



Original Article

A Genome-Wide Perspective on the Persistence of Red Wolf Ancestry in Southeastern Canids

Elizabeth Heppenheimer^o, Kristin E. Brzeski, Joseph W. Hinton^o,
Michael J. Chamberlain, Jacqueline Robinson, Robert K. Wayne, and
Bridgett M. vonHoldt^o

From the Department of Ecology and Evolutionary Biology, Princeton University, 106A Guyot Hall, Princeton, NJ 08544 (Heppenheimer and vonHoldt); Michigan Technological University, College of Forest Resources and Environmental Science, Houghton, MI 49931 (Brzeski); Department of Environmental and Forest Biology, College of Environmental Science and Forestry, State University of New York, 1 Forestry Drive, Syracuse, NY 13210 (Hinton); Warnell School of Forestry and Natural Resources, University of Georgia, 180 E Green Street, Athens, GA 30621 (Chamberlain); Institute for Human Genetics, University of California, San Francisco, 513 Parnassus Ave, San Francisco, CA 94143 (Robinson); and Department of Ecology and Evolutionary Biology, University of California, Los Angeles, 610 Charles E Young Dr East, Los Angeles, CA 90095 (Robinson and Wayne).

Address correspondence to Elizabeth Heppenheimer and Bridgett M. vonHoldt at the address above, or e-mail: elizabeth.heppenheimer@gmail.com; vonholdt@princeton.edu

Received June 16, 2019; First decision September 22, 2019 Accepted February 21, 2020.

Corresponding Editor: Eduardo Eizirik

Abstract

The red wolf (*Canis rufus*), a legally recognized and critically endangered wolf, is known to interbreed with coyotes (*Canis latrans*). Declared extirpated in the wild in 1980, red wolves were reintroduced to northeastern North Carolina nearly a decade later. Interbreeding with coyotes was thought to be restricted to a narrow geographic region adjacent to the reintroduced population and largely believed to threaten red wolf recovery. However, red wolf ancestry was recently discovered in canids along the American Gulf Coast, igniting a broader survey of ancestry in southeastern canid populations. Here, we examine geographic and temporal patterns of genome-wide red wolf ancestry in 260 canids across the southeastern United States at over 164 000 SNP loci. We found that red wolf ancestry was most prevalent in canids sampled from Texas in the mid-1970s, although non-trivial amounts of red wolf ancestry persist in this region today. Further, red wolf ancestry was also observed in a subset of coyotes inhabiting North Carolina, despite management efforts to limit the occurrence of hybridization events. Lastly, we found no evidence of substantial red wolf ancestry in southeastern canids outside of these 2 admixture zones. Overall, this study provides a genome-wide survey of red wolf ancestry in canids across the southeastern United States, which may ultimately inform future red wolf restoration efforts.

Subject areas: Conservation genetics and biodiversity

Keywords: admixture, ancestry, conservation, genomics, introgression

Characterizing geographic and temporal patterns of hybridization is critical for evaluating the conservation status and subsequent restoration of threatened species. In particular, the role of hybridization has been a predominant and contentious issue in the conservation of wolf populations in eastern North America (Waples et al. 2018). The red wolf (*Canis rufus*), a legally recognized wolf of the American southeast, was declared extirpated in the wild by 1980 as a result of habitat loss, predator control efforts, disease, and interbreeding with sympatric coyotes (*Canis latrans*) (Nowak 2002; Hinton et al. 2013; Waples et al. 2018). Between 1973 and 1977, up to 400 canids were trapped from the Gulf Coast region of Louisiana and Texas and evaluated for canonical red wolf features, which include cranial morphology and vocalizations (US Fish and Wildlife Service 1989; Hedrick and Fredrickson 2008). Only 17 of these individuals were classified as red wolves and retained as founders for the Species Survival Plan's captive breeding program, and most of the remaining "capture zone" canids were euthanized (Carley 2000). Twelve animals were reproductively successful in captivity, and following nearly a decade of captive breeding, red wolves were reintroduced to the Albemarle Peninsula in northeastern North Carolina in 1987.

At the time of red wolf reintroduction, coyotes were believed to be absent from the reintroduction area, with the first interbreeding documented in 1993 as coyote populations expanded eastward (Parker et al. 1995; Kelly et al. 1999). Consequently, various active management strategies (e.g., coyote sterilization, removal of hybrid litters) were implemented to reduce the impact of coyote introgression on the red wolf genome (Gese et al. 2015). Coyote and red wolf interbreeding in North Carolina is most commonly observed following the anthropogenic breakup stable wolf breeding pairs (Bohling and Waits 2015; Hinton et al. 2017), suggesting that human-caused mortality and mate availability are major factors in determining the frequency of these hybridization events (Hinton et al. 2018).

Interbreeding between red wolves and coyotes has been detected with microsatellite markers ascertained for red wolf ancestry, where ancestry is inferred from STRUCTURE's Q -values (Adams et al. 2003, 2007). Reciprocal introgression has been detected, with coyote ancestry present within the North Carolina red wolf population and minor red wolf ancestry in coyote populations adjacent to the North Carolina reintroduction area (Adams et al. 2007; Bohling and Waits 2015; Bohling et al. 2016). Recently, 2 studies independently concluded that canonical red wolf genotypes have persisted in canids along the American Gulf Coast of Louisiana and Texas, where red wolves have been considered extirpated for nearly 4 decades (Heppenheimer et al. 2018a; Murphy et al. 2019). Prior to these studies, attempts to identify remnant red wolf populations in Texas were unsuccessful (Hailer and Leonard 2008), although phenotypic evidence suggested that red wolf-like canids may persist in this area (Giordano and Pace 2000; Mech and Nowak 2010). Additionally, Heppenheimer et al. (2018a) noted that 2 admixed individuals from Galveston Island, Texas, also carried alleles that were absent from all other North American *Canis* species, and speculated that these alleles may represent unique red wolf diversity that was lost as a result of captive breeding and genetic drift. Heppenheimer et al. (2018a) also suggested that Gulf Coast canids may be an important reservoir of unique red wolf diversity, but that further research was needed to determine the persistence of red wolf ancestry in this region.

Such genome-wide ancestry surveys have not been exhaustively conducted on canid populations across the southeast, particularly within the geographic locations of the red wolf founders captured in the 1980s. The goal of our study was to evaluate broad geographic

patterns of red wolf genome-wide ancestry in southeastern canids on a recent timescale (i.e., within the last 5 decades). Such genome-wide methods are preferred to microsatellite-based methods due to the lack of ascertainment bias and increased representation across the genome that supports ancestry analyses (e.g., Muñoz et al. 2017). Our objectives were to survey the southeastern United States for evidence of previously undetected red wolf ancestry and compare levels of red wolf allele sharing and ancestry among the distinct red wolf-coyote admixture zones. Further, we evaluate canid samples originating from Texas in the 1970s that were trapped as part of the red wolf captive breeding initiative but ultimately excluded from the program based on phenotypic criteria. Previous mtDNA analyses of 1970s samples from this region indicated that this population primarily carried coyote-like or gray wolf haplotypes (Wayne and Jenks 1991). Genome-wide ancestry, however, has not been evaluated in these historical Texas samples. We predict that these canids will harbor allelic diversity that is absent from the contemporary southeastern coyote and captive red wolf population, with the origin of these alleles either derived from admixture events with historically sympatric gray wolves or representing unique red wolf diversity.

Methods

Study Areas and Sample Collection

Blood and tissue samples from 339 canids were collected from 22 states in the United States, one Canadian province, and 4 US captive breeding facilities (Supplementary Table S1). Of these, 35 samples originated from the coyote historical range (Hinton et al. 2019) and are the best available representation of putatively unadmixed coyotes (hereafter, referred to as reference coyotes). Additionally, 260 coyotes were collected post-1980 from 10 southeastern US states that included geographic regions that likely experienced gene flow with red wolves (e.g., Albemarle Peninsula of North Carolina $n = 46$; Louisiana and Texas $n = 14$) (Figure 1; Supplementary Table S1), with the Texas samples including the 2 individuals from Galveston Island that were previously identified to carry unique red wolf alleles (Heppenheimer et al. 2018a). All captive red wolf samples were collected in 2014 or earlier under active management efforts in North Carolina (Supplementary Table S1). Coyote samples were obtained through state management programs, government organizations archives, on-going research efforts, donation of legally hunted or trapped animals, or opportunistic sampling of roadkill.

In addition to the contemporary samples, we included 2 sets of historical samples. The first set included 10 canids collected in Texas during the mid-1970s as part of the red wolf captive breeding initiative but were not selected to be founders (hereafter, 1970s Texas canids). No genetic criteria, however, were considered in this initial evaluation. We completed mitochondrial DNA (mtDNA) sequencing for these 10 canids in Supplementary Appendix S1. The second set included 7 captive-born and wild-released red wolves (hereafter, 1991 red wolves), which were included as positive controls to evaluate the accuracy of ancestry inference methods.

A total of 40 captive red wolf samples (hereafter, reference red wolves) were obtained from the archives of the Point Defiance Zoo and Aquarium (Washington, DC, United States), the Know Wonder Museum of Life and Science (North Carolina, United States), the Wolf Conservation Center (New York, United States), and the Reflection Riding Arboretum (Tennessee, United States). While no original, wild-born genetic founders of the captive red wolf population were available for this study, the reference red wolves were

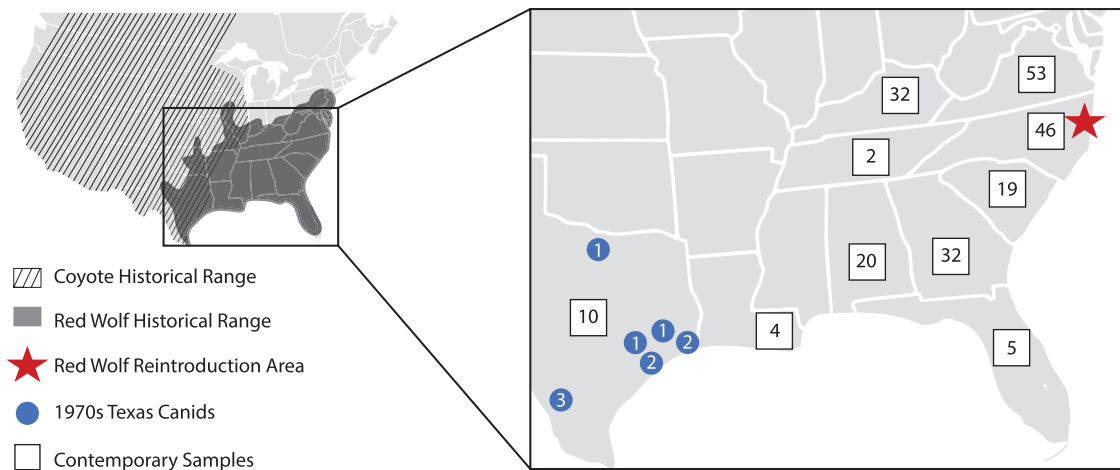


Figure 1. Map of southeastern canid sampling locations and sample sizes. For simplicity, contemporary sampling locations are shown at the state level. Complete sampling details for all individuals are given in [Supplementary Table S1](#). The coyote historical range was adapted from (Hody and Kays 2018) and the red wolf historical range was adapted from <https://www.fws.gov/southeast/wildlife/mammals/red-wolf/>.

representative of all 12 founder lineages. Red fox (*Vulpes vulpes*) samples ($n = 35$), which were included for analyses that required an outgroup, were collected from Zurich, Switzerland (sampling details described in [DeCandia et al. 2019](#)). Prior publication information for RADseq data on all samples is given in [Supplementary Table S1](#). All samples were collected under Princeton University IACUC #1961A-13.

DNA Extraction and RAD Sequencing

We extracted high molecular weight DNA from blood or tissue samples with the Qiagen DNeasy Blood and Tissue Kit protocol or the BioSprint 96 DNA Blood Kit in combination with a Thermo Scientific KingFisher Flex Purification platform, following manufacturer instructions. DNA quality was assessed on a 1% agarose gel and only samples with high molecular weight DNA were retained for sequencing. We quantified DNA with PicoGreen or Qubit 2.0 fluorometry and samples were standardized to 5 ng/ μ L.

We prepared restriction-site associated DNA (RADseq) libraries following a modified version of the protocol described by [Ali et al. \(2016\)](#), with full details given in [Heppenheimer et al. \(2018b\)](#). In brief, DNA was digested with *Sbf1* restriction enzyme, followed by the ligation of a unique 8 bp barcode adapter to each sample. We sheared DNA to 400 bp using a Covaris LE220 and a streptavidin bead binding assay was used to enrich for adapter-ligated fragments. We used the NEBnext Ultra or NEBnext Ultra II DNA library Prep Kit and selected for 300–400 bp insert sizes. Final libraries were sequenced on 2 lanes of an Illumina HiSeq 2500 platform (2x150nt) in pools of 96 samples per lane.

Variant Calling

We filtered forward and reverse raw sequencing reads to retain those containing the remnant *Sbf1* restriction enzyme cut site and one of the possible unique 8-mer barcodes using a custom perl script (`flip_trim_sbf1_170601.pl`, see [Supplementary Material](#)). We used the Stacks pipeline for paired-end RADseq data to discover variants ([Catchen et al. 2013](#)). Sequence data was demultiplexed and for polymerase chain reaction (PCR) duplicates were removed in Stacks v. 1.42, as in [Heppenheimer et al. \(2018b\)](#). This reduced computational time and allowed for greater modularity in sample

selection, as the Stacks 2.0 pipeline ([Rochette et al. 2019](#)) combines the PCR duplicate filtering with SNP calling, and therefore samples with low read counts could be removed prior to variant calling. Samples with >500 000 read pairs assigned to their unique barcode following the removal of PCR duplicates were mapped as read pairs to the CanFam3.1 assembly of the dog genome in STAMPY v 1.0.21 ([Lunter and Goodson 2011](#)). We then sorted the resulting SAM files, filtered for a minimum MAPQ score of 96, and exported as BAM files in samtools v 0.1.18 ([Li et al. 2009](#)). We used the *gstacks* module of Stacks v. 2.2 with the Maruki model of SNP calling for low coverage data (`--marukilow`; [Maruki and Lynch 2017](#)). To further increase confidence in inferences from low coverage data, we decreased the alpha threshold from 0.05 to 0.01 for both SNP discovery and genotype calling (`--var-alpha 0.01` and `--gt-alpha 0.01`, respectively).

We filtered genotype calls using VCFtools v. 0.1.12b ([Danecek et al. 2011](#)) to remove the following: 1) loci with >10% missing data across all individuals, 2) singletons and private doubletons, and 3) individuals with >20% missing data. For the subset of analyses that required genotypes to be filtered for linkage disequilibrium (LD), we removed loci with >50% correlation using Plink v. 1.90b3i (`--indep-pairwise 50 5 0.5`) (e.g., [Heppenheimer et al. 2018b](#)). We completed the calling and filtering of genotypes for 2 separate datasets, with and without an outgroup: 1) red wolf and coyote samples only, hereafter, *Canis* dataset, and 2) red wolf, coyote, and red fox samples, hereafter, *Canis/Vulpes* dataset.

Diversity Statistics

We estimated observed and expected heterozygosity (H_o and H_e , respectively), and private allele counts using Stacks for all *Canis* sampling groups over the unlinked SNP dataset. We calculated private allelic richness using a rarefaction approach implemented in ADZE ([Szpiech et al. 2008](#)), with a maximum standardized sample size (g) of 14, the highest obtainable given a tolerance of 10% missing data per group.

Population Structure and Allele Sharing Analyses

We conducted Principal Component Analyses (PCA) in flashpca ([Abraham and Inouye 2014](#)) to visualize clustering and identify

outliers. Outlier clusters were removed and PCAs were repeated. PCAs were then conducted for both the *Canis* and *Canis/Vulpes* LD-filtered datasets. Following outlier removal, we evaluated population structure in the *Canis* LD-filtered dataset using ADMIXTURE (Alexander et al. 2009) for $K = 1-10$ with the cross-validation (CV) flag, and considered the lowest CV value indicative of the best-fit K .

To evaluate the number of shared private alleles (S_{par}) among red wolves and various combinations of southeastern canids, we implemented a rarefaction approach in ADZE (Szpiech et al. 2008). This approach evaluated the number of shared alleles between a standardized sample size of the 2 focal groups that were absent from all other groups. The focal groups considered included: reference coyotes, Texas and Louisiana coyotes, North Carolina coyotes, and all other southeastern coyotes. The standardized sample size (g) was set to 18, the maximum obtainable given the tolerance threshold of 10% missing data per group.

Detection of Introgression

To detect red wolf introgression among groups of southeastern canids, we calculated the D -statistic (also known as ABBA/BABA test; Durand et al. 2011; equation 1). This method determines if the proportion of shared-derived alleles between putatively admixed coyotes and red wolves exceeds that which would be expected under incomplete lineage sorting alone (i.e., no recent introgression). Specifically, the D -statistic measures the excess of shared-derived sites between a potential introgressor (P3, reference red wolves) and a putatively admixed group (P2, southeastern canids; ABBA sites) over the shared-derived sites between P3 and a third group that is assumed to be unadmixed and sister to the P2 group (P1, reference coyotes; BABA sites). Under an incomplete lineage sorting scenario with no subsequent gene flow, the frequency of ABBA and BABA sites are expected to be equal, and D should be approximately zero. Alternatively, when interbreeding has occurred between the P3 and P2 groups, D is expected to be positive.

As we had multiple samples from each putatively admixed group, we implemented an allele frequency-based approach to define ABBA and BABA states, following Martin et al. (2013). In brief, using the *Canis/Vulpes* dataset, we first retained sites fixed in the red fox outgroup. We then calculated the frequency of the derived allele in the P1, P2, and P3 groups. ABBA and BABA states were calculated as a number between 0 and 1 to reflect the extent to which the population-level allele frequencies were skewed towards either the ABBA or BABA configuration. We evaluated significance using a block jackknife approach with a block size of 1.75 Mb. The D -statistic was calculated separately over 4 combinations of southeastern canids as P2 (Figure 5; Supplementary Table S3).

To directly evaluate the proportion of the genome shared between each possible P2 and red wolves, we calculated the admixture proportion, f (Durand et al. 2011; equation 8), for all significant values of D . This approach compares the observed excess of ABBA sites to what would be expected under complete genome replacement, that is, if P2 and P3 represented the same genetic population. To accomplish this, reference red wolves were randomly split into 2 equal-sized groups (P3a & P3b), with P3b used to represent a complete genome replacement population.

Local Ancestry Inference

We inferred local ancestry for all SNPs in each putatively admixed individual using a 2-layer hidden Markov model implemented in ELAI (Guan 2014). This approach evaluates LD within and between

the pre-defined parental groups to return allele dosage scores between 0 and 2 (for 2-way admixture) at each SNP, reflecting the most likely ancestry proportion at that site. That is, an allele dosage of 0 or 2 indicate the homozygous states, and an allele dosage of 1 is the heterozygous state.

We analyzed 96 532 unphased autosomal SNPs that had not been filtered for LD and applied only a filter for global minor allele frequency $> 1\%$, which improved accuracy in preliminary tests. The number of upper-layer clusters (-C) was set to 2 (i.e., red wolves and reference coyotes) and the lower-layer clusters (-c) was set to 10 (5 \times the value of C, as recommended). To account for both uncertainty in the precise timing of admixture and increased complexity in admixture scenarios (i.e., beyond the “single pulse” assumption), we analyzed 4 possible values of the admixture generation parameter (-mg): 5, 10, 15, and 20 generations since admixture. We implemented ELAI 3 times for each admixture generation value with 30 EM steps, and averaged results over all 12 independent runs. We considered sites with allele dosage between 0.8 and 1.8 to be heterozygous, and sites with allele dosage > 1.8 to be homozygous for the red wolf allele, following thresholds proposed by Seixas et al. (2018).

Results

Of the final *Canis* and *Vulpes* samples retained for analyses ($n = 350$), raw read counts ranged 551 390 – 16 558 569 (average: 2 059 232.21) per sample, with an average of 18.21% of reads discarded as PCR duplicates (Supplementary Table S2). Average mappability (i.e., percentage of reads mapped) to the dog genome was 86.93% and 69.46% for all *Canis* and *Vulpes* samples, respectively. For the *Canis* dataset, we retained 221 322 unfiltered biallelic SNPs and 182 667 SNPs after filtering for LD. Further, for the *Canis/Vulpes* genotype set, 164 597 unlinked autosomal SNPs were fixed in red foxes and retained for analysis. In all cases, total genotyping rate was approximately 96%. Heterozygosity was highest in the reference coyotes ($H_E = 0.035$), lowest in both groups of red wolves (1991 red wolves, $H_E = 0.016$; reference red wolves, $H_E = 0.018$), and moderate in all southeastern canid groups (e.g., North Carolina Coyotes, $H_E = 0.028$; Table 1). These relative trends were also reflected in both allelic richness and private allelic richness (Table 1) and were consistent with the known demographic history of each group. Furthermore, the private allelic richness rarefaction curve for the reference red wolves converged at a minimum sample size of 8, indicating that unique red wolf diversity in the captive population was sufficiently represented, despite not including original wild-born genetic founders in the sample set (Supplementary Figure S1).

Population Structure and Allele Sharing Analyses

PCA revealed species delineation along PC1 (1.4% of variance explained), while PC2 (1.2% of variance explained) primarily reflected intra-specific variation (Figure 2A). Overall, southeastern coyotes tended to have intermediate placement along PC1, as did most canid samples from the 1970s Texas canids, whereas the 1991 wild red wolf samples clustered entirely within the reference red wolf group (Figure 2A). ADMIXTURE analysis indicated that 2 genetic populations were best represented by the data (CV = 0.094; Supplementary Figure S2), corresponding to reference and 1991 red wolves and all other samples, with notable canid admixtures of coyotes and red wolves across several southeastern sampling locations (Figure 2B). Assignments to the red wolf cluster were highest in the 1970s Texas

Table 1. Summary statistics for all *Canis* sampling groups ($n = 312$) genotyped for 182 570 unlinked biallelic SNPs

Group	Location ^d	N	Count	Private allele			
				Richness	H _O	H _E	A _R
Reference Coyotes	AZ, CA, ID, NE, NM, NV, OK, SK, WA, WY, MO, KS	32	9374	0.091	0.033	0.035	1.184
Reference Red Wolves	Captive	40	2169	0.010	0.018	0.018	1.041
Southeastern Coyotes	AL, FL, GA, SC, KY, TN, VA	163	26 293	0.068	0.028	0.029	1.157
North Carolina Coyotes	NC	46	3237	0.066	0.028	0.028	1.153
Texas and Louisiana Coyotes	TX and LA	14	1690	0.055	0.027	0.027	1.139
1970s Texas Canids	TX	10	881	0.057	0.028	0.027	1.150
1991 Red Wolves	Captive	7	122	0.005	0.019	0.016	1.031

n , sample size; H_O, Observed Heterozygosity; H_E, Expected Heterozygosity; A_R, Allelic Richness. Unless otherwise noted, samples were collected post-1980.

^dAdditional details regarding sample collection are available in [Supplementary Table S1](#).

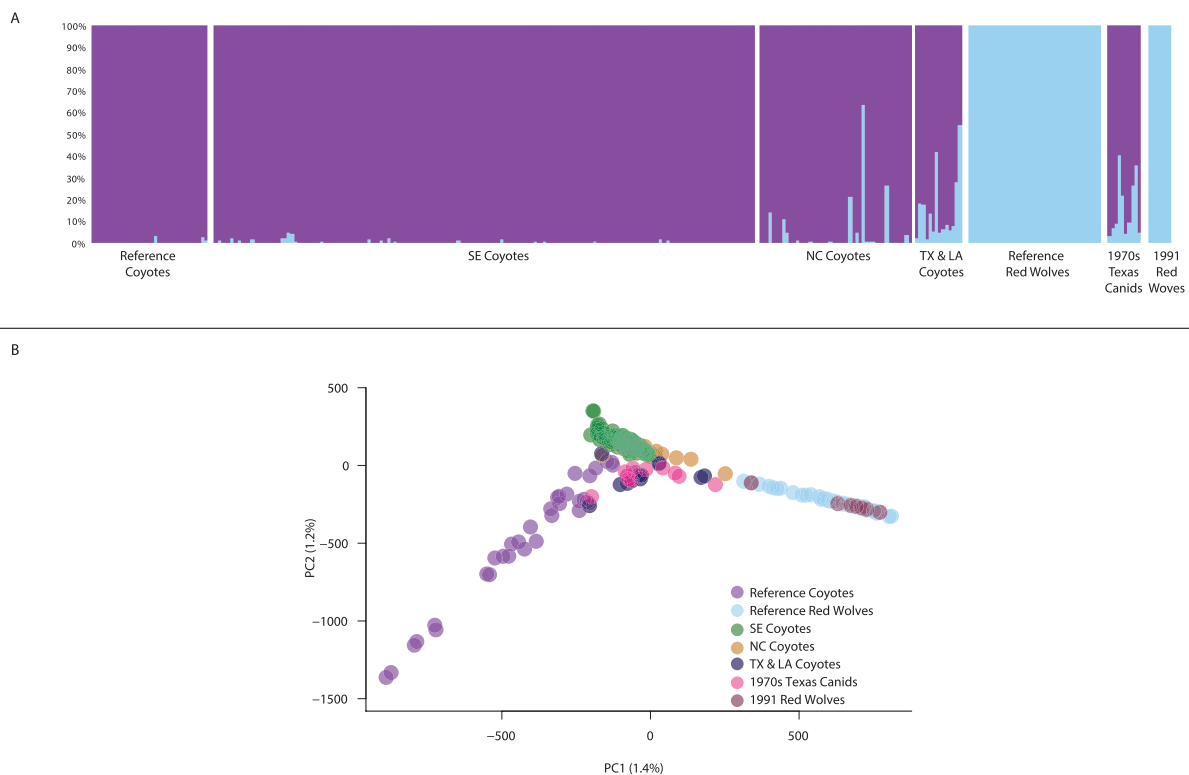


Figure 2. (A) Admixture and (B) Principal Component Analysis (PCA) of all 315 canids genotyped for 182 667 genome-wide SNPs. SE Coyotes refers to all Southeastern coyotes, except those sampled from NC, LA, and TX.

canids (Average $Q_{\text{red wolf}} = 0.16$), followed by Texas and Louisiana (Average $Q_{\text{red wolf}} = 0.15$), and North Carolina coyotes (Average $Q_{\text{red wolf}} = 0.03$). In the remaining southeastern sampling areas, assignment to the red wolf cluster was minimal (Average $Q_{\text{red wolf}} = 0.002$; [Figure 2B](#)). Furthermore, the 1991 wild red wolves had 100% assignment to the reference red wolf cluster. Three reference coyotes exhibited small assignments to red wolf cluster ($Q_{\text{red wolf}} = 0.03, 0.03$, and 0.01 , respectively; [Figure 2B](#)). To prevent circular analyses resulting from on-going gene flow between southeastern and historical range coyotes, these individuals were removed as reference individuals from the subsequent ancestry analyses.

We observed similar trends in the private allele sharing analysis, as reference red wolves shared the most private alleles with the 1970s Texas canids (SPAr = 0.006), followed by the Texas and Louisiana

coyotes (SPAr = 0.004), and North Carolina coyotes (SPAr = 0.003; [Figure 3](#)). Private allele sharing between reference red wolves and all other southeastern sampling locations was lower (SPAr = 0.002) and approximately comparable to that observed for reference coyotes ([Figure 3](#)).

Detection of Introgression

When all contemporary southeastern coyotes were considered as a single group, D was weakly positive, indicating a slight excess of shared derived sites with reference red wolves, though this value was not significantly different from zero ($D = 9.5 \times 10^{-3}$, $P = 0.345$; [Figure 4](#); [Supplementary Table S3](#)). When coyotes sampled from areas known to have experienced gene flow with red wolves (i.e., North Carolina, Texas, Louisiana) were removed, D decreased

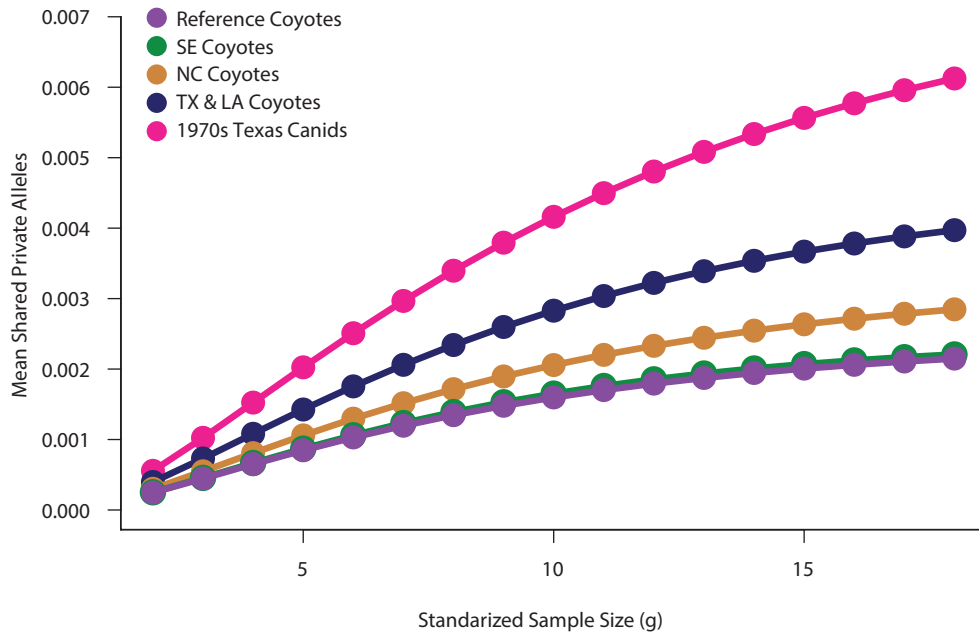


Figure 3. Private allele sharing with captive red wolves as a function of standardized sample size across 5 groups of coyotes.

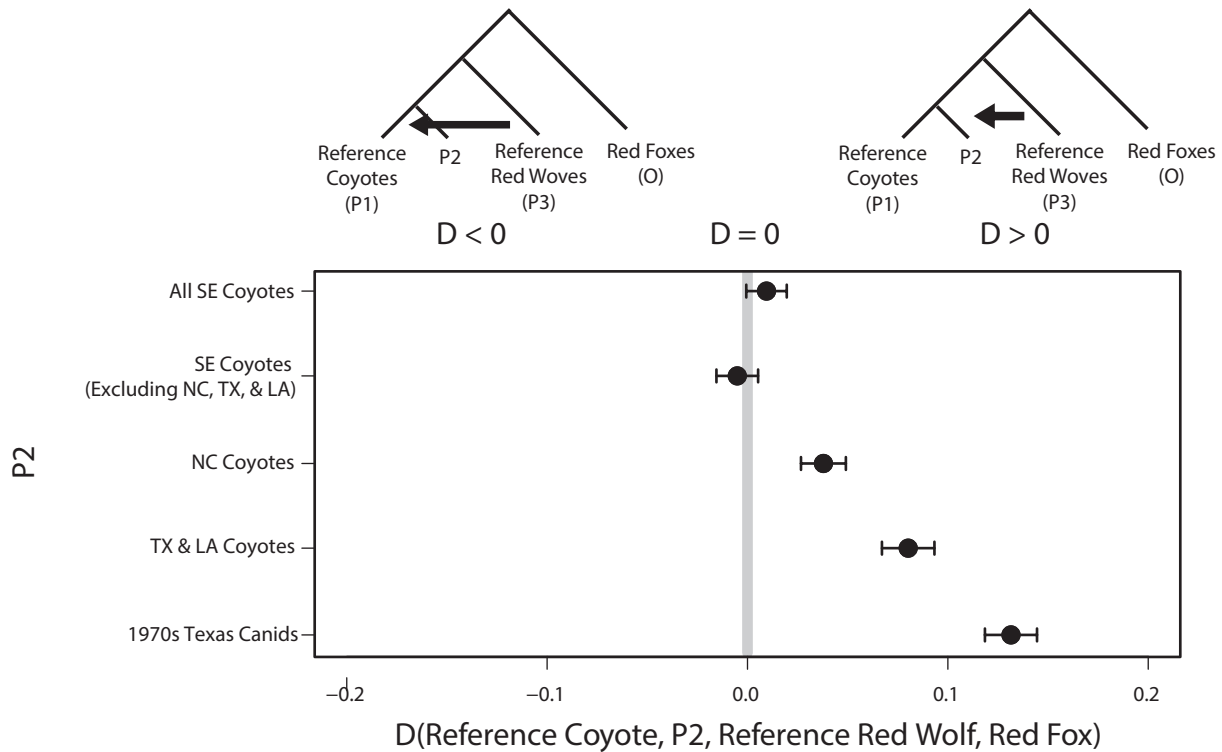


Figure 4. *D*-Statistics for all possible P2 groups. Error bars represent the standard error estimated via jackknife resampling with a block size of 1.75 MB. Under incomplete lineage sorting, *D* is expected to be zero. Positive values of *D* indicate introgression between P3 (reference red wolves) and P2. Significant negative values of *D* were not observed but would indicate introgression between P3 (reference red wolves) and P1 (reference coyotes).

slightly and remained nonsignificant ($D = -5.1 \times 10^{-3}$, $P = 0.628$; Figure 4; Supplementary Table S3). However, when North Carolina and Texas and Louisiana coyotes were each considered separately, we observed a significant signal of red wolf introgression in each group ($D_{NC} = 0.038$, $P < 0.001$; $D_{TX\&LA} = 0.080$,

$P < 0.001$; Figure 4; Supplementary Table S3). Interestingly, the admixture proportion was approximately twice as high in Texas and Louisiana coyotes ($f = 0.047$) than in North Carolina coyotes ($f = 0.021$; Supplementary Table S3). Furthermore, we also detected a significant excess of shared derived alleles between red

wolves and the 1970s Texas canids ($D = 0.132$, $P < 0.001$), at an even higher level of genome replacement ($f = 0.085$; Figure 4; Supplementary Table S3).

Local Ancestry Inference

Overall, inferred genome-wide ancestry values were highly correlated across both methods (ELAI and ADMIXTURE, Pearson's $r = 0.98$, $P < 0.001$; Supplementary Figure S3). We found the highest inferred red wolf ancestry in the 1970s Texas canids, with intermediate estimates in Louisiana and Texas, and in North Carolina, and lowest in all other southeastern sampling locations (Supplementary Table S4). However, we observed substantial individual variation in genome-wide ancestry. In the 1970s Texas canids, genome-wide red wolf ancestry ranged from 2 to 51% (mean = 15%, standard deviation [SD] = 17%; Supplementary Table S4). Similarly, in Louisiana and Texas, genome-wide red wolf ancestry ranged from 1% to 45% (mean = 10%, SD = 14%; Supplementary Table S4). Within this region, the highest levels of red wolf ancestry were observed in the 2 Galveston Island animals that were previously documented to carry red wolf alleles, at 45% and 39%, respectively. Further, 2 Louisiana coyotes from Plaquemines Parish exhibited moderate levels of red wolf ancestry (18% and 16%, respectively), while all other animals from Texas and Louisiana had relatively low levels of red wolf ancestry (i.e., 1–7%; Supplementary Table S4). In North Carolina, genome-wide ancestry ranged from 1% to 71% (mean = 8%,

SD = 12%), with 26 individuals exhibiting >5% red wolf ancestry, and 6 individuals with >10% red wolf ancestry. Outside of the regions known to have experienced recent gene flow with red wolves, red wolf ancestry was consistently low, ranging from 1% to 9% (mean = 5%, SD = 2%; Supplementary Table S4).

In the 1970s Texas canids, red wolf ancestry tracts were an average of 7780 Kb and 4195 Kb, for heterozygous and homozygous segments, respectively (Supplementary Table S4). In the Texas and Louisiana coyotes, heterozygous ancestry tracts were shorter (5091 Kb), whereas homozygous tracts were similar in length (4503 Kb; Supplementary Table S4). In North Carolina, heterozygous red wolf ancestry tracts were comparable to the 1970s Texas canids, with an average length of 8017 Kb, and homozygous tracts were longer than in the 1970s Texas canids (5804 Kb; Supplementary Table S4). In the remaining southeastern coyotes, average red wolf ancestry tracts were 6348 Kb and 4850, for heterozygous and homozygous segments, respectively (Supplementary Table S4). Local ancestry assignments for representative individuals from all sampled regions across the 38 autosomes are given in Figure 5A–C. Selected individuals exhibited overall ancestry scores within 1 SD of the mean for each respective sampling region and had a known county of origin.

Discussion

We evaluated genome-wide red wolf ancestry in canids across the southeastern United States. Though red wolves have been declared

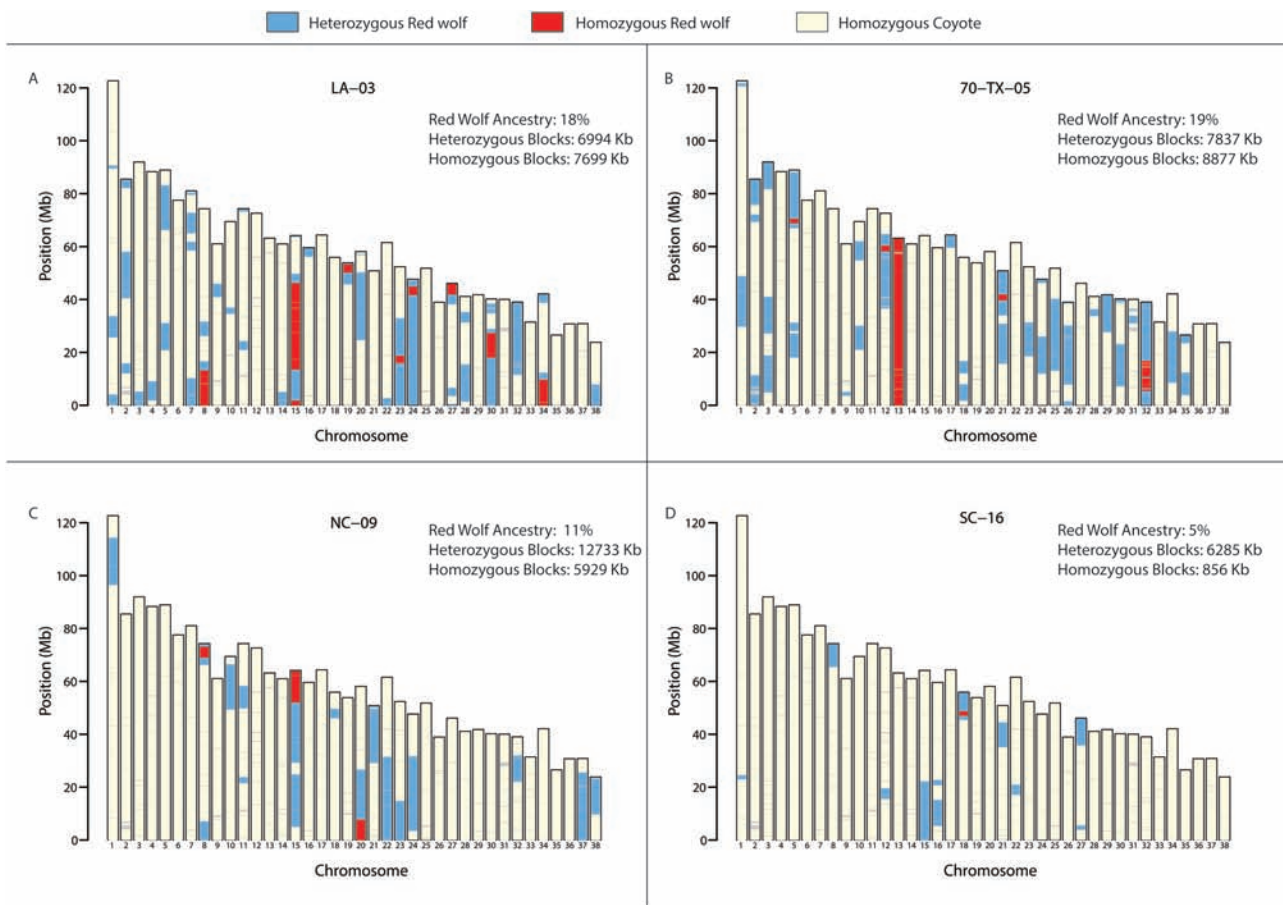


Figure 5. Local ancestry assignments across 38 autosomes for representative individuals sampled from (A) Louisiana (Plaquemines Parish), (B) 1970s Texas canid (Harris Co), (C) North Carolina (Tyrell Co.), and (D) South Carolina (McCormick Co.).

extirpated throughout most of this region for nearly 4 decades, the persistence of red wolf ancestry could be facilitated by isolating landscape features (e.g., islands; Hargreaves et al. 2010) or by assortative mating with respect to ancestry fractions (e.g., Murphy et al. 2019), which has been observed in the North Carolina admixture zone (e.g., Bohling et al. 2016; Hinton et al. 2018). Further, some red wolf alleles may be locally adapted and maintained by selection (i.e., adaptive introgression, Wellenreuther et al. 2018). We consistently identified 2 distinct geographic occurrences of red wolf ancestry in southeastern coyotes, concordant with previous reports of historical and on-going hybridization between coyotes and red wolves across Texas and Louisiana, and in North Carolina, respectively (Bohling et al. 2016; Heppenheimer et al. 2018a; Murphy et al. 2019). We found no compelling evidence of substantial red wolf ancestry in southeastern canids outside of the 2 known admixture zones, which have not been previously evaluated via molecular methods. We acknowledge these results are contingent upon the individuals we sampled, where more systematic sampling north of the historical admixture zones could demonstrate a spread of red wolf ancestry. However, our sampling east of the historical admixture zones was extensive, providing limited evidence of admixed coyotes carrying red wolf alleles eastward.

Though explicit analyses to determine the precise timing of admixture events between red wolves and coyotes were beyond the scope of this study, our observed trends were consistent with theoretical expectations for red wolf ancestry tracts to shorten with each subsequent generation as a result of recombination (e.g., Pool and Nielsen 2009). For instance, overall ancestry tract length was slightly shorter in the Texas and Louisiana samples (4797 kb across heterozygous and homozygous regions), where red wolves have been extirpated for nearly 4 decades, compared with the 1970s Texas canids and the North Carolina samples, where red wolves and coyotes were sympatric at the time of sampling (5969 kb and 6911 kb, respectively). Future studies may incorporate higher density SNP sets and complex statistical models (e.g., Liang and Nielsen 2014) to precisely estimate the timing of admixture in southeastern canid populations based on the mean and variance of ancestry tract lengths.

We observed that southeastern coyotes originating from outside of the 2 known admixture zones exhibited limited, but detectable levels of red wolf ancestry assignments. Shared ancestry can be due to interspecific gene flow mediated by long-distance dispersal events (Sasmal et al. 2019) or incomplete lineage sorting (i.e., shared ancestral polymorphisms; Degnan and Rosenberg 2009). Our analyses support the latter mechanism, although the *D*-statistic may lack the sensitivity to detect low rates of gene flow among the admixture zones given the large population size of southeastern coyotes (Durand et al. 2011; Kilgo et al. 2017). Within the 2 southeastern admixture zones, we observed considerable individual and geographic variation in red wolf ancestry. Approximately 11% of the North Carolina coyotes exhibited substantial genome-wide red wolf ancestry (ancestry > 15%). In contrast, 29% of contemporary Louisiana and Texas canids had similarly high levels of red wolf ancestry. This higher level of red wolf ancestry in contemporary Louisiana and Texas than in North Carolina likely reflects the success of management efforts to limit hybridization events in the latter region (e.g., Gese et al. 2015).

Historical Comparison of Genomic Ancestry

To determine how red wolf ancestry has changed since extirpation in the 1970s, we examine red wolf ancestry in 10 canids collected in Texas in the mid-1970s. On average, allelic diversity was marginally

higher in the 1970s Texas canids relative to their contemporaries, suggesting that both red wolf ancestry and unique alleles have been lost from Texas and Louisiana over the last 4 decades. Alleles private to this area (both from the 1970s and present day) may represent variation that was unique to the historical red wolf population, which were lost in the bottlenecked of establishing their captive breeding colony (i.e., ghost alleles; Bouzat et al. 1998; Groombridge et al. 2000). Our analysis did not evaluate whether these ghost alleles originated in the red wolf genome or if they derive from admixture events with historically sympatric gray wolves that are now extirpated (Hailer and Leonard 2008).

The existence of unique red wolf alleles that are absent from reference populations of North American canids has recently been used as partial evidence for the species status of red wolves (National Academies of Sciences, Engineering, and Medicine 2019). It is critical to emphasize that unique alleles may occur among populations of conspecifics under many evolutionary and demographic scenarios, including selection associated with environmental variables (e.g., gray wolves, Schweizer et al. 2016) or reduced effective population sizes (e.g., coyote range expansion, Heppenheimer et al. 2018b). Overall, red wolves are expected to carry unique alleles regardless of evolutionary origin and the interpretation of our results neither confirms nor rebuts any of the proposed evolutionary scenarios (for review, Waples et al. 2018).

Admixtures of red wolves and coyotes outside of the North Carolina recovery area likely represent an important reservoir of red wolf genetic diversity that has been lost in the captive breeding and wild populations due to demographic contraction over the last century (Heppenheimer et al. 2018b). Furthermore, it is important to recognize that estimates of red wolf ancestry are based entirely on the genetic diversity that is found in the captive population; thus, estimates of red wolf ancestry in wild canids are likely to be underestimated as the captive population is not expected to represent the full diversity of wild red wolves prior to extirpation. Given our findings that genome-wide red wolf ancestry in the southeast is highest in Texas and Louisiana, these regions may represent a candidate for the location for future red wolf reintroduction or genetic augmentation efforts.

Lastly, one aspect that remains unevaluated is the selective benefit of red wolf-coyote hybridization events and subsequent introgression. Though hybridization with coyotes is largely viewed as a threat to red wolf recovery (e.g., Hinton 2013), additional evolutionary scenarios are possible. For instance, hybridization could be adaptive for expanding coyotes by introducing genetic material that has already been filtered by natural selection as beneficial in the southeastern environment (for review: Pfennig et al. 2016). This process of adaptive introgression as a facilitator of coyote range expansion has been demonstrated in northeastern coyotes, where hybridization with remnant wolf populations in eastern Canada (gray wolves, *Canis lupus* and eastern wolves, *Canis lycaon*) has resulted in selection on genes associated with body size and skeletal proportions (vonHoldt et al. 2016). Additionally, for the red wolf population, which is characterized by severe inbreeding (Brzeski et al. 2014), coyote genetic material may increase the adaptive potential of the population (i.e., genetic rescue; for review: Whiteley et al. 2015). In this study, sample sizes were prohibitively small and characterized by highly related individuals (e.g., the Galveston Island samples; identity-by-state = 0.93; Heppenheimer et al. 2018a) to confidently evaluate patterns of selective introgression. Future studies should incorporate finer-scale sampling of canids along the Gulf Coast, along

with wild reintroduced red wolves from North Carolina, to address these intriguing outstanding questions.

Supplementary Material

Supplementary data are available at *Journal of Heredity* online.

Funding

This work was supported by the Point Defiance Zoo & Aquarium Dr. Holly Reed Conservation Fund to B.M.v.H., and the National Science Foundation Postdoctoral Research Fellowship in Biology under Grant No. 1523859 to K.E.B.

Acknowledgments

We thank William Waddell for facilitating the collection of captive red wolf samples. We are grateful to Dr. Linda Rutledge for her initial contributions to this project, and to Alexandra DeCandia for contributing the red fox samples. We thank the Oklahoma Museum of Natural History, which provided samples under loan agreement G.2016.3, 4318. We also thank Laura Palmer, Don Shumaker, Rebecca Bose, Sherry Samuels, Trish Gailmard, Ron Wooten, Heather Montgomery, Margaret Earthman, Steve Fain, and numerous hunters and trappers who donated coyote samples.

Data Availability

The primary data underlying these analyses has been deposited as follows:

- Demultiplexed reads: NCBI SRA (PRJNA605471)
- CanFam3.1 Mapped Bam files: NCBI SRA (PRJNA605471)
- mtDNA haplotypes: Genbank (MT075538-MT075545)
- Sampling locations: [Supplementary Table S1](#).

References

- Abraham G, Inouye M. 2014. Fast principal component analysis of large-scale genome-wide data. *PLoS One*. 9:e93766.
- Adams JR, Kelly BT, Waits LP. 2003. Using faecal DNA sampling and GIS to monitor hybridization between red wolves (*Canis rufus*) and coyotes (*Canis latrans*). *Mol Ecol*. 12:2175–2186.
- Adams JR, Lucash C, Schutte L, Waits LP. 2007. Locating hybrid individuals in the red wolf (*Canis rufus*) experimental population area using a spatially targeted sampling strategy and faecal DNA genotyping. *Mol Ecol*. 16:1823–1834.
- Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 19:1655–1664.
- Ali OA, O'Rourke SM, Amish SJ, Meek MH, Luikart G, Jeffres C, Miller MR. 2016. RAD Capture (Rapture): flexible and efficient sequence-based genotyping. *Genetics*. 202:389–400.
- Bohling JH, Dellinger J, McVey JM, Cobb DT, Moorman CE, Waits LP. 2016. Describing a developing hybrid zone between red wolves and coyotes in eastern North Carolina, USA. *Evol Appl*. 9:791–804.
- Bohling JH, Waits LP. 2015. Factors influencing red wolf – coyote hybridization in eastern North. *Biol Conserv*. 184:108–116.
- Bouzat JL, Lewin HA, Paige KN. 1998. The ghost of genetic diversity past: historical DNA analysis of the greater prairie chicken. *Am Nat*. 152:1–6.
- Brzeski KE, Rabon DR Jr, Chamberlain MJ, Waits LP, Taylor SS. 2014. Inbreeding and inbreeding depression in endangered red wolves (*Canis rufus*). *Mol Ecol*. 23:4241–4255.
- Carley CJ. 2000. The red wolf (*Canis rufus*) recovery program: things they didn't tell me in school. In: Choate JR, editor. *Reflections of a naturalist: papers honoring professor Eugene D. Fleharty*. Hays (KS): Fort Hays State University. pp. 125–141.
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. 2013. Stacks: an analysis tool set for population genomics. *Mol Ecol*. 22:3124–3140.
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al.; 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. *Bioinformatics*. 27:2156–2158.
- DeCandia AL, Brzeski KE, Heppenheimer E, Caro CV, Camenisch G, Wandeler P, Driscoll C, vonHoldt BM. 2019. Urban colonization through multiple genetic lenses: the city-fox phenomenon revisited. *Ecol Evol*. 9:2046–2060.
- Degnan JH, Rosenberg NA. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol Evol*. 24:332–340.
- Durand EY, Patterson N, Reich D, Slatkin M. 2011. Testing for ancient admixture between closely related populations. *Mol Biol Evol*. 28:2239–2252.
- Gese EM, Knowlton FF, Adams JR, Beck K, Fuller TK, Murray DL, et al. 2015. Managing hybridization of a recovering endangered species: the red wolf *Canis rufus* as a case study. *Curr Zool*. 61:191–205.
- Giordano MR, Pace RM III. 2000. Morphometrics and movement patterns of coyote-like Canids in a southwest Louisiana marsh complex. *Proc Southeast Assoc Fish and Wildl Agencies*. 54:424–435.
- Groombridge JJ, Jones CG, Bruford MW, Nichols RA. 2000. 'Ghost' alleles of the Mauritius kestrel. *Nature*. 403:616.
- Guan Y. 2014. Detecting structure of haplotypes and local ancestry. *Genetics*. 196:625–642.
- Hailer F, Leonard JA. 2008. Hybridization among three native North American *Canis* species in a region of natural sympatry. *PLoS One*. 3:e3333.
- Hargreaves S, Maxted N, Hirano R., Abberton, M., Sköt, L., Ford-Lloyd, B. V. 2010. Islands as refugia of *Trifolium repens* genetic diversity. *Conservation Genetics*. 11:1317–1326.
- Hedrick PW, Fredrickson RJ. 2008. Captive breeding and the reintroduction of Mexican and red wolves. *Mol Ecol*. 17:344–350.
- Heppenheimer E, Brzeski KE, Hinton JW, Patterson BR, Rutledge LY, DeCandia AL, Wheelton T, Fain SR, Hohenlohe PA, Kays R, et al. 2018b. High genomic diversity and candidate genes under selection associated with range expansion in eastern coyote (*Canis latrans*) populations. *Ecol Evol*. 8:12641–12655.
- Heppenheimer E, Brzeski KE, Wooten R, Waddell W, Rutledge L, Chamberlain M, et al. 2018a. Rediscovery of red wolf ghost alleles in a canid population along the American gulf coast. *Genes*. 9:618.
- Hinton JW, Ashley AK, Dellinger JA, Gittleman JL, vanManen F, Chamberlain MJ. 2017. Using diets of *Canis* breeding pairs to assess resource partitioning between sympatric red wolves and coyotes. *J Mammal*. 98:475–488.
- Hinton JW, Chamberlain MJ, Rabon DR. 2013. Red Wolf (*Canis rufus*) recovery: a review with suggestions for future research. *Animals (Basel)*. 3:722–744.
- Hinton JW, Gittleman JL, van Manen FT, Chamberlain MJ. 2018. Size-assortative choice and mate availability influences hybridization between red wolves (*Canis rufus*) and coyotes (*Canis latrans*). *Ecol Evol*. 8:3927–3940.
- Hinton JW, Heppenheimer E, West KM, Caudill D, Karlin ML, Kilgo JC, Mayer JJ, Miller KV, Walch M, vonHoldt B, et al. 2019. Geographic patterns in morphometric and genetic variation for coyote populations with emphasis on southeastern coyotes. *Ecol Evol*. 9:3389–3404.
- Hody JW, Kays R. 2018. Mapping the expansion of coyotes (*Canis latrans*) across America. *ZooKeys*. 97:81–97.
- Kelly BT, Miller PS, Seal US. 1999. *Population and habitat viability assessment workshop for the red wolf (Canis rufus)*. Apple Valley (MN): IUCN/SSC Conservation Breeding Specialist Group.
- Kilgo JC, Shaw CE, Vukovich M, Conroy MJ, Ruth C. 2017. Reproductive characteristics of a coyote population before and during exploitation. *J Wildl Manag*. 81:1386–1393.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 25:2078–2079.

- Liang M, Nielsen R. 2014. The lengths of admixture tracts. *Genetics*. 197:953–967.
- Lunter G, Goodson M. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res*. 21:936–939.
- Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, Simpson F, Blaxter M, Manica A, Mallet J, Jiggins CD. 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res*. 23:1817–1828.
- Maruki T, Lynch M. 2017. Genotype calling from population-genomic sequencing data. *G3 (Bethesda)*. 7:1393–1404.
- Mech LD, Nowak RM. 2010. Systematic status of wild *Canis* in north-central Texas. *Southeast Nat*. 9:587–594.
- Muñoz I, Henriques D, Jara L, Johnston JS, Chávez-Galarza J, De La Rúa P, Pinto MA. 2017. SNPs selected by information content outperform randomly selected microsatellite loci for delineating genetic identification and introgression in the endangered dark European honeybee (*Apis mellifera mellifera*). *Mol Ecol Resour*. 17:783–795.
- Murphy SM, Adams JR, Cox JJ, Waits LP. 2019. Substantial red wolf genetic ancestry persists in wild canids of southwestern Louisiana. *Conservation Letters*. 12:e12621.
- National Academies of Sciences, Engineering, and Medicine. 2019. *Evaluating the taxonomic status of the Mexican gray wolf and the red wolf*. Washington (DC): The National Academies Press. doi:10.17226/25351.
- Nowak RM. 2002. The original status of wolves in eastern North America. *Southeast Nat*. 1:95–130.
- Parker G. 1995. *Eastern coyote: the story of its success*. Halifax (NS): Nimbus Publishing.
- Pfennig KS, Kelly AL, Pierce AA. 2016. Hybridization as a facilitator of species range expansion. *Proc Royal Soc B*. 283:20161329. doi:10.1098/rspb.2016.1329
- Pool JE, Nielsen R. 2009. Inference of historical changes in migration rate from the lengths of migrant tracts. *Genetics*. 181:711–719.
- Rochette R, Rivera-Colon AG, Crachen JM. 2019. Stacks 2: analytical methods for paired-end sequencing improve radseq-based population genomics. *Mol Ecol*. doi:10.1111/mec.15253
- Sasmal I, Moorman CE, Swingen MB, Datta S, DePerno CS. 2019. Seasonal space use of transient and resident coyotes in North Carolina. *Can J Zool*. 97:326–331.
- Schweizer RM, Robinson J, Harrigan R, Silva P, Galverni M, Musiani M, Green RE, Novembre J, Wayne RK. 2016. Targeted capture and resequencing of 1040 genes reveal environmentally driven functional variation in grey wolves. *Mol Ecol*. 25:357–379.
- Seixas FA, Boursot P, Melo-Ferreira J. 2018. The genomic impact of historical hybridization with massive mitochondrial DNA introgression. *Genome Biol*. 19:91.
- Szpiech ZA, Jakobsson M, Rosenberg NA. 2008. ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*. 24:2498–2504.
- US Fish and Wildlife Service. 1989. *Red wolf recovery plan*. Atlanta (GA): U S. Fish and Wildlife Service. 110 pp.
- vonHoldt BM, Kays R, Pollinger JP, Wayne RK. 2016. Admixture mapping identifies introgressed genomic regions in North American canids. *Mol Ecol*. 25:2443–2453.
- Waples RS, Kays R, Fredrickson RJ, Pacifici K, Mills LS. 2018. Is the red wolf a listable unit under the US endangered species Act? *J Hered*. 109:585–597.
- Wayne RK, Jenks SM. 1991. Mitochondrial DNA analysis implying extensive hybridization of the endangered red wolf *Canis rufus*. *Nature*. 351:565–568.
- Wellenreuther M, Muñoz J, Chávez-Ríos JR, Hansson B, Cordero-Rivera A, Sánchez-Guillén RA. 2018. Molecular and ecological signatures of an expanding hybrid zone. *Ecol Evol*. 8:4793–4806.
- Whiteley AR, Fitzpatrick SW, Funk WC, Tallmon DA. 2015. Genetic rescue to the rescue. *Trends Ecol Evol*. 30:42–49.