Ancient Hybridization with an Unknown Population Facilitated High-Altitude Adaptation of Canids

Ming-Shan Wang,[†,1,2,3,4] Sheng Wang,[†,1,2] Yan Li,[†,5] Yadvendradev Jhala,[†,6] Mukesh Thakur,[†,7] Newton O. Otecko,[1,2] Jing-Fang Si,[8] Hong-Man Chen,[1] Beth Shapiro,[*,3,4] Rasmus Nielsen (iD),[*,9,10] Ya-Ping Zhang,[*,1,5,11] and Dong-Dong Wu[*,1,11]

[1]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

[2]Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, China

[3]Howard Hughes Medical Institute, University of California Santa Cruz, Santa Cruz, CA

[4]Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA

[5]State Key Laboratory for Conservation and Utilization of Bio-Resource, Yunnan University, Kunming, China

[6]Wildlife Institute of India, Chandrabani, Dehradun, Uttarakhand, India

[7]Zoological Survey of India, New Alipore, Kolkata, West Bengal, India

[8]Laboratory of Animal Genetics, Breeding and Reproduction, Ministry of Agriculture of China, National Engineering Laboratory for Animal Breeding, College of Animal Science and Technology, China Agricultural University, Beijing, China

[9]Departments of Integrative Biology and Statistics, University of California Berkeley, Berkeley, CA

[10]Globe Institute, University of Copenhagen, Copenhagen, Denmark

[11]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China

[†]These authors contributed equally to this work.

Corresponding authors: E-mails: beth.shapiro@gmail.com; rasmus_nielsen@berkeley.edu; zhangyp@mail.kiz.ac.cn; wudongdong@mail.kiz.ac.cn.

Associate editor: John Parsch

Abstract

Genetic introgression not only provides material for adaptive evolution but also confounds our understanding of evolutionary history. This is particularly true for canids, a species complex in which genome sequencing and analysis has revealed a complex history of admixture and introgression. Here, we sequence 19 new whole genomes from high-altitude Tibetan and Himalayan wolves and dogs and combine these into a larger data set of 166 whole canid genomes. Using these data, we explore the evolutionary history and adaptation of these and other canid lineages. We find that Tibetan and Himalayan wolves are closely related to each other, and that ∼39% of their nuclear genome is derived from an as-yet-unrecognized wolf-like lineage that is deeply diverged from living Holarctic wolves and dogs. The _EPAS1_ haplotype, which is present at high frequencies in Tibetan dog breeds and wolves and confers an adaptive advantage to animals living at high altitudes, was probably derived from this ancient lineage. Our study underscores the complexity of canid evolution and demonstrates how admixture and introgression can shape the evolutionary trajectories of species.

Key words: genetic introgression, phylogeny, _EPAS1_, high-altitude adaptation, dog and wolf.

Introduction

With the development of high-throughput sequencing and advances in comparative population genomics, gene flow between diverging lineages has increasingly been shown to play an important role in species' evolution (Taylor and Larson 2019). In several cases, past episodes of gene flow following admixture have allowed persistence of genetic segments from lineages that have otherwise become extinct (Supple and Shapiro 2018; Taylor and Larson 2019). For example, admixture with archaic hominins after the dispersal of anatomically modern humans out of Africa resulted in introgression of alleles derived in archaic lineages that are today observed at high frequency in some human populations, indicating that these alleles provide an adaptive advantage (Gittelman et al. 2016; McCoy et al. 2017). A classic example of this is the high frequency of the Denisovan-derived _EPAS1_ allele in present-day Tibetans, which induced blunted hypoxic responses in Tibetans and allowed adaptation to a life in the hypoxic environment on the Tibetan Plateau (Huerta-Sanchez et al. 2014).

Admixture and genomic introgression is particularly pervasive within the family Canidae, which includes dogs, wolves, coyotes, jackals, dingoes, and many other extant and extinct dog-like canids (Wang et al. 2004; Koepfli et al. 2015;
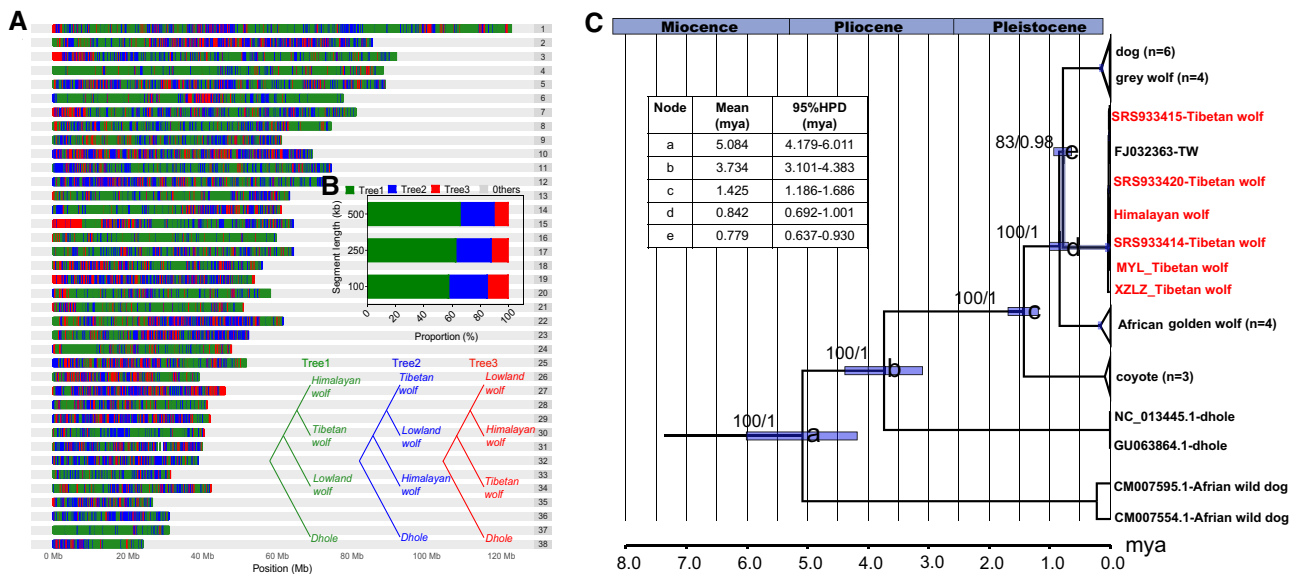
FIG. 1. Phylogenetic analyses of nuclear and mitochondrial genomes showed that Tibetan wolf and Himalayan wolf are closely related. (*A*) Distributions of tree topology constructed using segments with length of 100 kb across each chromosome. (*B*) Bar-plot depicting the proportion of three major topologies estimated with 100-, 250-, and 500-kb segments. (*C*) Chronogram estimated from protein-coding, rRNA and tRNA genes of mitochondrial genomes. Values at nodes (a–e) are bootstrap support based on maximum-likelihood analyses (RAxML) with 1,000 pseudoreplicates and posterior probability from Bayesian inference (BEAST), respectively. Node bars indicate 95% HPD for divergence times, with values shown in inner table. Mitochondrial sequences for HW and TW marked in red were de novo assembled in this study, for other canids were downloaded from NCBI GenBank with accession IDs shown in supplementary figure S6, Supplementary Material online.

vonHoldt, Cahill, et al. 2016; Gopalakrishnan et al. 2018; Werhahn et al. 2020). The resulting reticulated evolutionary history of canids has complicated both disentangling their evolutionary relationships and identifying the timing and location of dog domestication (Freedman et al. 2014; Skoglund et al. 2015; Fan et al. 2016; vonHoldt, Cahill, et al. 2016; Sinding et al. 2018). Among the most enigmatic canid lineages whose origins and evolutionary relationships remain unresolved are the high-altitude wolves (HAWs) of the Qinghai–Tibet Plateau (QTP). These HAWs are often divided into two groups, the Tibetan gray wolf (TW) and the Himalayan wolf (HW). HWs are distributed across the Trans-Himalayan region of Nepal, northern India in the Ladakh region of eastern Kashmir, Himachal Pradesh, and neighboring regions, whereas TWs are found in the provinces of Gansu, Qinghai, and Tibet (Sharma et al. 2004; Werhahn et al. 2018, 2020). Both wolf populations are critically small and vulnerable, and HWs are currently protected by wildlife legislation in India and Nepal.

Despite morphological and genetic research to better understand their evolutionary histories, the origin of HAWs and their relationships to other gray wolves remain unclear. Early morphological work found that the HAWs to be distinct from Eurasian wolves and considered them to be a distinct subspecies (Hodgson 1847). Mitochondrial DNA analyses found that HW cluster with TW in a clade that is distinct from other gray wolves (Sharma et al. 2004; Werhahn et al. 2017, 2020) and suggested that these HAWs could be assigned as a separate species (Sharma et al. 2004; Shrotriya et al. 2012; vonHoldt, Kays, et al. 2016; Werhahn et al. 2017, 2018, 2020). Analyses of whole nuclear genomes, however, found

that TW clustered with gray wolves from northern China in a clade that is sister to European wolves (Fan et al. 2016). This study suggested that TW was derived from a lineage of Asian wolves and recolonized the QTP after divergence from North Chinese wolves around 25 thousand years ago (ka) (Fan et al. 2016). More recently, another analysis of nuclear genomes showed that the phylogenetic placement of TW differed depending on what approach is used to estimate the consensus phylogeny (Zhang et al. 2020). Specifically, a tree estimated using the Neighbor-Joining approach (Zhang et al. 2020: fig. 1e and fig. S5) placed TW as sister to North Chinese wolves, as in the previous analysis (Fan et al. 2016), whereas a phylogeny estimated using maximum likelihood (Zhang et al. 2020: fig. S6) placed TW basal to dogs and all European and Asian wolves. As the nuclear genome for HW has not yet been sequenced, the genetic relationship based on nuclear genome between TW and HW remains unclear.

Both Tibetan wolves and HWs are genetically adapted to the harsh conditions of life at high altitude. One gene, *EPAS1*, which is a transcription factor that induces a physical response to hypoxia, is estimated to have evolved under strong positive selection in both Tibetan dogs (TDs) and TW (Gou et al. 2014; Zhang et al. 2014). Both TW and HW are known carry a unique *EPAS1* haplotype that is thought to have evolved under positive selection (Werhahn et al. 2018, 2020), but the origin of the adaptive *EPAS1* allele remains unknown. Previous studies assumed that *EPAS1* introgressed into TD from TW after dogs were introduced to QTP, probably after human settlement in this region during the Paleolithic (Miao et al. 2017; vonHoldt et al. 2017). This hypothesis is based on the presumed sister relationship between

TW and Asian wolves, and the assumption that TW inhibited the QTP before dogs were introduced. However, these evolutionary relationships remain unresolved, and alternative hypotheses are also possible. For example, the high-altitude *EPAS1* allele may have evolved in dogs and later introgressed into wolves. Introgression from dogs into wolves is known to occur; for example, a three base-pair (bp) mutation in the *K*-locus (a canine beta-defensin gene) causing black coat color and conferring higher fitness was transmitted from dogs into American gray wolves through hybridization (Anderson et al. 2009). Because the evolutionary relationships among the two HAWs of the QTP, other wolves, and dogs remain unclear, so does the origin of the adaptive *EPAS1* allele.

Here, we sequence complete nuclear genomes for HW, TW, and dogs and analyze these along with published genomes of Chinese wolves, dogs, and other canids including the golden jackal and dhole. We infer the evolutionary relationships among HAWs and other canid lineages by considering explicitly admixture among them. We find that TW and HW- are closely related, and that both derive from admixture with Eurasian gray wolves, domestic dogs, and a now-extinct or unknown lineage of wolf-like canid. We then explore specifically the evolutionary origin of the *EPAS1* allele, which we find is shared by these lineages living on QTP and probably derived from the deeply divergent previously undescribed canid lineage.

Results

Genomic Sequencing and Evolutionary Relationship

We sequenced 19 wolf and dog whole genomes to an average coverage of ∼8.5-fold. This includes two TWs, with one sampled from Linzhi of Tibet and another from Qinghai of China, one HW collected from Ladakh, Changtang of India, and 16 dogs from the QTP (>2,000 m; supplementary table S1, Supplementary Material online). Using these and previously published genomic data from wolves and dogs (Auton et al. 2013; Wang et al. 2013; Freedman et al. 2014; Li et al. 2014; Zhang et al. 2014; Fan et al. 2016), we compiled a data set of 166 complete genomes including 9 TWs, 1 HW, 34 TDs, 11 lowland wolves (LWs; one Croatian wolf, one Israeli wolf, three Russian wolves, five Mongolian wolves, and one Chinese wolf), 109 lowland dogs (LDs), 1 golden jackal, and 1 dhole (supplementary tables S1 and S2, Supplementary Material online). We mapped these to the dog reference genome (CanFam3.1) and identified more than 26.4 million single-nucleotide polymorphisms (SNPs).

To estimate the evolutionary relationships among the canids in our alignment, we first performed principal component analysis (PCA) and genetic clustering by ADMIXTURE, which is a tool for maximum-likelihood estimation of individual ancestries (Alexander et al. 2009). Concordant with previous results (Gou et al. 2014; Zhang et al. 2014), our PCA shows that dogs, TWs, and LWs fell into distinct clusters. HW falls close to the TW cluster (supplementary fig. S1, Supplementary Material online). Admixture analyses reflect similar clustering to the PCA and suggest that TWs, LWs and LDs can be assigned into distinct groups (supplementary fig.

S2, Supplementary Material online). It is important to note that both PCA and admixture clustering are likely confounded by several potentially factors including the number of samples for each group, genomic sequencing depth, and demographic history (i.e., bottleneck and admixture) (Schraiber and Akey 2015; Lawson et al. 2018). Since very limited wolf samples included in our study and the variance explained by principal components is low (supplementary fig. S1, Supplementary Material online), we caution that any definite conclusion could not be drawn without further evidences.

We then constructed maximum-likelihood tree graphs based on whole autosome variants using TreeMix (Pickrell and Pritchard 2012) by assuming zero to five migration events. Despite pervasive admixture between dog and wolf populations, all inferred admixture graphs show that HW and TW cluster together in the tree (supplementary fig. S3, Supplementary Material online). Next, we divided genomes from HW, TW, and LW (all LWs, or limited to either Chinese or European wolves) and dhole into segments of length 100, 250, and 500 kb and constructed gene trees, respectively. We used ASTRAL (Mirarab and Warnow 2015) and DiscoVista (Sayyari et al. 2018) to summarize these gene trees. The results of all analyses support that TW and HW are closely related (supplementary figs. S4 and S5, Supplementary Material online). Furthermore, $f4$-statistics (Patterson et al. 2012) in the forms of $f4$(Golden jackal, HW; Pop2, LW/Dog) and $f4$(Golden jackal, Pop2; HW, LW/Dog) where Pop2 represents each Tibetan wolf gave significantly negative values ($Z < -53$; supplementary tables S3 and S4, Supplementary Material online), strongly indicating that HW is genetically closer to TW than it is to LWs.

We next estimated a mitochondrial phylogeny using the above data set and published mitochondrial genomes from dog, gray wolf, coyote, dhole, golden jackal, and African wild dog. HW and six TW mitochondrial lineages formed a distinct clade that is basal to dog and LW (fig. 1C and supplementary fig. S6, Supplementary Material online). We analyzed previously published HW mitochondrial segments including *cytb*, *d-loop*, and *16S-RNA* genes with our data set including HW, TW, and Mongolian wolves and found these HW sequences clustered together with our HW and TW (supplementary fig. S7, Supplementary Material online). We estimated that the TW/HW clade diverged from other present-day dog and gray wolf lineages ∼779 ka (95% highest posterior density, HPD 637–930 ka; fig. 1C). Together, the mitochondrial and nuclear results indicate that HW and TW are closely related.

A phylogeny constructed in a previous study from concatenated nuclear sequences showed that TW and wolves from northern China (including Mongolian wolves) constitute a clade that is sister to European and Middle-Eastern wolves (Fan et al. 2016). This finding is incongruous with our and published mitochondrial phylogenies (fig. 1C and supplementary figs. S6 and S7, Supplementary Material online). However, our whole nuclear genome indicates that TW and HW form a lineage basal to dog and LW (supplementary fig. S3, Supplementary Material online). To further explore the evolutionary relationship of present-day HAWs (including

TW and HW) with dogs and other wolves, we constructed phylogenic trees based on three data sets (Data set 1: HAW, LW [all LWs], LD, TD, and dhole; Data set 2: LW in Data set 1 only includes EUW [European wolf]; and Data set 3: LW in Data set 1 only includes CHW [Chinese wolf]). As above, we divided these genomes into 100-, 250-, and 500-kb segments along each chromosome. Results from different size of segments were generally consistent and revealed three major topologies (T1: HAW is basal to dog and LW; T2: LW is basal to HAW and dog; and T3: HAW is sister to LW) that account more than 90% of the inferred ancestry (supplementary fig. S8, Supplementary Material online). These topologies vary in frequency among different groups of LWs. Specifically, for Data set 1, the most common topology is T1 (38.3–43.1%), followed by T3 (31.5–32.3%) and T2 (23.4–24.0%); for Data set 2, the most common is T1 (40.6–50.8%), followed by T2 (30.2–34.2%) and T3 (12.9–19.7%); and for Data set 3, the most common is T3 (39.6–46.4%), followed by T1 (30.8–33.0%) and T2 (19.2–22.2%). Unlike mitochondrial genome, phylogenies from nuclear genome suggest a complex pattern that the frequency for these topologies slightly differs from each other, and we could not pinpoint a clear relationship between HAWs and LWs relative to dogs.

Next, we used $f4$-statistics (Patterson et al. 2012) to test for admixture between wolves and dogs on the Tibetan Plateau. $f4$-Statistics in the form of $f4$(outgroup, HAW; LW, dog) found that HAWs share more alleles with LW than with dogs ($f4 < 0$; supplementary table S5, Supplementary Material online); in the form of $f4$(outgroup, HAW; TD and LD) revealed evidence of gene flow between HAW and highland dogs ($f4 < 0$; supplementary table S6, Supplementary Material online). And $f4$ in the form of $f4$(outgroup, dog; LW, HAW) showed that dogs from both highland and lowland share more ancestry with LW than with HAW ($f4 < 0$; supplementary table S7, Supplementary Material online). The latter finding is unexpected if HAW cluster with North Asian wolves as previously reported (Fan et al. 2016) and have gene flow with TD. This resembles a pattern observed in the study of Denisovan and Neanderthal genomes (Prufer et al. 2014) and reminds us to think about a scenario in which HAWs probably have admixed ancestry derived a more deeply diverged lineage. Further support for the hypothesis comes from estimate of the pairwise fixation-index (F_{ST}), which shows that HAW is nearly twice as genetically differentiated from dog than is LW (fig. 2A and supplementary fig. S9A, Supplementary Material online). It is also possible, however, that these patterns are the result of repeated episodes of gene flow between LW and dogs. To explore this further, we used qpGraph (Patterson et al. 2012) to construct admixture graphs and reconcile the divergence and gene flow among dog and wolf populations. HAW could be modeled (no $f4$ outliers) as an admixed population derived initially from admixture between a diverged lineage (∼39%) and LW (61%), and then this population (admixed) later derived 18% of ancestry from a lineage related to TD (fig. 2B). This is generally consistent with the three major topologies based on genetic segments across the genomes (supplementary fig. S8, Supplementary Material online). However, admixture graphs

that assume HAW and LW as sister clades and allow gene flow between LW and dog produced an excess of $f4$ outliers, indicative of poor fit (supplementary fig. 10, Supplementary Material online). Given the complicated nature of divergence and population substructure among LDs and Eurasian wolves, we performed a series of additional analyses in which we identify discrete groups comprising EUW, CHW, South Chinese dogs, and African and European dogs and ran qpGraph on these groups in multiple combinations, so as to test the effect of heterogeneity within these groups. As shown in the supplementary figure S11, Supplementary Material online, although the amount of gene flow between these groups and HAW changes slightly depending on what combination of LD and LW groups are included, admixture graphs based on model where HAWs carry diverged genetic input have no $f4$ outliers, suggesting reasonably good model fit.

The analyses described above support the hypothesis that HAW has some amount of ancestry derived from a divergent canid lineage. To explore this further, we searched for genomic regions within HAW with a signature of "divergent" origin using a hidden Markov model (HMM) approach (Skov et al. 2018). The approach looks for short regions of the genome that contain an unexpectedly high density of private mutations in a given individual by comparing the genome sequence in question with an unadmixed reference population. We choose LD and LW as the putatively unadmixed reference panel and measured the density of mutations in 1-kb windows along each chromosome. Because of divergence, substructure, and admixture among dog and wolf populations, LD and LW may not be perfect nonadmixed populations. However, previous admixture among LD and LW with HAW would lead to underestimation of the amount of divergence in HAW. Assuming a posterior probability cutoff of 0.8 as suggested by the author (Skov et al. 2018), we found several putatively highly divergent regions in the HAW genome. These results suggest that HAW contains admixed ancestry with a wolf lineage that is considerably more deeply diverged than LD and LW. To further explore the evolutionary history of these highly divergent regions of the genome, we created an alignment limited to these regions of the genomes that included the North American wolf, coyote, dhole, and African wild dog and constructed a Neighbor-Joining tree. The resulting phylogeny shows the two North American wolves basal to dog and Eurasian wolves and HAW basal to all Holarctic wolves and dogs (fig. 2C), similarly to the mitochondrial phylogeny (fig. 1C). The same result is supported by a Neighbor-Joining tree inferred from divergent fragments recovered using a stricter cutoff of 0.9 posterior probability (supplementary fig. S9B, Supplementary Material online). Overall, these analyses demonstrate that the genomes of present-day wolves on the QTP contain a proportion of ancestry shared with LW, a proportion of ancestry shared with domestic dogs, and a proportion of ancestry shared with at least one other more deeply diverged lineage that is as yet undescribed.
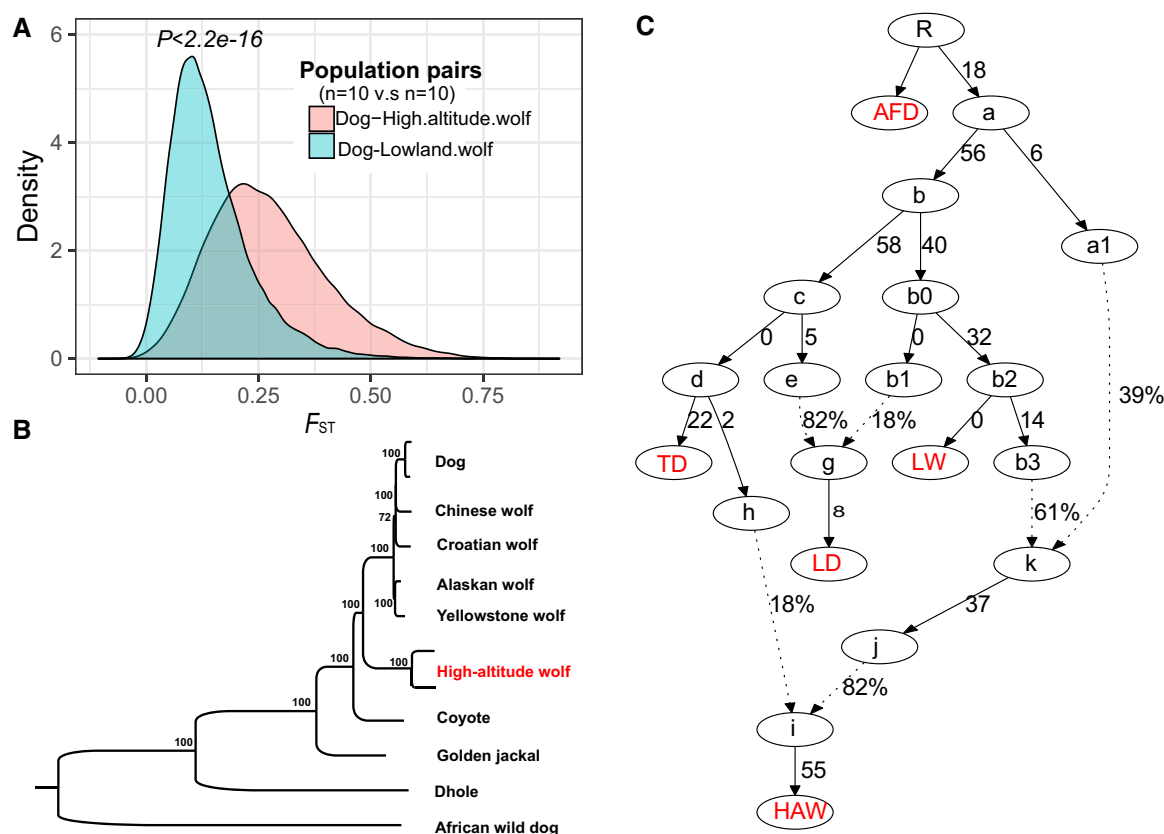
FIG. 2. High-altitude wolves likely carry ancestries from a diverged novel lineage. (*A*) Pairwise F_{ST} showed that high-altitude wolf (merging Tibetan wolf and Himalayan wolf) is more diverged with dog than is lowland. In this analysis, ten randomly selected samples for each population were used. Statistical significance was measured by Wilcoxon signed-rank test. (*B*) Best fitted model (without any *f4* outliers) inferred by *qpGraph* shows that high-altitude wolf could be modeled as a population carrying 39% of ancestry from a diverged lineage that is basal to dog and lowland wolves (one Chinese wolf and one European wolf). African wild dog (AFD) was used as an outgroup. Branch lengths are shown in units of $F_{ST} \times 1,000$ and dashed lines indicate inferred admixture events with admixture proportion reported beside the dashed lines. HAWs, high-altitude wolves; LDs, lowland dogs; TDs, Tibetan dogs; LWs, lowland wolves. Because this analysis assumed no gene flow between outgroup and inside population, and dhole and golden jackal have been reported to have gene flow with dog and gray wolf, we selected African wild dog to construct the admixture graph. (*C*) Neighbor-Joining tree constructed for genomic regions in high-altitude wolf showing of signal of "divergent" origin revealed by HMM with a cutoff of 0.8 posterior probability. Values at nodes indicate support in 1,000 bootstrap analyses.

Demographic History

To further explore the divergence between highland wolves and LWs, we next estimated the evolutionary trajectories of members of these lineages using pairwise sequentially Markovian coalescent (PSMC) (Li and Durbin 2011) and multiple sequentially Markovian coalescent (MSMC) (Schiffels and Durbin 2014). A previous nuclear genomic analysis reported that TW and North Chinese wolves formed a monophyletic clade outside the diversity of dogs (Fan et al. 2016). If this relationship is correct, we would expect TW to have a similar evolutionary trajectory as these LW. However, we find that the evolutionary trajectory of TW diverged from that of LW around ~100 ka, when TW begins to decrease in population size, whereas LW does not decline until later (fig. 3A).

To estimate the timing of the divergence between dogs and wolves, we leveraged MSMC's estimate of the relative cross coalescent rate (CCR) and used the rule of thumb of a 50% CCR as a rough estimate of the divergence time. We found the dog lineage separated from the LW lineage ~52.5 ka (fig. 3B). If TW and LW are monophyletic clades without

differential introgression or admixture, we would expect the divergence of TW and dog to be similar to that of LW and dog. However, we found that TW diverged from LW and dog ~68 and 81 ka, respectively (fig. 3B).

Simulation Analysis

Previously, the observed pattern of divergence between TW and Eurasian wolves by PSMC was suggested to arise from recent declines in the TW population size (Zhang et al. 2014; Fan et al. 2016). To test whether our results are a consequence of recent population decline in TW, we next performed a series of simulations in which two populations of varying sizes were assumed separated 30 ka (to reflect a commonly hypothesized time of divergence between dogs and wolves) (Skoglund et al. 2015; Frantz et al. 2016; Wang et al. 2016), after which one experiences a bottleneck of varying severity (supplementary fig. S12, Supplementary Material online). We then estimated evolutionary trajectories and divergence time from these simulated populations using both PSMC and MSMC and compared the results with the plots of the real
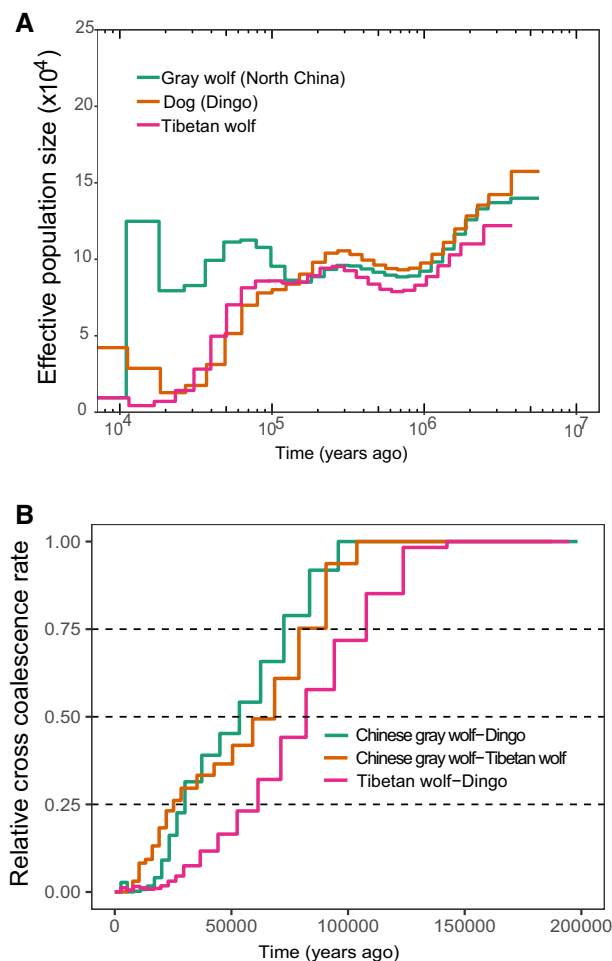
FIG. 3. Different evolutionary trajectory between Tibetan wolf and lowland gray wolf. (*A*) Dynamic changes of effective population size over past year inferred by PSMC. (*B*) Splitting time for population pairs estimated by MSMC. Sequencing depth for each of the genome is >20×. NCBI-SRA accession IDs for these genomes in this analysis are Dingo, SAMN03366709; Chinese gray wolf, SAMN03366711; and Tibetan wolf, SRS933420.

data (supplementary fig. S13, Supplementary Material online). None of these scenarios resulted in an older divergence time as depicted in the real data (fig. 3), suggesting that the decline in the TW population does not explain its divergent evolutionary trajectory.

Another potential explanation for the older estimate of the time of divergence between TW and dogs than between LW and dogs is that TW genomes contain some ancestry component from an older lineage. To test this hypothesis, we performed additional simulations in which two populations diverged 30 ka, with one received a varying level of gene flow from more deeply diverged population (we assumed they diverged 90 ka; supplementary fig. S12, Supplementary Material online). We found that gene flow from the diverged population influenced the estimated evolutionary trajectories from PSMC and MSMC in line with the amount (supplementary fig. S13, Supplementary Material online), mirroring the observed pattern in TW (fig. 3). This analysis supports the hypothesis that HAW genomes have a proportion of ancestry derived from a more deeply diverged wolf lineage.

Admixture between HAW and Dog Populations

Next, we looked for evidence of adaptive introgression using relative identical by descent (rIBD) (Bosse et al. 2014) and the locus-specific branch length (LSBL) statistic (Shriver et al. 2004). We identified three genes that exhibited the strongest signal of adaptive introgression between TD and HAW: *EPAS1*, *PRKCE*, and *TMEM247* (supplementary fig. S14, Supplementary Material online). All three occur in the same block on chromosome 10. In particular, the region containing *EPAS1* displayed the most significant signature of introgression (supported by "ABBA/BABA" [Green et al. 2010], *fd*-statistic [Martin et al. 2015], and RFMix [Maples et al. 2013]) and the most pronounced population differentiation (F_{ST}) between TD/HAW and LD/LW compared with the neighboring region (supplementary figs. S15 and S16, Supplementary Material online). Together, these results suggest that selection probably occurred at this gene.

EPAS1 in Present-Day HAW and Dog Is from an Older Lineage

Because HAWs are endemic to the Tibetan Plateau, previous analyses have suggested that the direction of introgression of the adaptive *EPAS1* allele was from HAW to TD after the introduction of domestic dogs to the high plateau (Miao et al. 2017; vonHoldt et al. 2017). However, we observe a higher frequency of the adaptive allele in TD (60 of 68; 88.2%) than in HW and TW (75%; [2 + 13]/[2 + 18]) (fig. 4A and supplementary fig. S17, Supplementary Material online).

We next constructed trees by dividing sequence across *EPAS1* and the neighboring regions spanning chromosome 10 into length of 100, 250, 500, and 1,000 kb. We found that the phylogeny across the region that includes *EPAS1* has a topology in which TD and HAW are basal to LW and other dogs (supplementary fig. S18, Supplementary Material online). This is in contrast to the topology for regions surrounding the *EPAS1* region, in which HAW is basal to all dogs and LW. Further, we constructed haplotype trees for *EPAS1* and its flanking regions by incorporating genomes from coyote, African wild dog, dhole, and North American wolf. A Neighbor-Joining tree showed that the adaptive *EPAS1* sequences carried by HW, TW, and TDs cluster as a distinct clade basal to Holarctic LW and dog (fig. 4B and supplementary fig. S19, Supplementary Material online). We note that *EPAS1* from HAW are assigned within the diversity of TD (supplementary fig. S19, Supplementary Material online), which may be due to limited sampling of HAW in this study and/or a recent population decline in HAW leading to a large diversity decrease in this region.

As shown in figure 4B, the placement of the HAW + TD clade varied by the length of *EPAS1* adjacent regions included in the alignment. The tree generated using data including the *EPAS1* gene and 50-kb regions flanking both sides suggests that unique adaptive sequences carried by HAW and TD are basal to the golden jackal. However, when the flanking regions are increased from 50 kb on both sides to 100 and 300 kb, this clade moves into the golden jackal lineage. Finally, when the
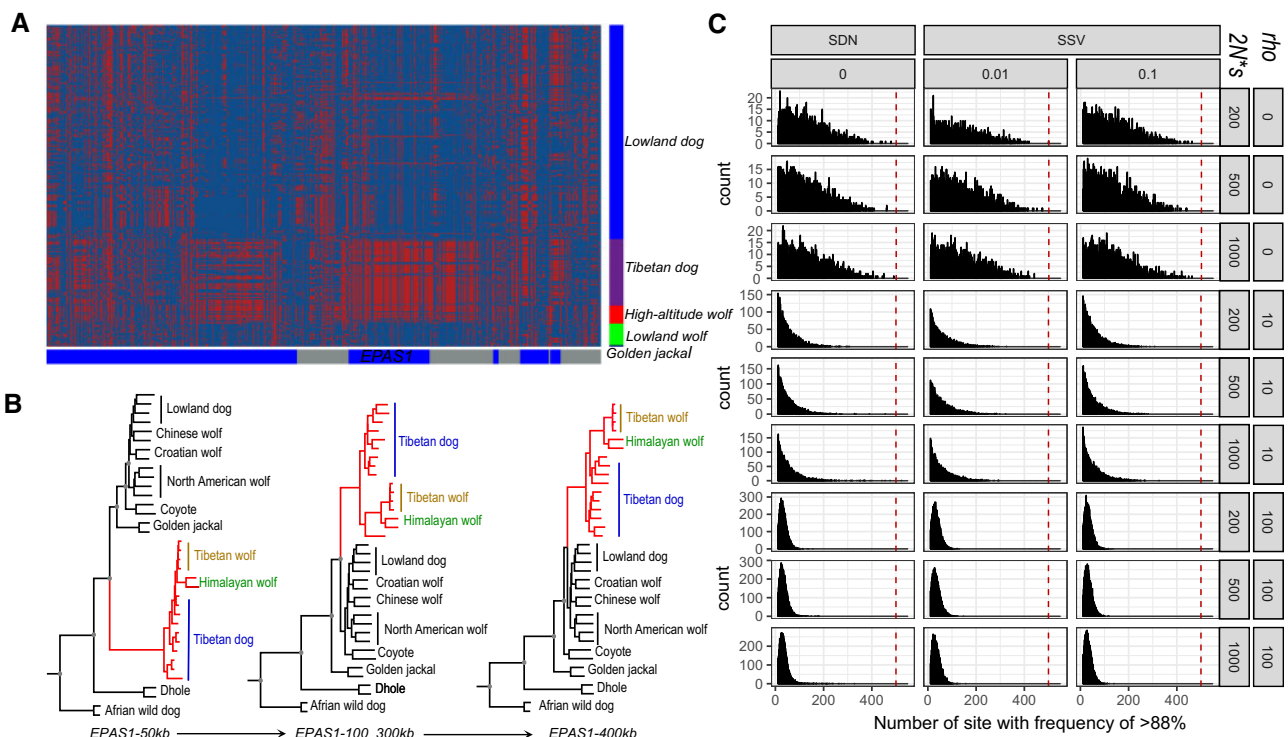
FIG. 4. Adaptive *EPAS1* in both high-altitude wolves and dogs was from a third divergent canid population. (*A*) Haplotype pattern for the region containing *EPAS1* (chr10:48,000,000–49,000,000). Each column indicates a polymorphic genomic location, each row indicates a phased haplotype (328 sequences in total), and the colored column on the right denotes the population identities of the individuals. Red signifies alternative alleles and blue the same allele relative as in the dog. The color panel below the heat map depicts gene (blue) and intergenic regions (gray). (*B*) Neighbor-Joining trees for sequences containing *EPAS1* and its 50 kb (left), 300 kb (middle), and 400 kb (right) flanking segment from both sides. Gray dots at nodes of tree indicate ≥99% support from 1,000 runs in bootstrap analysis. These trees are constructed using MEGA6 based on phased sequences, which are inferred using BEAGLE (V4.1). (*C*) Total number of highly differentiated sites observed under two assumed models (SDN and SSV with initial frequency of 1% and 5% for beneficial mutation) under varying strength of selection ($2N \times s = 200$, 500, and 1,000, where s is the selection coefficient for beneficial mutations) and recombination ($\rho = 0$, 10, and 100). Red lines indicate the number of observed highly differentiated site in Tibetan dog within the 143.5-kb block. Mutation rate used for simulation here is 4e-9. Across the simulations, the divergence and high frequencies for *EPAS1* would not be produced by realistic values of mutation and recombination from a wolf or dog ancestor, suggesting an origin outside this clade.

flanking regions are increased to 400 kb, the HAW + TD clade becomes sister to the wolf and dog clade, similarly to the tree constructed using the "divergent" sequences (fig. 2C). This is unsurprising that the differences in these topologies can probably be attributed to the combined effects of strong selection on *EPAS1* in HAWs and dogs and the breaking down of fragments spanning *EPAS1* by recombination. We hypothesize based on these data that the adaptive *EPAS1* alleles present in high-altitude canids introgressed neither from LWs nor from LDs but instead from a lineage that was ancestral to both.

Because *EPAS1* is known to evolve under strong positive selection, the emergence and sweep of adaptive mutations may confound phylogenetic inference. We identified a total of 1,242 variants in the region at chromosome 10 that contains *EPAS1* (chr10:48,551,410–48,694,966; 143.5 kb), of which 500 are at a frequency >88.2% in TD and <10% in LD (fig. 4A). We found no other region across the genome where so many linked variants reached similarly extreme differences in frequency between the two populations. To test whether selection acting on either de novo mutations (SDN) or on

standing variation (SSV) could lead to this pattern (many highly differentiated sites), we simulated regions of 143.5 kb by assuming the two possible scenarios under different strengths of selection and recombination rates. We found the number of highly divergent sites in the *EPAS1* region was significantly larger than those in the simulated data (fig. 4C), indicating that neither models of SDN nor models of SSV could explain the observed pattern in *EPAS1*. As the phylogeny including *EPAS1* has both wolves and dogs from the Tibetan plateau falling basal to dog and all other gray wolves (fig. 4B and supplementary figs. S18 and S19, Supplementary Material online), we hypothesize that the allele arose in an ancient wolf-like canid, which most likely occupied habitat at high altitude in the QTP.

Discussion

Tibetan Wolves and HWs Are Closely Related
In this study, we provide the first whole-genome sequence for an HW. We analyze this genome together with genomes of Tibetan wolves, other Eurasian gray wolves, and dogs and find that HW and Tibetan wolf are genetically closely related. This

result is consistent with previous studies of mitochondrial DNA, several nuclear loci, and microsatellites (Ersmark et al. 2016; Werhahn et al. 2017, 2018, 2020). Molecular clock-based dating of the mitochondrial phylogeny inferred a divergence time of around 779 ka between the clade including Tibetan wolves and HWs and the clade including Holarctic gray wolves and dogs, which is consistent with previous reports (Sharma et al. 2004; Werhahn et al. 2018).

Given that no geographic barrier exists between the reported ranges of the two HAWs across the Himalayas and Tibetan Plateau, it is not unexpected that TW and HW are closely genetically related. Our results are consistent with previous data showing that HAWs have a wide distribution, probably spanning regions across the Qinghai–Tibet Plateau and the Himalayas (Werhahn et al. 2018, 2020). PCA of nuclear genomic data shows that although TW and HW cluster closely together, our single HW individual falls outside the diversity of TW (supplementary fig. S1, Supplementary Material online). This is as expected for an individual with a distinct population history. Due to the challenges of collecting samples from small and highly endangered populations, we were only able to generate genomic data from a single HW individual; data from additional HWs will be necessary to understand finer levels of differentiation between HAW populations.

Wolves in the Tibetan Plateau Have Admixed Ancestries

All HAWs included in this study have admixed ancestries. In the mitochondrial phylogeny, HAWs are basal to the present-day Holarctic gray wolf complex. However, analyses of the nuclear genomes revealed a more complicated evolutionary history. Phylogenies inferred using genetic segments with length of 100, 250, and 500 kb revealed that three most frequent phylogenies that account for more than 90% of genome and the frequency of each topology differed slightly among the three data sets (supplementary fig. S8, Supplementary Material online). HAWs relatively more frequently cluster with low altitude Chinese wolves than with low altitude European wolves, probably due to more admixture between HAWs and nearby LWs (Werhahn et al. 2020). The topology showing HAWs basal to dogs and Eurasian wolves (similar to mitochondrial phylogeny) is more frequent compared with that showing Eurasian wolves basal to dog and HAWs, suggesting that highland-altitude wolf genomes include a certain amount of ancestries from a lineage that is more deeply diverged than Eurasian wolves. This observation is in accordance with estimations from MSMC, which suggests that TW and LW diverged ~68 ka, whereas TW and dog diverged ~81 ka (fig. 3). Thus, the previous conclusion showing TW sister to wolves from northern China (Fan et al. 2016; Wang et al. 2019) is an incomplete description of evolutionary history for these wolves complex.

Further analyses including TreeMix, $f4$ statistics, and admixture graph collectively indicate that the HAW genome comprises ancestry from at least three distinct sources (fig. 2): lowland gray wolves, dogs, and an as-yet-undescribed and possibly deeply divergent canid lineage, implying a

dynamically and reticulately genetic history among the canids of Tibetan Plateau. Reports of admixture between HAWs and feral dogs in their contacting zones support this conclusion (Hennelly et al. 2015), as does the extensive admixture known to occur among canid species worldwide (Wang et al. 2004; Koepfli et al. 2015; vonHoldt, Cahill, et al. 2016; Gopalakrishnan et al. 2018; Werhahn et al. 2020). The North American red wolf, for example, has mixed ancestry due to hybridization between coyote and North American wolf (vonHoldt, Cahill, et al. 2016; vonHoldt, Kays, et al. 2016), and the African golden wolf also has a hybrid origin (Gopalakrishnan et al. 2018). However, our data do not reveal the full history of HAWs; it remains unclear whether wolves of the Tibetan Plateau originated because of admixture between discrete lineages or are an old lineage that has experienced ongoing or irregular gene flow into their population throughout their evolutionary history. Additional whole genomes for HAWs, including wolves collected from across a broader geographic range and spanning different elevations of the Tibetan Plateau and the Himalayans, may help to resolve this question.

A history of admixture may also explain the distinct demographic trajectories estimated for Tibetan wolves compared with other wolves. Previously, two studies hypothesized that the distinct placement of Tibetan wolves in PCA and its separation from other wolves in PSMC may have been the consequence of severe habitat loss in QTP during Pleistocene glaciations and early human activities (Zhang et al. 2014; Fan et al. 2016). Our simulation analyses, however, indicate that population declines cannot explain such deep separation of Tibetan wolves from other wolves. Instead, we find the observed trajectories are better explained by a model in which genomes of present-day HAWs include a significant ancestry component derived from a more diverged lineage (fig. 3 and supplementary figs. S12 and S13, Supplementary Material online).

Adaptive *EPAS1* Alleles Were Likely Derived from an Unknown Source

Previous studies assumed, based on the timing of appearance of dogs on the Tibetan Plateau and the phylogenetic inference of Tibetan wolves as sister to Asian wolves, that the adaptive *EPAS1* alleles shared by present-day wolves and dogs on the Tibetan Plateau were introgressed into dogs from HAWs (Miao et al. 2017; vonHoldt et al. 2017). We find that these adaptive *EPAS1* sequences are deeply diverged from that found in LDs and LWs, and that the extent of this divergence cannot be explained by classic models of natural selection. The divergence depth and phylogeny depicting the adaptive *EPAS1* sequences places high-altitude dog and wolf as a distinct clade at the base of the LD and present gray wolves including North American wolves (fig. 4B), similar to the position of the "divergent" segments recovered by HMM (fig. 2C and supplementary fig. S9, Supplementary Material online). We suggest, therefore, that the adaptive *EPAS1* alleles were likely evolved in a presently unknown diverged lineage, and then persisted to the present day because of the significant benefits conferred by these alleles to high-altitude

canids. This would indicate that this diverged lineage probably has inhabited highlands and adapted to a life on the QTP. It is not unexpected since present ecological investigation on QTP is insufficient and it is possible that there may exist other canid species including those extinct in this area has not been reported. Additionally, previous studies also suggested that the existence of as-yet-undefined ("ghost") ancestries contributed the evolution of present wolf populations (Gopalakrishnan et al. 2018). Our study resembles an observations in other species, from example, Tibetan people's adaptive *EPAS1* was attributable to introgression from a Denisovan-like archaic hominid (Huerta-Sanchez et al. 2014); North Chinese pigs also have got adaptive genes from an unknown wild boar for surviving in cold environment (Ai et al. 2015). These collectively suggest that a convergent mechanism leveraged by these species to adapt to a harsh environment and highlight the important role of interspecies hybridization has played in shaping the adaptive evolution of species.

In conclusion, our genomic analyses suggest that Tibetan wolves and HWs are closely related, reconciling the long-standing debate regarding their taxonomic status. In addition, our results reveal that the main factor leading to inconsistent phylogenetic placement of Tibetan wolves in past studies is the highly admixed ancestry of HAWs, which contain genomic signatures of at least three origins: a deeply diverged as-yet-undefined population, a more recent population of gray wolves, and dogs. The genomic legacy of the ghost population includes the adaptive *EPAS1* allele, which is found in HAWs and dogs and allows them to persist at high altitudes. However, due to recombination and admixture, the evolutionary history for this ghost population including phylogenic position, to what extent it diverged with present dog and wolf, as well as when and how hybridization occurred and shaped genome of present wolf and dog on the Tibetan Plateau are not fully resolved. Further studies by incorporating genomes from more samples including fossils covering wider their ranges across the QTP and the Himalayas would be necessary to test our findings and to deep investigate how adaptive *EPAS1* evolved among Canid in this area. Overall, our analyses provide a result and case for future investigating the role of hybridization in shaping the evolution of Tibetan canids and other wolf-like canids.

Materials and Methods

Genomic Sequencing Data

We collected 19 samples for whole-genome sequencing: two TWs, three Tibetan Terriers, six Tibetan Mastiffs, six Tibetan indigenous dogs, one Tibetan Kyi Apso, and one HW. Genome sequencing was performed on the Illumina Genomic sequencing platform targeting genome coverage ranging from ∼6.4- to 10.6-fold. HW genome was sequenced at Xcelris Labs Ltd, Ahmedabad, Gujarat in India. By incorporating genomic data for dogs and gray wolves from previous studies (Auton et al. 2013; Wang et al. 2013; Freedman et al. 2014; Gou et al. 2014; Li et al. 2014; Zhang et al. 2014; Fan et al. 2016), we compiled a data set including 109 LDs, 11 lowland

gray wolves (LWs), 9 TWs, 34 TDs, 1 HW, 1 golden jackal, and 1 dhole. Detailed information for these samples is shown in supplementary tables S1 and S2, Supplementary Material online.

Calling SNPs

We used the standard Genome Analysis Toolkit (version 2.6-4, GATK) pipeline (McKenna et al. 2010) for calling SNPs as previously described (Wang et al. 2017). Raw genome sequencing reads were filtered by removing low-quality bases using Btrim (Kong 2011). Reads meeting quality filtration thresholds were mapped to the dog reference genome (CanFam3.1, 2.41 Gb; available at: https://www.ncbi.nlm.nih.gov/assembly/GCF_000002285.3/; last accessed May 13, 2020) by bwa (version: 0.7.5a-r405) using the "mem" algorithm (Li 2014) with default settings except the "-t 8 -M" options. A series of postprocesses were employed to process and filter the BAM format alignment file. These include alignment position sorting, duplicated reads marking, local realignment, and base quality recalibration. These procedures were carried out using the available tools in Picard (version 1.56; http://picard.sourceforge.net) and GATK packages (McKenna et al. 2010), including SortSam, MarkDuplicates, RealignerTargetCreator, IndelRealigner, and BaseRecalibrator. The "depth" function in samtools software (version: 1.3.1) (Li et al. 2009) was used to calculate genome sequencing depth for each sample. SNPs were called for all subjects together using UnifiedGenotyper function in GATK. To increase the reliability of the data, SNPs were further filtered using VariantFiltration command in GATK with parameters "QUAL < 40.0 MQ < 25.0 MQ0 >= 4 && ((MQ0/(1.0*DP)) > 0.1) cluster 3 -window 10."

Additionally, to infer the phylogeny of adaptive *EPAS1* carried by TD, genomes for one coyote (*Canis latrans*; SRA accession IDs: SRR1518485, SRR1518486, SRR1518487, and SRR3574870), one African wild dogs (*Lycaon pictus*; SRA accession ID: SRR2971425), two North American wolves (*Canis lupus*; SRA accession IDs: SRR8066602 andSRS661497), and one dhole (*Cuon alpinus*; SRR2827618–SRR2827626) were also incorporated in our analyses. Genotypes for these samples were called using the pipeline described above.

Phylogeny, PCA, and Admixture Clustering Analyses

To infer the phylogenic relationship between HW and TW and with Holarctic, we divided genomes into 100-, 250-, and 500-kb nonoverlapping segments, respectively. We excluded SNPs with missing rate >10% and segments with fewer than 100 SNPs. Gene trees were constructed using TreeMix (Pickrell and Pritchard 2012) for each segment. Because only one HW is included in this study, we used one TW and one dog with dhole as outgroup. Next, to estimate the evolutionary placement of HAWs among other wolves and dogs, we merged HW and nine TWs into a single HAW group and assigned other wolves and dogs as LW, LD, and TD, with dhole as outgroup. We estimated gene trees using 100-, 250-, and 500-kb segments across genome as mentioned above. We counted the number of tree topologies and painted tree distributions on all autosomes using APE (Popescu

et al. 2012) and karyoploteR (Gel and Serra 2017) in R (https://www.r-project.org/), respectively. Then, we used ASTRAL-II v4.11.2 (Mirarab and Warnow 2015) to analyze and summarize these gene trees and infer a coalescent-based consensus species tree. Nodal support (bootstrap) was estimated based on 1,000 resampled gene trees with options "-gene-only –r 1000." The consensus phylogenic trees were viewed and edited with MEGA6 (Tamura et al. 2013). DiscoVista (Sayyari et al. 2018) is also used to analyze our gene trees and measure the frequency for bipartitions. We performed PCA using GCTA (Yang et al. 2011) and genetic clustering analysis using ADMIXTURE (Alexander et al. 2009) based on the data set pruned using PLINK (Chang et al. 2015) with parameters "–indep-pairwise 50 10 0.1."

PSMC and MSMC Analyses

We used PSMC to infer population size (Ne) changes along with the time based a single diploid genome (Li and Durbin 2011). PSMC requires higher coverage to ensure the accurate calling of heterozygotes (Nadachowska-Brzyska et al. 2016), so we only used individuals with coverage >20-fold. Consensus sequences for each individual were called using samtools (version : 1.3.1) with "mpileup" command (Li et al. 2009). Loci with less than one-third or more than two times of average read depth were excluded, and those with consensus quality below 20 were also removed. We ran PSMC with parameters "psmc -N25 -t15 -r5 -p $4 + 25*2 + 4 + 6$." We then used the MSMC method to infer the divergence time for dog and wolf population pairs (Schiffels and Durbin 2014). Genotypes for all dogs and wolves were phased together using Beagle V.4.1 (Browning BL and Browning SR 2016). Fifty percent of relative CCR was to define the splitting time (Malaspinas et al. 2016). In our runs, four haplotypes were used, as was a generation time (g) of 3 years and a mutation rate (μ) of 4×10^{-9} substitutions per site per generation (Frantz et al. 2016). This rate was comparable to a recent estimation based on pedigree.

We ran simulations to test whether a decrease in population size or mosaic ancestry from an older lineage could generate a similar pattern produced by PSMC and MSMC in Tibetan wolf. We used two models: 1) two populations split 30 ka with one undergoing varying levels of population decline and 2) two populations split 30 ka with one undergoing varying levels of migration from a diverged lineage (assumed to have diverged 90 ka). We simulated a total of 20 chromosomes with length of 30 Mb, corresponding to the models by using msms (Ewing and Hermisson 2010). PSMC and MSMC were run using the simulated data based on same procedures above and compared with real data.

Scanning for Adaptive Introgressed Genomic Fragments

To retrieve the genomic regions putatively introgressed between HAWs and TD, we used the method described in a previous study that investigated introgressed regions between European and Asian pigs (Bosse et al. 2014) and the LSBL statistic (Shriver et al. 2004). First, we used Beagle v.4.1 to impute and phase the genotypes (Browning BL and Browning SR 2016). Second, we counted and computed

frequency of haplotypes (fIBD) shared between TD and HAW and HAW and LW populations. The relative shared haplotype of HAW in TD relative to LW was defined as

$$\text{rIBD } (\text{HAW}; \text{TD}, \text{LW}) = \text{fIBD}(\text{HAW} - \text{TD}) - \text{fIBD}(\text{HAW} - \text{LW}).$$

Sliding-widow analyses were performed to count haplotype frequency for each bin (window size of 10 kb with 5 kb step size). As the total number of pairwise comparisons differed between the groups, the rIBD values were Z-transformed as follows:

$$\text{ZrIBD} = (\text{rIBD}–\mu)/\sigma\text{rIBD},$$

where μ is the genome-wide mean for rIBD and σ is standard error. The threshold for extreme IBD with the TD from HAW compared with LW was set to two SD from the mean in the far-right tail of the distribution in different regions of the genome. Additionally, we also used robust local-ancestry inference (RFMix) (Maples et al. 2013), D-statistic (ABBA/ABAB statistic) (Green et al. 2010), and fd-statistic (Martin et al. 2015) to verify the signature of the identified introgressed segments. RFMIX was run using TD as target population and LW and HAW as the presumed source populations. D- and fd-statistics were performed in the form of D/fd (Golden jackal, TD; HAW, LW). Although these approaches can characterize admixture events, they could not define the direction of admixture. We therefore constructed a phylogenic tree using haplotypes of the EPAS1 gene region using MEGA 6.0 (Tamura et al. 2013).

Around ~0.695% of the genome sequences spanning 15.3 Mb per individual supported by rIBD showed putative signals of introgression between HAW and TD (supplementary fig. S14, Supplementary Material online). To test whether genes in these introgressed region subjecting selection, we used LSBL statistic (Shriver et al. 2004) to scan whole genome for signal of selection. First, we used vcftools (Danecek et al. 2011) to compute Weir and Cockerham's pairwise F_{ST} by each site for each pair of populations. Second, LSBL, measured as LSBL(A; B, C) = $(F_{ST}[AB] + F_{ST}[AC] – F_{ST}[BC])/2$, was used to investigate population differentiation for HAW and TD compared with LD and LW. We calculated LSBL for "TD; LD, LW" and "HAW; LW, LD" by setting a 50-kb window with 25-kb stepwise increments in each combination. We employed a common empirical 99.9th percentile as the cutoff to retrieve candidate selective sweeps, which were then annotated using the variant effect predictor (McLaren et al. 2016). Genes were supported by both rIBD and LSBL are considered adaptively introgressed.

Estimating Splitting and Admixture Graphs

To investigate the admixture and splitting among dog and wolf populations, we performed $f4$ statistics (Patterson et al. 2012) and TreeMix analyses (Pickrell and Pritchard 2012). We ran plink2treemix.py (available in TreeMix v1.2 package) to count the allele frequency for each site. The treemix program (available in TreeMix v1.2 package) was leveraged to infer population admixture networks by allowing 0–5 migration

events, with parameters "-k 1000 -global -root dhole." $f4$ statistics was computed using *fourpop* (available in TreeMix v1.2 package). Wolves were grouped as HAW and LW, and dogs were grouped as LD and TD.

We also used the *qpGraph* program available in the ADMIXTOOLS package (Patterson et al. 2012) to model admixture graph by reconciling the split and migration among these populations. Because of large computational costing (memory), we selected a few individuals for the analysis, including one African wild dogs (used as outgroup; because both Golden jackal and dhole have been reported having admixture with gray wolf [vonHoldt, Cahill, et al. 2016; vonHoldt, Kays, et al. 2016]), five TDs, six LDs, two LWs (one from Europe and another from North China), and three HAWs. Estimation was performed based on LD pruned the data set by PLINK (Chang et al. 2015) with parameters "–indep-pairwise 50 10 0.2."

Analyzing Genomic Regions Showing Signature of "Divergent" Origin

We used HMM (Skov et al. 2018) to excavate genomic regions in HAW showing signature of "divergent" lineage origin. HMM is the approach that detects genetic segments from individual genomes showing archaic origin by comparing with unadmixed populations without needing an archaic source reference genome. A window size of 1,000 bp was used to generate callability files (by MakeMask.py) and estimate the average mutation rate. The mount of regions showing signature of diverged origin varied by the cutoff for posterior probability, it is hard to choose a proper threshold value. Here, we used 0.8 of posterior probability as author suggested showing a good tradeoff between precision and sensitivity and a stricter one-0.9 of posterior probability. Because the overlapping of segment with diverged origin is limited. To reconcile more variants for phylogeny construction, we only choose two HAWs with high coverage (SRA accession IDs: SRS933415 and SRS933420) for the analysis. Under posterior probability of 0.8, 45.3M and 48.9M of segments of diverged origin were recovered in SRS933415 and SRS933420, respectively, with 29.8M overlapped. Under posterior probability of 0.9, 19.0M and 20.0M were recovered in SRS933415 and SRS933420, respectively, with 12.0M shared with each other.

To get a conserved result, regions showing diverged origin shared by the two HAWs (i.e., the 29.8M and 12.0M segments based on 0.8 and 0.9 of posterior probability, respectively) were used to perform phylogenic analysis. We also included genomes from North American wolf, coyote, dhole, African wild dog, and golden jackal to construct Neighbor-Joining tree using MEGA6 (Tamura et al. 2013). One thousand bootstraps were analyzed.

Inferring Population Size with fastsimcoal2

To estimated population size for lowland and TDs, we used joint site frequency spectrum method implemented in fastsimcoal2 (Excoffier et al. 2013). We considered introgression between dog and wolf populations (Ostrander et al. 2017).The unfolded site frequency spectrum was generated

using a modified script from *dadi* (Gutenkunst et al. 2009). For each model, 100 repeats with varying starting points were performed to ensure convergence. The best fit of each run was kept based on the maximum likelihoods calculated for independent repeats. In our estimation, we used a mutation rate of 4e-9 substitution/site/year, and a generation time of 3 years (Frantz et al. 2016). Estimation was ran with the following parameters: "-n 100000 -N 100000 -d -M 0.001 -l 10 -L40 -q -c 12."

Simulation and SDN and on SSV

We used msms program (Ewing and Hermisson 2010) to simulate genomic variants considering SDN and SSV, assuming consistency in population size, recombination rate, and mutation rate. In TD, we found that most variants with frequency of 60/68 and 61/68 that were highly differentiated with LD (<10%). So, we performed simulations conditioning on 1,242 separating sites observed in a 143.5-kb haplotype block. Effective population size for TD based on fastsimcoal2 was estimated to be ∼14,000. We presumed three varying strength degrees of selection for the homozygote ($2Ns = 200$, 500, 1.000, where s is the selection coefficient for beneficial mutations; s for the heterozygote is half that for the homozygote) and a varying recombination rate ($4Nr = 0$, 10, and 100; mutation rate used here is 4e-9 per substitution per generation). For simulating selection, we assumed the initial frequency of 0%, 1%, and 10% for the beneficial mutation when selection started. A total of 27 runs of simulation were performed. For a rough comparison of the number of beneficial mutations observed and simulated data under different models and parameters, we counted the number of beneficial mutations with frequency equal to or >60/68. Simulation for each run was performed 10,000 independent times.

Testing the Probability of the 143.5-kb Haplotype Being Result of Incomplete Lineage Sorting

Following the method described elsewhere (Huerta-Sanchez et al. 2014), we computed the probability that the adaptive haplotype with a length of 143.5 kb is shared by TD and HAW as a result of incomplete lineage sorting. The expected length (L) of a shared sequence for dog and wolf is $1/(t \times r)$ since divergent length of t, whereas r is the recombination rate per generation time per site. The probability of seeing a sequence with a length of L shared by dog and wolf follows a Gamma distribution with shape $= 2$, and a rate parameter lambda of $1/L$. Here, we used the conserved splitting age of 40,000 years as estimated by MSMC. We used a recombination rate (r) of 0.78 cM/Mb (Miao et al. 2017). Thus, $L = 1/(7.8e-9 \times 80,000/3) = 4,807.69$ bp. The probability of observing a sequence with length of 143.5 kb is (1-GammaCDF [143500,2, 1/4807.69] = 3.36e-12, which we calculated using R (https://www.r-project.org/).

Mitochondrial Phylogeny Construction

To generate mitochondrial sequences, we performed variant calling and filtering in GATK as described above for the nuclear analysis. We generated a consensus mitochondrial

genome for each sample using a custom-made Python script (available from the authors on request). We downloaded whole mitochondrial genomes and mitochondrial segments including the control region and cytochrome *b* for coyote, African wild dog, dhole, African wolf, and Tibetan wolves (GenBank accession IDs are provided in supplementary figs. S6 and S7, Supplementary Material online). We aligned these using MUSCLE-3.8.31 (Edgar 2004) and checked alignments by eye. We estimated a maximum-likelihood tree using RAxML-8.1.17 (Stamatakis 2006) with 1,000 pseudoreplicates. The command line for running the analysis was: raxmlHPC -x 12345 -# 1000 -p 321 -f b -m GTRGAMMAI -T 20 -s input.fas -o output.tre. We viewed and edited the resulting phylogeny using MEGA6 (Tamura et al. 2013).

Estimating the Divergence Time for Mitochondrial Sequences

To estimate the timing of divergence among mitochondrial lineages, we de novo assembled mitochondrial genomes for HAW using NOVOPlasty2.7.2 (Dierckxsens et al. 2017) using with dog mitochondrial genome (GenBank accession: NC_002008.4) as a seed sequence. Three K-mers—25, 31, and 39—were used to perform the assembly independently. We aligned these assemblies to dog and wolf mitochondrial genomes and checked the alignments by eye. Apart from the control region, which was assembled incompletely, six HAW mitogenomes were successfully recovered with all three K-mers, which performed nearly equally well. We selected the assembly that was generated using a K-mer of 31 for subsequent analysis.

We downloaded mitochondrial genomes for coyote, African dog, dhole, African wolf, and other gray wolves and dogs from NCBI (GenBank accession IDs are provided in supplementary figs. S6 and S7, Supplementary Material online) and analyzed these with our new assembles. We retrieved the 13 protein-coding genes, 12S and 16S rRNAs, and tRNA genes from the mitochondrial genomes and defined them as separate partitions for phylogenic inference and divergence estimation. Each partition was aligned independently using MUSCLE-3.8.31 (Edgar 2004) and alignments were checked by eye. We estimated the best-fitting DNA substitution model using Akaike information criteria as implemented in jModelTest-2.1.10 (Darriba et al. 2012). The DNA substitution model for protein-coding gene is HKY + I + G, for rRNA is GTR + I + G, and for tRNA is HKY + I. We then used BEAST v1.10.4 (Drummond and Rambaut 2007) to estimate tree topology and divergence times under a strict molecular clock model. As a first step, we used BEAUti program (available in BEAST v1.10.4 package) to set up the XML with the following parameters: unlinked substitution models using partitions (protein-coding gene was further partitioned into first + second codons and third codons) and substitution models described above, link trees enabled, strict clock, and tree prior specified as Yule process of speciation. Because the fossil record is incomplete in canid species with some reported fossil ages spanning large intervals, we used divergence time estimates from a recent study (Koepfli et al. 2015) as softbound priors on the following divergences: 1) a rooting age of 4.02

Ma for African wild dog + wolf and 2) 1.28 Ma for coyote + wolf. The calibration priors were approximated with normal distributions: mean = 4.01 and stdev = 0.4 (95% HPD: 3.36–4.68) for prior 1, and mean = 1.28 and stdev = 0.2 (95% HPD: 0.95–1.609) for prior 2. We ran each MCMC chain for 200,000,000 steps, logging trees and model parameters every 2,000 steps, and discarded the first 10% as burn-in. Posterior distributions of tree likelihoods and other estimated parameters were analyzed using Tracer v1.7.1 (https://github.com/beast-dev/tracer/releases) to ensure estimated sample size (ESS) for each statistic >500. TreeAnnotator (available in BEAST v1.10.4 program package) was used to summarize the posterior sample of trees with setting: burn-in of 100 states and posterior probability of 0.5. The estimated tree with node heights was visualized using FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/).

Data Availability

Assembled mitochondrial sequences and alignment and genomic genotype files (in plink format) were deposited in Dryad with the accession doi:10.5061/dryad.0vt4b8gvt.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Author Contributions

D.-D.W., M.-S.W., and Y.-P.Z. conceived the project and designed the research. R.N., B.S., Y.J., and D.-D.W. supervised the analysis. M.-S.W., S.W., Y.L., N.O.O., and J.-F.S. performed the analysis. M.T. and Y.J. provided and coordinated genome sequencing of Himalayan wolf sample in India. M.-S.W. and Y.L. prepared the DNA samples. M.-S.W., M.T., and B.S. drafted the manuscript. B.S., M.-S.W., Y.J., R.N., M.T., and D.-D.W. revised the manuscript. All authors read and improved the manuscript.

References

Ai H, Fang X, Yang B, Huang Z, Chen H, Mao L, Zhang F, Zhang L, Cui L, He W, et al. 2015. Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nat Genet.* 47(3):217–225.

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19(9):1655–1664.

Anderson TM, vonHoldt BM, Candille SI, Musiani M, Greco C, Stahler DR, Smith DW, Padhukasahasram B, Randi E, Leonard JA, et al. 2009. Molecular and evolutionary history of melanism in North American gray wolves. *Science* 323(5919):1339–1343.

Auton A, Rui Li Y, Kidd J, Oliveira K, Nadel J, Holloway JK, Hayward JJ, Cohen PE, Greally JM, Wang J, et al. 2013. Genetic recombination is targeted towards gene promoter regions in dogs. *PLoS Genet.* 9(12):e1003984.

Bosse M, Megens HJ, Frantz LA, Madsen O, Larson G, Paudel Y, Duijvesteijn N, Harlizius B, Hagemeijer Y, Crooijmans RP, et al. 2014. Genomic analysis reveals selection for Asian genes in European pigs following human-mediated introgression. *Nat Commun.* 5(1):4392.

Browning BL, Browning SR. 2016. Genotype imputation with millions of reference samples. *Am J Hum Genet.* 98(1):116–126.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4(1):7.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27(15):2156–2158.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods.* 9(8):772–772.

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45(4):e18.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol.* 7(1):214.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32(5):1792–1797.

Ersmark E, Klutsch CFC, Chan YL, Sinding MHS, Fain SR, Illarionova NA, Oskarsson M, Uhlen M, Zhang YP, Dalen L, et al. 2016. From the past to the present: wolf phylogeography and demographic history based on the mitochondrial control region. *Front Ecol Evol.* 4(134).

Ewing G, Hermisson J. 2010. MSMS: a coalescent simulation program including recombination, demographic structure and selection at a single locus. *Bioinformatics* 26(16):2064–2065.

Excoffier L, Dupanloup I, Huerta-Sanchez E, Sousa VC, Foll M. 2013. Robust demographic inference from genomic and SNP data. *PLoS Genet.* 9(10):e1003905.

Fan Z, Silva P, Gronau I, Wang S, Armero AS, Schweizer RM, Ramirez O, Pollinger J, Galaverni M, Ortega Del-Vecchyo D, et al. 2016. Worldwide patterns of genomic variation and admixture in gray wolves. *Genome Res.* 26(2):163–173.

Frantz LA, Mullin VE, Pionnier-Capitan M, Lebrasseur O, Ollivier M, Perri A, Linderholm A, Mattiangeli V, Teasdale MD, Dimopoulos EA, et al. 2016. Genomic and archaeological evidence suggest a dual origin of domestic dogs. *Science* 352(6290):1228–1231.

Freedman AH, Gronau I, Schweizer RM, Ortega-Del Vecchyo D, Han E, Silva PM, Galaverni M, Fan Z, Marx P, Lorente-Galdos B, et al. 2014. Genome sequencing highlights the dynamic early history of dogs. *PLoS Genet.* 10(1):e1004016.

Gel B, Serra E. 2017. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* 33(19):3088–3090.

Gittelman RM, Schraiber JG, Vernot B, Mikacenic C, Wurfel MM, Akey JM. 2016. Archaic hominin admixture facilitated adaptation to Out-of-Africa environments. *Curr Biol.* 26(24):3375–3382.

Gopalakrishnan S, Sinding MS, Ramos-Madrigal J, Niemann J, Samaniego Castruita JA, Vieira FG, Caroe C, Montero MM, Kuderna L, Serres A, et al. 2018. Interspecific gene flow shaped the evolution of the genus *Canis*. *Curr Biol.* 28(21):3441–3449.e3445.

Gou X, Wang Z, Li N, Qiu F, Xu Z, Yan D, Yang S, Jia J, Kong X, Wei Z, et al. 2014. Whole-genome sequencing of six dog breeds from continuous altitudes reveals adaptation to high-altitude hypoxia. *Genome Res.* 24(8):1308–1315.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, et al. 2010. A draft sequence of the Neandertal genome. *Science* 328(5979):710–722.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* 5(10):e1000695.

Hennelly LM, Habib B, Lyngdoh S. 2015. Himalayan wolf and feral dog displaying mating behaviour in Spiti Valley, India, and potential conservation threats from sympatric feral dogs. *Canid Biol Conserv.* 18:27–30.

Hodgson B. 1847. Description of the wild ass and wolf of Tibet, with illustrations. *Calcutta J Nat Hist.* 7:469–477.

Huerta-Sanchez E, Jin X, Asan Bianba Z, Peter BM, Vinckenbosch N, Liang Y, Yi X, He M, Somel M, et al. 2014. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* 512(7513):194–197.

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

Kong Y. 2011. Btrim: a fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. *Genomics.* 98(2):152–153.

Lawson DJ, van Dorp L, Falush D. 2018. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nat Commun.* 9(1):3258.

Li H. 2014. Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics* 30(20):2843–2851.

Li H, Durbin R. 2011. Inference of human population history from individual whole-genome sequences. *Nature* 475(7357):493–496.

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

Li Y, Wang GD, Wang MS, Irwin DM, Wu DD, Zhang YP. 2014. Domestication of the dog from the wolf was promoted by enhanced excitatory synaptic plasticity: a hypothesis. *Genome Biol Evol.* 6(11):3115–3121.

Malaspinas AS, Westaway MC, Muller C, Sousa VC, Lao O, Alves I, Bergstrom A, Athanasiadis G, Cheng JY, Crawford JE, et al. 2016. A genomic history of Aboriginal Australia. *Nature* 538(7624):207–214.

Maples BK, Gravel S, Kenny EE, Bustamante CD. 2013. RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. *Am J Hum Genet.* 93(2):278–288.

Martin SH, Davey JW, Jiggins CD. 2015. Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Mol Biol Evol.* 32(1):244–257.

McCoy RC, Wakefield J, Akey JM. 2017. Impacts of Neanderthal-introgressed sequences on the landscape of human gene expression. *Cell* 168(5):916–927.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20(9):1297–1303.

McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F. 2016. The Ensembl variant effect predictor. *Genome Biol.* 17(1):122.

Miao B, Wang Z, Li Y. 2017. Genomic analysis reveals hypoxia adaptation in the Tibetan mastiff by introgression of the gray wolf from the Tibetan Plateau. *Mol Biol Evol.* 34(3):734–743.

Mirarab S, Warnow T. 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* 31(12):i44–i52.

Nadachowska-Brzyska K, Burri R, Smeds L, Ellegren H. 2016. PSMC analysis of effective population sizes in molecular ecology and its application to black-and-white *Ficedula* flycatchers. *Mol Ecol.* 25(5):1058–1072.

Ostrander EA, Wayne RK, Freedman AH, Davis BW. 2017. Demographic history, selection and functional diversity of the canine genome. *Nat Rev Genet.* 18(12):705–720.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. *Genetics* 192(3):1065–1093.

Pickrell JK, Pritchard JK. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* 8(11):e1002967.

Popescu AA, Huber KT, Paradis E. 2012. ape 3.0: new tools for distance-based phylogenetics and evolutionary analysis in R. *Bioinformatics* 28(11):1536–1537.

Prufer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, Heinze A, Renaud G, Sudmant PH, de Filippo C, et al. 2014. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505(7481):43–49.

Sayyari E, Whitfield JB, Mirarab S. 2018. DiscoVista: interpretable visualizations of gene tree discordance. *Mol Phylogenet Evol.* 122:110–115.

Schiffels S, Durbin R. 2014. Inferring human population size and separation history from multiple genome sequences. *Nat Genet.* 46(8):919–925.

Schraiber JG, Akey JM. 2015. Methods and models for unravelling human evolutionary history. *Nat Rev Genet.* 16(12):727–740.

Sharma DK, Maldonado JE, Jhala YV, Fleischer RC. 2004. Ancient wolf lineages in India. *Proc Biol Sci.* 271(Suppl 3):S1–S4.

Shriver MD, Kennedy GC, Parra EJ, Lawson HA, Sonpar V, Huang J, Akey JM, Jones KW. 2004. The genomic distribution of population substructure in four populations using 8,525 autosomal SNPs. *Hum Genomics.* 1(4):274–286.

Shrotriya S, Lyngdoh S, Habib B. 2012. Wolves in Trans-Himalayas: 165 years of taxonomic confusion. *Curr Sci.* 103(8):885–887.

Sinding MS, Gopalakrishan S, Vieira FG, Samaniego Castruita JA, Raundrup K, Heide Jorgensen MP, Meldgaard M, Petersen B, Sicheritz-Ponten T, Mikkelsen JB, et al. 2018. Population genomics of grey wolves and wolf-like canids in North America. *PLoS Genet.* 14(11):e1007745.

Skoglund P, Ersmark E, Palkopoulou E, Dalen L. 2015. Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. *Curr Biol.* 25(11):1515–1519.

Skov L, Hui R, Shchur V, Hobolth A, Scally A, Schierup MH, Durbin R. 2018. Detecting archaic introgression using an unadmixed outgroup. *PLoS Genet.* 14(9):e1007641.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22(21):2688–2690.

Supple MA, Shapiro B. 2018. Conservation of biodiversity in the genomics era. *Genome Biol.* 19(1):131.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol.* 30(12):2725–2729.

Taylor SA, Larson EL. 2019. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nat Ecol Evol.* 3(2):170–177.

vonHoldt B, Fan Z, Ortega-Del Vecchyo D, Wayne RK. 2017. *EPAS1* variants in high altitude Tibetan wolves were selectively introgressed into highland dogs. *PeerJ* 5:e3522.

vonHoldt BM, Cahill JA, Fan Z, Gronau I, Robinson J, Pollinger JP, Shapiro B, Wall J, Wayne RK. 2016. Whole-genome sequence analysis shows that two endemic species of North American wolf are admixtures of the coyote and gray wolf. *Sci Adv.* 2(7):e1501714.

vonHoldt BM, Kays R, Pollinger JP, Wayne RK. 2016. Admixture mapping identifies introgressed genomic regions in North American canids. *Mol Ecol.* 25(11):2443–2453.

Wang GD, Zhai W, Yang HC, Fan RX, Cao X, Zhong L, Wang L, Liu F, Wu H, Cheng LG, et al. 2013. The genomics of selection in dogs and the parallel evolution between dogs and humans. *Nat Commun.* 4(1):1860.

Wang GD, Zhai W, Yang HC, Wang L, Zhong L, Liu YH, Fan RX, Yin TT, Zhu CL, Poyarkov AD, et al. 2016. Out of southern East Asia: the natural history of domestic dogs across the world. *Cell Res.* 26(1):21–33.

Wang GD, Zhang M, Wang X, Yang MA, Cao P, Liu F, Lu H, Feng X, Skoglund P, Wang L, et al. 2019. Genomic approaches reveal an endemic subpopulation of gray wolves in Southern China. *iScience* 20:110–118.

Wang M-S, Otecko NO, Wang S, Wu D-D, Yang M-M, Xu Y-L, Murphy RW, Peng M-S, Zhang Y-P. 2017. An evolutionary genomic perspective on the breeding of dwarf chickens. *Mol Biol Evol.* 34(12):3081–3088.

Wang X, Tedford RH, Van Valkenburgh B, RK W. 2004. Evolutionary history, molecular systematics, and evolutionary ecology of *Canidae*. In: Macdonald DW, Sillero-Zubiri C, editors. Biology and conservation of wild canids. Oxford (United Kingdom): Oxford University Press. p. 39–54.

Werhahn G, Liu Y, Meng Y, Cheng C, Lu Z, Atzeni L, Deng Z, Kun S, Shao X, Lu Q. Forthcoming 2020. Himalayan wolf distribution and admixture based on multiple genetic markers. *J Biogeogr.* 00:1–14.

Werhahn G, Senn H, Ghazali M, Karmacharya D, Sherchan AM, Joshi J, Kusi N, Lopez-Bao JV, Rosen T, Kachel S, et al. 2018. The unique genetic adaptation of the Himalayan wolf to high-altitudes and consequences for conservation. *Global Ecol Conserv.* 16:e00455.

Werhahn G, Senn H, Kaden J, Joshi J, Bhattarai S, Kusi N, Sillero-Zubiri C, Macdonald DW. 2017. Phylogenetic evidence for the ancient Himalayan wolf: towards a clarification of its taxonomic status based on genetic sampling from western Nepal. *R Soc Open Sci.* 4(6):170186.

Yang J, Lee SH, Goddard ME, Visscher PM. 2011. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet.* 88(1):76–82.

Zhang SJ, Wang GD, Ma P, Zhang LL, Yin TT, Liu YH, Otecko NO, Wang M, Ma YP, Wang L, et al. 2020. Genomic regions under selection in the feralization of the dingoes. *Nat Commun.* 11(1):671.

Zhang W, Fan Z, Han E, Hou R, Zhang L, Galaverni M, Huang J, Liu H, Silva P, Li P, et al. 2014. Hypoxia adaptations in the grey wolf (*Canis lupus chanco*) from Qinghai–Tibet Plateau. *PLoS Genet.* 10(7):e1004466.