Molecular Ecology (2007) **16**, 1823–1834 doi: 10.1111/j.1365-294X.2007.03270.x

Locating hybrid individuals in the red wolf (*Canis rufus*) experimental population area using a spatially targeted sampling strategy and faecal DNA genotyping

JENNIFER R. ADAMS,* CHRIS LUCASH,† LESLIE SCHUTTE† and LISETTE P. WAITS*
**Department of Fish and Wildlife, University of Idaho, College of Natural Resources, Moscow, Idaho 83844-1136, USA †Red Wolf Recovery Program, PO Box 1969, Manteo, North Carolina 27954, USA*

Abstract

Hybridization with coyotes (*Canis latrans*) continues to threaten the recovery of endangered red wolves (*Canis rufus*) in North Carolina and requires the development of new strategies to detect and remove coyotes and hybrids. Here, we combine a spatially targeted faecal collection strategy with a previously published reference genotype data filtering method and a genetic test for coyote ancestry to screen portions of the red wolf experimental population area for the presence of nonred wolf canids. We also test the accuracy of our maximum-likelihood assignment test for identifying hybrid individuals using eight microsatellite loci instead of the original 18 loci and compare its performance to the Bayesian approach implemented in NEWHYBRIDS. We obtained faecal DNA genotypes for 89 samples, 73 of which were matched to 23 known individuals. The performance of two sampling strategies — comprehensive sweep and opportunistic spot-check was evaluated. The opportunistic spot-check sampling strategy required less effort than the comprehensive sweep sampling strategy but identified fewer individuals. Six hybrids or coyotes were detected and five of these individuals were subsequently captured and removed from the population. The accuracy and power of the genetic test for coyote ancestry is decreased when using eight loci; however, nonred wolf canids are identified with high frequency. This combination of molecular and traditional field-based approaches has great potential for addressing the challenge of hybridization in other species and ecosystems.

Keywords: *Canis rufus*, faecal genotyping, hybrid detection, red wolf, sampling strategies

Received 19 August 2006; revision received 26 November 2006; accepted 11 December 2006

Introduction

The advent of new molecular techniques has resulted in the discovery of hybridization in many mammal and bird species (Grant & Grant 1992; Mallet 2005). In some cases, hybridization is thought to occur naturally (Arnold 1992) while in others hybridization may be the result of anthropogenic factors such as habitat modification, species eradication or species introduction (Rhymer & Simberloff 1996). When hybridization is believed to be the result of anthropogenic causes, the conservation goal is often to stop or reverse its progress. Hybridization can be a serious problem where populations of endangered species come into contact with a species that is abundant (Allendorf *et al*. 2001). Backcrossing of hybrid individuals to the abundant parental population may eventually result in the extinction of the endangered species through hybridization (Rhymer & Simberloff 1996). With no clear policy regarding protection of hybrid individuals under the Endangered Species Act or international legislation (Allendorf *et al*. 2001), the decision to intervene when hybridization occurs becomes critical to ensuring the continued existence of the species.

The use of molecular methods makes the identification of F_1 hybrids relatively straightforward for closely related mammalian species (Beaumont *et al*. 2001; Randi & Lucchini 2002; Pierpaoli *et al*. 2003; Vilà *et al*. 2003; Halbert *et al*. 2005; Lecis *et al*. 2006). The key to preventing introgression in populations of endangered species is the efficient detection and removal of F_1 hybrids (Miller *et al*. 2003),

Correspondence: Jennifer R. Adams, Fax: 208 8859080, E-mail: adam2483@uidaho.edu

which can be challenging for species that are wide-ranging and difficult to capture. Noninvasive genetic sampling is a promising alternative for locating hybrids within a population (Adams *et al*. 2003; Schwartz *et al*. 2004; Adams & Waits 2007). Noninvasive techniques are beneficial because individuals do not need to be captured and sampling can occur over a large geographical area. When combined with GIS technology, precise locations of hybrid individuals can be obtained, permitting a more focused approach to hybrid removal. This approach has been used to detect hybrids of endangered red wolves (*Canis rufus*) and coyotes (*Canis latrans*).

The red wolf was reintroduced to the wild in northeastern North Carolina in 1987 (USFWS 1989; Phillips *et al*. 2003). As red wolf numbers increased and their range expanded westward, they began encountering coyotes in the experimental population area (Phillips *et al*. 2003). Hybridization between red wolves and coyotes was confirmed by the United States Fish and Wildlife Service (USFWS) in 1999 and was quickly determined to be the biggest threat to the recovery of the red wolf (Kelly *et al*. 1999). Since then, recovery efforts have focused extensively on the removal of hybrid and coyote individuals from the experimental population area (Phillips *et al*. 2003; Stoskopf *et al*. 2005). This has been facilitated by the development of a genetic 'test' that detects coyote ancestry in newly captured individuals (Miller *et al*. 2003). However, the application of this test to the 6000-km^2 experimental population area has been limited by the number of individuals that can be live-captured and sampled using traditional techniques. Thus, effective methods for screening the experimental population area for coyotes and hybrids are needed (Stoskopf *et al*. 2005; Adams & Waits 2007).

Previously, we developed a faecal DNA reference genotype data filtering method and applied it to faecal samples (scats) collected in the Alligator River National Wildlife Refuge, North Carolina to locate new, potentially hybrid, individuals (Adams & Waits 2007). Here, we combine the use of a spatially targeted scat collection strategy with the reference genotype data filtering method and a likelihood-ratio assignment based approach for detecting coyote ancestry [canid assignment test (CAT), Miller *et al*. 2002] to locate hybrid and coyote individuals. This combined approach to hybrid detection is applied to scat samples collected throughout the red wolf experimental population area. The goals of this research are to (i) evaluate the success of this approach in locating hybrid and coyote individuals, (ii) assess the ability of two sampling strategies with different levels of effort to detect new unsampled individuals, and (iii) compare the accuracy and power of the canid assignment test and the Bayesian model implemented in NEWHYBRIDS to discriminate coyotes, F_1 hybrids, F_2 hybrids and backcrosses using eight loci of microsatellite data.

Methods

Study area

The red wolf experimental population area (6000 km^2) is situated on a peninsula in northeastern North Carolina and includes five counties: Beaufort, Dare, Hyde, Tyrrell, and Washington (Phillips *et al*. 2003, Fig. 1). The habitat includes a mixture of land developed for agriculture, partially drained pocosin wetlands and native pocosin wetlands (Richardson 1991). Many roads cover the experimental population area (Fig. 1); however, the majority of the roads are gravel or dirt and serve as access roads to agricultural fields. At the time of this study, there were 86 radiocollared individuals occupying 30 sites across the experimental population area (Fig. 1). Radiotelemetry flights occurred one to two times a week, and nine packs were located almost daily using ground radiotelemetry. USFWS field biologists also continually screened unoccupied areas for the presence of new individuals through the use of sign surveys (scats and footprints).

Sample collection

Scats were collected in March and April 2003 on roads throughout the experimental population area (Fig. 1). The goal of scat collection was to detect known, collared individuals and to locate new, noncollared hybrids, coyotes or red wolves. To minimize the chance of missing a new individual, sample collectors were instructed to collect all scats that could have originated from a large canid, regardless of sample age and condition, and to record suspected species of origin.

Two sampling strategies were employed: a comprehensive sweep strategy and an opportunistic spot-check strategy. In comprehensive sweeps, all of the roads within an area were searched using all-terrain vehicles. Comprehensive sweeps were employed within four areas occupied by known individuals (occupied areas) to monitor for the presence of new individuals (Table 1, Fig. 2). Comprehensive sweeps were also employed within four areas of concern (Table 1, Fig. 2). An area of concern is an area in which a red wolf pair does not currently reside but which contains habitat that could support a red wolf, hybrid or coyote pair. Finally, opportunistic spot-checks were performed using trucks during trapping or sign surveys in two occupied areas and five areas of concern (Table 1, Fig. 2). Opportunistic spot-checks occurred where USFWS field biologists had observed signs that an unknown individual was present. Scats were collected as in Adams *et al*. (2003) and their locations were recorded using hand-held GPS units. Scats were transferred to 95% ethanol solution in a 1 : 4 scat to ethanol ratio within 24 h of collection to

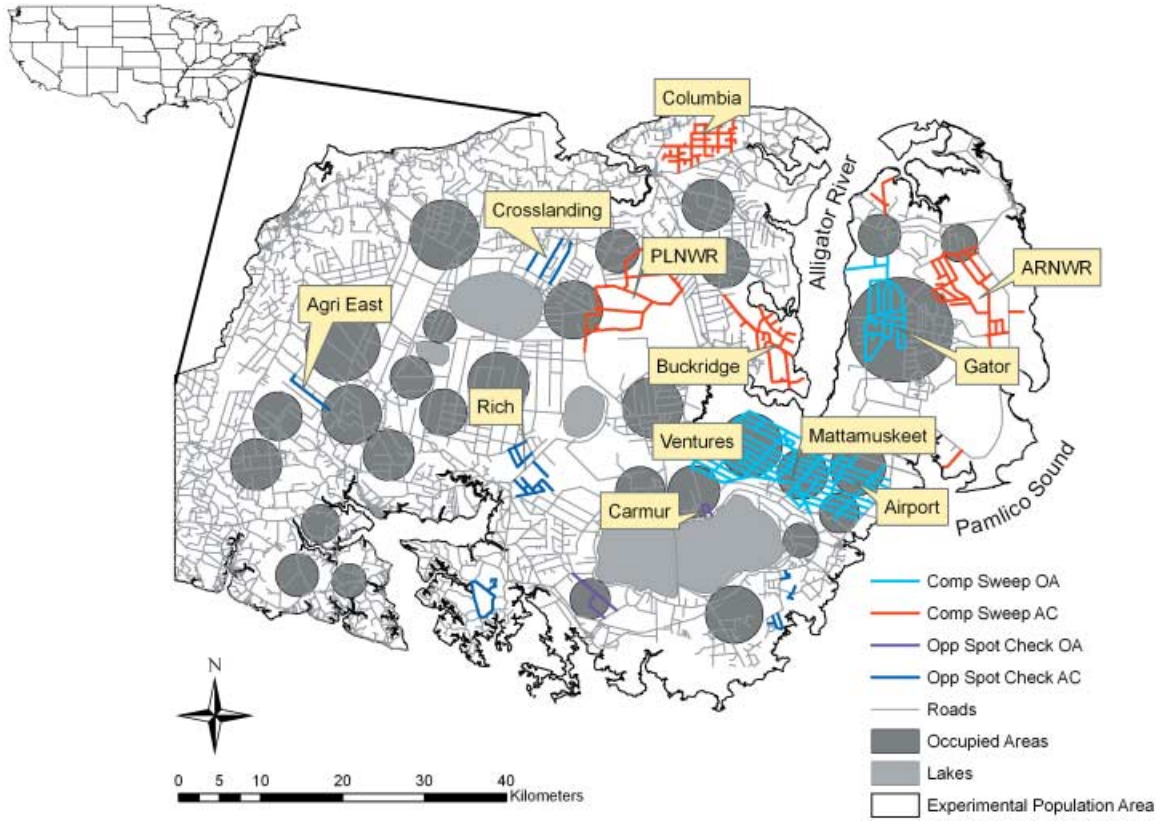

Fig. 1 Occupied areas within the red wolf experimental population area as of March 2003 and the locations and roads searched using each sampling strategy. Abbreviations: Comp, comprehensive; Opp, opportunistic; OA, occupied areas; AC, areas of concern.

preserve the DNA until extraction (Murphy *et al*. 2002; Adams *et al*. 2003).

To assess differences in effort required between the comprehensive sweep and opportunistic spot-check sampling strategies, the total length (km) of roads sampled was estimated using the XTools extension (Oregon Department of Forestry) in ArcView 3.2 (ESRI). A record was kept of the roads searched during the comprehensive sweep surveys; however, no records exist of the roads sampled using the opportunistic spot-check strategy. To generate an estimate of the distance searched during opportunistic spot-check sampling, we assumed that the entire length of the road on which the scat was found had been searched. In a few cases justified by field data notes, access roads were considered to have been searched as well.

DNA extraction and amplification

DNA was extracted from all scat samples using a QIAGEN stool kit protocol. To avoid contamination, all scat extractions were performed in a laboratory containing no concentrated forms of DNA. Contamination of extraction reagents was monitored using negative controls. The microsatellite loci and the polymerase chain reaction (PCR) conditions are described in detail in Adams & Waits (2007). Briefly, they include the amplification of seven loci (AHT103, Holmes *et al*. 1995; CXX.172, C09.173, CXX.20, CXX.225, Ostrander *et al*. 1993; FH2054, Mellersh *et al*. 1997; CXX.403, Ostrander *et al*. 1995) in four multiplexes using Ampli*Taq* Gold DNA polymerase (Applied Biosystems) and 55 PCR cycles. An eighth locus (CXX.200, Ostrander *et al*. 1993) was added to multiplex three (0.6 μM of each primer) from Adams & Waits (2007) to increase the resolving power of the canid assignment test (Miller *et al*. 2003) when identifying hybrid or coyote individuals. Fragments were separated by size on an ABI 377 sequencer and data were analysed using GENESCAN 3.1.2 (Applied Biosystems) and GENOTYPER 2.5 (Perkin Elmer).

Individual identification

The expected probability of identity (PID_{THEO}), observed probability of identity (PID_{OBS}) and the expected probability of identity between siblings (PID_{SIBS}) were calculated according to Waits *et al*. (2001) from 175 wild red wolves (unpublished) for the eight loci together, locus

Table 1 The sampling strategy used (Samp), number of scat samples collected (# scats), scat genotypes detected (Scat G), known individuals (KI), and new individuals (NI) identified within each area sampled

Area	Km*	Area	Samp	# scats	Scat G	KI	NI
Gator	71	OA	CSW	84	33	3	0
Airport	85†	OA	CSW	38	12	4	1
Mattamuskeet	85†	OA	CSW	91	11	2	2
Ventures	85†	OA	CSW	92	11	2	0
PLNWR	52	AC	CSW	30	8	6	1‡
Columbia	76	AC	CSW	26	0	0	0
Buckridge	45	AC	CSW	20	0	0	0
ARNWR	89	AC	CSW	72	4	3	0
Carmur	4	OA	OSC	2	2	1	0
Swindell	11	OA	OSC	13	2	2	0
Agri East	7	AC	OSC	7	1	0	1
Englehardt	18	AC	OSC	10	2	0	2§
Crosslanding	15	AC	OSC	3	0	0	0
Rich	25	AC	OSC	6	1	0	1‡
Scranton	14	AC	OSC	7	2	0	2
Total	682		Total	501	89	23	10

Abbreviations: OA, occupied area; AC, area of concern; CSW, comprehensive sweep; OSC, opportunistic spot-check.
*Distances were rounded to the nearest kilometre.
†There are no clearly delineated boundaries between these packs so the length of road sampled was divided equally between the three.
‡Fox.
§One fox, one dog.

CXX.200 (PID values for the other loci can be found in Adams & Waits 2007) and each combination of seven and six loci used. For locus CXX.200, the probability of identity (PID_{THEO}) was 0.30, the PID_{OBS} was 0.34, and the PID_{SIBS} was 0.55. Because the main goal of this study was to detect new hybrid or coyote individuals, we also calculated PID_{THEO}, PID_{OBS} and PID_{SIBS} for hybrids ($n = 80$) and coyotes ($n = 53$) captured in the experimental population area.

Data filtering to detect recaptures of known individuals

Scat genotypes were screened using the reference genotype data filtering method of Adams & Waits (2007). Briefly, homozygous results were accepted after two replicates, and each allele had to be observed twice for heterozygous results. Scat genotypes containing six or more loci of data were then compared to a reference genotype database of individuals captured in the red wolf experimental population area (unpublished) using the program GIMLET (Valiere 2002). A one-allele, one-locus mismatch was allowed between scat genotypes and reference genotypes if it occurred at the seventh or eighth locus of data, at least six loci matched and the mismatch could have been the result of allelic dropout.

Genotyping errors

The genotyping error rate per locus was calculated as in Broquet & Petit (2004). The genotyping error rate consisted of allelic dropout events observed at heterozygous loci and the number of false alleles detected at both homozygous and heterozygous loci. A false allele was classified as the amplification of a third allele at a heterozygous locus or the amplification of a second allele at a homozygous locus. Correct genotypes were determined by comparison to known individuals in our reference database, or through repeated amplifications of newly captured individuals until a reliability score of ≥ 95% was reached (Miller *et al*. 2002). We tallied the total number of PCR amplifications performed per locus to determine the average number of PCR replicates per locus needed to obtain the data.

Hybrid identification and DNA sequencing

Scat genotypes that did not match any individual in the reference genotype database were replicated further under the maximum-likelihood ratio (MLR) method of Miller *et al*. (2002) until a 95% reliability score was obtained. These genotypes were then matched to one another using GIMLET (Valiere 2002). All unique scat genotypes were classified as having originated from new, undocumented individuals. Because the microsatellite loci used were originally developed for dogs (Ostrander *et al*. 1993) and can cross-amplify in many canid species (Gottelli *et al*. 1994; Girman *et al*. 2001; Ralls *et al*. 2001; Wandeler *et al*. 2003; Iyengar *et al*. 2005; Kitchen *et al*. 2005), all new individuals were sequenced to determine the species of origin. These samples were amplified using the canid-specific ScatSeq primers (control region, mitochondrial DNA) as in Adams *et al*. (2003). In addition, 62 samples that failed in microsatellite analysis but were collected in areas of concern were amplified and sequenced at a portion of the cytochrome *b* region (Farrell *et al*. 2000) to determine if any red wolf, hybrid or coyote were missed. PCR products were purified using ExoSAP-It (USB) and then sequenced using BigDye version 3.1 (ABI). Final sequencing products were purified further using a Sephadex (Sigma-Aldrich) clean-up step. Species of origin was determined through a BLAST search on the NCBI website (www.ncbi.nlm.nih.gov). As noted in Adams *et al*. (2003), red wolves have a unique control region haplotype that is not shared by coyotes.

Scats that contained mitochondrial DNA (mtDNA) haplotypes consistent with either red wolves or coyotes were analysed using the canid assignment test (CAT) to determine whether they were red wolves, coyotes or hybrids (Miller *et al*. 2003). The canid assignment test is a maximum-likelihood-assignment-based approach that uses nuclear genotype data to assign an individual to one of six

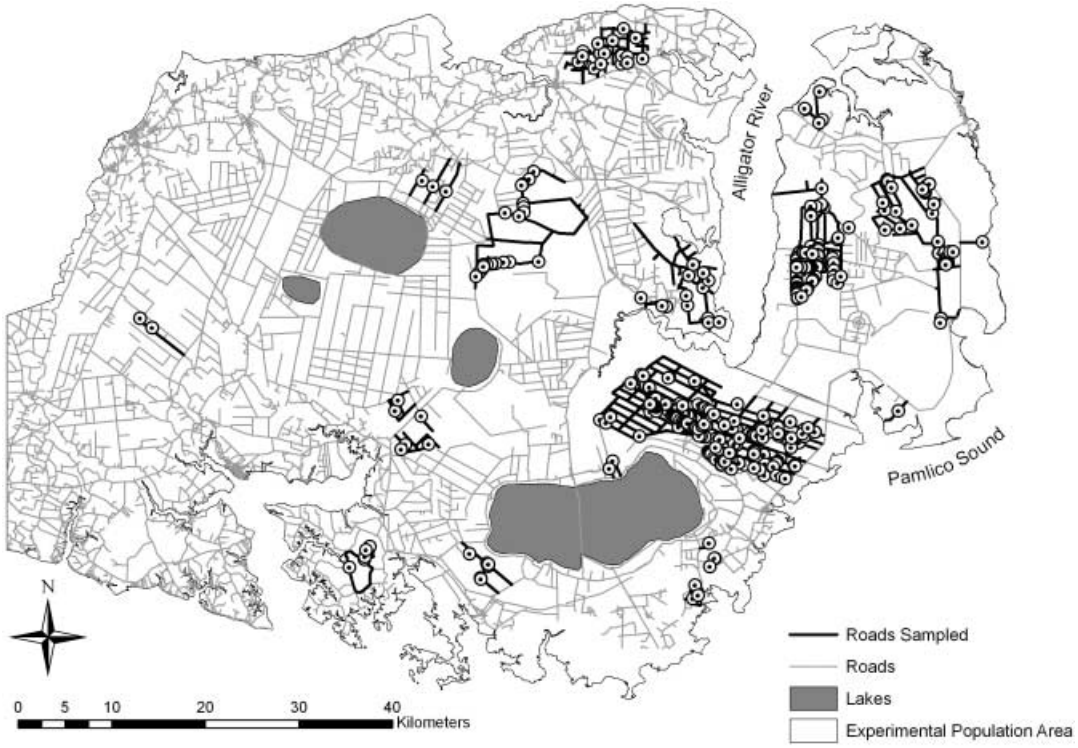

Fig. 2 The roads searched within the red wolf experimental population area and the locations of 501 collected scat samples.

ancestry categories [red wolf, 75% red wolf, 50% red wolf (F_1), 50% red wolf (F_2), 25% red wolf and coyote]. The first step in the CAT is to estimate the likelihood of observing the genotype in each of the six ancestry groups based on allele frequencies and the presence of coyote specific alleles (CSA). A CSA is an allele that is found in coyotes, but does not occur in the red wolf founders; presence in a red wolf can either be the result of a mutation event or a hybridization event (Miller *et al*. 2003). After determining the most likely ancestry category, a likelihood-ratio test is used to evaluate the ability to reject alternative ancestry categories (Miller *et al*. 2003). For additional details on the theory and implementation of the CAT, please see Miller *et al*. (2003). The locations of all red wolf, coyote and hybrid individuals were then placed on a map of the experimental population area using ArcView 3.2 (ESRI).

Because the CAT was developed for use with 18 microsatellite loci, it is likely that using only eight loci will affect the ability of the test to accurately discriminate between the different ancestry categories. To assess this, 1000 individuals were simulated for each ancestry category (red wolf, 75% red wolf, 50% red wolf (F_1), 50% red wolf (F_2), 25% red wolf and coyote) using the program HYBRIDLAB (Nielsen *et al*. in press). Genotypes from 40 red wolves present in the experimental population area during scat sampling and 40 coyotes captured from 1999 to 2003 were used to initiate the simulations. First F_1 hybrid individuals were simulated. Two groups of 40 individuals were then randomly selected from these F_1 individuals for use in generating the 75% red wolf, 50% red wolf (F_2) and 25% red wolf individuals. The 6000 simulated individuals were then analysed using the CAT (Miller *et al*. 2003).

In addition, we tested the performance of an alternative Bayesian model-based approach to detecting hybrids as implemented in the program NEWHYBRIDS (Anderson & Thompson 2002). Under this approach, the posterior probabilities of individuals belonging to one of six genotype frequency classes are calculated (Anderson & Thompson 2002). The unit-information prior was used for both the mixing proportions and allele frequencies because there are several informative alleles that are at low frequencies in both species. The allele frequency distributions for red wolves and coyotes used in the CAT were used as allele frequency priors. The MCMC was run for 100 000 iterations after a burn-in period of 10 000 iterations. If an individual had a posterior probability of belonging to a specific genotype frequency class of 80% or higher, then it was scored as belonging to that particular class (Pierpaoli *et al*. 2003). If an individual had a posterior probability less than 80% for any genotype frequency class, then it was scored as belonging to multiple classes (those classes with a posterior probability of greater than 10%).

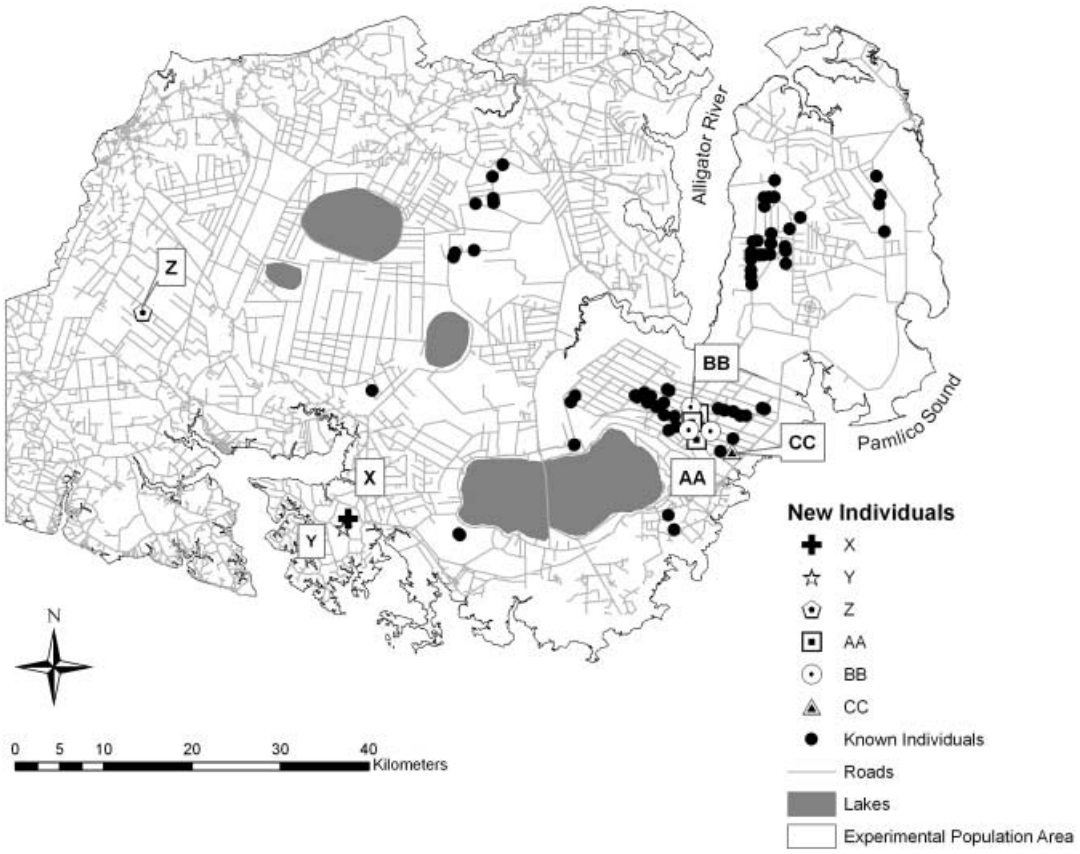

Fig. 3 The locations of the 89 scats for which a 6–8 locus microsatellite genotype was obtained. The six new hybrid and coyote individuals (V through AA) are highlighted.

Results

Sampling strategies

A total of 682 km of road was searched for scats during the course of this study (Table 1). An average of 82 km of road per occupied area and 66 km per unoccupied area of concern was searched using the comprehensive sweep sampling strategy (Table 1). Conversely, an average of 8 km of road per occupied area and 16 km per area of concern was searched using the opportunistic spot-check sampling strategy (Table 1). Using the comprehensive sweep strategy, four new individuals were detected in occupied areas while no new individuals were detected in areas of concern (Table 1). Using the opportunistic sampling strategy, no new individuals were detected in occupied areas and six new individuals were detected in areas of concern (Table 1).

Amplification success and genotyping errors

A sample of 501 scats was collected from across the experimental population area, 320 in occupied areas and 181 in areas of concern (Table 1, Fig. 2). All 501 scats were analysed to minimize the possibility that new individuals would be missed. Based upon faecal sample size and morphology, 326 scats were believed to have originated from a large canid as classified by sample collectors. We obtained genotypes from 89 scats for an amplification success rate of 27.9% (Fig. 3). Genotyping error rates per locus ranged from 15.3 to 61.1% with an average genotyping error rate of 35.2% (Table 2). When a genotyping error occurred, it was more likely to be allelic dropout ($n = 316$) than a false allele ($n = 11$). The average number of PCR amplifications per locus ranged from 2.7 to 4.0 (Table 2). The overall average number of PCR amplifications per sample per locus was 3.3 (Table 2).

Probability of identity

We obtained all eight loci of data for 67 faecal samples. The overall observed probability of identity (PID_{OBS}) and the probability of identity for siblings (PID_{SIBS}) for the eight loci used in this study were 0.000066 and 0.0040, respectively. Scat samples with genotypes based on seven loci ($n = 16$) and six loci ($n = 6$) were allowed to remain in the data set. For the combinations of seven loci, the PID_{OBS} ranged from 0.000065 to 0.000013 and the PID_{SIBS} ranged

Table 2 The size in base pairs, number of PCR amplifications (# PCR), correct PCR amplifications (Correct), allelic dropout events (Dropout), false alleles (False), the genotyping error rate (Error rate) and the average number of PCR amplifications per locus per sample (Avg amps) for the eight loci used

Locus	Size (bp)	# PCR	Correct	Dropout	False	Error rate	Avg amps
225	162–166	72	28	44	6	61.1%	3.9
403	267–287	140	83	57	4	40.7%	3.4
103	72–80	124	105	19	0	15.3%	2.7
20	122–132	110	52	58	0	52.7%	3.9
2054	145–161	114	74	40	1	35.1%	2.9
173	97–111	97	81	16	0	16.5%	2.9
200	214–224	169	101	68	0	40.2%	4.0
172	142–156	70	56	14	0	20.0%	2.7
Total		896	580	316	11	35.2%	3.3

from 0.0073 to 0.016. For the combinations of six loci, the PID_{OBS} ranged from 0.00033 to 0.00059 and the PID_{SIBS} ranged from 0.014 to 0.016. For all eight loci, the PID_{OBS} and PID_{SIBS} for hybrid individuals were 3.5×10^{-7} and 0.0017 and for coyote individuals were 7.5×10^{-10} and 0.00043. These loci should have sufficient discriminatory power as 80–100 red wolves and less than 15 coyotes or hybrids are believed to reside in the experimental population area.

Individual identification

Seventy-three scat genotypes matched 23 known individuals with the number of detections per individual ranging from 1 to 23 (Table 3). Forty-six scat genotypes matched a known individual at eight loci. Nine scat genotypes matched a known individual at seven loci (no data were obtained at the eighth locus), and six scat genotypes matched a known individual at six loci (no data were obtained for the seventh and eighth locus). Twelve scat genotypes matched a known individual at seven loci with a homozygous mismatch at the eighth locus. There were 12 known individuals detected in occupied areas and 11 known individuals detected within areas of concern (Table 1). The remaining 16 scat genotypes were determined to have originated from 10 new individuals (Table 4).

Table 3 The sex, age (adult, 2 years old; juvenile, between 1 and 2 years old) and number of scat genotypes obtained (# detections) for each known individual sampled

Individual	Sex	Age	# detections
A	M	Adult	23
B	M	Adult	9
C	M	Adult	1
D	F	Juvenile	1
E	F	Adult	1
F	F	Adult	6
G	M	Adult	5
H	F	Adult	1
I	M	Adult	2
J	M	Juvenile	3
K	F	Adult	2
L	M	Adult	5
M	M	Juvenile	1
N	M	Adult	2
O	M	Adult	1
P	M	Adult	2
Q	F	Juvenile	1
R	M	Juvenile	2
S	F	Adult	1
T	M	Adult	1
U	M	Juvenile	1
V	M	Juvenile	1
W	F	Juvenile	1

Table 4 For the new individuals detected, the number of samples (# scat), mitochondrial DNA haplotype, canid assignment test (CAT) results at eight loci (nonexcluded ancestry categories in parentheses) and the number of coyote specific alleles (CSAs) attributed to each individual. For the individuals captured and removed the date of capture, CAT results at 18 loci and number of CSAs are given

Ind	# scat	mtDNA	CAT eight loci	CSAs	Date	CAT 18 loci	CSAs
X	1	Coyote	Coy (25% RW)	5	11/04	25% RW	8
Y	1	Coyote	Coy (25% RW)	5	03/03	25% RW	10
Z	1	Coyote	25% RW (Coy, F_2)	4	12/03	25% RW	12
AA	3	Coyote	F_2	4	12/03	25% RW	12
BB	5	Red wolf	F_2	4	06/03	F_2	4
CC	1	Coyote	25% RW (Coy, F_2)	3			
DD	1	Red fox					
EE	1	Dog					
FF	1	Red fox					
GG	1	Red fox					

Table 5 Accuracy of the CAT (Miller *et al*. 2003) and NEWHYBRIDS (Anderson & Thompson 2002) using eight loci of microsatellite data

True identity	D TID*	TID ML†	TID NE‡	TID E§	Incorrect RW¶
Red wolf					
CAT	53.8	39.3	0	6.9	
NEWHYBRIDS	83.6		16	0.4	
Coyote					
CAT	36.1	28.5	16.2	19.2	0.1
NEWHYBRIDS	57.9		40.7	1.4	0
F_1					
CAT	23.2	36.5	17.1	23.2	5.5
NEWHYBRIDS	9.6		87.5	2.9	0.2
75% RW					
CAT	12.0	22	30.1	35.9	14.3
NEWHYBRIDS	0.2		83	16.8	10.6
25% RW					
CAT	5.0	26	37.5	31.5	0.8
NEWHYBRIDS	0.8		82.2	17	0.1
F_2					
CAT	12.2	19.3	39.4	29.1	8.6
NEWHYBRIDS	3.8		88.3	7.9	0.6

*The percentage of time the true identity was definitively assigned.
†The percentage of time the true identity was the most likely category, but other categories could not be excluded for the CAT.
‡The percentage of time the true identity was not the most likely category, but could not be excluded as a possibility.
§The percentage of time the true identity was excluded as a possibility.
¶The percentage of time the true identity was miscategorized as red wolf.

Hybrid detection and DNA sequencing

Mitochondrial DNA sequencing of the 10 new individuals revealed that one individual was a dog and three individuals were red foxes. Of the remaining six new individuals, one had the red wolf haplotype and five had coyote haplotypes (Table 4). The canid assignment test indicated that individuals AA and BB had 50% coyote and 50% red wolf ancestry and were F_2 hybrids (all other identities excluded). Z and CC were most likely 25% red wolf and the coyote and F_2 identities could not be excluded. X and Y were most likely coyotes, but the 25% red wolf identity could not be excluded (Table 4). Three of the new individuals were found within occupied areas while the other three were found in areas of concern (Fig. 3).

Sequencing of a cytochrome *b* fragment for the 62 samples that were collected from areas of concern, but failed in the microsatellite analyses, revealed 36 bobcats (*Felis rufus*), five grey foxes (*Urocyon cinereoargenteus*), three red foxes (*Vulpes vulpes*), three dogs, three rabbits (Leporidae), two black bears (*Ursus americanus*), two rats (Muridae) and one sample with a sequence that matched both coyotes and red wolves (seven samples failed to amplify). For the sample with an mtDNA sequence matching either a coyote or red wolf, we were able to generate five loci of microsatellite data but the genotype did not make it through the data filter. Further investigation of this partial genotype revealed it was a close match to a red wolf female which later became the breeding female in Columbia.

CAT assessment

Using eight vs. 18 loci of data affects the ability of the CAT to differentiate between different ancestry categories (Table 5). The major effect appears to be a decrease in the ability of the CAT to make a definitive assignment as alternate categories cannot be excluded (Table 5). Another important error occurs within the 75% red wolf category where, with eight loci, the CAT incorrectly assigns 14.3% of the individuals to the red wolf category (Table 5). The percentage of time individuals are misassigned to the red wolf category drops to 0.1–8.6% for the other four ancestry categories (Table 5).

The program NEWHYBRIDS makes definitive calls with greater frequency than the CAT for the red wolf and coyote categories (Table 5). However, NEWHYBRIDS is less able than the CAT to make definitive assignments for all four hybrid categories (Table 5). Another difference is NEWHYBRIDS excludes the true category with less frequency than the CAT but rather than making a definite call, it lumps the true identity in with other possible categories (Table 5). Finally, NEWHYBRIDS identifies red wolves as hybrid or coyote with less frequency than the CAT (Table 5).

Discussion

Sampling strategies

Hybridization is a threat to native species around the globe (Rhymer & Simberloff 1996; Allendorf *et al*. 2001), and new methods are needed to detect hybrids and curb hybridization. These methods may include noninvasive genetic sampling approaches, and need to be both time- and cost-efficient. In this study, we employed two different scat sampling methods, a comprehensive sweep strategy and an opportunistic spot-check strategy to locate coyote and hybrid faecal samples in the 6000-km² red wolf experimental population area. Direct comparisons between the two strategies are not possible because they were not performed in identical spatial areas. However, we are able to evaluate the costs and benefits associated with each strategy. The comprehensive sweep strategy should increase the probability of detecting all individuals in an area because all roads are thoroughly searched. In this study, all known red wolves were detected within the

areas they normally occupy. However, this approach utilizes many resources, as it requires launching a specific effort and coordinating multiple individuals and pieces of equipment.

In contrast, the opportunistic spot-check strategy is less likely to sample every individual within an area, but it requires only one person collecting samples during the course of normal field activities such as sign surveys or checking trap lines. Thus, it is much more time-efficient for field researchers. It is possible that we underestimated the length of roads sampled using the opportunistic spot-check sampling strategy. Yet, if the length of the roads searched under the opportunistic spot-check strategy were doubled, it would still only represent one-third of the length of road searched using the comprehensive sweep strategy (Table 1). Thus, there is still a large difference in the effort between the two strategies. Furthermore, the samples were collected during normal field activities and did not require additional effort to obtain. In addition, the turnaround time for DNA analysis is faster using the spot-check strategy, as fewer samples need to be processed (48 vs. 453, Table 1). The two sampling strategies did differ in the total number of individuals detected (comprehensive sweep = 23 and opportunistic spot-check = 6, Table 1), but both sampling strategies successfully detected three unknown hybrid or coyote individuals. No new red wolf individuals were detected using either sampling strategy (Table 4). This is likely due to the fact that most new red wolf individuals are sampled for genetic material and marked while they are still puppies in the den.

For the red wolf recovery program, the opportunistic spot-check sampling strategy may provide the most efficient method for detecting new individuals. It requires fewer resources than the comprehensive sweep sampling strategy (i.e. one person and truck vs. multiple people and all-terrain vehicles). Furthermore, sampling effort can be spread across the entire experimental population area, thereby increasing the probability of encountering hybrid or coyote individuals. The opportunistic spot-check sampling strategy will ultimately allow the red wolf program to spatially focus management efforts, such as extensive trapping, on areas where hybrid and coyote individuals exist.

There were five areas of concern where no large canids were detected; two were sampled using the comprehensive sweep strategy (Columbia and Buckridge) and three were sampled using the opportunistic spot-check strategy (Englehardt, Crosslanding and Rich; Fig. 1, Table 1). There are two plausible explanations for this result. One explanation is that there were large canids present in these areas but DNA failed to amplify from their scats. The other is that there were no large canids in these areas and all scats collected were from other species. To differentiate between these two explanations, we successfully sequenced 55 of the 62 samples from which microsatellites failed to amplify. Our results revealed that most samples (95%) were other carnivores and only one scat originated from a red wolf or coyote. Additional microsatellite analyses reveal that this individual from the Columbia area was likely a known female red wolf. Therefore, in the Buckridge, Englehardt, Crosslanding and Rich areas (Fig. 1), no large canids were detected using faecal genotyping and our data screening approach because few or no large canids were present. In Columbia, however, one red wolf and three dogs were missed using faecal genotyping likely due to poor DNA quality.

Amplification success and genotyping errors

The amplification success rate reported in this study for an eight-locus genotype (27.9%) is much lower than the rate (50.0%) we reported in Adams & Waits (2007). This is likely due to a number of different factors. First, the data set collected in 2000 (Adams & Waits 2007) was screened using mtDNA sequencing to detect canid scats while this study made no attempt to remove nonred wolf scats prior to microsatellite amplification because we decided it was more time and cost efficient to proceed directly to microsatellite analysis. Based on mtDNA results from faecal samples collected in 2000, we would estimate that ~20% of our samples did not originate from a large canid (Adams *et al*. 2003). Removal of these samples prior to genetic analysis would have increased our success rates. Second, success rates in this study may also have also been depressed by inclusion of more older, degraded samples in our current sampling approach (see below).

The range of genotyping errors reported in this study (15.3% to 61.1%) is slightly higher than the range we reported for the same loci (6.6% to 52.1%) in Adams & Waits (2007). This may be explained by three main differences in sample collection and DNA analysis between the 2000 data set (Adams & Waits 2007) and this one. First, scat samples were collected in late spring (April and May) for the 2000 data set while samples were collected in early spring (March and April) for this data set. Differences in climatic conditions between seasons may have had an effect on DNA quality in faecal samples (Lucchini *et al*. 2002; Maudet *et al*. 2004; Piggott 2004). The second difference relates to DNA extraction. A modified QIAGEN tissue protocol was used in 2000 (Adams *et al*. 2003) while the QIAGEN stool kit was used in this data set. Other studies have demonstrated that extraction method can have an effect on the quality and quantity of DNA obtained from faecal samples (Wasser *et al*. 1997; Flagstad *et al*. 1999; Piggott & Taylor 2003).

The final difference that may have affected success rates and error rates in the two data sets involves our sampling approach. Multiple studies have shown that success rates

decrease and error rates increase as a function of time that a faecal sample remains in the field (Lucchini *et al*. 2002; Piggott 2004; Murphy *et al*. in press). In the 2000 data set, samples were collected from the same locations three times over 42 days (Adams *et al*. 2003). This resulted in relatively fresh samples being collected in the second and third sampling periods. In contrast, all samples collected in this data set were collected in a single sweep and could have included many older samples. In fact, our field notes suggest that 61% of scats believed to be large canid were greater than 2 weeks old.

The increase in genotyping errors in this data set affected the average number of PCR amplifications per locus per sample. In the 2000 data set, the average was 2.4 (Adams & Waits 2007) while in this data set the average was 3.3; more PCR replicates were required to ensure genotype quality. These higher error rates might also affect reliability scores for new genotypes obtained using the MLR method of Miller *et al*. (2002). The MLR method assumes equal genotyping error rates across loci and has been shown to be robust to moderate deviations of this assumption (Miller *et al*. 2002). Major deviations from this assumption (as observed in this study) could result in the reliability of individual genotypes being overestimated (C. Miller, personal communication) and lead us to overstate our confidence in the six new hybrid and coyote individuals detected. However, an additional line of evidence suggests the six new individuals did, in fact, exist. Five of the individuals (X, Y, Z, AA and BB, Table 4) were subsequently captured and removed from the population and their blood genotypes matched the scat genotypes at eight loci.

CAT assessment and hybrid detection

When using eight microsatellite loci, the CAT, as expected, does not perform as well as when using 18 loci of data (Table 5). The low success rates and high genotyping error rates observed in this data set, however, would make it financially unfeasible to use all 18 loci when analysing scats. Because of the effects that hybridization has on the red wolf population, the red wolf program decided it is less acceptable to misclassify nonred wolves as red wolves (Buddy Fazio, personal communication). In essence, the accidental loss of a few red wolves due to misclassification is better than letting hybrid and coyote individuals remain in the population. Although the CAT is affected by the use of fewer loci, hybrids would have a 15% or less chance of being classified as red wolves and coyotes would almost never be classified as red wolves (Table 5). In addition, it is the policy of the red wolf program to trap an area where a new individual has been identified using faecal genotyping, regardless of its ancestry, so a classification error by the CAT would be adjusted once the individual was captured and analysed at 18 loci (Arthur Beyer, personal communication). Although the resolution decreased when using eight loci with this data set, the CAT was still able to detect coyote ancestry as verified by the presence of multiple coyote specific alleles and the subsequent capture of five individuals and analysis of their genotypes at 18 loci (Table 4).

The model-based approach as implemented in NEWHYBRIDS (Anderson & Thompson 2002) appears to perform better than the CAT at assigning individuals to the red wolf and coyote ancestry categories and worse than the CAT at assigning individuals to the hybrid categories (Table 5). NEWHYBRIDS appears to suffer the same issue as the CAT that fewer loci provide less resolution between hybrid categories which has been documented by others (Vähä & Primmer 2006). A benefit of NEWHYBRIDS as applied to this study is it assigns red wolves as hybrids and coyotes with less frequency than does the CAT (Table 5). A potential drawback of NEWHYBRIDS is the amount of time needed to process the data. It required almost 48 h to analyse 6000 individuals using NEWHYBRIDS whereas it required about 3 h using the CAT. The usefulness of NEWHYBRIDS to other situations where hybridization is being examined using noninvasive genetic samples with fewer microsatellite loci will depend upon the goals of hybrid management. If the goal is to maintain a pure parental population one, as is the case with the red wolf population, where hybrids and members of parental population two are removed then NEWHYBRIDS performs quite well at identifying individuals where some introgression is present. If, however, the goal is to simply document the presence of backcrossing or determine how many individuals of each hybrid class are present in the population then neither the CAT nor NEWHYBRIDS would perform very well using only 8 loci. In this case it would be necessary to use more microsatellite loci (Vähä & Primmer 2006).

By combining scat genotyping and the canid assignment test (Miller *et al*. 2003), we located the presence of new individuals in the red wolf population and determined the percentage of coyote ancestry in those individuals. As a result, six new hybrid and coyote individuals were identified, five of which were subsequently captured and removed from the population. This new tool will facilitate the location of hybrid and coyote individuals and allow the red wolf program to identify areas in which to focus management efforts.

Conclusions

Hybridization is a documented threat to recovery efforts of other carnivores such as wildcats (Randi *et al*. 2001; Pierpaoli *et al*. 2003; Lecis *et al*. 2006), grey wolves (Roy *et al*. 1994; Randi & Lucchini 2002; Verardi *et al*. 2006), and Algonquin wolves (Kyle *et al*. 2006) and may be a threat to the persistence of Canadian lynx at the southern end of its

distribution in North America (Schwartz *et al*. 2004). New approaches to detecting hybrids are needed for assessment and monitoring of hybridization risk for these and other species. Here, we present a powerful and efficient approach for detecting hybrids using faecal genotyping combined with assignment test approaches and GPS technology. Application of this method to scat samples collected throughout the red wolf experimental population area resulted in the identification of six new hybrid and coyote individuals five of which were captured and removed. The traditional field-based methods were critical to the implementation of a management plan to curb hybridization but were more targeted and efficient when coupled with the noninvasive genetic sampling approach. The opportunistic spot check sampling strategy may be the optimal sampling strategy for red wolf program, as it requires less sampling effort than the comprehensive sweep sampling strategy. This novel approach has had multiple positive impacts on red wolf recovery efforts and will continue to be used as a tool to minimize hybridization between red wolves and coyotes.

Acknowledgements

We thank Buddy Fazio, project coordinator for the red wolf program, and the Red Wolf Recovery Implementation Team for their continued support of our research efforts. The Waits lab group and two anonymous reviewers provided helpful comments on this manuscript. Members of the red wolf field crew, Arthur Beyer, Scott McLellan, Michael Morse, and Kathy Whidbee, and program volunteers Chris Crowe, Kyle Krzywicki, Rob Lonsinger, Carrie Pomfrey assisted with scat collection. Bruce Creef and the ARNWR maintenance facility staff provided help with the all-terrain vehicles. Paul Joyce and Craig Miller assisted with simulation analyses. Dustene Cummings and Becky Ternes provided laboratory support. Eva Strand assisted with GIS analysis. The United States Fish and Wildlife Service provided funding for this project.

References

Adams JR, Kelly BT, Waits LP (2003) Using faecal DNA sampling and GIS to monitor hybridization between red wolves (*Canis rufus*) and coyotes. (*Canis latrans*). *Molecular Ecology*, **12**, 2175–2186.

Adams JR, Waits LP (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.

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids; setting conservation guidelines. *Trends in Ecology & Evolution*, **16**, 613–622.

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, **160**, 1217–1229.

Arnold ML (1992) Natural hybridization as an evolutionary process. *Annual Review of Ecology and Systematics*, **23**, 237–261.

Beaumont M, Barratt EM, Gottelli D *et al.* (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology*, **10**, 319–336.

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

Farrell LE, Roman J, Sunquist ME (2000) Dietary separation of sympatric carnivores identified by molecular analysis of scats. *Molecular Ecology*, **9**, 1583–1590.

Flagstad Ø, Roed K, Stacy JE, Jakobsen KS (1999) Reliable noninvasive genotyping based on excremental PCR of nuclear DNA purified with a magnetic bead protocol. *Molecular Ecology*, **8**, 879–883.

Girman DJ, Vila C, Geffen E *et al.* (2001) Patterns of subdivision, gene flow and genetic variability in the African wild dog. (*Lycaon Pictus*). *Molecular Ecology*, **10**, 1703–1723.

Gottelli D, Sillero-Zubiri C, Applebaum GD (1994) Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*. *Molecular Ecology*, **3**, 301–312.

Grant PR, Grant BR (1992) Hybridization of bird species. *Science*, **256**, 193–197.

Halbert ND, Ward TJ, Schnabel RD, Taylor JF, Derr JN (2005) Conservation genomics; disequilibrium mapping of domestic cattle chromosomal segments in North American bison populations. *Molecular Ecology*, **14**, 2343–2362.

Holmes NG, Dickend HF, Parker HL, Binns MM, Mellersh CS, Sampson J (1995) Eighteen canine microsatellites. *Animal Genetics*, **26**, 132–133.

Iyengar A, Babu VN, Hedges S *et al.* (2005) Phylogeography, genetic structure, and diversity in the dhole. (*Cuon Alpinus*). *Molecular Ecology*, **14**, 2281–2297.

Kelly BT, Miller PS, Seal US (1999) *Population and habitat viability assessment workshop for the red wolf (*Canis rufus*)*. Conservation Breeding Specialist Group (SSC/IUCN), Apple Valley.

Kitchen AM, Gese EM, Waits LP, Karki SM, Schauster ER (2005) Genetic and spatial structure within a swift fox population. *Journal of Animal Ecology*, **74**, 1173–1181.

Kyle CJ, Johnson AR, Petterson BR *et al.* (2006) Genetic nature of eastern wolves: past, present and future. *Conservation Genetics*, in press.

Lecis R, Pierpaoli M, Birò ZS *et al.* (2006) Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. *Molecular Ecology*, **15**, 119–131.

Lucchini V, Fabbri E, Marucco F *et al.* (2002) Noninvasive molecular tracking of colonizing wolf (*Canis lupus*) packs in the western Italian Alps. *Molecular Ecology*, **11**, 857–868.

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, **20**, 229–237.

Maudet C, Luikart G, Dubray D, Von Hardenberg A, Taberlet P (2004) Low genotyping error rates in wild ungulate faeces sampled in winter. *Molecular Ecology Notes*, **4**, 772–775.

Mellersh CS, Langston AA, Acland GM *et al.* (1997) A linkage map of the canine genome. *Genomics*, **46**, 326–336.

Miller CR, Adams JR, Waits LP (2003) Pedigree-based assignment tests for reversing coyote (*Canis latrans*) introgression into the wild red wolf (*Canis rufus*) population. *Molecular Ecology*, **12**, 3287–3301.

Miller CR, Joyce P, Waits LP (2002) Assessing allelic dropout and genotype reliability using maximum likelihood. *Genetics*, **160**, 357–366.

Murphy MA, Kendall KC, Robinson A, Waits LP (in press) The impact of time and field conditions on brown bear (*Ursus arctos*) faecal DNA amplification. *Conservation Genetics*.

Murphy MA, Waits LP, Kendall KC *et al.* (2002) Long-term preservation methods for brown bear (*Ursus arctos*) faecal DNA samples. *Conservation Genetics*, **3**, 435–440.

Nielsen EEG, Bach LA, Kotlick P (in press) HYBRIDLAB (version 1.0): a program for generating simulated hybrids from population samples. *Molecular Ecology Notes*.

Ostrander EA, Mapa FA, Yee M, Rine J (1995) One hundred and one new simple sequence repeat-based markers for the canine genome. *Mammalian Genome*, **6**, 192–195.

Ostrander EA, Sprague GF, Rine J (1993) Identification and characterization of dinucleotide repeat $(CA)_n$ markers for genetic mapping in dogs. *Genomics*, **16**, 207–213.

Phillips MK, Henry VG, Kelly BT (2003) Restoration of the red wolf. In: *Wolves: Behavior, Ecology and Conservation* (eds Mech LD, Boitani L), pp. 272–288. University of Chicago Press, Chicago, Illinois.

Pierpaoli M, Birò ZS, Herrmann M *et al.* (2003) Genetic distinction of wildcat (*Felis silvestris*) populations in Europe, and hybridization with domestic cats in Hungary. *Molecular Ecology*, **12**, 2585–2598.

Piggott MP (2004) Effect of sample age and season of collection on the reliability of microsatellite genotyping of faecal DNA. *Wildlife Research*, **31**, 485–493.

Piggott MP, Taylor AC (2003) Extensive evaluation of faecal preservation and DNA extraction methods in Australian native and introduced species. *Australian Journal of Zoology*, **51**, 341–355.

Ralls K, Pilgrim KL, White PJ *et al.* (2001) Kinship, social relationships and den sharing in kit foxes. *Journal of Mammalogy*, **82**, 858–866.

Randi E, Lucchini V (2002) Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) populations by Bayesian admixture analyses of microsatellite variation. *Conservation Genetics*, **3**, 31–45.

Randi E, Pierpaoli M, Beaumont M, Ragni B (2001) Genetic identification of wild and domestic cats (*Felis silvestris*) and their hybrids using Bayesian clustering methods. *Molecular Biology and Evolution*, **18**, 1679–1693.

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, **27**, 83–109.

Richardson CJ (1991) Pocosins: an ecological perspective. *Wetlands*, **11**, 335–354.

Roy MS, Geffen E, Smith D, Ostrander EA, Wayne RK (1994) Patterns of differentiation and hybridization in North American Wolflike Candids, revealed by analysis of microsatellite loci. *Molecular Biology and Evolution*, **11**, 553–570.

Schwartz MK, Pilgrim KL, McKelvey KS *et al.* (2004) Hybridization between Canada lynx and bobcats: genetic results and management implications. *Conservation Genetics*, **5**, 349–355.

Stoskopf MK, Beck K, Fazio BB *et al.* (2005) Implementing recovery of the red wolf — integrating research scientists and managers. *Wildlife Society Bulletin*, **33**, 1145–1152.

USFWS (1989) *Red Wolf Recovery Plan*. USFWS, Atlanta.

Vähä J-P, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, **15**, 63–72.

Valiere N (2002) GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes*, **2**, 377–379.

Verardi A, Lucchini V, Randi E (2006) Detecting introgressive hybridization between free-ranging domestic dogs and wild wolves (*Canis lupus*) by admixture linkage disequilibrium analysis. *Molecular Ecology*, **15**, 2845–2855.

Vilà C, Walker C, Sundqvist A-K *et al.* (2003) Combined use of maternal, paternal and bi-parental genetic markers for the identification of wolf-dog hybrids. *Heredity*, **90**, 17–24.

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.

Wandeler P, Funk SM, Largiadèr CR, Gloor S, Breitenmoser U (2003) The city-fox phenomenon: genetic consequences of a recent colonization of urban habitat. *Molecular Ecology*, **12**, 647–656.

Wasser SK, Houston CS, Koehler GM, Cadd GG, Fain SR (1997) Techniques for application of faecal DNA methods to field studies of Ursids. *Molecular Ecology*, **6**, 1091–1097.

Jennifer Adams is a PhD candidate studying hybridization between red wolves and coyotes in the southeast United States. Chris Lucash is a senior wildlife biologist with the USFWS red wolf program. Leslie Schutte is a wildlife biologist with the USFWS red wolf program whose interests include radiotelemetry and GIS. Lisette Waits focuses on the molecular ecology and conservation genetics of a variety of vertebrate taxa. She also codirects the Laboratory for Conservation and Ecological Genetics and the Center for Research on Invasive Species and Small Populations.
