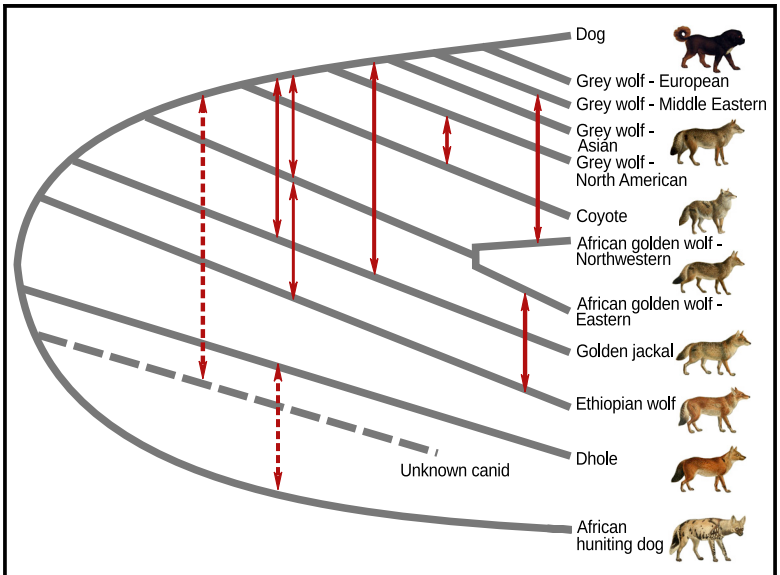
Report

Current Biology

Interspecific Gene Flow Shaped the Evolution of the Genus *Canis*

Graphical Abstract



Summary of the relationships between all the species included in this study. The red arrows indicate inferred gene flow between the two connected species. A solid arrow is corroboration of a previously reported finding, whereas a dashed arrow is a novel finding in this study.

Highlights

- Extensive gene flow in the genus *Canis*, especially among the crown group
- Genetic contribution from an unknown canid into the ancestor of the gray wolf and coyote
- The African golden wolf possibly a hybrid species, from the gray wolf and Ethiopian wolf
- Possible ancient admixture between the dhole and African hunting dog

Authors

Shyam Gopalakrishnan, Mikkel-Holger S. Sinding, Jazmín Ramos-Madrigal, ..., Øystein Wiig, Anders J. Hansen, M. Thomas P. Gilbert

Correspondence

shyam@snm.ku.dk

In Brief

Gopalakrishnan et al. present evidence of pervasive gene flow among species of the genus *Canis*. In addition to previously known admixture events, they find evidence of gene flow from a "ghost" canid, related to the dhole, into the ancestor of the gray wolf and coyote. Further, they suggest that the African golden wolf is a species of hybrid origin.

Gopalakrishnan et al., 2018, Current Biology *28*, 3441–3449
November 5, 2018
https://doi.org/10.1016/j.cub.2018.08.041

CellPress

Current Biology
Report

Interspecific Gene Flow Shaped the Evolution of the Genus *Canis*

Shyam Gopalakrishnan,[1,21,22,*] Mikkel-Holger S. Sinding,[1,2,3,4,21] Jazmín Ramos-Madrigal,[1,21] Jonas Niemann,[1] Jose A. Samaniego Castruita,[1] Filipe G. Vieira,[1] Christian Carøe,[1] Marc de Manuel Montero,[5] Lukas Kuderna,[5] Aitor Serres,[5] Víctor Manuel González-Basallote,[5] Yan-Hu Liu,[9] Guo-Dong Wang,[10] Tomas Marques-Bonet,[5,6,7,8] Siavash Mirarab,[11] Carlos Fernandes,[12] Philippe Gaubert,[13] Klaus-Peter Koepfli,[14,15] Jane Budd,[16] Eli Knispel Rueness,[17] Mads Peter Heide-Jørgensen,[1,3] Bent Petersen,[18,19] Thomas Sicheritz-Ponten,[18,19] Lutz Bachmann,[2] Øystein Wiig,[2] Anders J. Hansen,[1,3,4] and M. Thomas P. Gilbert[1,20]

[1]Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark
[2]Natural History Museum, University of Oslo, Oslo, Norway
[3]The Qimmeq Project, University of Greenland, Nuussuaq, Greenland
[4]University of Greenland, Manuutoq 1, Nuuk, Greenland
[5]Institute of Evolutionary Biology (UPF-CSIC), PRBB, Barcelona, Spain
[6]Catalan Institution of Research and Advanced Studies (ICREA), Passeig de Lluís Companys, 23, 08010, Barcelona, Spain
[7]CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Baldiri i Reixac 4, 08028 Barcelona, Spain
[8]Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Edifici ICTA-ICP, c/ Columnes s/n, 08193 Cerdanyola del Vallès, Barcelona, Spain
[9]State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan University, Kunming, Yunnan, China
[10]State Key Laboratory of Genetic Resources and Evolution and Yunnan Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China
[11]Department of Electrical and Computer Engineering, University of California, San Diego, San Diego, CA, USA
[12]Centre for Ecology, Evolution and Environmental Changes (CE3C), Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal
[13]Institut des Sciences de l'Evolution de Montpellier (ISEM), UM-CNRS-IRD-EPHE, Université de Montpellier, Montpellier, France
[14]Smithsonian Conservation Biology Institute, National Zoological Park, 3001 Connecticut Avenue NW, Washington, DC 20008, USA
[15]Theodosius Dobzhansky Center for Genome Bioinformatics, St. Petersburg State University, 41A Sredniy Prospekt, St. Petersburg 199034, Russia
[16]Breeding Centre for Endangered Arabian Wildlife, Sharjah, United Arab Emirates
[17]Centre for Ecological and Evolutionary Synthesis (CEES), University of Oslo, Oslo, Norway
[18]DTU Bioinformatics, Department of Bio and Health Informatics, Technical University of Denmark, Lyngby, Denmark
[19]Centre of Excellence for Omics-Driven Computational Biodiscovery (COMBio), Faculty of Applied Sciences, AIMST University, Kedah, Malaysia
[20]Norwegian University of Science and Technology, University Museum, Trondheim, Norway
[21]These authors contributed equally
[22]Lead Contact
*Correspondence: shyam@snm.ku.dk
https://doi.org/10.1016/j.cub.2018.08.041

SUMMARY

The evolutionary history of the wolf-like canids of the genus *Canis* has been heavily debated, especially regarding the number of distinct species and their relationships at the population and species level [1–6]. We assembled a dataset of 48 resequenced genomes spanning all members of the genus *Canis* except the black-backed and side-striped jackals, encompassing the global diversity of seven extant canid lineages. This includes eight new genomes, including the first resequenced Ethiopian wolf (*Canis simensis*), one dhole (*Cuon alpinus*), two East African hunting dogs (*Lycaon pictus*), two Eurasian golden jackals (*Canis aureus*), and two Middle Eastern gray wolves (*Canis lupus*). The relationships between the Ethiopian wolf, African golden wolf, and golden jackal were resolved. We highlight the role of interspecific hybridization in the evolution of this charismatic group. Specifically, we find gene flow between the ancestors of the dhole and African hunting dog and admixture between the gray wolf, coyote (*Canis latrans*), golden jackal, and African golden wolf. Additionally, we report gene flow from gray and Ethiopian wolves to the African golden wolf, suggesting that the African golden wolf originated through hybridization between these species. Finally, we hypothesize that coyotes and gray wolves carry genetic material derived from a "ghost" basal canid lineage.

RESULTS AND DISCUSSION

The genome dataset analyzed in this study contains 12 gray wolves and 14 dogs, chosen from regions overlapping the current ranges of the other basal canids included in this study,

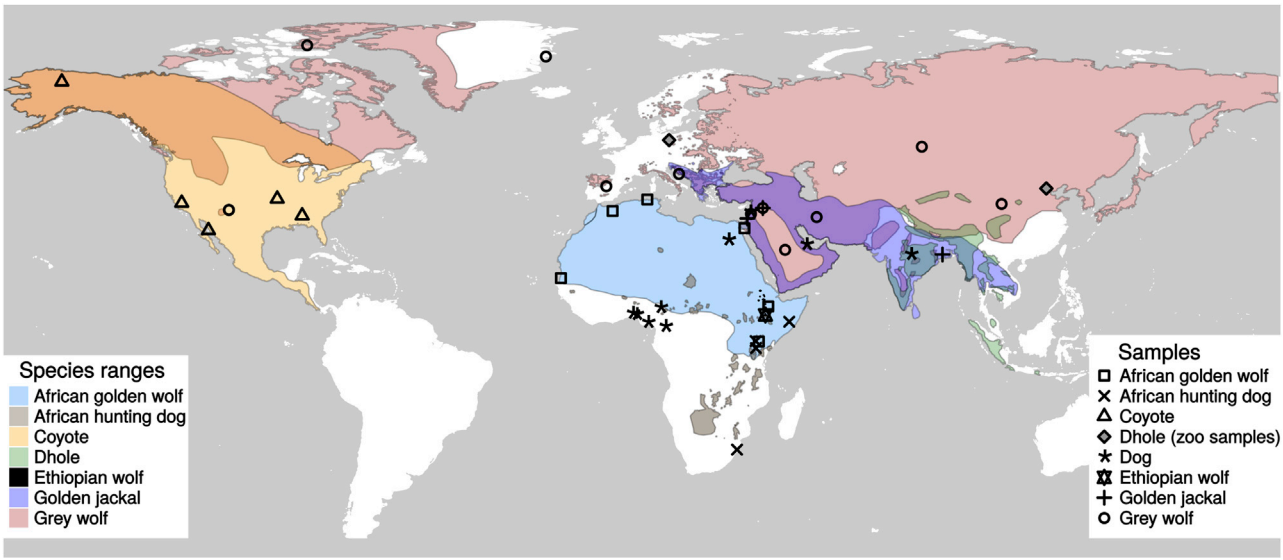


Figure 1. Map Showing the IUCN Ranges, Range Overlaps, and Sampling Locations of the Canids Included in This Study

The overlaps in ranges are shown in blended colors (orange, dark purple, dark olive green, light teal, etc.). Since IUCN does not have range information for African golden wolf, the IUCN range of golden jackal has been split in two; the Eurasian part is shown as the range of golden jackal, and the African part is shown as the range of African golden wolf. Further details on the samples, including their sampling location and source, can be found in Data S1, and their estimated heterozygosities—which are inversely proportional to their population sizes—are shown in Figure S1.

five coyotes, one Ethiopian wolf, three golden jackals, six African golden wolves (originally *Canis anthus*, but recently reclassified as *Canis lupaster* [1]), two dholes, four African hunting dogs, and one Andean fox (*Lycalopex culpaeus*) (Figure 1). Short-read sequencing of the samples and subsequent alignment to the recently published wolf genome assembly [7] resulted in genome-wide coverages ranging from 0.6–26.6× (for details, see Data S1). The genome-wide heterozygosity estimates (Figure S1) clearly show reduced levels in the Ethiopian wolf, African hunting dog, and dhole, an observation that is consistent with their small population sizes. The reconstructed phylogenetic relationships within this group of canids (Figure 2B) are of considerable relevance in light of extensive prior debate on the relationships between the Ethiopian wolf, golden jackal, and African golden wolf [2–5]. Our results corroborate the recent proposition based on both mitochondrial [2, 3] and nuclear [4, 6] data that the African golden wolf is evolutionarily distinct from the golden jackal (Figure 2C, panel labeled 16), but also that the Ethiopian wolf falls basal to both (Figure 2C, panel labeled 12) [5]. For convenience, we henceforth refer to five canid species, *viz*. the Ethiopian wolf, African golden wolf, golden jackal, gray wolf, and coyote, as "the crown group" in order to distinguish them from the more basal dholes and African hunting dogs. The placement of the Ethiopian wolf as the basal group in this clade is consistent with tree topologies obtained in previous phylogenetic analyses based on concatenated gene sequences [5] and more recent multispecies coalescent analyses [4] of datasets consisting of a subset of exonic and intronic sequences, but differs from the topology based on concatenated analyses in the latter study. We note that this nuclear-DNA-based phylogeny also places dogs as a sister clade to European gray wolves. However, we caution that this placement has only moderate support (0.86 mean local posterior probability); moreover, the gene tree quartet frequencies of alternate resolutions within the dog-gray wolf branches are comparable to that recovered in the main tree (Figure 2B, panel labeled 20–22), and thus no conclusion can be drawn about which wolf population gave rise to dogs. Indeed, our findings are not incompatible with previously suggested hypotheses [9] that either (1) the dog was domesticated from a now-extinct wolf population and/or (2) Eurasian gray wolf population genomic diversity has been reduced since the domestication event.

Mitochondrial genomes were *de novo* assembled from all species studied, using MtArchitect [10], which accounts for presence of numts in the reference genome. A maximum-likelihood phylogeny based on these mitochondrial genomes (Figure 2A) is largely consistent with that obtained from the nuclear genome analysis, with one obvious exception—the coyote mitochondrial genomes fall basal to all the other crown canids. This is consistent with Koepfli and colleagues' [4] results on near-complete mitochondrial genomes and thus contradicts the findings of numerous previous studies that used partial mitochondrial DNA sequences and placed coyotes (1) as sister to gray wolves [11], (2) in an unresolved clade with African golden wolves and Ethiopian wolves [2, 3], (3) as sister to Ethiopian wolves [1, 2, 12, 13], or, finally, (4) as sister to a clade containing Ethiopian wolves and golden jackals [14].

We subsequently explored the degree of interspecific gene flow between the various species. Many publications have reported interspecies gene flow between members of the canid crown group (dog-gray wolf complex, coyotes, Ethiopian wolves, golden jackals, and African golden wolves) [4, 5, 9, 13, 15–19]—something perhaps unsurprising, given the large geographic overlap of many of the populations. Initial analyses

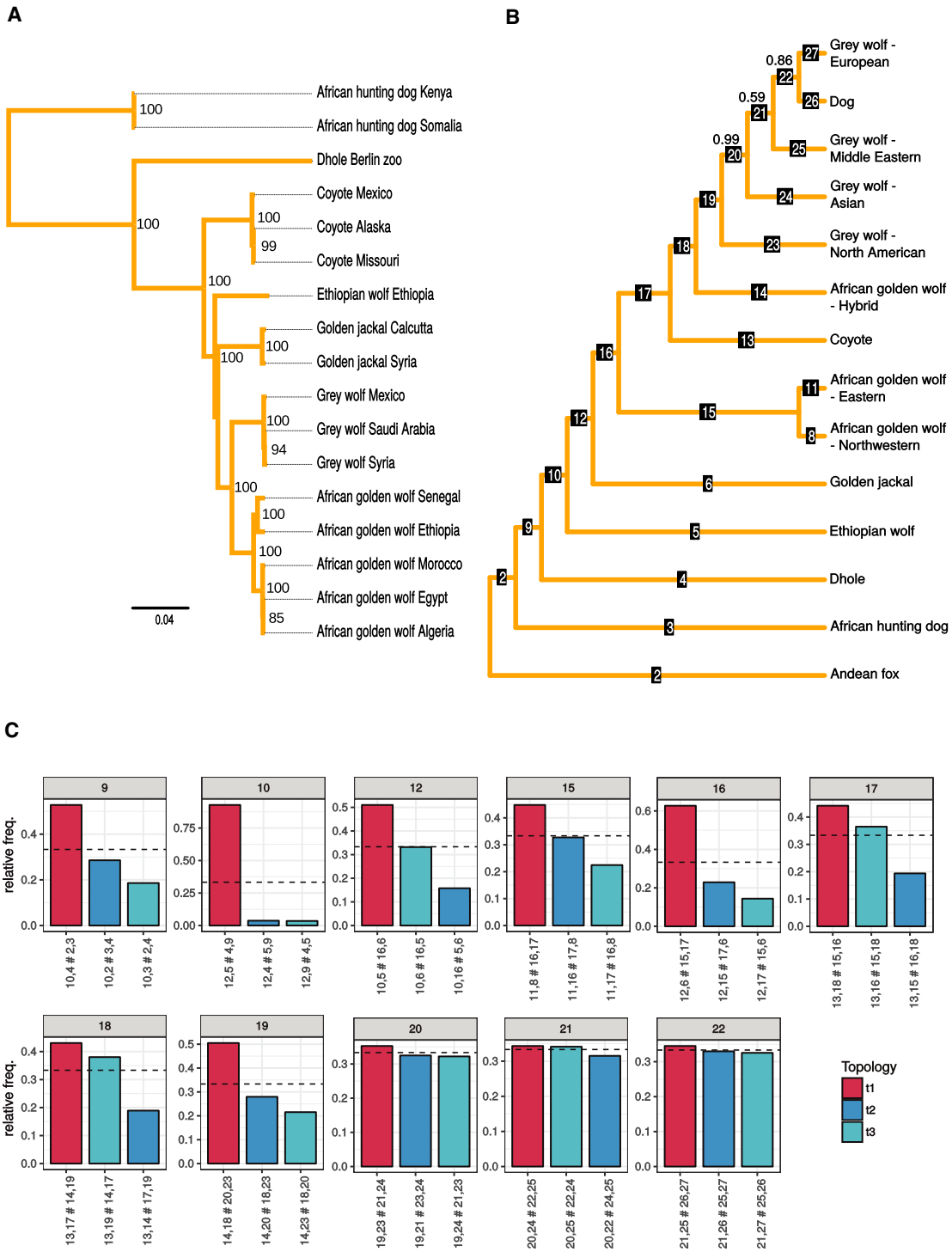


Figure 2. Nuclear and Mitochondrial Phylogeny of Basal Canids

(A) The maximum-likelihood estimate of the mitochondrial phylogeny for a subset of the samples, using *de novo* mitochondrial assemblies obtained with MtArchitect. The node labels show the bootstrap support for the node.

(B) The phylogeny estimated from nuclear DNA by ASTRAL-II, where monophyletic clusters have been collapsed into a single leaf node. The tip labeled "African golden wolf-hybrid'' represents a single known hybrid from the Sinai Peninsula—labeled "African golden wolf Egypt" in the mtDNA phylogeny—as described in the main text. The mean local posterior probabilities are shown for branches where this value is less than 1. The full nuclear phylogeny containing the sample relationships, branch supports, branch lengths proportional to divergence times, and estimated split times can be found in Figures S2A and S2B and Table S2.

(C) For a subset of the internal branches in the nuclear phylogeny, the quartet frequencies of the three possible configurations around each branch in the underlying unrooted tree are shown. The red bar represents the configuration shown in the phylogeny, and the two blue bars represent the two alternative

(legend continued on next page)

of genetic structure among these canids using NGSadmix [20] (Figure S3A) revealed that the individuals partition according to expected species structure. However, more details became apparent as the number of estimated clusters (K) was increased. For example, at higher values of K, gray wolves form five principal groups (Mexico, Ellesmere-Greenland, East Asia, the Middle East, and the remaining Eurasia), whereas African golden wolves are split into an Eastern and a Northwestern clade, as previously shown [4, 6, 16]. We note that similar east-west population differentiation is observed for several other African mammalian species [21], thus pointing to a general trend that the African golden wolves follow. The NGSadmix analyses also suggest the presence of admixture between the different species. For example, we detected not only dog introgression in the gray wolves from Spain and Israel, but also, perhaps of greater interest, gene flow between African golden wolves, golden jackals, and gray wolves. One example is a highly admixed African golden wolf from the Egyptian Sinai Peninsula, whose genome contains contributions from both Middle Eastern gray wolves and dogs (Figure S3A).

Previous studies that have reported admixture between canid species [9] and mitochondrial evidence for overlap of the gray wolf, African golden wolf, and golden jackal in eastern Egypt [4]. This points to the importance of the Sinai Peninsula and the Southwest Levant in canid evolution [4, 9], presumably due to its role as the land bridge between the African and Eurasian continents. We used TreeMix [22], *D* statistics [23], and admixture graphs [23] to examine signals of admixture between these species. The results confirmed that, in general terms, the level of gene flow between the three species is high, although varying across space in a manner consistent with their natural ranges (Figures 3B and S3A–S3E). For example, gene flow between golden jackals and gray wolves and between African golden wolves and gray wolves is lowest when North American gray wolves are considered, somewhat higher for Asian and European gray wolves, and highest with the gray wolves from the Middle East (e.g., Israel, Syria, and Saudi Arabia) (Figure S3E). Although the latter is not surprising in light of the natural ranges of the species, the evidence of golden jackal ancestry in North American wolves is intriguing. One possible explanation could be that gene flow happened before the divergence of the North American and Eurasian gray wolves. The fact that interspecific gene flow is considerably higher in Middle Eastern than in other gray wolves may also explain the distinctness of this population. The structure between Northwestern and Eastern African golden wolves can be explained using a similar argument—the former have highest levels of golden jackal and gray wolf admixture (Figures 3B, S3A, and S3B), whereas the latter show higher levels of gene flow from Ethiopian wolves. Overall, it is clear that individuals sampled in this land bridge region will be particularly informative for future studies that wish to study canid admixture in greater detail.

Furthermore, *D* statistics were used to test for gene flow between the dhole and African hunting dog, using members of the crown group as ingroup and the Andean fox as outgroup. Although no gene flow was detected between species of the crown group and the African hunting dogs, the analyses provided strong evidence of gene flow between the African hunting dog and dhole (Figure S3C). This is a surprising finding, since the ranges of the two species do not overlap. However, it is well documented that the dhole existed as far west as Europe during the Pleistocene [24]. Thus, one possible explanation could be the presence of dholes in the Middle East in the past, from where they could have encountered and mixed with African hunting dogs in North Africa. It must, however, be stressed that given that there has never been any reported evidence of dholes in either the Middle East or North Africa, our hypothesis is purely speculative. The timing and location of this admixture event remain unresolved.

Although there have been several reports of hybridization between dogs and Ethiopian wolves [13, 15], the genetic history of the Ethiopian wolf has not previously been investigated using nuclear genomic data. The *D*-statistics-based analyses provided evidence for gene flow between Ethiopian wolves and not only African golden wolves, but also golden jackals, gray wolves, and coyotes (Figure S3). The finding of considerable gene flow between the Ethiopian and Eastern African golden wolf lineages is not surprising, given their geographical co-occurrence in Africa. We consistently also observed a Northwestern-Eastern split in the African golden wolves and note that this correlates with our finding that the Ethiopian wolf contributes a higher amount to the Eastern African golden wolves. This suggests that admixture from the Ethiopian wolf may be a key factor contributing to African golden wolf population structure.

The presence of gene flow between the Ethiopian wolf and the other crown canid species is more surprising, given their lack of range overlap. However, this might be explained through the previously reported extensive evidence of admixture between African golden wolves and gray wolves, coyotes, and golden jackals [4, 9]. In short, we hypothesize that the signal of Ethiopian wolf admixture into the other crown canid species is mediated by African golden wolves. A summary of all the admixture events inferred in this study is shown in Figure 3A.

The uncertain placement of the African golden wolf (Figure 2C, panel labeled 17), combined with evidence of gene flow from the Ethiopian wolf, led us to investigate whether the African golden wolf is a species of hybrid origin, derived from a mixture between gray and Ethiopian wolves or close relatives. The current distribution ranges of Ethiopian and gray wolves do not overlap, and indeed, the known historical distribution of Ethiopian wolves is restricted to the Ethiopian highlands [15]. However, extensive gene flow with other canids, combined with the two distinct levels of Ethiopian wolf gene flow into the two distinct populations of African golden wolves, suggests that either Ethiopian wolves or a close (now extinct) relative had, in the past, a much larger range within Africa and thus greater opportunity to admix with other canid species. Additionally, mitochondrial analyses of African golden wolves, in this and previous studies,

configurations. For every quartet, the frequency of the true bipartition has previously been shown to be at least one-third [8], indicated here by a dotted line. Each alternative configuration is labeled by the bipartition it creates, with labels corresponding to those in (A). For example, the second bar of the panel labeled 12 swaps the positions of golden jackal (6) and Ethiopian wolf (5), whereas the third bar puts them as sister to each other. This plot summarizes the gene tree incongruence around examined branches.

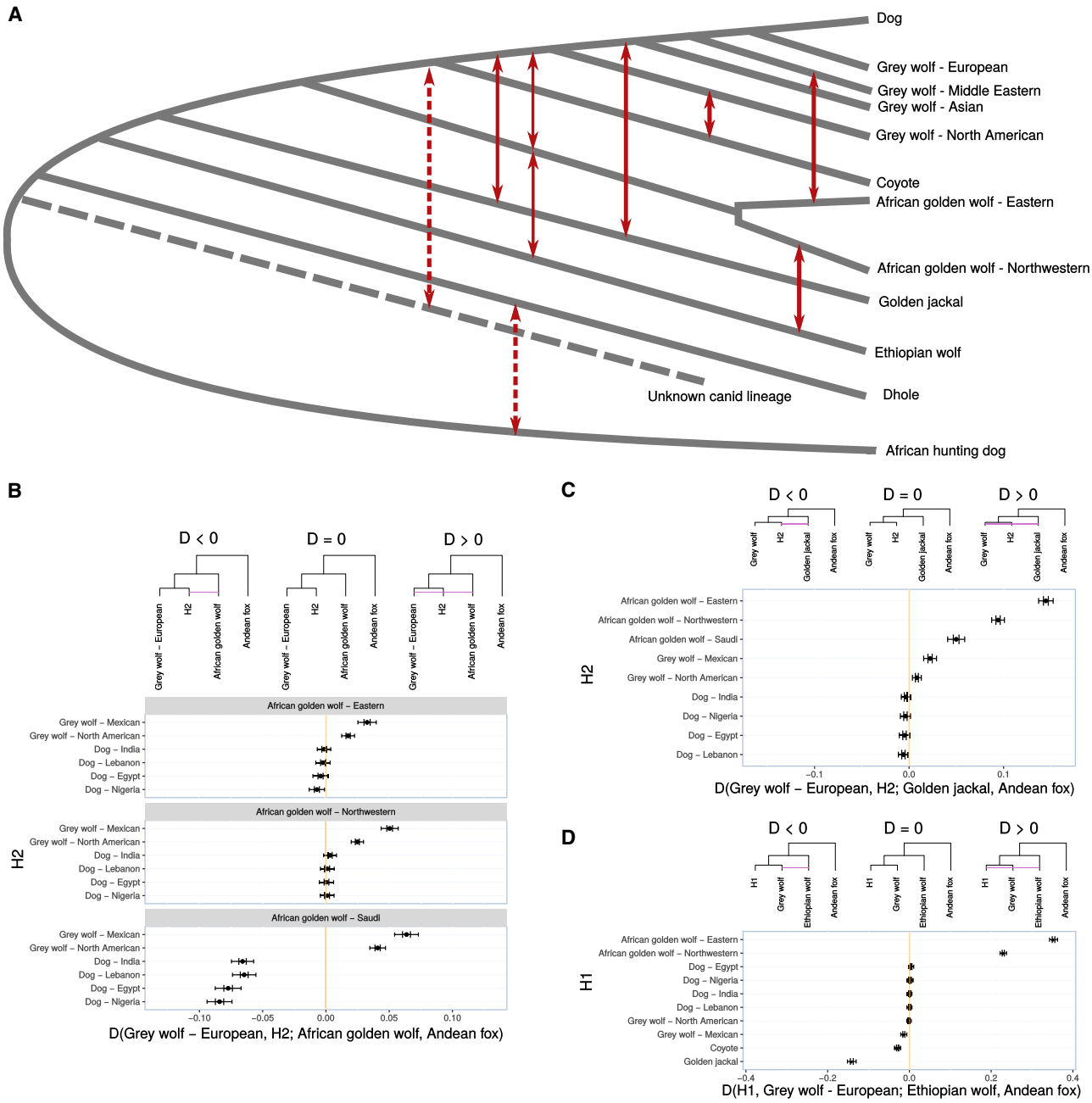


Figure 3. Gene Flow among the Crown Canid Species

(A) This figure summarizes the relationships among the species (phylogeny) and the various gene flow events inferred from the samples included in this study. Gene flow events are indicated with red arrows, and dotted red arrows show possible gene flow events that have been inferred in this study but have not been previously reported.

(B–D) These figures show the gene flow among the different crown canid species using *D* statistics. These *D* statistics show significant gene flow between the gray wolf, African golden wolf, golden jackal, and Ethiopian wolf. One principal new finding is structure within the African golden wolves, splitting into Northwestern and Eastern clades, which show genetic affinity to gray wolves and Ethiopian wolves, respectively. A second principal finding is inferred gene flow from an unknown canid lineage, related to the dhole, into the ancestor of the coyote and the gray wolves. We hypothesize this may explain the unexpected basal placement of the coyote in the mitochondrial tree. Further evidence of gene flow in the crown canids is shown in Figure S3.

find them to be most closely related to gray wolves [2–4, 25]. Further, African golden wolves are a sister clade to gray wolves and coyotes in the nuclear phylogeny, whereas they are a sister group to the Middle Eastern gray wolves in the mitochondrial phylogeny. We explored the relationships between the golden jackal, Ethiopian wolf, and African golden wolf using G-PhoCS [26] (Table S1), which supported the finding of gene flow into the Ethiopian wolf from the African golden wolf. To further

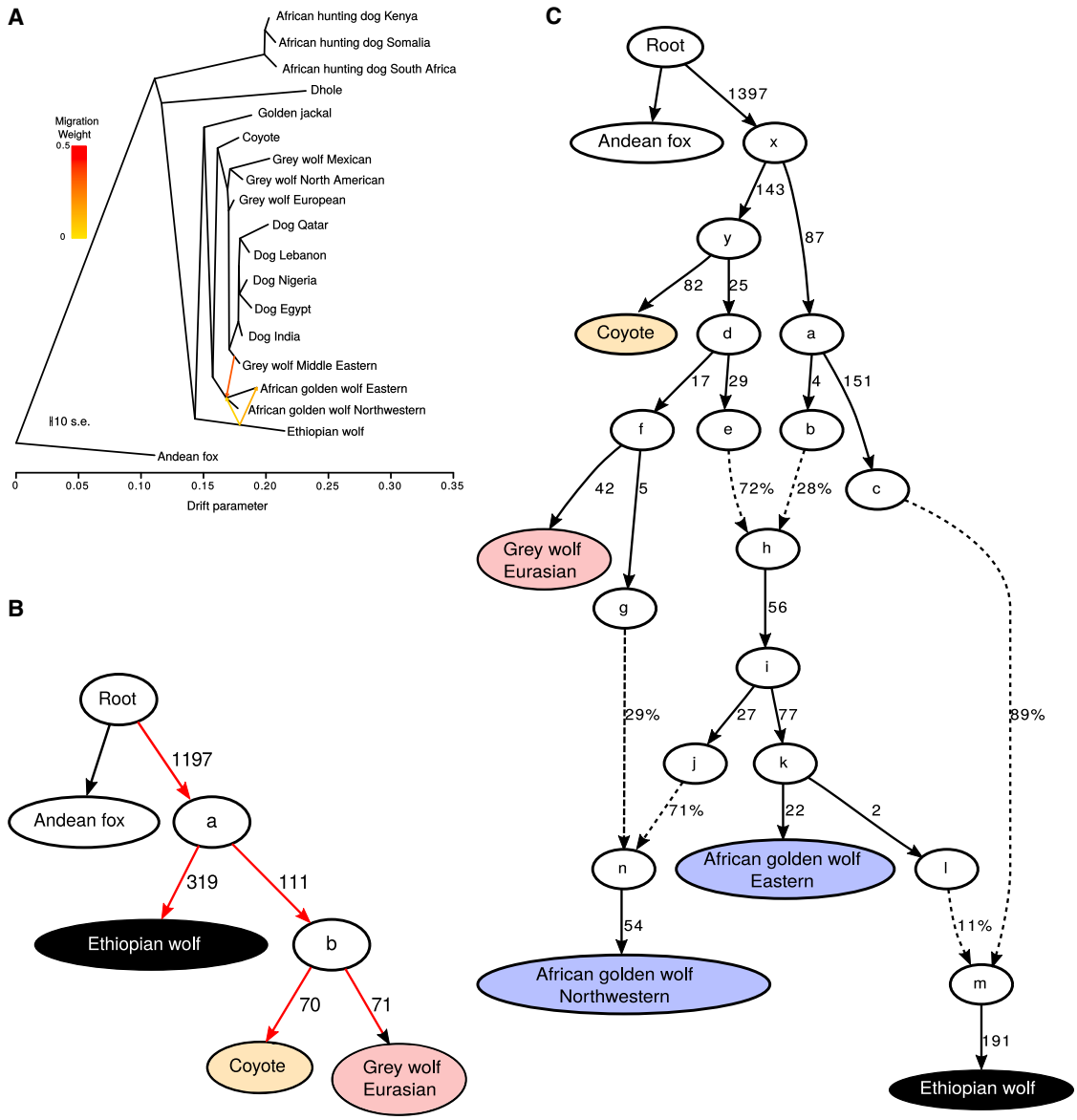


Figure 4. Modeling the Ancestry of African Golden Wolves

(A) TreeMix tree with all samples, estimated using the pairwise correlation of allele frequencies between all groups of samples. This tree is fit with three migration edges. The first three migration edges all indicate extensive gene flow from the gray and Ethiopian wolves into the African golden wolves, suggesting a hybrid origin for this species.

(B and C) The QP graph is an admixture graph estimated using all pairwise *D* statistics between samples. Estimated genetic drift is shown along the solid lines in units of f2 distance (parts per thousand), and estimated mixture proportions are given along the dotted lines. Names of specific modern populations are shown in full, whereas hypothetical ancestral individuals are represented by letters.

(B) This tree shows all the possible placements—highlighted in red—for the Northwestern African golden wolf, chosen due to their low levels of gene flow with the Ethiopian wolf. These were modeled as possible internal and external nodes and as an admixed group from all possible node pairs.

(C) The best fitting graph with a *Z* value closest to 0, modeling the Ethiopian wolf-like and gray wolf-like ancestry of Northwestern and Eastern African golden wolves, as well as gene flow into modern Ethiopian wolves from the Eastern African golden wolves. This admixture graph suggests that the African golden wolves are probably a species of hybrid origin, derived from the gray wolf and Ethiopian wolf as the parental species. Further, Figure S4 shows admixture graphs showing potential gene flow from a "ghost" basal canid lineage into the ancestor of wolves and dogs.

explore the relationship between these species and the gray wolf, we used TreeMix [22] and admixture graphs [23] to obtain trees, which were used to assess whether the African golden wolf is a hybrid species (Figures 4B and 4C). We initially constructed a graph including the coyote, Ethiopian wolf, gray wolf, and Andean fox and assessed the most likely position for the African golden wolf in this graph. The placement of the two African golden wolf populations in this tree was further investigated by modeling them as sister to all possible nodes and as admixed populations deriving ancestry from two possible nodes.

Finally, the model was extended to account for African golden wolf admixture into the Ethiopian wolf. We found that the common ancestor of the African golden wolf populations is best modeled as admixed between a component related to the Ethiopian wolf (~28%) and another related to the gray wolf (~72%) (worst-fitting f statistic Z value = −1.086; Figure 4C). Finally, the northwestern African golden wolf population is more closely related to the gray wolf, which is best explained in our model through admixture from gray wolves.

Lastly, our attention was drawn to the curious result of potential gene flow between the lineage representing the ancestor of the coyote and gray wolves and that representing all other canid species, excluding the African hunting dog (Figure S4), in all D statistics analyses computed with the coyote or gray wolf in the ingroup, namely position H2. Notably, these signals disappeared when the sister clade—H3—was replaced with the African hunting dog, leading us to hypothesize that the coyote and gray wolf genomes may contain a basal ancestral component derived from an as-yet-unidentified species that evolved after the divergence of the African hunting dog branch from the other canid species and that the signal of gene flow can be attributed to outgroup attraction of the coyote and gray wolf lineage. Note that such a hypothetical ancient admixture event would also explain the unexpectedly basal position of the coyote mitochondrial genome—the coyote may simply have retained the mitogenome from this unidentified ancestor. We acknowledge that the existence of an unknown ancestral component would be controversial—previous analyses of coyotes and the fossil records from their direct ancestors argue that they have been strictly restricted to North America for over a million years [27, 28]. However, within North America, the coyote has coexisted alongside several now extinct canids, including the American dhole (*Cuon* sp.) and dire wolf (*Canis dirus*) [29]. Although the unknown ancestral component to cannot be attributed to any of the known fossil species at this time, future paleogenomic analyses on such materials (if any can be found with surviving DNA) may provide exciting possibilities to test our hypothesis.

In conclusion, our results highlight how interspecific gene flow has played an important role in shaping the species and population structure of gray wolves, coyotes, African golden wolves, golden jackals, and Ethiopian wolves and that African golden wolves, coyotes, and gray wolves may have been greatly affected by hybridization events. In particular, we conclude not only that African golden wolves arose through hybridization between a Ethiopian-wolf-like and gray-wolf-like ancestral population, but that subsequently the resulting northwestern and eastern African golden wolf populations underwent continuous admixture with modern gray and Ethiopian wolves, respectively. We furthermore argue that the common ancestor of gray wolves and coyotes differentiated from the lineage leading to golden jackals, in part by admixing with a dhole-like canid. Finally, the robust signal of gene flow observed between African hunting dogs and dholes testifies to an as-yet-undiscovered prehistoric overlap between the two lineages. This underscores how much remains to be discovered about the history of the wolf-like canids and how paleogenomic approaches may be required to advance our understanding of this group. Lastly, our study adds to the growing evidence for the importance of gene flow and hybridization in the evolution of mammalian species in general [23, 30–32] and that rather than being isolated entities that evolve along tree-like phylogenies, they are interlinked and evolve through interactions in network-like topologies.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Whole-genome sequencing
  - Read mapping
  - Genotype calling
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Heterozygosity
  - Admixture
  - Nuclear genome phylogeny
  - Species split times
  - Mitochondrial reconstruction using *de novo* assembly
  - D statistics
  - TreeMix
  - qpGraph
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, two tables, and one data file and can be found with this article online at https://doi.org/10.1016/j.cub.2018.08.041.

ACKNOWLEDGMENTS

The authors would like to acknowledge the assistance of the Danish National High-Throughput Sequencing Centre for assistance in Illumina data generation. We also gratefully acknowledge the Danish National Supercomputer for Life Sciences, Computerome (https://www.computerome.dk), for the computational resources to perform the sequence analyses. For making sample material available, we would like to thank Jörns Fickel from Leibniz-Institut für Zoo- und Wildtierforschung and Kristian Gregersen from the Natural History Museum of Denmark. We also acknowledge the following for funding our research: the Qimmeq project funded by The Velux Foundations and Aage og Johanne Louis-Hansens Fond; Carlsbergfondet grant CF14–0995 and Marie Skłodowska-Curie Actions grant 655732-WhereWolf to S.G.; grant 676154-ArchSci2020 to J.N.; NSFC grant 91531303 to G.-D.W.; Danish National Research Foundation grant DNRF94, Lundbeckfonden grant R52–5062, and ERC Consolidator grant 681396-Extinction Genomics to M.T.P.G.; and the Universities of Oslo and Copenhagen for a PhD stipend awarded to M.-H.S.S. T.M.-B. is supported by MINECO/FEDER, UE, grant BFU2017-86471-P, NIMH grant U01 MH106874, a Howard Hughes Medical Institute International Early Career grant, Obra Social "La Caixa," and Secretaria d'Universitats i Recerca and CERCA Programme del Departament d'Economia i Coneixement de la Generalitat de Catalunya.

AUTHOR CONTRIBUTIONS

S.G., M.-H.S.S., A.J.H., and M.T.P.G. conceived the study. M.-H.S.S. and C.C. did the DNA lab work for high-throughput sequencing. S.G., J.R.-M., J.N., J.A.S.C., F.G.V., M.d.M.M., L.K., A.S., V.M.G.-B., Y.-H.L., and S.M. performed analyses. C.F., P.G., K.-P.K., J.B., E.K.R., and M.P.H.-J. contributed with sample collection. B.P. and T.S.-P. provided computation expertise and support. L.B., Ø.W., T.M.-B., A.J.H., and M.T.P.G. supervised the work.

S.G., M.-H.S.S., J.R.-M., and M.T.P.G. wrote the manuscript. All authors contributed to the preparation and editing of the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 9, 2017
Revised: April 30, 2018
Accepted: August 16, 2018
Published: October 18, 2018

REFERENCES

1. Viranta, S., Atickem, A., Werdelin, L., and Stenseth, N.C. (2017). Rediscovering a forgotten canid species. BMC Zoology *2*, 6.
2. Rueness, E.K., Asmyhr, M.G., Sillero-Zubiri, C., Macdonald, D.W., Bekele, A., Atickem, A., and Stenseth, N.C. (2011). The cryptic African wolf: Canis aureus lupaster is not a golden jackal and is not endemic to Egypt. PLoS ONE *6*, e16385.
3. Gaubert, P., Bloch, C., Benyacoub, S., Abdelhamid, A., Pagani, P., Djagoun, C.A.M.S., Couloux, A., and Dufour, S. (2012). Reviving the African wolf Canis lupus lupaster in North and West Africa: a mitochondrial lineage ranging more than 6,000 km wide. PLoS ONE *7*, e42740.
4. Koepfli, K.-P., Pollinger, J., Godinho, R., Robinson, J., Lea, A., Hendricks, S., Schweizer, R.M., 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*, 2158–2165.
5. Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas, E.J., 3rd, Zody, M.C., et al. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature *438*, 803–819.
6. Fan, Z., Silva, P., Gronau, I., Wang, S., Armero, A.S., Schweizer, R.M., 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*, 163–173.
7. Gopalakrishnan, S., Samaniego Castruita, J.A., Sinding, M.S., Kuderna, L.F.K., Räikkönen, J., Petersen, B., Sicheritz-Ponten, T., Larson, G., Orlando, L., Marques-Bonet, T., et al. (2017). The wolf reference genome sequence (Canis lupus lupus) and its implications for Canis spp. population genomics. BMC Genomics *18*, 495.
8. Allman, E.S., Degnan, J.H., and Rhodes, J.A. (2011). Identifying the rooted species tree from the distribution of unrooted gene trees under the coalescent. J. Math. Biol. *62*, 833–862.
9. Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E., Silva, P.M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., et al. (2014). Genome sequencing highlights the dynamic early history of dogs. PLoS Genet. *10*, e1004016.
10. Lobon, I., Tucci, S., de Manuel, M., Ghirotto, S., Benazzo, A., Prado-Martinez, J., Lorente-Galdos, B., Nam, K., Dabad, M., Hernandez-Rodriguez, J., et al. (2016). Demographic history of the genus Pan inferred from whole mitochondrial genome reconstructions. Genome Biol. Evol. *8*, 2020–2030.
11. Ostrander, E.A., and Wayne, R.K. (2005). The canine genome. Genome Res. *15*, 1706–1716.
12. Wayne, R.K., and Ostrander, E.A. (1999). Origin, genetic diversity, and genome structure of the domestic dog. BioEssays *21*, 247–257.
13. Gottelli, D., Sillero-Zubiri, C., Applebaum, G.D., Roy, M.S., Girman, D.J., Garcia-Moreno, J., Ostrander, E.A., and Wayne, R.K. (1994). Molecular genetics of the most endangered canid: the Ethiopian wolf Canis simensis. Mol. Ecol. *3*, 301–312.
14. Yumnam, B., Negi, T., Maldonado, J.E., Fleischer, R.C., and Jhala, Y.V. (2015). Phylogeography of the golden jackal (Canis aureus) in India. PLoS ONE *10*, e0138497.
15. Marino, J., and Sillero-Zubiri, C. (2011). *Canis simensis*. The IUCN Red List of Threatened Species 2011: e.T3748A10051312. https://doi.org/10.2305/IUCN.UK.2011-1.RLTS.T3748A10051312.en.
16. vonHoldt, B.M., Pollinger, J.P., Earl, D.A., Knowles, J.C., Boyko, A.R., Parker, H., Geffen, E., Pilot, M., Jedrzejewski, W., Jedrzejewska, B., et al. (2011). A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. Genome Res. *21*, 1294–1305.
17. Galov, A., Fabbri, E., Caniglia, R., Arbanasić, H., Lapalombella, S., Florijančić, T., Bošković, I., Galaverni, M., and Randi, E. (2015). First evidence of hybridization between golden jackal (Canis aureus) and domestic dog (Canis familiaris) as revealed by genetic markers. R. Soc. Open Sci. *2*, 150450.
18. vonHoldt, B.M., Cahill, J.A., Fan, Z., Gronau, I., Robinson, J., Pollinger, J.P., Shapiro, B., Wall, J., and Wayne, R.K. (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*, e1501714.
19. vonHoldt, B.M., Kays, R., Pollinger, J.P., and Wayne, R.K. (2016). Admixture mapping identifies introgressed genomic regions in North American canids. Mol. Ecol. *25*, 2443–2453.
20. Skotte, L., Korneliussen, T.S., and Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. Genetics *195*, 693–702.
21. Lorenzen, E.D., Heller, R., and Siegismund, H.R. (2012). Comparative phylogeography of African savannah ungulates. Mol. Ecol. *21*, 3656–3670.
22. Pickrell, J.K., and Pritchard, J.K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet. *8*, e1002967.
23. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065–1093.
24. Ripoll, M.P., Morales Pérez, J.V., Sanchis Serra, A., Aura Tortosa, J.E., and Montañana, I.S. (2010). Presence of the genus Cuon in upper Pleistocene and initial Holocene sites of the Iberian Peninsula: new remains identified in archaeological contexts of the Mediterranean region. J. Archaeol. Sci. *37*, 437–450.
25. Werhahn, G., Senn, H., Kaden, J., Joshi, J., Bhattarai, S., Kusi, N., Sillero-Zubiri, C., and Macdonald, D.W. (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*, 170186.
26. Gronau, I., Hubisz, M.J., Gulko, B., Danko, C.G., and Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. Nat. Genet. *43*, 1031–1034.
27. Kurtén, B. (1974). A history of coyote-like dogs (Canidae, Mammalia). Acta Zool. Fenn. *140*, 1–38.
28. Nowak, R.M. (1979). North American Quaternary Canis (Museum of Natural History, University of Kansas).
29. Tedford, R.H., Wang, X., and Taylor, B.E. (2009). Phylogenetic systematics of the North American fossil Caninae (Carnivora: Canidae). Bull. Am. Mus. Nat. Hist. *325*, 1–218.
30. Jónsson, H., Schubert, M., Seguin-Orlando, A., Ginolhac, A., Petersen, L., Fumagalli, M., Albrechtsen, A., Petersen, B., Korneliussen, T.S., Vilstrup, J.T., et al. (2014). Speciation with gene flow in equids despite extensive chromosomal plasticity. Proc. Natl. Acad. Sci. USA *111*, 18655–18660.
31. Figueiró, H.V., Li, G., Trindade, F.J., Assis, J., Pais, F., Fernandes, G., Santos, S.H.D., Hughes, G.M., Komissarov, A., Antunes, A., et al. (2017). Genome-wide signatures of complex introgression and adaptive evolution in the big cats. Sci. Adv. *3*, e1700299.
32. Kumar, V., Lammers, F., Bidon, T., Pfenninger, M., Kolter, L., Nilsson, M.A., and Janke, A. (2017). The evolutionary history of bears is characterized by gene flow across species. Sci. Rep. *7*, 46487.
33. Auton, A., Rui Li, Y., Kidd, J., Oliveira, K., Nadel, J., Holloway, J.K., Hayward, J.J., Cohen, P.E., Greally, J.M., Wang, J., et al. (2013).

Genetic recombination is targeted towards gene promoter regions in dogs. PLoS Genet. *9*, e1003984.

34. Campana, M.G., Parker, L.D., Hawkins, M.T.R., Young, H.S., Helgen, K.M., Szykman Gunther, M., Woodroffe, R., Maldonado, J.E., and Fleischer, R.C. (2016). Genome sequence, population history, and pelage genetics of the endangered African wild dog (Lycaon pictus). BMC Genomics *17*, 1013.
35. Liu, Y.-H., Wang, L., Xu, T., Guo, X., Li, Y., Yin, T.-T., Yang, H.-C., Yang, H., Adeola, A.C., Sanke, J.O., et al. (2017). Whole-genome sequencing of African dogs provides insights into adaptations against tropical parasites. Mol. Biol. Evol. *35*, 287–298.
36. Wang, G.-D., Zhai, W., Yang, H.-C., Fan, R.-X., Cao, X., Zhong, L., Wang, L., Liu, F., Wu, H., Cheng, L.-G., et al. (2013). The genomics of selection in dogs and the parallel evolution between dogs and humans. Nat. Commun. *4*, 1860.
37. Wang, G.-D., Zhai, W., Yang, H.-C., Wang, L., Zhong, L., Liu, Y.-H., Fan, R.-X., Yin, T.-T., Zhu, C.-L., Poyarkov, A.D., et al. (2016). Out of southern East Asia: the natural history of domestic dogs across the world. Cell Res. *26*, 21–33.
38. Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. *2010*, t5448.
39. Schubert, M., Ermini, L., Der Sarkissian, C., Jónsson, H., Ginolhac, A., Schaefer, R., Martin, M.D., Fernández, R., Kircher, M., McCue, M., et al. (2014). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. Nat. Protoc. *9*, 1056–1082.
40. Schubert, M., Lindgreen, S., and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res. Notes *9*, 88.
41. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.
42. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. *43*, 491–498.
43. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.
44. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: analysis of next generation sequencing data. BMC Bioinformatics *15*, 356.
45. Mirarab, S., and Warnow, T. (2015). ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics *31*, i44–i52.
46. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.
47. Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972–1973.
48. Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS ONE *5*, e9490.
49. Sayyari, E., Whitfield, J.B., and Mirarab, S. (2017). DiscoVista: interpretable visualizations of gene tree discordance. arXiv, arXiv: 1709.09305, https://arxiv.org/abs/1709.09305.
50. Yamada, K.D., Tomii, K., and Katoh, K. (2016). Application of the MAFFT sequence alignment program to large data-reexamination of the usefulness of chained guide trees. Bioinformatics *32*, 3246–3251.
51. Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2–a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189–1191.
52. Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods *9*, 772.
53. Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307–321.
54. Gilbert, M.T.P., Tomsho, L.P., Rendulic, S., Packard, M., Drautz, D.I., Sher, A., Tikhonov, A., Dalén, L., Kuznetsova, T., Kosintsev, P., et al. (2007). Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. Science *317*, 1927–1930.
55. Carøe, C., Gopalakrishnan, S., Vinner, L., Mak, S.S.T., Sinding, M.-H.S., Samaniego, J.A., Wales, N., Sicheritz-Pontén, T., and Gilbert, M.T.P. (2017). Single-tube library preparation for degraded DNA. Methods Ecol. Evol. *9*, 410–419.
56. Allentoft, M.E., Sikora, M., Sjögren, K.-G., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlström, T., Vinner, L., et al. (2015). Population genomics of Bronze Age Eurasia. Nature *522*, 167–172.
57. Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., and Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. USA *110*, 15758–15763.
58. Nielsen, R., Paul, J.S., Albrechtsen, A., and Song, Y.S. (2011). Genotype and SNP calling from next-generation sequencing data. Nat. Rev. Genet. *12*, 443–451.
59. Sayyari, E., and Mirarab, S. (2016). Fast coalescent-based computation of local branch support from quartet frequencies. Mol. Biol. Evol. *33*, 1654–1668.
60. Schlebusch, C.M., Malmström, H., Günther, T., Sjödin, P., Coutinho, A., Edlund, H., Munters, A.R., Vicente, M., Steyn, M., Soodyall, H., et al. (2017). Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago. Science *358*, 652–655.
61. Busing, F.M.T.A., Meijer, E., and Van Der Leeden, R. (1999). Delete-M jackknife for unequal M. Stat. Comput. *9*, 3–8.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
8 *Canid* blood or tissue samples	This paper	Data S1
Chemicals, Peptides, and Recombinant Proteins		
Proteinase K	Sigma-Aldrich	Cat# 3115844001
Phenol	Bionordika	Cat# A0447,0500
Chloroform	Sigma-Aldrich	Cat# 288306-1L
Critical Commercial Assays		
DNeasy Blood & Tissue Kit	QIAGEN	Cat# 69506
MinElute PCR Purification Kit	QIAGEN	Cat# 28006
NEBNext DNA Sample Prep Master Mix Set 2	New England Biolabs	Cat# E6070
Deposited Data		
10 *Canid* genomes	[33]	Data S1
2 *Canid* genomes	[34]	Data S1
3 *Canid* genomes	[6]	Data S1
5 *Canid* genomes	[9]	Data S1
1 *Canid* genomes	[4]	Data S1
4 *Canid* genomes	[35]	Data S1
2 *Canid* genomes	[18]	Data S1
1 *Canid* genomes	[36]	Data S1
5 *Canid* genomes	[37]	Data S1
1 African golden wolf	This article	NCBI SRA sample accession number: SAMN10199001
2 African hunting dogs	This article	NCBI SRA sample accession numbers: SAMN10180432, SAMN10180433
3 Coyotes	This article	NCBI SRA sample accession numbers: SAMN10180421, SAMN10180422, SAMN10180423
1 Dhole	This article	NCBI SRA sample accession number: SAMN10180424
1 Ethiopian wolf	This article	NCBI SRA sample accession number: SAMN10180425
2 Golden jackals	This article	NCBI SRA sample accession numbers: SAMN10180426, SAMN10180427
5 Gray wolves	This article	NCBI SRA sample accession numbers: SAMN10180428, SAMN10180429, SAMN10180430, SAMN10180431, SAMN10180511
Gray wolf reference genome	[7]	N/A
Oligonucleotides		
Illumina-compatible adapters	[38]	N/A
Software and Algorithms		
PALEOMIX	[39]	https://github.com/MikkelSchubert/paleomix; RRID:SCR_015057
AdapterRemoval2	[40]	https://github.com/MikkelSchubert/adapterremoval; RRID:SCR_011834
bwa v0.7.10	[41]	http://bio-bwa.sourceforge.net/; RRID:SCR_010910
Picard v1.128	N/A	https://broadinstitute.github.io/picard; RRID:SCR_006525
GATK v3.3.0	[42, 43]	https://broadinstitute.github.io/picard; RRID:SCR_001876
ANGSD	[44]	https://github.com/ANGSD/angsd

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
samtools v1.2	[41]	http://samtools.sourceforge.net/; RRID:SCR_002105
realSFS	[44]	https://github.com/ANGSD/angsd
NGSadmix	[20]	http://www.popgen.dk/software/index.php/NgsAdmix; RRID:SCR_003208
ASTRAL-II	[45]	https://github.com/smirarab/ASTRAL
RAxML	[46]	https://sco.h-its.org/exelixis/software.html; RRID:SCR_006086
trimal	[47]	http://trimal.cgenomics.org/
FastTree2	[48]	http://www.microbesonline.org/fasttree/; RRID:SCR_015501
DiscoVista	[49]	https://github.com/esayyari/DiscoVista
MtArchitect	[10]	http://biologiaevolutiva.org/tmarques/mtarchitect/
MAFFT	[50]	https://mafft.cbrc.jp/alignment/software/; RRID:SCR_011811
Jalview	[51]	http://www.jalview.org/; RRID:SCR_006459
jmodeltest2	[52]	https://github.com/ddarriba/jmodeltest2; RRID:SCR_015244
phyML	[53]	http://www.atgc-montpellier.fr/phyml/; RRID:SCR_014628
ADMIXTOOLS	[23]	https://github.com/DReichLab/AdmixTools
TreeMix	[22]	https://bitbucket.org/nygcresearch/treemix/wiki/Home

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Shyam Gopalakrishnan (shyam@snm.ku.dk).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The current study uses short read sequencing data from the full genomes of 47 canids spanning 8 different species (when the domestic dog is considered a different species from the gray wolf) from Africa, Eurasia and North America, to address questions about the genetic affinities of these species to each other, and the role of interspecific gene flow in shaping the evolution of the genus *Canis*. All known information on the context and sequencing coverage of the samples is provided in Data S1.

METHOD DETAILS

Whole-genome sequencing

DNA was extracted from 10 modern samples of fresh blood or tissue using the DNeasy Blood & Tissue Kit (QIAGEN, Hilden, Germany) following the manufacturer's protocol. Three samples ('African hunting dog Kenya 1', 'African hunting dog Somalia' and 'Golden jackal Calcutta') are from historical museum hides and were digested in a proteinase K-containing buffer following [54]; these digests were subsequently treated in a phenol chloroform step following [55]. The supernatant was then mixed 1:10 with a binding buffer following [56] in a binding apparatus following [57], including a Minelute column (QIAGEN, Hilden, Germany) that was then washed and DNA was eluted according to the manufacturer's guidelines. All extracts were incorporated into double-stranded DNA libraries build using the NEBNext DNA Sample Prep Master Mix Set 2 (E6070 - New England Biolabs, Beverly, MA, USA) following the manufacturer's protocol and Illumina-compatible adapters [38]. Libraries were sequenced using 50 base pair single (Golden jackal Calcutta, Hunting dog Kenya 1 and Hunting dog Somalia) or 100 base pair paired end (remaining samples) read chemistry on Illumina HiSeq 2000 and 2500 (Illumina, San Diego, CA, USA) platforms.

Read mapping

The short-read data from each sample, including samples from previous publications, was processed using the PALEOMIX pipeline [39]. As the first step of the pipeline, low quality and missing bases were trimmed from the reads, followed by removal of adapters using

AdapterRemoval2 [40]. Additionally, all paired end reads where the two reads overlapped by more than 10 base pairs were merged into a single read. Subsequently, the reads from each sample were mapped to the wolf reference genome [7] using bwa (v0.7.10; aln algorithm) [41]. The mapped reads were filtered for PCR and optical duplicates using Picard (v1.128, https://broadinstitute.github.io/picard), and reads that mapped to multiple locations in the genome were excluded. GATK (v3.3.0) [42, 43] was used to perform an indel realignment step to adjust for increased error rates at the end of short reads in the presence of indels. In the absence of a curated dataset of indels in wolves, this step relied on a set of indels identified in the specific sample being processed. After the initial mapping and quality control, the coverages of the samples ranged from 0.6 to 26.6x (for details see Data S1).

Genotype calling

The samples in this study span a wide range of genomic coverages. To avoid introducing biases in various analyses resulting from genotype calling in low coverage samples [58], the uncertainty in genotypes was instead propagated through to downstream analyses using genotype likelihoods. The genotype likelihoods at variant sites were computed in ANGSD [44] using the mapped reads, with the model for reads used by samtools (v1.2) [41]. Bases with base qualities lower than 20 and reads with mapping quality lower than 20 were discarded. Only sites with data present in at least 46 out of the 48 samples were retained. All sites with minor allele frequencies below 0.1 were excluded.

QUANTIFICATION AND STATISTICAL ANALYSIS

Heterozygosity

The heterozygosity for each sample was calculated using ANGSD, by estimating the per-sample folded site frequency spectrum (SFS) and using the fraction of singletons in the sample as a measure of heterozygosity. The variance of the estimate was obtained by bootstrapping the sites 100 times to obtain 100 bootstrapped estimates of the SFS. Briefly, for each sample, the site allele frequency for every site was estimated (“-doSaf 1 -fold 1”) using the reference genome as ancestral, while keeping all other parameters as above. Afterward, the SFS and their corresponding bootstraps was estimated for each sample using realSFS and, for each case, the fraction of singletons was calculated. The sample heterozygosities are shown in Figure S1.

Admixture

Using the genotype likelihoods obtained from the ANGSD pipeline, the ancestry clusters and admixture proportions for 48 samples representing all species (for details see Data S1) were estimated using NGSadmix [20] based on 5.7 million SNPs. Admixture analyses were performed using only markers with minor allele frequency greater than 0.1. We used a range of values for the number of clusters (2-15), to explore the structure in the dataset. To avoid convergence to local optima, the admixture analysis was repeated at least 200 times with different random initial parameter values, and the replicate with the highest likelihood was chosen.

Nuclear genome phylogeny

Using 28 individuals representing all species in this study (for details see Data S1), nuclear genome phylogenetic reconstruction based on coalescence of gene trees was performed using 100 ASTRAL-II trees [45], and an extended majority rule consensus tree was made with RAxML [46] using default parameters. Each tree was based on gene trees inferred from 5000 regions, each roughly 10 kb long sampled from a consensus genome sequence per individuals generated in ANGSD [44] using the “-doFasta 1” option. Regions with missing data were excluded using trimal [47] under the parameters “-gappyout -resoverlap 0.60 -seqoverlap 60.” Each gene tree was generated in FastTree2 [48] using a generalized time-reversible model for sequence evolution. A cut-off at a minimum of four samples per tree was selected, before generation of individual ASTRAL-II trees. Local posterior probabilities and quartet frequencies for the three possible unrooted resolutions around each internal branch were computed using ASTRAL [59] and visualized using DiscoVista [49]. Two support values are computed on the consensus ASTRAL tree: i) frequency of each branch in the 100 replicates and ii) means of local posterior probability across the 100 replicates. The local posterior probability is computed as the probability that the proportion of gene trees consistent with the bipartition shown in the full phylogeny is greater than 0.33, under a multinomial model with three possible outcomes, each representing a bipartition at the interior branch.

Since the branch lengths in the ASTRAL-II analysis are in terms of coalescent time units, another phylogeny was generated to get branch lengths proportional to evolutionary distances, from 1000 randomly sampled 1 kb regions across the genome using a concatenated analysis in RaxML [46], using a GTR-GAMMA model of sequence evolution.

Species split times

The divergence times between the different species were computed using the two plus two (TT) method [60], which uses a pair of samples, and the distribution of derived alleles at all sites, to compute the split time for a focal population from a contrast population. Specifically, the method uses the counts of sites in the genome where the samples fit into one of 9 configurations, i.e., both samples carry 0 derived alleles, one sample carries 1 derived allele and the other carries 0, and so on, to get an estimate of the time of either sample from the most recent common ancestor of the pair of samples. The method provides two estimates of split times for each pair

of samples, with one sample treated as the focal population and the other as the contrast population. One of the main advantages of this method is that it is not affected by the population size dynamics of the two populations after the split, but it does assume no migration and constant population size in the ancestor of the two populations (before the split).

In order to reduce the number of comparisons in this model, we chose one representative of each population for this analysis, viz., dhole – Beijing Zoo, African hunting dog – Kenya 1, golden jackal – Syria, African golden wolf Northwestern – Morocco, African golden wolf Eastern – Kenya, Ethiopian wolf – Ethiopia, coyote – California, gray wolf European – Spain, gray wolf Asian – Altai, gray wolf American – Greenland and Mexico 1, dog – India 1 and Qatar 2. The TT statistic was computed for each pair of samples, using only scaffolds longer than 1 Mb (705 in all), excluding sites with less than 5x coverage in either sample. The bootstrap estimate of the statistic and its variance was obtained treating each scaffold as a single block [61].

Mitochondrial reconstruction using *de novo* assembly

We used MtArchitect [10] to reconstruct *de novo* the mitochondrial genomes for 17 canids representing all species (for details see Data S1). The genomes were aligned using MAFFT [50] and curated with Jalview [51]. MtArchitect is designed to deal with the presence of numts, by aligning the reads to the mitochondrial and nuclear genome seperately, and including only read pairs (or single end reads), where both reads of the pair map unambiguously and with high mapping quality to the mitochondria. We tested a total of 56 phylogenetic models with jmodeltest2 [52] and chose HKY85 with gamma-distributed variation in the substitution rate and a fixed proportion of invariable sites as the most suitable model, which finally was used to construct maximum-likelihood tree using phyML [53]. We generally observed a small amount of undetermined sites, but the two African hunting dogs analyzed displayed poorer alignments and smaller genomes. This is most likely due to the reconstruction biases associated with using a distant reference and a lack of paired-end data to exploit the maximum potential of MtArchitect. Alignment visualization and tree inspection of the reconstructions confirmed that the phylogenetic clustering complied with previously reported data [4]. We observed, however, that the D-loop was particularly enriched in undetermined sites, and aligned notably worse than the remaining sequence. Given its potentially confounding nature and its small contribution to the phylogeny reconstruction when the rest of the sequence is well resolved [10], the D-loop, as well as minor positions containing the majority of the gaps, were manually discarded, resulting in a final 15.435 bp alignment.

D statistics

We used allele frequency-based *D* statistics as implemented in ADMIXTOOLS [23] to evaluate possible gene flow between the different lineages. *D* statistics are based on the observation that, if the given topology (((H1,H2), H3), Outgroup) is correct, then under the null hypothesis of no gene flow between any of the two lineages in the ingroup (H1, H2) and the lineage H3, the number of sites across the genome where the segregation patterns ABBA and BABA occur should be equal in number, as they can arise solely due to incomplete lineage sorting. But the presence of gene flow between H1 and H3 would lead to an increase in the number of BABA sites (H1 and H3 share the same allele B), while gene flow between H2 and H3 would lead to an increase in the number of ABBA sites (H2 and H3 share the same allele B). The *D* statistic measures the disparity between the number of ABBA and BABA sites across the genome to infer gene flow.

To account for the varying depth of coverage of the samples, we used a randomly sampled allele per site instead of called genotypes. Reads with mapping quality lower than 30, bases with quality lower than 20 and sites with coverage lower than 3 were discarded from the analysis. The significance of each test was estimated using a weighted block jackknife procedure over 1 Mb blocks. Deviations from $D = 0$ were presumed significant when the observed Z-score was above or below 3.3 ($|Z|>3.3$). To avoid inflating significance of the tests, only scaffolds 1 Mb or longer (~70% of the genome) were used in the analysis. Tests were performed with combinations of samples as individuals and samples were grouped into categories representing the main genetic clusters (for details see Data S1).

TreeMix

TreeMix [22] was used to infer potential admixture edges in the phylogeny. TreeMix models the correlation of allele frequencies at variable positions across the genome. The correlations that do not fit well under the modeled tree are then corrected for using migration events. We used a randomly sampled allele for each sample and a similar filtering approach as the one described for the *D* statistics tests. Tests were with combinations of samples as individuals and samples grouped into categories representing the main genetic clusters (for details see Data S1). Sites with at least one individual with coverage per group were kept. The final dataset consisted of a total of 834,537 segregating sites. We ran TreeMix on the final dataset assuming 0 to 4 migration edges (m = 0-4). For each value of m, we ran 100 replicates starting in different seed values and evaluated the replicate with the highest likelihood. Figure S3B shows the best replicate obtained for the graph modeled with four migration edges.

qpGraph

We used qpGraph from the ADMIXTOOLS package [23] to evaluate the relationships between the different species in our samples. In particular, we addressed the question of whether the African golden wolf can be modeled as a hybrid species. qpGraph uses the correlation on all possible f statistic tests in a given admixture graph to evaluate its overall fit. The same dataset and filtering parameters used for the *D* statistics tests were used in this analysis. Samples were grouped into clusters representing the main lineages in the admixture graph as indicated in Data S1. First, we started with a tree including the coyote, Ethiopian wolf, gray

wolf and Andean fox and evaluated the most likely branching point for the African golden wolf. Then, we modeled the African golden wolf as a sister clade to all possible internal and external nodes and as an admixed group from all possible node pairs. Finally, we extended our model with an admixture event to account for African golden wolf admixture in the Ethiopian wolf (Figure 4).

DATA AND SOFTWARE AVAILABILITY

The BioProject accession number for the short read sequences used in this paper is available at the NCBI short read archive under the accession PRJNA494815.

Current Biology, Volume *28*

Supplemental Information

Interspecific Gene Flow Shaped the Evolution of the Genus *Canis*

Shyam Gopalakrishnan, Mikkel-Holger S. Sinding, Jazmín Ramos-Madrigal, Jonas Niemann, Jose A. Samaniego Castruita, Filipe G. Vieira, Christian Carøe, Marc de Manuel Montero, Lukas Kuderna, Aitor Serres, Víctor Manuel González-Basallote, Yan-Hu Liu, Guo-Dong Wang, Tomas Marques-Bonet, Siavash Mirarab, Carlos Fernandes, Philippe Gaubert, Klaus-Peter Koepfli, Jane Budd, Eli Knispel Rueness, Mads Peter Heide-Jørgensen, Bent Petersen, Thomas Sicheritz-Ponten, Lutz Bachmann, Øystein Wiig, Anders J. Hansen, and M. Thomas P. Gilbert

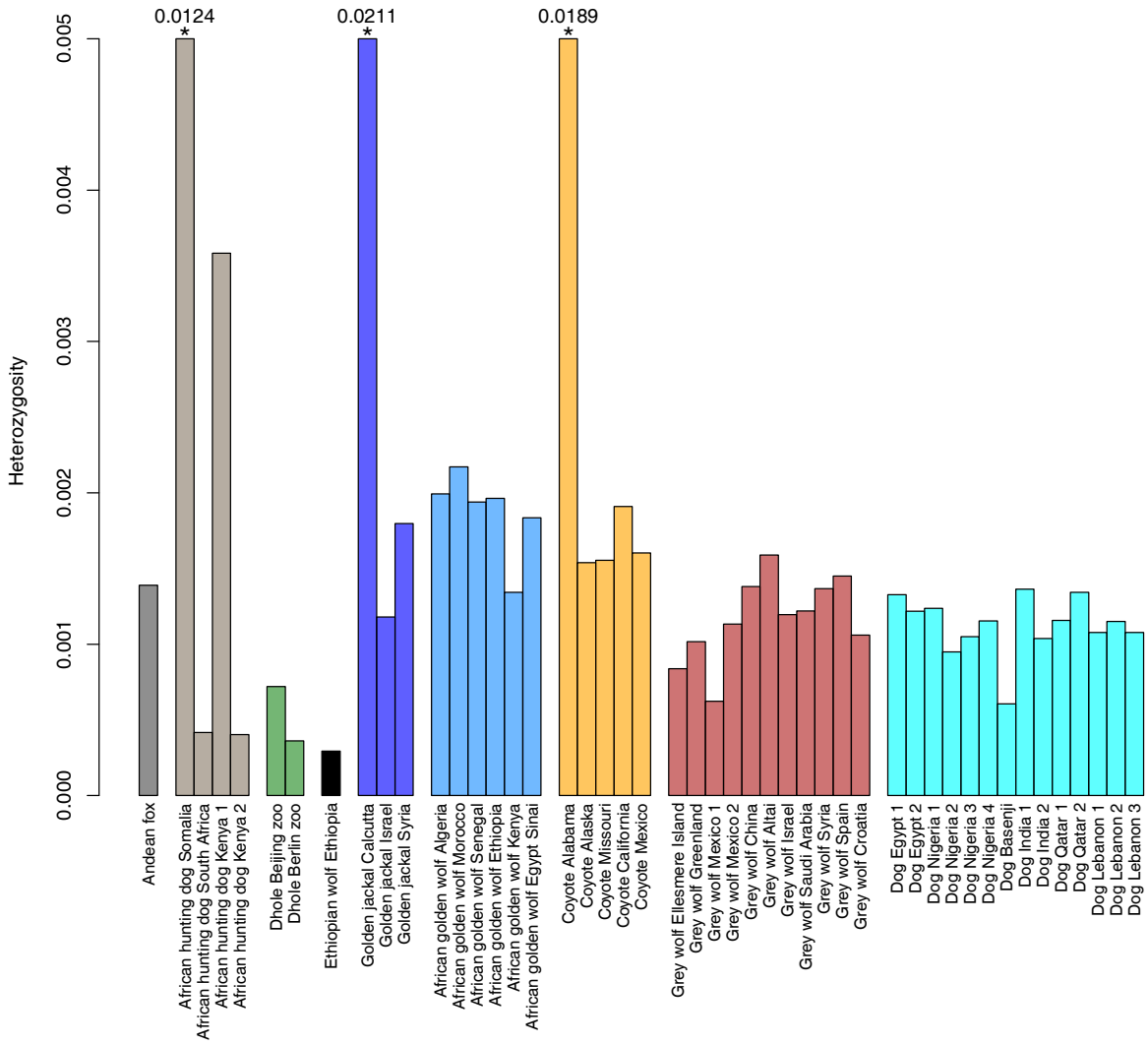


Figure S1: Heterozygosity estimates. Related to figure 1 in main article. Per sample heterozygosity, estimated under a genotype likelihood framework. Each sample is represented by a single bar. The heterozygosities range from 0.0003 to 0.0021. Heterozygosity was estimated per individual, with four individuals displaying high error rates in sequencing data (marked by '*' and with the calculated heterozygosity estimates next to them). The Ethiopian wolf, African hunting dog, and dhole are considered endangered and their wild populations are decreasing according to the IUCN [S1–S3], therefore insights into threats induced by low heterozygosity are valuable knowledge for conservation and management efforts. African golden wolves had the highest heterozygosity, ranging from 0.0013-0.0022, coyotes and golden jackals showed high values, ranging from 0.0015-0.0019 and 0.0012-0.0018 (excluding individuals with unreliable estimates), respectively. Grey wolves display a range of heterozygosity values overall, ranging from 0.0006-0.0018, the lowest being observed in a Mexican wolf. Two of the African hunting dogs that were sequenced are museum specimens, hence the DNA was degraded and unfortunately the estimated error rates in these genomes were high enough to reliably estimate heterozygosity. The other two hunting dogs display low heterozygosity values at 0.004. Both dholes show low values of heterozygosity (0.0004-0.0007), but since they are captive animals, their values might reflect low genetic variation in captive populations. Further, the pedigree of the Berlin zoo dhole indicates some level of

inbreeding. The Ethiopian wolf has the lowest heterozygosity at 0.0003, this extremely low figure being notable that it is comparable to the Mexican wolves, which descend from only 7-9 founders [S4, S5]. Thus, present-day Ethiopian wolf could potentially be subject to inbreeding depression as previously suggested [S6].

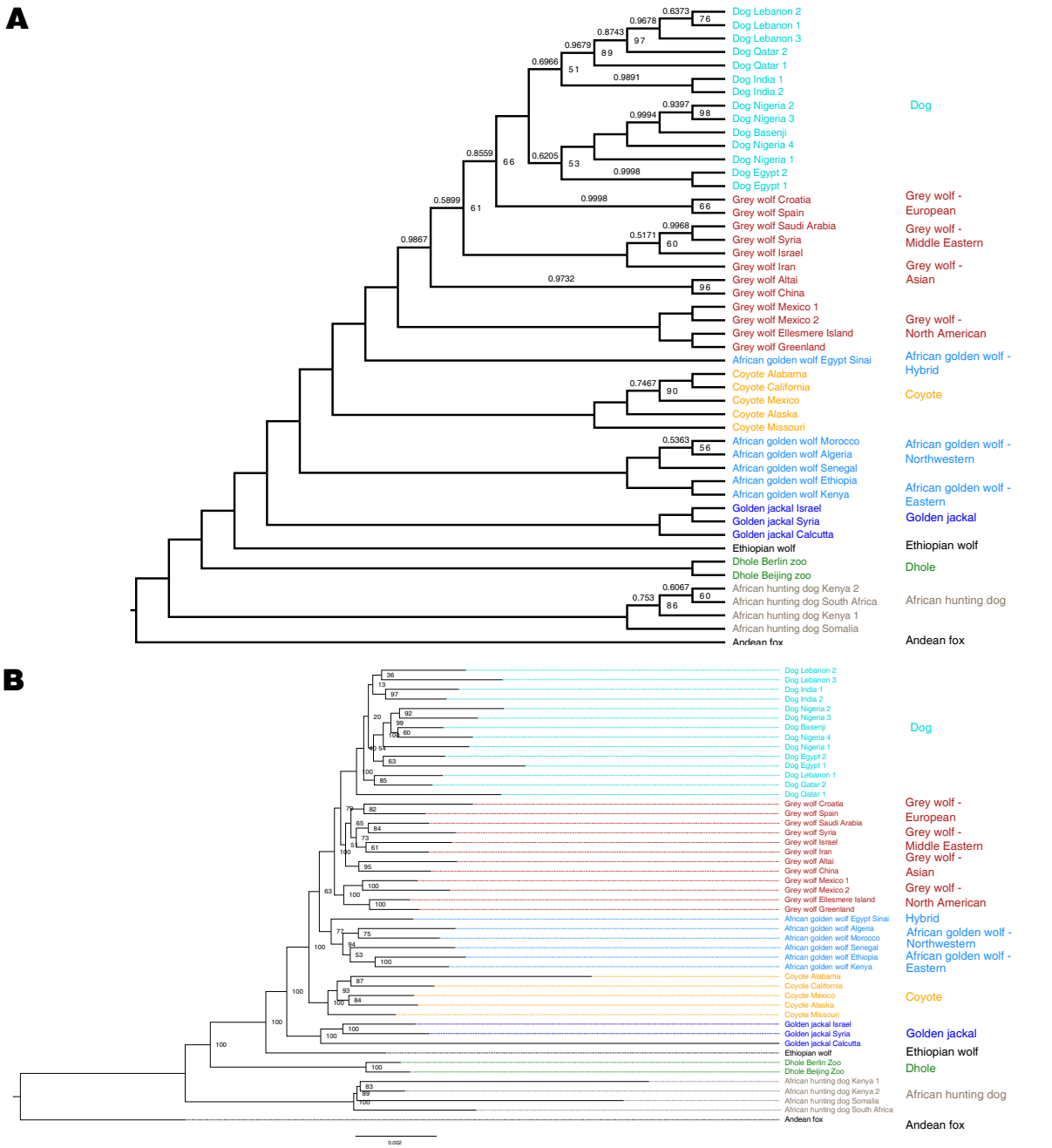
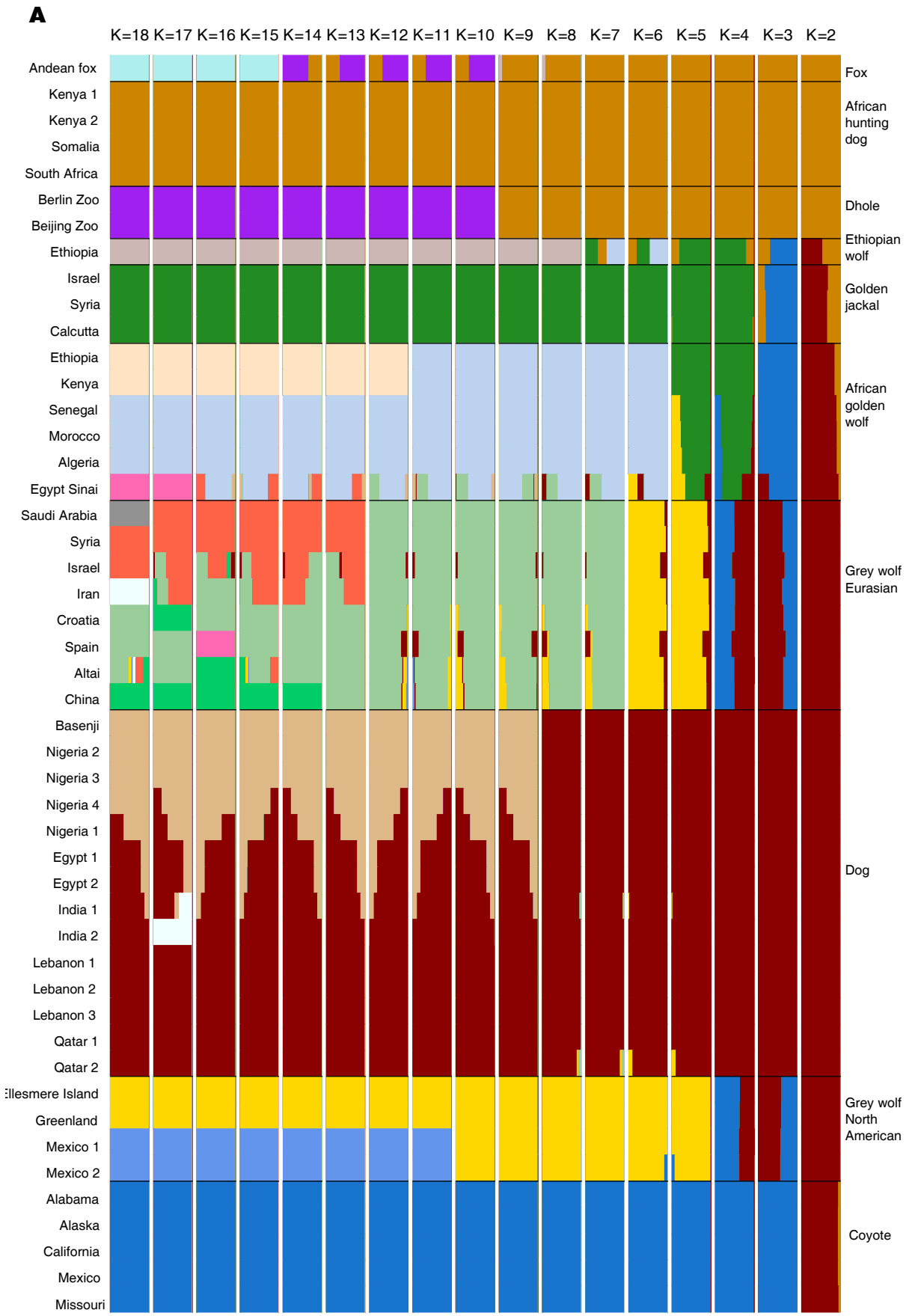


Figure S2: Phylogeny of all samples in the study. A. Full nuclear phylogeny computed using ASTRAL II. Related to figure 2 in main article. The phylogeny of all 48 samples, constructed as the consensus tree from 100 separate species trees estimated by Astral-II [S7] using gene trees obtained from FastTree2 [S8], under the GTR-GAMMA model of sequence evolution. The node labels show the bootstrap support computed by RAxML [S9] using 100 replicates obtained using Astral-II. The branch labels (decimal values above the branches) show the mean local posterior probability across the 100 Astral-II replicates. Only bootstrap

values less than 100 and mean local posterior probabilities less than 1.0 are shown in the figure. The different groups of samples that were used for collapsing in Figure 1 of the main text, are shown on the right side, with the appropriate labels and colors. The branches and node labels are colored by the species of the sample. The nuclear phylogeny is consistent with previous findings where the Ethiopian wolf was basal to both jackal species, the golden jackal and the African golden wolf. The low average local posterior probabilities and bootstrap supports for the nodes and branches leading to the split of the dogs from the Eurasian wolves shows a lack of power to resolve the topology at these nodes. Further investigation of the resolution at various branches of the phylogeny using bipartition frequencies can be found in figure 2C. Note that the branch lengths in this phylogeny are related to the coalescent times, and do not directly correspond to actual time between clades.
B. Full nuclear phylogeny using RaxML. The phylogeny of all 48 samples, constructed using the concated analyses of 1000 randomly chosen 1kb regions strewn across the genome. The tree and the bootstrap values are obtained using RAxML [9], under the GTR-GAMMA model of sequence evolution. In this tree, the branch lengths are proportional to the actual split times between the different species.


A
K=18 K=17 K=16 K=15 K=14 K=13 K=12 K=11 K=10 K=9 K=8 K=7 K=6 K=5 K=4 K=3 K=2
Andean fox
Kenya 1
Kenya 2
Somalia
South Africa
Berlin Zoo
Beijing Zoo
Ethiopia
Israel
Syria
Calcutta
Ethiopia
Kenya
Senegal
Morocco
Algeria
Egypt Sinai
Saudi Arabia
Syria
Israel
Iran
Croatia
Spain
Altai
China
Basenji
Nigeria 2
Nigeria 3
Nigeria 4
Nigeria 1
Egypt 1
Egypt 2
India 1
India 2
Lebanon 1
Lebanon 2
Lebanon 3
Qatar 1
Qatar 2
Ellesmere Island
Greenland
Mexico 1
Mexico 2
Alabama
Alaska
California
Mexico
Missouri
Fox
African hunting dog
Dhole
Ethiopian wolf
Golden jackal
African golden wolf
Grey wolf Eurasian
Dog
Grey wolf North American
Coyote

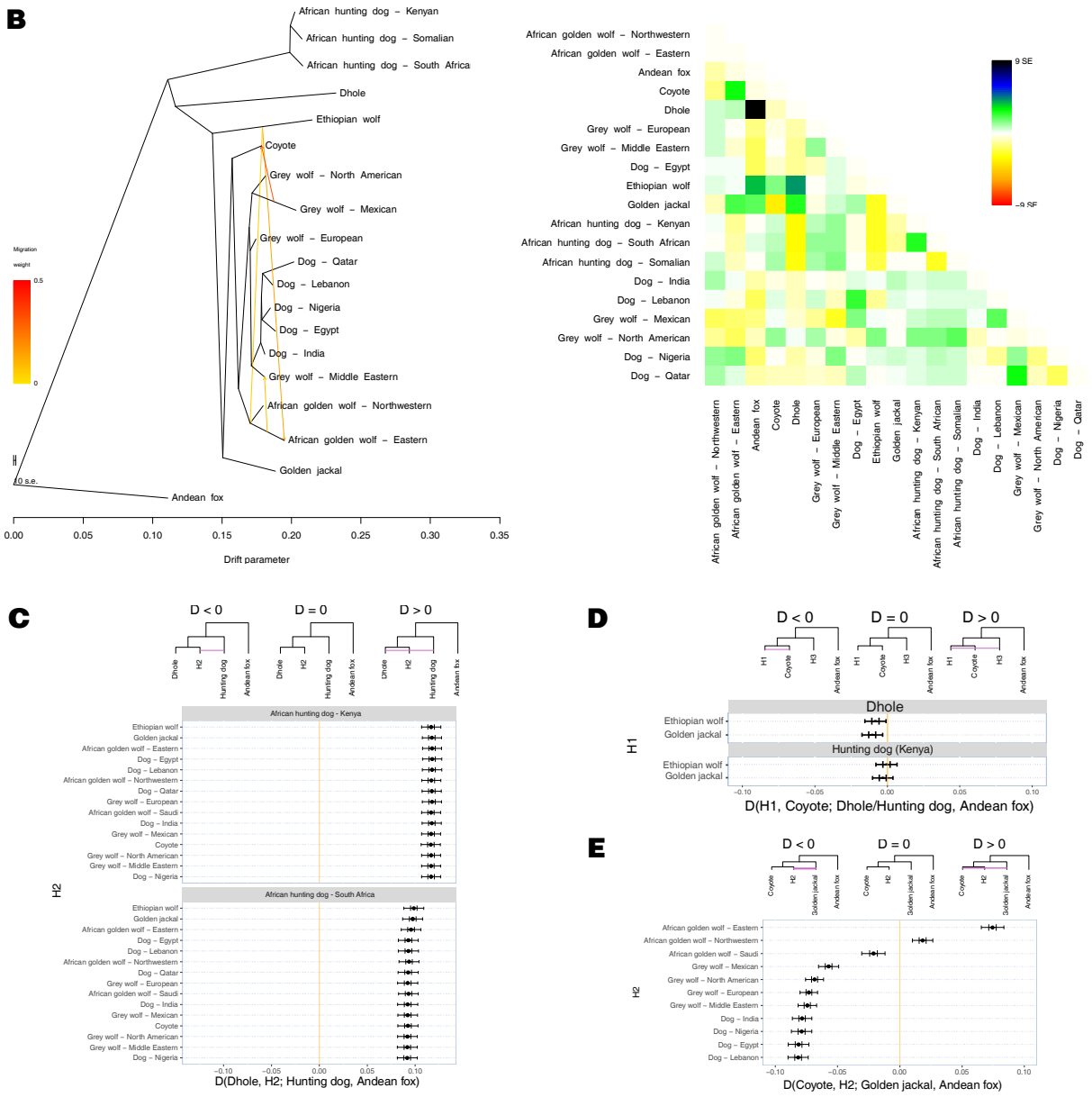


Figure S3: Evidence for extensive gene flow in the genus *Canis*. A. Related to figure 3 in main article. Admixture plots for K=2-18. Admixture analysis of all samples included in the manuscript. The plots show the ancestry proportions of the samples estimated using NGSAdmix, which performs an unsupervised clustering analysis using the genotype likelihood data for the samples, when estimating K=2-18 ancestry clusters as shown at top of figure. Each sample is represented by a line, and the admixture proportions are represented in different colors. The major clusters of populations and species are separated by solid black lines and detailed on the right of the figure. Name and location of specific samples are given on the left of the figure. One of the main findings of this analysis is the clear separation between the two clades of the African golden wolf, viz. the Eastern clade consisting of the Ethiopian and the Kenyan sample, and the Northwestern clade, which contains the Moroccan, Algerian and the Senegalese samples. Further, the African golden wolf from Sinai is a hybrid between the Middle Eastern grey wolves and the Eastern African golden wolves. Note that the clusters from K=16-18 are not stable, but they do show that the hybrid sample from Sinai

separating from the rest of the African golden wolves. **B.** A maximum likelihood tree obtained using treemix, fitting 4 migraiton edges. Treemix uses the correlation between allele frequencies in the different groups to estimate the maximum likelihood tree. Further, it uses the discrepancies between the estimated and observed allele frequency correlations to infer migration events between the groups. The tree with migration edges is shown in the left panel, whereas the residual error between the observed and fitted allele frequency correlations is shown on the right. The tree shows significant gene flow first of all between coyotes and Mexican grey wolves. Additional migration events are inferred between Ethiopian wolf and African golden woves, indicating both a migration to the root of African golden wolves, and also a second migration to the eastern population. Finally, there is a migration from African golden wolves to the dog/grey wolf/African golden wolf hybrid individual, fitting its admixed background. **C-E. *D*-statistics suggesting admixture between different species.** *D*-statistics measure the differential amount of allele sharing between two members of an ingroup (H1 and H2) with a sister group (H3), using an outgroup as a control (H4 or O). Tree-like topologies at the top of each figures represent the null hypothesis in each test, where pink lines indicate gene flow between the respective branches. D-statistics can be used to measure the deviation from the tree shown in the middle - the null hypothesis of no gene flow - using the relative abundance of different topologies, as would be induced by the trees on the left and the right - the trees with gene flow. Points indicate the D value obtained from the respective tests. Horizontal lines represent 1 (larger vertical line) and ~3.3 (smaller vertical line) standard errors estimated from a weighted block-jackknife procedure. **C. Gene flow between African hunting dog and dhole.** D-statistics tests showing significant gene flow ($|Z|>3.3$) between the African hunting dog and the dhole. D-statistics can be used to measure the deviation from the tree shown in the middle - the null hypothesis of no gene flow - using the relative abundance of different topologies, as would be induced by the trees on the left and the right - the trees with gene flow. Points indicate the *D* value obtained from test and horizontal bars show standard errors. Trees in the top are used to illustrate the null and alternative hypotheses, where pink lines indicate gene flow between the respective branches. Irrespective of the sample used in the ingroup along with the dhole, we see a statistically significant signal of gene flow between the African hunting dog and dhole. **D. D-statistics for grey wolves, coyotes and Canid X.** D-statistics tests to compute gene flow between coyote and the dhole or the African hunting dog. The trees at the top show the configuration of samples tested along with the gene flow suggested by the D-statistics. D-statistics can be used to measure the deviation from the tree shown in the middle - the null hypothesis of no gene flow - using the relative abundance of different topologies, as would be induced by the trees on the left and the right - the trees with gene flow. The D-statistic and its estimated standard error bars are plotted in the panel. Note that only the Ethiopian wolf and the golden jackal were used in the position H1, due to evidence of gene flow between the grey wolves, African golden wolf and golden jackal, which obfuscate the findings presented here. The upper two D-statistics, with the dhole in the position H3, coyote in position H2, and either the Ethiopian wolf or the golden jackal in position H2, show a small but significant signal for gene flow between coyote and dhole. However, when replacing the dhole with the African hunting dog, the signal of gene flow between the coyote and the species in H3, the African hunting dog disappears. We suggest that, rather than actual gene flow, the patterns are

evidence of admixture between coyote and an unknown canid taxon, which would be phylogenetically placed between the dhole and the African hunting dog. Such an admixture event would lead to outgroup attraction of the coyote lineage as long as the species used in H3 is closer to the ingroup than the unknown canid. Although the grey wolf is not used in this analysis, the results in subsequent analysis suggest that the grey wolves show a signal of gene flow from a similar unknown canid taxon. **E.** D-statistic showing a significant gene flow (|Z| >3.3) between golden jackal and grey wolf, dog and the African golden wolf compared to coyote. We note that in some tests where the African golden wolf is placed in H2 yielded positive results suggesting introgression between coyote and golden jackal, however we explain this as a consequence of the gene flow between the African golden wolf and the Ethiopian wolf (a more basal species), which results in outgroup attraction.

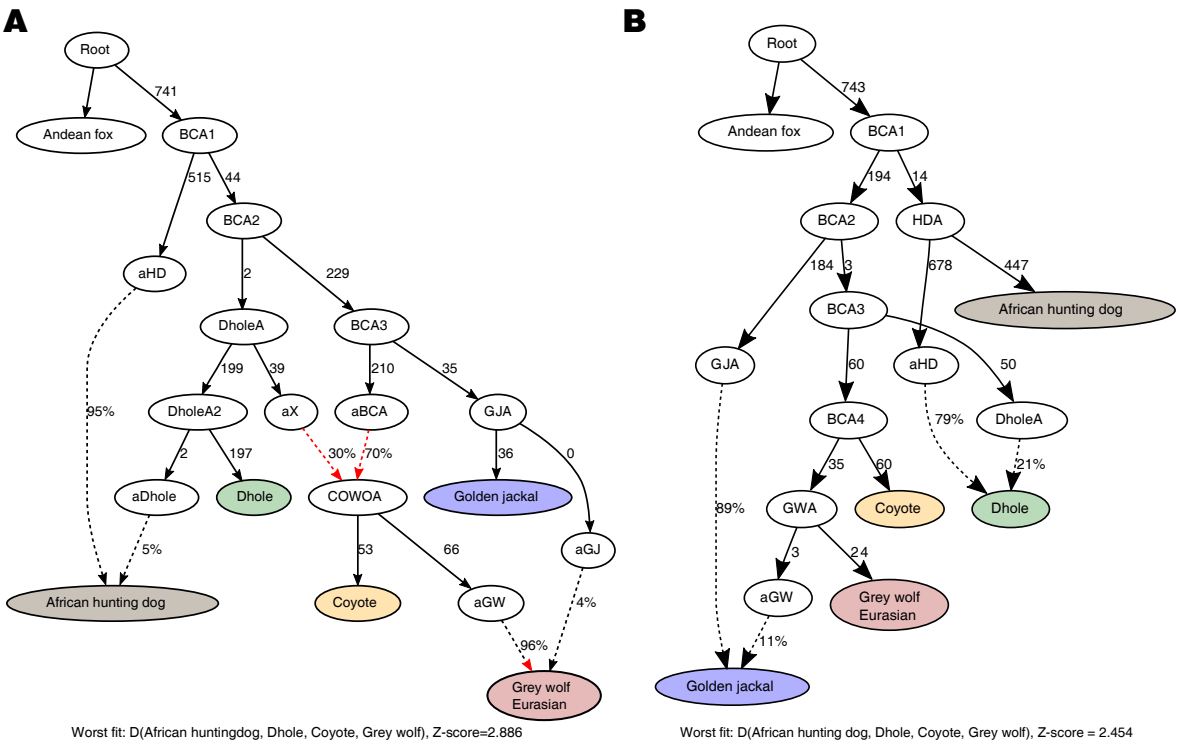


Figure S4: Two possible admixture graphs modelling excess basal ancestry in coyotes and grey wolves. Related to figure 4 in main article. QPgraph estimates the admixture graph using all pairwise D-statistics for a given set of populations. In these graphs, the excess genetic affinity between the coyote/grey wolf and the dhole, as shown in figure S8, is explained using two different admixture graphs. The first graph **A.** explains this excess ancestry by modelling the ancestor of the coyotes and the grey wolves as an admixture (arrows shown in red) between a lineage related to golden jackals and another related to population X, which is closely related to the dholes. The second admixture graph **B.** explains the excess basal ancestry in coyotes and grey wolves by incorporating an admixture edge from the ancestor of coyotes and grey wolves into the dhole, with the dhole deriving ~21% of its ancestry from this admixture event. Both these graphs fit the data well, in terms of outlier f-statistics, suggesting that there are multiple models that explain the data. In both these cases, we see a population related to the dholes sharing excess ancestry with the ancestor of coyotes and grey wolves.

Parameter	Model without migrations	Model with migrations
Population size parameters		
Ethiopian wolf (EW)	7.076×10^{-4}	6.406×10^{-4}
African golden wolf (AGW)	0.407	0.436
Golden jackal (GJ)	0.229	0.247
AGW - GJ ancestor	56.112	49.347
EW - AGW - GJ ancestor	3.747	3.618
Split time parameters		
AGW - GJ	6.443×10^{-2}	7.055×10^{-2}
EW - (AGW - GJ)	0.998	1.049
Migration parameters		
AGW → GJ	-	11.219
GJ → AGW	-	5.984
EW → AGW	-	9.402
AGW → EW	-	**1161.618**
EW → GJ	-	6.769
GJ → EW	-	**53.037**
EW → (AGW - GJ)	-	2.535
(AGW - GJ) → EW	-	**250.97**

Table S1: G-PhoCS estimates of populations sizes, migration rates and split times for Ethiopian wolf, African golden wolf and golden jackal. Related to figures 2 and 3 in the main article. Two different models demographic models were estimated for the three species included in the G-PhoCS analysis, viz. Ethopian wolf, African golden wolf and golden jackal. No migration rates were included in the first model, whereas the second model allowed for migrations between all pairs of species. The population size parameters are the effective population sizes (N_e) estimates, given in number of chromosomes, using G-PhoCS. The split times are given in coalescent scaled time units ($2N_e$ generations) and the migration rates are presented in scaled migration rates ($4N_em$, where N_em is the number of migrants per generation). All the parameter estimates are scaled by population sizes and mutation rates. High migration rates are shown in **bold** in the table.

Group 1	Group 2	T1 (ky)	SE(T1) (ky)	T2 (ky)	SE(T2) (ky)
Dhole	Coyote	1747.49	14.06	1972.11	14.06
Dhole	Ethiopian wolf	2009.06	46.73	2236.95	46.73
Dhole	Grey wolf American	1765.02	31.33	2016.93	31.33
Dhole	African hunting dog	2503.99	29.16	2328.75	29.16
Dhole	Grey wolf European	1730.51	20.18	1974.19	20.18
Dhole	Dog India	1763.38	28.72	2006.97	28.72
Dhole	African golden wolf Eastern	1804.97	25.81	1991.71	25.81
Dhole	Grey wolf Mexican	1910.02	23.42	2150.49	23.42
Dhole	African golden wolf Northwestern	1660.32	24.10	1890.30	24.10
Dhole	Dog Qatar	1648.25	81.48	1864.63	81.48
Dhole	Golden jackal	1811.13	17.92	1974.51	17.92
Dhole	Grey wolf Asian	1678.14	20.11	1949.04	20.11
Coyote	Ethiopian wolf	789.10	12.70	792.12	12.70
Coyote	Grey wolf American	231.82	0.99	261.56	0.99
Coyote	African hunting dog	2125.65	1.68	1713.55	1.68
Coyote	Grey wolf European	234.30	0.83	259.15	0.83
Coyote	Dog India	248.85	0.48	274.62	0.48
Coyote	African golden wolf Eastern	384.66	0.93	350.29	0.93
Coyote	Grey wolf Mexican	258.85	5.19	275.92	5.19
Coyote	African golden wolf Northwestern	303.94	0.47	310.25	0.47
Coyote	Dog Qatar	208.84	0.67	239.56	0.67
Coyote	Golden jackal	595.32	1.72	527.24	1.72
Coyote	Grey wolf Asian	222.16	0.55	269.11	0.55
Ethiopian wolf	Grey wolf American	865.05	52.44	891.65	52.44
Ethiopian wolf	African hunting dog	2467.13	36.70	2048.28	36.70

Ethiopian wolf	Grey wolf European	773.04	51.94	794.11	51.94
Ethiopian wolf	Dog India	792.86	28.12	813.43	28.12
Ethiopian wolf	African golden wolf Eastern	691.25	29.93	655.69	29.93
Ethiopian wolf	Grey wolf Mexican	961.68	13.78	977.06	13.78
Ethiopian wolf	African golden wolf Northwestern	659.32	7.63	661.33	7.63
Ethiopian wolf	Dog Qatar	726.49	43.64	746.98	43.64
Ethiopian wolf	Golden jackal	962.56	36.30	893.38	36.30
Ethiopian wolf	Grey wolf Asian	717.96	13.09	761.77	13.09
Grey wolf American	African hunting dog	2186.56	1.97	1739.41	1.97
Grey wolf American	Grey wolf European	21.15	4.01	15.99	4.01
Grey wolf American	Dog India	22.42	1.01	18.59	1.01
Grey wolf American	African golden wolf Eastern	345.22	6.69	280.97	6.69
Grey wolf American	Grey wolf Mexican	20.21	36.09	7.38	36.09
Grey wolf American	African golden wolf Northwestern	204.57	0.52	179.97	0.52
Grey wolf American	Dog Qatar	16.72	1.24	21.16	1.24
Grey wolf American	Golden jackal	598.10	5.78	500.47	5.78
Grey wolf American	Grey wolf Asian	-4.65	0.49	12.89	0.49
African hunting dog	Grey wolf European	1725.73	0.59	2161.78	0.59
African hunting dog	Dog India	1732.71	2.83	2175.63	2.83
African hunting dog	African golden wolf Eastern	1740.22	4.18	2113.24	4.18
African hunting dog	Grey wolf Mexican	1822.73	5.78	2255.17	5.78
African hunting dog	African golden wolf Northwestern	1694.51	1.07	2115.61	1.07
African hunting dog	Dog Qatar	1621.68	6.82	2027.96	6.82
African hunting dog	Golden jackal	1846.80	3.85	2202.99	3.85
African hunting dog	Grey wolf Asian	1715.43	0.23	2188.19	0.23
Grey wolf European	Dog India	-5.26	0.67	-3.94	0.67

Grey wolf European	African golden wolf Eastern	321.28	2.32	262.28	2.32
Grey wolf European	Grey wolf Mexican	39.88	15.54	31.53	15.54
Grey wolf European	African golden wolf Northwestern	188.25	0.46	169.50	0.46
Grey wolf European	Dog Qatar	-10.83	1.19	-0.25	1.19
Grey wolf European	Golden jackal	570.18	3.75	478.52	3.75
Grey wolf European	Grey wolf Asian	-22.80	0.37	-0.99	0.37
Dog India	African golden wolf Eastern	308.65	1.76	248.78	1.76
Dog India	Grey wolf Mexican	30.73	3.59	21.56	3.59
Dog India	African golden wolf Northwestern	196.73	0.33	175.93	0.33
Dog India	Dog Qatar	-119.24	0.60	-109.49	0.60
Dog India	Golden jackal	567.63	1.26	476.04	1.26
Dog India	Grey wolf Asian	-2.89	0.28	18.07	0.28
African golden wolf Eastern	Grey wolf Mexican	310.84	4.17	362.18	4.17
African golden wolf Eastern	African golden wolf Northwestern	108.44	0.31	147.67	0.31
African golden wolf Eastern	Dog Qatar	234.83	1.86	297.45	1.86
African golden wolf Eastern	Golden jackal	601.53	5.44	570.59	5.44
African golden wolf Eastern	Grey wolf Asian	225.87	1.11	307.53	1.11
Grey wolf Mexican	African golden wolf Northwestern	210.11	3.61	199.22	3.61
Grey wolf Mexican	Dog Qatar	0.81	3.53	17.05	3.53
Grey wolf Mexican	Golden jackal	588.54	6.56	503.33	6.56
Grey wolf Mexican	Grey wolf Asian	0.84	4.14	30.77	4.14
African golden wolf Northwestern	Dog Qatar	152.09	0.58	177.06	0.58

African golden wolf Northwestern	Golden jackal	577.03	0.92	503.35	0.92
African golden wolf Northwestern	Grey wolf Asian	156.31	0.29	199.62	0.29
Dog Qatar	Golden jackal	519.46	2.45	433.07	2.45
Dog Qatar	Grey wolf Asian	-7.92	0.21	3.02	0.21
Golden jackal	Grey wolf Asian	457.99	1.60	574.52	1.60

Table S2: Split time estimates from TT method. Related to figures 2, 3 and 4 in the main article. The sample pairwise split time estimates obtained from the two plus two method [S16] using the counts of the different configurations of number of derived alleles carried per site (9 in all, with 7 being variable). The Andean fox was used as outgroup to figure out the derived allele. The column T1 is the block jackknife estimate of the split time using the species/population in the first column as the focus population and the species/population in the second column as the contrast population, while the column T2 is the split time obtained when the roles of the populations are reversed. Negative values (shown in red) indicate that these split times cannot be accurately estimated using this method, either due to gene flow or due to changing population sizes in the ancestral population. The estimates and their standard errors are scaled to 1000 years (ky).

Supplemental references

[S1] Marino, J., and Sillero-Zubiri, C. (2011). Canis simensis. The IUCN Red List of Threatened Species 2011: e.T3748A10051312. http://dx.doi.org/10.2305/IUCN.UK.2011-1.RLTS.T3748A10051312.en. Downloaded on 05 April 2017.

[S2] Woodroffe, R., and Sillero-Zubiri, C. (2012). Lycaon pictus. The IUCN Red List of Threatened Species 2012: e.T12436A16711116. http://dx.doi.org/10.2305/IUCN.UK.2012.RLTS.T12436A16711116.en. Downloaded on 05 April 2017.

[S3] Kamler, J.F., Songsasen, N., Jenks, K., Srivathsa, A., Sheng, L., and Kunkel, K. (2015). Cuon alpinus. The IUCN Red List of Threatened Species 2015: e.T5953A72477893. http://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T5953A72477893.en. Downloaded on 05 April 2017.

[S4] Garcia-Moreno, J., Matocq, M.D., Roy, M.S., Geffen, E., and Wayne, R.K. (1996). Relationships and Genetic Purity of the Endangered Mexican Wolf Based on Analysis of Microsatellite Loci. Conserv. Biol. 10, 376–389.

[S5] Hedrick, P.W., and Fredrickson, R.J. (2008). Captive breeding and the reintroduction of Mexican and red wolves. Mol. Ecol. 17, 344–350.

[S6] Gottelli, D., Sillero-Zubiri, C., Marino, J., Funk, S.M., and Wang, J. (2013). Genetic structure and patterns of gene flow among populations of the endangered Ethiopian wolf. Anim. Conserv. 16, 234–247.

[S7] Mirarab, S., and Warnow, T. (2015). ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics 31, i44–52.

[S8] Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One 5, e9490.

[S9] Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

[S10] Auton, A., Rui Li, Y., Kidd, J., Oliveira, K., Nadel, J., Holloway, J.K., Hayward, J.J., Cohen, P.E., Greally, J.M., Wang, J., et al. (2013). Genetic recombination is targeted towards gene promoter regions in dogs. PLoS Genet. 9, e1003984.

[S11] Fan, Z., Silva, P., Gronau, I., Wang, S., Armero, A.S., Schweizer, R.M., 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, 163–173.

[S12] Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E., Silva, P.M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., et al. (2014). Genome sequencing highlights the dynamic early history of dogs. PLoS Genet. 10, e1004016.

[S13] Liu, Y.-H., Wang, L., Xu, T., Guo, X., Li, Y., Yin, T.-T., Yang, H.-C., Yang, H., Adeola, A.C., J Sanke, O., et al. (2017). Whole-genome sequencing of African dogs provides insights into adaptations against tropical parasites. Mol. Biol. Evol. DOI: 10.1093/molbev/msx258.

[S14] Koepfli, K.-P., Pollinger, J., Godinho, R., Robinson, J., Lea, A., Hendricks, S., Schweizer, R.M., 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, 2158–2165.

[S15] Schweizer, R.M., vonHoldt, B.M., Harrigan, R., Knowles, J.C., Musiani, M., Coltman, D., Novembre, J., and Wayne, R.K. (2016). Genetic subdivision and candidate genes under selection in North American grey wolves. Mol. Ecol. 25, 380–402.

[S16] Schlebusch, C.M., Malmström, H., Günther, T., Sjödin, P., Coutinho, A., Edlund, H., Munters, A.R. et al. 2017. “Southern African Ancient Genomes Estimate Modern Human Divergence to 350,000 to 260,000 Years Ago.” Science 358 (6363): 652–55.

[S17] Wang, G.-D., Zhai, W., Yang, H.-C., Fan, R.-X., Cao, X., Zhong, L., Wang, L., Liu, F., Wu, H., Cheng, L.-G., et al. (2013). The genomics of selection in dogs and the parallel evolution between dogs and humans. Nat. Commun. 4, 1860.

[S18] Wang, G.-D., Zhai, W., Yang, H.-C., Wang, L., Zhong, L., Liu, Y.-H., Fan, R.-X., Yin, T.-T., Zhu, C.-L., Poyarkov, A.D., et al. (2016). Out of southern East Asia: the natural history of domestic dogs across the world. Cell Res. 26, 21–33.

[S19] Campana, M.G., Parker, L.D, Hawkins, M.T.R, Young, H.S., Helgen K.M., Gunther, M.S., Woodroffe, R., Maldonado, J.E., Fleischer, R.C. (2016). Genome sequence, population history, and pelage genetics of the endangered African wild dog (Lycaon pictus). BMC Genomics 17:1013.