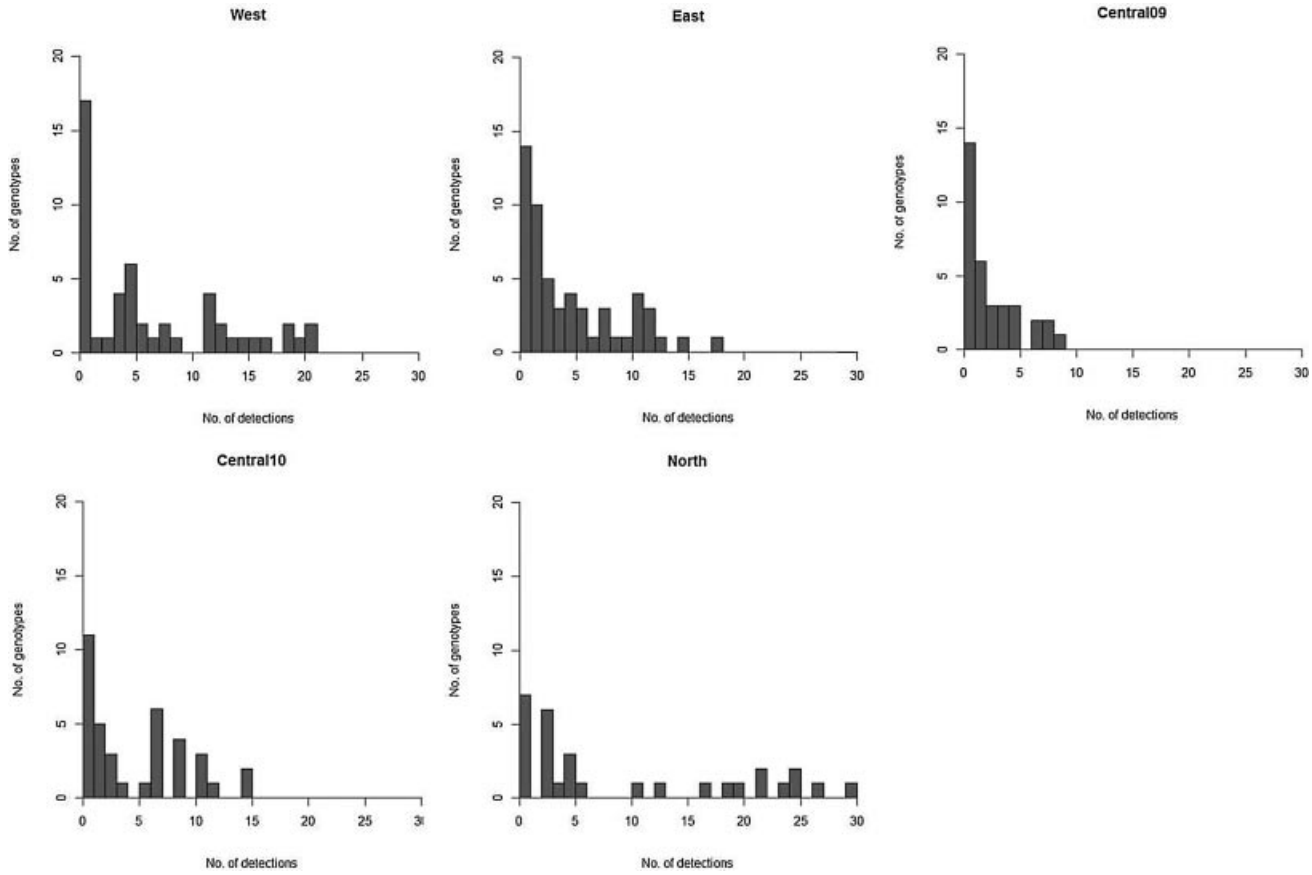

Figure 2. Capture distributions from noninvasive genetic sampling (NGS) data for gray wolves in each study area in Idaho.

Table 2. Species-identification (SpID) and individual-identification (IndID) success rates, error rates, and average number of polymerase chain reactions (PCRs) per sample by year and study area for gray wolves in Idaho. We present allelic drop-out (ADO) and false alleles (FA) error rates for scat only.

		SpID success (%)		IndID success (%)		Error rate (%)		PCRs[a]	
Year	Area	Scat	Hair	Scat	Hair	ADO	FA	Scat	Hair
2009	West	95	53	80	57	4.2	0.5	3.7	4.0
	East	92	95	78	60	4.3	0.8	3.2	3.5
	Central	95	67	50	100	5.6	0.7	3.2	4.8
	Average	94	72	69	72	4.7	0.67	3.4	4.1
2010	Central	97	90	66	36	9.3	0.9	3.4	2.9
	North	99	94	72	63	6.7	0.7	3.8	3.4
	Average	98	92	69	50	8.0	0.8	3.6	3.2

[a] Calculated only for samples that reached a consensus genotype.

overlap with the telemetry-derived estimate occurred twice (Central09 and North).

DISCUSSION

Our study evaluates an NGS approach to surveying gray wolf rendezvous sites using a previously developed predictive habitat model that we applied to new study areas with different climate and habitat. We demonstrate that this approach can successfully detect gray wolf presence across a large spatial area, and provide an accurate minimum count of individuals. The degree of agreement between NGS- and telemetry-derived population estimates varied by recapture-coding method and was affected by changes in our field effort.

Performance of NGS Across Study Areas

Evaluating genotyping success rates is important for genetic-based mark-recapture studies. Each sample that fails to amplify is either a new individual or a recapture of an existing individual, both of which influence the average number of detections per individual and population estimates. A number of factors influence PCR success rates, including seasonality (Lucchini et al. 2002, Hájková et al. 2006), exposure to moisture and ultraviolet light (Lindahl 1993, Friedberg 2003), sampling location on scat (Stenglein et al. 2010*b*), storage method (Roon et al. 2003, Nsubuga et al. 2004), and age of the sample (Murphy et al. 2007, Santini et al. 2007). For wolf scats, Lucchini et al. (2002) reported significantly better PCR success rates for mtDNA and nDNA in winter compared to summer, and Santini et al. (2007) reported a significant decrease in nDNA PCR success rates from DNA degradation after 3 days. By sampling from active rendezvous sites in the summer months, our survey method targets fresher, less degraded, samples. As expected, ADO error rates increased and individual-identification success rates declined in the wetter and more humid study areas. We documented a 16% increase in individual-identification PCR success in the wetter Central study area from 2009 to 2010 suggesting our PCR re-optimization efforts improved success rates. However, because we do not know the age of our samples and did not quantify freshness of scats, we are unable to rule out the possibility that the increased success rate was a result of collecting fresher (i.e., less degraded) samples in 2010.

Table 3. Average number of detections per individual for gray wolves in Idaho under 2 data-coding methods for genetic samples: CAP = restricted recaptures between data types (hair and scat) and rendezvous sites, ALL = all genetic detection data used and considered a separate recapture.

Year	Area	CAP	ALL
2009	West	1.8	7.0
	East	1.7[a]	5.0
	Central	1.2[a]	3.0
2010	Central	1.4[a]	5.2
	North	2.5	10

[a] Average number of detections per individual ≤1.7 threshold Stenglein et al. (2010*b*).

Field Effort and Sampling Strategy

Our strategy of using models to identify habitats suitable for use as rendezvous sites and NGS approaches using scat and hair remains was effective and efficient. We improved

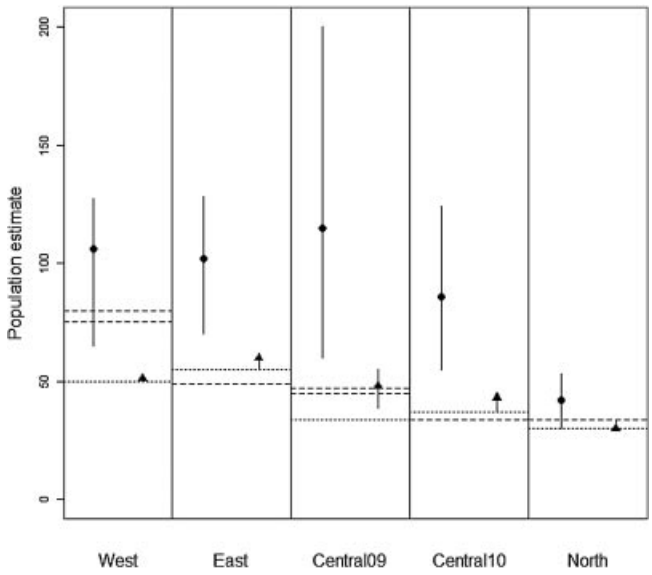

Figure 3. Noninvasive genetic sampling (NGS) population estimates (with 95% CI) using 2 different recapture coding methods (CAP = captures between sample types and rendezvous sites. ALL = each sample with the same genotype counts as a recapture) for all datasets in both years. Included for each dataset is the genetic minimum count (NGS min) and telemetry population estimate (Telemetry). Thicker horizontal dashed lines indicate a range in the population sizes for some telemetry estimates.

efficiency from an earlier study (Stenglein et al. 2010*a*) by reducing sampling time per active site from 5 hours to 3 hours and implementing a strategic and standardized sampling scheme. Reducing time spent at rendezvous sites can substantially affect the efficiency of this method by reducing cost of technicians, sampling supplies, and DNA analysis. Minimizing time surveying active rendezvous sites is extremely important to keep our method as noninvasive as possible to reduce our impact on these pup-rearing homesites (Frame et al. 2007). Also, rendezvous sites are used by wolves for only a few months in the summer so, minimizing time spent at 1 site provides more time to survey and sample additional sites allowing technicians to cover a larger area and potentially identify more packs. However, changes in the field effort and sampling strategy affected the average number of detections per individual and population estimates more than expected. Sub-sampling simulations conducted by Stenglein et al. (2010*a*) predicted that hair collection could be reduced without affecting population estimates. Altering the field protocol by eliminating collection of hair in daybeds reduced the total number of hair samples from 637 in 2008 (Stenglein et al. 2010*a*) to 62 in 2009. This drastically reduced the average number of detections per individual under CAP and, as expected, affected the precision and accuracy of the corresponding population estimates. Incorporating hair collection for a limited period of time in 2010 increased the average number of detections per individual (Table 3). Alterations to improve efficiency must be done carefully, especially when using multiple data types like hair and scat in a mark-recapture framework because they may generate extrinsic heterogeneity (i.e., resulting from study design) in the dataset (Crespin et al. 2008).

The sampling design implemented in 2010 allowed us to minimize extrinsic heterogeneity and sample active sites in less time while addressing heterogeneity arising from biological differences such as sex, age, and social influences (i.e., intrinsic heterogeneity; Piggott and Taylor 2003, Crespin et al. 2008). We attempted to minimize intrinsic heterogeneity by expanding collection of scat and hair samples from individuals resting away from the activity center, moving on trails, and localized at the activity center. Collecting both hair and scats can reveal more individuals in a population and allow for mark-recapture analysis between data types (Boulanger et al. 2008, Stenglein et al. 2010*a*). However, in our study, we identified all individuals with scat and did not detect new individuals with hair alone.

Population Estimation and Single-Session Estimators

Minimum counts and population estimates are valuable metrics in conservation and management. This study and others have shown agreement between NGS- and telemetry-derived minimum counts and population estimates for large carnivores (Stenglein et al. 2010*a*, Sawaya et al. 2011) providing support for researchers and managers to implement NGS population monitoring techniques. Mark-recapture methods are traditionally conducted in a multi-session framework in which animals are captured and marked in the first session, and in following sessions, the proportion of marked to unmarked animals captured is recorded (Otis et al. 1978). Mark-recapture using NGS data allows for passive sampling and single-session sampling where all samples may be efficiently obtained in 1 sampling occasion (Lukacs and Burnham 2005, Miller et al. 2005); this advantage has led to frequent use of NGS data and single-session estimators on a variety of species (Puechmaille and Petit 2007, Ruell et al. 2009, Mowry et al. 2011).

One of the challenges of using NGS and single-session estimation models is determining whether each detection following the first detection can be counted as a recapture or if detections should be coded into defined capture occasions. A major reason to develop recapture-coding methods is to ensure capture data fit model assumptions (e.g., independence) and thereby optimize performance of estimators. For example, Stenglein et al. (2010*a*) restricted recaptures to be between sample types and rendezvous sites (i.e., CAP) and showed that CAPWIRE performs well on these data when the average detections are >1.7 per observed individual. In assessing the performance of 2 estimation methods (i.e., ALL, CAP), we found CAP often yielded average number of detections per individual <1.7 (East, Central09, and Central10) and generally resulted in large confidence intervals and high point estimates compared to telemetry data. The low average number of detections per individual is driven by low sample sizes for hair in the East and Central09 datasets and the decrease in individual-identification success rates, observed primarily in the wettest study area (Central). Similar to Stenglein et al. (2010*a*), we found reduced performance in CAPWIRE when the average number of detections dropped below 1.7. Study areas (West and North) where the number of detections per individual was >1.7 had confidence intervals overlapping the telemetry-derived estimate indicating agreement between the 2 methods.

Marking each sample exhibiting the same genotype as a recapture following the first detection (ALL) is a common estimation method for NGS single session datasets (e.g., Williams et al. 2009, Mowry et al. 2011). In this study, the ALL method produced high average number of detections per individual and small confidence intervals relative to CAP (Fig. 3). However, telemetry data indicated that not all estimates are reliable. The West and North study areas produced NGS estimates equivalent to the minimum count, suggesting we detected all individuals in the area. Nevertheless, telemetry data revealed we missed active rendezvous sites for some packs indicating the ALL method is likely to be underestimating the population size in those areas.

A closer examination of the ALL datasets potentially explains why this coding method is problematic for CAPWIRE. One obvious problem with the ALL method is that it can easily violate the assumption that recaptures should be derived from independent events and can produce biased estimates of population size. Many of the datasets in this study have capture counts per individual ranging from singles into the teens, with a few individuals observed 20 or even 30 times (Fig. 3). These overdispersed counts are highly

unlikely to be observed if samples were independently drawn from the two innate rates capture class CAPWIRE model (TIRM). Furthermore, when a few individuals are detected an unusually large number of times, it can create severe bias in estimating α, the capture probability ratio between the 2 capture classes (Miller et al. 2005). Together, these issues appear to have led to poor performance of the CAPWIRE estimator on some of the ALL datasets.

We attempted to address these problems by developing a method of partitioning counts, removing improbably large counts from the datasets, and then analyzing the remaining data using CAPWIRE. This resulted in reasonable population estimates for our datasets that were intermediate between CAP and ALL estimates and overlapped the telemetry estimate in 2 out of 5 cases. Extensive simulation work revealed, however, that the method yields inconsistent results across a wider range of parameter space. Although partitioning improved wolf population estimates in some regions of parameter space, it made them worse in others, and it was not always possible to discern from which region of parameter space the data arose. This leads us to conclude that in some study circumstances, different methods of filtering or coding datasets will not be enough to produce robust estimates. Instead, the development of single-session estimators that incorporate higher levels of capture heterogeneity is needed.

One strategy that has been used in carnivore monitoring projects is to sample different data sources, such as hair or scat, and code recaptures as instances where the same individual is observed with both sources (e.g., CAP; Boulanger et al. 2008, Gervasi et al. 2008, Stenglein et al. 2010*a*). This approach may not always be optimal for several reasons. First, it may not be feasible or cost-effective to sample with the level of field intensity necessary to obtain multiple data types or to analyze the large number of samples collected for such a study. Second, the approach throws away a large amount of the raw detection data, at least some of which contains information about the population size. Exacerbating this problem is the fact that different data types have different capture probabilities and different DNA degradation and PCR success rates which can lead to highly uneven sample sizes between data types. Third, eliminating collection of 1 data type could simplify and further improve efficiency in both field sampling and laboratory procedures. Finally, the critical assumption behind the approach is that individual capture probabilities are not strongly correlated across data types; if this assumption is violated, as it likely is in our study, then the multiple data type method does not solve the problem of non-independence and only reduces sample size. Although a multiple-data-types approach to mark-capture estimation can be effective, it is not likely to be a general solution applicable to every monitoring program. Thus, we stress the need for development of new methods of mark-recapture analysis, particularly for single-session population estimators, that can account for non-independence between observations, allow more realistic forms of capture heterogeneity, and ideally are flexible with regards to the number of data sources and the number of sample sessions.

MANAGEMENT IMPLICATIONS

Predictive habitat modeling of rendezvous sites combined with NGS and species identification using mtDNA is an effective method to detect gray wolves and can be used by managers to document species presence and distribution. This method is particularly applicable for social species that leave abundant sign, thus allowing for single-session sampling. The addition of nDNA microsatellite analysis provides the opportunity to identify unique individuals, assess group structure, estimate abundance and density, identify dispersers, and assess genetic diversity and gene flow. We can gain valuable data on pup survival and population age structure by identifying pups-of-the-year at rendezvous sites and conducting subsequent annual resampling of the population. The ability to recapture individuals over multiple years of NGS provides opportunities to estimate population parameters (Marucco et al. 2011, Caniglia et al. 2012) and evaluate turn-over of breeders, pack splitting and formation, prevalence of inbreeding, and pack longevity (Lucchini et al. 2002, Marucco et al. 2009, Stenglein et al. 2011). As gray wolves are now subject to harvest in multiple regions, population estimation could also be conducted using genotypes generated from a combination of NGS survey data and harvested individuals (Dreher et al. 2007, Williams et al. 2009). Overall, this approach provides a valuable alternative method to population monitoring and should enhance our understanding of gray wolf demographics and population dynamics in the northern Rocky Mountains and elsewhere.

ACKNOWLEDGMENTS

This study was made possible through a collaborative multi-year wolf monitoring project and funding was provided by the Nez Perce Tribe, Idaho Department of Fish and Game through the U.S. Fish and Wildlife Service Federal Aid to Wildlife Restoration Project W-160-R, Oregon Zoo Future for Wildlife Grants, Regina Bauer Frankenberg Foundation for Animal Welfare, Leonard X. Bosack and Bette M. Kruger Foundation, U.S. Fish and Wildlife Service, The Mountaineers Foundation, Wolf Recovery Foundation, Wilburforce Foundation, University of Idaho (UI) Student Grant Program, and UI Environmental Science Program. The authors thank the Waits lab group, D. Beck, K. Cleary, S. Woodruff, and P. Joyce for their comments and contributions. The authors also thank S. Bassing, M. Brinkmeyer, A. Fahnestock, B. Fannin, A. Greenleaf, Q. Harrison, J. Hayden, C. Henderson, J. Holyan, D. Hopkins, J. Husseman, C. Jacobs, J. Linch, T. Loya, M. Mitchell, M. Myrvold, S. Nadeau, K. Oelrich, C. Rosenthal, J. Ruprecht, D. Spicer, J. Struthers, R. Toldness, W. Wakkinen, C. White, and B. Wojcik for their support.

LITERATURE CITED

Adams, J. R., and L. P. Waits. 2007. An efficient method for screening faecal DNA genotypes and detecting new individuals and hybrids in the red wolf (*Canis rufus*) experimental population area. Conservation Genetics 8:123–131.

Ausband, D. E., M. S. Mitchell, K. Doherty, P. Zager, C. M. Mack, and J. Holyan. 2010. Surveying predicted rendezvous sites to monitor gray wolf populations. Journal of Wildlife Management 74:1043–1049.

Barea-Azcón, J. M., E. Virgós, E. Ballesteros-Duperón, M. Moleón, and M. Chirosa. 2007. Surveying carnivores at large spatial scales: a comparison of four broad-applied methods. Biodiversity and Conservation 16:1213–1230.

Borthakur, U., R. D. Barman, C. Das, A. Basumatary, A. Talukdar, M. F. Ahmed, B. K. Talukdar, and R. Bharali. 2011. Noninvasive genetic monitoring of tiger (*Panthera tigris tigris*) population of Orang National Park in the Brahmaputra floodplain, Assam, India. European Journal of Wildlife Research 57:603–613.

Boulanger, J., K. C. Kendall, J. B. Stetz, D. A. Roon, L. P. Waits, and D. Paetkau. 2008. Multiple data sources improve DNA-based mark-recapture population estimates of grizzly bears. Ecological Applications 18:577–589.

Breen, M., S. Jouquand, C. Renier, C. S. Mellersh, C. Hitte, N. G. Holmes, A. Chéron, N. Suter, F. Vignaux, A. E. Bristow, C. Priat, E. McCann, C. André, S. Boundy, P. Gitsham, R. Thomas, W. L. Bridge, H. F. Spriggs, E. J. Ryder, A. Curson, J. Sampson, E. A. Ostrander, M. M. Binns, and F. Galibert. 2001. Chromosome-specific single-locus FISH probes allow anchorage of an 1800-marker integrated radiation-hybrid/linkage map of the domestic dog genome to all chromosomes. Genome Research 11:1784–1795.

Broquet, T., and E. Petit. 2004. Quantifying genotyping errors in noninvasive population genetics. Molecular Ecology 13:3601–3608.

Caniglia, R., E. Fabbri, S. Cubaynes, O. Gimenez, J.-D. Lebreton, and E. Randi. 2012. An improved procedure to estimate wolf abundance using non-invasive genetic sampling and capture–recapture mixture models. Conservation Genetics 13:53–64.

Crespin, L., R. Choquet, M. Lima, J. Merritt, and R. Pradel. 2008. Is heterogeneity of catchability in capture–recapture studies a mere sampling artifact or a biologically relevant feature of the population? Population Ecology 50:247–256.

Dalén, L., A. Götherström, and A. Angerbjörn. 2004. Identifying species from pieces of faeces. Conservation Genetics 5:109–111.

De Barba, M., L. P. Waits, E. O. Garton, P. Genovesi, E. Randi, A. Mustoni, and C. Groff. 2010. The power of genetic monitoring for studying demography, ecology and genetics of a reintroduced brown bear population. Molecular Ecology 19:3938–3951.

Dreher, B. P., S. R. Winterstein, K. T. Scribner, P. M. Lukacs, D. R. Etter, G. J. M. Rosa, V. A. Lopez, S. Libants, and K. B. Filcek. 2007. Noninvasive estimation of black bear abundance incorporating genotyping errors and harvested bear. Journal of Wildlife Management 71:2684–2693.

Frame, P. F., H. D. Cluff, and D. S. Hik. 2007. Response of wolves to experimental disturbance at homesites. Journal of Wildlife Management 71:316–320.

Frantzen, M. A., J. B. Silk, J. W. Ferguson, R. K. Wayne, and M. H. Kohn. 1998. Empirical evaluation of preservation methods for faecal DNA. Molecular Ecology 7:1423–1428.

Friedberg, E. C. 2003. DNA damage and repair. Nature 421:436–440.

Gervasi, V., P. Ciucci, J. Boulanger, M. Posillico, C. Sulli, S. Focardi, E. Randi, and L. Boitani. 2008. A preliminary estimate of the Apennine brown bear population size based on hair-snag sampling and multiple data source mark–recapture Huggins models. Ursus 19:105–121.

Hájková, P., B. Zemanová, J. Bryja, B. Hájek, K. Roche, E. Tkadlec, and J. Zima. 2006. Factors affecting success of PCR amplification of microsatellite loci from otter faeces. Molecular Ecology Notes 6:559–562.

Harrington, F. H., and L. D. Mech. 1982. Patterns of homesite attendance in two Minnesota wolf packs. Pages 81–105 in F. H. Harrington and P. C. Paquet, editors Wolves of the world: perspectives of behavior, ecology, and conservation. Noyes, Park Ridge, New Jersey, USA.

Holmes, N. G., N. J. Strange, M. M. Binns, C. S. Mellersh, and J. Sampson. 1994. Three polymorphic canine microsatellites. Animal Genetics 25:200.

Kunkel, K., C. Mack, and W. Melquist. 2005. An assessment of current methods for surveying and monitoring wolves. Nez Perce Tribe, Lapwai, Idaho, USA.

Lindahl, T. 1993. Instability and decay of the primary structure of DNA. Nature 362:709–715.

Lucchini, V., E. Fabbri, F. Marucco, S. Ricci, L. Boitani, and E. Randi. 2002. Noninvasive molecular tracking of colonizing wolf (*Canis lupus*) packs in the western Italian Alps. Molecular Ecology 11:857–868.

Lukacs, P. M., and K. P. Burnham. 2005. Review of capture-recapture methods applicable to noninvasive genetic sampling. Molecular Ecology 14:3909–3919.

Marucco, F., L. Boitani, D. H. Pletscher, and M. K. Schwartz. 2011. Bridging the gaps between non-invasive genetic sampling and population parameter estimation. European Journal of Wildlife Research 57:1–13.

Marucco, F., D. H. Pletscher, L. Boitani, M. K. Schwartz, K. L. Pilgrim, and J.-D. Lebreton. 2009. Wolf survival and population trend using non-invasive capture-recapture techniques in the Western Alps. Journal of Applied Ecology 46:1003–1010.

Miller, C. R., P. Joyce, and L. P. Waits. 2002. Assessing allelic dropout and genotype reliability using maximum likelihood. Genetics 160: 357–366.

Miller, C. R., P. Joyce, and L. P. Waits. 2005. A new method for estimating the size of small populations from genetic mark-recapture data. Molecular Ecology 14:1991–2005.

Mowry, R. A., M. E. Gompper, J. Beringer, and L. S. Eggert. 2011. River otter population size estimation using noninvasive latrine surveys. Journal of Wildlife Management 75:1625–1636.

Murphy, M. A., K. C. Kendall, A. Robinson, and L. P. Waits. 2007. The impact of time and field conditions on brown bear (*Ursus arctos*) faecal DNA amplification. Conservation Genetics 8:1219–1224.

Murphy, M. A., L. P. Waits, and K. C. Kendall. 2000. Quantitative evaluation of fecal drying methods for brown bear DNA analysis. Wildlife Society Bulletin 28:951–957.

Nsubuga, A. M., M. M. Robbins, A. D. Roeder, P. A. Morin, C. Boesch, and L. Vigilant. 2004. Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. Molecular Ecology 13:2089–2094.

Onorato, D., C. White, P. Zager, and L. P. Waits. 2006. Detection of predator presence at elk mortality sites using mtDNA analysis of hair and scat samples. Wildlife Society Bulletin 34:815–820.

Ostrander, E. A., G. F. Sprague, Jr., and J. Rine. 1993. Identification and characterization of dinucleotide repeat (CA)n markers for genetic mapping in dog. Genomics 16:207–213.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs 62:3–135.

Peakall, R., and P. E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology 6:288–295.

Pennell, M. W., C. R. Stansbury, L. P. Waits, and C. R. Miller. 2012. Capwire: a R package for estimating population census size from non-invasive genetic sampling. Molecular Ecology Resources 13:154–157.

Piggott, M. P., and A. C. Taylor. 2003. Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. Wildlife Research 30: 1–13.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Puechmaille, S. J., and E. J. Petit. 2007. Empirical evaluation of non-invasive capture-mark-recapture estimation of population size based on a single sampling session. Journal of Applied Ecology 44:843–852.

Roon, D. A., L. P. Waits, and K. C. Kendall. 2003. A quantitative evaluation of two methods for preserving hair samples. Molecular Ecology Notes 3:163–166.

Rudnick, J. A., T. E. Katzner, E. A. Bragin, O. E. Rhodes, and J. A. DeWoody. 2005. Using naturally shed feathers for individual identification, genetic parentage analyses, and population monitoring in an endangered eastern imperial eagle (*Aquila heliaca*) population from Kazakhstan. Molecular Ecology 14:2959–2967.

Ruell, E. W., S. P. D. Riley, M. R. Douglas, J. P. Pollinger, and K. R. Crooks. 2009. Estimating bobcat population sizes and densities in a fragmented urban landscape using noninvasive capture–recapture sampling. Journal of Mammalogy 90:129–135.

Santini, A., V. Lucchini, E. Fabbri, and E. Randi. 2007. Ageing and environmental factors affect PCR success in wolf (*Canis lupus*) excremental DNA samples. Molecular Ecology Notes 7:955–961.

Sawaya, M. A., T. K. Ruth, S. Creel, J. J. Rotella, J. B. Stetz, H. B. Quigley, and S. T. Kalinowski. 2011. Evaluation of noninvasive genetic sampling methods for cougars in Yellowstone National Park. Journal of Wildlife Management 75:612–622.

Stenglein, J. L., L. P. Waits, D. E. Ausband, P. Zager, and C. M. Mack. 2010*a*. Efficient, noninvasive genetic sampling for monitoring reintroduced wolves. Journal of Wildlife Management 74:1050–1058.

Stenglein, J. L., M. De Barba, D. E. Ausband, and L. P. Waits. 2010*b*. Impacts of sampling location within a faeces on DNA quality in two carnivore species. Molecular Ecology Resources 10:109–114.
Stenglein, J. L., L. P. Waits, D. E. Ausband, P. Zager, and C. M. Mack. 2011. Estimating gray wolf pack size and family relationships using noninvasive genetic sampling at rendezvous sites. Journal of Mammalogy 92:784–795.
U.S. Fish and Wildlife Service. 2011. Endangered and threatened wildlife and plants; reissuance of final rule to identify the Northern Rocky Mountain population of gray wolf as a distinct population segment and to revise the list of endangered and threatened wildlife. Federal Register 76:25590–25592.
Valière, N., L. Fumagalli, L. Gielly, C. Miquel, B. Lequette, M.-L. Poulle, J.-M. Weber, R. Arlettaz, and P. Taberlet. 2003. Long-distance wolf recolonization of France and Switzerland inferred from non-invasive genetic sampling over a period of 10 years. Animal Conservation 6:83–92.
Waits, L. P., G. Luikart, and P. Taberlet. 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. Molecular Ecology 10:249–256.
Waits, L. P., and D. Paetkau. 2005. Noninvasive genetic sampling tools for wildlife biologists: a review of applications and recommendations for accurate data collection. Journal of Wildlife Management 69:1419–1433.
Western Regional Climate Center. 2011. Historical climate information. http://www.wrcc.dri.edu. Accessed 29 Nov 2011.
Williams, B. W., D. R. Etter, D. W. Linden, K. F. Millenbah, S. R. Winterstein, and K. T. Scribner. 2009. Noninvasive hair sampling and genetic tagging of co-distributed fishers and American martens. Journal of Wildlife Management 73:26–34.
Wilson, G. J., and R. J. Delahay. 2001. A review of methods to estimate the abundance of terrestrial carnivores using field signs and observation. Wildlife Research 28:151–164.

Associate Editor: Kevin McKelvey.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.