Whole-genome sequence analysis shows that two endemic species of North American wolf are admixtures of the coyote and gray wolf

Bridgett M. vonHoldt,1 James A. Cahill,2 Zhenxin Fan,3 Ilan Gronau,4 Jacqueline Robinson,5 John P. Pollinger,5 Beth Shapiro,2 Jeff Wall,6 Robert K. Wayne5*

Protection of populations comprising admixed genomes is a challenge under the Endangered Species Act (ESA), which is regarded as the most powerful species protection legislation ever passed in the United States but lacks specific provisions for hybrids. The eastern wolf is a newly recognized wolf-like species that is highly admixed and inhabits the Great Lakes and eastern United States, a region previously thought to be included in the geographic range of only the gray wolf. The U.S. Fish and Wildlife Service has argued that the presence of the eastern wolf, rather than the gray wolf, in this area is grounds for removing ESA protection (delisting) from the gray wolf across its geographic range. In contrast, the red wolf from the southeastern United States was one of the first species protected under the ESA and was protected despite admixture with coyotes. We use whole-genome sequence data to demonstrate a lack of unique ancestry in eastern and red wolves that would not be expected if they represented long divergent North American lineages. These results suggest that arguments for delisting the gray wolf are not valid. Our findings demonstrate how a strict designation of a species under the ESA that does not consider admixture can threaten the protection of endangered entities. We argue for a more balanced approach that focuses on the ecological context of admixture and allows for evolutionary processes to potentially restore historical patterns of genetic variation.

INTRODUCTION

Two well-accepted species of wolf-like canids inhabit North America: the Holarctic gray wolf (Canis lupus) and the endemic coyote (Canis latrans). However, two other entities have been advanced as evolutionarily distinct species of North American origin: the red wolf (Canis rufus) of the southeastern United States and the eastern wolf (Canis lycaon), now found in the eastern Great Lakes (Algonquin Provincial Park and adjacent areas in Ontario) but were historically thought to inhabit a wider area, including the eastern United States (1, 2) (Fig. 1). However, an alternative hypothesis suggests that the red wolf is a hybrid between coyotes and gray wolves that historically inhabited the southeastern United States before gray wolves were eliminated through private and public bounty (3–5) (Fig. 1). Similarly, the eastern wolf may have been generated through admixture between gray wolves and coyotes as they expanded eastward into the Great Lakes region at the end of the last century, concurrent with the near extirpation of wolves in the conterminous United States (6–8). Both red and eastern wolves are intermediate in body size between coyotes and gray wolves, which is consistent with an admixture scenario, and recent evidence has shown that gray wolves and coyotes can produce viable offspring in captivity (9).

The U.S. Fish and Wildlife Service (USFWS) accepts the species status of both red and eastern wolves, with markedly divergent conservation implications. The red wolf is protected by the U.S. Endangered Species Act (ESA). However, the endangered eastern wolf, which was only recently recognized as a distinct species (8–14) and is currently restricted to a small portion of its historic range, would not be listed under the current political landscape. Instead, the acceptance of the eastern wolf species has led the USFWS to propose the delisting of the gray wolf. The reasoning for this action is that the historical range of the eastern wolf is hypothesized to include the Great Lakes region and 29 eastern states to the exclusion of the gray wolf (11, 15, 16). Because the geographic range of the gray wolf as originally listed in the 1975 ESA petition included these areas, the USFWS subsequently proposed that the entire original listing was invalid. Essentially, the presence of the eastern wolf, rather than the gray wolf, in the eastern United States would cause the original listing to be annulled. With the exception of the Mexican wolf, the gray wolf would be delisted (lose protection) from its entire North American range under the proposed USFWS rule change (17). These differing consequences of species listing, despite the possibility of similar admixed origin, provide a marked example of how taxonomy can both protect and threaten endangered species under the ESA.

Although there is extensive literature on the red wolf and eastern wolf [for example, (2, 11, 18, 19)], only recently have genome-wide data been analyzed to support an admixed or ancient origin hypothesis. A previous study genotyped and analyzed more than 42,000 single-nucleotide polymorphisms (SNPs) in a large panel of North American and Eurasian wolf-like canids that supported an admixed origin for both red and eastern wolves (5). However, a reanalysis of these data found evidence for a genetic cluster in central Ontario representing the eastern wolf and concluded that the SNP array data may suffer from ascertainment bias (20). A more recent study presented new, ascertainment-free genome-wide
Canis lupus baileyi (Mexican wolf)
Canis lupus (Gray wolf)
Canis lycaon (Great Lakes wolf)
Canis rufus (Red wolf)
Canis latrans (Coyote)

Fig. 1. Admixture proportions, hypothesized branching patterns, and the geographic distribution of Canis in North America. Top: Previously proposed phylogenetic relationships among Canis lineages, with gray lines indicating putative admixture events (5). Bottom: Geographic distributions of Canis in North America. Sample locations are indicated by dots and abbreviations are described in Table 1. Ancestry proportions from vonHoldt et al. (5) are indicated (proportion gray wolf/proportion coyote; see also new values in Table 3). IRNP, Isle Royale National Park; Ma, million years ago.
SNP data from the eastern wolf and showed through simulation that admixture alone cannot explain the unique positions of the eastern wolf in a principal components analysis (PCA) (21). Here, we use a genome sequencing approach to directly search for regions of unique ancestry in the genomes of red and eastern wolves that cannot be explained by admixture between coyotes and gray wolves. We present 28 sequenced genomes from a diversity of large canids representing Eurasian and North American wolf populations, including North American regions where wolf/coyote admixture is currently absent and regions with suspected admixture. An exhaustive search of wolf genomes from the Great Lakes region including Algonquin Provincial Park, where pure eastern wolves are thought to exist, and from red wolves from the captive breeding colony reveals little unique ancestry and instead demonstrates a distinct geographic pattern of admixture between gray wolves and coyotes. We argue strongly for a less typologically oriented implementation of the ESA that allows interim protection of hybrids while encouraging the restoration of historic patterns of variation through habitat protection.

RESULTS

Genome sequencing
We sequenced 28 canid genomes to varying coverage ranging from 4 to 29-fold and mapped reads to the domestic dog reference genome (Table 1). After filtering for quality and minimum coverage, we retained 5,424,934 SNPs (referred to as 5.4 million SNPs) genotyped across all sequenced genomes. From these data, we estimated that heterozygosity (π) was highest in the Indian wolf (π = 1.71/kb) and lowest in the endangered Mexican wolf (π = 0.48/kb), which is consistent with previous observations of low diversity in the inbred captive Mexican wolf colony (22). We note that the fraction of missing data is negatively correlated with π, although we could not quantify the extent of this effect given the heterogeneous nature of the samples (Table 1).

We quantified genetic differentiation using $F_{ST}$, confining our coyote representatives to the three individuals (California, Alabama, and Quebec coyotes) most likely to be nonadmixed (see the “Estimating admixture by D and $f$” section). We found that the $F_{ST}$ between wolves of the Great Lakes region (which include putative eastern wolves from Algonquin Provincial Park) and gray wolves or coyotes is nearly half that between red wolves and gray wolves or coyotes (North American gray wolf–Great Lakes wolves, $F_{ST} = 0.057$; coyote–Great Lakes region wolf, $F_{ST} = 0.045$; North American gray wolf–red wolf, $F_{ST} = 0.177$; coyote–red wolf, $F_{ST} = 0.108$) (Table 2). The highest value of divergence was between the red wolf and the Eurasian gray wolf and may reflect, in part, the limited founding size and enhanced drift in the small population of captive red wolves. These estimates of interpopulation genetic differentiation (as measured by $F_{ST}$) are comparable to those found among human populations (23), suggesting that previously hypothesized divergence time estimates of hundreds of thousands of years between wolf-like canid lineages are overestimates and/or that these lineages have experienced a substantial amount of recent admixture. Using a simple isolation model and a summary likelihood approach, we estimated a Eurasian gray wolf–coyote divergence time of $T = 0.38$ N generations (95% confidence interval, 0.376 to 0.386 N), where N is the effective population size. If we assume a generation time of 3 years, and an effective population size of 45,000 (24, 25), then this corresponds to a divergence time of 50.8 to 52.1 thousand years ago (ka), roughly the same as previous estimates of the divergence time of extant gray wolves (26–28). Thus, the amount of genetic differentiation between gray wolves and coyotes is low and not much greater than the amount of differentiation within each species (for example, Eurasian versus North American gray wolf, $F_{ST} = 0.099$; Table 2 and fig. S1). This result contradicts molecular clock calculations based on short mitochondrial control region sequences, which were calibrated using a 1-Ma (million years ago) divergence time between gray wolves and coyotes (10). Despite body size and other phenotypic differences between the two species [for example, (11)] and a long history of coyote- and wolf-like forms in North America (1, 29), the genomic data suggest that modern coyotes and gray wolves are very close relatives with a recent common ancestry.

Cluster and ancestry analysis
We first assessed the general pattern of sequence similarity across the observed 5.4 million SNPs using PCA and found distinct groups that corresponded to gray wolf, coyote, and putatively admixed populations, including the Algonquin wolf and red wolf (Fig. 2 and fig. S2). The overall PC space identified two clusters of wolves that can be explained by continental divergence (Eurasian and North American wolves) and identified the California coyote as the most distinct coyote and the Mexican wolf as a distinct North American wolf. The intermediate position on the first PC of Great Lakes region wolves and red wolves is consistent with a model of admixture between gray wolves and coyotes, although it is also consistent with the hypothesis that red and eastern wolves represent distinct conspecific populations [for example, Wilson et al. (10)]. The PCA also shows that coyotes in populations outside the present admixture zone are genetically distinct (5). These general patterns were found in a PCA of downsampled sequences that represented equivalent sampling of all the genomes (fig. S2).

We found 16,184 fixed differences between three nonadmixed coyotes (California, Alabama, and Quebec) and Eurasian gray wolves and used these to estimate wolf versus coyote ancestry proportions, scaled using simulations (Table 3). The Great Lakes region wolf genomes showed a majority of wolf-derived alleles (propwolf = 0.61 to 0.67), unlike the eastern wolves from Algonquin Provincial Park (propwolf = 0.39 to 0.47) and red wolves (propwolf = 0.09 to 0.20). The three nonreference coyote genomes had no estimated wolf admixture (propwolf = 0.00). We note that the results presented here are robust to changes in the specific genomes used for the wolf and coyote “reference” panels. For example, if we replace the three coyote genomes used with all six coyote genomes, the correlation in propwolf estimates in the two analyses is $r^2 = 0.988$.

Estimating admixture by D and $f$
We tested for admixture among wolves, dogs, and coyotes using the D statistic (also known as the ABBA-BABA test) and quantified the proportion of ancestry using $f$ (30, 31). We performed all possible D statistic comparisons among our samples and used a San Nicolas Channel Islands fox (Urocyon littoralis dickeyi) [table S1, (32)] as the outgroup to identify derived alleles. Below, we report tests as $D(P1, P2, candidate introgressor, outgroup).

Because of the substantial controversy regarding the proper classification of eastern wolves and red wolves, we began by testing which North American canids, regardless of species assignment in the field, shared the most derived alleles with each Eurasian wolf using $D$(North America1, North America2, Eurasian wolf, fox), where “North America1, North America2” refers to all combinations of North American canids.
We found that samples that were morphologically identified as coyotes tended to share the fewest derived alleles with Eurasian wolves (fig. S3), which is consistent with the expectation that gray wolves are a monophyletic species (33). The results of analyses with the Quebec wolf indicate that, like the California coyote and the Alabama coyote, it lacks detectable Eurasian wolf ancestry (Fig. 3 and fig. S3). This finding could reflect either an error in the field classification of the Quebec wolf specimen or a transcription error in the field or laboratory. Given the large variation in phenotype across the admixture zone [for example, Kolenosky et al. (34)] and the potential difficulty in making species-level assignments of hybrids under field conditions, we suggest that this genetic assignment of the “Quebec wolf” as a coyote is a more reliable guide to the ancestry of the sample than the field/laboratory assignment.

To determine the proportion of coyote ancestry in North American canids, we next calculated $D$ statistics using $D$(Eurasian wolf, North American wolf, California or Alabama coyote, fox) and quantified the

<table>
<thead>
<tr>
<th>Genome (reference)</th>
<th>Species, common name</th>
<th>Location</th>
<th>Group code</th>
<th>Average coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska (26)</td>
<td>C. lupus, gray wolf</td>
<td>Alaska, USA</td>
<td>NAGW</td>
<td>6.37</td>
</tr>
<tr>
<td>Alabama</td>
<td>C. latrans, coyote</td>
<td>Alabama, USA (REF)</td>
<td>COY</td>
<td>5.10</td>
</tr>
<tr>
<td>Algonquin1</td>
<td>C. lupus, eastern wolf*</td>
<td>Algonquin Provincial Park, Canada (HZ)</td>
<td>GLW</td>
<td>5.60</td>
</tr>
<tr>
<td>Algonquin2</td>
<td>C. lupus, eastern wolf*</td>
<td>Algonquin Provincial Park, Canada (HZ)</td>
<td>GLW</td>
<td>4.48</td>
</tr>
<tr>
<td>Boxer (26)</td>
<td>C. familiaris, dog (boxer)</td>
<td>Domestic dog</td>
<td>DOG</td>
<td>22.31</td>
</tr>
<tr>
<td>Basenji (63)</td>
<td>C. familiaris, dog (Basenji)</td>
<td>Domestic dog</td>
<td>DOG</td>
<td>21.93</td>
</tr>
<tr>
<td>California (26)</td>
<td>C. latrans, coyote</td>
<td>California, USA (REF)</td>
<td>COY</td>
<td>24.25</td>
</tr>
<tr>
<td>Florida</td>
<td>C. latrans, coyote</td>
<td>Florida, USA</td>
<td>COY</td>
<td>7.10</td>
</tr>
<tr>
<td>Minnesota (26)</td>
<td>C. lupus, Great Lakes wolf*</td>
<td>Minnesota, USA (HZ)</td>
<td>GLW</td>
<td>24.90</td>
</tr>
<tr>
<td>GShep (63)</td>
<td>C. familiaris, dog (German shepherd)</td>
<td>Domestic dog</td>
<td>DOG</td>
<td>23.86</td>
</tr>
<tr>
<td>Illinois</td>
<td>C. latrans, coyote</td>
<td>Illinois, USA</td>
<td>COY</td>
<td>6.21</td>
</tr>
<tr>
<td>India (26)</td>
<td>C. lupus, gray wolf</td>
<td>India (REF)</td>
<td>EuGW</td>
<td>15.88</td>
</tr>
<tr>
<td>Iran (54)</td>
<td>C. lupus, gray wolf</td>
<td>Iran (REF)</td>
<td>EuGW</td>
<td>28.70</td>
</tr>
<tr>
<td>IRNP</td>
<td>C. lupus, Great Lakes wolf*</td>
<td>IRNP, USA (HZ)</td>
<td>GLW</td>
<td>4.62</td>
</tr>
<tr>
<td>Kenya (26)</td>
<td>C. aureus, golden jackal</td>
<td>Kenya, Africa</td>
<td>JACK</td>
<td>25.74</td>
</tr>
<tr>
<td>Mexican (26)</td>
<td>C. lupus baileyi, Mexican wolf</td>
<td>New Mexico, USA</td>
<td>NAGW</td>
<td>23.88</td>
</tr>
<tr>
<td>Mongolia (54)</td>
<td>C. lupus, gray wolf</td>
<td>Mongolia (REF)</td>
<td>EuGW</td>
<td>25.65</td>
</tr>
<tr>
<td>Ohio</td>
<td>C. latrans, coyote</td>
<td>Ohio, USA</td>
<td>COY</td>
<td>5.92</td>
</tr>
<tr>
<td>Qinghai (54)</td>
<td>C. lupus, gray wolf</td>
<td>China (REF)</td>
<td>EuGW</td>
<td>27.03</td>
</tr>
<tr>
<td>QuebecCoy</td>
<td>C. latrans, coyote</td>
<td>Quebec, Canada (REF)</td>
<td>COY</td>
<td>6.41</td>
</tr>
<tr>
<td>QuebecWolf</td>
<td>C. lupus, gray wolf*</td>
<td>Quebec, Canada (HZ)</td>
<td>GLW</td>
<td>3.88</td>
</tr>
<tr>
<td>Redwolf1 (54)</td>
<td>C. rufus, red wolf</td>
<td>Captive breeding program, USA (HZ)</td>
<td>RW</td>
<td>28.30</td>
</tr>
<tr>
<td>Redwolf2</td>
<td>C. rufus, red wolf</td>
<td>Captive breeding program, USA (HZ)</td>
<td>RW</td>
<td>5.57</td>
</tr>
<tr>
<td>Redwolf3</td>
<td>C. rufus, red wolf</td>
<td>Captive breeding program, USA (HZ)</td>
<td>RW</td>
<td>6.29</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>C. lupus, Great Lakes wolf*</td>
<td>Wisconsin, USA (HZ)</td>
<td>GLW</td>
<td>7.49</td>
</tr>
<tr>
<td>Yellowstone1†</td>
<td>C. lupus, gray wolf</td>
<td>Yellowstone National Park, USA</td>
<td>NAGW</td>
<td>12.90</td>
</tr>
<tr>
<td>Yellowstone2 (26)</td>
<td>C. lupus, gray wolf</td>
<td>Yellowstone National Park, USA</td>
<td>NAGW</td>
<td>26.39</td>
</tr>
<tr>
<td>Yellowstone3† (26)</td>
<td>C. lupus, gray wolf</td>
<td>Yellowstone National Park, USA (REF)</td>
<td>NAGW</td>
<td>24.32</td>
</tr>
</tbody>
</table>

*Species designation uncertain: the Great Lakes region may contain a wolf-like endemic North American species, C. lycaon, only a mixture of gray wolf and coyote genes, or mixtures of all three taxa.
†Because of their parent-offspring relationship (Yellowstone1 and Yellowstone2 are the parents of Yellowstone3), we excluded these two individuals.
result using the related $f$ statistic. In this test, we use the California and Alabama coyotes as potential introgressors because they had the fewest derived gray wolf alleles (they were the most coyote-like samples in the data set) and originated from outside of one or both admixture zones. Although the Alabama coyote is from the southeastern United States and may therefore be admixed with the red wolf, tests to detect red wolf ancestry in this sample were consistently nonsignificant (table S1).

We found that all North American wolves and coyotes have significant amounts of coyote ancestry (table S1). In addition, we detect a strong geographic cline in the proportion of coyote ancestry across North American canids. Alaskan and Yellowstone wolves have 8 to 8.5% coyote ancestry, Great Lakes wolves have 21.7 to 23.9% coyote ancestry, Algonquin wolves have at least 32.5 to 35.5% coyote ancestry, and Quebec sequences have more than 50% coyote ancestry (Fig. 3). As expected, Eurasian wolves and dogs, which are allopatric to coyotes, do not have coyote ancestry (table S1).

Finally, we estimated the amount of gray wolf introgression into other canids in our data set. We detected significant amounts of gray wolf ancestry in Illinois coyotes ($\text{prop}_{\text{wolf}} \geq 0.06$), Florida coyotes ($\text{prop}_{\text{wolf}} \geq 0.09$), Ohio coyotes ($\text{prop}_{\text{wolf}} \geq 0.10$), and red wolves ($\text{prop}_{\text{wolf}} \geq 0.20$). The two Quebec individuals have different amounts of gray wolf ancestry: The Quebec wolf has no detectable gray wolf ancestry, consistent with mislabeling (see above), and the Quebec coyote has at least 15.8% gray wolf ancestry. The highest inferred proportion of gray wolf ancestry among nonadmixed individuals was 61.1% in the Basenji, suggesting that these proportions may underestimate the amount of wolf contribution to the coyote gene pool (table S2). However, these proportions are similar in magnitude and ranking to those independently estimated using diagnostic SNPs from the canine genotyping array (Fig. 1 and see below). These results also highlight the mixed ancestry of red wolves and wolves from the Great Lakes region including Algonquin Provincial Park, with the latter having a substantial proportion of gray wolf ancestry.

**Demographic analysis**

To better assess the demographic implications of a separate species origin, rather than one due entirely to admixture, we performed demographic inference by applying G-PhoCS (Generalized Phylogenetic Coalescent Sampler) to simple branching models (35). Notably, our models assume that red and eastern wolves have a phylogenetically distinct origin followed by admixture. Because of computational and coverage limitations, we focused on high-coverage genomes from nine individuals, each from a different population or species: a red wolf (Redwolf1), a Great Lakes region wolf with admixed history (Minnesota),

<table>
<thead>
<tr>
<th></th>
<th>NAGW</th>
<th>Coyote (REF)</th>
<th>Red wolf</th>
<th>GLW</th>
</tr>
</thead>
<tbody>
<tr>
<td>EuGW</td>
<td>0.099</td>
<td>0.160</td>
<td>0.188</td>
<td>0.076</td>
</tr>
<tr>
<td>NAGW</td>
<td>0.153</td>
<td>0.177</td>
<td>0.045</td>
<td>0.057</td>
</tr>
<tr>
<td>Coyote (REF)</td>
<td></td>
<td>0.108</td>
<td>0.045</td>
<td></td>
</tr>
<tr>
<td>Red wolf</td>
<td></td>
<td></td>
<td></td>
<td>0.085</td>
</tr>
</tbody>
</table>

**Fig. 2. PCA of 5.4 million unphased SNPs and 23 Canis genomes.** The dashed line contains genomes that are considered admixed.
Table 3. Estimated fraction of wolf-like versus coyote-like alleles at the 16,184 fixed differences between wolves and putatively unadmixed coyotes. See Table 1 for sample abbreviations.

<table>
<thead>
<tr>
<th>Sample</th>
<th>$N_{\text{coyote}}$</th>
<th>$N_{\text{heterozygous}}$</th>
<th>$N_{\text{wolf}}$</th>
<th>% Wolf alleles</th>
<th>% Wolf ancestry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida</td>
<td>746</td>
<td>4189</td>
<td>7114</td>
<td>23.6</td>
<td>0</td>
</tr>
<tr>
<td>Illinois</td>
<td>619</td>
<td>2577</td>
<td>7623</td>
<td>17.6</td>
<td>0</td>
</tr>
<tr>
<td>Ohio</td>
<td>717</td>
<td>3300</td>
<td>6103</td>
<td>23.4</td>
<td>0</td>
</tr>
<tr>
<td>Redwolf1</td>
<td>2,465</td>
<td>4657</td>
<td>8998</td>
<td>29.7</td>
<td>9.3</td>
</tr>
<tr>
<td>Redwolf2</td>
<td>2,106</td>
<td>3136</td>
<td>4864</td>
<td>36.4</td>
<td>20.3</td>
</tr>
<tr>
<td>Redwolf3</td>
<td>2,211</td>
<td>3983</td>
<td>5633</td>
<td>35.5</td>
<td>18.8</td>
</tr>
<tr>
<td>Alaska</td>
<td>8,544</td>
<td>2510</td>
<td>1064</td>
<td>80.9</td>
<td>93.9</td>
</tr>
<tr>
<td>Yellowstone1</td>
<td>10,400</td>
<td>2654</td>
<td>1536</td>
<td>80.4</td>
<td>93.1</td>
</tr>
<tr>
<td>Yellowstone2</td>
<td>11,150</td>
<td>3052</td>
<td>1873</td>
<td>78.9</td>
<td>60.6</td>
</tr>
<tr>
<td>Mexican</td>
<td>11,639</td>
<td>1114</td>
<td>3304</td>
<td>76</td>
<td>85.8</td>
</tr>
<tr>
<td>Algonquin1</td>
<td>2,516</td>
<td>5508</td>
<td>3021</td>
<td>47.7</td>
<td>39</td>
</tr>
<tr>
<td>Algonquin2</td>
<td>2,393</td>
<td>4855</td>
<td>1878</td>
<td>52.8</td>
<td>47.4</td>
</tr>
<tr>
<td>Minnesota</td>
<td>6,714</td>
<td>6165</td>
<td>3174</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>Isle Royale</td>
<td>4,372</td>
<td>2257</td>
<td>1987</td>
<td>63.8</td>
<td>65.6</td>
</tr>
<tr>
<td>QuebecWolf</td>
<td>131</td>
<td>807</td>
<td>1576</td>
<td>21.3</td>
<td>0</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>6,408</td>
<td>5501</td>
<td>2207</td>
<td>64.9</td>
<td>67.4</td>
</tr>
<tr>
<td>Boxer</td>
<td>12,001</td>
<td>1026</td>
<td>2219</td>
<td>82.1</td>
<td>95.9</td>
</tr>
<tr>
<td>GShep</td>
<td>12,684</td>
<td>1317</td>
<td>2115</td>
<td>82.8</td>
<td>97</td>
</tr>
<tr>
<td>Basenji</td>
<td>13,102</td>
<td>832</td>
<td>2172</td>
<td>83.9</td>
<td>98.8</td>
</tr>
</tbody>
</table>

Fig. 3. Estimates of ancestry proportions using the $\hat{f}$ statistic. Sequences grouped as coyotes are from Alabama, California, Illinois, Ohio, and Florida. (*Individual labeled as wolf but is likely to derive from a coyote; see discussion in the text.)
a California coyote, two North American wolves (Yellowstone2 and Mexican wolf), two Eurasian wolves (Mongolia and Croatia), Basenji, and a golden jackal (Table 1). Our objective was to infer rates of gene flow into red and Great Lakes region wolves in the context of a complete demographic model that includes population divergence and changes in ancestral population sizes. We assumed four different plausible topologies for the population phylogeny (fig. S4); however, to capture the large contribution of genomic ancestry from gray wolves into Great Lakes region wolves and coyotes into red wolves, we focused on a model in which the eastern wolf branched from the population ancestral to North American wolves and the red wolf branched from the coyote lineage (Fig. 4). We analyzed 13,647 previously determined putative neutral loci of 1 kb in length (26).

Under the assumed branching structure, we inferred high rates of gene flow from gray wolves and coyotes into the red wolf and the Great Lakes region wolf (fig. S4 and table S3). We converted these rates to admixture proportions and observed high proportions of coyote gene flow into the red wolf (48 to 88%) and high proportions of gene flow from the Yellowstone wolf into the Great Lakes region wolf (37 to 48%) (Fig. 4). High admixture proportions were also inferred from the coyote into the Great Lakes region wolf (25 to 34%) and between the red wolf and the Great Lakes region wolf (21 to 35%). We obtained similarly high estimated admixture proportions in the other three topologies examined (fig. S4), confirming that high rates of inferred gene flow were not a result of incorrect assumptions of the population phylogeny. Finally, the inferred divergence time under the model in Fig. 4 is relatively short for the Great Lakes region wolf (τ,GL,NA, 27 to 32 ka) (fig. S5). The estimated divergence time between the red wolf and the coyote is greater (τ,GL,NA, 55 to 117 ka), which may reflect the use of the high-coverage California coyote sequence, which is the most divergent sequence of our coyote samples (Fig. 2). This sequence is unlikely to represent the source of admixture for the red wolf in the American Southwest. Values of divergence time are even shorter in models without gene flow, as might be predicted (fig. S6). Consequently, these results show that even under the assumption of a distinct species origin, extensive gene flow into a recently diverged Great Lakes region wolf is needed to account for its genomic composition. Similarly, assuming a distinct origin for the red wolf still requires substantial gene flow from gray wolves and coyotes, although the inferred divergence time is greater. However, this divergence time might be inflated relative to that in which populations more directly ancestral to those comprising the red wolf admixture zone were used.

Novel genome ancestry
To specifically assess evidence for unique ancestry in red and Great Lakes region wolves, we tabulated the proportion of alleles in each genome sequence that was not found in our reference sample of four Eurasian gray wolves plus three largely nonadmixed (based on D statistics) coyotes (Fig. 3; California, Alabama, and Quebec). The California coyote is located outside known admixture zones, and the Alabama coyote, although within the red wolf hybrid zone, shows no excess ancestry with red wolves using D statistics (Fig. 1 and table S1; see above). The Eurasian wolves represent a conservative comparison because they will have diagnostic alleles from the common gray wolf ancestor, as well as uniquely derived alleles found only in Eurasian wolves that define a separate clade within gray wolves (33). Our initial analyses found that sequences with higher coverage have a larger proportion of unique alleles (Table 4 and table S4). Presumably, this is because the nonreference wolf and dog alleles tend to be rare and are less likely to be called by analysis of next-generation sequencing data (ANGSD) given the low coverage. We then reran our calculations after downsampling the high-coverage coyote, red wolf, and Great Lakes region wolf genomes to sixfold the average coverage (Table 4). We found that the proportion of new alleles was highest in the nonreference coyote samples (mean, 5.13%), followed by red wolves (mean, 4.41%), Algonquin wolves (mean, 3.82%), Great Lakes wolves (mean, 3.61%), and North American gray wolves (mean, 3.30%, including only samples with <10 times the average coverage). If we assumed that the red and eastern wolves were distinct species that hybridized with gray wolves and coyotes with proportions estimated as in Table 3, then the expectation is that they would have more novel alleles than actually observed. The fact that they do not provides additional support for our claim that these groups are recent gray wolf–coyote hybrid populations. Consequently, our results do not support an ancient (for example, >250 ka) independent ancestry for red or Great Lakes region wolves (including those for Algonquin Provincial Park), because such a history would have led to the observation of more “novel” alleles than found in our complete genome sequences from six Great Lakes region wolves and two red wolves (Table 1). We find higher levels of novelty in nonreference coyotes, presumably conspecific with our reference group of coyotes, than in red wolves and Algonquin wolves, thought to represent two distinct North American wolf-like species.

DISCUSSION
Our results suggest a surprisingly recent and admixed history of North American canids from the Great Lakes and southeastern regions of the United States (Fig. 1). Using five distinct approaches (FST, PCA, D statistics, G-PhoCS, and unique alleles), we find that a substantial proportion of the canid genomes from these geographic areas have admixed
ancestry. Further, even models that assume a distinct origin of Great Lakes region wolves and red wolves (Fig. 4) require both a large amount of admixture with canonical gray wolves and coyotes and a relatively recent origin of the former two species. These results contradict claims that red wolves and wolves of the Great Lakes region have ancestry from native North American wolves that share common ancestry with coyotes more than 250 ka (10). Our analyses suggest that all of the North American canids diverged from a common ancestor less than 6 to 117 ka and that both Great Lakes region wolves and red wolves are highly admixed with different proportions of gray wolf and coyote ancestry.

We found that coyote-derived ancestry is highest in individuals identified as red wolves from the southeastern United States and lowest among wolves of the Great Lakes region (for example, Minnesota, Isle Royale National Park, Wisconsin, and Algonquin Provincial Park) (Fig. 3). The south-to-north gradient of coyote ancestry (Fig. 3) is consistent with a known historical process in which wolf-like canids disappeared first from the American South and East, concurrent with early European colonization and the conversion of woodland habitat to agricultural landscape. Extirpation of wolves in the southeast followed shortly after the advent of private, state, and federal bounty beginning in the 1880s (1). As wolves became scarce, dispersing individuals would have a low probability of finding conspecific mates, resulting in an increase in coyote-wolf admixture. Only at the turn of the last century would a similar process occur in the Great Lakes region, as local gray wolf populations declined and coyotes expanded into the region (8). Coyotes and their hybrid descendants advanced eastward through Ontario, Quebec, the Maritime Provinces, and New England (8, 36).

Both the timetable of wolf extirpation and its thoroughness likely explain the observed gradient in coyote-gray wolf admixture. The early and complete extermination of wolves in the American South provided opportunities for admixture, which has created a varied, dominantly coyote ancestry mosaic across the genome [for example, vonHoldt et al. (5) and Tang et al. (37)]. In contrast, the more recent entrance of coyotes into the Great Lakes region and the continued abundance of gray wolves in much of the region have maintained a higher proportion of gray wolf ancestry (Figs. 1 to 3 and Table 3).

**Origins of the North American canids**

Consistent with the above results, Great Lakes region wolves and red wolves are admixed populations composed of various proportions of gray wolf and coyote ancestry. The red wolf may have captured genomic elements that were unique to gray wolves and coyotes of the American South (12). In contrast, the Great Lakes region wolves largely sample lineages of gray wolves and coyotes that have descendants in the extant population, which may include a distinct gray wolf ecotype (7, 8). We find little evidence of distinct genomic elements in either red or Great Lakes region wolves that would support separate evolutionary legacies. In general, the uniformly low values of uniqueness among all admixed samples (Table 4) are not consistent with the presence of a distinct, wolf-like canid in the American South or Eastern Canada and the eastern United States. The Mexican wolf is the most distinct North American gray wolf, and the California coyote is the most distinct coyote sequence in our data set.

Conceivably, an increased genomic sampling of red wolves and Great Lakes region wolves might reveal genomic segments from a distinct (and now-extinct) North American canid. However, a single genome captures much of the evolution history of a species [for example, Gronau et al. (35) and Durbin et al. (36)]. For example, the number of ancestors represented by a genome in a genealogy increases by an exponent of two with each generation; thus, a single genome represents more than 1000 ancestors from 20 generations ago or, alternatively, contains ~2 Mb of sequence from each of these ancestors. Consequently, our individual genome sequences have sampled a population of ancestors that would likely have included genomic contributions from a much larger historic population of a distinct wolf-like species. An empirical example is the presence of Neandertal ancestry in nearly all Europeans, which is detected at levels of less than 2% and has been shown to be virtually absent in sub-Saharan Africans (30, 39). This contribution from Neandertals is detected despite their extinction more than 30,000 years ago. Therefore, the historic presence of a distinct North American canid in the Great Lakes region seems unlikely, given our analysis of six Great Lakes region wolves from a wide geographic area, including Algonquin Provincial Park where pure eastern wolves are thought to exist (Fig. 1).

**Conservation implications**

The red wolf was listed as an endangered species in 1973, initiating a captive breeding program by the USFWS. The program began with 12 of 14 founding individuals that reproduced, selected from a panel of several hundred captured individuals that were thought to represent the ancestry spectrum ranging from coyote to pure red wolf and various admixtures of the two forms. These 12 founders were considered to be pure red wolves based on phenotypic characteristics and the lack of segregation of “coyote-like” traits in their offspring [for example, Wayne et al. (3)]. The descendants of these founders defined the ancestry of the several hundred red wolves produced by the captive breeding program and have been the source for a single reintroduced population in eastern North Carolina. Our results suggest that the red wolves selected for the captive breeding program had a higher fraction of gray wolf ancestry.
than is apparent in the current coyote population in the southern United States (Fig. 3). However, although this gray wolf genome may have been derived from a now-extinct population of southern wolves, the recent coalescence of all modern gray wolves (Fig. 3) (26–28) suggests that even these distinct wolf ancestors were not so genetically divergent from modern conspecifics to be considered a distinct species.

Although the historical scenario for the Great Lakes region wolves is similar to that for red wolves, there are several important differences. First, the admixture zone appeared more recently as coyotes arrived in the Great Lakes area primarily within the past century, although both genetic and ancient DNA analyses suggest that they may have been there earlier [for example, (5, 8, 10, 13, 40)]. Moreover, admixture between gray wolves and coyotes occurred throughout the eastward expansion, because gray wolves were never completely extirpated in this region and remained abundant in some areas. This population of recently admixed individuals has been documented as an important regional predator of deer throughout much of eastern Canada, and their large admixed coyote-like relatives may have a similar role in New England and the eastern United States (8).

Recently, the USFWS (17) has recognized an eastern wolf species that historically inhabited the Great Lakes region and 29 eastern states to the exclusion of the gray wolf. Thus, designating the eastern wolf would imply that this entity could be considered endangered under the ESA, because it is now restricted to a fraction of its historic range and is threatened by hybridization. Rather than formally recognizing this threatened status, the USFWS has accepted the revised taxonomy and proposed that the original listing of the gray wolf is incorrect because it included the geographic range of the eastern wolf. Consequently, the USFWS concluded that the gray wolf listing under the ESA is invalid and should be revoked. However, our results suggest that the genomes from the Great Lakes region show little taxonomic distinction and that only two distinct North American species (coyote and gray wolf) are supported as inhabiting the Great Lakes region.

Our findings provide a critical heuristic lesson in endangered species management. The overly strict application of taxonomy to support endangered species status is antiquated. Species and taxonomic concepts are varied, complex, and difficult to apply in practice (41–44). Of greater importance are the preservation of evolutionary and ecological processes and the role of an endangered taxon in this dynamic. Admixture is one critical example of a process that may enhance adaptation and evolution in the rapidly changing environment of the modern world (44–47). Smaller wolf-life canids, such as the Great Lakes region wolf and the red wolf, may be more appropriate predators in the increasingly fragmented habitats of eastern North America than larger western gray wolves that require extensive pristine habitats with massive ungulates. We maintain that the ESA could be interpreted in a modern evolutionary framework, devaluing the Victorian typological concept in exchange for a more dynamic view that allows for natural selection to occur on admixed genomes and to evolve phenotypes that are adapted to human-altered habitats and changing climates. These suggestions follow the "ecological authenticity" concept, in which admixed individuals that have an ecological function similar to that of the native endangered taxon, and that maintain a portion of the endangered genetic ancestry, warrant protection (48). Additionally, we suggest that there should be a possibility of recovery of the endangered species gene pool (49). For the Great Lakes region wolf, preservation of areas that favor abundant self-sustaining populations of wolves to the exclusion of coyotes may allow the wolf population to become increasingly gray wolf in genetic composition and approach the historic population (40). For example, after the reintroduction of the gray wolf to Yellowstone National Park, an area that has ideal wolf habitat, coyote populations decreased initially and no wolf/coyote hybridization has been observed (50). Preservation of such high-value wolf habitat in the Great Lakes may likewise provide the best option for genetic restoration of the population through natural processes (40). Unfortunately, as a consequence of the extirpation of gray wolves in the American Southeast, the reintroduced population of red wolves in eastern North Carolina is doomed to genetic swamping by coyotes without the extensive management of hybrids [for example, Gese et al. (51)] as is currently practiced by the USFWS. Further, the absence of the ancestral population of gray wolves that once existed in the American South means that the historical gene pool cannot be readily reconstructed by conservation actions.

Finally, our results inform the larger debate on the importance of genomic analysis in conservation. We demonstrate the utility of complete genome analysis over techniques that assay variation in a limited fraction of the genome. Namely, we were able to evaluate specific evolutionary and demographic hypotheses and test the notion of unique ancestry as opposed to hybrid origin. Such genomic analyses have direct relevance to the protection and management of admixed species and the pending decisions about the fate of the red wolf and the eastern wolf (17, 45, 52, 53).

MATERIALS AND METHODS

Genomic library construction and sequencing

We used previously published genomes (n = 3) and sequenced additional genomes (n = 25) to address the controversy concerning species origin and admixture in North American canids (Table 1). The 28 canine genomes were all sequenced on an Illumina HiSeq platform and include four Eurasian gray wolves (EuGW) (India, n = 1; Mongolia, n = 1; China, n = 1; and Iran, n = 1) and five North American gray wolves (NAGW) (Mexico, n = 1; Alaska, n = 1; and Yellowstone National Park, n = 3). Six Great Lakes region wolves (GLW) (including two Algonquin Provincial Park wolves) and three red wolves (RW) were also represented (Table 1). Additionally, six coyotes (COY) and three domestic dogs (DOG) were included, with one golden jackal (Canis aureus) (JACK). Given that the samples from the Great Lakes region (Fig. 1) may represent gray wolves, eastern wolves, or their hybrids with each other and coyotes, we designated all these samples as coming from a hybrid zone. However, this grouping includes the two specimens from Algonquin Provincial Park thought to contain the last population of relatively pure eastern wolves (20). Therefore, of the North American samples, 10 were derived from nonadmixed populations residing outside the hybrid zone and 7 were derived from the putatively admixed Great Lakes region population, including one coyote (Table 1).

Following our previous library preparation and sequencing protocol (26), we constructed genomic paired- and single-end sequencing libraries with an average insert size of 300 to 500 bp. Approximately 5 µg of purified genomic DNA was fragmented by sonication using the Covaris Adaptive Focused Acoustics system. Both 3’ and 5’ overhangs of the recovered genomic DNA fragment were converted into blunt ends using T4 DNA polymerase and Klenow enzyme (New England BioLabs). After end repair, an adenine was added through ligation using Klenow 3’-to-5’ Exo-(New England BioLabs). The standard TruSeq paired-end adapters were also ligated using the Quick DNA Ligation Kit (New
England BioLabs). The ligated products were separated on a 2% agarose gel, and the desired DNA fragments were recovered from the gel using the QIAquick Gel Extraction Kit (Qiagen). After the initial denaturation at 98°C for 30 s, the polymerase chain reaction was carried out for eight cycles of 98°C for 10 s, 65°C for 30 s, and 72°C for 30 s, with the final extension for 5 min at 72°C using Phusion DNA Polymerase (Finnzymes). Libraries were sequenced on an Illumina HiSeq 2000, following the manufacturer’s standard cluster generation with a V2 Paired End Cluster generation kit and sequencing protocol with TruSeq SBS sequencing reagents. Base calling was performed with the on-instrument computer using RTA (version 1.7). Reads were trimmed and clipped using Trim Galore (version 0.3.7), discarding reads that were <20 bp in length (www.bioinformatics.babraham.ac.uk/projects/trim_galore/) to exclude sites of low quality (<30) and remnant TruSeq adapter sequence. All reads were mapped to the recent dog genome assembly (CanFam3.1, GenBank Assembly ID GCA_000002285.2) generated from a boxer breed individual using stampy (54), discarding reads that were present. We used a minimum inclusion threshold of 10-fold sequence coverage per site. SNP's were called using ANGSD (56) and monomorphic sites were excluded. The remaining 5.4 million SNPs represent the 38 canine autosomes, the X chromosome, and the mitochondrial genome. Finally, all sequences in this study have high proportions of reads mapping to the dog genome (>95%). As found in previous analyses, the genus Canis that includes all North American large canids are very recently diverged and have high similarity at the genome level, which is generally similar to that found within many species, such as humans (26, 28, 33).

Statistical analysis

Diversity and divergence analysis. We used a series of established pipelines to analyze the sequence data with regard to relationships of the eastern wolf and red wolf to gray wolf and coyote genome sequences (5, 26). SNP genotypes were called by ANGSD to tabulate the number of heterozygous sites and the fraction of un-called genotypes for each genome (Table 1). Assuming that the autosomal reference genome is ~2 Gb in size, we calculated heterozygosity as the number of heterozygous sites and the fraction of un-called genotypes for each parameter combination. Finally, we constructed a profile likelihood curve for T using standard asymptotic likelihood assumptions and exponential interpolation of log-likelihood values to estimate a 95% confidence interval. We decided to interpret the results in a frequentist context rather than a Bayesian one, because we felt that we had no meaningful knowledge regarding the priors for any of the parameters.

Cluster and ancestry analysis. We performed PCA to identify putatively admixed genomes. For the Yellowstone National Park wolves with known pedigree relationships, we excluded Yellowstone3 because it is the known offspring of the two other individuals, Yellowstone1 and Yellowstone2 (Table 1). We first used the PCA function in the Genome Diversity toolbox in Galaxy (58–60) to conduct a PCA of all North American unphased SNP genotypes. We excluded the golden jackal to avoid the potential strong polarizing influence of a distant outgroup.

Under a simple two-way admixture model of gray wolves and coyotes, we calculated the proportion of "wolf-like" versus "coyote-like" alleles that were present in our sequences. Specifically, we used 16,184 fixed differences between the three reference coyotes that showed no evidence of admixture with gray wolves based on D statistics (Alabama, California, and Quebec) and four Eurasian gray wolves to estimate wolf versus coyote ancestry proportions (Table 3; see discussion in the text). For all other North American canid samples, we examined the called genotypes from each of these sites and tabulated the number of "wolf" and "coyote" alleles present. We estimated wolf allele proportions in each North American canid sample as the number of "wolf" alleles divided by the total number of called alleles. To correct for incomplete lineage sorting, we tabulated the expected proportion of "wolf" alleles for an additionally sampled coyote or wolf, under a simple isolation model. Then, we transformed the allele proportions into wolf ancestry proportion estimates using linear interpolation. Individuals with smaller proportions of "wolf" alleles than found in the simulated additional coyote were assigned a wolf ancestry fraction of 0.

This analysis implicitly assumes low levels of admixture within the reference coyote and wolf genomes. Our D statistic analyses (Fig. 3) provide evidence that the three reference coyotes are nonadmixed and that the Alabama coyote, although within the red wolf hybrid zone, does not show excess ancestry with red wolves. The Eurasian wolves are presumed nonadmixed with North American canid groups because of geography. Nonetheless, we tested other potential sets of reference populations (for example, including all coyotes, only the California coyote, and/or all gray wolves) and found in all cases a strong correlation ($r^2 > 0.9$) between the resulting estimates of wolf ancestry proportions.

Admixture detection and ancestry estimation. We used D statistics (also known as ABBA-BABA tests) and $f$ statistics to test and then quantify the impact of admixture between wolves and coyotes. To test for admixture between wolves and coyotes, we conducted D statistic tests involving three samples plus an outgroup for admixture [D(P1, P2, I, O)] for all possible combinations of wolves, coyotes, and dogs (Table S1). All comparisons used a San Nicolas Channel Islands fox (U. littoralis dickeyi) as the outgroup for determining the ancestral state at variable sites. The golden jackal (C. aureus) was not used, because the sequence has significant evidence of admixture with gray wolves (26). To avoid comparisons between nonhomologous sequences, we excluded sites with coverage in excess of the 95th percentile for each individual. We reduced the impact of coverage variation between samples by downsampling each individual to a single high-quality (base quality ≥ 30 and read mapping quality ≥ 30) base call chosen randomly from all of the high-quality base calls at each site. Such downsampling allows unbiased comparisons between samples of different coverages
where either heterozygous site calling or consensus base calling might be biased toward the reference genome, in this case a dog, and therefore potentially introduce bias into admixture estimates.

To test for possible impacts of reference genome bias on our study, we examined $D$ statistic tests for nonadmixed coyote (California and Alabama) introgression into the Eurasian dogs and wolves [for example, $D$(Eurasian wolf, dog, coyote, fox)], which are allopatric to coyotes and therefore expected to be free of coyote ancestry. As the outgroup, the San Nicolas Channel Islands fox is the most divergent from the reference genome and thus the most susceptible to reference genome bias (61). Impacts of this bias on $D$ statistic calculation would manifest as individuals that are more closely related to the reference genome individual [the boxer; Reference NCBI (National Center for Biotechnology Information) no. GCF_000002285.3], sharing more alleles than expected with the outgroup. In terms of $D$ statistics, this would result in $D$(Eurasian wolf, dog, reference genome, coyote, fox) < 0 because allele sharing between the outgroup and the reference genome creates a false excess of shared derived alleles between the Eurasian wolf and the coyote. We found that the specific reference genome individual, the boxer, exhibited a slight lack of coyote ancestry of $D = 0.03$ but that $D$ statistic comparisons involving other dogs in our data set $D$(Eurasian wolf, dog, coyote, fox) were not significantly different from zero. A difference would be expected if reference bias were significantly affecting the $D$ statistic calculation in those individuals. To avoid reference genome bias affecting our analyses, we excluded the boxer sequence that was used to generate the dog reference genome (24) from admixture estimates, and although there was no evidence for bias affecting dogs, we used Eurasian wolves for $f$ calculations.

To quantify the impact of gene flow on the hybrid populations, we used the $f$ statistic (30, 31), which is related to the $D$ statistic but uses population diversity to estimate the proportion of the genome derived from admixture. $f$ estimates the amount of introgressed ancestry in a candidate hybrid by dividing the excess of derived alleles shared between a candidate introgressor and a candidate hybrid by the excess of shared derived alleles that would be present if the candidate hybrid were 100% derived from the other species. To estimate $f$, we calculated the numerator of a typical $D$ statistic comparison [for example, $D$(wolf, hybrid, coyote, fox)] divided by the numerator of a $D$ statistic comparison, where the hybrid was replaced by a second nonadmixed member of the introgressing species [for example, $D$(wolf, coyote, coyote, fox)] (30, 31). For our $f$ estimates of wolf and coyote ancestry, we used nonadmixed Eurasian wolves to represent nonadmixed wolf samples and the coyotes that had the least allele sharing with wolves, the California coyote and the Alabama coyote, to represent nonadmixed coyotes (table S2). A known limitation of $f$ is that it is biased toward underestimation of the introgressed fraction of the genome, and underestimation is more severe for cases of older introgression (31). The results reported here should be interpreted therefore as a lower-bound estimate of the amount of introgression.

We measured the statistical significance of both $D$ and $f$ using a weighted block jackknife. We divided the genome into 5-Mb nonoverlapping blocks, removed each block in turn, and recalculated the $D$ or $f$ statistic on the remaining blocks to determine an SE estimate (30, 62). This estimates how the signal of admixture is uniformly distributed across the genome, making it possible to assess whether a small number of anomalous regions drive the apparent signal of admixture. We considered estimates of $D$ or $f$ that are more than three times the SE ($Z \geq 3$) to be significant evidence of admixture.

Demographic analysis. We performed demographic inference by applying G-PhoCS to canid genomes of nine different populations, focusing on high-coverage genomes (Fig. 4 and Table 1): coyote (California), red wolf (Redwolf1), Great Lakes region wolf (Minnesota), Mexican wolf, Yellowstone wolf (Yellowstone2), Mongolian gray wolf, Croatian gray wolf (26), Basenji, and golden jackal (26). Standard filters were applied to the genome sequences to reduce the influence of sequencing errors and strong natural selection, and 13,647 previously determined putative neutral loci of 1 kb in length were extracted (26). The main analysis, based on a population phylogeny with topology as shown in Fig. 4, was supplemented with three additional analyses based on alternative topologies (fig. S4A). All four models assumed that dogs diverged from the population ancestral to all Eurasian wolves and that coyotes diverged from the population ancestral to all gray wolves, and assumed golden jackals to be an earlier divergence from all other canids (26, 28). In two of these models, the red wolf was assumed to be a sister species to coyote, with the Great Lakes region wolf branching from the population ancestral to the two North American wolves (model 1; main analysis) or from the population ancestral to all gray wolves (model 1a). In the other two models, the red and Great Lakes region wolves were assumed to be sister lineages, either as a sister clade to coyote (model 2) or as a sister clade to all gray wolves (model 2a).

In all four analyses, we allowed gene flow between the red wolf and the remaining eight canid populations and gene flow between the Great Lakes region wolf and the remaining eight canid populations. We also modeled ancestral gene flow from the golden jackal outgroup to the population ancestral to all other canids (COY_ANC) to assess ancient admixture, as previous described (26). In total, our models had 17 migration parameters including admixture and migration between all canid populations. For our best estimate of gene flow with the coyote, with the Great Lakes region wolf branching from the population ancestral to the two North American wolves (model 1; main analysis) or from the population ancestral to all gray wolves (model 1a). In the other two models, the red and Great Lakes region wolves were assumed to be sister lineages, either as a sister clade to coyote (model 2) or as a sister clade to all gray wolves (model 2a).

Novel genomic sequences. To further test for admixture between North American canid populations, we quantified how much novel nongray wolf genetic variation was present in each North American canid sample. We started with a gray wolf reference group of putatively nonadmixed samples consisting of four Eurasian wolves (India, Iran, Mongolia, and China) and three coyotes (Alabama, California, and Quebec). Then, for each other sample, we tabulated the fraction of SNP positions representing novel variation not found in the reference group (gray wolves or coyotes). We sampled one allele at random from each individual to control for recent inbreeding. We then divided by the total number of SNPs called in the reference plus test samples (seven individuals) to obtain the fraction of “new SNP positions” (table S4).

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/2/7/e1501714/DC1
Supplementary Results
15. C. J. Kyle, A. R. Johnson, B. R. Patterson, P. J. Wilson, K. Shami, S. K. Grewal, B. N. White,
16. R. J. Fredrickson, P. W. Hedrick, Dynamics of hybridization and introgression in red wolves
18. R. J. Fredrickson, P. W. Hedrick, Dynamics of hybridization and introgression in red wolves
19. vonHoldt
20. to six times (in parenthesis).
22. Profile likelihood curve for
23. estimates for wolf and coyote ancestry proportions in each sample across the 38 autosomes.
25. American canids, following
26. of northeastern coyotes via
31. K. Wayne, S. M. Jenks, Mitochondrial DNA analysis implying extensive hybridization of 
32. R. Kays, A. Curtis, J. J. Kirchman, Rapid adaptive evolution of northeastern coyotes via
33. R. Kays, Curtis, J. J. Kirchman, Rapid adaptive evolution of northeastern coyotes via 
34. Canis lupus/lycaon
36. N. Stange-Thomann, S. Stavropoulos, C. Stone, S. Stone, S. Sykes, P. Tchuinga, 
38. C. Jones, A. Kamat, B. Kanga, C. Kells, D. Khazanovich, A. Kieu, P. LeVine, J. Liu, X. Liu, 
42. A. Freedman, I. Gronau, R. M. Schweizer, D. Ortega-Del Vecchio, E. Han, P. M. Silva, M. 
43. Galaverni, Z. Fan, P. Marx, B. Lorente-Galdos, H. Beale, O. Ramirez, F. Hormozdari, 
45. F. Haller, J. A. Leonard, Hybridization among three native North American Canis species in 
47. T. Wheeler, B. N. White, Genetic analysis of historic western Great Lakes region wolf 
49. B. J. Theberge, M. T. Theberge, D. R. Voigt, W. Waddell, DNA profiles of the eastern Canadian wolf and the 
51. P. Skogland, E. Ermakov, E. Palkopoulou, L. Dalén, Ancient wolf genome reveals an early 
52. M. M. Gray, J. M. Granka, C. D. Bustamante, N. B. Sutter, A. R. Boyko, L. Zhu, E. A. Ostrander, 
54. L. Y. Rutledge, S. Devillard, J. Q. Boone, P. A. Hohenlohe, B. N. White, RAD sequencing and 
57. R. K. Wayne, S. M. Jenks, Mitochondrial DNA analysis implying extensive hybridization of 
58. J. K. Young. Production of hybrids between western gray wolves and western coyotes. PLOS One 9, 
61. R. W. Hedrick, R. J. Fredrickson, Captive breeding and the reintroduction of Mexican and 


Acknowledgments: We express our gratitude to the Morris Animal Foundation for their pioneering availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Genome sequence and sample metadata are archived and available from RKW upon request.

Submitted 25 November 2015
Accepted 28 June 2016
Published 27 July 2016
10.1126/sciadv.1501714

Whole-genome sequence analysis shows that two endemic species of North American wolf are admixtures of the coyote and gray wolf
Bridgett M. vonHoldt, James A. Cahill, Zhenxin Fan, Ilan Gronau, Jacqueline Robinson, John P. Pollinger, Beth Shapiro, Jeff Wall and Robert K. Wayne

Sci Adv 2 (7), e1501714.
DOI: 10.1126/sciadv.1501714