2023

Genomic analyses of gray fox lineages suggest ancient divergence and secondary contact in the southern Great Plains

Elizabeth M. Kierepka

Sophie Preckler-Quisquater

Dawn M. reding

Antoinette Piaggio

Seth P.D. Riley

See next page for additional authors

Authors

Elizabeth M. Kierepka, Sophie Preckler-Quisquater, Dawn M. reding, Antoinette Piaggio, Seth P.D. Riley, and Benjamin N. Sacks

Journal of Heredity, 2023, **114**, 110–119
https://doi.org/10.1093/jhered/esac060
Advance access publication 3 November 2022
Original Article

American Genetic Association

OXFORD

Original Article

Genomic analyses of gray fox lineages suggest ancient divergence and secondary contact in the southern Great Plains

Elizabeth M. Kierepka[1,2], Sophie Preckler-Quisquater[2], Dawn M. Reding[3], Antoinette J. Piaggio[4], Seth P. D. Riley[5,6], Benjamin N. Sacks[2,7]

[1]North Carolina Museum of Natural Sciences, Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, United States,
[2]Mammalian Ecology and Conservation Unit, Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California-Davis, Davis, CA, United States,
[3]Department of Biology, Luther College, Decorah, IA, United States,
[4]USDA, Wildlife Services, National Wildlife Research Center, Wildlife Genetics Lab, Fort Collins, CO, United States,
[5]National Park Service, Santa Monica Mountains National Recreation Area, Thousand Oaks, CA, United States,
[6]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles, CA, United States,
[7]Department of Population Health and Reproduction, School of Veterinary Medicine, University of California-Davis, Davis, CA, United States

Address correspondence to E.M. Kierepka at the address above, or e-mail: emkierep@ncsu.edu.

Corresponding Editor: Eduardo Eizirik

Abstract

The gray fox (*Urocyon cinereoargenteus*) lineage diverged from all other extant canids at their most basal node and is restricted to the Americas. Previous mitochondrial analysis from coastal populations identified deeply divergent (up to 1 Mya) eastern and western lineages that predate most intraspecific splits in carnivores. We conducted genotyping by sequencing and mitochondrial analysis on gray foxes sampled across North America to determine geographic concordance between nuclear and mitochondrial contact zones and divergence times. We also estimated the admixture within the contact zone between eastern and western gray foxes based on nuclear DNA. Both datasets confirmed that eastern and western lineages met in the southern Great Plains (i.e. Texas and Oklahoma), where they maintained high differentiation. Admixture was generally low, with the majority of admixed individuals carrying <10% ancestry from the other lineage. Divergence times confirmed a mid-Pleistocene split, similar to the mitochondrial estimates. Taken together, findings suggest gray fox lineages represent an ancient divergence event, far older than most intraspecific divergences in North American carnivores. Low admixture may reflect a relatively recent time since secondary contact (e.g. post-Pleistocene) or, alternatively, ecological or reproductive barriers between lineages. Though further research is needed to disentangle these factors, our genomic investigation suggests species-level divergence exists between eastern and western gray fox lineages.

Key words: genomics, gray fox, secondary contact, speciation, *Urocyon cinereoargenteus*

Introduction

Pleistocene glaciation events caused repeated distributional changes and isolation across numerous temperate taxa, which largely shaped current patterns of biodiversity in North America and Europe over the past 2 million years (Hewitt 2004; Shafer et al. 2010). Glaciers forced taxa into refugia where populations experienced prolonged isolation and independent evolutionary trajectories. Following glacial recession, many once-isolated refugial lineages were able to recolonize newly accessible habitats, often meeting in secondary contact. Depending on the time of isolation and evolutionary forces acting on isolated populations, the consequences of secondary contact have ranged from high levels of interbreeding (Latch et al. 2011; Slager et al. 2020) to reproductive isolation and speciation (Weir and Schluter 2004; Hope et al. 2016).

Secondary contact between lineages of wide-ranging mammals, such as carnivores, often entails high levels of admixture and little evidence of reproductive isolation, particularly in mid-continent North America (Cullingham et al. 2008; Aubry et al. 2009; Barton et al. 2012; Reding et al. 2012; Puckett et al. 2015). These intraspecific patterns tend to reflect high dispersal ability during interglacial periods, which results in shorter periods of isolation. Secondary contact among carnivores also can result in narrow hybrid zones with low levels of admixture and a high degree of reproductive isolation (i.e. interspecific), but usually reflects longer periods of isolation. Typically, these more divergent cases result in sister species, such as kit fox (*Vulpes macrotis*) and swift fox (*V. velox*; Mercure et al. 1993); marten species (*Martes* spp.; Stone et al. 2002; Dawson and Cook 2012),

Received June 20, 2022; Accepted November 2, 2022

and spotted skunk species (*Spilogale* spp.; Ferguson et al. 2017). In this study, we sought to better characterize a secondary contact zone between eastern and western lineages of gray fox (*Urocyon cinereoargenteus*) that potentially reflect cryptic sister species (Goddard et al. 2015).

Most North American carnivore species were identified over a century ago based on clear morphological characters, so it is unusual to discover deep divergence in a widespread carnivore based initially on genetic criteria. Nevertheless, such cryptic divergence was recently hypothesized to account for a >500 kya divergence between gray fox mitochondrial DNA sampled on either coast of the North American continent (Goddard et al. 2015). Additional mtDNA sequencing of samples from the 3,800 km of intervening habitat identified the location of the contact zone at the Great Plains (Reding et al. 2021). Because these studies were restricted to mitochondrial DNA, the magnitude of genome-wide divergence and mito-nuclear concordance remain unknown.

Carnivore contact zones in North America often occur either along geographic barriers (e.g. mountain ranges; Mercure et al. 1993; Shafer et al. 2010) or habitat transitions (e.g. Reding et al. 2012). Gray foxes have a large geographic range that encompasses the United States, Mexico, and northern South America across multiple geographic barriers and habitat transitions. This large geographic range includes 16 recognized subspecies (Fritzell and Haroldson 1982), 7 of which occur in the United States where the eastern and western lineages were identified by Goddard et al. (2015). One such subspecies boundary (*U. c. floridanus* and *U. c. scotti*) lies within the southern Great Plains where eastern forests transition into western scrub habitats. Western gray foxes are known to be associated with scrub habitats whereas eastern gray foxes are typically found in forested habitats (Fritzell and Haroldson 1982), which suggests the Great Plains may be the location of a gray fox contact zone (Reding et al. 2021).

If gray fox lineages meet in a contact zone within the Great Plains as suggested by mtDNA, confirming this with nuclear data would help clarify the biogeographic mechanisms underlying their divergence and would enable exploration of the magnitude of post-contact admixture. Therefore, our first objective was to determine the location of the contact zone as reflected in nuclear DNA between eastern and western gray fox lineages. Multiple forest-dwelling vertebrates in North America meet in forested or scrub habitats encroaching into the Great Plains (Rising 1983), where separation between them is reinforced by habitat-based (e.g. scrub vs. forest) differences (Swenson 2006). Following nuclear positioning of the gray fox contact zone, we sought to estimate genomic divergence between lineages and quantify admixture between them.

Materials and methods

Sampling and DNA extraction

We obtained 376 gray fox samples during 2013 to 2019 from 3 sources across North America: 211 DNA extracts from Texas (USDA Wildlife Services) in and around the southern Great Plains, 26 DNA extracts from California (Goddard et al. 2015), and 139 tissue samples from fur-trapped foxes in several eastern, central, and western states. We extracted the tissue samples using Qiagen DNEasy Blood and Tissue kits (Qiagen Inc., Valencia, CA). All extracts were normalized to 10 ng/µL prior to mitochondrial sequencing and library preparation or culled due to insufficient DNA concentration. After this quality control step, we narrowed our sample set to a geographically representative subset of gray foxes from the eastern (n = 100) and western (n = 181) regions (Supplementary Table S1).

Mitochondrial DNA methods

To investigate concordance between mtDNA and nuclear patterns, we amplified a 441 bp region of the mitochondrial cytochrome b region (RF14724, RF15149; Perrine et al. 2007), which was sufficient to unambiguously distinguish between eastern and western gray foxes (Goddard et al. 2015). For example, a 100-bp region in the center of the amplicon contained 10 diagnostic sites. Polymerase chain reactions (PCR) were conducted in 11 µL volumes that contained 2 µL of template DNA, 1.1 µL of 10× Buffer, 1.1 µL of 25 mM $MgCl_2$, 1.1 µL of 2 mM dNTPs, 0.11 µL of 10 µg/mL of bovine serum albumin, 0.28 µL of each 20 µM forward and reverse primer, and 0.2 µL of 5 U/µL of AmpliTaq DNA polymerase (ThermoFisher Scientific), and conditions followed Aubry et al. (2009). We cleaned all PCR and sequencing reaction products with a modified low sodium precipitation method (Latch and Rhodes 2005) and Sanger sequenced using a BigDye terminator v3.1 cycle sequencing kit (Applied Biosystems) on an ABI 3730 platform at the Veterinary Genetics Laboratory at the University of California, Davis.

Library preparation and sequencing

We used a genotype by sequencing (GBS) method modified from Elshire et al. (2011) to construct reduced representation libraries for gray fox. To minimize introducing batch effects to downstream analyses, we divided samples from the eastern and western United States evenly across 3 separate libraries. We obtained and annealed top ACACTCTTTCCCTACAC GACGCTCTTCCGATCT{4 to 7 bp barcode}TGCA and bottom {barcode compliment}AGATCGGAAGAGCGTCGT GTAGGGAAAGAGTGT sequences of the first 96 barcoded adapters and 1 common adapter (top: AGATCGGAAGAG CGGTTCAGCAGGAATGCCGAG, bottom: CTCGGCAT TCCTGCTGAACCGCTCTTCCGATCTTGCA) designed for *Pst*I, *Sbf*I, and *Eco*T22I restriction enzymes (Wallace and Mitchell 2017). We then digested 100 ng of DNA at 37 °C for 2 h with a high-fidelity restriction enzyme to maximize consistency of digestions. Each sample was digested in 20-µL reactions containing 2 µL of 10× NEB Cutsmart Buffer, 1 U of *Nsil*-HF restriction enzyme (an equivalent to *Eco*T22I; New England BioLabs Inc.), 0.2 µL of 100× bovine serum albumin, 10 µL of DNA, and 7.75 µL of deionized water. We then ligated both the common and a uniquely barcoded adapter to each DNA sample by adding 5 µL of 10× NEB Ligase buffer, 4.5 µL each of adapter (0.3 ng/µL), 1 µL (400 CEU) of T4 DNA Ligase (400 U/µL), and 15 µL of deionized water to the previously digested DNA samples. We incubated the resulting 50 µL reactions at 22 °C for 60 min and heated to 65 °C for 30 min to deactivate the ligase. To minimize the risk of a few high-quality samples dominating the final pooled libraries, we PCR-amplified small pools of 12 samples each before combining the PCR products in equal concentrations to the final library. Each sub pool was PCR amplified in 50 µL reaction volumes composed of 10 µL of purified library, 25

µL Taq 2x Master Mix (New England Biolabs Inc., Ipswich, MA), 13 µL of deionized water, and 1 µL of both forward and reverse primer (50 µM GBSEcoT22IPrimer-F: aatgatac ggcgaccaccgagatctacactctttccctacacgacgctcttccg atct; 50 µM GBSEcoT22IPrimer-R: caagcagaagacggcatacgagatcggtctcggc attcctgctgaaccgctcttccgatct). The thermal profile was 5 min at 72 °C, 30 s at 98 °C, followed by 18 cycles of 10 s at 98 °C, 30 s at 65 °C, and 30 s at 72 °C, with a final extension at 72 °C for 5 min. We purified the PCR products with QiaQuick PCR purification kits, and quantified library concentrations with a Qubit fluorometer (Qiagen Inc.). All libraries were run on 1% agarose gels and a Bioanalyzer (Agilent Technologies, Santa Clara, CA) trace before being sent to the University of California, Davis Genome Center DNA Technologies core for 100-bp single-end sequencing on an Illumina HiSeq4000.

SNP calling and filtering

We demultiplexed reads and trimmed adapters using GBSX_v1.3.jar (Herten et al. 2015). Although no reference genome was available for the gray fox, previous studies indicated high mapping success of *Urocyon* reads to the dog genome (e.g. Robinson et al. 2016). Therefore, we aligned trimmed reads to the dog genome (canFam3.1) using BWA-MEM (Li et al. 2010; Li 2013) as was used previously (Robinson et al. 2016). Although little reference bias was expected, we nevertheless applied a relaxed mapping quality filter (-q 10) in SAMTOOLS v1.9 (Li et al. 2009) to minimize any bias against lower mapping quality of derived alleles in gray foxes.

We called single nucleotide polymorphisms (SNPs) using the reference genome "ref_map.pl" pipeline in Stacks (v2.53, Catchen et al. 2011). The "gstacks" module was run with default SNP model parameters (--model marukilow, --var-alpha 0.05, and –gt-alpha 0.05) for unpaired reads, which assesses the statistical likelihood of each genotype call, reducing the need for subsequent coverage filters. We used the "populations" module in Stacks to filter out low frequency variants (--min_maf 0.02) and SNPs with excess levels of heterozygosity (--max_obs_het 0.60) that could reflect paralogs. We then used a stepwise filtering approach in PLINK version 1.9.0 (Purcell et al. 2007). We initially removed the lowest quality individuals from the dataset that were missing >90% of SNPs (--mind 0.9). We then filtered out SNPs that were called in <80% of remaining individuals (--geno 0.2). Based on this SNP set, we identified a maximum-missing rate of 32% as one that would retain most individuals while removing those with a significant amount of missing data (--mind 0.32).

Population structure and admixture between lineages

We calculated nucleotide diversity (π) and genetic distance (F_{ST}) for the eastern and western lineages using the "populations" module in Stacks. We used 3 complementary methods to quantify population structure and admixture. First, we examined admixture via the Bayesian program FastStructure (Raj et al. 2014) assuming 1 to 5 genetic clusters (K = 1 to 5) over 5 independent runs. We then chose the most likely K with the "chooseK.py" function. This program calculates q values for each individual, which correspond to the proportions of an individual's genome assigning to each inferred genetic cluster. Second, we performed a principal components analysis (PCA) in Plink (Purcell et al. 2007). Third, we used fineRADStructure (Malinsky et al. 2018), which utilizes local haplotypes (i.e. linked SNPs) and employs individual-based coancestry and a MCMC clustering algorithm (Lawson et al. 2012) to identify the optimal partition of individuals. This method reveals both global (i.e. between lineage) and local population structure in a single analysis, which typically requires multiple analyses to infer (Malinsky et al. 2018).

Phylogenetic relationships and divergence

To assess the phylogenetic relationship of eastern and western gray foxes and estimate nuclear divergence time, we selected a subset of putatively unadmixed samples (based on FastStructure assignments) from California (n = 6), Texas (n = 6), and the East (n = 6), and additionally incorporated GBS data from other members of the family, Canidae, including members of the genera, *Vulpes* and *Canis* (Supplementary Table S2). We aligned and called SNPs for all canid samples together (including gray foxes) using a similar approach to that described above with the following modifications: In the Stacks populations module, we used a minor allele count (--min-mac 2) to remove only singletons while avoiding systematic removal of alleles specific to a smaller subset of taxa. Additionally, we only included sites that were called in all 15 taxa (-p 15). In Plink, we then removed individuals with >70% missing data (--mind 0.7), and filtered out SNPs that were called in <95% of individuals (--geno 0.05).

To construct a maximum likelihood tree, we converted SNP files (.ped/.fam) to fasta files with heterozygotes coded using IUPAC ambiguity codes using a perl script (source code available at https://github.com/squisquater/Cryptic-Gray-Fox-Lineages-Secondary-Contact). We then created a phydat object from the fasta file using the "phangorn" package (version 2.5.3; Schliep 2011) in R (version 4.1.1), and built the tree by computing the pairwise distances between sequences in the "ape" package (version 5.5, Paradis et al. 2019). To assess the best-fitting nucleotide substitution model, we used the function modelTest() in "phangorn." The tree was refitted with the best substitution model (GTR) with 100 bootstrap replicates to assess node support.

We calibrated estimates of divergence within the *Urocyon* clade to those in the *Vulpes* clade because the 2 fox lineages exhibit similar life histories and generation times and because they were similarly divergent from the dog (~9.7 and 8.7 MY, respectively; Lindblad-Toh et al. 2005; Koepfli et al. 2015); the 2 genera would therefore be expected to manifest similar, if any, reference biases (Duchen and Salamin 2021). We applied a relaxed clock approach implemented in Mega (v. 11; Tamura et al. 2021) that utilizes hierarchical differences between sister branch lengths to infer relative rates within and among clades (RelTime, Tamura et al. 2012, 2018). Because the split between *Urocyon* and other canids represents the most basal node among the extant Canidae, there was no true outgroup in our sample (Lindbladh-Toh et al. 2005; Nyakatura et al. 2012). Therefore, we rooted the tree to *Canis*, which resulted in a distorted tree, for which the bias in evolutionary rates between *Vulpes* and *Urocyon* lineages was quantifiable (see below). In both cases, we applied maximum likelihood estimation (RelTime-ML), along with 2 normally distributed calibrations for *Vulpes* nodes: *V. vulpes* vs. *V. macrotis*/*V. lagopus* (2.91 MY, 95% CI: 2.09 to 3.84) and *V. macrotis* vs. *V. lagopus* (0.97 MY, 95% CI: 0.55 to 1.45) (Perini et al. 2010). Confidence intervals around each divergence estimate were calculated according to Tao et al. (2020).

Since improper rooting to *Canis* may negatively impact our estimates of divergence, we multiplied a correction factor of 1.23 to adjust our calculated divergence time and confidence intervals between eastern and western *Urocyon* clades. This adjustment factor was calculated as the ratio of independently estimated divergence between *Urocyon* and the node ancestral to *Canis* and *Vulpes* (9.7 + [9.7 – 8.7] = 10.7 MY) divided by the divergence between *Canis* and *Vulpes* (8.7 MY; Koepfli et al. 2015). This ratio agreed (to 2 decimal places, i.e. 1.23) with that calculated directly from the empirically estimated branch lengths in our original unrooted maximum likelihood tree (0.2521 for *Urocyon*; 0.2038 and 0.2066 for *Vulpes* and *Canis*, respectively, averaging 0.2052; 0.2521/0.2052 = 1.2289).

Results

After filtering, we retained 259 gray foxes for the intraspecific SNP set, and used 18 of these, along with 18 *Vulpes* and 41 *Canis* samples for the interspecific SNP set. Average (±SE) sequencing depth was 34× (±1.3×) across individuals. We observed no qualitative differences in sequencing depth associated with sequencing lane or source population (Supplementary Fig. S1). On average, 88.9% (±0.2%) of gray fox reads aligned to the dog reference genome, compared to 92.7% (±0.5%) of *Vulpes* reads and 93.6% (±0.3%) of *Canis* reads. Our primary (intraspecific) genomic dataset contained 30,431 variable sites (SNPs) across 30,410 GBS loci (i.e. read stacks).

Population structure and admixture

Nucleotide diversity was nearly twice as high in the western lineage (π = 0.0033) than the eastern lineage (π = 0.0018), and F_{ST} between them was 0.12. We also obtained cytochrome b haplotypes for 231 of the 259 gray foxes, resulting in 86 eastern and 145 western haplotypes (Fig. 1A). For SNP genotypes, FastStructure found the highest support for K = 3 genetic clusters. Specifically, K = 3 divided samples into eastern and western clusters and further divided the western lineage into 2 clusters (Fig. 1B and C). Based on these nuclear assignments, 165 gray fox genotypes were assigned to the 2 western clusters and 94 gray fox genotypes were assigned to the eastern cluster (Fig. 1B and C). The mitochondrial haplotypes and nuclear genetic cluster assignments were geographically concordant, revealing an abrupt contact zone between eastern and western lineages in the southern Great Plains. We recorded only 8 cases where the mitochondrial type disagreed with the predominant nuclear ancestry, including 7 individuals that assigned primarily to the western lineage in Texas, but had eastern haplotypes and 1 individual with a western haplotype in Oklahoma that was assigned primarily to the eastern lineage. The PCA indicated substantial divergence between the eastern and western lineages (Fig. 2). The first PC axis split the 2 lineages and explained 20.1% of the variance. The second PC axis explained only 2.7% of the variance, separating individuals geographically along a cline from California to Texas, similarly to the transition observed in the fastStructure analysis.

FineRADstructure revealed patterns consistent with introgression between the eastern and western lineages (Fig. 3). The clustered coancestry matrix and the dendrogram mirrored the PCA results with the strongest differentiation occurring between the eastern and western individuals. The dendrogram also captured the substructure within the western lineage, with Colorado/New Mexico, Nevada, and California all forming separate clusters with 100% bootstrap support.

Phylogenetic relationships and divergence

The family-wide (i.e. Canidae) dataset included 125,044 variable sites across 77 individuals. The maximum likelihood tree indicated 100% bootstrap support for differentiation among eastern and western gray foxes and recovered all previously identified phylogenetic groups (e.g. genera and species; Fig. 4A). The RelTime tree improperly rooted to *Canis* and calibrated to 2 nodes within the *Vulpes* clade resulted in a biased estimate of 0.71 MY (95% CI: 0.33 to 1.50) MY, which when corrected for misrooting, resulted in the final estimate of 0.87 (0.41 to 1.86) MY separating eastern and western gray fox clades (Fig. 4B). Divergence between gray fox lineages was intermediate between that of Eurasian/North American red fox and kit/arctic foxes, which supports a mid-Pleistocene divergence in line with the mitochondrial estimate (Goddard et al. 2015).

Discussion

We sought to characterize the magnitude and geographic pattern of genomic divergence between eastern and western gray fox lineages. Consistent with mitochondrial divergence (Goddard et al. 2015; Hofman et al. 2015; Reding et al. 2021), our findings demonstrate that these lineages diverged prior to the penultimate (Illinoian) Pleistocene glaciation. We also identified a well-defined contact zone and low overall levels of admixture, consistent with long-term isolation followed by recent contact or longstanding forms of reproductive isolation. Finally, the location of the contact zone was the same for both mitochondrial and nuclear markers, both of which occurred in the southern Great Plains. These findings suggest that Pleistocene habitat rather than physical barriers such as the Mississippi River Delta (e.g. Barton et al. 2012) separated these lineages. If so, these habitat affinities could have allowed intermittent breeding opportunities without high levels of intermixing.

Eastern and western gray foxes diverged in the mid-Pleistocene

Our SNP dataset indicated that eastern and western gray foxes were differentiated where they came into contact, concordant with mitochondrial findings. The PCA and fineRADStructure results, in particular, showed that eastern and western individuals were assignable to their geographic population 100% of the time. The first principal component in our PCA showed strong differentiation (20.1% of variance explained) between these lineages, whereas the second principal component that separated individuals in the western lineage explained much less variance (2.7%) within the western lineage.

Further, in agreement with previous mitochondrial patterns (Goddard et al. 2015; Hofman et al. 2015; Reding et al. 2021), nuclear phylogenetic analyses indicated that this differentiation reflected deep divergence between lineages. These divergence time estimates corresponded to the Early-Middle Pleistocene Transition (1.4 to 0.4 Mya; Head et al. 2015) in the Irvingtonian land mammal age, which is substantially

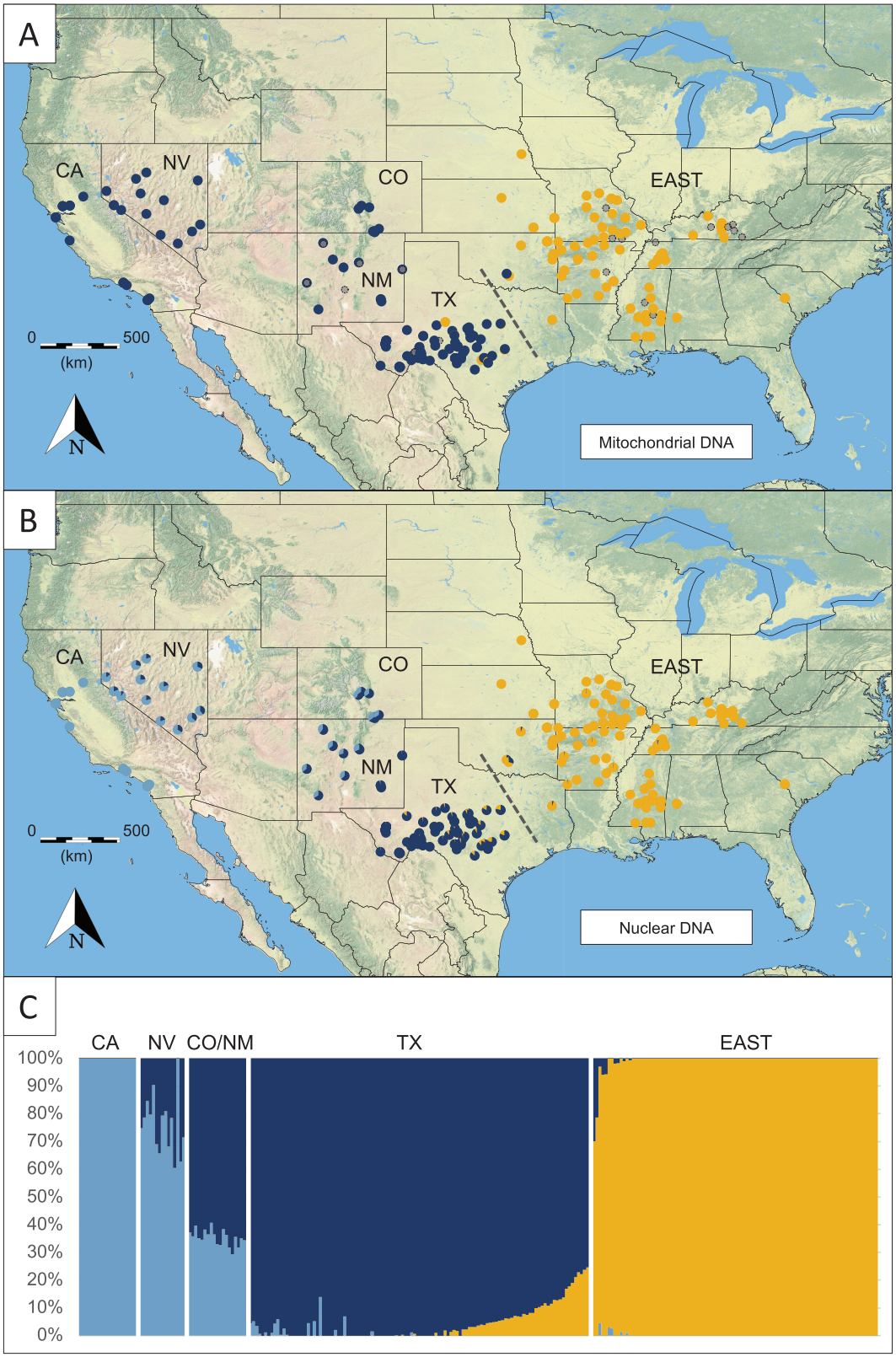

Fig. 1. Population assignments of North American gray fox samples based on (A) mtDNA cytochrome b (n = 231) and (B, C) 44,931 nuclear SNPs (n = 259 foxes) from FastStructure K = 3. Both datasets suggest gray fox undergo secondary contact in the southern Great Plains, and experience low levels of admixture as revealed by calculated ancestries in FastStructure. Samples for which we failed to recover cytochrome b haplotypes are noted in gray. Nuclear analyses identified the primary split between the eastern (yellow) and western (all blues) lineages with additional substructure exhibited within the western lineage from California (CA; light blue) to Texas (TX; dark blue). Samples collected from Nevada (NV), Colorado (CO), and New Mexico (NM) showed varying levels of these 2 western ancestry types. Pie charts indicate a single individual and their inferred ancestries of the 3 genetic clusters (yellow, dark blue, and light blue). We detected 8 instances of mito-nuclear discordance where 7 individuals with primarily western nuclear ancestry had eastern cytochrome b haplotypes, and a single individual with primarily eastern nuclear ancestry had a western cytochrome b haplotype.

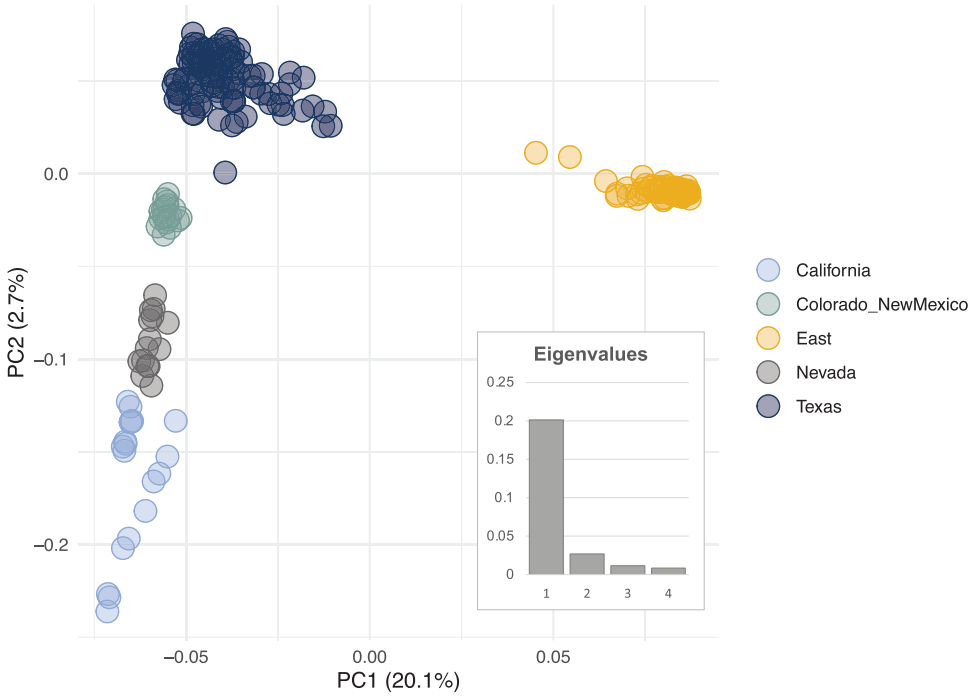

Fig. 2. PCA of gray fox SNPs based on nuclear genotyping by sequencing. The first PC axis split the eastern (yellow) and western (all blues) lineages and explained 20.1% of the variance. The second PC axis explained 2.7% of the variance and revealed population structure within the western cluster (shades of blue). Each circle represents an individual fox colored by its geographic location (California = light blue, Nevada = blue-gray, Colorado/New Mexico = blue-green, Texas = dark blue, and East = yellow).

earlier than those ascribed to intraspecific divisions for most North American carnivores (e.g. Aubry et al. 2009; Reding et al. 2012; Puckett et al. 2015). Fossils of *Urocyon* species dating to the mid-Hemphillian land mammal age (~6 to 8 Mya) are known from Florida and by the Blancan age (>1.8 Mya) occurred in Florida, Texas, New Mexico, Arizona, and California (Tedford et al. 2009). Although fragmentary, the fossil evidence is consistent with continuous presence of some form of *Urocyon* spanning the southeast to the southwest regions of the continent throughout the Pleistocene (Kurtén and Anderson 1980). Current subspecies boundaries within Texas correspond to the eastern (*U. c. floridanus*) and western (*U. c. scotti*) lineages, but the level of divergence recorded here is similar to that between other fox sister species, such as swift and kit foxes (Mercure et al. 1993).

Secondary contact in the southern Great Plains

We found that eastern and western gray foxes form an area of secondary contact within the southern Great Plains. All individuals with appreciable east–west admixture (e.g. >10%) were from Oklahoma and Texas. Although our sampling gap was wider further north, the complete absence of admixture in samples from Colorado, Kansas, and Nebraska suggest even less admixture to the north. The relatively narrow span we documented between predominantly western vs. predominantly eastern ancestry suggests either that secondary contact was recent (e.g. mid- to late Holocene) or that partial reproductive barriers operate to minimize the rate of gene flow. We acknowledge it is difficult to precisely measure the span of this contact zone due to low sampling density within eastern Texas and Oklahoma. However, despite a larger gap in Arizona and Utah, the zone of admixture observed between California and Texas-centered clusters spanned a much broader geographic distance (e.g. Fig. 1B). Thus, if the area was wider like the western lineage contact zone, we expect to have detected it given our intense sampling in the east central United States. Instead, the admixture within the western lineage is consistent with much more ancient isolation-by-distance (or ancient secondary contact) or potentially recent secondary contact, but followed by high gene flow.

Given that interglacial periods have occurred multiple times since the divergence of the eastern and western lineages, it is unclear what would have changed to only recently allow their secondary contact. One possibility is anthropogenic changes in the landscape, such as conversion of grasslands and woody plant invasions. Like many forest-associated species, records indicate the gray fox were largely absent in the plains-dominated areas of western Kansas, Oklahoma, and Texas prior to the 1900s (Choate and Krause 1974; Fritzell and Haroldson 1982; Fritzell 1987), but are now found throughout the southern Great Plains (Davis and Schmidly 1994). Thus, very recent range expansions could have initiated interbreeding between previously isolated eastern and western forms.

An alternative scenario is that long-term isolation created pre- or post-zygotic reproductive barriers that limit successful hybridization. Strong habitat preferences, for example, could prevent foxes from interbreeding as observed in multiple North American carnivores (i.e. habitat-biased dispersal; Sacks et al. 2004; Reding et al. 2012). When strong habitat-biased dispersal occurs, populations are associated with separate habitats and differentiation corresponds to breaks in habitat (e.g. Sacks et al. 2004, 2008). Whereas western gray foxes often occur in scrub habitats, which are common in western and central Texas, eastern gray foxes tend to be associated with woodlands (Fritzell and Haroldson 1982). Additionally,

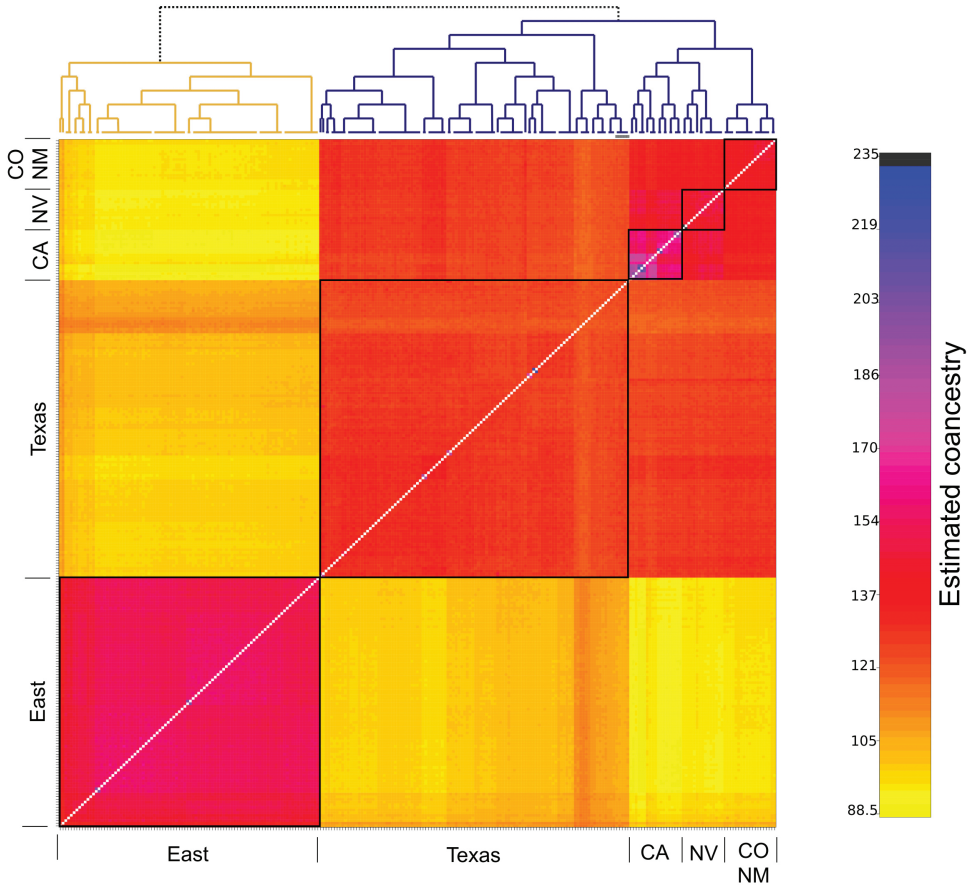

Fig. 3. Symmetrical coancestry matrix of 259 individual gray foxes and dendrogram of inferred groups from fineRADStructure. The heat map depicts variation in pairwise coancestry among individuals according to the scale shown on the right. The primary split between groups corresponds to the eastern (yellow) and western (blue) lineages. In addition to the primary split, all individuals sampled from the highlighted geographic regions clustered together with 100% bootstrap support. Sample locations are listed as follows: California = CA, Nevada = NV, Colorado = CO, New Mexico = NM, Texas = TX, and East = All Eastern Samples.

the long-term isolation and accumulation of genomic differentiation via natural selection and/or drift in separate habitats could have created genomic incompatibilities (Delph et al. 2016; Kuo et al. 2019). Thus, admixed individuals could be less fit than parental types, which could reduce introgression with or without habitat preferences. Further work is needed to accurately time the age of the contact zone and determine whether observed levels of gene flow are indicative of neutral forces, or whether reproductive barriers may in fact be acting to maintain boundaries between lineages.

Substructure within the western lineage

In addition to the main eastern and western split, we documented additional substructure within the western lineage in FastStructure, PCA, and fineRADStructure. Several well-recognized glacial refugia occurred in western North America, which could explain the divergence observed in the western lineage. For example, California served as a glacial refugium for multiple taxa and often contains distinct genetic populations owing to the barrier effect of the Sierra Nevada Mountains (e.g. Hull et al. 2010; Barrows et al. 2014; Puckett et al. 2015). Furthermore, Northern Mexico and the desert southwest also served as a glacial refugium for mammals like western spotted skunks (*Spilogale gracilis*; Ferguson et al. 2017), another widespread mesocarnivore in western North America. The western lineage may also simply reflect an isolation-by-distance relationship. Overall, substructure in the western lineage is much less pronounced than the main split between the eastern and western lineages, and requires further sampling, particularly to the south in Mexico and Central America, to examine if the western lineage underwent further divergence during the Pleistocene or simply exhibits isolation-by-distance.

Conclusions

Our genomic assessment of the gray fox indicates that divergence between eastern and western lineages represents ancient isolation during the mid-Pleistocene. The relatively narrow zone of hybridization in the Great Plains suggests either that secondary contact was very recent or that reproductive barriers maintain distinctiveness between lineages. Gray fox join a growing number of wide-ranging species that exhibit strong divergence in North America (e.g. Dawson and Cook 2012; Ferguson et al. 2017), but the Great Plains represents a relatively rare area for secondary contact in carnivores. In the future, more extensive geographic sampling, particularly in Central and South America, and sequencing of whole genomes can help to further clarify the evolutionary history and systematics of *Urocyon*. Such high-resolution genomes are particularly important for examining isolating mechanisms that

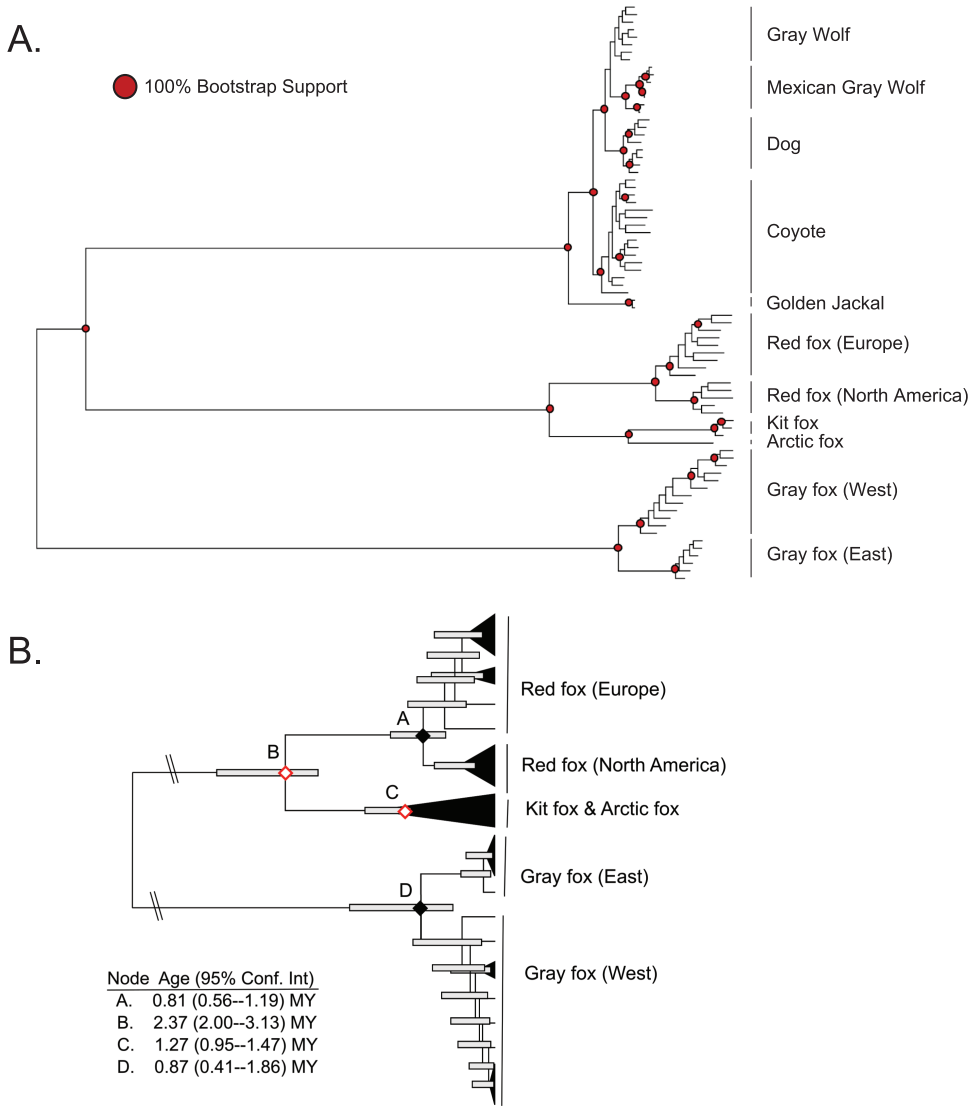

Fig. 4. Phylogenetic tree built using (A) maximum likelihood with bootstrap support (red nodes = bootstrap support of 100%), and (B) timetree analysis using the RelTree method (Tamura et al. 2012; Tao et al. 2020) based on 2 calibration points within the *Vulpes* clade (red nodes), indicating a bias-corrected divergence estimate between eastern and western gray fox clades at 870,000 (95% CI: 410,000 to 1,860,000) yr ago, during the mid-Pleistocene. A bias correction factor of 1.23× was necessary to correct raw (biased) estimates within the *Urocyon* clade (e.g. 706,000, 95% CI: 333,000 to 1,500,000 yrs splitting eastern and gray foxes) to accommodate rooting to an ingroup, *Canis*, which was the sister taxon to *Vulpes* (see Methods).

underlie continued divergence in gray fox and potentially other species that undergo secondary contact in the Great Plains.

Supplementary material

Supplementary material is available at *Journal of Heredity* online.

Funding

Funding was provided by the University of California, Davis, Mammalian Ecology and Conservation Unit.

Acknowledgments

We extend gratitude to the landowners, fur-trappers, USDA Wildlife Services personnel, and J. Beasley (Savannah River Ecology Laboratory, University of Georgia) and E. Boydston (USGS) for providing tissue and DNA extract samples for this study. Critical laboratory support was provided by S. Vanderzwan and the University of California, Davis Veterinary Genetics Laboratory staff. We thank the associate editor and 2 reviewers for helpful comments on an earlier draft of this manuscript.

Conflict of interest

The authors declare no conflict of interest.

Data availability

The raw, demultiplexed fastq files used in these analyses are publicly available in the NCBI SRA under BioProject PRJNA938958. All scripts and R code can be accessed at https://github.com/squisquater/Cryptic-Gray-Fox-Lineages-Secondary-Contact.

References

Aubry KB, Statham MJ, Sacks BN, Perrine JD, Wisely SM. Phylogeography of the North American red fox: vicariance in Pleistocene forest refugia. *Mol Ecol*. 2009;18(12):2668–2686.

Barrows CW, Fisher M, Fisher M. Past, present, and future distributions of a local assemblage of congeneric lizards in southern California. *Biol Conserv*. 2014;180:97–107.

Barton HD, Wisely SM, Wisely SM. Phylogeography of striped skunks (*Mephitis mephitis*) in North America: Pleistocene dispersal and contemporary population structure. *J Mammal*. 2012;93(1):38–51.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH, De Koning DJ. Stacks: building and genotyping loci de novo from short-read sequences. *G3*. 2011;1(3):171–182.

Choate JR, Krause JE. Historical biogeography of the gray fox (*Urocyon cinereoargenteus*) in Kansas. *Trans Kans Acad Sci*. 1974;77(4):231–236.

Cullingham CI, Kyle CJ, Pond BA, Rees EE, White BN. Differential permeability of rivers to raccoon gene flow corresponds to rabies incidence in Ontario, Canada. *Mol Ecol*. 2008;18(1):43–53.

Davis WB, Schmidly DJ. *The mammals of Texas*. Austin (TX): Texas Parks and Wildlife Press; 1994.

Dawson NG, Cook JA. Behind the genes: diversification of North American martens (*Martes americana* and *M. caurina*). In: Aubry KB, Zielinksi WJ, Raphael MG, Proulx G, Buskirk SW, editors. *Biology and conservation of martens, sables, and fishers: a new synthesis*. New York: Cornell University Press; 2012. p. 23–38.

Delph LF, Demuth JP, Demuth JP. Haldane's Rule: genetic bases and their empirical support. *J Hered*. 2016;107(5):383–391.

Duchen P, Salamin N. A cautionary note on the use of genotype callers in phylogenomics. *Syst Biol*. 2021;70(4):844–854.

Elshire RJ, Glaubitz JC, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. A robust, simple genotype-by-sequencing (GBS) approach for high diversity species. *PLoS One*. 2011;6(5):e19379.

Ferguson AW, McDonough MM, Guerra GI, Rheude M, Dragoo JW, Ammerman LK, Dowler RC. Phylogeography of a widespread small carnivore, the western spotted skunk (*Spilogale gracilis*) reveals temporally variable signatures of isolation across western North America. *Ecol Evol*. 2017;7(12):4229–4240.

Fritzell EK. Gray fox and island gray fox. In: Novak M, Baker JA, Obbard ME, Malloch B, editors. *Wild furbearer management and conservation in North America*. Ontario (Canada): Ontario Ministry of Natural Resources; 1987. p. 409–420.

Fritzell EK, Haroldson KJ. *Urocyon cinereoargenteus*. *Mammal Species*. 1982;189:1–8. doi:10.2307/3503957

Goddard NS, Statham MJ, Sacks BN. Mitochondrial analysis of the most basal canid reveals deep divergence between eastern and western North American gray foxes (*Urocyon* spp.) and ancient roots in Pleistocene California. *PLoS One*. 2015;10(8):e0136329.

Head MJ, Gibbard PL, Gibbard PJ. Early-middle Pleistocene transitions: linking terrestrial and marine realms. *Quat Int*. 2015;389:7–46. doi:10.1016/j.quaint.2015.09.042

Herten K, Hestrand MS, Vermeesch, JR, Van Houdt JKJ. GBSX: a toolkit for experimental design and demultiplexing genotyping by sequencing experiments. *BMC Bioinform*. 2015;16:73.

Hewitt GM. Genetic consequences of climatic oscillations in the Quaternary. *Philos Trans R Soc Lond B Biol Sci*. 2004;359(1442):183–195.

Hofman CA, Rick TC, Hawkins MTR, Funk WC, Ralls K, Boser CL, Collins PW, Coonan T, King JL, Morrison SA, et al. Mitochondrial genomes suggest rapid evolution of dwarf California Channel Island foxes (*Urocyon littoralis*). *PLoS One*. 2015;10(2):e0118240.

Hope AG, Malaney JL, Bell KC, Salazar-Miralles F, Chavez A, Cook JA. Widespread red squirrels (genus: *Tamiasciurus*) highlights the complexity of speciation within North American forests. *Mol Phylogenet Evol*. 2016;100:170–182. doi:10.1016/j.ympev.2016.04.014

Hull JM, Keane JJ, Savage WK, Godwin SA, Shafer J, Jepsen EP, Gerhardt R. Range-wide genetic differentiation among North American great gray owls (*Strix nebulosa*) reveals a distinct lineage restricted to the Sierra Nevada, California. *Mol Phylogenet Evol*. 2010;56(1)212–221.

Koepfli KP, Pollinger J, Godinho R, Robinson J, Lea A, Hendricks S, Schweizer RM, Thalmann O, Silva P, Fan Z, et al. Genome-wide evidence reveals that African and Eurasian golden jackals are distinct species. *Curr Biol*. 2015;25(16):2158–2165.

Kuo TH, Vanderzwan SL, Kasprowicz AE, Sacks BN. Using ancestry-informative SNPs to quantify introgression of European alleles into North American red foxes. *Heredity*. 2019;110:782–792.

Kurtén B, Anderson E. *Pleistocene mammals of North America*. New York: Columbia University Press; 1980.

Latch EK, Kierepka EM, Heffelfinger JR, Rhodes OE Jr. Hybrid swarm between divergent lineages of mule deer (*Odocoileus hemionus*). *Mol Ecol*. 2011;20(24):5265–5279.

Latch EK, Rhodes OE Jr. The effects of gene flow and population isolation on the genetic structure of reintroduced wild turkey populations: are genetic signatures of source populations retained? *Conserv Genet*. 2005;6:981–997. doi:10.1007/s10592-005-9089-2

Lawson DJ, Hellenthal G, Myers S, Falush D. Inference of population structure using dense haplotype data. *PLoS Genet*. 2012;81(1):e1002453.

Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM [Algorithm], arXiv, 2013. doi:10.48550/arXiv.1303.3997

Li H, Durbin R, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26(5):589–595.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup1000 Genome Project Data Processing Subgroup. The sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–2079.

Lindbladh-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, Clamp M, Chang JL, Kulbokas III EJ, Zody MC, et al. Genome sequence, comparative analysis, and haplotype structure of the domestic dog. *Nature*. 2005;438:803–819. doi:10.1038/nature04338

Malinsky M, Trucchi E, Lawson DJ, Falush D. RADpainter and fineRADstructure: population inference from RADseq data. *Biol Evol*. 2018;35(5):1284–1290.

Mercure A, Ralls K, Koepfli KP, Wayne RK. Genetic subdivisions among small canids: mitochondrial DNA differentiation of swift, kit, and arctic foxes. *Evolution*. 1993;47(5):1313–1328.

Nyakatura K, Bininda-Emonds ORP, Bininda-Emonds OR. Updating the evolutionary history of Carnivore (Mammalia): a new species-level supertree complete with divergence time estimates. *BMC Biol*. 2012;10:12. doi:10.1186/1741-7007-10-12

Paradis E, Schliep K, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*. 2019;35(3):526–528.

Perini FA, Russo CAM, Schrago CG. The evolution of South American endemic canids: a history of rapid diversification and morphological parallelism. *J Evol Biol*. 2010;23(2):311–322.

Perrine JD, Pollinger JP, Sacks BN, Barrett RH, Wayne RK. Genetic evidence for the persistence of the critically endangered Sierra Nevada red fox in California. *Conserv Genet*. 2007;8:1083–1095. doi:10.1007/s10592-006-9265-z

Puckett EE, Etter PD, Johnson EA, Eggert LS. Phylogeographic analyses of American black bear (*Ursus americanus*) suggest four glacial refugia and complex patterns of postglacial admixture. *Mol Biol Evol*. 2015;32(9):2338–2350.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *AJHG*. 2007;81(3):559–575.

Raj A, Stephens M, Pritchard JK. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*. 2014;197(2):573–589.

Reding DM, Bronikowski AM, Johnson WE, Clark WR. Pleistocene and ecological effects on continental-scale genetic differentiation in the bobcat (*Lynx rufus*). *Mol Ecol*. 2012;21(2):3078–3093.

Reding DM, Castaneda-Rico S, Shirazi S, Hofman CA, Cancellare IA, Lance SL, Beringer J, Clark WR, Maldonado JE. Mitochondrial genomes of the United States distribution of gray fox (Urocyon cinereoargenteus) reveal a major phylogeographic break at the Great Plains Suture Zone. *Front Ecol Evol*. 2021;9. doi:10.3389/fevo.2021.666800

Rising JD. The Great Plains hybrid zones. In: Johnston RF, editor. *Current ornithology*, Vol. 1. New York: Springer Publishing; 1983. p. 131–157.

Robinson JA, Ortega-Del Vecchyo D, Fan Z, Kim BY, Marsden CD, Lohmueller KE, Wayne RK. Genomic flatlining in the endangered island fox. *Curr Biol*. 2016;26(9):1183–1189.

Sacks BN, Bannasch DL, Chomel BB, Ernest HB. Coyotes demonstrate how habitat specialization by individuals of a generalist species can diversify populations in a heterogeneous ecoregion. *Mol Biol Evol*. 2008;25(7):1384–1394.

Sacks BN, Brown SK, Ernest HB. Population structure of California coyotes corresponds to habitat-specific breaks and illuminates species history. *Mol Ecol*. 2004;13(5):1265–1275.

Schliep KP. phangorn: phylogenetic analysis in R. *Bioinformatics*. 2011;27(4):592–593.

Shafer ABA, Cullingham CI, Côté SD, Coltman DW. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Mol Ecol*. 2010;19(21):4589–4621.

Slager DL, Epperly KL, Ha RR, Rohwer S, Wood C, Van Hemert C, Klicka J. Cryptic and extensive hybridization between ancient lineages of American crows. *Mol Ecol*. 2020;29(5):956–969.

Stone KD, Flynn RW, Cook JA. Post-glacial colonization of northwestern North America by the forest-associated American marten (*Martes americana*, Mammalia: Carnivore: Mustelidae). *Mol Ecol*. 2002;11(10):2049–2063.

Swenson NG. GIS-based niche models reveal unifying climatic mechanisms that maintain the location of avian hybrid zones in a North American suture zone. *J Evol Biol*. 2006;19(3):717–725.

Tamura K, Battistuzzi FU, Billing-Ross P, Murillo O, Filipski A, Kumar S. Estimating divergence times in large molecular phylogenies. *Proc Natl Acad Sci USA*. 2012;109(47):19333–19338.

Tamura K, Qiqing T, Kumar S. Theoretical foundation of the RelTime method for estimating divergence times from variable evolutionary rates. *Mol Biol Evol*. 2018;35(7):1770–1782.

Tamura K, Stecher G, Kumar S. MEGA 11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol*. 2021;38(7):3022–3027.

Tao Q, Tamura K, Mello B, Kumar S. Reliable confidence intervals for RelTime estimates of evolutionary divergence times. *Mol Biol Evol*. 2020;37(1):280–290.

Tedford RH, Wang X, Taylor BE. Phylogenetic systematics of the North American fossil Caninae (Carnivore: Canidae). *Bull Am Mus Nat Hist*. 2009;325:1–218. doi:10.1206/574.1

Wallace JG, Mitchell SE. Genotype-by-sequencing. *Curr Protoc Plant Biol*. 2017. doi:10.1002/cppb.20042

Weir JT, Schluter D. Ice sheets promote speciation in boreal birds. *Proc R Soc Lond B Biol Sci*. 2004;271(1551):1881–1887.