EVOLUTIONARY GENETICS

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. Wayne[5]*

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*).
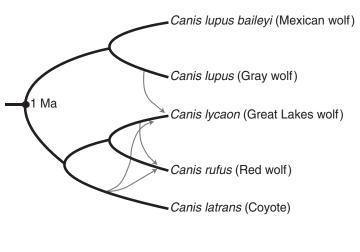
The U.S. Fish and Wildlife Service (USFWS) accepts the species status of both red and eastern wolves, with markedly divergent con-servation implications. The red wolf is protected by the U.S. Endangered Species Act (ESA). However, the endangered eastern wolf, which was only recently recognized as a distinct 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

[1]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA. [2]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Santa Cruz, CA 95064, USA. [3]Sichuan Key Laboratory of Conservation Biology on Endangered Wildlife, College of Life Sciences, Sichuan University, Chengdu 610064, People's Republic of China. [4]Efi Arazi School of Computer Science, Herzliya Interdisciplinary Center, Herzliya 46150, Israel. [5]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles, CA 90095–1606, USA. [6]Department of Epidemiology and Biostatistics, School of Medicine, University of California, San Francisco, San Francisco, CA 94143, USA.
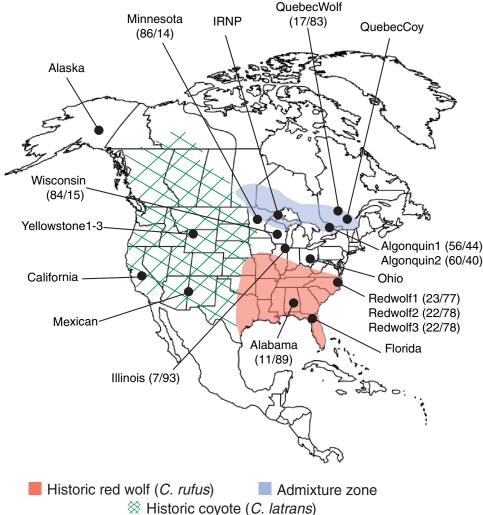*Corresponding author. Email: rwayne@ucla.edu

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 ($\pi = 1.71$/kb) and lowest in the endangered Mexican wolf ($\pi = 0.48$/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 $\hat{f}$" 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, (*1*)] 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 ($prop_{wolf} = 0.61$ to 0.67), unlike the eastern wolves from Algonquin Provincial Park ($prop_{wolf} = 0.39$ to 0.47) and red wolves ($prop_{wolf} = 0.09$ to 0.20). The three nonreference coyote genomes had no estimated wolf admixture ($prop_{wolf} = 0.00$). 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 $prop_{wolf}$ estimates in the two analyses is $r^2 = 0.988$.

Estimating admixture by D and $\hat{f}$

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 $\hat{f}$ (*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.

Table 1. Samples, origin, and genome code used in the article; average genome coverage; and ancestry proportions. If a population resides within the gray wolf and coyote hybrid zone, the location is indicated with "HZ"; reference populations are indicated by "REF." When previously sequenced, the appropriate citation is provided. EuGW, Eurasian gray wolf; NAGW, North American gray wolf; RW, red wolf; GLW, Great Lakes region wolf; COY, coyote; DOG, dog; JACK, golden jackal.

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

*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.

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

result using the related $\hat{f}$ 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).

Table 2. Pairwise F_{ST} estimates between canid lineages. Abbreviations are found in Table 1.

	NAGW	Coyote (REF)	Red wolf	GLW
EuGW	0.099	0.160	0.188	0.076
NAGW		0.153	0.177	0.057
Coyote (REF)			0.108	0.045
Red wolf				0.085

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 ($prop_{wolf} \geq 0.06$), Florida coyotes ($prop_{wolf} \geq 0.09$), Ohio coyotes ($prop_{wolf} \geq 0.10$), and red wolves ($prop_{wolf} \geq 0.20$). 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),
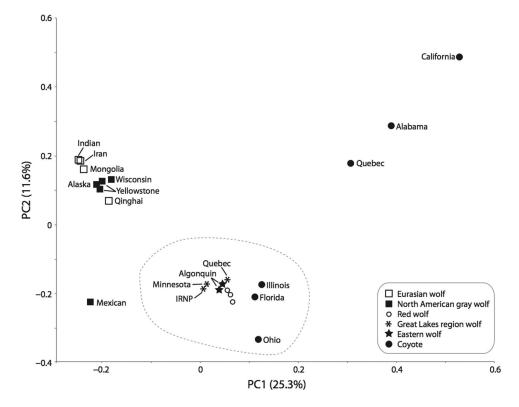

Fig. 2. PCA of 5.4 million unphased SNPs and 23 *Canis* genomes. The dashed line contains genomes that are considered admixed.

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.

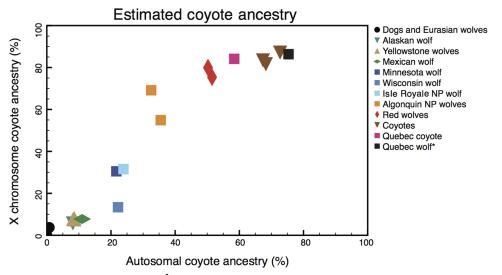| Sample	N_{coyote}	$N_{heterozygous}$	N_{wolf}	% Wolf alleles	% Wolf ancestry
Florida	746	4189	7114	23.6	0
Illinois	619	2577	7623	17.6	0
Ohio	717	3300	6103	23.4	0
Redwolf1	2,465	4657	8998	29.7	9.3
Redwolf2	2,106	3136	4864	36.4	20.3
Redwolf3	2,211	3983	5633	35.5	18.8
Alaska	8,544	2510	1064	80.9	93.9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

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
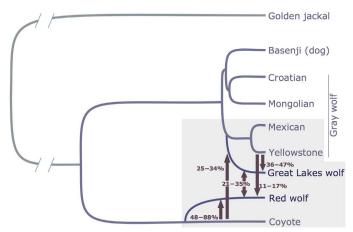
Southeast. Values of divergence time are even shorter in models without gene flow, as might be predicted (fig. S6). Consequently, these results show that even under the assumption of a distinct species origin, extensive gene flow into a recently diverged Great Lakes region wolf is needed to account for 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.

Fig. 4. Demographic history inferred using *G-PhoCS*. A schematic depiction of the population phylogeny assumed in the analysis. The phylogeny was augmented with migration bands from all canids to the red wolf and the Great Lakes region wolf. *G-PhoCS* infers significant rates of gene flow primarily from the gray wolf and the coyote to the red wolf and the Great Lakes region wolf (shaded box). Ninety-five percent Bayesian credible intervals are shown for the total rates transformed into proportions between 0 and 100% (see Materials and Methods). Similarly high rates were also inferred when assuming three alternative topologies for the population phylogeny (fig. S4).

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 (F_{ST}, PCA, D statistics, *G-PhoCS*, and unique alleles), we find that a substantial proportion of the canid genomes from these geographic areas have admixed

Table 4. Fraction of unique (non-EuGW + COY-ref) alleles in the complete sequence data and in sequences downsampled to six times.

Sample	New SNP%	New SNP% (downsampled)	Coverage
Basenji*	6.57	6.42	21.93
GShep	6.82	2.51	24.25
Redwolf1	8.78	4.62	28.30
Minnesota	7.13	3.96	24.90
Redwolf2	3.47	3.89	5.57
Redwolf3	4.28	4.73	6.29
Algonquin2	3.59	4.08	4.48
Algonquin1	3.19	3.55	5.60
IRNP	2.09	2.24	4.62
QuebecWolf	2.41	3.22	3.88
Wisconsin	4.37	4.64	7.49
FLcoy	4.18	5.19	3.88
ILcoy	4.57	5.8	6.21
OHcoy	3.56	4.4	5.92

*Genome was not downsampled.

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. (38)]. 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 composi-

tion 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

vonHoldt et al. Sci. Adv. 2016;2:e1501714   27 July 2016

9 of 13

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*). BAM files were indexed and sorted, and VCF (variant call format) files were produced with SAMtools (*55*). We used a minimum inclusion threshold of 10-fold sequence coverage per site. SNPs 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 uncalled genotypes for each genome (Table 1). Assuming that the autosomal reference genome is ~2 Gb in size, we calculated heterozygosity as π, the proportion of segregating sites in a specified length of DNA. Next, we stratified the samples into groups and calculated F_{ST} between Eurasian gray wolves, North American wolves, coyotes, red wolves, and wolves from the Great Lakes region using a custom script. Each calculation used only those SNPs where genotypes were called in all of the samples.

We also compared the observed value of F_{ST} between Eurasian wolves and coyotes with the values expected under a simple isolation model that describes wolf and coyote history. We used a summary likelihood approach to estimate the scaled divergence time T (in units of N generations, where N is the effective population size) using F_{ST} (*57*) as a summary of the data. The scaled mutation and recombination rates θ (= $4N\mu$, where μ is the mutation rate per base pair per generation) and ρ (= $4Nr$, where r is the recombination rate per base pair per generation) were included in the model as nuisance parameters. For each collection of model parameters (T, θ, and ρ), we simulated 40,000 different 1-Mb regions using the standard coalescent simulator ms. We then repeatedly subsampled 2000 of these regions and tabulated the probability that the value of F_{ST} for the subsampled regions was approximately equal to (that is, between 0.159 and 0.161) the observed F_{ST} of 0.160. We ran simulations over a grid of parameter values (T = 0.36 to 0.40, θ = 0.5 to 2/kb, and ρ = 0.2 to 1/kb) and resampled 10,000 subsets of the simulated data for each parameter combination. Finally, we constructed a profile likelihood

curve for T using standard asymptotic likelihood assumptions and linear 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 genotype 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 wolf 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 $\hat{f}$ 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

vonHoldt *et al. Sci. Adv.* 2016;2:e1501714   27 July 2016

10 of 13

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 $\hat{f}$ calculations.

To quantify the impact of gene flow on the hybrid populations, we used the $\hat{f}$ statistic (30, 31), which is related to the D statistic but uses population diversity to estimate the proportion of the genome derived from admixture. $\hat{f}$ 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 $\hat{f}$, 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 $\hat{f}$ 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 $\hat{f}$ 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 $\hat{f}$ using a weighted block jackknife. We divided the genome into 5-Mb nonoverlapping blocks, removed each block in turn, and recalculated the D or $\hat{f}$ 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 $\hat{f}$ 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 from nine different popula-

tions, 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 bands augmenting the population phylogeny. We implemented G-PhoCS using the standard setting described by Freedman et al. (26) for 100,000 burn-in iterations and an additional 200,000 sampling iterations. We recorded sampled parameter values every 50 iterations, resulting in a total of 4001 sampled values for each parameter, which we summarized using the mean value and 95% Bayesian credible intervals. Gene flow was initially measured using the total migration rate, which is equal to the per-generation rate times the number of generations that migration is allowed. When the total rate is low, it approximates the probability of gene flow between the two populations. For higher rates, we converted probabilities into rates using the formula $p = 1 - e^{-m}$ (where p is the probability of gene flow and m is the total migration rate). Other demographic parameters were calibrated by assuming a per-generation mutation rate of $\mu = 4 \times 10^{-9}$/bp (27) and an average generation time of 3 years (26).

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

table S1. The *D* statistic test for admixture in coyotes and North American hybrids. Results are calculated separately for the X chromosome and the autosomes.

table S2. $\hat{f}$ estimates for wolf and coyote ancestry proportions in each sample across the 38 autosomes.

table S3. Estimates of total migration rates and migration proportions.

table S4. Fraction of nonwolf alleles in the full-sequence data and in sequences downsampled to six times (in parenthesis).

fig. S1. Profile likelihood curve for *T*, the time of split between gray wolves and coyotes under a simple isolation model (blue).

fig. S2. PCA of downsampled genomes.

fig. S3. Results of all possible *D* statistic tests for Eurasian wolf introgression into North American canids, following *D*(P1, P2, Eurasian wolf, fox), where P1 and P2 are all possible combinations of North American canids.

fig. S4. Estimates of total migration rates under four different models assumed in four separate *G-PhoCS* runs.

fig. S5. Estimates of population divergence times and effective population sizes in the four separate *G-PhoCS* runs.

fig. S6. Estimates of divergence time obtained in eight separate runs of *G-PhoCS*.

REFERENCES AND NOTES

1. R. M. Nowak, Evolution and taxonomy of coyotes and related Canis, in *Coyotes: Biology, Behavior, and Management,* M. Bekoff, Ed. (Academic Press, New York, 1978), pp. 3–16.

2. S. M. Chambers, S. R. Fain, B. Fazio, M. Amaral, An account of the taxonomy of North American wolves from morphological and genetic analyses. *N. Am. Fauna* **77**, 1–67 (2012).

3. R. K. Wayne, S. M. Jenks, Mitochondrial DNA analysis implying extensive hybridization of the endangered red wolf *Canis rufus. Nature* **351**, 565–568 (1991).

4. M. S. Roy, E. Geffen, D. Smith, R. K. Wayne, Molecular genetics of pre-1940 red wolves. *Conserv. Biol.* **10**, 1413–1424 (1996).

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

6. N. Lehman, A. Eisenhawer, K. Hansen, L. D. Mech, R. O. Peterson, P. J. P. Gogan, R. K. Wayne, Introgression of coyote mitochondrial DNA into sympatric North American gray wolf populations. *Evolution* **45**, 104–119 (1991).

7. S. Koblmüller, M. Nord, R. K. Wayne, J. A. Leonard, Origin and status of the Great Lakes wolf. *Mol. Ecol.* **18**, 2313–2326 (2009).

8. R. Kays, A. Curtis, J. J. Kirchman, Rapid adaptive evolution of northeastern coyotes via hybridization with wolves. *Biol. Lett.* **6**, 89–93 (2010).

9. L. D. Mech, B. W. Christensen, C. S. Asa, M. Callahan, J. K. Young, Production of hybrids between western gray wolves and western coyotes. *PLOS One* **9**, e88861 (2014).

10. P. J. Wilson, S. Grewal, I. D. Lawford, J. N. M. Heal, A. G. Granacki, D. Pennock, J. B. Theberge, M. T. Theberge, D. R. Voigt, W. Waddell, DNA profiles of the eastern Canadian wolf and the red wolf provide evidence for a common evolutionary history independent of the gray wolf. *Can. J. Zool.* **78**, 2156–2166 (2000).

11. S. K. Grewal, P. J. Wilson, T. K. Kung, K. Shami, M. T. Theberge, J. B. Theberge, B. N. White, A genetic assessment of the eastern wolf (*Canis lycaon*) in Algonquin Provincial Park. *J. Mammal.* **85**, 625–632 (2004).

12. F. Hailer, J. A. Leonard, Hybridization among three native North American Canis species in a region of natural sympatry. *PLOS One* **3**, e3333 (2008).

13. T. Wheeldon, B. N. White, Genetic analysis of historic western Great Lakes region wolf samples reveals early *Canis lupus/lycaon* hybridization. *Biol. Lett.* **5**, 101–104 (2009).

14. L. Y. Rutledge, C. J. Garroway, K. M. Loveless, B. R. Patterson, Genetic differentiation of eastern wolves in Algonquin Park despite bridging gene flow between coyotes and grey wolves. *Heredity* **105**, 520–531 (2010).

15. C. J. Kyle, A. R. Johnson, B. R. Patterson, P. J. Wilson, K. Shami, S. K. Grewal, B. N. White, Genetic nature of eastern wolves: Past, present and future. *Conserv. Genet.* **7**, 273–287 (2006).

16. P. J. Wilson, S. K. Grewal, F. F. Mallory, B. N. White, Genetic characterization of hybrid wolves across Ontario. *J. Hered.* **100**, S80–S89 (2009).

17. Federal Registrar, C. F. R.: 50 CFR 17. Document Citation: 79 FR 7627. Document Number: 2014–02817 (2014).

18. R. J. Fredrickson, P. W. Hedrick, Dynamics of hybridization and introgression in red wolves and coyotes. *Conserv. Biol.* **20**, 1272–1283 (2006).

19. P. W. Hedrick, R. J. Fredrickson, Captive breeding and the reintroduction of Mexican and red wolves. *Mol. Ecol.* **17**, 344–350 (2008).

20. L. Y. Rutledge, P. J. Wilson, C. F. C. Klütsch, B. R. Patterson, B. N. White, Conservation genomics in perspective: A holistic approach to understanding Canis evolution in North America. *Biol. Conserv.* **155**, 186–192 (2012).

21. L. Y. Rutledge, S. Devillard, J. Q. Boone, P. A. Hohenlohe, B. N. White, RAD sequencing and genomic simulations resolve hybrid origins within North American *Canis. Biol. Lett.* **11**, 20150303 (2015).

22. P. W. Hedrick, P. S. Miller, E. Geffen, R. K. Wayne, Genetic evaluation of the three captive Mexican wolf lineages. *Zoo Biol.* **16**, 47–69 (1997).

23. J. D. Wall, M. P. Cox, F. L. Mendez, A. Woerner, T. Severson, M. F. Hammer, A novel DNA sequence database for analyzing human demographic history. *Genome Res.* **18**, 1354–1361 (2008).

24. K. Lindblad-Toh, C. M. Wade, T. S. Mikkelsen, E. K. Karlsson, D. B. Jaffe, M. Kamal, M. Clamp, J. L. Chang, E. J. Kulbokas III, M. C. Zody, E. Mauceli, X. Xie, M. Breen, R. K. Wayne, E. A. Ostrander, C. P. Ponting, F. Galibert, D. R. Smith, P. J. deJong, E. Kirkness, P. Alvarez, T. Biagi, W. Brockman, J. Butler, C.-W. Chin, A. Cook, J. Cuff, M. J. Daly, D. DeCaprio, S. Gnerre, M. Grabherr, M. Kellis, M. Kleber, C. Bardeleben, L. Goodstadt, A. Heger, C. Hitte, L. Kim, K.-P. Koepfli, H. G. Parker, J. P. Pollinger, S. M. J. Searle, N. B. Sutter, R. Thomas, C. Webber, J. Baldwin, A. Abebe, A. Abouelleil, L. Aftuck, M. Ait-zahra, T. Aldredge, N. Allen, P. An, S. Anderson, C. Antoine, H. Arachchi, A. Aslam, L. Ayotte, P. Bachantsang, A. Barry, T. Bayul, M. Benamara, A. Berlin, D. Bessette, B. Blitshteyn, T. Bloom, J. Blye, L. Boguslavskiy, C. Bonnet, B. Boukhgalter, A. Brown, P. Cahill, N. Calixte, J. Camarata, Y. Cheshatsang, J. Chu, M. Citroen, A. Collymore, C. Cooke, T. Dawoe, R. Daza, K. Decktor, S. DeGray, N. Dhargay, K. Dooley, K. Dooley, P. Dorje, K. Dorjee, L. Dorris, N. Duffey, A. Dupes, O. Egbiremolen, R. Elong, J. Falk, A. Farina, S. Faro, D. Ferguson, P. Ferreira, S. Fisher, M. FitzGerald, K. Foley, C. Foley, A. Franke, D. Friedrich, D. Gage, M. Garber, G. Gearin, G. Giannoukos, T. Goode, A. Goyette, J. Graham, E. Grandbois, K. Gyaltsen, N. Hafez, D. Hagopian, B. Hagos, J. Hall, C. Healy, R. Hegarty, T. Honan, A. Horn, N. Houde, L. Hughes, L. Hunnicutt, M. Husby, B. Jester, C. Jones, A. Kamat, B. Kanga, C. Kells, D. Khazanovich, A. C. Kieu, P. LeVine, J. Liu, X. Liu, Y. Lokyitsang, T. Lokyitsang, A. Lui, J. Macdonald, J. Major, R. Marabella, K. Maru, C. Matthews, S. McDonough, T. Mehta, J. Meldrim, A. Melnikov, L. Meneus, A. Mihalev, T. Mihova, K. Miller, R. Mittelman, V. Mlenga, L. Mulrain, G. Munson, A. Navidi, J. Naylor, T. Nguyen, N. Nguyen, C. Nguyen, T. Nguyen, R. Nicol, N. Norbu, C. Norbu, N. Novod, T. Nyima, P. Olandt, B. O'Neill, K. O'Neill, S. Osman, L. Oyono, C. Patti, D. Perrin, P. Phunkhang, F. Pierre, M. Priest, A. Rachupka, S. Raghuraman, R. Rameau, V. Ray, C. Raymond, F. Rege, C. Rise, J. Rogers, P. Rogov, J. Sahalie, S. Settipalli, T. Sharpe, T. Shea, M. Sheehan, N. Sherpa, J. Shi, D. Shih, J. Sloan, C. Smith, T. Sparrow, J. Stalker, N. Stange-Thomann, S. Stavropoulos, C. Stone, S. Stone, S. Sykes, P. Tchuinga, P. Tenzing, S. Tesfaye, D. Thoulutsang, Y. Thoulutsang, K. Topham, I. Topping, T. Tsamla, H. Vassiliev, V. Venkataraman, A. Vo, T. Wangchuk, T. Wangdi, M. Weiand, J. Wilkinson, A. Wilson, S. Yadav, S. Yang, X. Yang, G. Young, Q. Yu, J. Zainoun, L. Zembek, A. Zimmer; for Broad Sequencing Platform members, E. S. Lander, Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803–819 (2005).

25. M. M. Gray, J. M. Granka, C. D. Bustamante, N. B. Sutter, A. R. Boyko, L. Zhu, E. A. Ostrander, R. K. Wayne, Linkage disequilibrium and demographic history of wild and domestic canids. *Genetics* **181**, 1493–1505 (2009).

26. A. H. Freedman, I. Gronau, R. M. Schweizer, D. Ortega-Del Vecchyo, E. Han, P. M. Silva, M. Galaverni, Z. Fan, P. Marx, B. Lorente-Galdos, H. Beale, O. Ramirez, F. Hormozdiari, C. Alkan, C. Vilà, K. Squire, E. Geffen, J. Kusak, A. R. Boyko, H. G. Parker, C. Lee, V. Tadigotla, A. Siepel, C. D. Bustamante, T. T. Harkins, S. F. Nelson, E. A. Ostrander, T. Marques-Bonet, R. K. Wayne, J. Novembre, Genome sequencing highlights the dynamic early history of dogs. *PLOS Genet.* **10**, e1004016 (2014).

27. P. Skoglund, E. Ersmark, E. Palkopoulou, L. Dalén, Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. *Curr. Biol.* **25**, 1515–1519 (2015).

28. K.-P. Koepfli, J. Pollinger, R. Godinho, J. Robinson, A. Lea, S. Hendricks, R. M. Schweizer, O. Thalmann, P. Silva, Z. Fan, A. A. Yurchenko, P. Dobrynin, A. Makunin, J. A. Cahill, B. Shapiro, F. Álvares, J. C. Brito, E. Geffen, J. A. Leonard, K. M. Helgen, W. E. Johnson, S. J. O'Brien, B. Van Valkenburgh, R. K. Wayne, Genome-wide evidence reveals that African and Eurasian golden jackals are distinct species. *Curr. Biol.* **25**, 2158–2165 (2015).

29. J. A. Meachen, A. C. Janowicz, J. E. Avery, R. W. Sadleir, Ecological changes in coyotes (*Canis latrans*) in response to the Ice Age megafaunal extinctions. *PLOS One* **9**, e116041 (2014).

30. R. E. Green, J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M. H.-Y. Fritz, N. F. Hansen, E. Y. Durand, A.-S. Malaspinas, J. D. Jensen, T. Marques-Bonet, C. Alkan, K. Prüfer, M. Meyer, H. A. Burbano, J. M. Good, J. Schultz, A. Aximu-Petri, A. Butthof, B. Höber, B. Höffner, M. Siegemund, A. Weihmann, C. Nusbaum, E. S. Lander, C. Russ, N. Novod, J. Affourtit, M. Egholm, C. Verna, P. Rudan, D. Brajkovic, Ž. Kucan, I. Gušic, V. B. Doronichev, L. V. Golovanova, C. Lalueza-Fox, M. de la Rasilla, J. Fortea, A. Rosas, R. W. Schmitz, P. L. F. Johnson, E. E. Eichler, D. Falush, E. Birney, J. C. Mullikin, M. Slatkin, R. Nielsen, J. Kelso, M. Lachmann, D. Reich, S. Pääbo, A draft sequence of the Neanderthal genome. *Science* **328**, 710–722 (2010).

31. E. Y. Durand, N. Patterson, D. Reich, M. Slatkin, Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* **28**, 2239–2252 (2011).

32. J. A. Robinson, D. Ortega-Del Vecchyo, Z. Fan, B. Y. Kim, B. M. vonHoldt, C. D. Marsden, K. E. Lohmueller, R. K. Wayne, Genomic flatlining in the endangered island fox. *Curr. Biol.* **26**, 1183–1189 (2016).

33. Z. Fan, P. Silva, I. Gronau, S. Wang, A. S. Armero, R. M. Schweizer, O. Ramirez, J. Pollinger, M. Galaverni, D. Ortega Del-Vecchyo, L. Du, W. Zhang, Z. Zhang, J. Xing, C. Vilà, T. Marques-Bonet, R. Godinho, B. Yue, R. K. Wayne, Worldwide patterns of genomic variation and admixture in gray wolves. *Genome Res.* **26**, 163–173 (2016).

34. G. B. Kolenosky, R. O. Standfield, Morphological and ecological variation among gray wolves (*Canis lupus*) of Ontario, Canada, in The Wild Canids. Their Systematics, Behavioral Ecology and Evolution, M. W. Fox, Ed. (Van Nostrand Reinhold Co., New York, 1975), pp. 62–72.

35. I. Gronau, M. J. Hubisz, B. Gulko, C. G. Danko, A. Siepel, Bayesian inference of ancient human demography from individual genome sequences. *Nat. Genet.* **43**, 1031–1034 (2011).

36. J. Monzón, R. Kays, D. E. Dykhuizen, Assessment of coyote–wolf–dog admixture using ancestry-informative diagnostic SNPs. *Mol. Ecol.* **23**, 182–197 (2014).

37. H. Tang, M. Coram, P. Wang, X. Zhu, N. Risch, Reconstructing genetic ancestry blocks in admixed individuals. *Am. J. Hum. Genet.* **79**, 1–12 (2006).

38. H. Li, R. Durbin, Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).

39. M. G. Llorente, E. R. Jones, A. Eriksson, V. Siska, K. W. Author, J. W. Author, M. C. Curtis, J. T. Stock, M. Coltorti, P. Pieruccini, S. Stretton, F. Brock, T. Higham, Y. Park, M. Hofreiter, D. G. Bradley, J. Bhak, R. Pinhasi, A. Manica, Ancient Ethiopian genome reveals extensive Eurasian admixture in Eastern African. *Science* **350**, 820–822 (2015).

40. J. A. Leonard, R. K. Wayne, Native Great Lakes wolves were not restored. *Biol. Lett.* **4**, 95–98 (2008).

41. S. J. O'Brien, E. Mayr, Bureaucratic mischief: Recognizing endangered species and subspecies. *Science* **251**, 1187–1188 (1991).

42. G. M. Mace, The role of taxonomy in species conservation. *Philos. Trans. R. Soc. London Ser. B* **359**, 711–719 (2004).

43. N. J. B. Isaac, J. Mallet, G. M. Mace, Taxonomic inflation: Its influence on macroecology and conservation. *Trends Ecol. Evol.* **19**, 464–469 (2004).

44. K. A. Crandall, O. R. P. Bininda-Emonds, G. M. Mace, R. K. Wayne, Considering evolutionary processes in conservation biology. *Trends Ecol. Evol.* **15**, 290–295 (2000).

45. F. W. Allendorf, R. F. Leary, P. Spruell, J. K. Wenburg, The problems with hybrids: Setting conservation guidelines. *Trends Ecol. Evol.* **16**, 613–622 (2001).

46. M. Becker, N. Gruenheit, M. Steel, C. Voelckel, O. Deusch, P. B. Heenan, P. A. McLenachan, O. Kardailsky, J. W. Leigh, P. J. Lockhart, Hybridization may facilitate in situ survival of endemic species through periods of climate change. *Nat. Clim. Chang.* **3**, 1039–1043 (2013).

47. J. A. Hamilton, J. M. Miller, Adaptive introgression as a resource for management and genetic conservation in a changing climate. *Conserv. Biol.* **30**, 33–41 (2016).

48. B. M. Fitzpatrick, J. R. Johnson, D. K. Kump, J. J. Smith, S. R. Voss, H. B. Shaffer, Rapid spread of invasive genes into a threatened native species. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 3606–3610 (2010).

49. B. M. vonHoldt, D. R. Stahler, D. W. Smith, D. A. Earl, J. P. Pollinger, R. K. Wayne, The genealogy and genetic viability of reintroduced Yellowstone grey wolves. *Mol. Ecol.* **17**, 252–274 (2008).

50. B. M. vonHoldt, D. R. Stahler, E. E. Bangs, D. W. Smith, M. D. Jimenez, C. M. Mack, C. C. Niemeyer, J. P. Pollinger, R. K. Wayne, A novel assessment of population structure and gene flow in grey wolf populations of the Northern Rocky Mountains of the United States. *Mol. Ecol.* **19**, 4412–4427 (2010).

51. E. M. Gese, P. A. Terletzky, Using the "placeholder" concept to reduce genetic introgression of an endangered carnivore. *Biol. Conserv.* **192**, 11–19 (2015).

52. S. Piett, H. A. Hager, C. Gerrard, Characteristics for evaluating the conservation value of species hybrids. *Biodivers. Conserv.* **24**, 1931–1955 (2015).

53. R. K. Wayne, H. B. Shaffer, Hybridization and endangered species protection in the molecular era. *Mol. Ecol.* **25**, 2680–2689 (2016).

54. G. Lunter, M. Goodson, Stampy: A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res.* **21**, 936–939 (2011).

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

56. T. S. Korneliussen, A. Albechtsen, R. Nielsen, ANGSD: Analysis of next generation sequencing data. *BMC Bioinf.* **15**, 356 (2014).

57. R. R. Hudson, M. Slatkin, W. P. Maddison, Estimation of levels of gene flow from DNA sequence data. *Genetics* **132**, 583–589 (1992).

58. B. Giardine, C. Riemer, R. C. Hardison, R. Burhans, L. Elnitski, P. Shah, Y. Zhang, D. Blankenberg, I. Albert, J. Tayler, M. Miller, W. J. Kent, A. Nekrutenko, Galaxy: A platform for interactive large-scale genome analysis. *Genome Res.* **15**, 1451–1455 (2005).

59. D. Blankenberg, G. Von Kuster, N. Coraor G. Ananda, R. Lazarus, M. Mangan, A. Nekrutenko, J. Tayler, Galaxy: A web-based genome analysis tool for experimentalists. *Curr. Protoc. Mol. Biol.* **89**, 19.10.1–19.10.21 (2010).

60. J. Goecks, A. Nekrutenko, J. Taylor; The Galaxy Team, Galaxy: A comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* **11**, R86 (2010).

61. K. Prüfer, U. Stenzel, M. Hofreiter, S. Pääbo, J. Kelso, R. E. Green, Computational challenges in the analysis of ancient DNA. *Genome Biol.* **11**, R47 (2010).

62. H. R. Künsch, The jackknife and the bootstrap for general stationary observations. *Ann. Stat.* **17**, 1217–1241 (1989).

63. C. D. Marsden, D. Ortega-Del Vecchyo, D. P. O'Brien, J. F. Taylor, O. Ramirez, C. Vilà, T. Marques-Bonet, R. D. Schnabel, R. K. Wayne, K. E. Lohmueller, Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 152–157 (2016).

The following resources related to this article are available online at
http://advances.sciencemag.org. (This information is current as of August 15, 2016):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://advances.sciencemag.org/content/2/7/e1501714.full

Supporting Online Material can be found at:
http://advances.sciencemag.org/content/suppl/2016/07/25/2.7.e1501714.DC1

This article **cites 60 articles,** 21 of which you can access for free at:
http://advances.sciencemag.org/content/2/7/e1501714#BIBL