Received: 29 June 2021 | Revised: 24 September 2021 | Accepted: 27 October 2021

DOI: 10.1111/1755-0998.13552

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

An efficient method for simultaneous species, individual, and sex identification via in-solution single nucleotide polymorphism capture from low-quality scat samples

Lillian D. Parker[1,2] | Michael G. Campana[1,2,3] | Jessica D. Quinta[1] | Brian Cypher[4] | Isabel Rivera[1] | Robert C. Fleischer[1] | Katherine Ralls[1] | Tammy R. Wilbert[1] | Ryan Boarman[5] | William I. Boarman[5] | Jesús E. Maldonado[1,2,3]

[1]Center for Conservation Genomics, Smithsonian Conservation Biology Institute and National Zoological Park, Washington, District of Columbia, USA

[2]School of Systems Biology, George Mason University, Fairfax, Virginia, USA

[3]Department of Environmental Science and Policy, George Mason University, Fairfax, Virginia, USA

[4]Endangered Species Recovery Program, California State University, Turlock, California, USA

[5]Conservation Science Research and Consultation, Spring Valley, California, USA

Correspondence
Lillian D. Parker, Center for Conservation Genomics, Smithsonian Conservation Biology Institute and National Zoological Park, Washington, DC, USA.
Email: lilly.parker@gmail.com

Funding information
This work was supported by the Desert Conservation Program of Clark County Nevada

Abstract

Understanding predator population dynamics is important for conservation management because of the critical roles predators play within ecosystems. Noninvasive genetic sampling methods are useful for the study of predators like canids that can be difficult to capture or directly observe. Here, we introduce the FAECES* method (Fast and Accurate Enrichment of Canid Excrement for Species* and other analyses) which expands the toolbox for canid researchers and conservationists by using in-solution hybridization sequence capture to produce single nucleotide polymorphism (SNP) genotypes for multiple canid species from scat-derived DNA using a single enrichment. We designed a set of hybridization probes to genotype both coyotes (*Canis latrans*) and kit foxes (*Vulpes macrotis*) at hundreds of polymorphic SNP loci and we tested the probes on both tissues and field-collected scat samples. We enriched and genotyped by sequencing 52 coyote and 70 kit fox scats collected in and around a conservation easement in the Nevada Mojave Desert. We demonstrate that the FAECES* method produces genotypes capable of differentiating coyotes and kit foxes, identifying individuals and their sex, and estimating genetic diversity and effective population sizes, even using highly degraded, low-quantity DNA extracted from scat. We found that the study area harbours a large and diverse population of kit foxes and a relatively smaller population of coyotes. By replicating our methods in the future, conservationists can assess the impacts of management decisions on canid populations. The method can also be adapted and applied more broadly to enrich and sequence multiple loci from any species of interest using scat or other noninvasive genetic samples.

KEYWORDS
canid, conservation genetics, coyote (*Canis latrans*), hybridization capture, kit fox (*Vulpes macrotis*), noninvasive genetic sampling

1 | INTRODUCTION

Studying the population dynamics of predators, including canids, is critical for conservation and ecosystem management because of the direct and indirect impacts predators have on prey species and the status of predators as sentinels, that is, they are sensitive to changes in the environment and can serve as indicators of ecosystem health (Sergio et al., 2008). Both coyotes (*Canis latrans*) and kit foxes (*Vulpes macrotis*, Figure 1) are known to consume the federally threatened Mojave desert tortoise (*Gopherus agasizzii*; Cypher et al., 2018). Coyotes are the primary cause of predation on adult tortoises (Esque et al., 2010), while kit foxes are the primary predators of tortoise nests (Bjurlin & Bissonette, 2004). It is therefore important for managers to monitor these canids in desert areas where the tortoise occurs.

Because canids tend to be elusive and difficult to study, noninvasive genetic sampling methods have been utilized extensively, for example, to estimate abundance and genetic diversity of coyotes (e.g., Morin et al., 2016; Prugh et al., 2005; Woodruff et al., 2021) and kit foxes (e.g., Lonsinger et al., 2018; Lonsinger et al., 2018; Wilbert et al., 2015, 2019). Scat, the most commonly collected material in noninvasive genetic studies (Waits & Paetkau, 2005), is particularly easy to identify in the field because canids defecate to mark territories (Morin et al., 2016). In areas where visual detection is difficult, scat searching dogs (*C. lupus familiaris*) have proven effective at locating scats (Smith et al., 2001; Ralls et al., 2003). In addition, scats both enable high-resolution analysis of host genetics and allow researchers to study predation by identifying DNA from consumed species (Banks et al., 2003).

Historically, most noninvasive studies targeted mitochondrial DNA for host species identification (e.g., Bozarth et al., 2010; Dalén et al., 2004; Paxinos et al., 1997) and nuclear microsatellites to reliably identify individuals and estimate population genetic structure (Lampa et al., 2007; Schwartz et al., 2006; Smith et al., 2006). However, single nucleotide polymorphism (SNP) loci are being used more often as new genotyping methods have made it more efficient and cost-effective to simultaneously genotype hundreds of individuals at hundreds to thousands of SNPs (Carroll et al., 2018). To date, the majority of noninvasive studies that have used SNP markers generated genotypes by using polymerase chain reaction (PCR) to amplify loci followed by measurement of fluorescence (e.g., with the Fluidigm platform; von Thaden et al., 2017) or direct amplicon sequencing (Natesh et al., 2019). However, the use of multiplex PCRs in noninvasive applications presents some challenges, including the need for species-specific references to design primer pairs, the potential for PCR inhibition due to coextracted inhibitors in scat-derived DNA, complex optimization of the multiplex reaction to avoid interactions between primers, and the requirement for DNA extracts >0.2 ng/μl (von Thaden et al., 2020).

In-solution DNA hybridization capture ("capture" hereafter) is an alternative to multiplex PCR amplification that is also well-suited to noninvasive DNA applications. First, unlike the primers used in PCR amplification, hybridization probes can be as much as 20% divergent from the target sequences in some cases (Li et al., 2018), precluding the need to have a reference genome for the target species. Second, capture methods do not require specialized equipment to generate genotypes (e.g., a Fluidigm machine). Finally, bait design is highly flexible, allowing one to minimize allelic dropout by probe tiling of the target region (Cruz-Dávalos et al., 2017) and including probes that match alternate alleles. It allows one to target as many regions as desired, including whole nuclear genome sequences or sequences from multiple taxa in a single assay (Campana, Hawkins, et al., 2016).

Despite the similar challenges posed by ancient and noninvasive DNA, capture methods have been widely adopted in ancient DNA studies but have rarely been used for noninvasive applications (Carroll et al., 2018). This may be due in part to the complexity of

FIGURE 1 Trail camera photographs of a coyote (left) and kit fox (right) in the BCCEA, Boulder City, Nevada, USA. Animals are shown next to Mojave Desert tortoise models. BCCEA, Boulder County Conservation Easement Area

bioinformatics involved in probe design, the high cost of generating probes (Meek & Larson, 2019), or the incidence of off-target capture (Jensen et al., 2020; von Thaden et al., 2020). Although several authors have demonstrated that it is possible to enrich and sequence primate DNA from scat samples (e.g., Hernandez-Rodriguez et al., 2018; Perry et al., 2010; Snyder-Mackler et al., 2016; White et al., 2019), to our knowledge these methods have not been applied to any taxonomic group outside of primates, have not been used to enrich for SNP loci, and have not targeted multiple species in a single enrichment.

We aimed to show that it is possible to generate SNP genotypes for individuals of two different canid species by using capture methods, and that these SNP data are capable of differentiating individuals, assigning sex, and estimating kinship, genetic structure, and genetic diversity. Accordingly, our first objective was to characterize and validate a set of informative SNP markers capable of identifying individual coyotes, desert kit foxes (*V. macrotis arsipus*, DKF), and San Joaquin kit foxes (*V. macrotis mutica*, SJKF).

Our second objective was to assess coyote and kit fox populations in the Boulder County Conservation Easement Area (BCCEA) in the Mojave Desert (Nevada, USA) by using capture methods to enrich and genotype by sequencing field-collected scat samples for our newly developed SNP loci. We aimed to estimate genetic diversity, genetic population structure, and effective population sizes for both coyotes and kit foxes.

The methods described here can be replicated to monitor canid population dynamics to evaluate the effects of management actions including minimizing anthropogenic subsidies (e.g., refuse and water), the presence of which can result in elevated predation on federally threatened tortoises (Esque et al., 2010) and other prey species. The methods can also be adapted to enable the use of scats and other noninvasive samples for other applied or basic genetic research.

2 | MATERIALS AND METHODS

2.1 | Study site and sample collection

We collected 340 scat samples at randomly selected points within the BCCEA and adjacent areas between September 2015 and April 2018 (Figure 2). The BCCEA is a 34,800-hectare area of public land in the northeastern Mojave Desert within the Eldorado Valley in southern Nevada, USA that was established in 1995 for the conservation of the Mojave desert tortoise and other desert wildlife. During the sampling period, weather was typical for the eastern Mojave Desert, with a mean annual temperature of 19.6°C (ranged from a mean minimum temperature of 0.5°C in January to a mean maximum of 32.5°C in July), and mean annual precipitation of 130 mm. August was the wettest month, with a mean monthly precipitation of 21.6 mm, and June was the driest (2.7 mm mean rainfall).

We used stratified random sampling (Ratti & Garton, 1994) to select locations for scat collection. Using ArcGIS, we selected 84 random points that were at least 1 km from any other sampling point. We included multiple sampling locations within three areas that we defined based on dominant land-use: city, BCCEA, and desert reference. City consisted of urbanized areas within Boulder City boundaries, the BCCEA zone included areas within and close to the BCCEA boundaries, and the desert reference area consisted of lands managed by the Bureau of Land Management south of the BCCEA, not including the highland mountains. Additional samples were collected opportunistically at power towers and fences. For consistency, at each sampling point circles of 10-m diameter were searched for scats and all scats within this area were collected. We recorded the collection date and location using a Garmin Montana 650 GPS device.

Collectors visually estimated and recorded the species that produced the scat based on size—coyotes are approximately 4–5× larger than kit foxes (Golightly & Ohmart, 1984). Coyotes and kit foxes are the most commonly found canids in the BCCEA; domestic dogs and grey foxes have also been observed, but they occur at very at low densities (W.I. Borman, personal communication). Scats were bagged individually unless they were found as part of a putative latrine site (Ralls & Smith, 2004); that is, if multiple pieces of scat were found at the same site, they were placed in a single bag. Based on visual inspection, the scats were rarely fresh and some may have been several weeks old.

Samples were stored dry in sealable plastic bags and shipped to the Center for Conservation Genomics (CCG), Smithsonian Institution, Washington, D.C. Upon receipt, subsamples of approximately 1.5 cm^3 were taken and stored in 15 ml conical tubes until DNA extraction. In cases where multiple pieces of scat were collected at the same locality during the same sampling session (i.e. multiple samples were taken from a putative latrine), we collected one subsample from each individual piece of scat. Scat subsamples were stored in a cool, dry location away from direct sunlight to preserve DNA (Goossens & Salgado-Lynn, 2013).

2.2 | FAECES* probe design

We designed a set of probes to enrich scat-derived DNA samples for a subset of SNP loci that are polymorphic in both coyotes and kit foxes and a subset that are fixed (or with a low minor allele frequency) in one species but polymorphic in the other to confirm species identification, identify individuals of each species, and calculate population genetics parameters. We also designed and included probes targeting the zinc finger-Y and -X genes (*ZFX*/*ZFY*) for sex identification (Figure 3). We refer to this method of using capture to genotype multiple canid species using DNA extracted from scat as FAECES*, or Fast and Accurate Enrichment of Canid Excrement for Species* and other analyses.

First, we generated a reference data set from which to select SNPs by enriching and genotyping coyote and kit fox tissue-derived DNA samples for a set of ~20,000 SNP loci that we previously found to be polymorphic in African wild dogs (*Lycaon pictus*, Campana, Parker, et al., 2016; Supporting Information: *Lycaon pictus* myBaits RNA probe

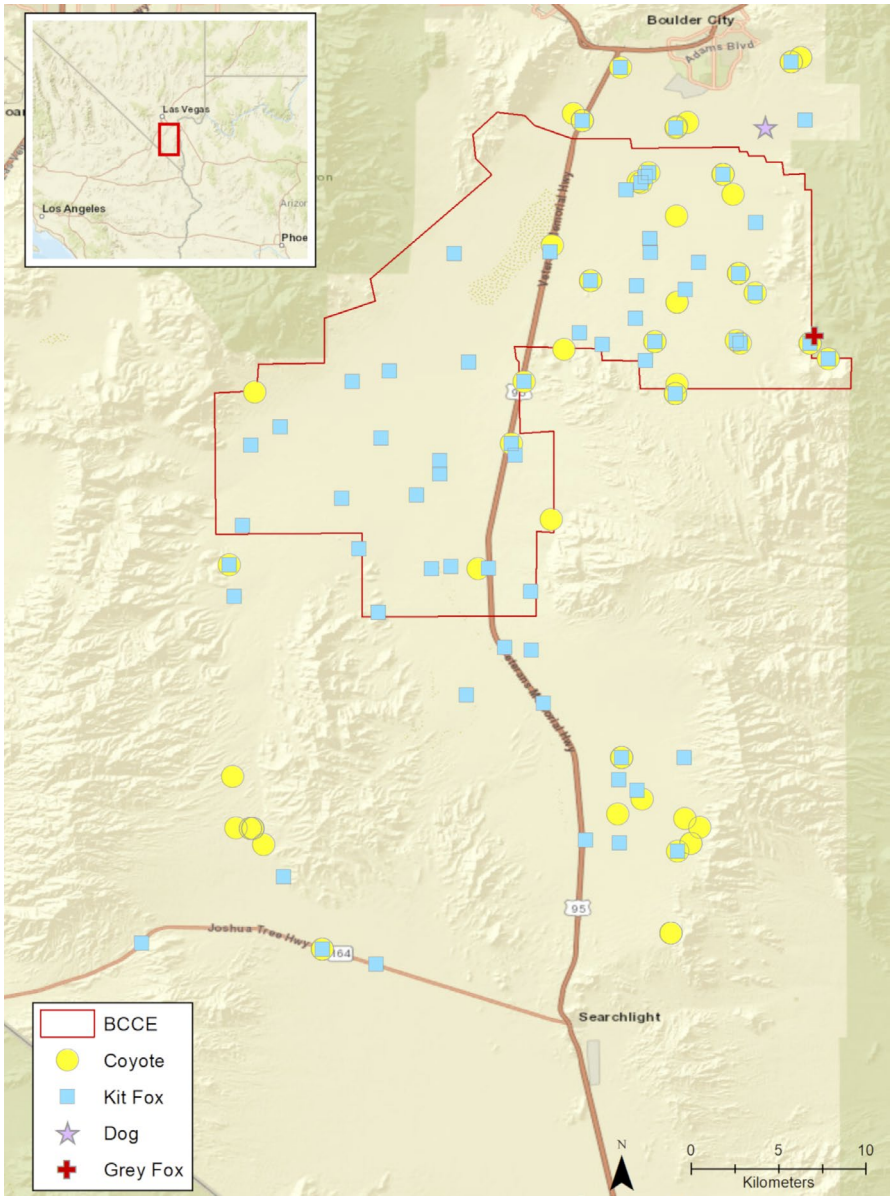

FIGURE 2 Map of sample collection localities. Data organized and map generated using Esri ArcMap 10.4.1. *Base map Source:* Esri © OpenStreetMap contributors 2021

design, Probes available in Figshare, 10.25573/data.14633298). We extracted DNA from eight western coyote tissue samples obtained from the Museum of Southwestern Biology (Accession numbers in Table S1) and one eastern coyote obtained from the tissue collection at the CCG using a DNeasy Blood and Tissue Kit (Qiagen), following the manufacturer's protocol for tissue. We also obtained nine DNA samples that were previously extracted from SJKF tissues collected in Bakersfield, CA (Wilbert et al., 2015, Table S1). We quantified DNA samples with a Qubit 4 fluorometer (Life Technologies) using a 1× dsDNA HS assay. We sheared DNA to an average length of 250 base pairs (bp) using a Bioruptor Pico sonicator (Diagenode, Inc.). The number of cycles required for adequate shearing varied by sample and ranged from 45 to 60; each cycle was 30 s on followed by 30 s off. We visualized the sheared DNA with a TapeStation 4200 System (Agilent Technologies) using High Sensitivity D1000 reagents. We then prepared dual-indexed libraries using the "BEST" single-tube method described in Carøe et al. (2017) with revisions as described in Mak et al. (2017). We quantified libraries after index PCR with the Qubit® fluorometer and enriched samples in single-plex reactions following the myBaits Manual v3 standard protocol. We quantified the enriched libraries using Qubit and visualized them on the TapeStation as described above. Finally, we pooled samples equimolarly and sequenced with paired-end 150 bp reads on an Illumina MiSeq at the CCG (Illumina, Inc.).

We downloaded demultiplexed sequence data from the BaseSpace Server (Illumina) and used the program FastQC v0.11.8 (Andrews, 2010) to check for sequence quality and adapter content. We removed adapter sequences using TrimGalore v0.6.4 (Krueger, 2019). We aligned reads to the domestic dog (*C. familiaris*) reference genome (Hoeppner et al., 2014; GenBank: CanFam3.1) using the "mem" algorithm in BWA v0.7.17 (Li, 2013). We used SAMTools v1.3.1 (Li et al., 2009) to sort BAM files and convert to the SAM format. Following the SAMtools variant calling workflow v1.0 (http://www.htslib.org/workflow/#mapping_to_variant), we then marked

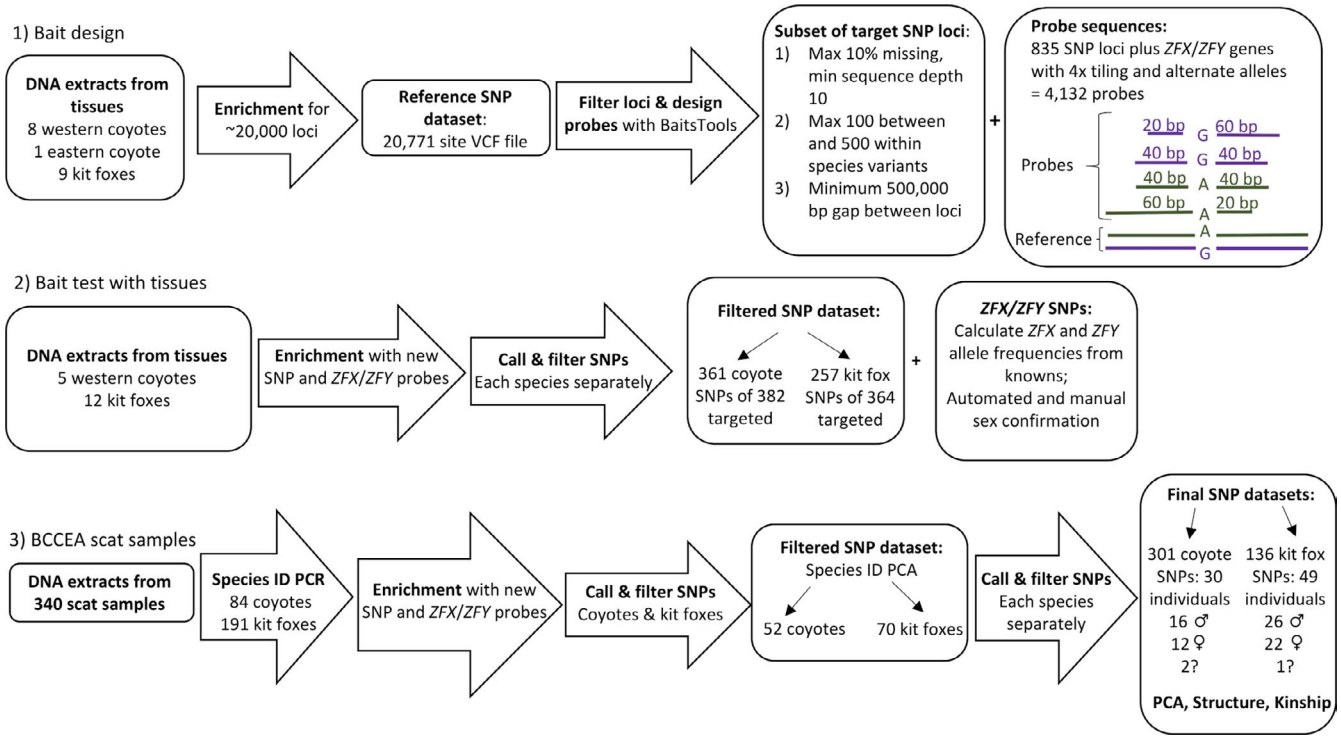

FIGURE 3 Workflow for FAECES* probe design, testing, and implementation

duplicates with Picard Tools v 2.20.6 (Picard Toolkit, 2019, https://github.com/broadinstitute/picard) and realigned reads around indels (insertions/deletions) using the GenomeAnalysisToolKit (GATK v3.7, McKenna et al., 2010). We identified sequence variants (minimum quality 20) using the SAMtools 'mpileup' command (-C50 option) and the BCFtools v1.4.1 "call" command (Li et al., 2009). We removed indels, any site with more than 10% missing data, and sites with average coverage <10 reads using VCFtools v0.1.15 (Danecek et al., 2011). We then used the "vcf2baits" command in BaitsTools v1.2.0 (Campana, 2018) to generate probe sequences. We utilized three options to improve the likelihood of enrichment success: (1) generating short (80 bp) probes (option -L 80), (2) designing probes with alternate alleles represented (option -a), (3) tiling probes to cover each variant site with ~4× average depth with an offset of 20 bp between probes (option -O 20). Probably because wild dogs are more closely related to coyotes than to kit foxes, and our initial probe set was designed to target wild dog variants, we initially recovered more coyote variants (1155 SNPs) than kit fox variants (454 SNPs). We then incorporated an option in BaitsTools to balance the number of probes by taxon, i.e. so that we would target approximately equal numbers of polymorphic coyote and kit fox loci. We set a maximum of 500 loci for each species, including some loci that are fixed within species in our samples but that vary between, as well as loci that are variable within each species. We used the options "taxacount" and "popcategories" to select the first 500 variants for each species.

To test the ability of the selected variants to differentiate species and individuals, we ran principal component analysis (PCAs) and pairwise relatedness analyses with the SNPRelate v1.8.0 package (Zheng et al., 2012) in R v4.0.3 (R Core Team, 2020, applies to all uses of R). First, we filtered the multisample VCF file to restrict loci to only those selected by BaitsTools. We then ran three PCA analyses: one with coyotes only, one with kit foxes only, and one with both species. We calculated pairwise kinship values by maximum likelihood estimation (MLE) with minor allele frequency ≥0.1. To simulate the effect of locus dropout, we randomly removed 20% of loci from each data set and ran all analyses again to ensure that the patterns were consistent.

Finally, we downloaded reference sequences for *ZFX* and *ZFY* genes for coyotes (Williams et al., 2003, GenBank AY145847 and AY145848, respectively) and kit foxes (Ortega et al., 2004, GenBank AY310919 and AY310920, respectively) and aligned them using the *MAFFT* v7.450 plugin (Katoh, 2005) in Geneious v9.1.2 (Biomatters Ltd). We exported the alignment in FASTA format and generated probes using the "aln2baits" command in BaitsTools with the same options as above (80 bp probes with 20 bp offset and c. 4× tiling). The final set of probe sequences, targeting both autosomal and sex-linked loci, was further filtered by Arbor Biosciences using their standard pipeline to remove loci that overlapped >25% with repeat-masked regions in the dog genome. Finally, we purchased myBaits RNA probes (myBaits-1 kit) from Arbor Biosciences (Probes available in FigShare, 10.25573/data.146332988).

2.3 | Scat samples

2.3.1 | Laboratory methods: DNA extraction

We extracted DNA from scat samples using a Mag-Bind Stool DNA kit (Omega Bio-Tek) following the manufacturer's standard

protocol with several modifications. First, bead bashing was omitted and replaced with overnight digestion with 30 μl DS buffer, 20 μl proteinase k (>600 mAU/ml), and enough SLX-Mlus Buffer to completely cover the sample. Samples were incubated at 56°C with agitation at 40 RPM. After digestion, samples were centrifuged for 2 min at 3000 *g*, and approximately 700 μl supernatant was transferred to a 2.2 ml 96-deepwell plate. The plate was then centrifuged again at 4000 *g* for 10 min, and 600 μl supernatant was transferred to a new 2.2 ml 96-deepwell plate, taking care not to disturb the pellet. We then added 1.2 ml Mag-Bind particles in XP2 buffer to each sample and mixed by pipetting. After a 5-min incubation, the plate was centrifuged for 2 min at 3000 *g* and then placed on a magnetic separation device. After elution with 100 μl of water, DNA extracts were quantified using a Qubit fluorometer with a 1× dsDNA HS assay.

2.3.2 | Library preparation and enrichment using FAECES* probes

Before making genomic libraries from the scat-derived DNA samples, we performed a PCR-based species identification assay following Bozarth et al. (2010). Briefly, we amplified a small fragment of the mitochondrial control region that is a different length in each canid species in the study area. We included both positive and negative controls in each set of reactions. Fragment length was determined by running PCR products on an ABI 3130xl at CCG. We excluded from further analysis all samples that produced PCR products for more than one canid species, that were identified as a nontarget canid species (i.e., dog or grey fox), or that failed to amplify. We proceeded with library preparation on the remaining samples that were identified as either coyote or kit fox.

To validate our SNP genotyping and sexing methodology, we included five of the tissue-derived coyote DNA samples used for probe design and 21 additional SJKF DNA samples collected in LoKern, CA from previous studies: seven tissue-derived samples and, for each individual, two additional fecal-derived samples that were previously identified genotype matches (Ralls et al., 2003) at six microsatellite loci shown to be sufficient to identify individuals (Smith et al., 2006). The kit fox faecal and tissue samples were previously sexed following PCR-based protocols developed by Ortega et al. (2004). Finally, we included previously extracted DNA from hair samples of five SJKF individuals from Bakersfield, CA (Bremner-Harrison et al., 2006) to test whether our methodology would result in SNP genotypes capable of discriminating SJKF from different populations.

We sheared all DNA samples to an average length of 250 bp using a Q800R sonicator (QSonica). We then prepared dual-indexed libraries using the single-tube method as described above (Carøe et al., 2017; Mak et al., 2017). We quantified libraries after index PCR with a Qubit fluorometer and pooled three samples equimolarly into each capture reaction. The total amount of starting DNA in each capture reaction ranged from approximately 100 to 2.5 μg—we added the maximum amount available for each library after index PCR. We diluted the probes three-fold and performed target enrichment using the standard protocol in the myBaits Manual v4. Briefly, the capture process involved heating to denature the libraries in solution with adapter blockers, followed by a cooling step to allow the blockers to hybridize. The baits were added when the solution reached the hybridization temperature (60°C). We allowed the hybridization to continue for at least 24 h. We then used streptavidin-coated magnetic beads to bind the biotinylated probes and pull the targeted molecules out of solution with a magnet. Nontarget DNA was washed away, followed by a PCR amplification step (14–18 cycles). After post-capture PCR, we quantified the enriched library pools using a Qubit fluorometer and visualized them on a 4200 TapeStation System (Agilent Technologies) with High Sensitivity D1000 reagents. Finally, we pooled captured libraries equimolarly and sequenced with paired-end 150 bp reads. We sequenced a random subset of 23 samples on an Illumina MiSeq (CCG) to validate that our methods were working; all remaining samples were then sequenced on one Illumina NovaSeq 6000 SP lane (Vincent J. Coates Genomics Sequencing Laboratory at the University of California Berkeley).

2.4 | Data analysis

2.4.1 | SNP calling

We trimmed raw reads for adapter content and quality, aligned reads to the dog reference genome (using BWA v0.7.17), and identified variants following the SAMtools workflow as described above—first, we identified variants in the combined coyote and kit fox data set in order to confirm species identification. We removed indels (option --remove-indels) and filtered sites to include only autosomal variants in the 140 bp regions targeted by the baits (option --bed) using VCFtools. We then filtered the all-sample VCF file (including coyotes and kit foxes) for sites with a minimum depth of five reads and no more than 80% missing data, and then used the option "thin −500000" to select only one variant per baited region. We confirmed species identification by performing a PCA analysis with the SNPRelate package in R as described above, and for subsequent analyses, we retained only samples which could be identified as either coyote or kit fox according to separation by PC1. We then separated samples according to species and called variants for coyotes and kit foxes separately using SAMtools mpileup and BCFtools v1.9 call as described above. We repeated this for coyote tissue samples ($n = 5$) and for kit fox tissue samples ($n = 12$) only so that we could compare locus dropout between tissue and scat samples. Using VCFtools, we filtered each all-sample VCF file to include one variant per autosomal baited region and only sites with a minimum depth of five reads, minor allele count ≥2, and no more than 25% missing data. Finally, for the scat sample data, we removed all sites that were out of Hardy-Weinberg equilibrium after Bonferroni correction ($\alpha = .05$).

2.4.2 | Identification of individuals and recaptures

Using the SNPRelate package in R, we performed PCA on both the kit fox and coyote variant data. For kit foxes, we performed analyses with both DKFs and SJKFs included and also on each subspecies separately. We then calculated identity-by-state (IBS) and identity-by-descent (IBD) with the MLE, which has been shown to be more accurate than the method of moments procedure (Blue et al., 2016). We characterized recaptures and resampling events (i.e., samples from the same individual) as pairs of samples with IBD >0.4 (IBS >0.95). Theoretically, two samples from the same individual should produce identical genotypes with IBD = 0.5 (IBS = 1); however, we used IBD >0.4 as the cutoff for individuals because we expect some level of allelic dropout due to DNA degradation, as well as stochasticity in genotyping. For microsatellite genotypes generated from noninvasive genetic samples, it is common to consider pairs of samples with one, two, or three differences as the same individual (e.g. Eggert et al., 2014; Roon et al., 2004). For SNP genotypes derived from hair samples via PCR and genotyping with the Fluidigm system, approximately five to six mismatches (6% dissimilarity) has been used as the maximum allowed dissimilarity to consider two samples from the same individual (von Thaden et al., 2020). Here, we used a genotyping-by-sequencing method following hybridization enrichment with probes for alternate alleles and applied genotyping quality controls—for these methods, there is not an established best practice for determining unique individuals. We chose the IBD >0.4 cutoff based on the range of kinship values expected for duplicate (or monozygotic twin) samples in human SNP data sets (Manichaikul et al., 2010).

We considered pairs of samples to represent first-order relatives (parent-offspring or full-siblings) if pairwise IBD ≥0.2 and ≤0.4 (Manichaikul et al., 2010; Milligan, 2003). We classified resampling events as occasions when the same individual was identified from multiple samples collected at the same site on a given date (i.e., multiple scats from a single individual at the same site during one sampling session), because we cannot rule out the possibility that multiple samples from the same individual were deposited on the same day. We defined recapture events as multiple samples from the same individual collected on different days and/or at different sites (Johnson et al., 2019). Using GenAlEx v6.503 (Peakall & Smouse, 2012), we calculated the probability of identity assuming siblings are present in the data (p_{IDsibs}, Waits et al., 2001).

We used VCFtools to calculate observed (H_o) and expected heterozygosity (H_e) and the inbreeding coefficient (F_{IS}) for all unique individuals and all "unrelated individuals", that is, with first degree relatives removed. For kit foxes, we calculated diversity metrics for SJKFs and DKFs separately. To determine if our SNPs are capable of discriminating between kit fox subspecies and between SJKF populations, we ran STRUCTURE v2.3.4 (Pritchard et al., 2000) on the data set including all unique individuals and for SJKFs and DKFs separately. For each STRUCTURE run, we used a burnin of 250,000 steps followed by 1,000,000 recorded steps. We used the admixture model, no location priors, and assumed correlated allele frequencies (Falush et al., 2003). We performed simulations with K = 1–8 with five replicates each, and identified meaningful K values using the ΔK method (Evanno et al., 2005) implemented in STRUCTURE HARVESTER v0.6.94 (Earl & vonHoldt, 2012). We combined replicate runs using CLUMPP v1.1.2 (Jakobsson & Rosenberg, 2007). We quantified the differentiation between kit fox groups by running an analysis of molecular variance (AMOVA) and estimating pairwise F_{ST} in GenAlEx. We assessed statistical significance with a permutation test of 10,000 replicates.

We estimated the effective population sizes (N_e) of coyotes and kit foxes in the BCCEA using the linkage disequilibrium model with random mating (Waples & Do, 2008) in NeEstimator v2.1 (Do et al., 2014). We report N_e values using minor allele frequency $p_{\text{crit}} = 0$, because our variants were already filtered for frequency, and 95% confidence intervals generated by the "parametric method" (Waples & Do, 2008). For coyotes, we then estimated the census population size based on the number of recaptures of different frequencies using the maximum likelihood program CAPWIRE v1.1.4 in R (Pennell et al., 2013). We used the likelihood ratio test to select between the null even capturability model (ECM) and the two innate rates model with default parameters.

For kit foxes, population turnover is generally high—previous studies have reported mean annual survival to be 0.42 (Cypher et al., 2000). Because our total sampling period spans multiple years, leading to a violation of the assumption of population closure, we estimated the census population size of kit foxes by using the estimated N_e/N_c ratio as 0.55 (Wilbert et al., 2019).

Finally, we investigated local spatial structure in both species. Using a Mantel test implemented in GenAlEx, we tested for a correlation between pairwise genotypic distance and Euclidean geographic distance, with 9999 permutations to generate the null distribution. Also using GenAlEx, we generated a Mantel correlogram to test for spatial correlation between pairs of individuals at different distance classes. We used the Sturges rule (Sturges, 1926) to determine the number of classes, and defined each class to ensure a sufficient number of comparisons within each class. We selected the "Spatial" option and performed 9999 permutations.

2.4.3 | Sex identification

For each sample positively identified as either a kit fox or coyote by PCA, we aligned trimmed reads to the appropriate species reference for *ZFX* and *ZFY* separately. We followed the steps outlined above for alignment and variant calling except for the "mpileup" command, during which we omitted the -C50 option. This option downgrades map quality for reads with excessive mismatches. Although omission of the -C50 option increases the likelihood of spurious alignments, we disabled it because the multiple SNPs separating the *ZFX* and *ZFY* gene copies in the short reference region produced strong bias towards the reference gene copy (e.g., inclusion of the -C50 option yielded 99.25% kit fox Y alleles when aligned against *ZFY* but only 7.62% Y alleles when mapped against *ZFX*).

Preliminary sex designations were automatically assigned using a voting algorithm (script available: https://github.com/campanam/

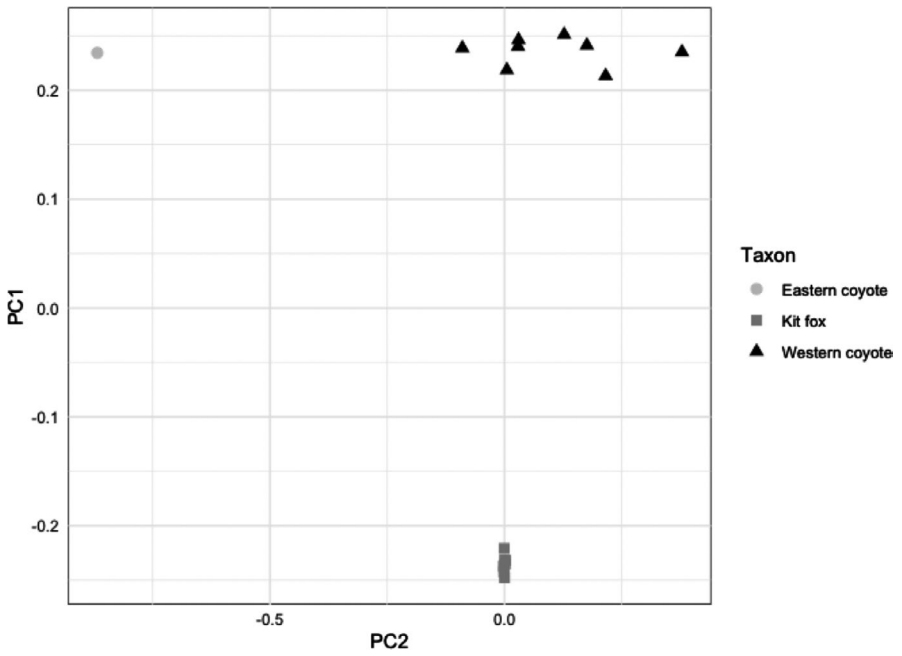

FIGURE 4 PCA of SNPs derived from coyote ($n = 9$) and San Joaquin kit fox ($n = 9$) tissues. PC1 accounts for 6.3% of variance; PC2 accounts for 4.1%. Eastern coyotes are shown in light grey circles, kit foxes in charcoal squares, and western coyotes in black triangles. PCA, principal component analysis; SNP, single nucleotide polymorphism

FAECES). For the automated sexing assignments, we discarded the *ZFY* alignments because mapping against the coyote *ZFY* yielded 100% Y alleles after removing the -C50 option, indicating significant asymmetric alignment bias. We catalogued SNP alleles that separate canid *ZFY* (male) alleles from their *ZFX* (female) homologues. At each sex-specific site, each read "voted" for the "male" or "female" allele. A minimum of 10 allele copies (across all sex-specific sites) was required to call sex. Since our data were unphased, we assumed that each allele and every site was independent. We recognize that alleles are linked on individual reads within the short region examined, which could potentially bias these preliminary assignments (e.g., through dropout of multiple Y SNPs if a single male DNA sequence is not retained). For each sample, we determined false negative (Y dropout) and false positive (Y drop-in/sequence misalignment) male determinations using the binomial probability of the deviation of the Y allele vote distributions from the expected Y allele frequency. We used an uncorrected $\alpha = .05$ for these statistical tests. Hybridization capture bias and reference bias can produce strong deviations from the expected 0.5 Y allele frequencies. Therefore, we empirically estimated the expected Y allele frequencies including these biases for each species using samples of known sex. These known samples were derived from tissues, where we can expect a minimum of allelic dropout and empirical frequencies closer to the expected 0.5. This penalizes male sex estimation in the scat samples, where dropout is more likely and statistical artefacts due to low DNA concentrations are much more likely to generate strong deviations from expected allele frequencies. To maximize sensitivity for males, we used the minimum observed Y allele frequency for each species in the known data sets as the expected Y allele frequency for the unknown samples. While we also experimented using the mean Y allele frequency and hard cut-offs at set *z*-scores, we found that the minimum Y frequency produced the results most consistent with the known samples (Table S2).

After initial automated sex estimation, we manually checked all results against known individuals and between sample replicates. We also checked all samples preliminarily identified as "female" where a significant number of Y alleles were detected (>10) as these probably represent Y dropout events, skewing the Y allele distribution from that expected from the known tissue samples. First, we imported BAM files into Geneious and generated consensus sequences with options "Assign Quality Total", call "N" if coverage <2 reads, and call heterozygotes >30%. We then aligned *ZFX* and *ZFY* consensus sequences to the appropriate reference *ZFX*/*ZFY* alignment and confirmed sex by visual inspection.

3 | RESULTS

3.1 | FAECES* probe design

The nine coyote and nine San Joaquin kit fox tissue samples that were enriched for ~20,000 SNP loci using the African wild dog probe set (see Supporting Information) were sequenced with a mean of approximately 1 million reads per sample (range of 65,3640–1,254,195; SRA data PRJNA767395). Our final probe set targeted 835 autosomal SNP loci and the *ZFY*/*ZFX* genes for both coyotes and kit foxes and comprised 4132 unique baits (FigShare 10.25573/data.14633298). Of the 835 autosomal SNPs, 382 were polymorphic in coyotes and 364 were polymorphic in SJKFs—the remaining 89 were variable between species but not within either species (i.e., fixed in our data set between species). The PCA including all samples showed separation of kit foxes and coyotes, and eastern coyote from western coyote (Figure 4). Results of PCA and kinship analyses on coyote and kit fox data sets were similar after simulating the effect of dropout by randomly removing 20% of loci. The kinship analysis including all SNP loci with a minor allele frequency ≥0.1 included 249

SNPs, and after simulating 20% dropout, 89 SNPs; the coefficient of determination (R^2) between kinship estimates was 0.80.

3.2 | DNA sequencing of FAECES* probe enriched libraries

After excluding samples that failed to amplify in the species identification PCR, mixed-species samples, and nontarget canid samples (dog and grey fox), we proceeded with library preparation and sequencing on 275 scat samples (81% of the 340 collected) from the BCCEA. Failure of the species identification PCR could have been due to poor DNA quality, the presence of PCR inhibitors, or the scats may have been produced by a noncanine species. Of these 275 samples, 84 were identified as coyote and 191 kit fox by the species identification PCR (Bozarth et al., 2010). Collectors correctly identified the canid species (i.e., the visual classification matched the species identification based on PCR product length) in 82% of samples (225/275). Including the known individual SJKF and coyote tissue samples ($n = 31$), we sequenced a total of 306 samples with a mean of 1.5 million reads per sample (376–11.3 million; SD = 1,045,501). Across all samples, a mean of 75% of reads (0.19%–99%; SD = 21%) mapped to the dog reference genome. For scat samples, a mean of 73% (SD = 19%) of reads mapped; for tissues, a mean of 99% (SD = 0.2%) of reads mapped.

3.3 | Species identification of canid samples using SNP data

After filtering variants from the joint species all-sample VCF (including coyotes and kit foxes), 668 sites remained. Based on the ability of the PCA to discriminate species (PC1, 57% of variation in the data), we filtered out samples with fewer than 30 SNPs, leaving 70 kit fox samples and 52 coyote samples (122 samples/275, 44%, Table S3). The species identification based on PCA matched the known species for all tissue samples and matched the identification based on our PCR assay in all but two of the scat samples. These two scats were classified as kit foxes by the species ID assay, but clustered with coyotes in the PCA. For these two samples, we assembled reads that did not map to the dog reference genome to both a coyote and a red fox mitochondrial genome (GenBank Accessions NC_008093 and AM181037, respectively) using the Geneious algorithm (medium-low sensitivity and up to five iterations). We generated consensus sequences (options "Total" quality), assigning ambiguity codes if at least 40% of reads disagreed at a given site, and aligned consensus sequences and references using the *MAFFT* v7.450 (Katoh, 2005) plugin. Visual inspection of alignments at the *12S rRNA* gene revealed bases matching both the coyote and kit fox reference. These two mixed samples were either the result of contamination, perhaps due to a mixed-species latrine (Ralls & Smith, 2004) or coyote predation of kit fox (Ralls & White, 1995) and were removed from further analyses.

3.4 | Individual identification, genetic diversity, and estimates of coyote population size

For all samples confirmed as coyotes ($n = 52$), 301 polymorphic sites remained (out of 382 targeted) after filtering for missingness, minor allele count, depth of coverage, and Hardy-Weinberg equilibrium. For tissue samples only ($n = 5$), 361 SNPs remained; all filters were the same, except we did not filter this small data set for loci in Hardy-Weinberg equilibrium. The PCA showed no apparent pattern of clustering among coyote samples (Figure S1). Among the 47 coyote samples from the BCCEA, we identified five individuals that were recaptured in different locations during multiple sampling sessions (Figure 5), and five that were resampled in a single session (pairwise kinship >0.40 or IBS >95%), leaving 30 unique individuals. Two of the recaptured individuals were sampled from latrines; this suggests that latrine samples did not necessarily have sufficient cross-contamination to generate novel genotypes. We identified 11 individuals in 13 samples in 2015, four individuals in five samples in 2016, and 22 individuals in 23 observations in 2017; three individuals were recaptured in multiple years. The maximum distance observed between recaptured individuals was 25.5 km, between two samples collected in September 2015 and March 2017; the shortest distance was 1.1 km between two samples collected in January 2016 and May 2017.

For subsequent estimation of genetic diversity, we used only unique individuals, selecting the sample with the least missing data for each. We also identified several first-order relatives (parent-offspring or full sibling pairs, kinship >0.2)—after removing the individual with the most missing data in each related pair, 22 individuals remained. Including first-order relatives, the average observed heterozygosity across variable sites was 0.24 (SD 0.04) and the inbreeding coefficient, F_{IS} was 0.050 (SD 0.2); excluding these individuals, average observed heterozygosity was 0.25 (SD 0.04) and F_{IS} was 0.050 (SD 0.1, Table 1). Bartlett's test revealed that the observed and expected heterozygosity were not significantly different in either data set ($p > .1$). The probability of two individuals having identical genotypes, p_{ID}, was 1.30×10^{-65}; assuming siblings are present in the data, the probability (p_{IDsibs}) was 1.1×10^{-33}. The number of SNPs at which the p_{IDsibs} was <.0001 was 38.

Including first order relatives ($n = 30$), the NeEstimator estimated effective population size (N_e) of coyotes in the BCCEA was 38 individuals (95% confidence interval [CI] = 34–42). Inclusion of close relatives can result in artificially lower estimates of N_e, so we also ran the analysis on a data set with first order relatives removed ($n = 22$). As expected, the estimated N_e was higher when using this data set: N_e =65 (95% CI =54–82). We did not find significant isolation-by-distance (IBD, $R_{xy} = 0.14$, $p = .169$), or spatial autocorrelation among individuals.

Because of the small number of recaptures, we pooled recapture data across years for the CAPWIRE analysis. We were not able to reject the ECM ($p = .1$); based on this model, the estimated census population size in the BCCEA was 83 (95% CI = 48–210).

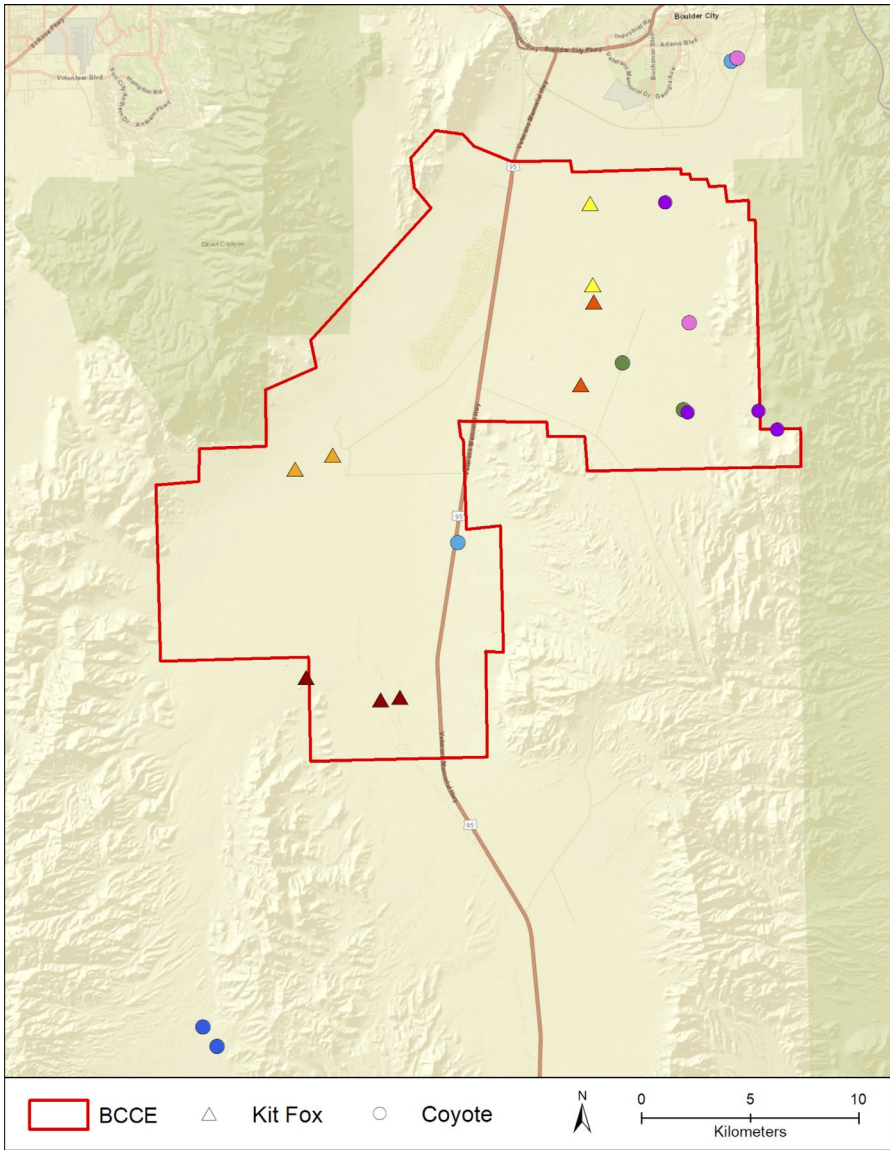

FIGURE 5 Map of coyote and kit fox recaptures coloured by individual. Triangles represent desert kit foxes and circles represent coyotes. Data organized and map generated using Esri ArcMap 10.4.1. *Base map Source:* Esri © OpenStreetMap contributors, 2021

	H_o	H_e	F_{IS}	N_e (95% CI)
All coyotes (n = 52)	0.24 (0.04)	0.25 (0.009)	0.037 (0.2)	N/A
Unique BCCEA individuals (n = 30)	0.24 (0.04)	0.25 (0.01)	0.049 (0.2)	38 (34–42)
Unrelated BCCEA individuals (n = 22)	0.25 (0.04)	0.26 (0.007)	0.050 (0.1)	65 (54–82)

TABLE 1 Mean observed and expected heterozygosity (H_o and H_e), inbreeding coefficients (F_{IS}), and effective population sizes (N_e) for coyotes

Note: Standard deviations are in parentheses (95% CIs for the N_e estimates).
Abbreviations: BCCEA, Boulder County Conservation Easement Area; CI, confidence interval.

3.5 | Individual identification, genetic diversity and structure, and estimates of kit fox population size

After filtering for missingness, minor allele count, depth of coverage, and Hardy-Weinberg equilibrium, 136 polymorphic SNPs remained in our full kit fox data set (of 364 targeted by our probe set). For tissue samples only (n = 12), 257 SNPs remained after filtering with the same filters as in the full data set except we did not filter for loci in Hardy-Weinberg equilibrium. Most of the filtered loci in the full data set were removed due to missingness.

These 136 loci in the full data set were able to differentiate between SJKFs and DKFs, as well as between SJKF from two localities in California. Including all 62 unique kit fox individuals (SJKFs and DKFs), PCA separated the two subspecies (PC1, accounting for 15.6% of variation, Figure 6). Analysing only the 13 SJKF samples, the PCA separated the SJKFs from the two localities, LoKern

and Bakersfield (PC1, accounting for 16.4% of variation, Figure S2). Including all unique kit foxes ($n = 62$), the results of the STRUCTURE analysis indicated the most likely number of population clusters is two, with individuals separated by subspecies (Table S4). Analysing only SJKF ($n = 13$, 12 tissue-derived and one scat-derived sample), the most likely number of population clusters was one when using no location priors, and two with location priors (Table S5). Individuals were divided between those from LoKern and Bakersfield. Pairwise F_{ST} between the two kit fox subspecies was 0.16 ($p = .0001$); the results of the AMOVA showed that between-subspecies variation accounts for 15% of total variation in the data set. Between the two SJKF localities, $F_{ST} = 0.094$ ($p = .005$). The average observed heterozygosity among SJKFs is 0.32 (SD 0.04), and Bartlett's test revealed no significant difference in variances between observed and expected heterozygosity ($p > .1$). The average F_{IS} was 0.04 (SD 0.13).

Of the seven sets of matching LoKern SJKF scat (two each) and tissue pairs, only one scat sample produced enough sequencing reads for individual identification. The estimated pairwise kinship between the tissue and its putative corresponding scat sample was 0.27, indicating probable first-degree relatives. Based on p_{IDsibs}, our SNP data set provides more statistical power than the microsatellite data previously used to distinguish between individual recaptures and first-degree relatives (1.2×10^{-18} and 7.95×10^{-3}, Smith et al., 2006, respectively).

Among the 57 BBCEA desert kit fox samples, we identified 49 unique individuals, four recaptured individuals (Figure 5), and three individuals that were resampled. Of the four recaptured individuals, two were captured in multiple years (2015 and 2017), and one was sampled from a latrine. The maximum geographical distance between observations of individuals was 4.4 km, sampled on October 2015 and January 2018. We identified 27 first-degree relative pairs; after removing the individual with the most missing data in each pair, 36 "unrelated" individuals remained. Average observed heterozygosity among the 49 individuals was 0.30 (SD 0.06), and average F_{IS} was −0.016 (SD 0.2). Excluding putative first-degree relatives, average observed heterozygosity was 0.29 (SD 0.05) and F_{IS} was −0.017 (SD 0.2, Table 2). Bartlett's test revealed that the variances in observed and expected heterozygosity were not significantly different in either data set ($p > .1$). p_{ID} was 1.30×10^{-34} and p_{IDsibs} was 1.2×10^{-18}. The number of SNPs at which the p_{IDsibs} was <.0001 was 34. Using genotypes from unrelated individuals captured between January and April 2017 ($n = 34$), the NeEstimator estimated effective population size was 179 (95% CI = 92–1644). Assuming an N_e/N_c ratio of 0.55, $N_c = 325$ (166–2989), or 0.4 kit foxes/km^2 (0.2–3.7 foxes/km^2). The Mantel test revealed no significant correlation between genetic distance and genotypic distance ($R_{xy} = -0.054$, $p = .250$), and the Mantel correlogram showed no evidence of spatial autocorrelation of individuals at any distance class.

3.6 | Sex identification

The estimated expected Y allele frequencies using samples of known sex were 0.8315 for kit foxes and 0.4382 for coyotes. The minimum observed Y allele frequencies in the known data set for each species, which we used for the expected Y allele frequency for unknown samples, were 0.7994 for kit fox and 0.4036 for coyotes. Of the 30 individual coyotes identified in our surveys of the BCCEA, 16 were male, 12 were female, and 2 were undetermined. The estimated sex across all samples representing resamples of individuals ($n = 5$ individuals, $n = 6$ resample events) matched; all recaptures ($n = 4$ individuals, $n = 6$ recapture events) also matched, except for one recapture event, where the sex of one sample could not be determined. Our sex assignments matched the known sex for all tissue references.

Of the 49 individual kit foxes from the BCCEA, 22 were female, 26 male, and one undetermined. The assigned sex matched across all

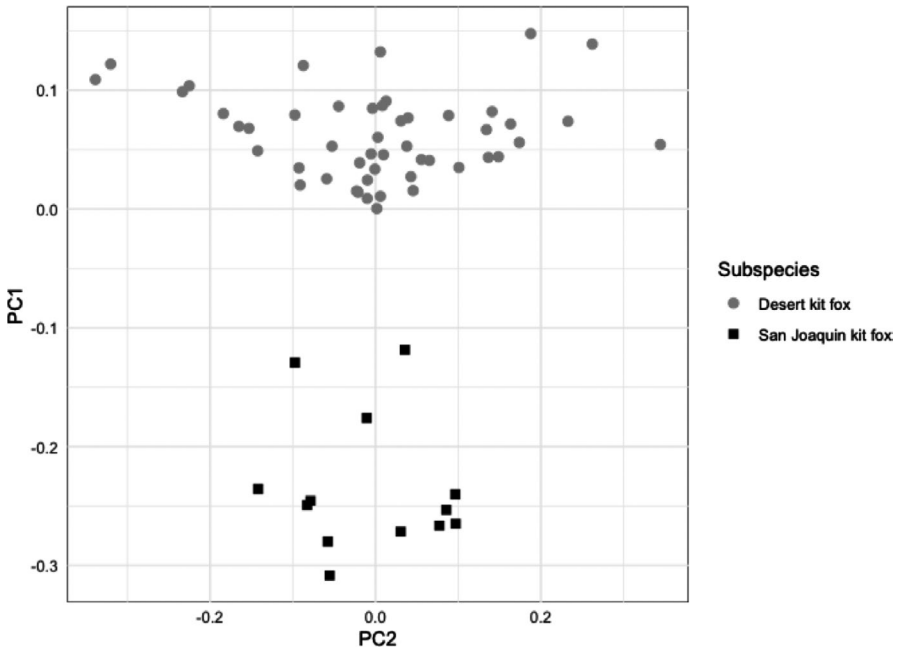

FIGURE 6 PCA of all kit fox individuals ($n = 62$); PC1 accounts for 15.6% of variation and PC2 accounts for 4.5% of variation. PCA, principal component analysis

	H_o	H_e	F_{IS}
All (n = 70)	0.27 (0.06)	0.29 (0.008)	0.078 (0.2)
SJKF (n = 13)	0.32 (0.04)	0.33 (0.002)	0.040 (0.1)
DKF (n = 57)	0.29 (0.06)	0.29 (0.008)	−0.0071 (0.2)
Unique DKF individuals (n = 49)	0.30 (0.06)	0.29 (0.008)	−0.016 (0.2)
Unrelated DKF individuals (n = 36)	0.29 (0.05)	0.29 (0.006)	−0.017 (0.2)

TABLE 2 Mean observed and expected heterozygosity (H_o and H_e) and inbreeding coefficients (F_{IS}) for San Joaquin kit foxes (SJFK) and desert kit foxes (DKF) in the BCCEA

Note: Standard deviations are in parentheses.

Abbreviation: BCCEA, Boulder County Conservation Easement Area.

samples for recaptured individuals (n = 4 individuals, n = 5 events) and resampled individuals (n = 3 individuals, n = 3 events). For all SJKF tissue samples, the sex estimated from the data matched the known sex of each individual (n = 12). Regarding the matching LoKern SJKF scat and tissue pairs, for which scat sexing was previously conducted using a PCR method (Ortega et al., 2004), only one scat produced enough sequencing reads for sex identification (the same sample that produced enough reads for individual identification). Although the estimated sex of this sample (F) did not match the expected sex based on the PCR method (M), as discussed above, based on pairwise kinship, these two samples most probably represent first-order relatives. Based on the high mean sequence depth, it is unlikely that this scat sample represents dropout of *ZFY* alleles; mean sequencing depth was 711 reads (0–1618) with 2577 reads mapped to the *ZFX* reference. It is more likely that the PCR-based method incorrectly identified the sample as male.

4 | DISCUSSION

4.1 | Canid populations in the BCCEA and implications for the conservation of Mojave desert tortoises

Given the proportion of scats that were identified as coyotes (31%) and estimated effective and census population sizes, our data suggest that the population of coyotes in the BCCEA is smaller than that of kit foxes and on the low end of estimates reported in previous studies. From our scat DNA analyses, we estimated the census population size (N_c) of the coyotes in this study area to be 83. Assuming that the total suitable habitat in the sampled region is 800 km^2, that is, the total area sampled not including the mountainous habitat in the southeast, the density of coyotes in the area is 0.10 individuals/km^2 (95% CI = 48–210 or 0.06–0.26 coyotes/km^2). Previous studies of western coyotes have reported values including 0.053–0.112 coyote/km^2 (Woodruff et al., 2021), 0.14 coyotes/km^2 (Ralls & White, 1995), and 0.07–0.08 coyotes/km^2 (Lonsinger, Lukacs, et al., 2018). In general, coyote densities tend to be lower in desert areas where they are sympatric with kit foxes because coyotes have higher water needs than kit foxes which are better adapted to arid environments (Lonsinger, Lukacs, et al., 2018). In undisturbed landscapes, coyote densities are expected to be lower than those of kit foxes because of their larger size (4–5×) and higher energetic requirements (Golightly & Ohmart, 1984). However, coyote densities can be higher than those of kit fox in landscapes with anthropogenic disturbances including invasive plant species and artificial water sources, both of which decrease available kit fox habitat and prey species while increasing the number of coyotes, the primary competitors (and occasional predators; Ralls & White, 1995) of kit foxes (Arjo et al., 2007).

Because we pooled recapture data across several years, we probably violated the assumption of population closure; that is, it is probable that there were immigration, emigration, birth, and/or death events during the total sampling period. However, it is unlikely that we underestimated the coyote population size because violating the closure assumption decreases the likelihood of recapture and increases the estimated population size. Another source of potential bias could come from how we identified unique individuals and recaptured/resampled individuals. As is the case with any genotypes derived from noninvasive samples and unknown individuals, it is possible that a portion of the samples that we characterized as unique individuals were actually recaptures, with variation between genotypes due to contamination (false alleles) or allelic dropout, which would also lead to an overestimation of the population size.

We estimated that the population of kit foxes is relatively large compared to that of coyotes. Sixty-nine percent of scats collected were identified as kit foxes; of 54 genotyped DKF scats, we identified 49 individuals. The estimated mean N_e of the DKF population is more than 2× that of coyotes (179 compared to 65). The estimated population density (0.4/km^2) overlaps with a previous study of SJKF based on mark-recapture at the Naval Petroleum Reserves in southern San Joaquin Valley, California (0.2–1.7/km^2, Cypher et al., 2000) but is higher than the density reported for SJKF in the Ciervo-Panoche Natural Area (CPNA, 0.12–0.24 foxes/km^2, Wilbert et al., 2019), and greatly exceeds contemporary estimates of DKF in UT (0.02 foxes/km^2, Lonsinger, Adams, et al., 2018).

While we found no evidence of isolation-by-distance or spatial autocorrelation among DKF individuals, Wilbert et al. (2019) reported that SJKF individuals in the CPNA found within 6 km have significantly higher relatedness than expected by chance. These authors also reported a strong signature of IBD and population structure probably caused by bisection of the landscape by a major highway and the complex, heterogenous habitat. This suggests that by contrast, the landscape around the BCCEA provides adequate kit fox habitat capable of supporting high and unimpeded gene flow.

Given the large DKF population and its potential impact on threatened Mojave desert tortoises by nest predation (Bjurlin & Bissonette, 2004), land managers should consider limiting factors that can support increased canid populations during times of low prey availability, for example, reducing supplemental, anthropogenic sources of food including garbage (Cypher et al., 2018). Both coyotes and kit foxes are known to consume anthropogenetic food sources during times of low prey availability such as during winter and after periods of low precipitation when rodent and lagomorph populations decline (Kelly et al., 2019). Anthropogenic subsidization could sustain canid abundance through these times of natural food shortages, which could in turn increase predation pressure on prey species including the Mojave desert tortoises (i.e., hyperpredation, Esque et al., 2010). By replicating the methodology we used here, managers in the BCCEA could monitor canid populations over time to assess the effects of actions to prevent subsidization of canid populations. Future studies could also estimate the frequency of canid predation on tortoises by developing efficient molecular methods to detect tortoise DNA in scats and thus evaluate the impact that canids are having on this protected species.

4.2 | Use of the FAECES* method to genotype scat samples

We showed that by using the FAECES* method, employing in-solution hybridization capture, it is possible to generate SNP genotypes capable of identifying individual canids and their sex using scat samples from multiple canid species in a single assay. We enriched for 382 and 364 polymorphic SNPs in coyotes and kit foxes, respectively, and successfully genotyped individuals of both species using even very low quantity scat-derived DNA extracts (<1 ng/μl). The average amount of starting DNA (ng) that went into library preparation was significantly higher (Wilcoxon two-sample t test, $p = .003$) for samples that successfully produced genotypes of at least 35 SNP loci than for those that did not (199.9 and 113.5 ng, respectively); however, we generated successful genotypes from scats of both species with starting DNA concentrations as low as 0.1 ng/μl (~3 ng total). Our method worked reliably for DNA extracted from tissue samples—all tissue-derived samples resulted in full genotypes with no missing data, even using low quantity DNA that was extracted more than 20 years ago and stored at −20°C (LoKern kit fox DNA, 1.34–5.5 ng/μl). However, only one of the 14 scat extracts of similar age yielded enough reads for individual and sex identification, suggesting poorer long-term preservation of scat DNA extracts relative to DNA extracted from tissue. The high concentration scat DNA extracts that failed to generate genotypes probably had low percentages of endogenous canid DNA content (i.e., they had a high percentage of prey or microbial DNA), which we did not quantify (Cruz-Dávalos et al., 2017).

Our final data sets consisted of 136 polymorphic kit fox SNPs and 301 coyote SNPs. Fifty-eight percent of coyote samples and 33% percent of kit fox samples were successfully genotyped. The lower success rate and smaller number of SNPs recovered from kit fox samples is most likely due to divergence between probe and target DNA sequences, as has been documented in previous studies (van der Valk et al., 2017). Dogs and kit foxes are separated by approximately 9–10 million years of evolution (Lindblad-Toh et al., 2005), while dogs and coyotes only diverged ~1 million years ago (vonHoldt et al., 2011). In this study, we recovered more of the targeted kit fox SNP loci (257/364, 71%) when we analysed the 12 kit fox tissue samples separately—the lower locus recovery rate (136/364, 37%), was only observed in the scat-derived kit fox DNA samples. We observed a higher recovery (lower dropout) rate in the coyote scats (301/384, 79%) than the kit fox scats, suggesting that capture efficiency using scat-derived DNA is diminished when there is a higher sequence divergence between the capture bait and the target DNA. However, our success rates for both species fall within the range of success previously reported in microsatellite studies on coyotes and kit foxes using scat (27.5%–91.4%, Eriksson et al., 2020; Lonsinger, Adams, et al., 2018).

Although our success rates were lower than previous studies that genotyped amplicons using the Fluidigm platform (80%–97%, von Thaden et al., 2017), by using the FAECES* method we were able to generate sequence data and include a larger number of loci than the Fluidigm platform in a single enrichment (i.e., >96) without the need for specialized equipment (only a thermocycler is needed in-house—samples can be sequenced by an off-site sequencing core). We were able to recover sufficient kit fox SNPs with enough power to differentiate between subspecies and populations of SJKFs, and to identify DKF individuals with a very low probability of identity ($p_{ID} = 1.3 \times 10^{-34}$).

Given that the aim of this study was to test the use of capture methods to generate SNP genotypes from field-collected scat samples which vary greatly in quality and quantity, we did not selectively collect fresh scat or prescreen DNA extracts for quality beyond species identification. Future studies could improve the FAECES* method success rate by (1) preferentially collecting fresh scat samples in the winter when DNA degrades at a slower rate due to lower temperatures and less UV radiation (Lonsinger, Lukacs, et al., 2018) and freezing subsamples prior to extraction, (2) prescreening samples for endogenous nuclear DNA content through qPCR assays, amplification success with microsatellite loci, or shotgun sequencing (Fontsere et al., 2021), and/or (3) performing multiple DNA extractions and/or library preparations on each sample and pooling prior to enrichment (Fontsere et al., 2021; Hernandez-Rodriguez et al., 2018; Perry et al., 2010).

Although van der Valk et al. (2017) showed no difference in consensus mitogenome sequences derived from multiple scat samples from the same individual, future studies should also investigate the average amount of allelic dropout in nuclear SNP genotypes derived from hybridization enrichment followed by genotyping-by-sequencing of noninvasive samples. This would ideally be done by collecting multiple scats from the same known individual at various time points post-defecation to allow for differential degredation in the field. These estimates could then be used as benchmarks for the

average number of mismatches to expect between SNP genotypes derived from multiple samples from the same individual. Future work is also needed to establish if replicates are necessary and, if so, how many and at what step (e.g., library preparation, index PCR, hybridization, and/or sequencing).

Capture methods provide greater flexibility for SNP genotyping than methods based on the generation of amplicons and subsequent genotyping by sequencing or fluorescence because of the ability for probes to hybridize with sequences as much as 25% divergent in some cases (van der Valk et al., 2017), precluding the need for a species-specific reference as well as the need to optimize PCR conditions for large multiplexes. Here, to select kit fox and coyote variants for probe design, we enriched with probes previously designed to target sites of known canid variation, that is, targetting polymorphic SNPs identified in African wild dogs, and used the domestic dog reference genome for SNP calling. We chose enrichment over RADseq because our tissue-derived DNA references were themselves relatively degraded, and DNA degradation significantly decreases the efficiency of RADseq (Graham et al., 2015). However, given higher quality reference samples, RADseq is another option for genome subsampling for SNP selection that is especially useful in the case where a genomic reference is not available.

Because of the flexibility of the in-solution capture approach, the methods we described here could be applied to population genetics studies of any species of interest for which scat or other nonivasive samples can be collected. For SNP identification and bait design, preliminary sequence data can be obtained from shotgun sequencing or RADseq (e.g., Hoffberg et al., 2016) performed on reference tissue samples. In addition to putatively neutrally evolving SNP loci, researchers could also include baits targeting genes that may be under selection, sex identification genes, etc. Designing enrichment probes and generating SNP genotypes using sequence capture data requires some bioinformatics skills (Meek & Larson, 2019) which can be a barrier to the use of capture methods. However, the use of BaitsTools (Campana, 2018), a fast and user-friendly software, automates and facilitates probe design.

Finally, we implemented several cost-saving measures that increased the economic feasibility of our methodology, including probe dilution, multiplexing three samples per capture reaction (Hernandez-Rodriguez et al., 2018), and using a single-tube library preparation method (Carøe et al., 2017; Mak et al., 2017). Including the cost of probes, our estimated per-sample cost (for 400 samples) including library preparation, capture, and sequencing was approximately half (47% less) of the estimated per-sample cost to genotype 96 samples for 96 SNPs using the Fluidigm platform (Table S6; Carroll et al., 2018).

4.3 | Conclusion

We described and validated the FAECES* method, utilizing in-solution DNA hybridization capture of SNPs to genotype canids from noninvasively collected scat samples. Using this method, we showed that the landscape in and surrounding the BCCEA in the Mojave Desert harbors a relatively large and genetically diverse population of desert kit foxes and a smaller population of coyotes. The FAECES* method can be replicated and adapted in the future to enable noninvasive population genetic studies, including capture-recapture, in multiple sympatric species using a single capture assay—thus expanding the toolbox available to researchers and conservation practitioners studying rare or elusive taxa including canids, felids, or any other species of interest for which scat or other noninvasive samples can be collected.

ACKNOWLEDGEMENTS

We thank Scott Cambrin and Larry Mata of the Desert Conservation Program of Clark County Nevada, and Nancy C. R. McInerney, Natalí Chavez, and Margad-Erdene Ochirbat of the Center for Conservation Genomics for logistical support and help with laboratory work. We thank Alison Devault and Brian Brunelle of Arbor Biosciences for assistance with probe design, and Matthew Hopken and three anonymous reviewers for helpful comments that improved the quality of this manuscript.

CONFLICT OF INTEREST

The authors declared that they have no conflicts of interest.

AUTHOR CONTRIBUTIONS

William I. Boarman conceived the project and William I. Boarman, Robert C. Fleischer, and Jesús E. Maldonado secured funding; Lillian D. Parker, Michael G. Campana, Tammy R. Wilbert, Katherine Ralls, Robert C. Fleischer, William I. Boarman, and Jesús E. Maldonado designed the study; William I. Boarman, Ryan Boarman, conducted the field work and collected samples; Lillian D. Parker, Jessica D. Quinta, and Isabel Rivera carried out laboratory work; Lillian D. Parker analysed and archived the data and wrote the manuscript with input from all authors; Michael G. Campana participated in data analysis and wrote the manuscript for sample sexing; Lillian D. Parker, William I. Boarman and Ryan Boarman drafted figures and maps; all authors participated in revisions and acceptance of the submitted version of the manuscript.

OPEN RESEARCH BADGES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at (https://doi.org/10.25573/data.14633298).

DATA AVAILABILITY STATEMENT

DNA sequences: SRA PRJNA767395. Probe sequences: Available on FigShare (https://doi.org/10.25573/data.14633298). Sexing script: https://github.com/campanam/FAECES

ORCID

Lillian D. Parker https://orcid.org/0000-0003-3370-9473
Michael G. Campana https://orcid.org/0000-0003-0461-6462
Jesús E. Maldonado https://orcid.org/0000-0002-4282-1072

REFERENCES

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data (v0.11.8) [Computer software]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Arjo, W. M., Gese, E. M., Bennett, T. J., & Kozlowski, A. J. (2007). Changes in kit fox-coyote-prey relationships in the Great Basin Desert, Utah. *Western North American Naturalist*, *67*(3), 389–401. https://doi.org/10.3398/1527-0904

Banks, S. C., Horsup, A., Wilton, A. N., & Taylor, A. C. (2003). Genetic marker investigation of the source and impact of predation on a highly endangered species. *Molecular Ecology*, *12*(6), 1663–1667. https://doi.org/10.1046/j.1365-294X.2003.01823.x

Bjurlin, C. D., & Bissonette, J. A. (2004). Survival during early life stages of the desert tortoise (*Gopherus agassizii*) in the south-central Mojave Desert. *Journal of Herpetology*, *38*(4), 527–535. https://doi.org/10.1670/94-02A

Blue, E. M., Brown, L. A., Conomos, M. P., Kirk, J. L., Nato, A. Q., Popejoy, A. B., Raffa, J., Ranola, J., Wijsman, E. M., & Thornton, T. (2016). Estimating relationships between phenotypes and subjects drawn from admixed families. *BMC Proceedings*, *10*(7), 42. https://doi.org/10.1186/s12919-016-0056-3

Bozarth, C. A., Alva-Campbell, Y. R., Ralls, K., Henry, T. R., Smith, D. A., Westphal, M. F., & Maldonado, J. E. (2010). An efficient noninvasive method for discriminating among faeces of sympatric North American canids. *Conservation Genetics Resources*, *2*(1), 173–175. https://doi.org/10.1007/s12686-010-9215-4

Bremner-Harrison, S., Harrison, S. W. R., Cypher, B. L., Murdoch, J. D., Maldonado, J. E., & Darden, S. K. (2006). Development of a single-sampling noninvasive hair snare. *Wildlife Society Bulletin*, *34*(2), 456–461. https://doi.org/10.2193/0091

Campana, M. G. (2018). BaitsTools: Software for hybridization capture bait design. *Molecular Ecology Resources*, *18*(2), 356–361. https://doi.org/10.1111/1755-0998.12721

Campana, M. G., Hawkins, M. T. R., Henson, L. H., Stewardson, K., Young, H. S., Card, L. R., Lock, J., Agwanda, B., Brinkerhoff, J., Gaff, H. D., Helgen, K. M., Maldonado, J. E., McShea, W. J., & Fleischer, R. C. (2016). Simultaneous identification of host, ectoparasite and pathogen DNA via in-solution capture. *Molecular Ecology Resources*, *16*(5), 1224–1239. https://doi.org/10.1111/1755-0998.12524

Campana, M. G., Parker, L. D., Hawkins, M. T. R., Young, H. S., Helgen, K. M., Szykman Gunther, M., Woodroffe, R., Maldonado, J. E., & Fleischer, R. C. (2016). Genome sequence, population history, and pelage genetics of the endangered African wild dog (*Lycaon pictus*). *BMC Genomics*, *17*(1), 1013. https://doi.org/10.1186/s12864-016-3368-9

Carøe, C., Gopalakrishnan, S., Vinner, L., Mak, S. S. T., Sinding, M. H. S., Samaniego, J. A., Wales, N., Sicheritz-Pontén, T., & Gilbert, M. T. P. (2017). Single-tube library preparation for degraded DNA. *Methods in Ecology and Evolution*, *9*(2), 410–419. https://doi.org/10.1111/2041-210X.12871

Carroll, E. L., Bruford, M. W., DeWoody, J. A., Leroy, G., Strand, A., Waits, L., & Wang, J. (2018). Genetic and genomic monitoring with minimally invasive sampling methods. *Evolutionary Applications*, *11*(7), 1094–1119. https://doi.org/10.1111/eva.12600

Cruz-Dávalos, D. I., Llamas, B., Gaunitz, C., Fages, A., Gamba, C., Soubrier, J., Librado, P., Seguin-Orlando, A., Pruvost, M., Alfarhan, A. H., Alquraishi, S. A., Al-Rasheid, K. A. S., Scheu, A., Beneke, N., Ludwig, A., Cooper, A., Willerslev, E., & Orlando, L. (2017). Experimental conditions improving in-solution target enrichment for ancient DNA. *Molecular Ecology Resources*, *17*(3), 508–522. https://doi.org/10.1111/1755-0998.12595

Cypher, B. L., Kelly, E. C., Westall, T. L., & Van Horn Job, C. L. (2018). Coyote diet patterns in the Mojave Desert: Implications for threatened desert tortoises. *Pacific Conservation Biology*, *24*(1), 44–54. https://doi.org/10.1071/PC17039

Cypher, B. L., Warrick, G. D., Otten, M. R. M., O'Farrell, T. P., Berry, W. H., Harris, C. E., Kato, T. T., McCue, P. M., Scrivner, J. H., & Zoellick, B. W. (2000). Population dynamics of San Joaquin Kit foxes at the naval petroleum reserves in California. *Wildlife Monographs*, *145*, 1–43 (JSTOR).

Dalén, L., Götherström, A., & Angerbjörn, A. (2004). Identifying species from pieces of faeces. *Conservation Genetics*, *5*(1), 109–111. https://doi.org/10.1023/B:COGE.0000014060.54070.45

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, *14*(1), 209–214. https://doi.org/10.1111/1755-0998.12157

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Eggert, L. S., Buij, R., Lee, M. E., Campbell, P., Dallmeier, F., Fleischer, R. C., Alonso, A., & Maldonado, J. E. (2014). Using genetic profiles of African forest elephants to infer population structure, movements, and habitat use in a conservation and development landscape in Gabon. *Conservation Biology*, *28*(1), 107–118. https://doi.org/10.1111/cobi.12161

Eriksson, C. E., Ruprecht, J., & Levi, T. (2020). More affordable and effective noninvasive SNP genotyping using high-throughput amplicon sequencing. *Molecular Ecology Resources*, *20*, 1505–1516. https://doi.org/10.1111/1755-0998.13208

Esque, T., Nussear, K., Drake, K., Walde, A., Berry, K., Averill-Murray, R., Woodman, A., Boarman, W., Medica, P., Mack, J., & Heaton, J. (2010). Effects of subsidized predators, resource variability, and human population density on desert tortoise populations in the Mojave Desert, USA. *Endangered Species Research*, *12*(2), 167–177. https://doi.org/10.3354/esr00298

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, *164*(4), 1567–1587. https://doi.org/10.1093/genetics/164.4.1567

Fontsere, C., Alvarez-Estape, M., Lester, J., Arandjelovic, M., Kuhlwilm, M., Dieguez, P., Agbor, A., Angedakin, S., Ayuk Ayimisin, E., Bessone, M., Brazzola, G., Deschner, T., Eno-Nku, M., Granjon, A.-C., Head, J., Kadam, P., Kalan, A. K., Kambi, M., Langergraber, K., … Lizano, E. (2021). Maximizing the acquisition of unique reads in noninvasive capture sequencing experiments. *Molecular Ecology Resources*, *21*(3), 745–761. https://doi.org/10.1111/1755-0998.13300

Golightly, R. T. Jr, & Ohmart, R. D. (1984). Water economy of two desert canids: Coyote and kit fox. *Journal of Mammalogy*, *65*(1), 51–58. https://doi.org/10.2307/1381199

Goossens, B., & Salgado-Lynn, M. (2013). Advances and difficulties of molecular tools for carnivore conservation in the tropics. *Raffles Bulletin of Zoology*, *28*, 43–53.

Graham, C. F., Glenn, T. C., McArthur, A. G., Boreham, D. R., Kieran, T., Lance, S., Manzon, R. G., Martino, J. A., Pierson, T., Rogers, S. M., Wilson, J. Y., & Somers, C. M. (2015). Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). *Molecular Ecology Resources*, *15*(6), 1304–1315. https://doi.org/10.1111/1755-0998.12404

Hernandez-Rodriguez, J., Arandjelovic, M., Lester, J., de Filippo, C., Weihmann, A., Meyer, M., Angedakin, S., Casals, F., Navarro, A., Vigilant, L., Kühl, H. S., Langergraber, K., Boesch, C., Hughes, D., & Marques-Bonet, T. (2018). The impact of endogenous content, replicates and pooling on genome capture from faecal samples. *Molecular Ecology Resources*, *18*(2), 319–333. https://doi.org/10.1111/1755-0998.12728

Hoeppner, M. P., Lundquist, A., Pirun, M., Meadows, J. R. S., Zamani, N., Johnson, J., Sundström, G., Cook, A., FitzGerald, M. G., Swofford, R., Mauceli, E., Moghadam, B. T., Greka, A., Alföldi, J., Abouelleil, A., Aftuck, L., Bessette, D., Berlin, A., Brown, A., … Grabherr, M. G. (2014). An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. *PLoS One*, *9*(3), e91172. https://doi.org/10.1371/journal.pone.0091172

Hoffberg, S. L., Kieran, T. J., Catchen, J. M., Devault, A., Faircloth, B. C., Mauricio, R., & Glenn, T. C. (2016). RADcap: Sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data. *Molecular Ecology Resources*, *16*(5), 1264–1278. https://doi.org/10.1111/1755-0998.12566

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jensen, M. R., Sigsgaard, E. E., Liu, S., Manica, A., Bach, S. S., Hansen, M. M., Møller, P. R., & Thomsen, P. F. (2020). Genome-scale target capture of mitochondrial and nuclear environmental DNA from water samples. *Molecular Ecology Resources*, *21*(3), 690–702. https://doi.org/10.1111/1755-0998.13293

Johnson, M. B., Parker, L. D., Vanthomme, H., Tchignoumba, L., Deichmann, J. L., Maldonado, J. E., Korte, L., & Alonso, A. (2019). Patterns of genetic diversity in African forest elephants living in a human-modified landscape in southwest Gabon. *Conservation Science and Practice*, *1*(8), e76. https://doi.org/10.1111/csp2.76

Katoh, K. (2005). MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, *33*(2), 511–518. https://doi.org/10.1093/nar/gki198

Kelly, E. C., Cypher, B. L., & Germano, D. J. (2019). Temporal variation in foraging patterns of desert kit foxes (*Vulpes macrotis arsipus*) in the Mojave Desert, California, USA. *Journal of Arid Environments*, *167*, 1–7. https://doi.org/10.1016/j.jaridenv.2019.04.008

Krueger, F. (2019). Trim Galore (v0.6.4) [Computer software]. https://github.com/FelixKrueger/TrimGalore

Lampa, S., Gruber, B., Henle, K., & Hoehn, M. (2007). An optimisation approach to increase DNA amplification success of otter faeces. *Conservation Genetics*, *9*(1), 201. https://doi.org/10.1007/s10592-007-9328-9

Li, C., Hofreiter, M., Straube, N., Corrigan, S., & Naylor, G. J. P. (2018). Capturing protein-coding genes across highly divergent species. *BioTechniques*, *54*(6), 321–326. https://doi.org/10.2144/000114039

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv*, 1303.3997v1.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lindblad-Toh, K., Wade, C. M., Mikkelsen, T. S., Karlsson, E. K., Jaffe, D. B., Kamal, M., Clamp, M., Chang, J. L., Kulbokas, E. J., Zody, M. C., Mauceli, E., Xie, X., Breen, M., Wayne, R. K., Ostrander, E. A., Ponting, C. P., Galibert, F., Smith, D. R., deJong, P. J., … Lander, E. S. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*, *438*, 803–819. https://doi.org/10.1038/nature04338

Lonsinger, R. C., Adams, J. R., & Waits, L. P. (2018). Evaluating effective population size and genetic diversity of a declining kit fox population using contemporary and historical specimens. *Ecology and Evolution*, *8*(23), 12011–12021. https://doi.org/10.1002/ece3.4660

Lonsinger, R. C., Lukacs, P. M., Gese, E. M., Knight, R. N., & Waits, L. P. (2018). Estimating densities for sympatric kit foxes (*Vulpes macrotis*) and coyotes (*Canis latrans*) using noninvasive genetic sampling. *Canadian Journal of Zoology*, *96*(10), 1080–1089.

Mak, S. S. T., Gopalakrishnan, S., Carøe, C., Geng, C., Liu, S., Sinding, M.-H.-S., Kuderna, L. F. K., Zhang, W., Fu, S., Vieira, F. G., Germonpré, M., Bocherens, H., Fedorov, S., Petersen, B., Sicheritz-Pontén, T., Marques-Bonet, T., Zhang, G., Jiang, H., & Gilbert, M. T. P. (2017). Comparative performance of the BGISEQ-500 vs Illumina HiSeq2500 sequencing platforms for palaeogenomic sequencing. *GigaScience*, *6*(8), 1–13. https://doi.org/10.1093/gigascience/gix049

Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics*, *26*(22), 2867–2873. https://doi.org/10.1093/bioinformatics/btq559

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Meek, M. H., & Larson, W. A. (2019). The future is now: Amplicon sequencing and sequence capture usher in the conservation genomics era. *Molecular Ecology Resources*, *19*(4), 795–803. https://doi.org/10.1111/1755-0998.12998

Milligan, B. G. (2003). Maximum-likelihood estimation of relatedness. *Genetics*, *163*(3), 1153–1167. https://doi.org/10.1093/genetics/163.3.1153

Morin, D. J., Kelly, M. J., & Waits, L. P. (2016). Monitoring coyote population dynamics with fecal DNA and spatial capture–recapture. *The Journal of Wildlife Management*, *80*(5), 824–836. https://doi.org/10.1002/jwmg.21080

Natesh, M., Taylor, R. W., Truelove, N. K., Hadly, E. A., Palumbi, S. R., Petrov, D. A., & Ramakrishnan, U. (2019). Empowering conservation practice with efficient and economical genotyping from poor quality samples. *Methods in Ecology and Evolution*, *10*(6), 853–859. https://doi.org/10.1111/2041-210X.13173

Ortega, J., Rosario Franco, M. D., Adams, B. A., Ralls, K., & Maldonado, J. E. (2004). A reliable, non-invasive method for sex determination in the endangered San Joaquin kit fox (*Vulpes macrotis mutica*) and other canids. *Conservation Genetics*, *5*(5), 715–718. https://doi.org/10.1007/s10592-004-1862-0

Paxinos, E., McIntosh, C., Ralls, K., & Fleischer, R. (1997). A noninvasive method for distinguishing among canid species: Amplification and enzyme restriction of DNA from dung. *Molecular Ecology*, *6*(5), 483–486. https://doi.org/10.1046/j.1365-294X.1997.00206.x

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research–an update. *Bioinformatics*, *28*(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Pennell, M. W., Stansbury, C. R., Waits, L. P., & Miller, C. R. (2013). Capwire: A R package for estimating population census size from non-invasive genetic sampling. *Molecular Ecology Resources*, *13*(1), 154–157. https://doi.org/10.1111/1755-0998.12019

Perry, G. H., Marioni, J. C., Melsted, P., & Gilad, Y. (2010). Genomic-scale capture and sequencing of endogenous DNA from feces. *Molecular Ecology*, *19*(24), 5332–5344. https://doi.org/10.1111/j.1365-294X.2010.04888.x

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

Prugh, L. R., Ritland, C. E., Arthur, S. M., & Krebs, C. J. (2005). Monitoring coyote population dynamics by

genotyping faeces. *Molecular Ecology*, *14*(5), 1585–1596. https://doi.org/10.1111/j.1365-294X.2005.02533.x

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ralls, K., Maldonado, J. E., & Smith, D. A. (2003). Dogs, scats, and DNA: A non-invasive approach for carnivore field studies [Unpublished raw data].

Ralls, K. S., & Smith, D. A. (2004). Latrine use by San Joaquin kit foxes (*Vulpes macrotis mutica*) and coyotes (*Canis latrans*). *Western North American Naturalist*, *64*(4), 544–547.

Ralls, K., & White, P. J. (1995). Predation on San Joaquin kit foxes by larger canids. *Journal of Mammalogy*, *76*(3), 723–729. JSTOR. https://doi.org/10.2307/1382743

Ratti, J. T., & Garton, E. O. (1994). Research and experimental design. In T. A. Bookhout (Ed.), *Research and management techniques for wildlife and habitats* (pp. 1–23). The Wildlife Society.

Roon, D. A., Thomas, M. E., Kendall, K. C., & Waits, L. P. (2004). Evaluating mixed samples as a source of error in non-invasive genetic studies using microsatellites. *Molecular Ecology*, *14*(1), 195–201. https://doi.org/10.1111/j.1365-294X.2004.02385.x

Schwartz, M. K., Cushman, S. A., McKelvey, K. S., Hayden, J., & Engkjer, C. (2006). Detecting genotyping errors and describing American black bear movement in northern Idaho. *Ursus*, *17*(2), 138–148. https://doi.org/10.2192/1537-6176(2006)17[138:DGEADA]2.0.CO;2

Sergio, F., Caro, T., Brown, D., Clucas, B., Hunter, J., Ketchum, J., McHugh, K., & Hiraldo, F. (2008). Top predators as conservation tools: Ecological rationale, assumptions, and efficacy. *Annual Review of Ecology, Evolution, and Systematics*, *39*(1), 1–19. https://doi.org/10.1146/annurev.ecolsys.39.110707.173545

Smith, D. A., Ralls, K., Davenport, B., Adams, B., & Maldonado, J. E. (2001). Canine assistants for conservationists. *Science*, *291*(5503), 435. https://doi.org/10.1126/science.291.5503.435B

Smith, D. A., Ralls, K., Hurt, A., Adams, B., Parker, M., Davenport, B., Smith, M. C., & Maldonado, J. E. (2003). Detection and accuracy rates of dogs trained to find scats of San Joaquin kit foxes (*Vulpes macrotis mutica*). *Animal Conservation*, *6*, 339–346. https://doi.org/10.1017/S136794300300341X

Smith, D. A., Ralls, K., Hurt, A., Adams, B., Parker, M., & Maldonado, J. E. (2006). Assessing reliability of microsatellite genotypes from kit fox faecal samples using genetic and GIS analyses: Testing reliability of faecal genotypes. *Molecular Ecology*, *15*(2), 387–406. https://doi.org/10.1111/j.1365-294X.2005.02841.x

Snyder-Mackler, N., Majoros, W. H., Yuan, M. L., Shaver, A. O., Gordon, J. B., Kopp, G. H., Schlebusch, S. A., Wall, J. D., Alberts, S. C., Mukherjee, S., Zhou, X., & Tung, J. (2016). Efficient genome-wide sequencing and low-coverage pedigree analysis from noninvasively collected samples. *Genetics*, *203*(2), 699–714. https://doi.org/10.1534/genetics.116.187492

Sturges, H. (1926). The choice of a class-interval. *Journal of the American Statistical Association*, *21*, 65–66. https://doi.org/10.1080/01621459.1926.10502161

van der Valk, T., Lona Durazo, F., Dalén, L., & Guschanski, K. (2017). Whole mitochondrial genome capture from faecal samples and museum-preserved specimens. *Molecular Ecology Resources*, *17*(6), e111–e121. https://doi.org/10.1111/1755-0998.12699

von Thaden, A., Cocchiararo, B., Jarausch, A., Jüngling, H., Karamanlidis, A. A., Tiesmeyer, A., Nowak, C., & Muñoz-Fuentes, V. (2017). Assessing SNP genotyping of noninvasively collected wildlife samples using microfluidic arrays. *Scientific Reports*, *7*(1), 10768. https://doi.org/10.1038/s41598-017-10647-w

von Thaden, A., Nowak, C., Tiesmeyer, A., Reiners, T. E., Alves, P. C., Lyons, L. A., Mattucci, F., Randi, E., Cragnolini, M., Galián, J., Hegyeli, Z., Kitchener, A. C., Lambinet, C., Lucas, J. M., Mölich, T., Ramos, L., Schockert, V., & Cocchiararo, B. (2020). Applying genomic data in wildlife monitoring: Development guidelines for genotyping degraded samples with reduced single nucleotide polymorphism (SNP) panels. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13136

vonHoldt, B. M., Pollinger, J. P., Earl, D. A., Knowles, J. C., Boyko, A. R., Parker, H., Geffen, E., Pilot, M., Jedrzejewski, W., Jedrzejewska, B., Sidorovich, V., Greco, C., Randi, E., Musiani, M., Kays, R., Bustamante, C. D., Ostrander, E. A., Novembre, J., & Wayne, R. K. (2011). A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. *Genome Research*, *21*(8), 1294–1305. https://doi.org/10.1101/gr.116301.110

Waits, L. P., Luikart, G., & Taberlet, P. (2001). Estimating the probability of identity among genotypes in natural populations: Cautions and guidelines. *Molecular Ecology*, *10*(1), 249–256. https://doi.org/10.1046/j.1365-294X.2001.01185.x

Waits, L. P., & Paetkau, D. (2005). Noninvasive genetic sampling tools for wildlife biologists: a review of applications and recommendations for accurate data collection. *Journal of Wildlife Management*, *69*(4), 1419–1433. https://doi.org/10.2193/0022-541X

Waples, R. S., & Do, C. (2008). ldne: A program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, *8*(4), 753–756. https://doi.org/10.1111/j.1755-0998.2007.02061.x

White, L. C., Fontsere, C., Lizano, E., Hughes, D. A., Angedakin, S., Arandjelovic, M., Granjon, A.-C., Hans, J. B., Lester, J. D., Rabanus-Wallace, M. T., Rowney, C., Städele, V., Marques-Bonet, T., Langergraber, K. E., & Vigilant, L. (2019). A roadmap for high-throughput sequencing studies of wild animal populations using noninvasive samples and hybridization capture. *Molecular Ecology Resources*, *19*(3), 609–622. https://doi.org/10.1111/1755-0998.12993

Wilbert, T. R., Woollett, D. A. S., Westphal, M. F., Whitelaw, A., Ralls, K., & Maldonado, J. E. (2019). Distribution, fine-scale subdivision, and population size of San Joaquin kit foxes in the Ciervo-Panoche Natural Area, California. *Conservation Genetics*, *20*(3), 405–417. https://doi.org/10.1007/s10592-018-1122-3

Wilbert, T., Woollett, D., Whitelaw, A., Dart, J., Hoyt, J., Galen, S., Ralls, K., Meade, D., & Maldonado, J. (2015). Non-invasive baseline genetic monitoring of the endangered San Joaquin kit fox on a photovoltaic solar facility. *Endangered Species Research*, *27*, 31–41. https://doi.org/10.3354/esr00649

Williams, C. L., Blejwas, K., Johnston, J. J., & Jaeger, M. M. (2003). A coyote in sheep's clothing: Predator identification from saliva. *Wildlife Society Bulletin (1973–2006)*, *31*(4), 926–932.

Woodruff, S. P., Eacker, D. R., & Waits, L. P. (2021). Estimating coyote densities with local, discrete bayesian capture-recapture models. *The Journal of Wildlife Management*, *85*(1), 73–86. https://doi.org/10.1002/jwmg.21967

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, *28*(24), 3326–3328. https://doi.org/10.1093/bioinformatics/bts606

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Parker, L. D., Campana, M. G., Quinta, J. D., Cypher, B., Rivera, I., Fleischer, R. C., Ralls, K., Wilbert, T. R., Boarman, R., Boarman, W. I., & Maldonado, J. E. (2022). An efficient method for simultaneous species, individual, and sex identification via in-solution single nucleotide polymorphism capture from low-quality scat samples. *Molecular Ecology Resources*, 22, 1345–1361. https://doi.org/10.1111/1755-0998.13552