Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids

Bridgett M. vonHoldt, John P. Pollinger, Dent A. Earl, et al.

Genome Res. published online May 12, 2011
Access the most recent version at doi:10.1101/gr.116301.110

Supplemental Material	**http://genome.cshlp.org/content/suppl/2011/04/08/gr.116301.110.DC1.html**
P<P	Published online May 12, 2011 in advance of the print journal.
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *Genome Research* go to:
http://genome.cshlp.org/subscriptions

Copyright © 2011 by Cold Spring Harbor Laboratory Press

Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids

Bridgett M. vonHoldt,[1] John P. Pollinger,[1] Dent A. Earl,[2] James C. Knowles,[1] Adam R. Boyko,[3] Heidi Parker,[4] Eli Geffen,[5] Malgorzata Pilot,[6] Wlodzimierz Jedrzejewski,[7] Bogumila Jedrzejewska,[7] Vadim Sidorovich,[7] Claudia Greco,[8] Ettore Randi,[8] Marco Musiani,[9] Roland Kays,[10] Carlos D. Bustamante,[3] Elaine A. Ostrander,[4] John Novembre,[1] and Robert K. Wayne[1,11]

[1]*Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California 90095, USA;* [2]*Department of Biomolecular Engineering, University of California, Santa Cruz, California 95064, USA;* [3]*Department of Genetics, Stanford School of Medicine, Stanford, California 94305, USA;* [4]*Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA;* [5]*Department of Zoology, Tel Aviv University, Tel Aviv 69978, Israel;* [6]*Museum and Institute of Zoology, Polish Academy of Sciences, 00-679 Warszawa, Poland;* [7]*Mammal Research Institute, Polish Academy of Sciences, 17-230 Bialowieza, Poland;* [8]*Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), 40064 Ozzano Emilia (BO), Italy;* [9]*Faculty of Environmental Design, University of Calgary, Calgary, Alberta T2N 1N4, Canada;* [10]*New York State Museum, CEC 3140, Albany, New York 12230, USA*

High-throughput genotyping technologies developed for model species can potentially increase the resolution of demographic history and ancestry in wild relatives. We use a SNP genotyping microarray developed for the domestic dog to assay variation in over 48K loci in wolf-like species worldwide. Despite the high mobility of these large carnivores, we find distinct hierarchical population units within gray wolves and coyotes that correspond with geographic and ecologic differences among populations. Further, we test controversial theories about the ancestry of the Great Lakes wolf and red wolf using an analysis of haplotype blocks across all 38 canid autosomes. We find that these enigmatic canids are highly admixed varieties derived from gray wolves and coyotes, respectively. This divergent genomic history suggests that they do not have a shared recent ancestry as proposed by previous researchers. Interspecific hybridization, as well as the process of evolutionary divergence, may be responsible for the observed phenotypic distinction of both forms. Such admixture complicates decisions regarding endangered species restoration and protection.

[Supplemental material is available for this article. The genotyping data are available at http://genome-mirror.bscb.cornell.edu/cgi-bin/hgGateway (see "SNPs" track under the Variations and Repeats heading).]

High-density single nucleotide polymorphism (SNP) genotyping arrays developed from domestic species can potentially enhance our understanding of population history and relationships of their close relatives. We use the Affymetrix Canine SNP Genome Mapping Array (version 2) to assess long-standing questions about diversification and admixture of charismatic wolf-like canids, including the gray wolf (*Canis lupus*), red wolf (*C. rufus*), Great Lakes wolf (*C. lycaon* or *C. lupus lycaon*), and coyote (*C. latrans*). These species are characterized by high mobility and weak patterns of intraspecific differentiation (e.g., Wayne et al. 1992; Roy et al. 1994, 1996; Forbes and Boyd 1997; Vilà et al. 1999). Similarly, large dispersal distances have led to the formation of extensive admixture zones in North America, where four morphologically distinguishable wolf-like canids can potentially interbreed: the gray wolf of Old World derivation, the coyote and red wolf (both of which originated in North America), and the Great Lakes wolf. The latter two taxa are of controversial ancestry and species status and readily hybridize with other wolf-like canids (Supplemental Table S1). Their evolutionary origin has been explained either as a consequence of admixture between coyotes and varieties of the gray wolf, or as parallel evolution of a wolf-like phenotype independently in the New World from a common coyote-like ancestor (Fig. 1; Supplemental Table S1). Origin through ancient hybridization or an independent New World evolution might warrant greater preservation efforts and legal protection, whereas origin through recent hybridization would suggest a dynamic evolutionary zone of questionable conservation status, although the ecological significance of such hybrids should also be considered (Crandall et al. 2000; Kyle et al. 2006; Leonard and Wayne 2008). Specifically, the red wolf is protected as a distinct endangered species under the US Endangered Species Act (ESA) and wildlife management agencies dedicate considerable resources to study, monitor, and protect the red wolf (Phillips et al. 2003).

The vast majority of genetic studies on wild populations have utilized a small number of genetic markers for evolutionary inference. To better resolve population structure and admixture within wolf-like canids, we assayed 48,036 SNPs in a panel of 208 gray wolves (*C. lupus*) representing their worldwide distribution (Eurasia and North America), 57 coyotes and 12 red wolves (Fig. 2; Supplemental Table S2). To our knowledge, this represents the most extensive SNP survey of any wild vertebrate group and provides

[11]Corresponding author.
E-mail rwayne@ucla.edu.
Article published online before print. Article, supplemental material, and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.116301.110.

Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

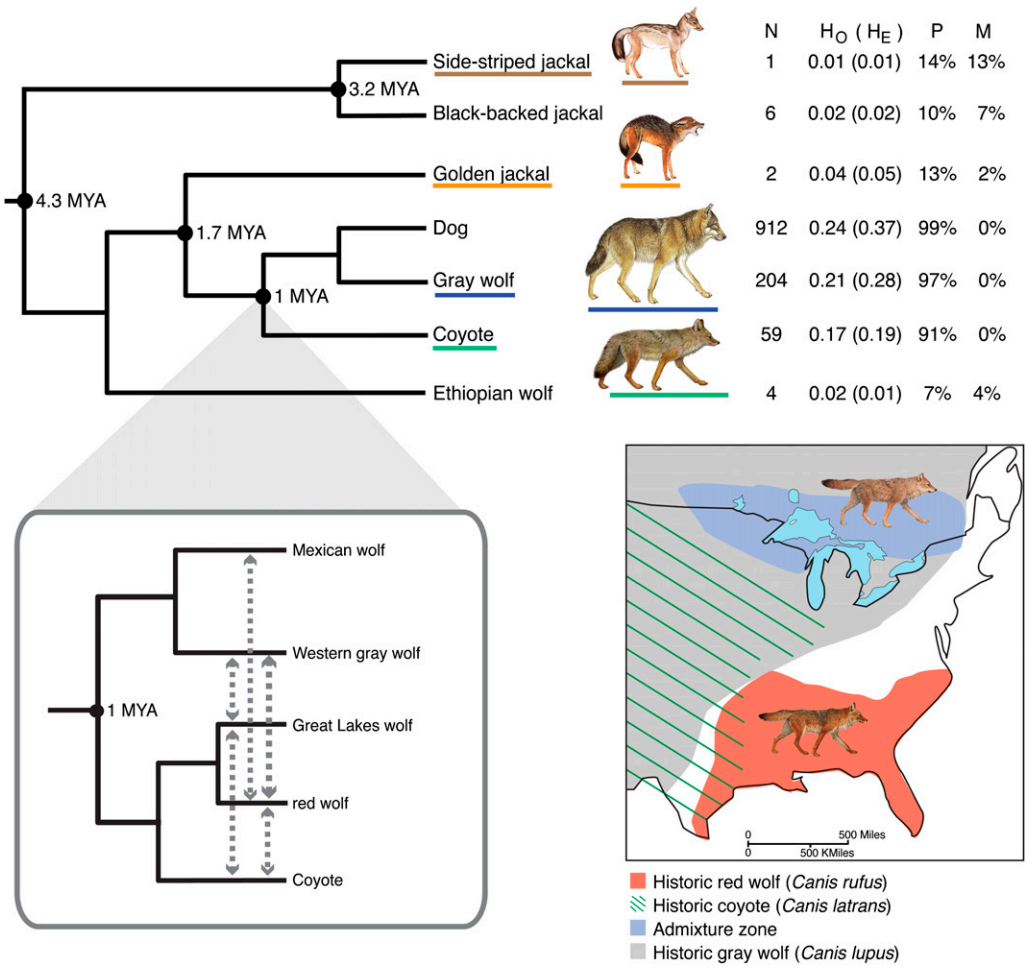

Figure 1. Genetic variation of wolf-like canids in the 48K SNP data set. Sample size (N), average observed and expected heterozygosity (H_O, H_E), percent of loci that are polymorphic (P), and could not be scored (missingness, M) are shown. Except for dog (SE = 0.01), all standard errors for (H_O, H_E) are <0.001. A modified relationship tree (Lindblad-Toh et al. 2005) and divergence dates are presented (Perini et al. 2009). (*Inset*) Tree depicts admixture hypotheses with dashed lines among North American canid populations (Supplemental Table S1). Map depicts historic species distributions and admixture zone (data from Nowak 2009; Rutledge et al. 2010b).

an important precedent for testing the power of inferences about population subdivision, admixture, and species origin based on previous more limited genetic surveys (Supplemental Table S1). Our analysis reveals extensive population subdivision despite the high mobility of wolf-like canids. Further, we find strong signals of admixture with coyotes and gray wolves in the genomes of the red wolf and Great Lakes wolf, respectively. This finding implies that these taxa do not represent separately evolving lineages, and suggest that admixture has led to substantial phenotypic variation. Our results demonstrate how genomic tools developed for model species can enable new insights into the evolutionary history of wild relatives.

Results

Utility of the canine array to assess variability in wolf-like canids

The canine genotyping array interrogates variation in SNPs primarily ascertained from comparisons of boxer and poodle genomes, with additional comparisons to sequences from nine other dog breeds, four gray wolves, and a coyote (Lindblad-Toh et al. 2005; vonHoldt et al. 2010). Less than 2% of the SNPs on the array were ascertained from comparison of genomic sequences from a domestic dog and a wild canid (vonHoldt et al. 2010). Hence, we predict an ascertainment bias of increasing monomorphism with genetic divergence from the discovery panel (e.g., Nielsen and Signorovitch 2003; Conrad et al. 2006; Rosenblum and Novembre 2007). To assess the severity of this predicted bias, we genotyped species from the wolf-like clade, spanning divergence dates to ~4 million years ago (Fig. 1; Lindblad-Toh et al. 2005; Perini et al. 2009). We found variation to decrease rapidly with phylogenetic distance from the domestic dog. The dog and its closest relatives, the gray wolf and coyote, had high levels of genomic variation (Fig. 1), but species having more than about one million years divergence from the domestic dog had little to no shared variation (e.g., black-backed and golden jackals) (Fig. 1). Given the decrease in polymorphism and increase in missingness with phylogenetic distance (Fig. 1), both SNP ascertainment as well as the failure of probe-target hybridization likely accounts for the limited utility of the array in distantly related species. Nonetheless, the high levels of segregating genotyped variants in gray wolves and coyotes support the use of the canine SNP array for quantifying genetic diversity in both of these species.

Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

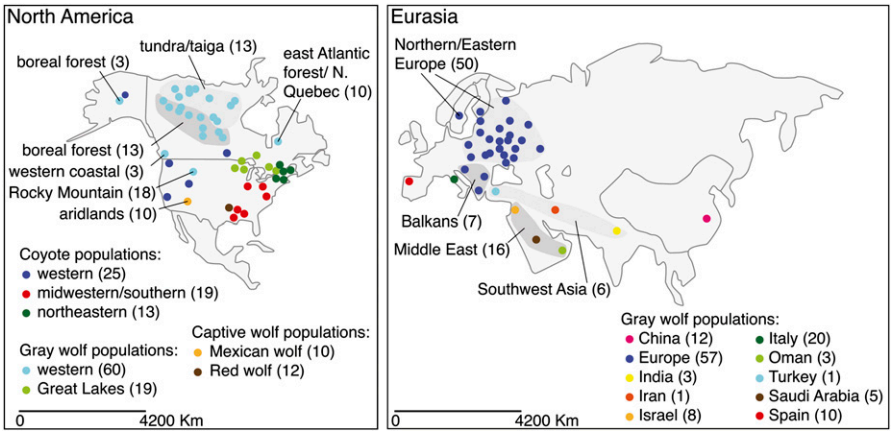

Figure 2. Distribution map of gray wolf and coyote sample locations and regions. Sample sizes are indicated in parentheses.

We found considerable genome-wide SNP variation at the population level in gray wolves and coyotes (Table 1). As predicted, large, and/or expanding populations had high levels of SNP variability (e.g., Canadian and Western wolves, H_E = 0.24–0.29), including those that are currently hunted (e.g., Russia and Poland/Belarus wolves, H_E = 0.25–0.26) (Table 1). Overall, the lowest observed heterozygosity was found in captive red and Mexican wolves and in historically bottlenecked populations (e.g., Spain, Italy, red, and Mexican wolves; H_E = 0.16–0.18). These results further support the use of SNP data to infer demographic history of wild canids (e.g., Gray et al. 2009).

Global patterns of population subdivision

We found principal component analysis (PCA) of individual SNPs clearly discriminates domestic and wolf-like canids on the first two PC axes (Fig. 3; Supplemental Fig. S1). PC1 represents a wild versus domestic canid axis, whereas PC2 separates wolves (n = 198) and dogs (n = 912) from coyotes (n = 57) and red wolves (n = 10). Our result shows dogs and gray wolves are genetically distinct (F_{ST} = 0.165). PC2 in this analysis and in a subset of the data demonstrate a geographically based population hierarchy within gray wolves and coyotes (Fig. 3; Supplemental Figs. S2, S3). First, within gray wolves, Old and New World wolves are distinct. Second, regional geographic groupings in the Old World are evident, such

Table 1. Average observed (H_O) and expected (H_E) heterozygosity for 48K SNPs in populations of differing demographic histories (n, sample size)

Demographic history	Population	n	H_O (H_E)	References
Old World gray wolf				
Large (recent) expanding population	Europe[a]	57	0.24 (0.26)	Pilot et al. 2006
Historic population bottleneck	Italy	20	0.15 (0.17)	Gray et al. 2009; Fabbri et al. 2007
	Poland and Belarus	15	0.24 (0.25)	Jedrzejewski et al. 2005; Aspi et al. 2009
Recent population bottleneck and subsequent expansion with continual hunting pressures	Russia	18	0.25 (0.26)	Jedrzejewski et al. 2005; Aspi et al. 2009
Recent population bottleneck	Spain	10	0.18 (0.17)	Gray et al. 2009; Ramirez et al. 2006
North American wolf				
Founding from a large source population	Yellowstone National Park	18	0.22 (0.22)	Gray et al. 2009; vonHoldt et al. 2008
Large constant population size	Canada	13	0.22 (0.24)	Gray et al. 2009; Carmichael et al. 2007; Musiani et al. 2007
Large (recent) expanding population	Western	60	0.21 (0.29)	Roy et al. 1994
Recent population bottleneck and subsequent expansion	Minnesota and Southern Quebec	12	0.19 (0.22)	Gray et al. 2009
Recent range expansion and potential hybridization	Great Lakes[b]	23	0.18 (0.21)	Roy et al. 1994; Kyle et al. 2006; Koblmüller et al. 2009
Recent population bottleneck with managed breeding; possible hybrid-species origin	Red wolf	12	0.16 (0.16)	Roy et al. 1994
Recent population bottleneck with managed breeding	Mexican wolf	10	0.12 (0.18)	Hedrick et al. 1997
Coyote				
Large (recent) expanding population	Western	25	0.14 (0.18)	Roy et al. 1994
	Northeastern	13	0.20 (0.25)	Roy et al. 1994; Kyle et al. 2006; Koblmüller et al. 2009
Recent range expansion and potential hybridization	Midwestern/Southern	19	0.18 (0.24)	Roy et al. 1994; Kyle et al. 2006; Koblmüller et al. 2009

All estimates have a standard error less than ±0.001.
[a]Excludes Italian and Spanish wolves.
[b]Excludes Minnesota and Quebec wolves.

Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

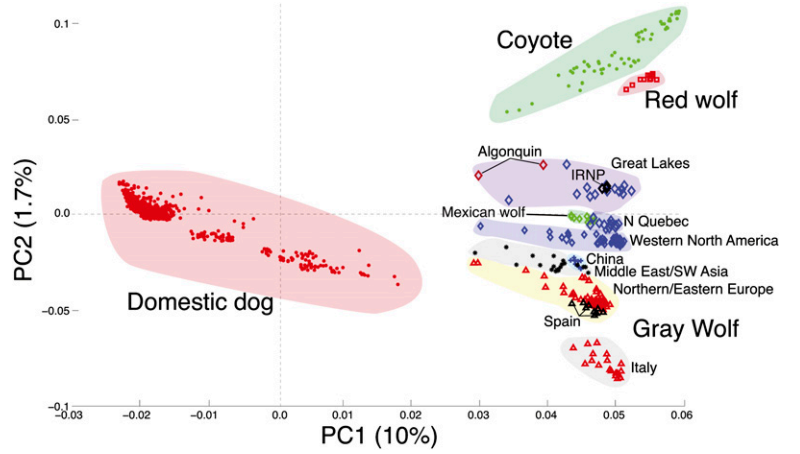

Figure 3. Principal component analysis of all wolf-like canids for the 48K SNP data set (IRNP, Isle Royale National Park).

as Italy, Spain, the Middle East, and China. Third, groupings in the New World, such as Mexican, Great Lakes, Northern Quebec, and Western North America wolves appear distinct from wolves elsewhere (Fig. 3; Supplemental Fig. S2A–C). Italian and Spanish wolf populations are resolved on multiple PC axes, as well as wolves from the Middle East, China, Mexico, and the British Columbian coast (Supplemental Fig. S3A–C). In contrast, coyotes appear divided roughly into only three primary groupings from the American West, Midwest/South, and Northeast (Supplemental Figs. S2D, S3D–F). Finally, we examined apparent partitions for a reduced set of 710 SNPs ascertained by comparison of only dog and wolf or coyote sequences (Supplemental Fig. S4). Although lower in resolution, this analysis finds similar groupings of wild and domestic canids, suggesting that the specific ascertainment scheme does not strongly affect the overall pattern of clustering, a conclusion consistent with previous SNP and haplotype clustering analyses (vonHoldt et al. 2010).

Bayesian ancestry inference using the program *structure* resembles the PCA results, with dogs and wolf-like canids forming the first partitions (K = 2), followed by coyotes and gray wolves (K = 3), Old and New World wolves (K = 4), Italian wolves (K = 5), Mexican wolves (K = 6), Spanish wolves (K = 7), Middle Eastern wolves (K = 8), red wolves (K = 9), and Great Lakes wolves (K = 10) (Fig. 4). However, the red and Great Lakes wolves show consistent signals of admixture with coyotes (light green) (Fig. 4). Midwestern/Southern and Northeastern coyotes show a small degree of admixture with gray wolves (blue) and also with dogs (red) (Fig. 4). All probability intervals for individuals from putatively admixed populations indicated that the admixed assignments were meaningful (90% PI interval >0).

To assess possible population subdivision of nonadmixed populations, we purged our data set of all wolf populations showing evidence of interspecific admixture in *structure*, and analyzed Old

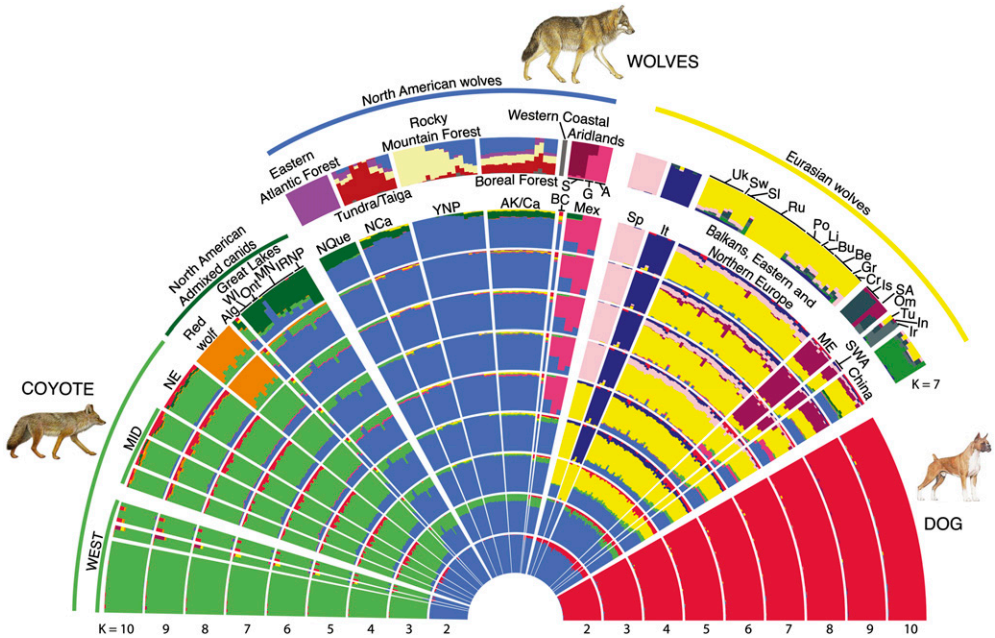

Figure 4. *structure* clustering analysis of domestic and wild canids for the 44K SNP pruned data set (see Supplemental Table S2). Separate analyses of North American (n = 68) and Eurasian (n = 74) wolves further resolves biologically informative clusters (K = 7; the *top* layer in plot). The three captive Mexican wolf colonies (Studbook, S; Ghost Ranch, G; and Aragon, A) are shown. See discussion in the Supplemental Material for an explanation of how K values were chosen for presentation. Coyote abbreviations: Midwest/Southern, MID; Northeastern, NE; and Western, WEST. North American abbreviations: Algonquin, Alg; Isle Royale NP, IRNP; Minnesota, MN; Ontario, Ont; Wisconsin, WI; Alaska/Canada, AK/Ca; British Columbia, BC; Mexico, Mex; Northern Canada, NCa; Northern Quebec, NQue; and Yellowstone, YNP. Eurasian abbreviations: Belarus, Be; Bulgaria, Bu; Croatia, Cr; Greece, Gr; India, In; Iran, Ir; Israel, Is; Italy, It; Lithuania, Li; Middle East, ME; Oman, Om; Poland, Po; Russia, Ru; Saudi Arabia, SA; Slovakia, Sl; Southwest Asia, SWA; Spain, Sp; Sweden, Sw; Turkey, Tu; and Ukraine, Uk.

Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

and New World wolves separately. In the Old World analysis, K = 7 resulted in distinct *structure* profiles of wolves from Spain, Italy, Eastern/Northern Europe, Southwest Asia and China, and an admixed grouping (two dominant colors) from the Middle East (Fig. 4). In the New World wolf analysis, at K = 7, differences in *structure* profiles suggest regional and habitat groupings of Northern Quebec (Eastern Atlantic forest), Northern Canada near and above the treeline (taiga and tundra), Yellowstone National Park (Rocky Mountain forest), Alaskan and Canadian Northern forest (boreal forest), British Columbian coast (Western Coastal forest), and Mexico (aridlands) (Fig. 4; Supplemental Fig. S7). Mexican wolves are further divided into Studbook and Aragon lineages with the Ghost Ranch lineage appearing as admixed. Similarly, the mixed profiles of Rocky Mountain and boreal forest wolves suggest that they are derived from heterogeneous migrants or are currently experiencing gene flow from other populations. Most of these regional subdivisions are also differentiated in F_{ST} and AMOVA analyses (Supplemental Tables S3, S4).

Finally, to examine topological relationships among population groupings, we constructed neighbor-joining trees based on genome-wide allele sharing (Supplemental Fig. S5). We excluded wolf populations for which *structure* analysis suggested substantial admixture with coyotes (see above), as admixture would be expected to cause reticulation in relationship trees (Posada and Crandall 2001; Susnik et al. 2007; Addison and Pogson 2009). Rooted with the coyote, Old and New World populations are the first division in the gray wolf tree (Supplemental Fig. S5). In the Old World, the Chinese population is the first branch (red), followed by groupings of wolves from the Middle East and Southwest Asia (blue), and then a group containing European wolves (green and magenta, Supplemental Fig. S5). Within Europe, Northeastern and Southern Europe define separate groupings, and in the latter Bulgaria/Croatia/Greece define a Balkans grouping, whereas Italy and Spain define sister taxa (magenta and orange). In the New World, Mexican wolves are most basal, followed by British Columbian coastal forest wolves, then the Northern Quebec Atlantic forest and tundra/taiga wolves, and finally a mixed grouping of boreal and Rocky Mountain forests wolves. These locality and habitat associations largely support ecotype designations based on previous microsatellite analyses (Carmichael et al. 2007; Musiani et al. 2007; Koblmüller et al. 2009; Muñoz-Fuentes et al. 2009). However, in North American wolves, bootstrap support for boreal forest and Rocky Mountain forest groupings is low, and define a somewhat mixed cluster (Supplemental Fig. S5). For coyotes, the Northeastern populations are the most divergent grouping (green), followed by clusters of Midwestern/Southern coyotes (red), and lastly, groupings of Western coyotes, for which the easternmost population in Manitoba is most divergent (blue, Supplemental Fig. S5).

A notable feature of all trees is the high concordance with which individuals cluster to their population of origin. Such precise assignment in a genetic similarity tree was not anticipated given the dispersal ability of wolves and coyotes, and likely reflects the heightened power of genome-wide studies to detect a shared genealogical history within single populations (Jakobsson et al. 2008; Li et al. 2008).

Relationships and admixture of enigmatic wolf-like canids

The evolutionary origin and relationships of the red wolf and Great Lakes wolf are controversial and clouded by admixture with related species (Supplemental Table S1). Our PCA and *structure* analyses suggest that the red wolf and Great Lakes wolf are genetically differentiated and are unlikely to share a common origin (F_{ST} = 0.11; Supplemental Table S3). In a plot of PC1 and PC2, red wolves are genetically similar to coyotes, and on PC2, Great Lakes wolves and Mexican wolves are more similar to North American gray wolves (Fig. 3). *structure* analysis consistently assigned ~80% of the red wolf genome to the coyote (light green) (Fig. 4), whereas about the same fraction of the Great Lakes wolf genome is assigned to the gray wolf (blue). These contrasting patterns imply that red wolves are predominantly of coyote ancestry, possibly with limited historic hybridization with gray wolves. Moreover, assignments are more variable in Great Lakes wolves, ranging from about 50% to 100% gray wolf ancestry (Fig. 4), suggesting a more heterogeneous process of admixture, with some individuals largely free of coyote ancestry, especially in more western populations (Koblmüller et al. 2009; Rutledge et al. 2010a). Notably, the distinctive wolves of Algonquin Provincial Park in central Ontario that have the archetypical C1 mtDNA haplotype thought to be representative of the Great Lakes wolf taxon (Wilson et al. 2000; Kyle et al. 2006) have the largest proportion of their genome assigned to coyotes. Finally, in the full analysis, at K = 9 and K = 10, the red wolf (orange) and Great Lakes wolf (dark green), respectively, show a distinct genetic signature that appears analogous to subspecific partitions such as the Mexican wolf (Fig. 4) or other distinct regional populations of wolves (Supplemental Figs. S1–S3).

Analysis of linkage disequilibrium and autozygosity segments

Linkage disequilibrium (LD) is expected to be greater in inbred and admixed populations (Pritchard and Przeworski 2001; Gaut and Long 2003; Tang et al. 2006; Gray et al. 2009). To assess LD patterns, we estimated $r^2_{0.5}$, the physical distance at which the pairwise genotypic association (r^2) in the 48K SNP data set decays below a threshold of 0.5 (Fig. 5). We found that populations of North American Western wolves (n = 43) and coyotes (n = 25) had low levels of LD ($r^2_{0.5}$ < 2 Kb) as expected for outbred, non-admixed populations. These LD values are comparable to previous estimates of LD derived from targeted sequencing (r^2 = 0.5 at <10 Kb) (Gray et al. 2009). In contrast, LD was greatest in the Mexican wolf (n = 4;

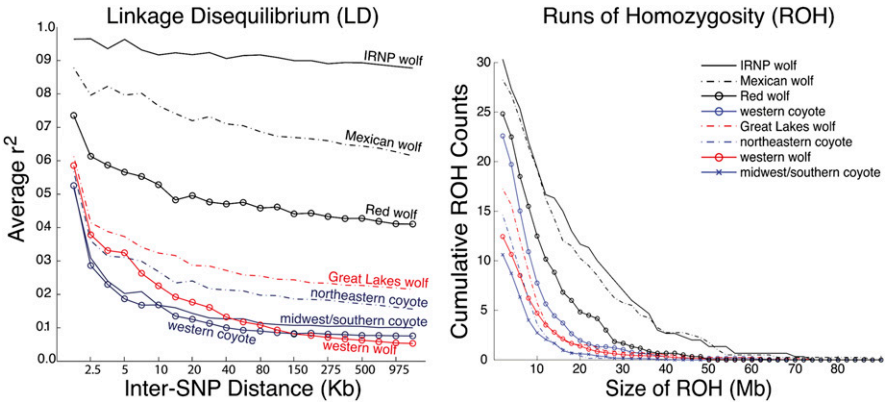

Figure 5. (*Left*) Extent of LD (genotypic association, r^2) as a function of inter-SNP distance (Kb) for all North American canid populations. (*Right*) Autozygosity frequency distribution of runs of homozygosity (ROH) for all North American canid populations.

Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

$r^2_{0.5}$ > 1 Mb, average inbreeding coefficient (*f*) = 0.19) (Kalinowski et al. 1999) and the Isle Royale National Park wolf (*n* = 3, $r^2_{0.5}$ > 1 Mb), consistent with known founding bottlenecks and subsequent inbreeding (Gray et al. 2009). Likewise, the red wolf has elevated levels of LD (*n* = 6, $r^2_{0.5}$ ≈ 10–20 Kb) (Fig. 5), reflecting a founding bottleneck and increased inbreeding in the captive population (average *f* = 0.04) (Kalinowski et al. 1999). Large non-inbred populations of wolves in the Great Lakes region have slightly lower LD (*n* = 10, $r^2_{0.5}$ ≈ 1–2 Kb) followed by the Northeastern (*n* = 13) and Midwestern/Southern coyotes (*n* = 19, $r^2_{0.5}$ < 2 Kb) (Fig. 5). Finally, LD estimates may be sensitive to sample size (Teare et al. 2002; England et al. 2006) and, consequently, we experimented with different sample sizes ranging from three to 10 individuals (Supplemental Fig. S6). The results suggest that the Great Lakes wolves have consistently inflated LD, especially long-range LD. However, the rank order of LD in inbred populations such as Isle Royale National Park wolves and the Mexican wolf are not consistent.

To separate the effects of inbreeding and admixture on LD, we examined genome-wide autozygosity, which more directly measures recent inbreeding (Boyko et al. 2010). We assessed autozygosity as runs of homozygosity (ROH), and expected proportionally longer ROH in recently inbred populations and the converse in more ancestrally inbred populations (Boyko et al. 2010). The Isle Royale National Park and Mexican wolves had the highest fraction of autozygous segments across all fragment sizes, as well as the highest $r^2_{0.5}$ (Fig. 5). The red wolf had elevated ROH relative to the outbred gray wolf and coyote populations; however, the curve was uniformly lower than the extremely inbred Mexican and Isle Royale National Park wolves (Fig. 5). Specifically, ROH was elevated for the Mexican and Isle Royale National Park wolves relative to outbred populations throughout the ROH curve, suggesting ancient and recent inbreeding. In contrast, Great Lakes wolves (*n* = 10) have low ROH, but slightly inflated linkage disequilibrium, suggesting that recent admixture rather than past inbreeding explains higher levels of LD (Fig. 5).

Ancestry blocks and assignments

We used the program SABER to assign ancestry of chromosomal blocks based on two or three putative reference ancestral populations. In Great Lakes wolves, red wolves, and Algonquin wolves, we utilized 12 Western gray wolves and 12 Western coyotes as the reference ancestral populations. Because significant dog ancestry was identified in Midwestern/Southern and Northeastern coyotes by *structure* analyses, we included 12 modern domestic dogs as a third reference ancestor for analysis of these populations (Figs. 3, 6). For each ancestral individual, all ancestry blocks were correctly assigned to the appropriate parental population (e.g., Fig. 6A,B,F; Supplemental Tables S5, S6). In contrast, we found Great Lakes wolves (*n* = 18) to have, on average, 14.9% ± 0.3 of their genome assigned to coyotes (range: 13.4%–19.5%, Supplemental Table S5). In contrast, red wolves (*n* = 12) have, on average, 76.1% ± 0.3 of

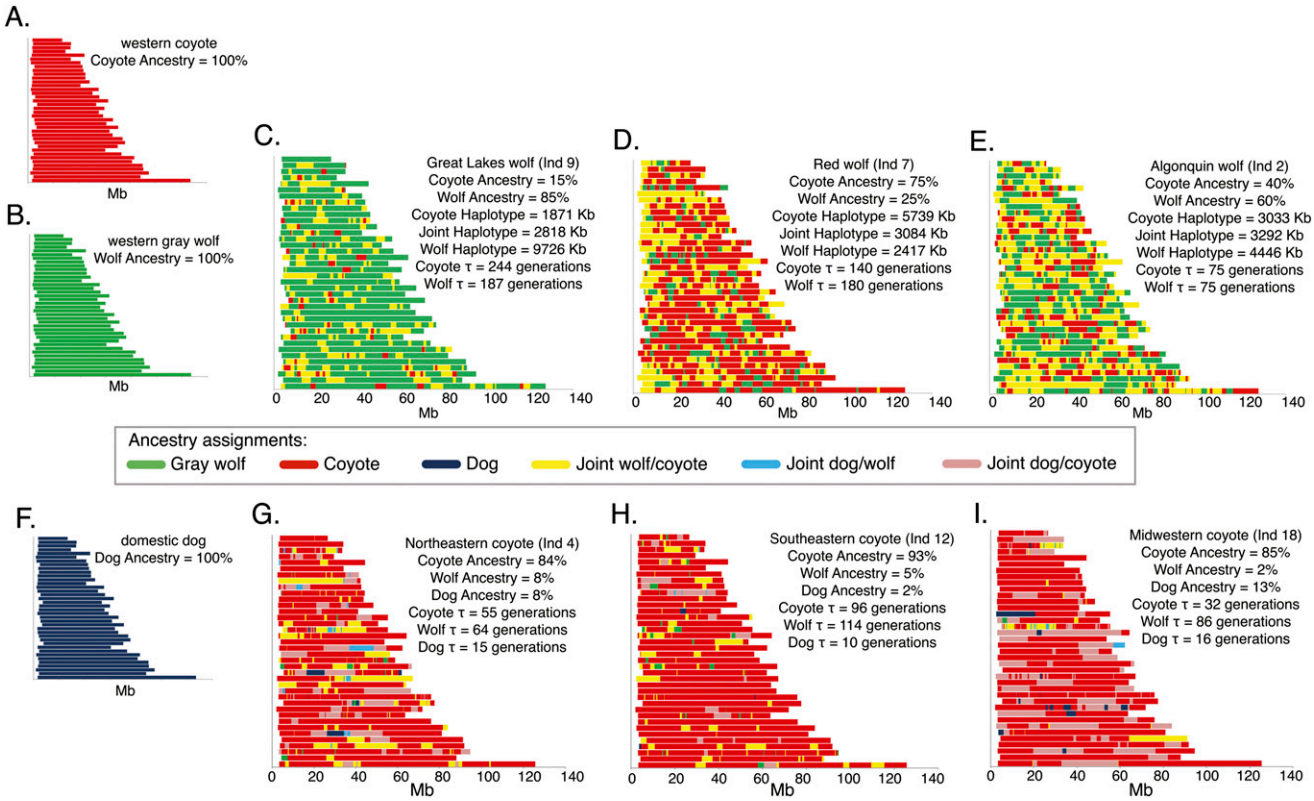

Figure 6. Ancestry analysis of 38 canine autosomes. Plots show ancestry blocks and their assignments for representative individuals of canid populations with average size of blocks, percent ancestry, and number of generations since most recent admixture (τ) indicated. Two-ancestor (coyote, *A*; gray wolf, *B*) analyses are presented for a Great Lakes wolf from Minnesota (*C*), a captive red wolf (*D*), and an Algonquin wolf (*E*). Three-ancestor analyses (coyote, *A*; gray wolf, *B*; dog, *F*) are presented for a Northeastern coyote from Vermont (*G*), a Southern coyote from Louisiana (*H*), and a Midwestern coyote from Ohio (*I*). Individual results are given in Supplemental Tables S5 and S6.

Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

their genomic composition assigned to coyotes (range: 74.3%–78.1%; Supplemental Table S5). These results suggest that the two entities have distinct genomic compositions and ancestries (e.g., Fig. 6C,D). Finally, the phenotypically distinct Algonquin wolf (n = 2), suggested as the last remaining population retaining a substantial component of the Eastern wolf genome (Rutledge et al. 2010a), had a highly admixed genome with an average of 58.1% ± 2.2 derived from gray wolves and the remainder from coyotes (e.g., Fig. 6E; Supplemental Table S5).

The influence of ancestry and admixture is also evident in Northeastern (n = 13) and Midwestern/Southern (n = 19) coyotes. In the former, 8.7% ± 0.6 (range: 6.4%–13.1%) of segments are assigned to gray wolves, suggesting a limited impact of hybridization with Great Lakes wolves as coyotes recently invaded this region (Fig. 6; Wilson et al. 2009; Kays et al. 2010; Supplemental Table S6). In the latter, although gray wolves have been absent from the Midwestern/Southern regions of the US for 100 yr or more (Nowak 1979), we also detected gray wolf ancestry, specifically, on average, 4.1% ± 0.4 (range: 2.1%–6.4%) of the Southern coyote (Alabama, Louisiana, and Mississippi) and 1.4% ± 0.2 (range: 0.3%–2.9%) of the Midwestern (Illinois, Ohio, and Virginia) coyote genome is assigned to gray wolf (Fig. 6H,I; Supplemental Table S6). These two coyote populations also have assignments to dog ancestry, as the Northeastern coyote has on average 9.1% ± 0.7 (range: 5.2%–12.8%) dog ancestry and Midwestern/Southern population has 4.4% ± 0.6 (range: 1.9%–8.1%). The highest dog ancestry was identified in Ohio and Virginia (16.7% and 16.9%, respectively; Supplemental Table S6). Moreover, the considerable range in chromosome assignments across Midwestern/Southern coyotes suggests a varied history of hybridization (Supplemental Table S6).

Timing of initial admixture

We used SABER to estimate the number of generations since admixture (τ) (Tang et al. 2006). SABER estimates τ as a function of the inverse of the average inferred chromosomal block size across all assigned blocks. We estimate 297 ± 24 generations (τ) since the initial admixture events between Great Lakes wolves and coyotes, which is equivalent to 546 to 963 yr ago assuming a 2–3-yr generation time (Supplemental Table S5). This ancient admixture suggests that hybridization events began prior to the recent invasion of coyotes into the Great Lakes area (Wilson et al. 2009; Kays et al. 2010). In the red wolf, the initiation of admixture with coyotes began about 144 ± 5 generations or 287–430 yr ago (Supplemental Table S5). In the Great Lakes wolves, segments assigned to wolf ancestry are nearly five times larger and more variable (8016 ± 135 Kb) than those assigned to coyotes (1735 ± 35 Kb) (Supplemental Fig. S8; Supplemental Table S5). This difference suggests historic admixture between the two species, followed by a predominance of backcrossing to gray wolf populations. Similarly, coyote ancestry block size is 2.6 times larger and more variable (6198 ± 115 Kb) than gray wolf blocks in the red wolf genome (2397 ± 49 Kb) (Supplemental Fig. S8A; Supplemental Table S5) and implies extensive backcrossing to coyotes after the initial admixture with gray wolves.

In Midwestern/Southern coyotes (n = 19), blocks of gray wolf and dog ancestry are small (874 ± 201 Kb and 2621 ± 412 Kb, respectively) and coyote ancestry blocks are much larger (18,837 ± 1501 Kb) (Supplemental Fig. S8). This disparity is consistent with a model of very limited admixture with gray wolves and dogs, followed by extensive backcrossing to coyotes. An average τ of 140 generations for wolf ancestry implies effective admixture with wolves occurring around 280 yr ago, assuming a generation time for coyotes of 2 yr (Bekoff and Wells 1986). However, individual values show considerable variability in τ (range: 39.0–595.6 generations), suggesting a varied history involving differing degrees of admixture and backcrossing at different times (Supplemental Table S6). In contrast, an average τ of 14.6 (range: 7.2–39.4) generations for dog ancestry implies much more recent admixture with dogs. Northeastern coyotes show similar average wolf and dog block sizes of 1110 ± 78 Kb and 2397 ± 373 Kb, respectively, on a background of an average coyote block size of 9060 ± 310 Kb, (Fig. 6). However, values of τ for wolf ancestry fragments in Northeastern coyotes are much less (τ = 50.6 ± 11.6, range: 48.5–142.0) than those of Midwestern/Southern coyotes, but estimates of time since dog admixture are comparable (τ = 15.6 ± 1.2) (Supplemental Table S6). These values imply the initial admixture between Northeastern coyotes and gray wolves occurred about 100 yr ago and with dogs about 30 yr ago, given a generation time of 2–3 yr.

Discussion

Population subdivision and relationships of wolf-like canids

Analysis with the canine SNP genotyping array adds considerable resolution at multiple levels to the population structure and relationships of wolf-like canids. First, we showed that even within gray wolves, a species with high dispersal abilities, regional and continental patterns of genetic subdivision are found. Many of these genetic partitions were not identified in previous studies of mtDNA sequence or microsatellite loci (Roy et al. 1994; Vilà et al. 1999), highlighting the ability of genome-wide SNP surveys to uncover variants unique to individual populations or that differ substantially in allele frequency (Rosenberg et al. 2002; Li et al. 2008; Novembre et al. 2008). Specifically, Old World wolf populations from Italy, Spain, and Eastern/Northern Europe comprised distinct units that correspond to three well-accepted Ice Age refugia (Fig. 4; Hewitt 1996). Italy and Spain were among the most divergent populations (Italy: mean F_{ST} = 0.15, Spain: mean F_{ST} = 0.11) (Supplemental Table S3C), which likely reflect historic Ice Age isolation as well as drift from recent population contractions (Randi and Lucchini 2002; Lucchini et al. 2004; Fabbri et al. 2007; Gray et al. 2009). Other genetically distinct populations include Eastern and Northern Europe, China, the Middle East, and Southwest Asia (Fig. 4; Supplemental Fig. S2B; Supplemental Table S3C).

Second, in the New World, Mexican wolves appear as the most genetically distinct group, corroborating the hypothesis that this subspecies is a remnant of an ancient invasion from Eurasia and of conservation importance (Fig. 4; Garcia-Moreno et al. 1996). Other genetic partitions were defined in North America as well, including distinct populations on the British Columbian coast, Northern Quebec, and interior North America (Fig. 4). *structure* and topological analyses provide limited support for habitat-related population structure that corresponds to Western Coastal, Eastern Atlantic, Rocky Mountain and Boreal forests, and tundra/taiga habitats (Fig. 4; Supplemental Fig. S5; Carmichael et al. 2007; Musiani et al. 2007; Koblmüller et al. 2009; Muñoz-Fuentes et al. 2009). These results highlight the importance of using genome-wide surveys to better define and evaluate genetic units for conservation, and further support the notion that in high mobile carnivores, ecology may have an important role in restricting gene flow among populations. Finally, within each population there is strong genealogical structure, with individuals generally assigned to the population from

Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

which they were sampled (Fig. 4; Supplemental Fig. S5). This pattern suggests that even for highly mobile species, whole-genome analysis can effectively resolve hierarchies on multiple levels and assay genealogical structure within and among populations.

In contrast to gray wolves, coyote populations are not well partitioned in a *structure* analysis. As in gray wolves, coyotes have high dispersal ability, and this trait is manifested in their recent colonization of North America over the last 100 yr from a historic geographic range in the American South and Southwest (Kays et al. 2008, 2010), which likely weakened the establishment of genetic structure. Nonetheless, Western, Midwestern/Southern and Northeastern coyote population groupings are suggested by our data (Fig. 4; Supplemental Fig. S2D; Supplemental Table S3). The Northeastern population has been established over the last 70 yr (Kays et al. 2008, 2010) and was likely affected by hybridization with Great Lakes wolves throughout the colonization process (see below). The Midwestern coyote population was probably the source for these colonists and is closely related to the Northeastern population in genetic similarity trees (Supplemental Fig. S5). More fine-scale patterns of differentiation may be evident given better sampling of specific areas and ecotypes (e.g., Sacks et al. 2004, 2005, 2008).

Admixture in enigmatic canids

The evolutionary history of the red wolf is controversial, with three primary theories of its origin as: (1) a distinct North American species most closely aligned with gray wolves of Old World origin; (2) a species derived independently from a coyote-like ancestor in the New World and possibly conspecific with the Great Lakes wolf; and (3) a hybrid population of coyotes and gray wolves (Supplemental Table S1). The red wolf was listed as an endangered species in 1967 because of dwindling populations in the American South and extensive hybridization with coyotes (McCarley 1962; Nowak 1979). Prior to extinction in the wild, 14 individuals from this population with supposed red wolf characteristics were chosen for captive breeding (Phillips and Parker 1988; Wayne and Jenks 1991; Hedrick and Fredrickson 2008). This process may have also unintentionally selected for admixed individuals with a higher proportion of gray wolf ancestry. Nonetheless, our genome-wide analysis of red wolves finds levels of divergence between them and coyotes (F_{ST} = 0.08–0.10) comparable to the level of genetic distinction between gray wolf populations such as European and North American wolves (F_{ST} = 0.08), Great Lakes and Western wolves (F_{ST} = 0.05), coastal B.C. and Northern Quebec wolves (F_{ST} = 0.11), or Mexican and Western gray wolves (F_{ST} = 0.10) (Supplemental Table S3). In the PC analysis, coyotes and red wolves are in close proximity on the first two axes, and *structure* analysis infers substantial coyote ancestry for red wolves (Figs. 3, 4). Red wolves are only resolved as distinct on more minor PC axes and at K = 9 in the *structure* analysis. However, caution needs to be used even in the interpretation of this result, because founder effect and inbreeding as well as small effective population size may inflate the probability of a cluster being defined in *structure* (Pritchard et al. 2000; Anderson and Dunham 2008).

In both *structure* and SABER analyses, red wolves appear to have an admixed ancestry with ~75%–80% of their genome attributed to coyotes and the remainder to gray wolves (Figs. 4, 6D; Supplemental Table S5). Detailed assignments of red wolf chromosomal segments found coyote ancestry blocks 2.6 times longer on average than those assigned to gray wolves, and more dispersed in size (Supplemental Fig. S8; Supplemental Table S5). These results support the hypotheses that red wolves are closely related to coyotes, but somewhat divergent from them due to a history of limited admixture with gray wolves. Such historic admixture between gray wolves and coyotes was followed by extensive backcrossing to coyotes, as the source population of gray wolves disappeared in the American South and the Southeast. We estimate admixture was initiated 144 generations (287–430 yr) ago, placing it approximately in a period when the Southeast U.S. was being converted to agriculture and predators were intensely hunted for fur or as pests (McCarley 1962; Paradiso 1968; McCarley and Carley 1979; Ferrell et al. 1980). Previous model-based analysis using microsatellite data also predicted a relatively recent hybridization between the two species, but the time interval was large (0–2500 yr ago) (Reich et al. 1999). The implications of our results are that a component of the phenotypic distinction of red wolves may be attributed to historic hybridization of distinct populations of gray wolves and coyotes. It has been suggested that hybrids are not clearly protected under the ESA (O'Brien and Mayr 1991), especially hybrids between nonlisted entities (U.S. Fish and Wildlife Service 1973). Since a critical aim of the red wolf recovery project is to maintain the introduced population free from hybridization (Hedrick and Fredrickson 2008), the rationale of the program may need reconsidering as the extant red wolves clearly derive from a process of admixture.

The Great Lakes wolf has generally been considered a subspecies of gray wolf (*Canis lupus lycaon*) with a distribution centered on the Northern Great Lakes region (for review, see Nowak 2002). However, recent genetic studies found unique mtDNA haplotypes in Great Lakes wolves that were allied with coyotes and red wolves, as well as shared microsatellite alleles, suggesting that Great Lakes and red wolves may be conspecific or that the former is a distinct species derived from coyotes (Wilson et al. 2000; Grewal et al. 2004; Fain et al. 2010). Our results show that the Great Lakes wolves are genetically distinct from Western gray wolves (F_{ST} = 0.05), although whether such distinction reflects subspecies, ecotype, or distinct population status is controversial (Cronin and Mech 2009; Koblmüller et al. 2009; Wheeldon and White 2009). However, we do not find genomic evidence of an association of Great Lakes wolves and red wolves. In fact, detailed reconstruction of ancestry segments across the genome finds dramatically different patterns in the two species and shows that the Great Lakes wolf genome largely consists of fragments assigned to gray wolf ancestry (>84%) (Fig. 6C; Supplemental Table S5). Such gray wolf–derived fragments tend to be much longer and more dispersed in size, suggesting extensive recent backcrossing to parental gray wolf populations (Supplemental Fig. S8). The high assignment to gray wolf agrees with the *structure* analysis that consistently shows admixture varying among individuals with the dominant component being derived from gray wolves. At K = 10, a distinct profile is evident that is typical of other gray wolf ecotypes or geographically distinct populations (Fig. 4). Further, Great Lakes wolves are clustered with gray wolves in PCA (Fig. 3). High LD and low ROH support admixture, rather than inbreeding in the Great Lakes wolves (Fig. 5). Consequently, our results suggest admixture between a variety of gray wolf and coyotes may have contributed to the distinct phenotype and intermediate size of the Great Lakes wolf (Nowak 2009). This intermediate-sized wolf might be better adapted for hunting deer, the dominant prey in the region, rather than larger ungulates found elsewhere, and is a unique example of hybridization contributing to adaptive differentiation in a mammal species (Kays et al. 2010).

Model-based analysis of chromosome fragments assigned to gray wolves and coyotes suggests that admixture of Great Lakes

Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

wolves occurred on the order of 600–900 yr ago, prior to the recent invasion of coyotes into the area (Supplemental Table S5) (Grewal et al. 2004; Wilson et al. 2009; Rutledge et al. 2010b). However, extinct coyote-like mtDNA sequences have been described in 400–500-yr-old canid remains from Quebec (Rutledge et al. 2010b), and the presence of coyote-like mtDNA haplotypes in historic Great Lakes wolves implies that coyote–wolf hybridization has a longer history in the area (Leonard and Wayne 2008; Koblmüller et al. 2009). Moreover, pre-Columbian archeological remains establish coyotes in the Great Lakes area, but the reasons for their subsequent disappearance are unclear (Voigt and Berg 1987). Therefore, Great Lakes wolves may have hybridized with pre-Columbian, as well as contemporary coyotes. More recently, admixture could be augmented by human activities such as predator control and habitat conversion that could favor hybridization with the more common coyote (Lehman et al. 1991).

Admixture in coyotes

Midwestern/Southern and Northeastern coyotes show a surprising fraction of admixture with gray wolves and dogs, which is maintained to some extent in their descendents who colonized the Great Lakes region and New England. The fraction of ancestry blocks that are associated with gray wolves in the SABER analysis is as much as 13.1% (Fig. 6I; Supplemental Table S6). *structure* analyses also suggest limited dog ancestry, and SABER analysis with dogs as a third ancestral population found similar levels overall of dog ancestry in Northeastern (9.1%) and Midwestern/Southern coyotes (7.5%) (Supplemental Table S6). Recent dog–coyote admixture is likely given the abundance of dogs associated with Native Americans and extensive human habitation of the Eastern US (Schwartz 1997; Morey 2010). Further, this possibility is supported by morphologic studies and genetic evidence of a dog-like mitochondrial DNA haplotype and dog-derived black coat color variants in coyotes of the Southeast (Mengel 1971; Freeman and Shaw 1979; Adams et al. 2003; Schmutz et al. 2007; Anderson et al. 2009; Kays et al. 2010; Supplemental Table S1). As in red wolves, our results suggest that historic admixture between gray wolves and coyotes began as long as 250–300 yr ago, coincident with the decline and extirpation of the gray wolf in the Midwestern and Southern US. This result reaffirms the importance of relative species abundance to admixture, and suggests that the processes that gave rise to the red wolf may be similar to those leading to admixture in Midwestern/Southern and Northeastern coyotes. The latter display the highest fraction of wolf ancestry and elevated LD relative to ROH (Fig. 5; Supplemental Table S5), suggesting the effect of recent admixture as coyotes expanded their range through the Great Lakes region beginning early in the last century.

In conclusion, we show how new genomic tools, such as SNP genotyping arrays developed from one species, greatly enhance an understanding of population subdivision and admixture in close relatives. Studies of wolf-like canids with the canine SNP microarray revealed a hierarchy of genetic partitions from individual populations to higher order regional and habitat-related groupings. We find a coyote–wolf admixture zone that stretched from Southern Texas to the Great Lakes and Northeastern US. This admixture zone is the largest in area ever described for a terrestrial vertebrate (Barton and Hewitt 1985; Arnold 1997) and is testimony to the dispersal ability of wolf-like canids as well as the influence of anthropogenic activities. In addition to possible natural and artificial selection for larger individuals (Kays et al. 2010), a shared history of coyote–wolf admixture likely contributed to the phenotypic and genetic similarity of the red wolf and Great Lakes wolf and their classification as conspecific or closely related unique species. Using a genome-wide approach, we show that the red and Great Lakes wolves have a distinct but admixed evolutionary history. This result has important implications for conservation policy, because current preservation efforts are focused on populations whose admixed genomes may be due in part to recent habitat changes and predator control efforts (Lehman et al. 1991; Wayne and Jenks 1991). However, these concerns must be weighed against the beneficial top-down ecosystem effects that admixed populations have in environments, which now may be unsuitable for large wolves. Such ecologic, rather than strictly taxonomic considerations are also integral to deciding which species and subspecies should be preserved (e.g., Crandall et al. 2000; Allendorf et al. 2001; Carroll et al. 2010).

Methods

Sample and data collection

We extracted genomic DNA from blood samples collected from 912 domestic dogs (*Canis familiaris*) and tissue and blood samples from gray wolves (*Canis lupus*), coyotes (*C. latrans*), red wolves (*C. rufus*), Mexican wolves (*C. l. baileyi*), black-backed jackals (*C. mesomelas*), Ethiopian wolves (*C. simensis*), golden jackals (*C. aureus*), and a side-striped jackal (*C. adustus*; Supplemental Table S2) (see also Boyko et al. 2010; vonHoldt et al. 2010). All samples were genotyped on the Affymetrix Canine version 2 genome-wide SNP mapping array, and quality control filters were applied for the genotyping algorithm (Boyko et al. 2010), from which we obtained a final set of 48,036 high-quality autosomal SNP loci excluding X-chromosome SNPs (48K).

Single-SNP measures of genetic diversity

Single-marker descriptive statistics (e.g., observed/expected heterozygosity, polymorphism, missingness, and pairwise genotype associations) were estimated using PLINK (Purcell et al. 2007) for the complete 48K SNP data set. When pedigree data was known (Mexican and red wolf studbooks, or Yellowstone wolf pedigree), only individuals unrelated by at least two generations were included. Additionally, we identified a set of 43,953 SNPs (44K) after exclusion of highly linked SNPs ($r^2 > 0.5$) for all canids, and an additional subset of 30,168 unlinked SNPs (30K) in coyotes and wolves.

PCA

To visualize the dominant relationships in the 48K SNP genotype data set of all canids, we used the smartpca program distributed in the Eigensoft package for principal component analysis (PCA) (Price et al. 2006). To determine how SNP ascertainment influences the PCA structure, we selected a set of 710 SNPs ascertained by comparison of only dog and wolf or coyote sequences for PCA with the same sample set (see vonHoldt et al. 2010).

Estimating genetic differentiation and variation

We used an in-house program to calculate Weir and Cockerham's estimate of the F_{ST} parameter between populations and species suggested by PCA, *structure*, and topological analyses (Weir and Cockerham 1984). An allele-similarity (IBS) matrix was constructed for wild canids using PLINK and used as input for ARLEQUIN version 3 in order to analyze molecular variance (AMOVA) within and among population and groups with 9999 permutations for significance testing (Excoffier et al. 2005).

Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

Genetic structure analysis

We used the Bayesian inference program *structure* (Pritchard et al. 2000) to assess genetic admixture for the 44K pruned SNP data set of 300 individuals: 53 dog breeds (one individual per modern breed), wolves (gray wolf, $n = 163$; red wolf, $n = 12$), and coyotes ($n = 57$) (vonHoldt et al. 2010). To limit the effect of the large number of dogs relative to wild canids, we reduced the sampling of dogs to include only one individual per modern breed (see vonHoldt et al. 2010). We utilized 10,000 burn-in iterations and 10,000 MCMC iterations in *structure*, with three repetitions of these parameter settings. The alpha and likelihood statistics were verified to reach convergence before the 10,000 burn-in iterations were completed during each repetition for each number of K populations analyzed.

Linkage disequilibrium analysis

We used PLINK (Purcell et al. 2007) to obtain genome-wide pairwise genotypic associations (r^2) as an estimate of linkage disequilibrium (LD). Analyses were performed independently on unrelated individuals from the three coyote populations (Western, $n = 25$; Midwestern/Southern, $n = 19$; Northeastern, $n = 13$), three wolf populations (Western, $n = 43$; Great Lakes, $n = 10$; and IRNP, $n = 3$), the red wolf ($n = 6$), and the Mexican wolf ($n = 4$). The physical distance at which LD decayed to $r^2 = 0.5$ ($r^2_{0.5}$) was estimated for each population grouping (see Gray et al. 2009). We assessed runs of homozygosity (ROH) from the pruned 30K SNP set using a hidden Markov model (Boyko et al. 2010) for the same sample set as the LD analysis. We assumed a 1.0-cM/Mb recombination rate and then assigned each ROH to either autozygosity (ROH > 1 Mb) or ancient population event (ROH < 1 Mb) (Boyko et al. 2010).

Modeling ancestry and timing of admixture

We used the program SABER (Tang et al. 2006) to assign ancestry to each haplotype block across the genome. SABER utilizes an extended Markov-Hidden Markov Model (MHMM) to account for background LD (Tang et al. 2006). We specified a 1.0-cM/Mb recombination rate (Boyko et al. 2010). We conducted two-ancestor and three-ancestor analyses. Both analyses included 12 individuals each of the Western coyote and Western North American wolf populations as the ancestral reference populations. Additionally, in a three-ancestor model, we included 12 dogs from 12 modern breeds (see Supplemental Material) for analysis of only the Midwestern/Southern and Northeastern coyote populations, because *structure* analyses suggested that these populations were admixed with dogs. Following previous studies that utilized subsets of ancestry informative markers (AIMs) that are diagnostic of parental populations to enhance haplotype assignment (Tian et al. 2006; Price et al. 2007; Rosenberg et al. 2010), we used a subset of 3102 (two-ancestor) and 7183 SNPs (three-ancestor) that had $F_{ST} \geq 0.4$ (see Supplemental Material).

Tree reconstruction

We generated neighbor-joining (NJ) trees based on allele-sharing distances among the subset of nonadmixed individuals for the 48K SNP data set (Supplemental Fig. S5) with 1000 bootstrap iterations using the Microsat program (written by E Minch and A Ruiz-Linares, Stanford University, 1996). The resulting pairwise matrices were used in neighbor from the PHYLIP package, and consensus trees were generated using the majority rule option in the program consense from the PHYLIP package (Felsenstein 1993). Trees were visualized using Dendroscope (Huson and Richter 2007).

Acknowledgments

Grants from NSF and NIH (C.D.B., J.N., M.M., and R.K.W.), the Intramural Program of the National Human Genome Research Institute (E.A.O.), the Searle Scholars program (J.N.), the Polish Ministry of Science and Higher Education (M.P. and W.J.), European Nature Heritage Fund EURONATUR (W.J.), New York State Museum (R.K.), National Sciences and Engineering Research Council (M.M.) supported this research. We thank Y. Zhang and team for their permission to include 10 Chinese wolf samples in this study. A NIH Training Grant in Genomic Analysis and Interpretation supported B.V.H.; the Foundation for Polish Science supported M.P. Wolf samples from Central/Eastern/Northern Europe and Turkey were collected as a result of an ongoing project on genetic differentiation in Eurasian wolves. We thank the project participants (M. Shkvyrya, I. Dikiy, E. Tsingarska, S. Nowak) for their permission to use 72 samples for this study. We acknowledge R. Hefner and the Zoological Collection at Tel Aviv University for Israel wolf samples. We also gratefully acknowledge Affymetrix Corporation and specifically Zuwei Qian from Affymetrix Asia Pacific for assistance. We also want to thank Adam Auton and Andy Reynolds for assistance with array genotyping, Scott Livingston for comments on the manuscript, and Hua Tang for advice with ancestry analyses.

References

Adams JR, Leonard JA, Waits LP. 2003. Widespread occurrence of a domestic dog mitochondrial DNA haplotype in southeastern US coyotes. *Mol Ecol* **12:** 541–549.

Addison JA, Pogson GH. 2009. Multiple gene genealogies reveal asymmetrical hybridization and introgression among strongylocentrotid sea urchins. *Mol Ecol* **18:** 1239–1251.

Allendorf FW, Leary RF, Spruell P, Wenburg JK. 2001. The problems with hybrids: setting conservation guidelines. *Trends Ecol Evol* **16:** 613–622.

Anderson EC, Dunham KK. 2008. The influences of family groups on inferences made with the program Structure. *Mol Ecol Resources* **8:** 1219–1229.

Anderson T, vonHoldt B, Candille S, Musiani M, Greco C, Stahler DR, Smith DW, Padhukasahasram B, Randi E, Leonard J, et al. 2009. Molecular and evolutionary history of melanism in North American gray wolves. *Science* **323:** 1339–1343.

Arnold ML. 1997. *Natural hybridization and evolution*. Oxford University Press, NY.

Aspi J, Roininen E, Kiiskila J, Ruokonen M, Kojola I, Bljudnik L, Danilov P, Heikkinen S, Pulliainen E. 2009. Genetic structure of the northwestern Russian wolf population and gene flow between Russia and Finland. *Conserv Genet* **10:** 815–826.

Barton NH, Hewitt GM. 1985. Analysis of hybrid zones. *Annu Rev Ecol Syst* **16:** 113–148.

Bekoff M, Wells MC. 1986. The social ecology and behavior of coyotes (*Canis latrans*). *Adv Stud Behav* **16:** 251–338.

Boyko AR, Quignon P, Li L, Schoenebeck J, Degenhardt JD, Lohmueller KE, Zhao K, Brisbin A, Parker HG, vonHoldt BM, et al. 2010. A simple genetic architecture underlies morphological variation in dogs. *PLoS Biol* **8:** e1000451. doi: 10.1371/journal.pbio.1000451.

Carmichael LE, Krizan J, Nagy J, Fuglei E, Dumond M, Johnson D, Veitch A, Berteaux D, Strobeck C. 2007. Historical and ecological determinants of genetic structure in arctic canids. *Mol Ecol* **16:** 3466–3483.

Carroll C, Vucetich JA, Nelson MP, Rohlf DJ, Phillips MK. 2010. Geography and recovery under the U.S. Endangered Species Act. *Conserv Biol* **24:** 395–403.

Conrad DF, Jakobsson M, Coop G, Wen X, Wall JD, Rosenberg NA, Pritchard JK. 2006. A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nat Genet* **38:** 1251–1260.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK. 2000. Considering evolutionary processes in conservation biology. *Trends Ecol Evol* **15:** 290–295.

Cronin MA, Mech LD. 2009. Problems with the claim of ecotype and taxon status of the wolf in the Great Lakes region. *Mol Ecol* **18:** 4991–4993.

England PR, Cornuet JM, Berthier P, Tallmon DA, Luikart G. 2006. Estimating effective population size from linkage disequilibrium: severe bias in small samples. *Conserv Genet* **7:** 303–308.

Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinform Online* **1:** 47–50.

Fabbri E, Miquel C, Lucchini V, Santini A, Caniglia R, Duchamp C, Weber JM, Lequette B, Marucco F, Boitani L, et al. 2007. From the Apennines to the Alps: colonization genetics of the naturally expanding Italian wolf (*Canis lupus*) population. *Mol Ecol* **16:** 1661–1671.

Fain SR, Straughan DJ, Taylor BF. 2010. Genetic outcomes of wolf recovery in the Western Great Lakes states. *Conserv Genet*. doi: 10.1007/s10592-010-0068-x.

Felsenstein J. 1993. *PHYLIP version 3.5c*. Department of Genetics, University of Washington, Seattle, WA.

Ferrell R, Morizot D, Horn J, Carley C. 1980. Biochemical markers in a species endangered by introgression: the red wolf. *Biochem Genet* **18:** 39–49.

Forbes SH, Boyd DK. 1997. Genetic structure and migration in native and reintroduced Rocky Mountain wolf populations. *Conserv Biol* **11:** 1226–1234.

Freeman R, Shaw J. 1979. Hybridization in Canis (Canidae) in Oklahoma. *Southwest Nat* **24:** 485–500.

Garcia-Moreno J, Matocq M, Roy M, Geffen E, Wayne RK. 1996. Relationships and genetic purity of the endangered Mexican wolf based on analysis of microsatellite loci. *Conserv Biol* **10:** 376–389.

Gaut BS, Long A. 2003. The lowdown on linkage disequilibrium. *Plant Cell* **15:** 1502–1506.

Gray MM, Granka J, Bustamante CD, Sutter N, Boyko A, Zhu L, Ostrander EA, Wayne RK. 2009. Linkage disequilibrium and demographic history of wild and domestic canids. *Genetics* **181:** 1493–1505.

Grewal S, Wilson JP, Kung TK, Shami K, Theberge M, Theberge JB, White BN. 2004. A genetic assessment of the eastern wolf (*Canis lycaon*) in Algonquin Provincial Park. *J Mammal* **85:** 625–632.

Hedrick PW, Fredrickson J. 2008. Captive breeding and the reintroduction of Mexican and red wolves. *Mol Ecol* **17:** 344–350.

Hedrick PW, Miller PS, Geffen E, Wayne RK. 1997. Genetic evaluation of the three captive Mexican wolf lineages. *Zoo Biol* **16:** 47–69.

Hewitt G. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol J Linn Soc Lond* **58:** 247–276.

Huson D, Richter D. 2007. Dendrosope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics* **8:** 460–465.

Jakobsson M, Scholz S, Scheet P, Giggs JR, Vanliere J, Fung H, Szpiech Z, Degnan J, Wang K, Guerreiro R, et al. 2008. Genotype, haplotype and copy-number variation in worldwide human populations. *Nature* **451:** 998–1003.

Jedrzejewski W, Branicki W, Veit C, Medugorac I, Pilot M, Bunevich AN, Jedrzejewska B, Schmidt K, Theuerkauf J, Okarma H, et al. 2005. Genetic diversity and relatedness within packs in an intensely hunted population of wolves *Canis lupus*. *Acta Theriol (Warsz)* **50:** 3–22.

Kalinowski ST, Hedrick PW, Miller PS. 1999. No inbreeding depression observed in Mexican and Red wolf captive breeding programs. *Conserv Biol* **13:** 1371–1377.

Kays RW, Gompper ME, Ray JC. 2008. Landscape ecology of eastern coyotes based on large-scale estimates of abundance. *Ecol Appl* **18:** 1014–1027.

Kays R, Curtis A, Kirchmann JJ. 2010. Rapid adaptive evolution of northeastern coyotes via hybridization with wolves. *Biol Lett* **6:** 89–93.

Koblmüller S, Nord M, Wayne RK, Leonard JA. 2009. Origin and status of the Great Lakes wolf. *Mol Ecol* **18:** 2313–2326.

Kyle CJ, Johnson AR, Patterson BR, Wilson PJ, Grewal SK, White BN. 2006. Genetic nature of eastern wolves: Past, present and future. *Conserv Genet* **7:** 273–287.

Lehman N, Eisenhawer A, Hansen K, Mech LD, Peterson RO, Gogan PJP, Wayne RK. 1991. Introgression of coyote mitochondrial DNA into sympatric North American gray wolf populations. *Evolution* **45:** 104–119.

Leonard JA, Wayne RK. 2008. Native Great Lakes wolves were not restored. *Biol Lett* **4:** 95–98.

Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, Cann HM, Barsh GS, Feldman M, Cavalli-Sforza LL, et al. 2008. Worldwide human relationships inferred from genome-wide patterns of variation. *Science* **319:** 1100–1104.

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

Lucchini V, Galov A, Randi E. 2004. Evidence of genetic distinction and long-term population decline in wolves (*Canis lupus*) in the Italian Apennines. *Mol Ecol* **13:** 523–536.

McCarley H. 1962. The taxonomic status of wild canids (Canidae) in the south central United States. *Southwest Nat* **7:** 227–235.

McCarley H, Carley C. 1979. Recent changes in distribution and status of wild red wolves (*Canis rufus*). Endangered Species Report no. 4. U.S. Fish and Wildlife Service, Albuquerque, NM.

Mengel RM. 1971. A study of dog-coyote hybrids and implications concerning hybridization in Canis. *J Mammal* **52:** 316–336.

Morey D. 2010. *Dogs: Domestication and the development of a social bond.* Cambridge University Press, NY.

Muñoz-Fuentes V, Darimont CT, Wayne RK, Paquet PC, Leonard JA. 2009. Ecological factors drive differentiation in wolves from British Columbia. *J Biogeogr* **36:** 1516–1531.

Musiani M, Leonard JA, Cluff HD, Gates CC, Mariani S, Paquet PC, Vilà C, Wayne RK. 2007. Differentiation of tundra/taiga and boreal coniferous forest wolves: genetics, coat colour and association with migratory caribou. *Mol Ecol* **16:** 4149–4170.

Nielsen R, Signorovitch J. 2003. Correcting for ascertainment biases when analyzing SNP data: applications to the estimation of linkage disequilibrium. *Theor Popul Biol* **63:** 245–255.

Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko A, Auton A, Indap A, King KS, Bergmann S, Nelson MR, et al. 2008. Genes mirror geography within Europe. *Nature* **456:** 98–101.

Nowak RM. 1979. *North American Quaternary* Canis. Monograph of the Museum of Natural History. University of Kansas, KS.

Nowak RM. 2002. The original status of wolves in eastern North America. *Southeast Nat* **1:** 95–130.

Nowak RM. 2009. Taxonomy, morphology, and genetics of wolves in the Great Lakes region of the United States. In *Recovery of gray wolves in the Great Lakes regions of the United States: An endangered species success story* (ed. AP Wydeven, et al.), pp. 233–250. Springer, NY.

O'Brien SJ, Mayr E. 1991. Bureaucratic mischief: recognizing endangered species and subspecies. *Science* **251:** 1187–1188.

Paradiso J. 1968. Canids recently collected in east Texas, with comments on the taxonomy of the red wolf. *Am Midl Nat* **80:** 529–535.

Perini FA, Russo CA, Schrago CG. 2009. The evolution of South American endemic canids: a history of diversification and morphological parallelism. *J Evol Biol* **23:** 311–322.

Phillips MK, Parker WT. 1988. Red wolf recovery: a progress report. *Conserv Biol* **2:** 139–141.

Phillips MK, Henry VG, Kelly BT. 2003. Restoration of the red wolf. In *Wolves: Behavior, ecology, and conservation* (ed. LD Mech, L Boitani), pp. 272–288. University of Chicago Press, Chicago, IL.

Pilot M, Jedrzejewski W, Branicki W, Sidorovich VE, Jedrzejewska B, Stachura K, Funk SM. 2006. Ecological factors influence population genetic structure of European grey wolves. *Mol Ecol* **15:** 4533–4553.

Posada D, Crandall KA. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends Ecol Evol* **16:** 37–45.

Price A, Patterson N, Plenge RM, Weinblatt ME, Shadick NA, Reich D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38:** 904–909.

Price AL, Patterson N, Yu F, Cox DR, Waliszewska A, McDonald GJ, Tandon A, Schirmer C, Neubauer J, Bedoya G, et al. 2007. A genomewide admixture map for Latino populations. *Am J Hum Genet* **80:** 1024–1036.

Pritchard JK, Przeworski M. 2001. Linkage disequilibrium in humans: models and data. *Am J Hum Genet* **69:** 1–14.

Pritchard J, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81:** 559–575.

Ramirez O, Altet L, Enseñat C, Vilà C, Sanchez A, Ruiz A. 2006. Genetic assessment of the Iberian wolf *Canis lupus signatus* captive breeding program. *Conserv Genet* **7:** 861–878.

Randi E, Lucchini V. 2002. Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) populations by Bayesian admixture analyses of microsatellite variation. *Conserv Genet* **3:** 31–45.

Reich D, Wayne RK, Goldstein DB. 1999. Genetic evidence for a recent origin by hybridization of red wolves. *Mol Ecol* **8:** 139–144.

Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, Feldman MW. 2002. Genetic structure of human populations. *Science* **298:** 2381–2385.

Rosenberg NA, Huang L, Jewett EM, Szpiech Z, Jankovic I, Boehnke M. 2010. Genome-wide association studies in diverse populations. *Nat Rev Genet* **11:** 356–366.

Rosenblum EB, Novembre J. 2007. Ascertainment bias in spatially structured populations: a case study in the eastern fence lizard. *J Hered* **98:** 331–336.

Roy M, Geffen E, Smith D, Ostrander EA, Wayne RK. 1994. Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. *Mol Biol Evol* **11:** 553–570.

Roy MS, Geffen E, Smith D, Wayne RK. 1996. Molecular genetics of pre-1940 red wolves. *Conserv Biol* **10:** 1413–1424.

Rutledge LY, Garroway CJ, Loveless KM, Patterson BR. 2010a. Genetic differentiation of eastern wolves in Algonquin Park despite bridging gene flow between coyotes and grey wolves. *Heredity* **105:** 520–531.

Downloaded from genome.cshlp.org on June 3, 2011 - Published by Cold Spring Harbor Laboratory Press

Rutledge LY, Bos KI, Pearce RJ, White BN. 2010b. Genetic and morphometric analysis of sixteenth century *Canis* skull fragments: implications for historic eastern and gray wolf distribution in North America. *Conserv Genet* **11:** 1273–1281.
Sacks BN, Brown SK, Ernest HB. 2004. Population structure of California coyotes correspond to habitat-specific breaks and illuminates species history. *Mol Ecol* **13:** 1265–1275.
Sacks BN, Mitchell BR, Williams CL, Ernest HB. 2005. Coyote movements and social structure along a cryptic population genetic subdivision. *Mol Ecol* **14:** 1241–1249.
Sacks BN, Bannasch DL, Chomel BB, Ernest HB. 2008. Coyotes demonstrate how habitat specialization by individuals of a generalist species can diversify population in a heterogeneous environment. *Mol Biol Evol* **25:** 1384–1394.
Schmutz SM, Berryere TG, Barta JL, Reddick KD. 2007. Agouti sequence polymorphisms in coyotes, wolves and dogs suggest hybridization. *J Hered* **98:** 351–355.
Schwartz M. 1997. *The history of dogs in the early Americas*. Yale University Press, New Haven, CT.
Susnik S, Weiss S, Odak T, Delling B, Treer T, Snoj A. 2007. Reticulate evolution: ancient introgression of the Adriatic brown trout mtDNA in softmouth trout *Salmo obtusirostris* (Teleostei: Salmonidae). *Biol J Linnean Soc* **90:** 139–152.
Tang H, Coram M, Wang P, Zhu X, Risch N. 2006. Reconstructing genetic ancestry blocks in admixed individuals. *Am J Hum Genet* **79:** 1–12.
Teare MD, Dunning AM, Durocher F, Rennart G, Easton DF. 2002. Sampling distribution of summary linkage disequilibrium measures. *Ann Hum Genet* **66:** 223–233.
Tian C, Hinds DA, Shigeta R, Kittles R, Ballinger DG, Seldin MF. 2006. A genomewide single-nucleotide-polymorphism panel with high ancestry information for African American admixture mapping. *Am J Hum Genet* **79:** 640–649.
U. S. Fish and Wildlife Service. 1973. Endangered Species Act. Department of the Interior, 108th Congress, Washington, DC.
Vilà C, Amorim IR, Leonard JA, Posada D, Castroviejo J, Petrucci-Fonseca F, Crandall KA, Ellegren H, Wayne RK. 1999. Mitochondrial DNA phylogeography and population history of the grey wolf *Canis lupus*. *Mol Ecol* **8:** 2089–2103.
Voigt DR, Berg WE. 1987. Coyote. In *Wild Furbearer Management and Conservation in North America* (ed. M Novak, et al.), pp. 345–357. Ontario Trappers Association, Peterborough, Ontario, Canada.
vonHoldt B, Stahler D, Smith DW, Earl DA, Pollinger JP, Wayne RK. 2008. The genealogy and genetic viability of reintroduced Yellowstone grey wolves. *Mol Ecol* **17:** 252–274.
vonHoldt B, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, Degenhardt JD, Boyko AR, Earl DA, Auton A, et al. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* **464:** 898–902.
Wayne RK, Jenks S. 1991. Mitochondrial DNA analysis implying extensive hybridization of the endangered red wolf *Canis rufus*. *Nature* **351:** 565–568.
Wayne RK, Lehman N, Allard MW, Honeycutt RL. 1992. Mitochondrial DNA variability of the gray wolf: genetic consequences of population decline and habitat fragmentation. *Conserv Biol* **6:** 559–569.
Weir B, Cockerham C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* **38:** 1358–1370.
Wheeldon T, White B. 2009. Genetic analysis of historic western Great Lakes region wolf samples reveals early *Canis lupus*/*Canis lycaon* hybridization. *Biol Lett* **5:** 101–104.
Wilson P, Grewal S, Lawford ID, Heal JNM. 2000. 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.
Wilson PJ, Grewal SK, Mallory FF, White BN. 2009. Genetic characterization of hybrid wolves across Ontario. *J Hered* **100:** S80–S89.

Received October 8, 2010; accepted in revised form April 4, 2011.