Report

Current Biology

Evolutionary history of the extinct Sardinian dhole

Highlights

- The Sardinian dhole lineage diverged from the Asian dholes around 885 ka
- Post-divergence gene flow between the dhole lineages ended between 560 and 310 ka
- Our sample also showed a reduced genome-wide diversity
- The gene flow between *Lycaon* and the dholes' ancestor ended around 1.05–0.83 Ma

Authors

Marta Maria Ciucani, Julie Kragmose Jensen, Mikkel-Holger S. Sinding, ..., Love Dalén, M. Thomas P. Gilbert, Shyam Gopalakrishnan

Correspondence

ciucani@sund.ku.dk (M.M.C.), shyam.gopalakrishnan@sund.ku.dk (S.G.)

In brief

Ciucani et al. sequence the genome of a 21,100-year-old Sardinian dhole. The sample represents a different lineage from the current Asian dholes and split from them *ca* 885 ka. Post-divergence gene flow between the two lineages ended later—*ca* 560–310 ka—followed by a stable but long-term population decline and extinction of the Sardinian dhole.

Ciucani et al., 2021, Current Biology *31*, 5571–5579
December 20, 2021
https://doi.org/10.1016/j.cub.2021.09.059

CellPress

Current Biology

CellPress

Report

Evolutionary history of the extinct Sardinian dhole

Marta Maria Ciucani,[1,*] Julie Kragmose Jensen,[2] Mikkel-Holger S. Sinding,[3] Oliver Smith,[1,4] Saverio Bartolini Lucenti,[5,6] Erika Rosengren,[7] Lorenzo Rook,[5] Caterinella Tuveri,[8] Marisa Arca,[8] Enrico Cappellini,[1] Marco Galaverni,[9] Ettore Randi,[10] Chunxue Guo,[11] Guojie Zhang,[11,12,13,14] Thomas Sicheritz-Pontén,[1,15] Love Dalén,[16,17] M. Thomas P. Gilbert,[1,18] and Shyam Gopalakrishnan[1,2,19,*]

[1]Section for Evolutionary Genomics, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark
[2]Department of Health Technology, Technical University of Denmark, Kongens Lyngby, Denmark
[3]Smurfit Institute of Genetics, Trinity College Dublin, Dublin 2, Ireland
[4]Micropathology Ltd, University of Warwick Science Park, Coventry, UK
[5]Dipartimento di Scienze della Terra, Paleo[Fab]Lab, Università di Firenze, Via G. La Pira 4, 50121 Firenze, Italy
[6]Sezione di Geologia e Paleontologia, Museo di Storia Naturale, Università degli Studi di Firenze, Via G. La Pira 4, 50121 Firenze, Italy
[7]Department of Archaeology and Ancient History, Lund University, Helgonavägen 3, Box 192, 221 00 Lund, Sweden
[8]Soprintendenza Archeologia, Belle Arti e Paesaggio per le province di Sassari e Nuoro (Ufficio Operativo di Nuoro), Via G. Asproni 8, 08100 Nuoro, Italy
[9]WWF Italy, Science Unit, Rome, Italy
[10]Department of Chemistry and Bioscience, Faculty of Engineering and Science, University of Aalborg, Aalborg, Denmark
[11]BGI-Shenzhen, Shenzhen 518083, China
[12]Villum Center for Biodiversity Genomics, Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Denmark
[13]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China
[14]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, 32 Jiaochang Donglu, Kunming 650223, China
[15]Centre of Excellence for Omics-Driven Computational Biodiscovery (COMBio), Faculty of Applied Sciences, AIMST University, Batu 3 1/2, Butik Air Nasi, 08100 Bedong, Kedah, Malaysia
[16]Centre for Palaeogenetics, Svante Arrhenius väg 20C, 10691 Stockholm, Sweden
[17]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Box 50007, 10405 Stockholm, Sweden
[18]University Museum, Norwegian University of Science and Technology, Trondheim, Norway
[19]Lead contact
*Correspondence: ciucani@sund.ku.dk (M.M.C.), shyam.gopalakrishnan@sund.ku.dk (S.G.)
https://doi.org/10.1016/j.cub.2021.09.059

SUMMARY

The Sardinian dhole (*Cynotherium sardous*)[1] was an iconic and unique canid species that was endemic to Sardinia and Corsica until it became extinct at the end of the Late Pleistocene.[2–5] Given its peculiar dental morphology, small body size, and high level of endemism, several extant canids have been proposed as possible relatives of the Sardinian dhole, including the Asian dhole and African hunting dog ancestor.[3,6–9] Morphometric analyses[3,6,8–12] have failed to clarify the evolutionary relationship with other canids.We sequenced the genome of a *ca*-21,100-year-old Sardinian dhole in order to understand its genomic history and clarify its phylogenetic position. We found that it represents a separate taxon from all other living canids from Eurasia, Africa, and North America, and that the Sardinian dhole lineage diverged from the Asian dhole *ca* 885 ka. We additionally detected historical gene flow between the Sardinian and Asian dhole lineages, which ended approximately 500-300 ka, when the land bridge between Sardinia and mainland Italy was already broken, severing their population connectivity. Our sample showed low genome-wide diversity compared to other extant canids—probably a result of the long-term isolation—that could have contributed to the subsequent extinction of the Sardinian dhole.

RESULTS AND DISCUSSION

We successfully re-sequenced the genome of a Sardinian dhole (SD) specimen from Corbeddu Cave (Sardinia) (Figures 1A and 1B) to an average coverage of *ca* 5× (Table S2). This sample has been radiocarbon dated to *ca* 21,000 calibrated years before present (Figure 1C), and the analyses of the mapped reads showed high DNA damage levels (Figures S1A–S1C).

Comparing the depths of coverage of the mapped reads on the sex chromosomes and autosomes of the domestic dog reference genome (CanFam3.1),[13] we found that our sample derived from a female Sardinian dhole (Figure S1E).

To place the Sardinian dhole in an evolutionary context along other canids, we analyzed it together with 49 previously published canid genomes (Table S3). The samples used for the reference dataset span from 4.4× to 28.2× genome-wide coverage and cover the species diversity of the Eurasian, African, and American wolf-like canids, including African hunting dogs (*Lycaon pictus*), Asian dholes (*Cuon alpinus*), an Ethiopian wolf (*Canis simensis*), coyotes (*Canis latrans*), African golden wolves

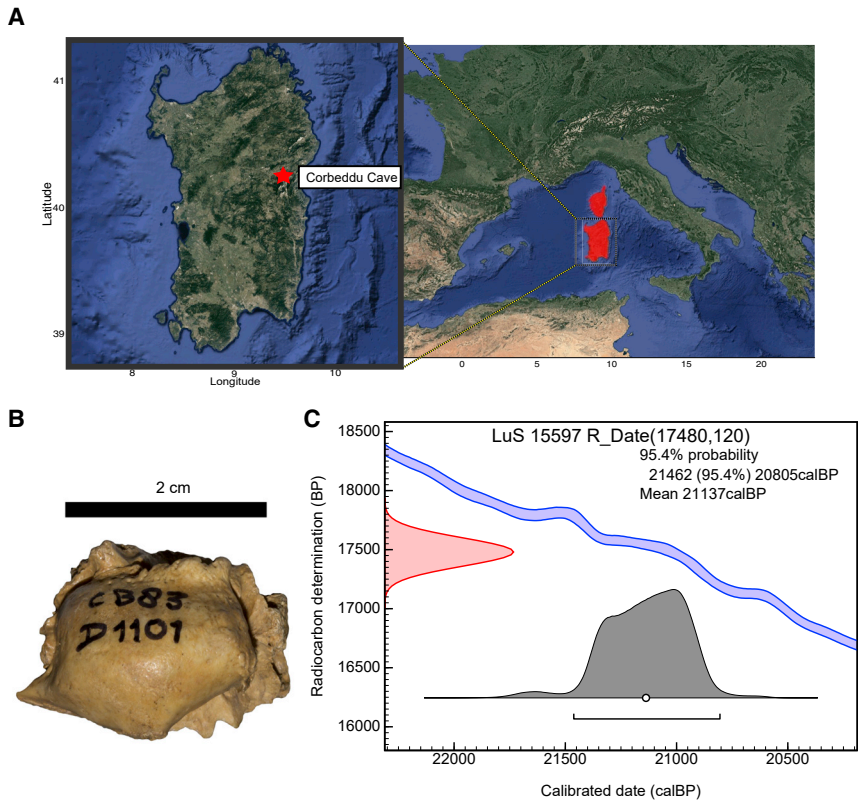

Figure 1. Information related to the sample

(A) Sampling location (inset) and hypothetical distribution range of *Cynotherium*. Inset: exact location of the archaeological site (Corbeddu Cave) in which the Sardinian dhole was excavated.

(B) Picture of the petrous bone analyzed in this study.

(C) Radiocarbon dating results of the petrous bone fragment. The y axis shows radiocarbon concentration expressed in years "before present" (BP), and the x axis shows calibrated years BP (derived from the tree ring data). The pair of blue curves shows the radiocarbon measurements on the tree rings (plus and minus one standard deviation). The red curve on the left indicates the radiocarbon concentration in the sample. The gray histogram shows possible ages for the sample (the higher the histogram, the more likely that age is).

(*Canis lupaster*), golden jackals (*Canis aureus*), gray wolves (*Canis lupus*), domestic dogs (*C. l. familiaris*), and an Andean fox (*Lycalopex culpaeus*) as the outgroup. Three more ancient samples of Dire wolf (*Canis dirus*) and Eurasian dhole (*Cuon alpinus*) were used only to estimate a mitochondrial phylogeny.

To explore the phylogenomic placement of the Sardinian dhole among other canids, and especially understand its relationship with the Asian dhole, we used ASTRAL-III[14] to estimate the species tree of the canids included in this study by combining 1,000 gene trees estimated from randomly chosen 5-kb regions across the nuclear genome. The estimated species tree was rooted using the Andean fox as the outgroup (Figures 2A and S2C). In the multispecies coalescent tree estimated by ASTRAL-III, the Sardinian dhole formed a distinct lineage inside the Asian dhole and sister to *Canis*. The mitochondrial phylogeny, computed using the 50 mitogenomes in the dataset, placed the Sardinian dhole basal to both modern and ancient dholes (Figure 2B), confirming the relatedness to the *Cuon* lineage and the divergence from it before the diversification between the ancient and modern populations of *Cuon*.

We subsequently tested the discordance between the species tree and the gene trees—based on the nuclear genome—to quantify the uncertainty of the branch that split the Sardinian dhole, Asian dhole, *Canis*, and basal canids. The frequencies of the three bipartitions induced by the aforementioned branch (identified as branch 16 in Figure 2A) are shown in Figure 2C, along with similar measures for all the internal branches of the species tree. Two of the three possible bipartitions induced by branch 16 have a frequency greater than 33%—the cutoff previously shown to be required for identifying the true topology.[15] Essentially, although gene trees clustering the Sardinian dhole and the Asian dholes are more likely, both topologies, i.e., Sardinian Dhole clustering with Asian dholes or *Canis*, are observed in more than 33% of all gene trees, implying that both topologies could represent the true phylogeny. Further, we conducted a Twisst[16] analysis, which performs topology weighting by iterative sampling of sub-trees to estimate the proportion of the genome supporting the different phylogenetic placements of the Sardinian dhole. Results from Twisst estimated that around 60% of the genome supports the clustering of the Sardinian dhole with the Asian dhole (Figures S3A and S3B).

We then performed a principal-component analysis (PCA) on the 46 individuals (excluding the outgroup). Consistent with previous studies,[17] we found that along PC1 (55.9%) and PC2 (8.45%), the African hunting dogs (AHDs) cluster together and are differentiated from the other canids included in this study. At the same time, the Sardinian dhole holds a distinct placement whereby, along the first two principal components, it is placed near the two representative modern Asian dholes (Figure 2D). The second principal component separates the genus *Canis* from the dhole samples. Similarly, when we performed PCA excluding the AHDs, the first component placed the Sardinian dhole between the Asian dhole and the *Canis* group (Figure S2A). To further investigate the relationships between the Sardinian dhole (*Cynotherium sardous*) and other canids, we performed an admixture analysis based on genotype likelihoods for all the samples, excluding the Andean fox (Figures 2E and S2B). Clearly, when only two ancestral clusters (K = 2) were estimated, the earliest branching lineage—AHDs—separated from the *Canis* genus,[17,18] with the two dhole species being represented as mixtures of these two clusters. Upon increasing the number of estimated ancestry clusters to four (K = 4), we observed a division between the AHDs, the dholes, gray wolves/dogs, and the rest of the canids. Specifically, the AHD maintains the same structure while the Sardinian dhole and the Asian dholes are grouped into a cluster of their own. Upon increasing the number of estimated ancestry clusters, the samples fall into clusters

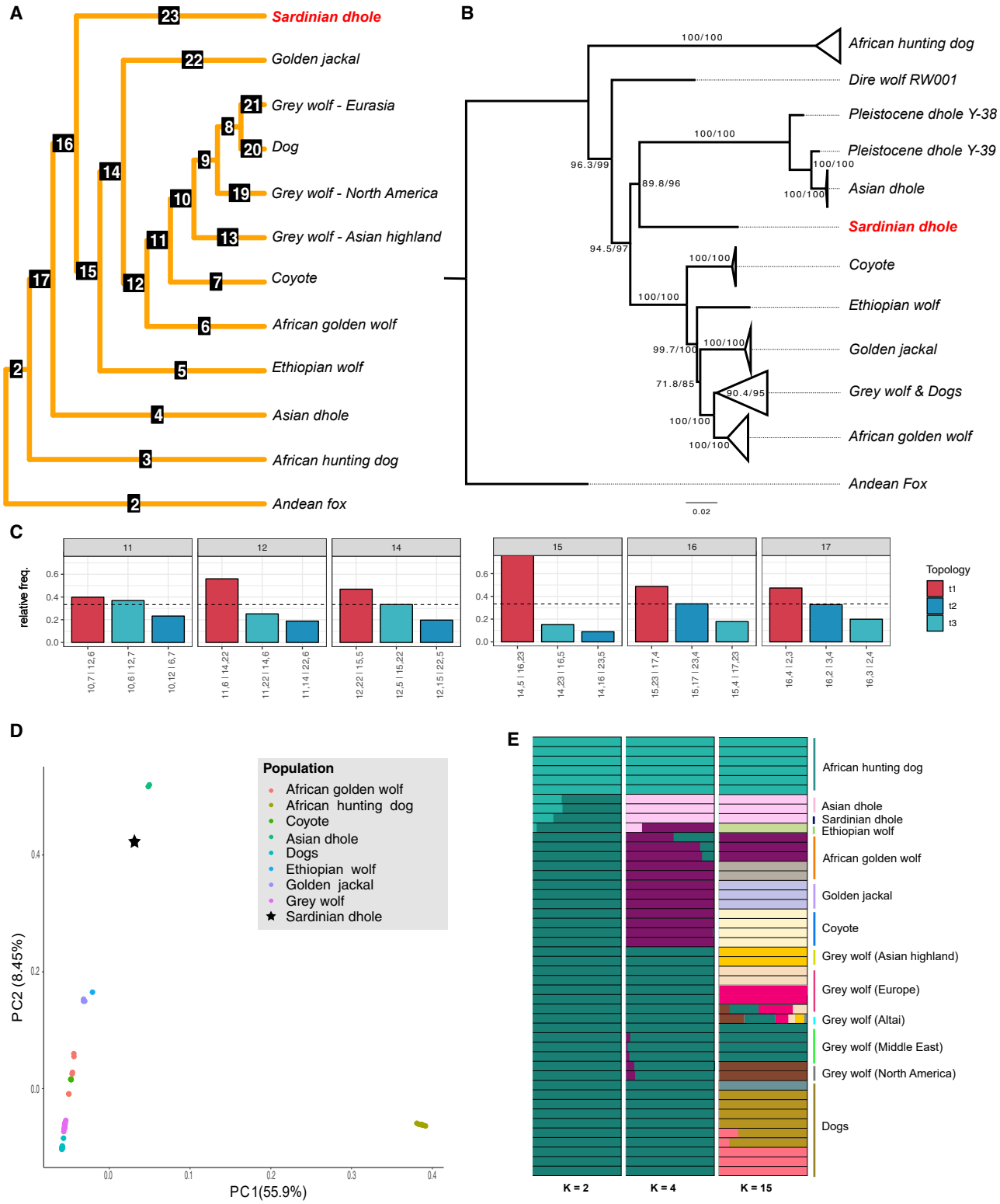

Figure 2. The structure of the canids' genomic diversity

(A) Species tree phylogeny generated by Astral-III estimated for 1,000 5-kb genomic regions. The tree was rooted on the Andean fox (*Lycalopex culpaeus*). Monophyletic clusters were collapsed into the same leaf node.

(B) Mitochondrial phylogeny generated using IQ-TREE2 on the canids in this study, including the dire wolf and two Pleistocene dholes. Numbers on the branches represent SH-aLRT (Shimodaira-Hasegawa approximate likelihood ratio test) support (%)/ ultrafast bootstrap support (%).

(legend continued on next page)

according to species and/or populations, with the Sardinian dhole clustering with the two Asian dholes.

The uncertain placement of the Sardinian dhole on the phylogenetic tree led us to investigate historical gene flow events between this species and the other canids. D-statistics, which use quartets of populations/samples, were applied to identify gene flow between canid lineages. We computed D-statistics using all possible triplets of samples with the Andean fox as the outgroup (Table S4). Considering the past presence of the AHD and dholes' ancestors in Europe during the Pleistocene,[19–23] we investigated the topology in which the AHD is the sister clade, with the Sardinian dhole and Asian dhole in the ingroup; i.e., (((Asian dhole, Sardinian dhole), AHD), Andean fox). Our results (D = −0.068, *Z* score = −19.41, Table S4) suggest a significant excess of allele sharing between the Asian dholes and the AHD compared to the allele sharing between the AHD and the Sardinian dhole (Figures 3A and 3B). Next, we looked for signs of gene flow between the Sardinian dhole and other species of *Canis*. When testing the following three different combinations—(1) (((*Canis*, Golden Jackal), Sardinian dhole), Andean fox); (2) (((*Canis*, Ethiopian wolf), Sardinian dhole), Andean fox); and (3) (((*Canis*, African golden wolf), Sardinian dhole), Andean fox)—the D-statistics suggest higher allele sharing between *Canis lupus* and the Sardinian dhole (Figures S3C–S3G). This result could arise from two scenarios: (1) there was indeed gene flow between the gray wolf and the Sardinian dhole, or (2) the trio of related species—Ethiopian wolf, African golden wolf, and Golden jackal—share ancestry with a species not represented in our study that falls outside the Sardinian dhole in the phylogeny. This scenario is consistent with the findings of a previous study on basal canids, which hypothesized gene flow from an unknown canid into the ancestor of gray wolves and coyotes.[17]

Assuming that the ancestor of *Cynotherium* arrived in Sardinia-Corsica through a terrestrial connection between the islands and mainland Europe, this sets the possible colonization of Sardinia-Corsica at *ca* 5 Ma, during the Messinian salinity crisis,[24] or around 3 Ma—close to the Plio-Pleistocene boundary.[12,25] However, since then, no evidence of a land bridge that would allow movements of species between the two islands and the continent has been found. Nevertheless, Pleistocene mammals of Sardinia are currently divided into two major faunal complexes representing the shift in taxa composition that happened by the end of the Early Pleistocene.[26,27] The older complex (*Nesogoral* Faunal Complex) is characterized by species like *Nesogoral melonii*, *Macaca maiori*, and *Sus sondaari*, while the younger complex (*Tyrrhenicola* Faunal Complex) was represented by *Tyrrenicola henseli* and the cervid *Megaloceros cazioti* [25]. This observed transition was probably the result of a long turn-over process in which only a few new species represent the descendants of pre-existing lineages while the replacement was mainly made up by the over-water colonization by continental species.[25,27] Given this complex scenario, a number of earlier studies considered the Sardinian dhole as a subgenus of *Xenocyon*, *Cuon*, or a derived form of *Canis*[3,10–12] that might have reached Sardinia and Corsica by sweepstake or passive dispersal at the transition between Early and Middle Pleistocene.[4,28] This phenomenon is especially feasible during periods of fluctuation of the sea level and has been known to contribute to the faunal turnover in Sardinia.[4,26,28]

Therefore, to further investigate the demographic history of dholes, we used admixture graphs to test whether the diffuse ancestries can be explained by AHD-like admixture in Asian dhole or *Canis*-like admixture in Sardinian dhole. First, we estimated the proportion of ancestry derived from the ancestor of the Asian dhole into the Sardinian dhole using AHD, dholes, Ethiopian wolf, and Andean fox as the outgroup. We found that the node representing the ancestral population of the Asian and Sardinian dholes derives 60% ancestry from the node that diverged from the ancestral AHD population in the past (Figure 3C). A second admixture event brings 25% of AHD ancestry into only the ancestor of the Asian dholes. Subsequently, when including the Portuguese wolf in the previous graph, we found that it was necessary to model one more admixture event (see Figure S4A) between the ancestors of the Ethiopian wolf and the dholes. The Ethiopian wolf was best modeled as a mixture of 4% from an ancestral population related to the dholes and 96% from the *Canis* lineage. This last admixture event confirmed the results obtained in admixture analysis using four ancestry components (Figure 2E) in which the Ethiopian wolf shares a proportion of its ancestry with the dholes.

Given the geological history of Sardinia and Corsica, along with cycles of long-term isolation and past colonization, we explored the effects of isolation on genetic diversity in the Sardinian dhole using estimates of genome-wide heterozygosity and effective population size. We thus first inferred the heterozygosity in sliding windows for all the representative canid species in our dataset and found that the Sardinian dhole shows remarkably low levels of heterozygosity across the whole genome, comparable to other isolated canids with small population sizes, such as the Ethiopian wolf and the AHD (Figure 4A).

The other two dholes showed a strongly bimodal distribution in their heterozygosity distribution with regions of high and very low heterozygosity, probably as a result of their recent demographic histories and captive breeding. Among the different canid species that went through bottlenecks or population reductions,[29–34] the Sardinian dhole showed decreased genetic diversity across the entire genome, comparable to the AHDs from Zimbabwe, Kenya, and South Africa (Figure 4A) that went through stable and long-term population declines.[35,36] This

(C) DiscoVista relative frequency analysis. Each box title indicates the corresponding branch on the tree in (A) and shows the frequency of three topologies around the focal internal branch of ASTRAL species tree. The first topology represented in red is the main topology followed by the other two alternatives in blue. On the y axis, the relative frequency is indicated and the dashed lines represent the one-third threshold. On the x axis, each quartet topology is shown using the neighboring branch labels.

(D) Principal-component analysis of the Sardinian dhole with canid species from Eurasia, Africa, and North America. The black star represents the individual of *Cynotherium sardous* analyzed in this study.

(E) Admixture ancestry component analysis selected for 46 individuals belonging to eight canid species. Only the run with the best likelihood out of 100 runs was selected to be displayed for each K (number of ancestry components).

See also Figure S2.

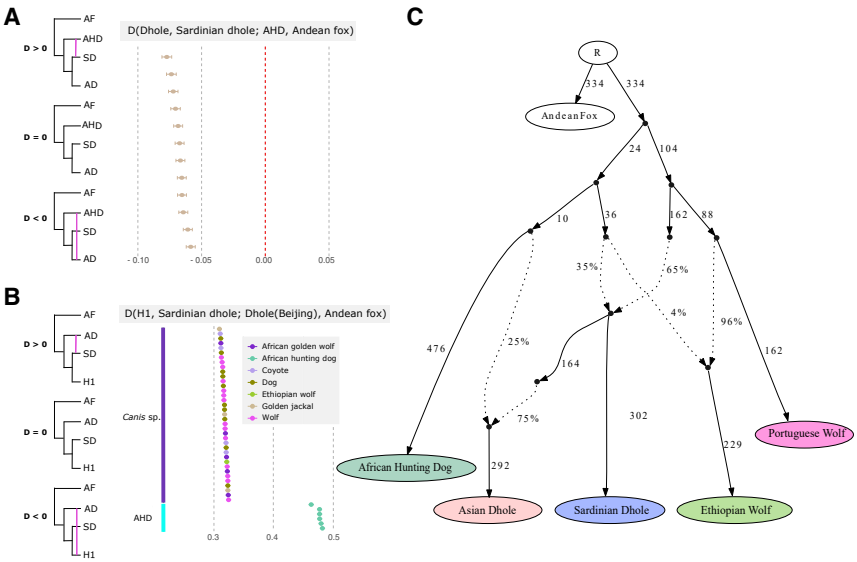

Figure 3. Gene flow and population ancestry model of the Sardinian dhole

(A and B) In these panels, the gene flow among different canids is shown using ABBA-BABA test in ANGSD. On the y axes of (A), different combinations of Asian dholes (AD) and AHD individuals are considered (e.g., from the top to the bottom dhole [Berlin zoo]-AHD [Kenya], dhole [Beijing zoo]-AHD [Kenya], and so on). In (A), there is significant gene flow between the Asian dhole (AD) and the African hunting dog (AHD) showing a higher degree of genetic affinity between these two groups compared to the Sardinian dhole (SD) and AHD. In (B), allele sharing between the AD and SD is higher than when considering AD with any of the other canid lineages in the study. In both panels, the error bars indicate 3 standard errors around the estimates of the D-statistics.

(C) Model of the phylogenetic relationships among canids augmented with admixture events. The qpGraph shown here was estimated considering the pairwise D-statistics. Dotted lines represent admixture events, and the estimated mixture proportion is shown along them (%). Genetic drift (expressed in drift units per 1,000) is shown along solid lines. This admixture graph represents the best fitting graph ($-3 > Z < +3$) to model African-hunting-dog-like ancestry into the Asian and Sardinian dhole and Ethiopian wolf.

See also Figures S3 and S4 and Table S4.

result represents a clear picture of the past population dynamics of the Sardinian dhole and how the long-term isolation has shaped its genome. The low levels of heterozygosity in the Sardinian dhole across all regions of its genome strongly suggest a long period of time with low effective population size, exhibiting population dynamics similar to that of the mountain gorilla.[37]

We then calculated the divergence time between the Sardinian dhole and the contemporary Asian dhole using the statistic F(A|B), which estimates the probability of an individual A (Sardinian dhole) carrying the derived allele at sites that are heterozygous in individual B (Asian dhole). The assumption behind this approach is that when two populations start to diverge, they will also accumulate mutations that—due to isolation—will not be shared with other populations. We estimated that the Sardinian dhole carried the derived allele at ~2% of sites that were heterozygous in the Asian dhole. The proportion of derived alleles was then used as a summary statistic to estimate the divergence time of the Sardinian dhole and the Asian dhole from simulations of multiple divergence times. The simulations were calibrated using the effective population size of Asian dhole estimated using the pairwise sequentially Markovian coalescent (PSMC)[38] model, which estimates effective population sizes from the abundance and density of heterozygous sites in a single diploid individual. The simulations were calibrated using the gray wolf mutation rate ($\mu = 4 \times 10^{-9}$ /bp/generation)[39] and a generation time (g) of 3 years.[39,40] Using F(A|B), we conclude that the Sardinian dhole and the Asian dhole diverged *ca* 885 ka (range 870 and 900 ka) (Figure 4B).

This time frame postdates the last appearance of a land bridge between the island and the continent; however, it is consistent with a previous study attesting the entrance of the ancestors of *Cynotherium* around 0.95 and 0.85 Ma during sea level lowering.[25] It is also possible that other dhole-like lineages entered the island before this time, as attested by a single Pliocene-Pleistocene fossil of an unknown species of *Cynotherium*. However, from the paucity of specimens until the Middle Pleistocene, we hypothesize that the last common ancestor between the Asian and the Sardinian dhole entered the island at the end of the Early Pleistocene (*ca* 885 ka) and since then the two lineages started diverging.

Lastly, we explored the timing of gene flow between the Sardinian/Asian dhole and Sardinian dhole/Eurasian gray wolves and domestic dogs by performing a pseudodiploid demographic analysis using hPSMC.[41] The results suggest that the gene flow between the Sardinian dhole and the ancestor of the modern Asian dhole ended between 510 and 310 ka, while gene flow between the Sardinian dhole and the Eurasian wolf-like ancestor ended between 510 and 360 ka. This shows that gene flow between the Sardinian dhole and the other two lineages might have ended around the same time, between 510 and 310 ka. We found that the demographic histories of the two pseudo-hybrids (Figure 4C) are almost identical, indicating that gene flow ceased after the Middle Pleistocene, a timing that is also consistent with a number of publications in which different faunal complexes that populated the island were taken into consideration.[26,28,42]

Further, we determined that estimates using Sardinian dhole/AHD and Asian dhole/AHD hybrids gave an interval between 1.05 Ma and 300 ka for the cessation of gene flow between AHD and the two dhole lineages (Figures S4B–S4E). Wide intervals, such as those reported here, could be the result of very low levels of post-divergence gene flow between two species.[41] Furthermore, the last fossil of *Lycaon* found in Europe is dated to 0.83 Ma, implying that this genus went locally extinct in Europe toward the end of the Early Pleistocene.[22]

Combining archaeological evidence and hPSMC analysis, we hypothesize that the divergence time between *Lycaon* and the lineage that gave rise to the Sardinian dhole and to the modern form of

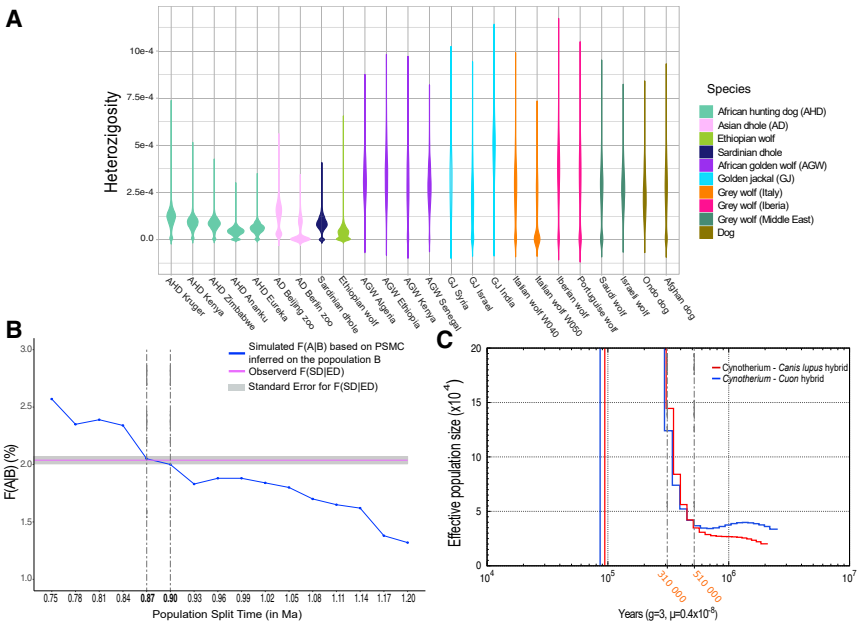

Figure 4. Population dynamics, split time, and end of gene flow

(A) Violin plot representing heterozygosity in sliding windows across the species of canids selected in this study.

(B) Estimate of the population split time between the Sardinian dhole and Asian dhole populations. F(A|B), represented on the y axis, is the probability to observe a derived allele in the population A (SD) when B (ED) is heterozygous at the same site. The numbers on the x axis represent the population split time expressed in Ma (derived from the number of generations assuming a generation time of 3 years).

(C) hPSMC plot based on the artificial hybrid genomes constructed using different species.

See also Figures S4B–S4E.

Cuon in Europe can be narrowed down to between 1.05 Ma and *ca* 0.83 Ma. Admixture events between the ancestors of the AHD and the Asian dhole lineages could have happened in other locations in the past when the two species were not confined to their present-day ranges. For instance, the first fossil of a *Lycaon* dating back to the Middle Pleistocene was found in Israel, suggesting that the ecological boundaries that separate these two species were not in place yet.[43] This could explain the 25% ancestry derived from the AHD lineage observed only in the Asian dhole (Figure 3C), which probably occurred after the Sardinian dhole lineage became isolated in Sardinia and Corsica.

In conclusion, by generating the first whole genome of a Sardinian dhole, we clarified the phylogenomic placement of this species among the genomic diversity of living canids from Eurasia, Africa, and North America. We detected historical gene flow between the Sardinian dhole and Asian dhole lineages, and evidence of past admixture from the ancestor of the AHD into the Asian dhole lineage. We found that the Sardinian dhole lineage diverged from the Asian dholes around *ca* 885 ka, followed by post-divergence gene flow ceasing later, between 560 and 310 ka, meaning that Sardinia could have been colonized several times during the Pleistocene. However, from these results, it can be inferred that probably sometime during the Middle Pleistocene, the physical barrier created by the sea separating Sardinia from the continent effectively ended further canid migrations. We also found that our sample showed reduced genome-wide diversity that—together with the long-term isolation on the island—could have contributed to its extinction. In this complex and delicate situation, the hypothesis that humans played a crucial role in the extinction of the last Pleistocene mammalian fauna cannot be excluded.[44,45] However, the direct or indirect anthropogenic pressure exerted on the Sardinian dhole population could have been a concause—together with the long-term isolation, loss of genetic diversity and reduction of Ne—leading to the extinction of this species. The results presented here also highlight the importance of sequencing ancient genomes from island ecosystems to improve our knowledge of the past evolution, colonization, and migration events of extinct species.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Sample information
- METHOD DETAILS
  - Data generation
  - Radiocarbon dating
  - Dataset
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Quality control and alignment
  - Sex determination
  - Genotype likelihoods
  - Principal components analysis
  - Admixture
  - Heterozygosity in sliding windows
  - Nuclear genome phylogeny
  - Topology weighting analysis
  - Mitochondrial phylogeny
  - Gene flow between the Sardinian dhole, AHD, and Asian dholes
  - Calling of polymorphism and filtering
  - qpGraph
  - Split time analysis - F(A|B)
  - hPSMC

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2021.09.059.

ACKNOWLEDGMENTS

This research was funded by the ERC Consolidator grant 681396, "Extinction Genomics." This work has been performed within a 5-year (2017–2022) scientific agreement between the "Soprintendenza Archeologia, Belle arti e Paesaggio per le province di Sassari, Olbia-Tempio e Nuoro" and the University of Florence Earth Sciences department and is framed within a wider project on Late Neogene vertebrate evolution granted by the University of Florence (Fondi di Ateneo) under the responsibility of L.R. The authors also acknowledge support from Science for Life Laboratory, the Knut and Alice Wallenberg Foundation, the National Genomics Infrastructure funded by the Swedish Research Council, and Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure. We would like to thank Ashot Margaryan, Davide Palumbo, Elisabetta Cilli, and Romolo Caniglia for useful intellectual discussions about this research and the two anonymous reviewers for their valuable suggestions on earlier versions of the manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization, M.M.C., M.-H.S.S., S.G., and M.T.P.G.; investigation, M.M.C. and S.G.; formal analysis, M.M.C. and J.K.J.; data curation, M.M.C. and J.K.J.; resources, C.T., E.C., E.R., L.D., L.R., G.Z., C.G., M.A., M.T.P.G., O.S., T.S.-P., S.B.L., and S.G.; writing – original draft, M.M.C.; writing –review & editing, M.M.C., M.-H.S.S., S.B.L., E.R., L.R., C.T., M.A., E.C., M.G., E.R., L.D., M.T.P.G., and S.G.; visualization, M.M.C.; supervision, M.-H.S.S. and S.G.; funding acquisition, M.T.P.G.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 26, 2021
Revised: June 23, 2021
Accepted: September 22, 2021
Published: October 15, 2021

REFERENCES

1. Malatesta, A. (1970). Cynotherium sardous Studiati an extinct canid from the Pleistocene of Sardinia. Mem. dell'Istituti Ital. di Paleontol. Umana, NS *1*, 1–72.
2. Vigne, J.-D., Bailon, S., and Cuisin, J. (1997). Biostratigraphy of Amphibians, Reptiles, Birds and Mammals in Corsica and the role of Man in the Holocene faunal turnover. Anthropozoologica *25-26*, 587–604.
3. Lyras, G.A., van der Geer, A.A.E., Dermitzakis, M.D., and De Vos, J. (2006). Cynotherium sardous, an insular canid (Mammalia: Carnivora) from the Pleistocene of Sardinia (Italy), and its origin. J. Vertebr. Paleontol. *26*, 735–745.
4. Palombo, M.R. (2018). Insular mammalian fauna dynamics and paleogeography: A lesson from the Western Mediterranean islands. Integr. Zool. *13*, 2–20.
5. Palombo, M.R. (2006). Biochronology of the Plio-Pleistocene terrestrial mammals of Sardinia: the state of the art. Hellenic Journal of Geosciences *41*, 47–66.
6. Eisenmann, V., and van der Geer, B. (1999). The Cynotherium from Corbeddu (Sardinia): comparative biometry with extant and fossil canids. Deinsea *7*, 147–168.
7. Abbazzi, L., Angelone, C., Arca, M., Barisone, G., Bedetti, C., Delfino, M., Kotsakis, T., Marcolini, F., Palombo, M.R., Pavia, M., et al. (2004). Plio-Pleistocene fossil vertebrates of Monte Tuttavista (Orosei, Eastern Sardinia, Italy), an overview. Riv. Ital. Paleontol. Stratigr. *110*, 681–706.
8. Lyras, G., and van der Geer, A. (2006). Adaptations of the Pleistocene island canid Cynotherium sardous (Sardinia, Italy) for hunting small prey. Cranium *23*, 51–60.
9. Madurell-Malapeira, J., Palombo, M.R., and Sotnikova, M. (2015). Cynotherium malatestai, sp. Nov. (Carnivora, Canidae) from the early middle Pleistocene deposits of Grotta dei Fiori (Sardinia, Western Mediterranean). J. Vert. Paleontol. *35*, e943400.
10. Malatesta, A. (1962). Il cane selvaggio del Pleistocene di Sardegna.
11. Bonifay, M.-F. (1971). Carnivores quaternaires du Sud-Est de la France (Editions du Muséum).
12. Abbazzi, L., Arca, M., Tuveri, C., and Rook, L. (2005). The endemic canid Cynotherium (Mammalia, Carnivora) from the Pleistocene deposits of Monte Tuttavista (Nuoro, Eastern Sardinia). Riv. Ital. Paleontol. Stratigr. *111*, 497–511.
13. 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.
14. Zhang, C., Rabiee, M., Sayyari, E., and Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics *19* (*Suppl 6*), 153.
15. 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.
16. Martin, S.H., and Van Belleghem, S.M. (2017). Exploring Evolutionary Relationships Across the Genome Using Topology Weighting. Genetics *206*, 429–438.
17. Gopalakrishnan, S., Sinding, M.S., Ramos-Madrigal, J., Niemann, J., Samaniego Castruita, J.A., Vieira, F.G., Carøe, C., de Manuel Montero, M., Kuderna, L., Serres, A., et al. (2019). Interspecific Gene Flow Shaped the Evolution of the Genus Canis. Curr. Biol. *29*, 4152.
18. Chavez, D.E., Gronau, I., Hains, T., Kliver, S., Koepfli, K.-P., and Wayne, R.K. (2019). Comparative genomics provides new insights into the remarkable adaptations of the African wild dog (Lycaon pictus). Sci. Rep. *9*, 8329.
19. Kurtén, B. (1968). Pleistocene Mammals of Europe (AldineTransaction).
20. 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.
21. Petrucci, M., Romiti, S., and Sardella, R. (2012). The Middle-Late Pleistocene Cuon Hodgson, 1838 (Carnivora, Canidae) from Italy. Boll. Soc. Paleontol. Ital. *51*, 138.
22. Madurell-Malapeira, J., Rook, L., Martínez-Navarro, B., Alba, D.M., Aurell-Garrido, J., and Moyà-solà, S. (2013). The latest European painted dog. J. Vertebr. Paleontol. *33*, 1244–1249.
23. Taron, U.H., Paijmans, J.L.A., Barlow, A., Preick, M., Iyengar, A., Drăguşin, V., Vasile, ., Marciszak, A., Roblíčková, M., and Hofreiter, M. (2021). Ancient DNA from the Asiatic Wild Dog (*Cuon alpinus*) from Europe. Genes (Basel) *12*, https://doi.org/10.3390/genes12020144.
24. Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J., and Wilson, D.S. (1999). Chronology, causes and progression of the Messinian salinity crisis. Nature *400*, 652–655.
25. Palombo, M.R., and Rozzi, R. (2014). How correct is any chronological ordering of the Quaternary Sardinian mammalian assemblages? Quat. Int. *328-329*, 136–155.
26. Sondaar, P.Y., Sanges, M., Kotsakis, T., and de Boer, P.L. (1986). The Pleistocene deer hunter of Sardinia. Geobios Mem. Spec. *19*, 17–31.
27. Palombo, M.R. (2009). Biochronology, paleobiogeography and faunal turnover in western Mediterranean Cenozoic mammals. Integr. Zool. *4*, 367–386.
28. Melis, R.T., Palombo, M.R., Ghaleb, B., and Meloni, S. (2016). A key site for inferring the timing of dispersal of giant deer in Sardinia, the Su Fossu de Cannas cave, Sadali, Italy. Quat. Res. *86*, 335–347.

29. Roy, M.S., Girman, D.J., Taylor, A.C., and Wayne, R.K. (1994). The use of museum specimens to reconstruct the genetic variability and relationships of extinct populations. Experientia *50*, 551–557.

30. 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.

31. Woodroffe, R., and Ginsberg, J.R. (1999). Conserving the African wild dog Lycaon pictus. II. Is there a role for reintroduction? Oryx *33*, 143–151.

32. Marsden, C.D., Woodroffe, R., Mills, M.G.L., McNutt, J.W., Creel, S., Groom, R., Emmanuel, M., Cleaveland, S., Kat, P., Rasmussen, G.S.A., et al. (2012). Spatial and temporal patterns of neutral and adaptive genetic variation in the endangered African wild dog (Lycaon pictus). Mol. Ecol. *21*, 1379–1393.

33. Zimen, E., and Boitani, L. (1975). Number and distribution of wolves in Italy. Z. Saugetierkd. *40*, 102–112.

34. Kamler, J., Songsasen, N., Jenks, K., Srivathsa, A., Sheng, L., and Kunkel, K.E. (2015). IUCN Red List of Threatened Species: Cuon alpinus (IUCN Red List of Threatened Species). http://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T5953A72477893.en.

35. 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.

36. Armstrong, E.E., Taylor, R.W., Prost, S., Blinston, P., van der Meer, E., Madzikanda, H., Mufute, O., Mandisodza-Chikerema, R., Stuelpnagel, J., Sillero-Zubiri, C., and Petrov, D. (2019). Cost-effective assembly of the African wild dog (Lycaon pictus) genome using linked reads. Gigascience *8*, https://doi.org/10.1093/gigascience/giy124.

37. van der Valk, T., Díez-Del-Molino, D., Marques-Bonet, T., Guschanski, K., and Dalén, L. (2019). Historical Genomes Reveal the Genomic Consequences of Recent Population Decline in Eastern Gorillas. Curr. Biol. *29*, 165–170.e6.

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

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

40. 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.

41. Cahill, J.A., Soares, A.E.R., Green, R.E., and Shapiro, B. (2016). Inferring species divergence times using pairwise sequential Markovian coalescent modelling and low-coverage genomic data. Philos. Trans. R. Soc. Lond. B Biol. Sci. *371*, https://doi.org/10.1098/rstb.2015.0138.

42. Van der Made, J. (1999). Biogeography and stratigraphy of the Mio-Pleistocene mammals of Sardinia and the description of some fossils. Deinsea *7*, 337–360.

43. Stiner, M.C., Howell, F.C., Martínez-Navarro, B., Tchernov, E., and Bar-Yosef, O. (2001). Outside Africa: Middle Pleistocene Lycaon from Hayonim Cave, Israel. Boll. Soc. Paleontol. Ital. *40*, 293–302.

44. Schule, W. (1993). Mammals, Vegetation and the Initial Human Settlement of the Mediterranean Islands: A Palaeoecological Approach. J. Biogeogr. *20*, 399–411.

45. Gippoliti, S., and Amori, G. (2006). Ancient introductions of mammals in the Mediterranean Basin and their implications for conservation. Mammal Rev. *36*, 37–48.

46. 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.

47. Chen, L., Zhang, H., Zhang, J., Zhao, C., and Sha, W. (2015). The complete mitochondrial genome of Cuon alpinus lepturus. Mitochondrial DNA *26*, 767–768.

48. Arnason, U., Gullberg, A., Janke, A., and Kullberg, M. (2007). Mitogenomic analyses of caniform relationships. Mol. Phylogenet. Evol. *45*, 863–874.

49. 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.

50. Schubert, M., Lindgreen, S., and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res. Notes *9*, 88.

51. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

52. Toolkit, P. (2018). Picard Toolkit (Broad Institute).

53. Poplin, R., Ruano-Rubio, V., DePristo, M., Fennell, T., Carneiro, M., Van der Auwera, G., Kling, D., Gauthier, L., Levy-Moonshine, A., Roazen, D., et al. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv. https://doi.org/10.1101/201178.

54. 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.

55. 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.

56. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics *15*, 356.

57. 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.

58. 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.

59. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics *29*, 1682–1684.

60. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.

61. Chang, C.C., Chow, C.C., Tellier, L.C.A.M., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience *4*, 7.

62. Meisner, J., and Albrechtsen, A. (2018). Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. Genetics *210*, 719–731.

63. Skotte, L., Korneliussen, T.S., and Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. Genetics *195*, 693–702.

64. RStudio Team (2021). RStudio: Integrated Development Environment for R (Boston, MA: RStudio, PBC). http://www.rstudio.com/.

65. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

66. Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. *47* (W1), W256–W259.

67. Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics *35*, 4453–4455.

68. 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.

69. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

70. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol. Biol. Evol. *37*, 1530–1534.

71. Sayyari, E., Whitfield, J.B., and Mirarab, S. (2018). DiscoVista: Interpretable visualizations of gene tree discordance. Mol. Phylogenet. Evol. *122*, 110–115.

72. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al.; 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

73. Hellenthal, G., and Stephens, M. (2007). msHOT: modifying Hudson's ms simulator to incorporate crossover and gene conversion hotspots. Bioinformatics *23*, 520–521.

74. Behr, A.A., Liu, K.Z., Liu-Fang, G., Nakka, P., and Ramachandran, S. (2016). pong: fast analysis and visualization of latent clusters in population genetic data. Bioinformatics *32*, 2817–2823.

75. Hudson, R.R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics *18*, 337–338.

76. Martini, F. (1992). In Il più antico popolamento umano delle isole: la Sardegna. I primi abitanti della Valle Padana: Monte Poggiolo, J. Book, ed. (Milan: Italie), pp. 175–187.

77. Sondar, P., Elburg, R., Hofmeijer, G.K., Spaan, A., De Visser, H., Sanges, M., et al. (1995). The Human Colonization of Sardinia: a Late-Pleistocene human fossil from Corbeddu cave. Comptes rendus de l'Académie des sciences*320* (Earth & planetary sciences), pp. 145–150.

78. Spoor, F. (1999). The human fossils from Corbeddu Cave, Sardinia: a reappraisal. Deinsea *7*, 297–302.

79. Palombo, M.R., Antonioli, F., Presti, V.L., Mannino, M.A., Melis, R.T., Orru, P., Stocchi, P., Talamo, S., Quarta, G., Calcagnile, L., et al. (2017). The late Pleistocene to Holocene palaeogeographic evolution of the Porto Conte area: Clues for a better understanding of human colonization of Sardinia and faunal dynamics during the last 30 ka. Quat. Int. *439*, 117–140.

80. Hofmeijer, G.K. (1997). Late Pleistocene Deer Fossils from Corbeddu cave: implications for human Colonization of the Island of Sardinia (British Archaeological Reports Ltd).

81. Boessenkool, S., Hanghøj, K., Nistelberger, H.M., Der Sarkissian, C., Gondek, A.T., Orlando, L., Barrett, J.H., and Star, B. (2017). Combining bleach and mild predigestion improves ancient DNA recovery from bones. Mol. Ecol. Resour. *17*, 742–751.

82. 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.

83. 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. (2018). Single-tube library preparation for degraded DNA. Methods Ecol. Evol. *9*, 410–419.

84. Mak, S.S.T., Gopalakrishnan, S., Carøe, C., Geng, C., Liu, S., Sinding, M.S., Kuderna, L.F.K., Zhang, W., Fu, S., Vieira, F.G., et al. (2017). Comparative performance of the BGISEQ-500 vs Illumina HiSeq2500 sequencing platforms for palaeogenomic sequencing. Gigascience *6*, 1–13.

85. Reimer, P.J., Austin, W.E.N., Bard, E., Bayliss, A., Blackwell, P.G., Ramsey, C.B., Butzin, M., Cheng, H., Lawrence Edwards, R., Friedrich, M., et al. (2020). The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). Radiocarbon *62*, 725–757.

86. Ramsey, C.B. (2009). Bayesian Analysis of Radiocarbon Dates. Radiocarbon *51*, 337–360.

87. 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.

88. 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.

89. 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*, e1004466.

90. 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.

91. Liu, Y.-H., Wang, L., Xu, T., Guo, X., Li, Y., Yin, T.-T., Yang, H.-C., Hu, Y., Adeola, A.C., Sanke, O.J., et al. (2018). Whole-Genome Sequencing of African Dogs Provides Insights into Adaptations against Tropical Parasites. Mol. Biol. Evol. *35*, 287–298.

92. 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.

93. 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.

94. Sinding, M.S., Gopalakrishan, S., Vieira, F.G., Samaniego Castruita, J.A., Raundrup, K., Heide Jørgensen, M.P., Meldgaard, M., Petersen, B., Sicheritz-Ponten, T., Mikkelsen, J.B., et al. (2018). Population genomics of grey wolves and wolf-like canids in North America. PLoS Genet. *14*, e1007745.

95. 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.

96. Perri, A.R., Mitchell, K.J., Mouton, A., Álvarez-Carretero, S., Hulme-Beaman, A., Haile, J., Jamieson, A., Meachen, J., Lin, A.T., Schubert, B.W., et al. (2021). Dire wolves were the last of an ancient New World canid lineage. Nature *591*, 87–91.

97. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (New York: Springer-Verlag).

98. Wickham, H., François, R., Henry, L., and Müller, K. (2018) (dplyr: A Grammar of Data Manipulation. Version 0.7.6).

99. Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics *20*, 289–290.

100. R Core Team (2020). R: A language and environment for statistical computing. (Vienna, Austria: R Foundation for Statistical Computing) https://www.R-project.org/.

101. Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., and Vinh, L.S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol. Biol. Evol. *35*, 518–522.

102. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermiin, L.S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods *14*, 587–589.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
1 canid bone paleontological remains	This paper	Sardinian dhole (CB83-D1101)
Chemicals, peptides, and recombinant proteins		
Proteinase K	Sigma-Aldrich	Cat#3115844001
Critical commercial assays		
MinElute PCR Purification Kit	QIAGEN	Cat#28006
PfuTurbo Cx Hotstart DNA Polymerase	Agilent	Cat#600414
Phusion® High-Fidelity PCR Master Mix with HF buffer	New England Biolabs Inc.	Cat#M0531S
T4 DNA ligase	New England Biolabs Inc.	Cat#M0202L
T4 Polynucleotide Kinase	New England Biolabs Inc.	Cat#M0201L
T4 DNA Polymerase	New England Biolabs Inc.	Cat#M0203S
BSt 2,0 warmstart polymerase	New England Biolabs Inc.	Cat#M0538S
Deposited data		
Sardinian dhole (CB83-D1101) sequencing data	This study	Table S3
Sequencing data for 1 Dire wolf		Table S3
Sequencing data for 2 Pleistocene dholes (Y-38, Y-39)		Table S3
Sequencing data for 10 dogs		Table S3
Sequencing data for 2 North American wolves		Table S3
Sequencing data for 1 Altai wolf		Table S3
Sequencing data for 5 European wolves		Table S3
Sequencing data for 4 Middle eastern wolves		Table S3
Sequencing data for 2 Asian wolves		Table S3
Sequencing data for 4 coyotes		Table S3
Sequencing data for 3 golden jackals		Table S3
Sequencing data for 4 African golden wolves		Table S3
Sequencing data for 1 African golden wolf x Grey wolf hybrid		Table S3
Sequencing data for 1 Ethiopian wolf		Table S3
Sequencing data for 2 Asian dholes		Table S3
Sequencing data for 6 African hunting dogs		Table S3
Andean fox sequencing data		Table S3
Grey wolf reference genome	Gopalakrishnan et al.[46]	https://sid.erda.dk/wsgi-bin/ls.py?share_id=f1ppDgUPQG
Dog reference genome	Lindblad-Toh et al.[13]	https://www.ncbi.nlm.nih.gov/assembly/GCF_000002285.3/
Asian dhole reference mitochondrial genome	Chen et al.[47]	GenBank: GU063864.1, https://www.ncbi.nlm.nih.gov/nuccore/GU063864.1/
Grey wolf reference mitochondrial genome	Arnason et al.[48]	N/A
Oligonucleotides		
Illumina-compatible adapters		N/A
Software and algorithms		
PALEOMIX v1.2.13.1	Schubert et al.[49]	RRID:SCR_015057, https://github.com/MikkelSchubert/paleomix
AdapterRemoval2	Schubert et al.[50]	RRID:SCR_011834, https://github.com/MikkelSchubert/adapterremoval

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
bwa-*backtrack* v0.7.12	Li and Durbin[51]	RRID:SCR_010910, http://bio-bwa.sourceforge.net/
Picard v2.9.1	Picard Toolkit[52]	RRID:SCR_006525, https://broadinstitute.github.io/picard
GATK v3.4 -Haplotype Caller	Poplin et al.[53]	RRID:SCR_001876, https://software.broadinstitute.org/gatk/
GATK v4.1	DePristo et al.,[54] McKenna et al.[55]	RRID:SCR_001876, https://software.broadinstitute.org/gatk/
ANGSD v0.933	Korneliussen et al.[56]	https://github.com/ANGSD/angsd
samtools v1.9	Li et al.[57]	RRID:SCR_002105, http://samtools.sourceforge.net/
ADMIXTOOLS	Patterson et al.[58]	RRID:SCR_018495, https://github.com/DReichLab/AdmixTools
mapDamage 2.0	Jónsson et al.[59]	RRID:SCR_001240, https://github.com/ginolhac/mapDamage
Plink 1.9	Purcell et al.,[60] Chang et al.[61]	https://www.cog-genomics.org/plink/2.0/
PCAngsd v0.98	Meisner and Albrechtsen[62]	https://github.com/Rosemeis/pcangsd
NGSadmix v32	Skotte et al.[63]	RRID:SCR_003208, http://www.popgen.dk/software/index.php/NgsAdmix
RStudio v.1.2.5	RStudio Team[64]	RRID:SCR_000432, https://www.rstudio.com
BEDTools	Quinlan and Hall[65]	RRID:SCR_006646, https://github.com/arq5x/bedtools2
Tree Of Life (iTOL) v4	Letunic and Bork[66]	RRID:SCR_018174, https://itol.embl.de/
RAxML-ng	Kozlov et al.[67]	https://github.com/amkozlov/raxml-ng
ASTRAL-III	Zhang et al.[14]	https://github.com/smirarab/ASTRAL
Twisst	Martin and Van Belleghem[16]	https://github.com/simonhmartin/twisst
MtArchitect	Lobon et al.[68]	https://www.upf.edu/web/comparative-genomics/mt-architect
MUSCLE	Edgar[69]	RRID:SCR_011812, https://www.ebi.ac.uk/Tools/msa/muscle/
IQ-TREE2 v2.1.2	Minh et al.[70]	http://www.iqtree.org/
DiscoVista	Sayyari et al.[71]	https://github.com/esayyari/DiscoVista
VCFtools v0.1.8	Danecek et al.[72]	RRID:SCR_001235, https://vcftools.github.io/index.html
PSMC	Li and Durbin[38]	https://github.com/lh3/psmc
msHOT	Hellenthal and Stephens[73]	https://bio.tools/mshot https://uchicago.app.box.com/s/l3e5uf13tikfjm7e1il1eujitlsjdx13
hPSMC	Cahill et al.[41]	https://github.com/jacahill/hPSMC
Pong	Behr et al.[74]	https://github.com/ramachandran-lab/pong
Ms	Hudson[75]	http://home.uchicago.edu/~rhudson1/source/mksamples.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for reagents and data may be directed to and will be fulfilled by the Lead Contact, Shyam Gopalakrishnan (shyam.gopalakrishnan@sund.ku.dk)

Materials availability

This study did not generate new unique reagents.

Data and code availability

Raw sequencing reads from whole genome sequencing of the Sardinian dhole have been deposited at the European Nucleotide Archive (ENA; study accession number PRJEB46831).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Sample information

A *Cynotherium sardous* petrous bone (CB83-D1101) from Corbeddu cave was selected at the collection of the National Archaeological Museum of Nuoro. Corbeddu cave is an archeo-paleontological site located in the Lanaittu Valley (Oliena, Nuoro Province, Northeastern Sardinia, 40.254921°, 9.485078°) famous for its lithic and osteological evidence of *H. sapiens*,[76,77] disputedly referred to the Late Pleistocene or Early Holocene.[78,79] *Cynotherium sardous* remains, among which a nearly-complete skeleton, come mainly from the upper level of layer 3 of hall II.[3,80] Such specimens represent the last occurrence of the species, recently re-dated by Palombo and colleagues[79] to 12,945 ± 75 years.

METHOD DETAILS

Data generation

The sample was processed under strict clean laboratory conditions at the GLOBE Institute, University of Copenhagen. Small petrous bone chunks of around 350 mg, were divided into 3 eppendorf tubes - A, B and C (*ca* 120 mg each) - and washed with diluted bleach, ethanol and ddH2O, following Boessenkool et al.[81] The subsampled material was extracted following Dabney et al. (2013), purified using modified PB buffer[82] and eluted using 2 washes in 22 μl buffer EB (EB) - with 10 min of incubation time at 37°C. The concentration of each extract was checked on a Qubit (ng/μl) and on an Agilent 2100 Bioanalyzer High-Sensitivity DNA chip (Agilent Technologies) for molar concentration and fragment size. Sequence data was generated using both BGISeq and Illumina sequencing platforms. Specifically, five Illumina libraries were built using 16 μl of DNA from each extract (A, B, C) in a final reaction volume of 40 μl following Carøe et al.[83] and amplified using PfuTurbo Cx HotStart DNA Polymerase (Agilent Technologies) and Phusion® High-Fidelity PCR Master Mix with HF buffer (New England Biolabs Inc) (see Table S1 for further details). From the extract C two aliquots of 16 μl were used to build BGIseq compatible libraries following Mak et al.[84] and Carøe et al.[83] and amplified with Phusion® High-Fidelity PCR Master Mix with HF buffer (New England Biolabs Inc). The appropriate number of cycles were determined using Mx3005 qPCR (Agilent Technologies) in which 1 μl of SYBRgreen fluorescent dye (Invitrogen, Carlsbad, CA, USA) was loaded in 20 μl indexing reaction volume using also 1 μl of template, 0.2 mM dNTPs (Invitrogen), 0.04 U/μl AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA), 2.5 mM MgCl2 (Applied Biosystems), 1X GeneAmp® 10X PCR Buffer II (Applied Biosystems), 0.2 μM forward and reverse primers mixture,[84] and 16.68 μL AccuGene molecular biology water (Lonza). qPCR cycling conditions were 95°C for 10 min, followed by 40 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 45 s. Library index amplifications were performed in 50 μL PCR reactions that contained 14 μL of purified library, 0.1 μM of each forward (BGI 2.0) and custom made reverse primers,[84] 2x Phusion® High-Fidelity PCR Master Mix with HF buffer and 8.6 μL AccuGene molecular biology water (Lonza, Basel, CH). PCR cycling conditions were: initial denaturation at 98°C for 45 s followed by 18 to 20 cycles of 98°C for 20 s, 60°C for 30 s, and 72°C for 20 s, and a final elongation step at 72°C for 5 min. Amplified libraries were then purified using 1.5x ratio of SPRI beads to remove adaptor dimers and eluted in 50 μL of EB (QIAGEN) after an incubation for 10 min at 37°C. Two indexed BGI libraries with the same index combination were pooled together before the beads purification and sequenced on 3 lanes BGIseq-500 using 100 bp single end sequencing reactions. Five Illumina libraries with different combinations of indexes were pooled together in equimolar concentration and sequenced on using 150 bp paired end chemistry on 5 lanes of the Hiseq X platform at SciLifeLab Data Centre, Sweden.

Radiocarbon dating

Fragments of the petrous bone of the Sardinian dhole were submitted for AMS ^{14}C dating at the Radiocarbon Dating Laboratory at the Department of Geology, Lund University, Sweden. The obtained radiocarbon age, 17,480 ± 120 BP (Lab code LuS-15597), was calibrated based on the INTCAL20 calibration dataset[85] using the OxCal v4.4.2 software.[86]

Dataset

The dataset used in this study is represented by 50 representative canid genomes (Table S3), 49 of which were previously published.[17,23,35,36,40,87–96] The genomes considered for this study were chosen to represent the genetic diversity of 9 different species (with the domestic dog considered as a different species from the gray wolf, Table S3) from Africa, Eurasia, and North-America.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quality control and alignment

Short reads obtained from BGI and Illumina sequencing platforms were processed using the PALEOMIX v1.2.13[49] pipeline. The same pipeline was run for each sample in the study. In the first step, the adapters were trimmed with AdapterRemoval2[50] with default settings and the alignments were performed against the gray wolf reference genome[46] and the domestic dog reference genome

(CanFam3.1)[13] using BWA v0.7.12 *backtrack* algorithm[51] with minimum base mapping quality set to 0 to ensure that all the reads were retained in this process. Mapping quality and base quality filters were applied in the later steps of the analysis. PCR duplicates were filtered out using Picard *MarkDuplicates* v2.9.1[52] and in the last step GATK v4.1.0.0[54,55] was used to perform the indel realignment step with no external indel database. Post-mortem DNA damage profiles and nucleotide misincorporation patterns (sub C>T and G>A) were computed using mapDamage2.0[59] (Figure S1) .

Sex determination

No Y chromosome reference sequence is available for CanFam3.1. Therefore, to infer the biological sex of the Sardinian dhole we used the statistics generated from mapping to the domestic dog reference genome. RStudio v1.2.5003[64] and ggplot2[97] were used to visualize the depth of coverage for all the chromosomes of the Sardinian dhole. We then identified as female the individual if the depth of coverage for the X chromosome was similar or exceeded the mean coverage of all the autosomal chromosomes.

Genotype likelihoods

The samples analyzed in the present study span from 4.4X to 28.2X coverage, including in this range also the ancient Sardinian dhole. In order to avoid biases, when possible, we used genotype likelihood over genotype calling. SAMtools model (-GL 1) in ANGSD[56] was used to estimate the genotype likelihood at variant sites for all the scaffolds above 1 MB. Bases with base quality lower than 20 and reads with mapping quality lower than 20 were discarded (-minQ 20 -minmapq 20). We retained sites with the default minimum depth for a minimum of 95% of the individuals in our dataset (-minInd 44) and the following parameters: -remove_bads 1 -baq 1 -C 50 -uniqueOnly 1. The resulting output, generated in beagle format, was then used for the PCA and Admixture analysis.

Principal components analysis

To explore the genetic affinities in our data we performed the principal component analysis (PCA) using PCAngsd v0.98[62] on the 46 individuals' genotype likelihood panels obtained using ANGSD. A covariance matrix was created and we then used RStudio to calculate the eigenvectors and eigenvalues on the covariance matrix file and used ggplot2 to plot the PCA.

Admixture

The genotype likelihood file generated with ANGSD was used as input for NGSadmix v32[63] to generate the ancestry cluster and proportion of admixture for 46 samples in the dataset using 3008607 SNPs. The outgroup, the Andean fox, was excluded from this analysis and we retained sites with the default minimum depth for a minimum of 95% of the individuals in our dataset (-minInd 44). The structure in the dataset was computed using 2 to 15 clusters (K) and for every cluster the analysis was repeated 100 times to ensure convergence to the global maximum. For each K the replicates with the best likelihood scores were chosen and used as input file list (option -i) in pong[74] to visualize the ancestry clusters. The options -n and -l were used to assign sample order and a color for each cluster, respectively.

Heterozygosity in sliding windows

The heterozygosity per sample was estimated using ANGSD, by calculating the per sample folded site frequency spectrum (SFS). We generated a saf.idx file based on individual genotype likelihoods using GATK (-GL 2) from the scaffolds larger than 1Mb (704 scaffolds) for each bam file (doSaf 1 -fold 1) and we excluded transitions (rmtrans 1) and reads with quality score and bases with mapping quality lower than 20 (-minQ 20 -minmapq 20). Since we chose the option -fold 1 to estimate the folded SFS, the gray wolf reference genome was used both as reference and as ancestral (- ref and -anc options). The repeat regions were masked using a repeat mask of the gray wolf reference genome.[46]

The scaffolds that were longer than 1 Mb were partitioned into overlapping windows of size 1 Mb with a step size of 500 kb using the BEDTools[65] *windows* tool. Windows shorter than 1Mb at the end of the scaffolds were discarded. The SFS for each window was estimated using the realSFS utility tool provided in ANGSD and subsequently the ratio of heterozygous sites/total sites was calculated to provide the final heterozygosity per window. RStudio and ggplot2 and dplyr[98] were used to visualize the heterozygosity level at each window and to create a violin plot for a subset of samples.

Nuclear genome phylogeny

All the individuals in the dataset, including the outgroup (Andean fox) were used to construct the nuclear genome phylogeny. First, ANGSD was used to generate a consensus sequence for each genome in our dataset using the gray wolf genome as reference. Each base was sampled based on the consensus base (-dofasta 2). Bases with base quality lower than 20 and reads with mapping quality lower than 20 were discarded (-minQ 20 -minmapq 20). The minimum coverage for each individual was set to 3x (-setminDepthInd 3) and the following additional filters were used: -doCounts 1 -remove_bads 1 -uniqueOnly 1 -baq 1 -C 50. We then selected 1000 random regions, each 5000bp long, from the gray wolf reference genome using BEDTools *random* with the following parameters: -l 5000 -n 1000.

Samtools[57] was used to generate a fasta file for each region using the consensus sequence generated by ANGSD. For each region, the consensus sequences across all samples were combined into a single multi-sequence fasta file. Subsequently, for each region, RAxML-ng[67] was used to reconstruct the phylogeny using the evolutionary model GTR+G. The gene trees for the 1000 regions were concatenated together and a species tree was estimated using Astral-III[14] with the default parameters, retaining all the branches in the gene trees, i.e., the different individuals of the same species were not collapsed into a single group. Interactive Tree Of Life (iTOL)

v4[66] online tool was used to visualize the species tree estimated by Astral-III. Further, we used DiscoVista[71] to visualize the discordance between the 1000 gene trees and the species tree generated with ASTRAL-III. For this step, the samples belonging to the same species were then collapsed together using an annotation file (option -a) and the Andean fox was specified as outgroup to root the tree by using the option -g. The option -o was used to create an output folder with the resulting DiscoVista tree and the plot showing the relative frequency analysis which were used to evaluate the three topologies for each internal branch of the tree.

Topology weighting analysis

To better understand the relative frequencies of the different topologies, particularly the topologies with the different placement of the Sardinian dhole with respect to the Asian dhole and the gray wolf, we performed a Twisst[16] analysis. For this analysis, we only used the 704 largest scaffolds of the gray wolf reference genome, which were larger than 1 Mb. Further, since we were interested only in the phylogenetic placement of the Sardinian dhoe, we restricted our analysis to only 5 groups of samples, viz. the Andean fox as the outgroup, the Sardinian dhole, the Asian dholes, the African hunting dogs, and Eurasian gray wolves.

For the analysis, we split the 704 scaffolds into 500 kb windows using BEDTools[65] (makewindows sub-command), discarding windows that were smaller than 400 kb. For each of these windows, we estimated the pairwise distance for each pair of samples using ANGSD. The pairwise distance matrix was computed using identity by state distances estimated by randomly sampling a single base at each position in the window (-doIBS 1 -doCounts 1 -makematrix 1). Subsequently, we used the pairwise distance matrix to obtain a neighbor joining tree using the library *ape*[99] in RStudio,[100] resulting in 2868 trees across the 704 scaffolds. Finally, we used the 2868 trees to perform the topology weighting analysis in Twisst.

Mitochondrial phylogeny

MtArchitect[68] was used to reconstruct *de novo* sequences of the mitochondrial genomes for 46 modern canids in the dataset. The raw files of the ancient Dire wolf (RW001),[96] Eurasian dholes (Y-38 and Y39)[23] and the Sardinian dhole were instead mapped using PALEOMIX to the Chinese dhole (*Cuon alpinus lepturus*) mitochondrial genome (GU063864.1)[47] and to the gray wolf mitochondrial genome (AM711902.1).[48] For the Sardinian dhole and Dire wolf, the highest coverage was obtained when mapping to the gray wolf, while for the two Pleistocene dholes the highest coverage was obtained when mapping to the Chinese dhole as reference. Thus, we used the alignments generated by mapping to the gray wolf for the Sardinian dhole and the Dire wolf, whereas for the two Pleistocene dholes, we used the alignments obtained by mapping to the Chinese dhole. The consensus sequence was called for each sample based on the highest number of reads mapped to the reference, following Taron and colleagues.[23] We used ANGSD to generate a consensus sequence for the four ancient mitogenomes and each base was sampled based on the consensus base (-dofasta 2). Bases with base quality lower than 20 and reads with mapping quality lower than 20 were discarded (-minQ 20 -minmapq 20) and the following additional filters were used: -doCounts 1 -remove_bads 1 -uniqueOnly 1. The fasta files were merged and aligned using MUSCLE.[69] The alignment was used as input in IQ-TREE2 (v.2.1.2)[70] to reconstruct the mitochondrial phylogeny using 1000 bootstrap replicates (-B 1000) with UFBoot2,[101] 1000 bootstrap replicates for SH-aLRT (-alrt 1000) and *ModelFinder Plus*[102](-m MFP) to search the best model and perform the remaining analysis using the selected model, in this case TPM2+F+R3.

Gene flow between the Sardinian dhole, AHD, and Asian dholes

D-statistics (ABBA-BABA)

The program ANGSD was used to investigate the presence and extent of gene flow between the Sardinian dhole and the other canids included in this study. We restricted our analyses to only the scaffolds over 1 MB. To compute the D-statistics using ANGSD, sites with base quality and mapping quality lower than 20 (-minQ 20 -minMapQ 20) were discarded, and at each site a single allele was randomly sampled (-doAbbababa 1). Transitions were also removed (-rmTrans 1) from the analysis to avoid biases introduced by ancient DNA damage and the following options were used: -doCounts 1 -useLast 1 -blockSize 1000000. We tested all triplets of samples, using the Andean fox as the outgroup. The subset of triplets representing the correct tree topology, as estimated by ASTRAL-III and DiscoVista, were considered for testing gene flow hypotheses between the Sardinian dhole and the other species. Using a similar approach, we also investigated gene flow between Asian dholes and AHDs. Finally, D-statistics with a Z-score between 3 and −3 were not considered significant.

Calling of polymorphism and filtering

We used GATK v3.4.0 with the option -T *Haplotype caller*[53] tool to call variants for all the samples in our dataset, using the gray wolf genome as reference (option -R). The option -L was used to specify the regions of our interest (scaffold above 1 Mb) and sites with base quality and mapping quality lower than 20 (-mmq 20 -mbq 20) were discarded. The resulting *vcf* files were then compressed and validated using VCFtools v0.1.8[72] option *VCF-validator*.

VCFtools was subsequently used for further filtering the VCF files by excluding indels, sites with minimum depth below 5 (-minDP 5). All the sites with more than 90% of missing genotypes (–max-missing-count) over all individuals were excluded and the option–plink was used to generate the plink files in MAP and PED format. The plink files were then merged using Plink v1.9[60,61] and used to generate the files BED BIM and FAM. The following options were applied:–merge-list,–allow-extra-chr and–keep-allele-order.

qpGraph

qpGraph is part of ADMIXTOOLS software[58] and we used it to reconstruct the different relationships across the species in the study by comparing the various f statistics (f2, f3, f4) and generating an admixture graph with the best fitting admixture proportions and branch length (in unit of genetic drift). In particular, the graph generated using this tool is the representation of the relationship between 6 main lineages: Sardinian dhole, Asian dhole, AHDs, Ethiopian wolf and gray wolves.

In order to run qpGraph the PLINK files (BED BIM and FAM) were converted into EIGENSTRAT format using the package ConvertF implemented in ADMIXTOOLS. Then a qpGraph par file was created to specify working directory, the genotypename, snpname and indivname files. The options hires: YES, lsqmode: YES, blgsize: 0.005 and diag: 0.0001 were applied.

The samples were clustered into different groups representing the main lineages: AHDs, Asian dholes, Sardinian dhole, Ethiopian wolf and gray wolf. The best strategy was to start with a small graph and later add the other populations. Therefore, we built a graph topology including only the outgroup (the Andean fox), 6 AHDs, 2 Asian dholes, and the Sardinian dhole. When constructing the graph, it is important to do so considering the tree topology already available (see ASTRAL and DiscoVista trees) in order to specify the *root*, the outgroup and every other lineage representing a leaf of the graph and connected through nodes. A graph with a Z-value within 3 and −3 was considered significant, meaning that the topology of the graph fits with the combination of the f-statistics. Instead, when the Z-value is higher than 3 or lower than −3 the topology does not fully represent the relationship between the samples in the study and changes are necessary. Therefore, when constructing the graph, the Sardinian dhole was modeled as sister clade to all possible internal and external nodes and as admixed from different node pairs and the Z values were then evaluated. Subsequently, once confirmed that the graph had a good fit (- 3 > Z < +3) we added the Ethiopian wolf and later a gray wolf (Portuguese wolf) and evaluated their fit in the phylogeny as explained above.

Split time analysis - F(A|B)

We also investigated the divergence time between the two dhole lineages by computing the probability F(A|B) that an (ancient) individual A (in this case the Sardinian dhole) carries a derived heterozygous allele in an individual B (Asian dhole). We estimated the standard error using a block jackknife estimate of the statistic, using a block size of 1 MB to partition the genome into non-overlapping regions. The assumption behind this approach is that when two populations start to diverge, they will also accumulate mutations that - due to isolation - will not be shared with other populations. Therefore, we first polarized the alleles to ancestral and derived using the Andean fox genome and we then called haplotypes with minimum base quality of 25 for the population B that we used to select only the heterozygous sites. We used these sites to compute the probability that the individual A would carry the derived allele, by randomly sampling each allele. To calibrate the probability F(A|B) to the population size history of the Asian dhole, we computed PSMC[38] on the higher coverage genome Asian dhole in our dataset (Beijing Zoo dhole). The PSMC inference was based on the parameters, -N25 -t15 -r5 -p "4+25*2+4+6," and the output was used to simulate 900 mb through msHOT.[73,75] Different divergence times were computed, every 10 ka, spanning from 30 ka to 400 ka (expressed in generation time). We found the divergence time range by identifying the intersection between the empirical value line and the expected decay of F(A|B) as a function of split time.

hPSMC

To estimate the end of gene flow between the Sardinian dhole and the Asian dhole we used hPSMC.[41] We first used ANGSD to generate haploid consensus sequences mapped to the domestic dog reference genome and considering only autosomes. Bases with base quality lower than 30 and reads with mapping quality lower than 30 were discarded (-minQ 30 -minmapq 30). The minimum depth was set to 2x (-setminDepth 2) and the following quality filters: -remove_bads 1, -uniqueOnly 1, -baq 1 and -C 50. The two fasta files generated were combined into a diploid sequence using the hPSMC tool psmcfa_from_2_fastas.py. Subsequently, we ran the psmcfa output through PSMC with the parameters (-p) "4+25*2+4+6," number of iterations = 25 (-N25), maximum 2N0 coalescent time = 15 (-t15), initial theta/rho ratio = 5 (-r5). We used psmc_plot.pl to translate this information into a plot assuming a mutation rate of 4×10^{-9} per base pair per generation[39] and 3 years generation time.[39,40] The pre-divergence effective population size (Ne) estimated from the output was used to run simulations using hPSMC_quantify_split_time.py script from hPSMC tool with different divergence times, between 100,000 and 700 ka in 50 ka intervals using ms.[75]

Current Biology, Volume *31*

Supplemental Information

Evolutionary history of the extinct Sardinian dhole

Marta Maria Ciucani, Julie Kragmose Jensen, Mikkel-Holger S. Sinding, Oliver Smith, Saverio Bartolini Lucenti, Erika Rosengren, Lorenzo Rook, Caterinella Tuveri, Marisa Arca, Enrico Cappellini, Marco Galaverni, Ettore Randi, Chunxue Guo, Guojie Zhang, Thomas Sicheritz-Pontén, Love Dalén, M. Thomas P. Gilbert, and Shyam Gopalakrishnan

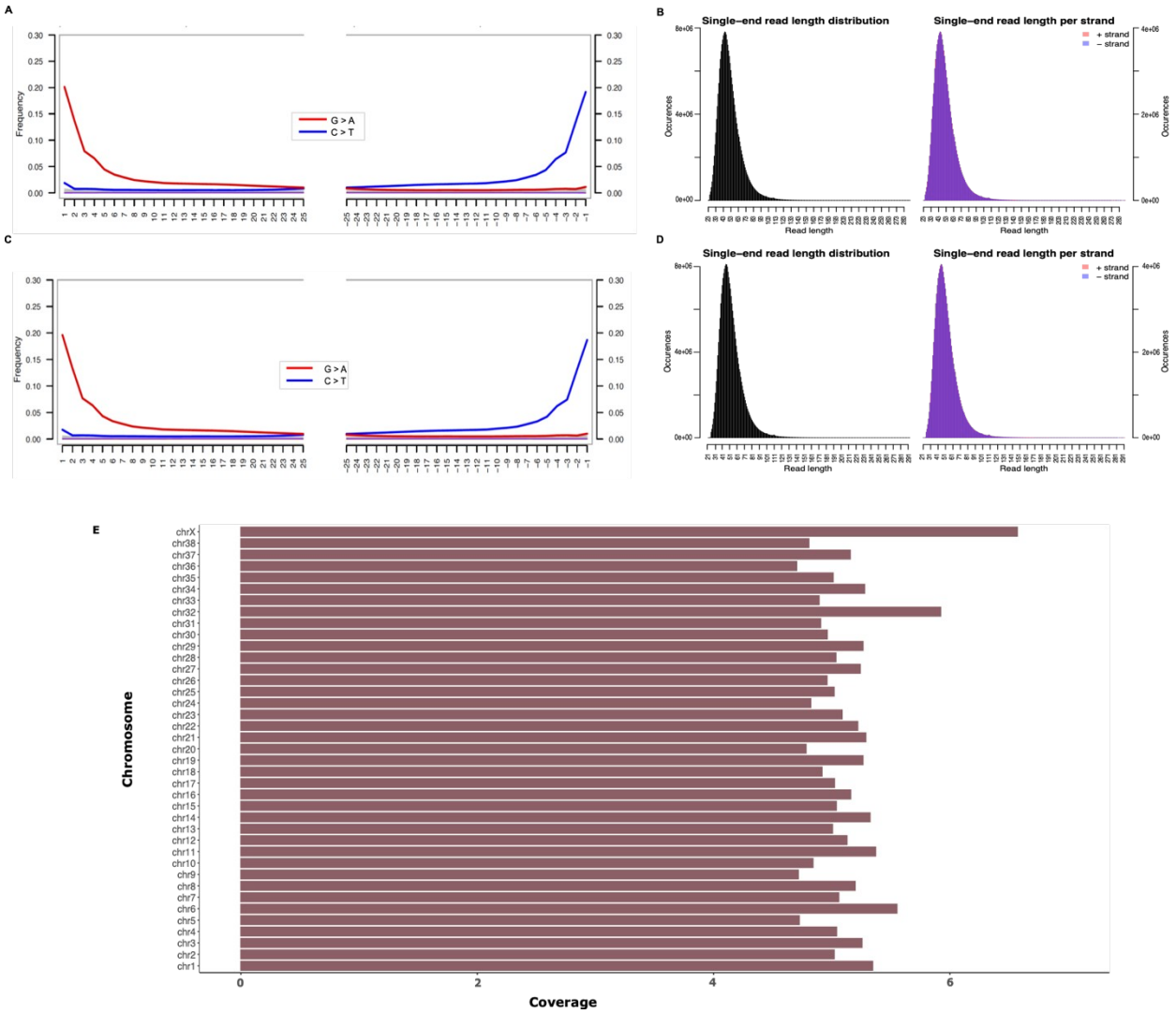

Figure S1. Nucleotide misincorporation rates, length distribution and chromosome coverage of the Sardinian dhole. Related to STAR Methods sections "Quality control and alignment", "Sex determination" and table S2.

A) Nucleotide misincorporation rates of the sample mapped to the wolf reference genome[S1].

B) Length distribution of the total reads mapped to the wolf reference genome.

C) Nucleotide misincorporation rates of the sample mapped to the dog reference genome (CanFam3.1)[S2].

D) Length distribution of the total reads mapped to the CanFam3.1 reference genome.

E) Biological sex determination. The sequencing reads sample were mapped to the dog reference genome (CanFam3.1), which present the genome divided in 38 autosomal chromosomes and the sex chromosome X. The plot here shows on the x axis the average depth of coverage at each chromosome of the Sardinian dhole which resulted to be ~ 5x across the whole genome. However, the average depth of coverage for the chromosome X was higher (> 6x) than the average depth of each chromosome and the whole genome, thus we concluded that our sample was female.

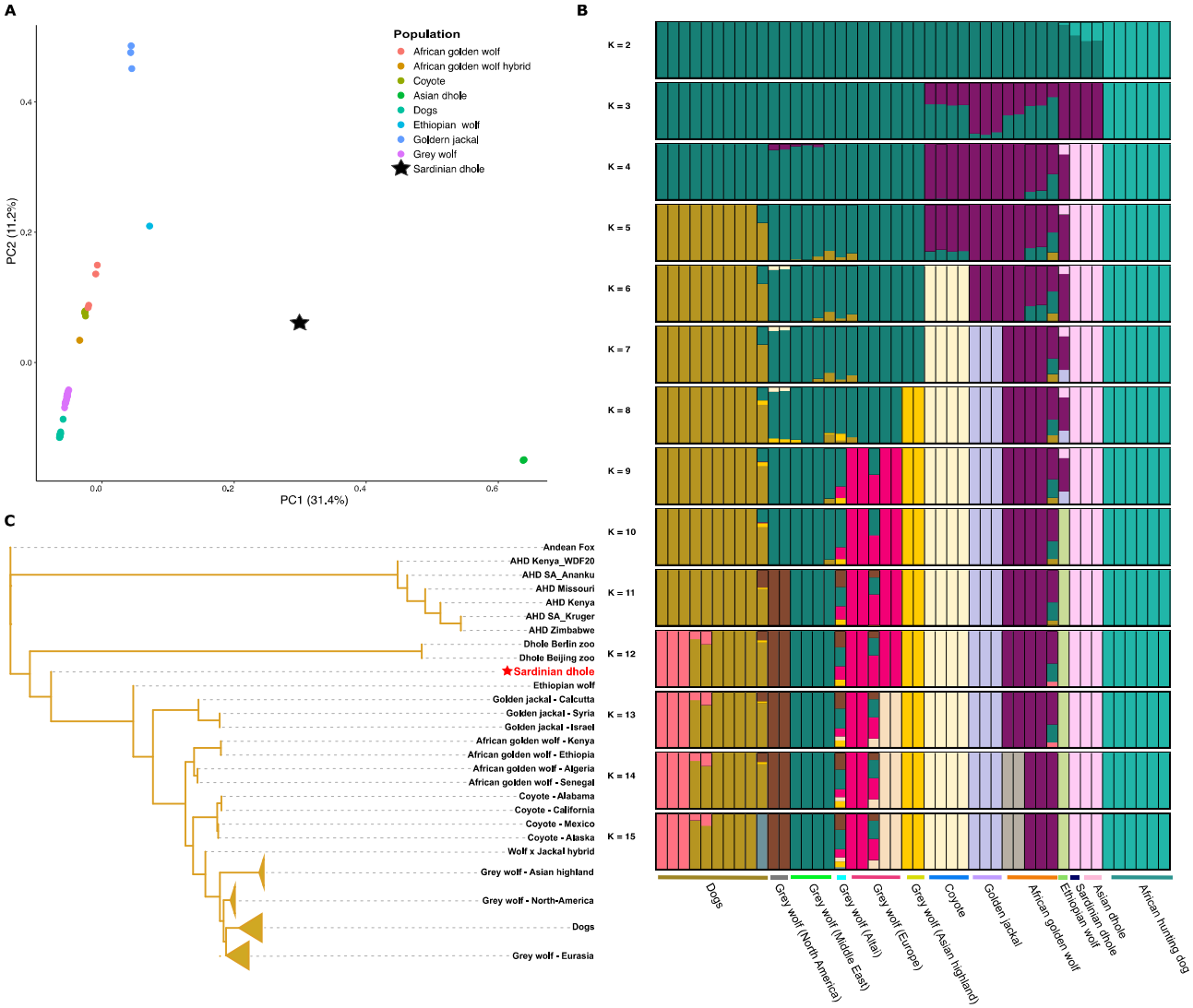

Figure S2. PCA, Admixture proportions, and Astral tree. Related to Figure 2.

A) Principal component analysis (PCA) of 41 canids excluding Andean fox and the African hunting dogs

B) The admixture proportions were estimated using NGSAdmix[S3], for a range of ancestry components from K=2 to K=15.

C) The phylogeny of 47 samples was computed by Astral III[S4], on 1000 gene trees generated with RAxML-ng[S5] under the evolutionary model GTR-GAMMA.

Dogs, Eurasian, North-American and Chinese highland wolf nodes were collapsed separately.

The node labels represent the posterior probability computed by RAxML-ng using 100 replicated in Astral-III. Branch nodes with a posterior probability above 0.75 are shown.

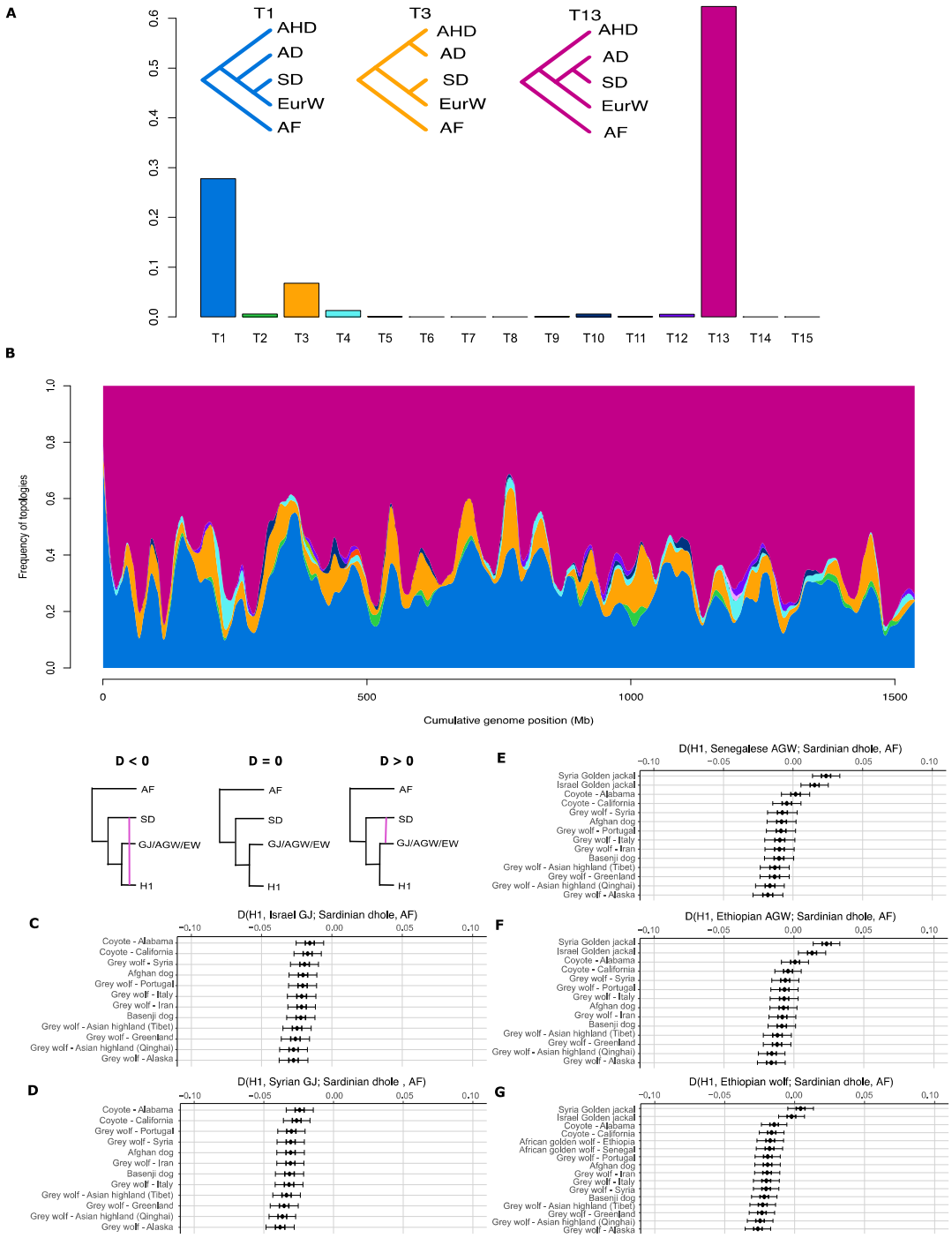

Figure S3. Topology weighting and D-stats. Related to Figure 3

In panel A-B) the results of Twisst[S6] are shown. This analysis was performed to better understand the relative frequencies of the different topologies, particularly the topologies with the different placement of the Sardinian dhole with respect to the Asian dhole and the grey wolf.

The analysis was restricted on 704 largest scaffolds of the wolf reference genome, which were larger than 1 Mb, and to only 5 groups of samples, viz. the Andean fox (AF) as the outgroup, the Sardinian dhole (SD), the Asian dholes (AD), the African hunting dogs (AHD), and Eurasian grey wolves (EurW). The main 3 topologies (T1, T2, T3) are represented in color blue, orange and

magenta respectively. The frequency of the topologies is shown on the y-axis and in panel B) the cumulative genome position (in Mb) of the windows is expressed on the x-axis.

In panels C-G of this figure the gene flow among different canids is shown using ABBA-BABA test in ANGSD [S7]. The following three different combinations were tested: 1) (((*Canis lupus*, Golden Jackal), Sardinian dhole), Andean fox), 2) (((*Canis lupus*, African golden wolf), Sardinian dhole), Andean fox) and 3) (((*Canis lupus*, Ethiopian wolf), Sardinian dhole), Andean fox).
In certain places in the figure the species names are abbreviated with AF, SD, GJ, AGW and EW for Andean for, Sardinian dhole, Golden jackal, African golden wolf and Ethiopian wolf respectively.
In the figure C-D) the first combination is shown and there is significant allele sharing ($D < 0$) between the coyotes, wolves and dogs to/from Sardinian dhole (SD) when considering two Golden jackals (GJ) (Syria and Israel golden jackals) in H2.
In panel E-F) the combination 2) was tested and the allele sharing between the AGW and SD when placing the GJ in H1. However, when considering *Canis lupus* in H1 the D-stats shift to a negative result ($D < 0$) indicating gene flow into the Sardinian dhole.
In figure G) the combination 3) is shown and when considering all the *Canis* crown species but Golden jackal (GJ) the D-statistics show a negative D ($D < 0$ - BABA) meaning that there is allele sharing between these species and the Sardinian dhole (SD).

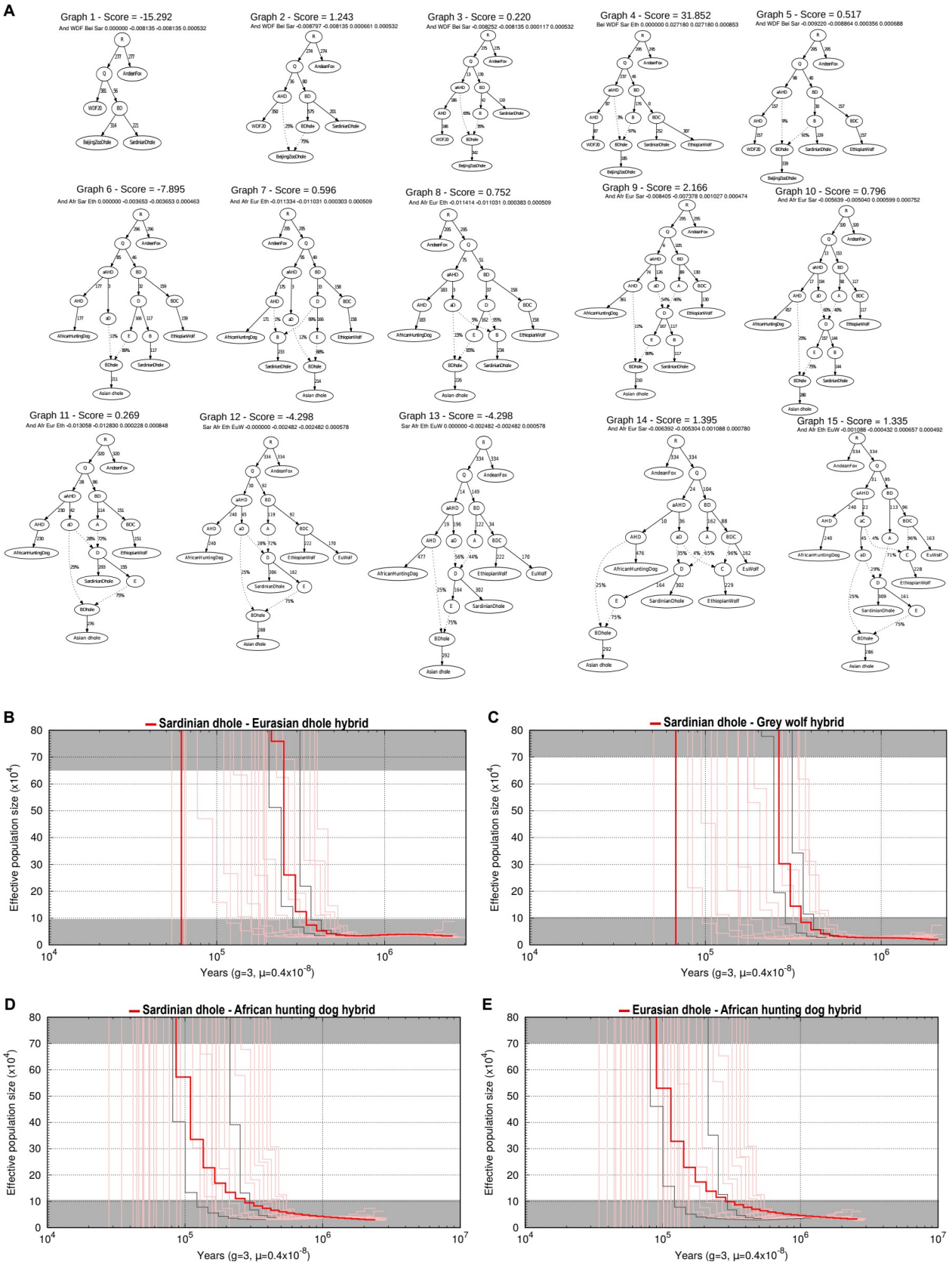

Figure S4. Admixture graph population history models and hybrid PSMC (hPSMC) plots. Related to Figure 3C and Figure 4C.

A) The demographic models of the Sardinian dhole, modern African hunting dogs, Eurasian dholes, Ethiopian wolves and Eurasian wolves, estimated using qpGraph of Admixtools[S8] with the observed and expected f-statistics to compute the admixture graph**.**
B-E) Hybrid PSMC (hPSMC)[S9] plot of the Sardinian with Asian dholes and African hunting dogs with simulations of different divergence times spanning 100,000-700,000 years in 50,000 years intervals. Greyed out regions represent 1.5x and 10x the pre-divergence effective population size. Red lines in bold represent the hPSMC results based on the real data while the remaining red thin lines represent the simulated data. The simulation closest to the real data that do not overlap with it are shown in dark grey and were used to infer the time range in which the gene flow ceased between B) the Sardinian dhole and the Eurasian dhole, C) the Sardinian dhole and European wolf (Portugal), D) the Sardinian dhole and African hunting dog (Zimbabwe) and E) the Eurasian dhole and African hunting dog (Zimbabwe).

Supplemental references

S1. Gopalakrishnan S, Samaniego Castruita JA, Sinding M-HS, Kuderna LFK, Räikkönen J, Petersen B, Sicheritz-Ponten T, Larson G, Orlando L, Marques-Bonet T, Hansen AJ, Dalén L and Gilbert MTP (2017). The wolf reference genome sequence (Canis lupus lupus) and its implications for Canis spp. population genomics. BMC Genomics *18*: 495.

S2. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, Clamp M, Chang JL, Kulbokas III EJ, Zody MC, Mauceli E, Xie X, Breen M, Wayne RK, Ostrander EA, Ponting CP, Galibert F, Smith DR, deJong PJ, Kirkness E, Alvarez P, Biagi T, Brockman W, Butler J, Chin C-W, Cook A, Cuff J, Daly MJ, DeCaprio D, Gnerre S, Grabherr M, Kellis M, Kleber M, Bardeleben C, Goodstadt L, Heger A, Hitte C, Kim L, Koepfli K-P, Parker HG, Pollinger JP, Searle SMJ, Sutter NB, Thomas R, Webber C, Broad Sequencing Platform members and Lander ES (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature *438*: 803–819.

S3. Skotte L, Korneliussen TS and Albrechtsen A (2013). Estimating individual admixture proportions from next generation sequencing data. Genetics *195*: 693–702.

S4. Zhang C, Rabiee M, Sayyari E and Mirarab S (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics *19*: 153.

S5. Kozlov AM, Darriba D, Flouri T, Morel B and Stamatakis A (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics *35*: 4453–4455.

S6. Martin SH and Van Belleghem SM (2017). Exploring Evolutionary Relationships Across the Genome Using Topology Weighting. Genetics *206*: 429–438.

S7. Korneliussen TS, Albrechtsen A and Nielsen R (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics *15*: 356.

S8. Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T and Reich D(2017). Ancient admixture in human history. Genetics *192*: 1065–1093.

S9. Cahill JA, Soares AER, Green RE and Shapiro B (2016). Inferring species divergence times using pairwise sequential markovian coalescent modelling and low-coverage genomic data. Philos. Trans. R Soc. Lond. B Biol. Sci. *371*. doi:10.1098/rstb.2015.0138