Dire wolves were the last of an ancient New World canid lineage

Angela R. Perri[1,*§], Kieren J. Mitchell[2,*§], Alice Mouton[3,*], Sandra Álvarez-Carretero[4,*], Ardern Hulme-Beaman[5,6], James Haile[7], Alexandra Jamieson[7], Julie Meachen[8], Audrey T. Lin[7,9,10], Blaine W. Schubert[11], Carly Ameen[12], Ekaterina E. Antipina[13], Pere Bover[14], Selina Brace[15], Alberto Carmagnini[4], Christian Carøe[16], Jose A. Samaniego Castruita[16], James C. Chatters[17], Keith Dobney[5], Mario dos Reis[4], Allowen Evin[18], Philippe Gaubert[19], Shyam Gopalakrishnan[16], Graham Gower[2], Holly Heiniger[2], Kristofer M. Helgen[20], Josh Kapp[21], Pavel A. Kosintsev[22,23], Anna Linderholm[7, 24], Andrew T. Ozga[25, 26, 27], Samantha Presslee[28], Alexander T. Salis[2], Nedda F. Saremi[21], Colin Shew[3], Katherine Skerry[26], Dmitry E. Taranenko[29], Mary Thompson[30], Mikhail V. Sablin[31],Yaroslav V. Kuzmin[32, 33], Matthew J. Collins[34, 35], Mikkel-Holger S. Sinding[16, 36], M. Thomas P. Gilbert[16, 37], Anne C. Stone[25 ,26], Beth Shapiro[21, 38], Blaire Van Valkenburgh[3], Robert K. Wayne[3], Greger Larson[7], and Alan Cooper[39], Laurent A. F. Frantz[4, 40§].

[1]Department of Archaeology, Durham University, Durham, UK
[2]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Australia
[3]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA
[4]School of Biological and Chemical Sciences, Queen Mary University of London, London, UK
[5]Department of Archaeology, Classics and Egyptology, University of Liverpool, Liverpool, UK
[6]School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool, UK
[7]The Palaeogenomics & Bio-Archaeology Research Network, Research Laboratory for Archaeology and History of Art, The University of Oxford, Oxford, UK
[8]Department of Anatomy, Des Moines University, Des Moines, IA, USA
[9]Department of Zoology, University of Oxford, Oxford, UK
[10] Department of Anthropology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA
[11]Center of Excellence in Paleontology & Department of Geosciences, East Tennessee State University, Johnson City, TN, USA
[12]Department of Archaeology, University of Exeter, Exeter UK
[13]Institute of Archaeology, Russian Academy of Sciences, Moscow, Russia
[14]ARAID Foundation, Instituto Universitario de Investigación en Ciencias Ambientales (IUCA) - Aragosaurus Group, Universidad de Zaragoza, Spain
[15]Department of Earth Sciences, Natural History Museum, London, UK
[16]Section for Evolutionary Genomics, Natural History Museum of Denmark, University of Copenhagen, Denmark
[17]Applied Paleoscience, Bothell, WA, USA
[18]Institut des Sciences de l'Evolution - Montpellier, CNRS, Université de Montpellier, IRD, EPHE, France
[19]Laboratoire Evolution & Diversité Biologique, UPS/CNRS/IRD, Université Paul Sabatier, Toulouse, France
[20]School of Biological Sciences and Environment Institute, University of Adelaide, Australia

[21]*Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA,USA*
[22]*Institute of Plant and Animal Ecology, Urals Branch of the Russian Academy of Sciences, Yekaterinburg, Russia*
[23]*Ural Federal University, Yekaterinburg, Russia*
[24]*Department of Anthropology, Texas A&M University, College Station, TX, USA*
[25]*Center for Evolution and Medicine, Arizona State University, Tempe, AZ, USA*
[26]*School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA*
[27]*Halmos College of Natural Sciences and Oceanography, Nova Southeastern University, Fort Lauderdale, FL, USA*
[28]*Department of Archaeology, University of York, York, UK*
[29]*Institute of Systematics and Ecology of Animals, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia*
[30]*Idaho Museum of Natural History, Idaho State University, Pocatello, ID, USA*
[31]*Zoological Institute of the Russian Academy of Sciences, Universitetskaya nab. 1, 199034, St. Petersburg, Russia*
[32]*Sobolev Institute of Geology and Mineralogy, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia*
[33]*Tomsk State University, Tomsk, Russia*
[34]*Department of Biology, University of Copenhagen, Copenhagen, Denmark*
[35]*McDonald Institute for Archaeological Research, University of Cambridge, Cambridge, UK*
[36]*Greenland Institute of Natural Resources, Nuuk, Greenland*
[37]*NTNU University Museum, Trondheim, Norway*
[38]*Howard Hughes Medical Institute, University of California Santa Cruz, Santa Cruz, CA, USA*
[39]*South Australian Museum, Adelaide 5000, Australia.*
[40] *Palaeogenomics Group, Department of Veterinary Sciences, Ludwig Maximilian University, Munich, Germany.*

§ Corresponding author. Email: angela.r.perri@durham.ac.uk (A.R.P.); kieren.mitchell@adelaide.edu.au (K.J.M.); laurent.frantz@qmul.ac.uk (L.A.F.F.)
* These authors contributed equally to this work.

Abstract

Dire wolves are considered one of the most common and widespread large carnivores in Pleistocene America, yet relatively little is known about their evolution or extinction. To reconstruct the evolutionary history of dire wolves, we sequenced five genomes from sub-fossil bones dating from 13,000 to over 50,000 years ago. Our results indicate that though they were similar morphologically to the extant gray wolf, dire wolves were a highly divergent lineage that split from living canids ~5.7 million years ago. In contrast to numerous examples of hybridization across Canidae, there is no evidence for gene flow between dire wolves and either North American gray wolves or coyotes. This suggests that dire wolves evolved in isolation from the Pleistocene ancestors of these species. Our results also support an early New World origin of dire wolves, while the ancestors of gray wolves, coyotes, and dholes evolved in Eurasia and only colonized North America relatively recently.

Main Text

Dire wolves (*Canis dirus*) were large (~68 kg) wolf-like canids and among the most common extinct large carnivores of the American Late Pleistocene megafauna [1]. Dire wolf remains are present in the North American paleontological record from at least ~250,000 to ~13,000 years ago, at the end of the Pleistocene, particularly in the lower latitudes [2] (Fig. 1A). Other canid species present in Late Pleistocene North America include the slightly smaller gray wolf (*C. lupus*), the much smaller coyote (*C. latrans*), and the dhole (or Asiatic wild dog; *Cuon alpinus*), though dire wolves appear to be more common overall [1]. For example, >4,000 individuals have been excavated in California's fossil-rich Rancho La Brea tar seeps alone, where they outnumber gray wolves more than 100-fold [3,4].

Despite the abundance of dire wolf fossils, the origin, taxonomic relationships, and ultimate driver of their extinction remain unclear. Dire wolves are generally described as a sister species to [5–8], or even conspecific with the gray wolf [9]. The leading hypothesis to explain their extinction is that, due to their larger body size than gray wolves and coyotes, dire wolves were more specialized for hunting large prey, and were unable to survive the extinction of their megafaunal prey (e.g. [10–12]). To test this hypothesis, we performed geometric morphometric analyses of >700 specimens. Our results indicate that although dire wolves and gray wolves species can be differentiated, their morphology is highly similar (Supplementary Information; Fig. 1B; Supplementary Fig. 1-6). Although this morphometric similarity may partly be driven by allometry (Supplementary Information; Fig. 1B), the lack of distinctiveness between gray wolves and dire wolves has been interpreted as a result of a close evolutionary relationship [7,9]. Alternatively, a competing hypothesis maintains that these morphological similarities are the result of convergence, and that dire wolves instead are a species belonging to a separate taxonomic lineage (classified in the monotypic genus *Aenocyon*; "terrible or dreadful wolf" [13]).

To resolve the evolutionary history of dire wolves, we screened 46 sub-fossil specimens for the presence of preserved genomic DNA (Supplementary Data 1). We identified five samples from Idaho (DireAFR & DireGB), Ohio (DireSP), Tennessee (DireGWC), and Wyoming (DireNTC) ranging in age from 12,900 to >50,000 years before present, that possessed sufficient endogenous DNA to obtain both mitochondrial genomes (between ~1x and 31x coverage) and low-coverage nuclear genome sequences (~0.01x to 0.23x coverage) using hybridization capture or shotgun sequencing methods (Supplementary Information). Although we did not successfully sequence DNA from the La Brea tar seeps dire wolf specimens, one specimen did contain type I collagen (*COL1*) suitable for sequencing using paleoproteomic methods (Supplementary Data 1; Supplementary Information).

Analyses of the dire wolf *COL1* sequence suggested that they were not closely related to gray wolves, coyotes, African wolves (*C. anthus*), and dogs (*C. familiaris*) (Supplementary Fig. 7). These data, however, could not confidently resolve the relationships between more distantly related canids due to a lack of lineage-specific amino acid changes among these species [14]. Phylogenetic analyses of the mitochondrial genomes indicated that dire wolves form a well-

supported monophyletic group that is highly divergent from gray wolves and coyotes (Supplementary Fig. 10; see Supplementary Data 13 and Supplementary Table 2 for a list of the 13 species used in this analysis), contradicting recent paleontological analyses [5–7] (Figure 1B). Canid mitochondrial phylogenies, however, may not represent the true species evolutionary relationships since both admixture and incomplete lineage sorting have been shown to affect canid phylogenetic topologies [15,16].

In order to resolve the phylogenetic relationships of dire wolves, we analyzed our dire wolves' nuclear genomic data with previously published genomic data from eight extant canids: gray wolf, coyote, African wolf, dhole, Ethiopian wolf (*C. simensis*), African wild dog (*Lycaon pictus*), Andean fox (*Lycalopex culpaeus*), and gray fox (*Urocyon cinereoargenteus*; an outgroup). Of these species, the geographical ranges of gray wolves, coyotes, dholes, and gray foxes overlapped with that of dire wolves during the Pleistocene (Fig. 1A). We also generated new nuclear genome sequences from a gray wolf from Montana and from the two endemic African jackals, the black-backed and side-striped jackal (*C. mesomelas* and *C. adustus*, respectively), in order to ensure representation of all extant members of the "wolf-like canid" clade (comprising *Canis*, *Lycaon*, *Cuon*, and their extinct relatives) (Supplementary Data 13). Supermatrix analyses, based on 70 Kb to 28 Mb nuclear sequence alignments (depending on overall coverage for each dire wolf genome, see Supplementary Table 5 & 7) confirmed a distant evolutionary relationship between dire wolves and the other wolf-like canids (Fig. 2A; Supplementary Fig. 11; Supplementary Fig. 15). This analysis, however, could not definitively resolve whether dire wolves were the basal members of the wolf-like canid clade, or the second lineage to diverge after the common ancestor of the two African jackals.

We investigated canid phylogenetic relationships in greater detail using a range of species tree analyses [17,18] and D-statistics (Supplementary Information). These approaches produced concordant trees that support the monophyly of three primary lineages: dire wolves, African jackals, and a clade comprising all other extant wolf-like canids (Fig. 2A; Supplementary Fig. 11-15). Although our species tree analyses provided equivocal results regarding the relationships among these lineages, gray wolves (genus *Canis*) are more closely related to dholes (Supplementary Fig. 21) (genus *Cuon*), African wild dogs (genus *Lycaon*) (Supplementary Fig. 19) and Ethiopian wolves (Supplementary Fig. 22) than to either dire wolves or African jackals (both genus *Canis*). This finding is consistent with previously proposed designations of genus *Lupulella* [19] for the African jackals and *Aenocyon* [13] for dire wolves.

To assess the timing of divergence among the major wolf-like canid lineages we performed a molecular clock analysis based on two fossil calibrations using MCMCtree[20]. Although the dire wolf sequences are low coverage and include post-mortem damage, extensive simulations indicated this is unlikely to affect the time of divergence estimates inferred by MCMCtree (Supplementary Information; Supplementary Fig. 17). This analysis confirmed that the initial divergences of the three primary wolf-like canid lineages occurred rapidly, contributing to the poor resolution of the tree as a result of incomplete lineage sorting (Fig. 2A). The dire wolf lineage last shared a common ancestor with extant wolf-like canids ~5.7 million years ago (95%

HPD=4.0-8.5 million years ago; Fig. 2A), followed by the divergence of African jackals ~5.1 million years ago (95% HPD=3.5-7.6 million years ago; Fig. 2B).

Given the propensity for sympatric canid species to interbreed [15,21,22], we tested for genomic signals of admixture between extant North American canids and dire wolves using D statistics[23] (Supplementary Information) in a data set that included 22 modern North American gray wolves and coyotes, three ancient dogs [24–26] , and a Pleistocene wolf [27] (Supplementary Data 13). Specifically, we computed statistics of the form D (outgroup [gray fox]; dire wolf; North American canid [gray wolf or coyote], African wolf/Eurasian wolf) and found no significant excess of shared derived alleles between dire wolves and any extant North American canid (Fig. 2B; Supplementary Fig. 18; Supplementary Data 14). This result indicates that the dire wolves sequenced in this study did not possess ancestry from gray wolves, coyotes, or their recent North American ancestors. Although we cannot exclude the possibility that some unsampled canid population has some dire wolf hybrid ancestry, the lack of signal of hybridization in our broad set of genomes suggests that admixture is unlikely to have occurred. While we did not find evidence of recent admixture, we did find that African wild dogs share fewer derived alleles with dire wolves than with gray wolves, coyote, African wolves, dhole, or Ethiopian wolves (Fig. 2C; Supplementary Fig. 20; Supplementary data 15). This indicates that an episode of ancient admixture between the ancestor of dire wolves and the ancestor of wolves, coyotes, and dhole occurred at least ~3 million years ago (based on the lower bound of the 95% HPD on the age of their common ancestor; Fig. 2A), which may have contributed challenges resolving the branching order of the basal wolf-like canid lineages (Fig. 2A).

Hybridization is common among wolf-like canid lineages when their ranges overlap. For example, modern gray wolves and coyotes hybridize readily in North America (e.g., [21]). Genomic data also suggest gene flow occurred between dholes and African wild dogs during the Pleistocene [15], millions of years after their divergence. Consequently, our finding of no evidence for gene flow between dire wolves and gray wolves, coyotes, or their common ancestor, despite substantial range overlap with dire wolves during the Late Pleistocene suggests that the common ancestor of gray wolves and coyotes probably evolved in geographical isolation from members of the dire wolf lineage. This result is consistent with the hypothesis that dire wolves originated in the Americas [1,4,28,29], likely from the extinct Armbruster's wolf (*C. armbrusteri* [5]).

Long term isolation of the dire wolf lineage in the Americas implies that other American fossil taxa, such as the Pliocene *C. edwardii*, a proposed relative of the coyote [5], may instead belong to the dire wolf lineage. Thus, the diversification of the extant wolf-like canids likely occurred in parallel outside of the Americas, and perhaps began earlier than hypothesized. The living *Canis* species may have descended from Old World members of the extinct genus *Eucyon*, which first appeared in the fossil record of Africa and Eurasia at the end of the Miocene (see [30]). Geographic isolation since the late Miocene is consistent with our molecular estimates for the age of the dire wolf lineage, and may have allowed dire wolves to evolve some degree of reproductive isolation prior to the Late Pleistocene North American arrival of gray wolves, coyotes, dholes, and *Xenocyon* (another extinct wolf-like canid).

Despite their overall phenotypic similarities, gray wolves and coyotes survived the Late Pleistocene megafaunal extinctions while dire wolves did not. One possible reason may be that both gray wolves and coyotes possessed greater morphological plasticity and dietary flexibility, thus allowing them to avoid extinction and become the dominant terrestrial predators in North America [12,31]. This scenario is supported by the date we obtained from the DireGWC specimen (12,820-12,720 calBP), which suggests that dire wolves survived until at least the Younger Dryas cold reversal, a period that also witnessed the latest known dates for other specialized North American mega-carnivores such as the American lion (*Panthera atrox*) and giant short-faced bear (*Arctodus simus*)[32,33]. Alternatively, gray wolves and coyotes may have survived as a result of their ability to hybridize with other canids. Through adaptive introgression with dogs, North American gray wolves are known to have acquired traits related to coat color, hypoxia, and immune response [34,35]. Specifically, enhanced immunity may have allowed gray wolves to resist novel diseases carried by newly arriving Old World taxa. Since our results demonstrate that dire wolves did not derive any ancestry from other wolf-like canid species, it is plausible that reproductive isolation prevented dire wolves from acquiring traits that may have allowed them to survive into the Holocene.

References

1. Dundas, R. G. Quaternary records of the dire wolf, Canis dirus, in North and South America. *Boreas* **28**, 375–385 (1999).
2. Meachen, J. A., Brannick, A. L. & Fry, T. J. Extinct Beringian wolf morphotype found in the continental U.S. has implications for wolf migration and evolution. *Ecol. Evol.* **6**, 3430–3438 (2016).
3. Leonard, J. A. *et al.* Megafaunal extinctions and the disappearance of a specialized wolf ecomorph. *Curr. Biol.* **17**, 1146–1150 (2007).
4. Kurtén, B. & Anderson, E. *Pleistocene mammals of North America*. (Columbia University Press, 1980).
5. Tedford, R. H., Wang, X. & Taylor, B. E. Phylogenetic Systematics of the North American Fossil Caninae (Carnivora: Canidae). *Bull. Am. Mus. Nat. Hist.* **325**, 1–218 (2009).
6. Prevosti, F. J. Phylogeny of the large extinct South American Canids (Mammalia, Carnivora, Canidae) using a ‘total evidence’ approach. *Cladistics* **26**, 456–481 (2010).
7. Zrzavý, J., Duda, P., Robovský, J., Okřinová, I. & Pavelková Řičánková, V. Phylogeny of

the Caninae (Carnivora): Combining morphology, behaviour, genes and fossils. *Zool. Scr.* **47**, 373–389 (2018).

8. Álvarez-Carretero, S., Goswami, A., Yang, Z. & Dos Reis, M. Bayesian Estimation of Species Divergence Times Using Correlated Quantitative Characters. *Systematic Biology* vol. 68 967–986 (2019).
9. Goulet, G. D. Comparison of temporal and geographical skull variation among Nearctic modern, Holocene and Late Pleistocene gray wolves (Canis lupus)(and selected Canis). (1993).
10. Graham, R. W. & Mead, J. I. Environmental fluctuations and evolution of mammalian faunas during the last deglaciation in North America. *North America and adjacent oceans during the last deglaciation* 371–402 (1987).
11. Barnosky, A. D. The late Pleistocene event as a paradigm for widespread mammal extinction. *Mass extinctions: processes and evidence* 235–254 (1989).
12. DeSantis, L. R. G. *et al.* Causes and Consequences of Pleistocene Megafaunal Extinctions as Revealed from Rancho La Brea Mammals. *Curr. Biol.* **29**, 2488–2495.e2 (2019).
13. Merriam, J. C. Note on the systematic position of the wolves of the Canis dirus group. *Bulletin of the Department of Geology of the University of California* **10**, 531–533 (1918).
14. Buckley, M., Harvey, V. L. & Chamberlain, A. T. Species identification and decay assessment of Late Pleistocene fragmentary vertebrate remains from Pin Hole Cave (Creswell Crags, UK) using collagen fingerprinting. *Boreas* **46**, 402–411 (2017).
15. Gopalakrishnan, S. *et al.* Interspecific Gene Flow Shaped the Evolution of the Genus Canis. *Curr. Biol.* **28**, 3441–3449.e5 (2018).
16. Koepfli, K.-P. *et al.* Genome-wide Evidence Reveals that African and Eurasian Golden Jackals Are Distinct Species. *Curr. Biol.* **25**, 2158–2165 (2015).
17. Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A. & RoyChoudhury, A. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full

coalescent analysis. *Mol. Biol. Evol.* **29**, 1917–1932 (2012).

18. Yang, Z. The BPP program for species tree estimation and species delimitation. *Current Zoology* vol. 61 854–865 (2015).
19. Geraads, D. A revision of the fossil Canidae (Mammalia) of north-western Africa. *Palaeontology* **54**, 429–446 (2011).
20. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
21. vonHoldt, B. M. *et al.* 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 (2016).
22. vonHoldt, B. M. *et al.* A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. *Genome Res.* **21**, 1294–1305 (2011).
23. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
24. Sinding, M.-H. S. *et al.* Arctic-adapted dogs emerged at the Pleistocene-Holocene transition. *Science* **368**, 1495–1499 (2020).
25. Ní Leathlobhair, M. *et al.* The evolutionary history of dogs in the Americas. *Science* **361**, 81–85 (2018).
26. Frantz, L. A. F. *et al.* Genomic and archaeological evidence suggest a dual origin of domestic dogs. *Science* **352**, 1228–1231 (2016).
27. Skoglund, P., Ersmark, E., Palkopoulou, E. & Dalén, L. Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. *Curr. Biol.* **25**, 1515–1519 (2015).
28. Nowak, R. M. North American Quaternary Canis, Monograph of the Museum of Natural History, The University of Kansas. (1979).
29. Nowak, R. M. Wolf evolution and taxonomy. *Wolves: Behavior, ecology, and conservation* 239–258 (2003).

30. Sotnikova, M. & Rook, L. Dispersal of the Canini (Mammalia, Canidae: Caninae) across Eurasia during the Late Miocene to Early Pleistocene. *Quat. Int.* **212**, 86–97 (2010).
31. Van Valkenburgh, B. & Hertel, F. *The decline of North American predators during the late Pleistocene*. (4320).
32. Cooper, A. *et al.* PALEOECOLOGY. Abrupt warming events drove Late Pleistocene Holarctic megafaunal turnover. *Science* **349**, 602–606 (2015).
33. Schubert, B. W. Late Quaternary chronology and extinction of North American giant short-faced bears (Arctodus simus). *Quat. Int.* **217**, 188–194 (2010).
34. Schweizer, R. M. *et al.* Natural Selection and Origin of a Melanistic Allele in North American Gray Wolves. *Mol. Biol. Evol.* **35**, 1190–1209 (2018).
35. Anderson, T. M. *et al.* Molecular and evolutionary history of melanism in North American gray wolves. *Science* **323**, 1339–1343 (2009).

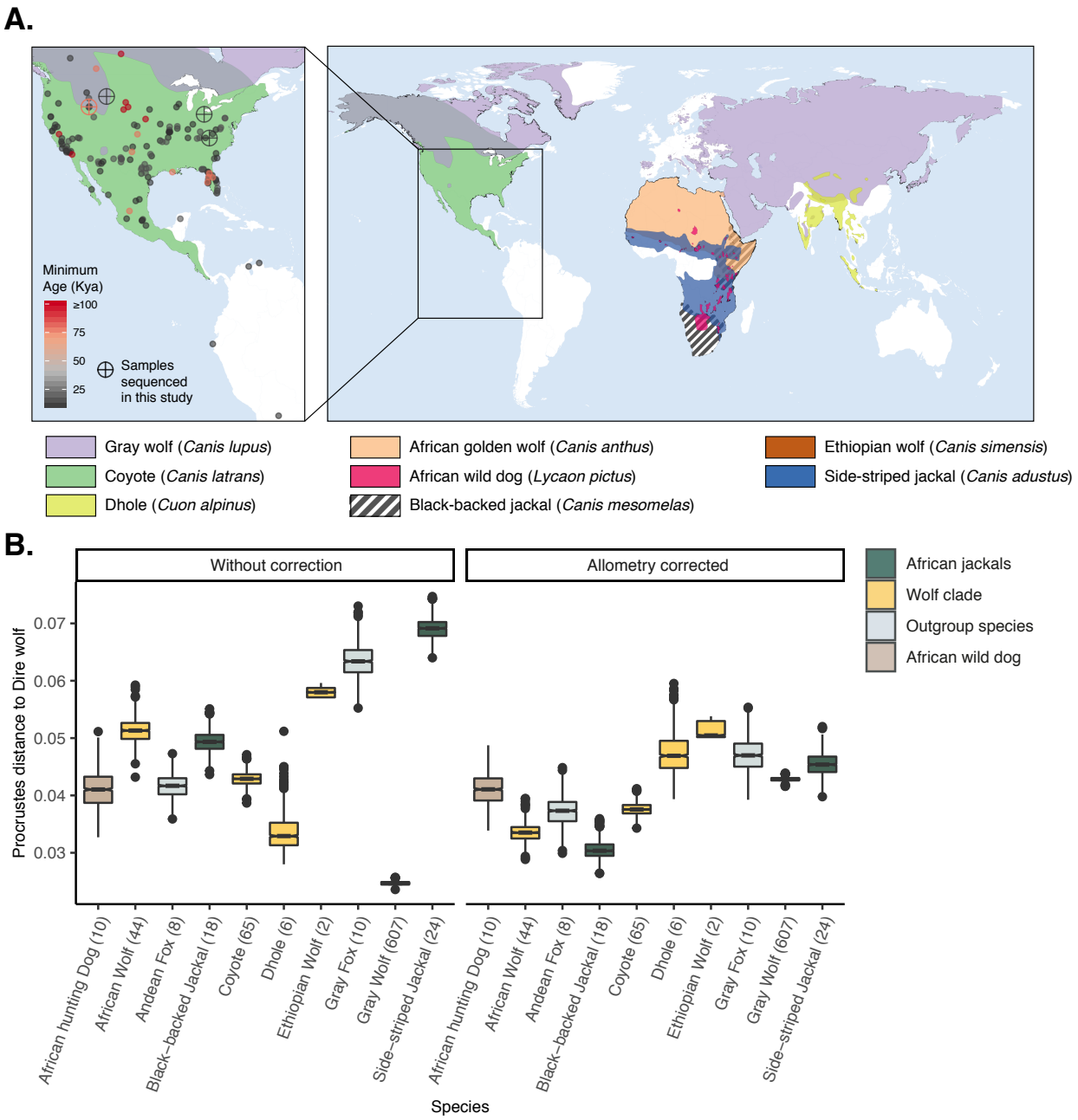


Figure 1. Map of dire wolf remains and morphological differentiation with wolf-like canids A. Right: map representing the geographic range (obtained from IUCN, https://www.iucnredlist.org) of the canid species investigated in this study. Left: map representing the distribution of sites in the Americas where dire wolf remains (*Canis dirus*) were identified (Supplementary Table 1). Colored circles represent the location and approximate age of the remains, with crossed circles representing the five samples from Idaho (2), Ohio (1), Tennessee (1), and Wyoming (1) that yielded sufficient endogenous DNA to reconstruct both mitochondrial genomes and low-coverage nuclear genome sequences. **B.** Procrustes distance between the combined mandible and M1 shape of dire wolf and other extant canid species. Pairwise procrustes distances were calculated by superimposing landmarks from molar and mandibular shapes between pairs of specimens and by computing the square root of the

330 squared differences between the coordinates of corresponding landmarks, with and without
331 correction for allometry (Supplementary Information).

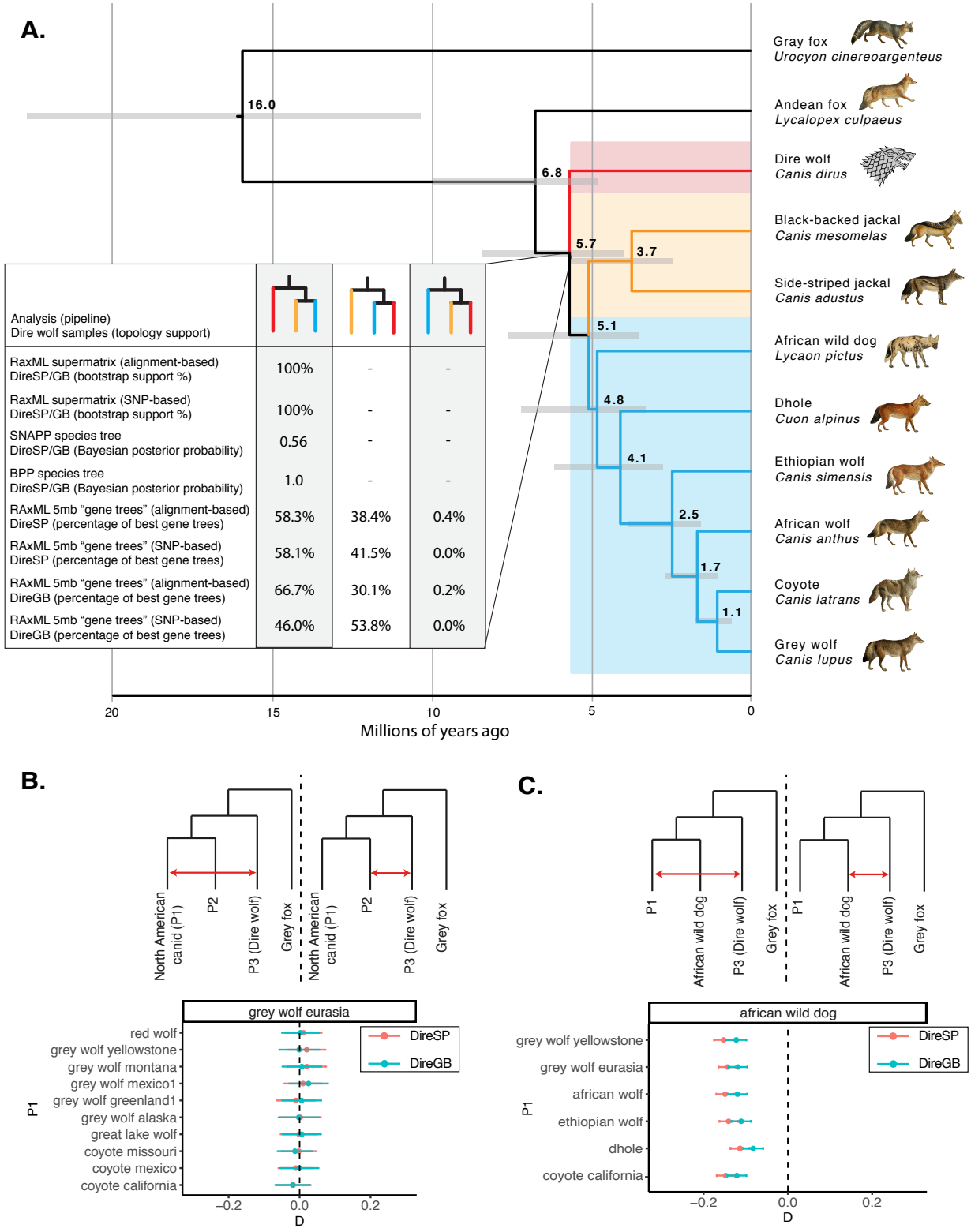


Figure 2. Relationships among living and extinct wolf-like canids A. Time-scaled nuclear phylogeny generated in MCMCtree based on the best species tree topology obtained from BPP and SNAPP. Values associated with nodes are mean age estimates (millions of years before present) while bars represent 95% Highest Posterior Densities. The inset table shows levels of

support for the three possible arrangements of the dire wolf (red), the African jackals (orange), and the remaining wolf-like canids (blue) that we obtained under different analytical frameworks (Supplementary Information) when including either one or both of our two highest coverage dire wolf samples (DireSP and DireGB). Although only one dire wolf branch is depicted in the tree, multiple dire wolf individuals were included they formed a monophyletic clade (Supplementary Fig. 12-13, and 15). **B.** Results of D statistics used to assess the possibility of gene flow between the dire wolf and extant North American canids. Each dot represents the mean D calculated along the genome and the error bar represents 3 standard deviations. These plots show that the dire wolf genomes do not share significantly more derived alleles with extant North American canids compared to Eurasian wolves (values of D not significantly different to zero), suggesting that no hybridization occurred between the dire wolf and the ancestor of extant North American canids. Non significant D-statistics were also obtained using an alternative reference genome and using the African wolf as P2 (Supplementary Fig. 18 and Supplementary Data 14). **C.** Results of D statistics showing the existence of an ancient gene flow event between the ancestor of the dhole, Ethiopian wolf, African wolf, gray wolf and coyotes and the lineage of the dire wolf (consistently non-zero values of D regardless of P1).

Acknowledgement

We thank the staff at the Carnegie Museum of Natural History, Cincinnati Museum Center, Danish Zoological Museum, Harrison Zoological Museum, Harvard Museum of Comparative Zoology, Idaho Museum of Natural History, Institute of Archaeology (Russian Academy of Sciences), Institute of Systematics and Animal Ecology (Russian Academy of Sciences), Institute of Zoology (Chinese Academy of Sciences), Instituto de Conservação da Natureza e das Florestas, Kansas Museum of Natural History, La Brea Tar Pits and Museum, Ludwig Maximilian University, McClung Museum, Museum of the Institute of Plant and Animal Ecology (Russian Academy of Sciences), Museum National d'Histoire Naturelle, National Museums Scotland, Natural History Museum London, Naturalis Biodiversity Center, Naturhistorisches Museum Bern, Smithsonian National Museum of Natural History, Swedish Naturhistoriska Riksmuseet, United States Bureau of Reclamation, University of California Museum of Paleontology, University of Texas at El Paso, University of Washington Burke Museum, and the Zoological Institute (Russian Academy of Sciences; state assignment № state assignment № AAAA-A19-119032590102-7) for access to specimens in their care. We also thank T. Barnosky, S. Bray, A. Farrell, R. Fischer, A. Harris, J. Harris, A. Henrici, P. Holroyd, R. MacPhee, T. Martin, A. Philpot, J. Saunders, J. Southon, G. Storrs, G. Takeuchi, X. Wang and C. Widga for their assistance. A.M. used computational and storage services associated with the Hoffman2 Shared Cluster provided by UCLA Institute for Digital Research and Education's Research Technology Group. DireGWC was sequenced using the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant. Further, the authors would like to acknowledge the assistance of the Danish National High-Throughput Sequencing Centre, BGI-Europe, the Garvan Institute of Medical Research and the Australian Cancer Research Foundation (ACRF) Cancer Genomics Facility for assistance in Illumina and BGIseq500 data generation. A.R.P. was supported by a Marie Curie COFUND Junior Research Fellowship (Durham University). A.M. was supported by an NSF grant (award number: 1457106) and the QCB Collaboratory Postdoctoral Fellowship (UCLA).

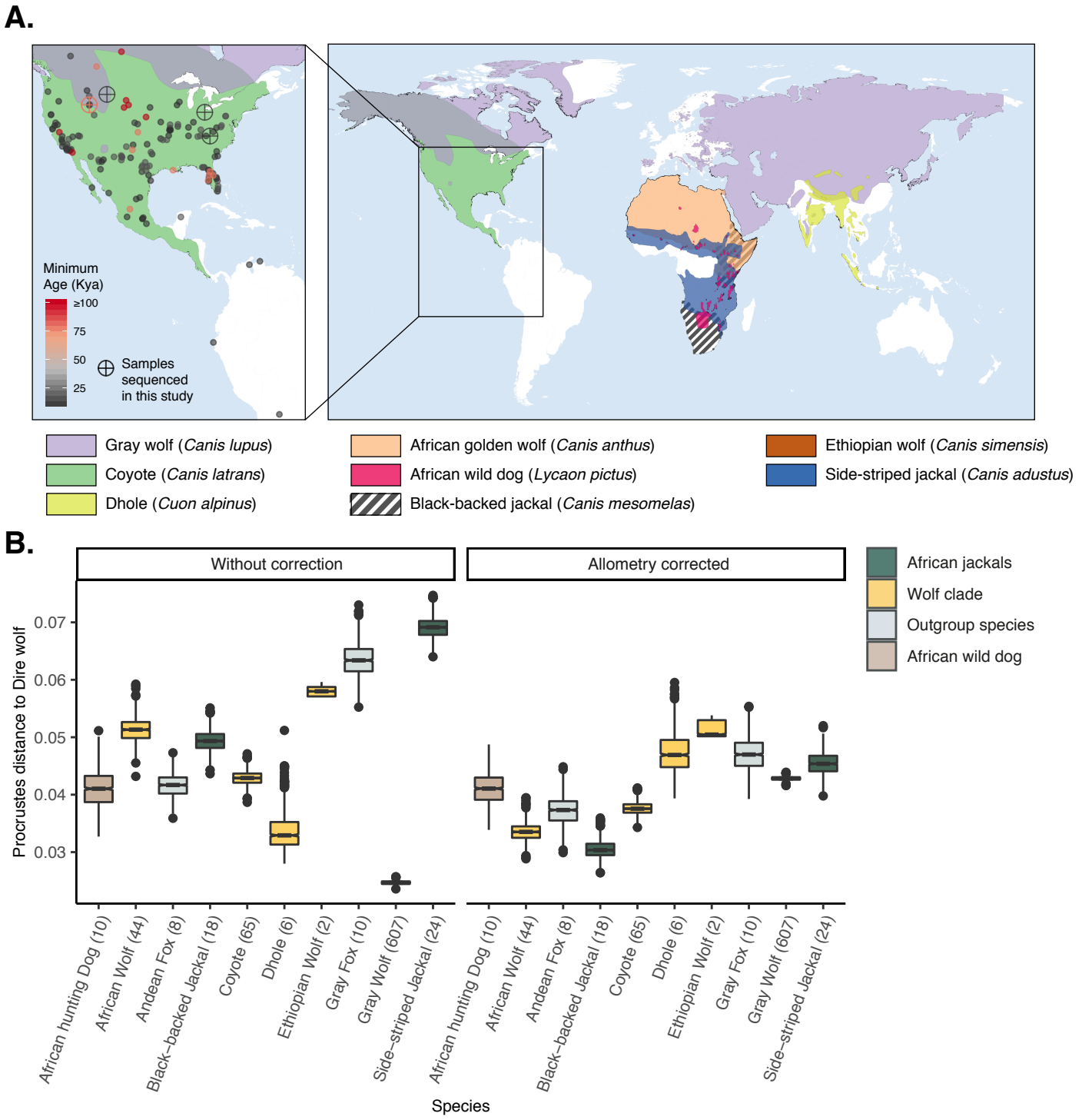
L.A.F.F., J.H., A.H-B. and G.L. were supported either by a European Research Council grant (ERC-2013-StG-337574-UNDEAD) or Natural Environmental Research Council grants (NE/K005243/1 and NE/K003259/1), or both. K.S. was supported by a grant from Barrett, the Honors College at Arizona State University. A.T.O. was supported by the Strategic Initiative Funds, Office of the President, Arizona State University to the Institute of Human Origins' DNA and Human Origins at Arizona State University project. L.A.F.F. was supported by a Junior Research Fellowship (Wolfson College, University of Oxford) and L.A.F.F. and A.Carmagnini were supported by the Wellcome Trust (210119/Z/18/Z). S.G. was supported by Carlsbergfondet grant CF14–0995 and Marie Skłodowska-Curie Actions grant 655732-WhereWolf. M.T.P.G was supported by ERC Consolidator grant 681396-Extinction Genomics. B.S. and J.K. were supported by IMLS MG-30-17-0045-17 and NSF DEB-1754451. A.H-B. was supported by the Leverhulme Trust. A.Cooper, K.J.M., and H.H. were supported by the Australian Research Council. A.S. and G.G. were supported by Australian Government Research Training Program Scholarships. A.T.L. is supported by the Peter Buck Postdoctoral Fellowship from the Smithsonian Institution's National Museum of Natural History. Finally, we thank the editors, Larisa DeSantis and the other reviewers for their helpful comments.

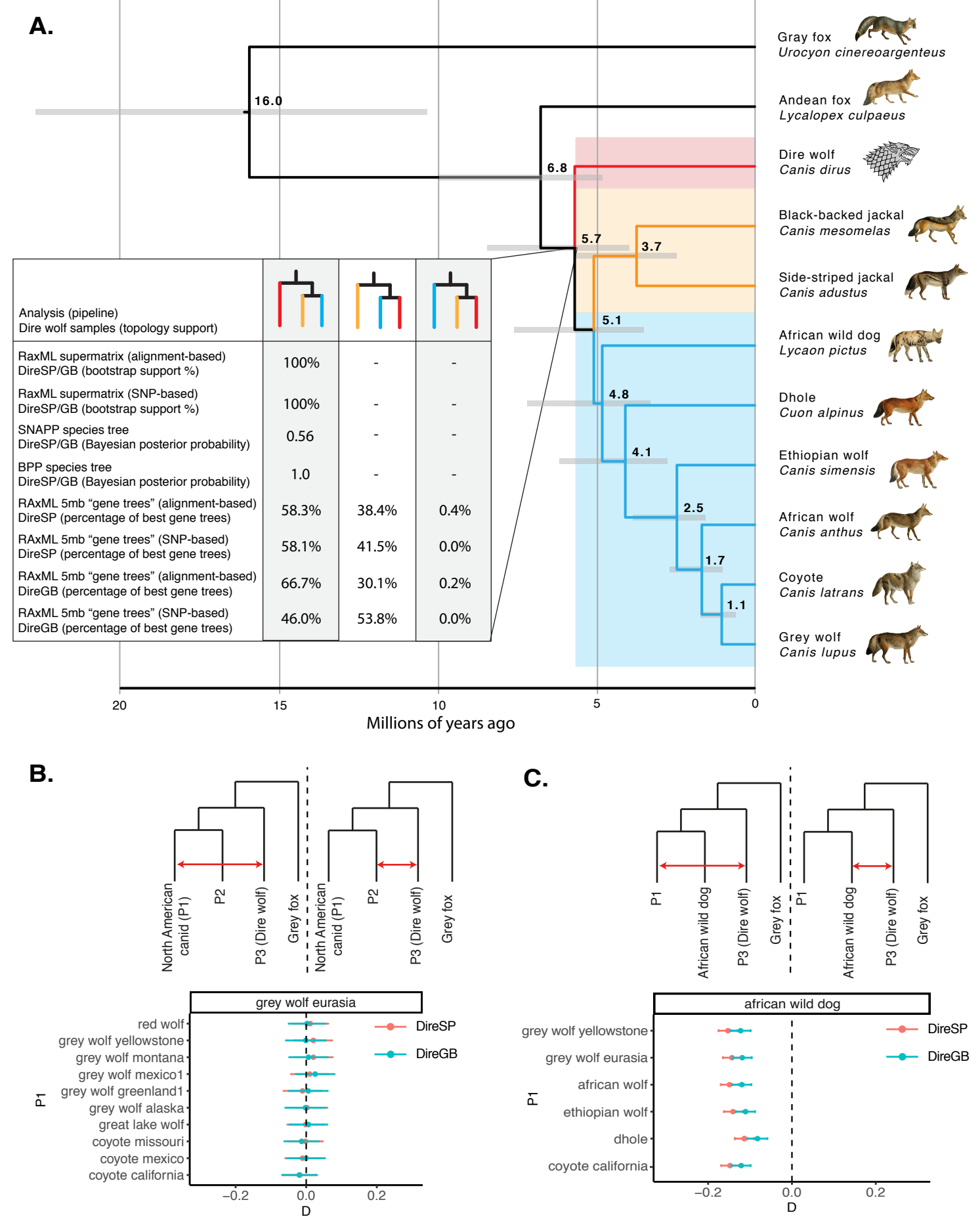
Author contributions

A.R.P., K.J.M., A.M., R.K.W., G.L., L.A.F.F., and A. Cooper conceived of the project and designed the research; A.R.P. and K.J.M. coordinated the sample collection efforts with input from R.K.W., G.L., L.A.F.F., K.M.H., and A. Cooper; A.R.P., K.J.M., A.H.-B., J.M., C.A., J.C.C., A.E., J.K., A.L., A.O., S.P., B.W.S., M.T., M.J.C., M.-M.H.S., M.T.P.G., A.C.S., B.S., B.V.V., R.W.K., and A. Cooper provided and/or collected samples; A.R.P., K.J.M., K.W., A.M., C.S., J.H., A.J., A.T.S., P.B., and H.H. conducted the genomic laboratory work; K.J.M., A.M., G.G., G.L., L.A.F.F. and A. Cooper conducted the analyses of the genomic data; S. A.-C., A.H.-B., J.M., C.A., and A.E. conducted the morphological analyses; A.R.P., K.J.M., A.M., S.A.-C., B.V.V., R.K.W., G.L., L.A.F.F. and A. Cooper wrote the paper with input from all other authors.

Data availability

The reads generated for this study have been deposited at the European Nucleotide Archive (ENA) with project number PRJEB31639. Geomorphometric data and collagen sequence were deposited onto a Dryad archive (TBD).


A.
Minimum
Age (Kya)
≥100
75
50
25
Samples
sequenced
in this study
Gray wolf (Canis lupus)
Coyote (Canis latrans)
Dhole (Cuon alpinus)
African golden wolf (Canis anthus)
African wild dog (Lycaon pictus)
Black-backed jackal (Canis mesomelas)
Ethiopian wolf (Canis simensis)
Side-striped jackal (Canis adustus)
B.
Without correction
Allometry corrected
African jackals
Wolf clade
Outgroup species
African wild dog
Procrustes distance to Dire wolf
0.07
0.06
0.05
0.04
0.03
African hunting Dog (10)
African Wolf (44)
Andean Fox (8)
Black−backed Jackal (18)
Coyote (65)
Dhole (6)
Ethiopian Wolf (2)
Gray Fox (10)
Gray Wolf (607)
Side−striped Jackal (24)
Species


A.
16.0
6.8
5.7
3.7
5.1
4.8
4.1
2.5
1.7
1.1
Gray fox
Urocyon cinereoargenteus
Andean fox
Lycalopex culpaeus
Dire wolf
Canis dirus
Black-backed jackal
Canis mesomelas
Side-striped jackal
Canis adustus
African wild dog
Lycaon pictus
Dhole
Cuon alpinus
Ethiopian wolf
Canis simensis
African wolf
Canis anthus
Coyote
Canis latrans
Grey wolf
Canis lupus
Analysis (pipeline)
Dire wolf samples (topology support)
RaxML supermatrix (alignment-based) DireSP/GB (bootstrap support %) 100% - -
RaxML supermatrix (SNP-based) DireSP/GB (bootstrap support %) 100% - -
SNAPP species tree DireSP/GB (Bayesian posterior probability) 0.56 - -
BPP species tree DireSP/GB (Bayesian posterior probability) 1.0 - -
RAxML 5mb "gene trees" (alignment-based) DireSP (percentage of best gene trees) 58.3% 38.4% 0.4%
RAxML 5mb "gene trees" (SNP-based) DireSP (percentage of best gene trees) 58.1% 41.5% 0.0%
RAxML 5mb "gene trees" (alignment-based) DireGB (percentage of best gene trees) 66.7% 30.1% 0.2%
RAxML 5mb "gene trees" (SNP-based) DireGB (percentage of best gene trees) 46.0% 53.8% 0.0%
20
15
10
5
0
Millions of years ago
B.
North American canid (P1)
P2
P3 (Dire wolf)
Grey fox
grey wolf eurasia
red wolf
grey wolf yellowstone
grey wolf montana
grey wolf mexico1
grey wolf greenland1
grey wolf alaska
great lake wolf
coyote missouri
coyote mexico
coyote california
P1
DireSP
DireGB
−0.2
0.0
0.2
D
C.
P1
African wild dog
P3 (Dire wolf)
Grey fox
african wild dog
grey wolf yellowstone
grey wolf eurasia
african wolf
ethiopian wolf
dhole
coyote california
P1
DireSP
DireGB
−0.2
0.0
0.2
D

Dire wolves were the last of an ancient New World canid lineage

Supplementary Information

Site descriptions

This section details the site descriptions for the five dire wolf samples from which we obtained mitochondrial and low-coverage nuclear genomes (Supplementary Data 1 and Supplementary Data 2). *Site name, state (sample repository;* ***repository code; lab code****;* ***publication code****; sample provenience)*

Natural Trap Cave, Wyoming (University of Kansas; ***KU48130; ACAD5529****;* ***DireNTC***; *fragment from mandible)*

Natural Trap Cave (NTC), a late Pleistocene fossil site, is a 25-meter-deep karst sinkhole on the western side of the Big Horn Mountains in north-central Wyoming. It includes a rich vertebrate biota, of mostly open-habitat associated species, including both Pleistocene-only taxa, and taxa that survived the extinction event, such as pronghorn antelope (*Antilocapra americana*) and bighorn sheep (*Ovis canadensis*) [1]. The earliest known radiocarbon date of wolves at Natural Trap Cave is approximately 25,800 years before present [2]. The wolves at Natural Trap Cave were identified as both gray wolves (*Canis lupus*) and dire wolves by [1]. But upon further morphological investigation, most of the wolves at this site have recently been identified as the extinct *Canis lupus* morph, the Beringian wolf [3]. The specimen that yielded dire wolf DNA (KU48130) was the anterior portion of a right mandible (with several premolars and molars still in situ) and has a date of 21773 calBP (19970 ± 110 BP; OxA-37752).

Gigantobison Bay, Idaho (Idaho Museum of Natural History; ***IMNH 48001/52; ACAD18742****;* ***DireGB****; petrous from complete cranium)*

Gigantobison Bay is an open Pleistocene fossil bearing site located within the American Falls Reservoir, Idaho. Seasonal changes in water level in the reservoir periodically expose and inundate the site, occasionally revealing fossils in the sediment (most commonly extinct long-horned bison - *Bison latifrons* - for which the site is named). While remains from the site have proven intractable to radiocarbon dating, a minimum age can be assigned to all specimens based on regional changes in geology. Fossils from Gigantobison Bay are found in old fluvial (river) sediments that were overlain by lacustrine (lake) sediments deposited following the damming of the Snake River by the volcanic Cedar Butte Basalt, which is dated to 72 ± 14 ky [4]. This sample is from a complete cranium with most teeth still *in situ*. We removed the left petrosal from the specimen, which successfully yielded aDNA.

Sheriden Pit, Ohio (Cincinnati Museum Center; ***VP1737; ACAD1735; DireSP***; *incisor root)*

The Sheriden Pit is part of a local cave system accessed via a sinkhole at Indian Trail Caverns in Wyandot County, northwestern Ohio. As opposed to the glacial till derived from the late Wisconsinan ice sheet that covers and fills most regional karst features, the sink contained alluvium. Glacial retreat from the region is estimated at 14,100 to 12,500 years B.P. [5] thus the sink, formed from subsurface dissolution and roof collapse, was open to the surface sometime during and/or after this period. The site was excavated by Cincinnati Museum Center (CMC) between 1990 and 1996 and has produced a diverse faunal assemblage of late Pleistocene (Rancholabrean) taxa, including extinct and extralimital species. Over 75 species have been identified to date from the site. Extinct taxa include *Arctodus*, *Castoroides*, *Cervalces*, and *Platygonus*. The assemblage is indicative of a mosaic habitat comprising open woodland and grassy ecosystems with one or more shallow, marshy ponds and streams in the area [6–8].

The sinkhole is developed in a stromatolitic reef of Middle Silurian Lockport Group dolomite. Approximately 10m deep, it formed a natural trap into which animals fell, were washed, or were carried by predators. Sediments indicate the past presence of running water into the sink and out through a horizontal cave connection at its bottom. When this connection became choked with sediment, a layer of laminated silts and clays indicative of standing water formed in the upper part of the pit. Four stratigraphic units can be identified, from top to bottom: a laminated diamicton, a layer of gray- and red- banded clay rhythmite, a lower diamicton, and a base unit of dolostone cobble colluvium. The rhythmite is the primary bone-bearing unit, although smaller amounts of fossils were recovered from the lower diamicton [9].

Although the *Canis dirus* specimen has no precise stratigraphic data associated, it almost certainly derives from the primary bone-bearing middle layer of rhythmite. Other bone material from this layer has produced radiocarbon dates of 11,060 ± 60 years BP (CAMS 10349) and 11,710 ± 220 years BP (PITT-0892).

Guy Wilson Cave, Tennessee (McClung Museum, University of Tennessee; ***MMNHC 0013****;* ***RW001****;* ***DireGWC****; 4th lower premolar root. Picture here https://viewshape.com/shapes/dfomd4trmbn)*

Guy Wilson Cave is a terminal Pleistocene fossil-bearing cave site located in Sullivan County, northeastern Tennessee. At least nine extinct large mammals have been recovered from the site: dire wolf (*Canis dirus*), caribou (*Rangifer tarandus*), flat-headed peccary (*Platygonus compressus*), long-nosed peccary (*Mylohyus* sp.), tapir (*Tapirus* sp.), horse (*Equus* sp.), ground sloth (*Megalonyx* sp.), mammoth (*Mammuthus* sp.), and mastodon (*Mammut americanum*). A sample from the cave was taken by Charles Coney (1970) and the material was donated to the McClung Museum, University of Tennessee, Knoxville [10]. Many bones are carnivore-gnawed, and it has been suggested that dire wolves used the site as a den [11].

A root of a lower right 4th premolar from a dire wolf (MMNHC 0013) was submitted to DirectAMS for radiocarbon dating (D-AMS 26659) and underwent standard gelatin extraction procedures. A portion of the dentine was digested in 0.5 molar HCl for three days at 4 °C with daily changes of acid, producing a strong, collagen pseudomorph. After multiple rinses in deionized H2O, the collagen was soaked in 5 g/L KOH at 4°C to remove organic contaminants and again underwent multiple rinses. It was then completely dissolved to gelatin in 0.05 molar HcL at 90°C, after which it was filtered through a 0.5 micron filter and lyophilized. The dried sample was then combusted in a vacuum and the resulting CO2 isolated and graphitized.This process was conducted twice on this sample in batches four months apart. Graphite targets were measured by an NEC Peletron accelerator mass spectrometer.

Resulting ages on two separate preparations were 10933 ± 44 and 10955 ± 25 years BP, which average to 10944 ± 22 years BP (12820-12720 cal BP) representing the latest reliable date for dire wolves prior to extinction.Two subsamples of the second extraction were submitted to the Washington State University Stable Isotope Laboratory for elemental and stable 13C and 15N analysis (sample G-168765). The extract had average C:N ratio of 3.0955, which is well within the range expected for well-preserved collagen.The ∂13C of -20.08 and ∂15N of 10.28 are indicative of a terrestrial carnivore in a C3 food chain.

American Falls Reservoir, Idaho (Idaho Museum of Natural History; ***IMNH 255/8007; AJ66; DireAFR;*** *petrous from partial cranium)*

Remains from the American Falls Reservoir locality are located nearby to the Gigantobison Bay locality from the same site, having the same geology. Despite multiple attempts, the dire wolf material from American Falls Reservoir has also proven intractable to radiocarbon dating. These remains were found in the same fluvial (river) sediments overlain by lacustrine (lake) sediments deposited following the damming of the Snake River by the volcanic Cedar Butte Basalt, which is dated to 72 ± 14 ky [4]. This sample is from a partial juvenile cranium with some teeth still *in situ* and was collected by Howard Emry in May 1988. We removed the right petrosal from the specimen, which successfully yielded aDNA.

Rancho La Brea Tar Seeps, California (La Brea Tar Pits and Museum; **LACMP23-1619***;* ***DireRLB****; left tibia fragment)*

The Rancho La Brea Tar Seeps (RLB), located in the Los Angeles Basin, California, are the result of asphalt originating from oil sands pushing to the surface and forming seeps that can reach several square meters in area and 9-11 m in depth [12]. The resulting pooled asphalt has led to the entrapment of local fauna over tens of thousands of years. Over two million skeletal elements have been recovered and housed at the George C. Page Museum of La Brea Discoveries in Los Angeles. Of these, over 4,000 dire wolf individuals have been excavated - the most common species within the RLB assemblage [12]. The specimen that yielded type I collagen (*COL1*) for this analysis comes from the La Brea Tar Pits and Museum. It was sourced from RLB locality P23-1, Grid B-1, Level 3.

Geometric Morphometrics

Geometric morphometric data

Geometric morphometrics was carried out on the mandibular first molar morphology of 735 specimens and the mandible morphology of 810 specimens representing 11 species (see Supplementary Table 1). Mandibles were photographed by C.A., A.H-B., A.E. and A.P. using a Nikon reflex camera with 60mm fixed lens. Positioning of the mandible standardized with a spirit level. First mandibular molars were photographed by A.H-B. and A.E.. Mandibles were recorded using 15 fixed landmarks (Supplementary Fig. 1A), while first molars were recorded with 3 fixed landmarks and two curves of sliding semilandmarks; one to capture the anterior outline with 29 equidistant points and the other capturing the posterior of the outline using 19 equidistant points (Supplementary Fig. 1B). TpsDig 2.3 software was used for digitising landmarks [13] and data was imported into the R environment [14] for further processing and analyses.

Geometric morphometric analyses

Shape data was standardized using Generalised Procrustes analysis [15] which scales, translates and rotates the data to minimize the least squared distances among the configuration of landmarks for each specimen [16]. Full Procrustes distance is regarded as the simplest and truest representation of similarities/relationships between shapes. This is because for highly dimensional data with low levels of covariance a plot of 2 PCs rarely presents an accurate picture [15]. At this stage, sliding semilandmarks of the molar shape dataset were slid using the minimum bending energy method [17]. The superimposition and sliding procedures were carried out using the Morpho package [18]. Size, defined as the square root of the sum of squared distances from each landmark to the configuration centroid (centroid size; CS), was extracted from the configuration of landmarks for each specimen. CS was then plotted as boxplots and tested using a pairwise t-test. Mean

shapes were calculated for each species (hereafter group) and full Procrustes distances were calculated among group means. Full Procrustes distances were also calculated among specimens to their respective group means and among all specimens in the dataset. Full Procrustes distances were calculated with the Shapes package [19].

Morphological variance was calculated as the dispersions around the group mean using the full Procrustes distance after Foote [20]. Homogeneity of dispersions between groups was tested pairwise for significant differences following Anderson [21]. All multiple comparison p-values were adjusted using the false discovery rate (FDR) method [22]. Differences among group shapes were tested using a Procrustes ANOVA [16]. Allometry, defined as shape variation that co-varies with size, was assessed using a Procrustes ANCOVA. Homogeneity of allometric slopes among groups was also assessed and if found to be sufficiently parallel allometry was removed from the dataset by regression. These tests and procedures were carried out using the Geomorph package [23].

The following analyses were carried out on both shape and allometry removed shape. A principal component analysis was carried out on the aligned coordinates of the specimens. Linear discriminant analysis (LDA) paired with leave-one-out correct cross validation (CCV) procedures were applied to subsets of principal components (PCs) [24]. The sub setting of PCs was carried out using a stepwise approach after [25] combined with resampling to equal sample size following [24]. These LDAs were applied both on a multiclass basis across the entire dataset, and also on a pairwise basis among groups. In each case of applying the LDA the number of PCs used was determined using the aforementioned methods, with numbers of PCs selected to optimise identification. The cross validation percentage for each LDA was used to assess the accuracy of identification among groups.

To further assess which groups were closest to the dire wolf, the distance from each group mean to all other means was assessed using bootstrapping and plotted as one standard deviation of bootstrapped values around the mean. This was achieved by resampling the specimens of each group by bootstrap procedures, then recalculating the group mean shape and finally calculating the full Procrustes distance among the mean shapes. This procedure was carried out for 999 iterations. Full Procrustes distances of all specimens to the mean of each group were visualized with a violin and boxplot (Fig. 1B) to assess the differences in groups sampling and variance and its effect on the results.

Morphological-molecular comparisons and tests

Size of both datasets (in the form of CS) was tested for phylogenetic signals. Unlike the multivariate shape data, for which there are few tests available, we could test the continuous univariate CS data against multiple evolutionary models following procedures similar to Meloro and Raia [26] and Piras et al. [27]. We compared the CS data to the following models: Brownian [28], Ornstein-Uhlenbeck [29], accelerating-decelerating [30], Pagel's lambda [31], and a white noise model (following [27]). This was carried out using the Geiger package [32]. We chose the model with the lowest Akaike Information Criterion (AIC) as the best fit. We also ran a test of significance for phylogenetic signal in size data using the K-statistic [33] with functions from the package phytools [34].

To assess the mandibular morphology for phylogenetic signals we used a multivariate generalized K-statistic (Kmult, [35]). This method compared the phylogenetic distances constructed from both mtDNA and nuclear DNA with the full Procrustes distance tree calculated among species' mean shapes. This test is a multivariate approach to the method developed by Blomberg et al. [33], which

assesses the data for a phylogenetic signal according to what is expected under a Brownian motion model of evolution. The phylogeny was then mapped to the PCA of mean shapes to visualise the correspondence among shape variables and phylogenetic distances [36]. As the correspondence was particularly poor among some species, and rapid divergence and/or convergence is known from canids (e.g. African wolves and jackals,[37]), and the phylomapping approach to identifying which species did not conform was visually difficult to decipher, it was therefore useful to investigate this further by examining each pairwise comparison.

To identify the species that deviated the most from how the morphological and molecular distances corresponded, the ratio of each respective pairwise distance was calculated. This metric presented a method for assessing molecular-morphological incongruence and was applied just to Canina species with outgroups (Andean fox and gray fox) removed. Raw incongruence scores could be interpreted as follows: lower scores between species represented high morphological similarity relative to a high genetic divergence; in contrast high incongruence scores represented morphologically dissimilar species that had a relatively low genetic divergence. For the purposes of plotting these values for visual assessment a heatmap was used. However, as the incongruence among some species was particularly large the distribution of the incongruence scores was skewed rendering the finer detailed incongruence poorly visualised. Log transforming these values adjusted the extremes of the incongruence scores for ease of visual assessment. Furthermore, as these values are only understandable relative to the dataset being examined we featured scaled them. The resulting visualization of the scores could therefore be interpreted as follows: distance ratios closer to 1 were more morphologically similar than expected given their genetic distance (possible convergence or stabilising selection); distance ratios closer to 0 were more morphologically different than expected (possible rapid divergence).

Centroid Size

Significant differences in centroid size were identified among many groups (FDR adjusted $p<0.05$), but not all (Supplementary Data 3). Dire wolves were the largest species in both datasets, particularly in molar size, but the range, particularly in mandible size, overlapped with wolves substantially (Supplementary Fig. 2).

Morphological variance

Morphological diversity varied among taxa and numerous pairwise differences were found to be significant (at the FDR corrected $p<0.05$ level, see Supplementary Data 4 for pairwise comparisons). This likely reflects the differing levels of ecological plasticity among some species (e.g. wolves are highly plastic and return the highest Procrustes variance scores). However, the variance may also be influenced by sampling, as wolves are extensively sampled across their entire range, whereas some other species, particularly dire wolves, were represented by specimens from only a fraction of their range. Such variation in morphological diversity should be investigated further in future studies for other species, but here we focus on dire wolves. Dire wolves were less morphologically diverse than gray wolves in both mandibular morphology and first mandibular molar morphology, but only mandible morphology was significantly different ($p<0.01$, see Supplementary Data 4). However, this could reflect the extensive spatial sampling of gray wolves (their entire range) compared with the restricted sampling of dire wolves (primarily La Brea).

Group differences and allometry

All Procrustes ANOVA pairwise comparisons were found to be significant (at the FDR corrected $p<0.05$ level see Supplementary Data 5 for pairwise comparisons). Of particular note, dire wolves and gray wolves were different, but as their variance differed and gray wolves occupied a greater area of morphospace in both mandible and first molar datasets, it is likely that this characteristic resulted in different central tendencies. CCV percentages varied across both datasets (Supplementary Data 6 & 7), but all were high (87.5–100%) and well above the CCV percentage achievable by random chance alone; this indicates that all species are morphologically different and distinguishable to a high degree of accuracy. Dire wolves and gray wolves could be identified between with a high level of confidence (90–97% CCV).

Allometry was found to be significant in both the mandibular and first molar datasets ($p<0.01$, Supplementary Fig. 3). Allometric slopes were found to be parallel (at the $p<0.01$ level) in both datasets and could therefore be removed. Removal of allometry increased the already high ability to identify between dire wolves and gray wolves in the first molar dataset (from ~91% to 100% Supplementary Data 8), but decreased the identification accuracy in the mandibular dataset (from 100% to ~98% Supplementary Data 9). The change in mandibular results is relatively small considering the already high level of accuracy. These allometric dental differences may indicate that large tooth size has been selected for in dire wolves, which may correspond with their specialization in consumption of megafauna and hard foods.

Dire wolf mean shape was closest to both gray wolf mean tooth shape and mandible shape; gray wolf mean shape was found to be the second nearest neighbor in both datasets (Fig. 1B & Supplementary Fig. 4). Dholes are also very close to dire wolf mandibular mean shape, but the sample size for dhole is also extremely small and the variability of shape is lower. When allometry is removed the allometry corrected dire wolf mandible mean shape was found to be closest to the black-backed jackal, while the first molar mean shape was closest to the African hunting dog. It was possible to calculate procrustes distances among both combined mean shapes (i.e. the distance between each species based on the combination of superimposed mandible and first molar landmark configurations), but because of the preservation of the dire wolf specimens the mean shapes for each morphological dataset was constructed from different sets of individuals and as a result it was not possible to calculate the combined shape distance of individuals to species means. The results from the combined shape and allometry corrected shape, identify wolves and black-backed jackals respectively, as the closest to dire wolf combined mean shape.

Phylogenetic signal and morphology

A brownian motion model best fit the phylogenetic signal in the CS data, however it was found to be non-significant in both datasets (mandible: $K=0.8109$, $p=0.092$; first molar: $K=0.8078$, $p=0.095$). Almost all of the results of phylogenetic signal testing in the morphological datasets were non-significant ($p>0.05$) except when allometry was removed from the mandibular dataset and the combined mandibular-molar dataset, which produced significant results (Kmult 0.4082, FDR adjusted $p=0.0168$ and Kmult 0.54247, FDR adjusted $p=0.0168$ respectively).

To assess the phylogenetic signal more thoroughly we employed a leave one out approach for each species to assess whether the fit of the brownian motion model improved when a species was removed from the analysis. To do so, we removed each species iteratively and assessed the change in phylogenetic signal with the removal of that species. We found that the phylogenetic signal in full shape only significantly improved with the removal of the Ethiopian wolf (Kmult 0.4463, uncorrected $p=0.0384$). In allometry corrected shape removing the gray fox dramatically changed the phylogenetic signal resulting in a non-significant result (Kmult 0.7181, uncorrected

p=0.0863); this likely indicates that improvement in phylogenetic signal in allometry corrected shape is not evenly distributed across all species and in fact the significant phylogenetic signal is almost entirely driven by the inclusion of the the gray fox in the allometry free analysis. This could indicate that selection pressures on size are strong and the tempo of this selective pressure variable, which could then obscure the phylogenetic signal in morphology. Such a scenario would not be surprising given the element being examined (mandibles) and its role in feeding behavior, which is highly variable among these species. None of the phylomapping exercises were clear and the majority had extensive overlap of internal phylogeny nodes (Supplementary Fig. 5). Furthermore, the phylomapping approach was not able to indicate which pairs of species deviated from the phylogenetic signal the most from visual inspection alone.

To investigate this further we examined the incongruence between the molecular distances compared with the morphological distances. These incongruences are the result of multiple evolutionary mechanisms and as such a low score can be the result of either stabilising selection or convergence, whereas a high score may be the result of rapid divergence. Therefore we do not attempt to use these scores as any form of test or to determine the process, rather we use them as a descriptive tool for examining the dataset (Supplementary Data 10-12). The morpho-molecular incongruence score revealed that when raw shape was analyzed the molar and combined datasets for the dire wolves and gray wolves had the lowest incongruence score, with high morphological similarity, despite deep genetic divergence (Supplementary Data10 & 12). In the mandibular dataset dire wolves *vs.* dhole had the lowest incongruence score and wolves *vs.* dire wolves were the second lowest (Supplementary Data 10, Supplementary Fig. 4 & 6). On the other side of the incongruence spectrum, where species were more morphologically dissimilar relative to their depth of genetic divergence and the resulting incongruence scores were high, coyotes v gray wolves were found to be the most incongruent in the mandible and combined datasets (Supplementary Data 10 & 12), while in the molar dataset African wolves v gray wolves had the highest incongruence scores (Supplementary Data 10). When allometry was removed gray wolves v black-backed jackals consistently had the lowest incongruence across all datasets (Supplementary Data 11-12). Gray wolves v coyotes had the highest incongruence scores in the allometry corrected mandible dataset (Supplementary Data 11). Coyotes v African wolves produced the highest incongruence scores in the molar and combined allometry corrected datasets (Supplementary Data 10&12). Incongruence scores between gray wolves and dire wolves in the allometry corrected datasets continue to be low, particularly for mandible morphology (Supplementary Data 11), suggesting that the increase in phylogenetic signal found in these datasets were more likely the result of improved correspondence of genetic and morphological distances among other species (e.g. it is likely the result of resolving the morphological relationship of the gray fox as identified by the leave one out phylogenetic signal analysis).

Proteomics

Collagen extraction and MS/MS

The *Canis dirus* (LACMP23-1619) purified collagen sample (extracted using the method stated in Fuller et al. [38], was prepared for proteomic analysis following a slightly modified version of the ZooMS protocol outlined in Welker et al. [39]. The collagen was removed from the Eppendorf and stored in the freezer, and 100 µl 50 mM Ambic was added to the empty Eppendorf and heated at 65 ^{o}C for 1 hour. This was followed by digestion overnight at 37 ^{o}C; 50 µl of the heated sample was digested using 1 µl of 0.5 µg/µl porcine trypsin in trypsin resuspension buffer (Promega, UK)

and the other 50 µl was dried down and resuspended in 50 µl 100 mM Tris solution to be digested with elastase (Worthington; USA) at the same concentration in 10% Tris solution. Two different enzymes were used to increase the protein sequence coverage for LC-MS/MS [40,41]. Digestion was stopped by the addition of trifluoroacetic acid (TFA) at a concentration of 0.5-1% of the total solution. Peptides were desalted using zip-tips [42] and eluted in 100 µl of 50% acetonitrile (ACN)/0.1% TFA (v/v).

The extracted peptides were analysed at the Discovery Proteomic Facility at Oxford University. The sample was analyzed on a Q-Exactive employing an Easyspray column (ES803, 500mmx75µm, Thermo) and a gradient of 2%-35% ACN in 0.1% FA/5%DMSO over 60 minutes. The MS1 resolution was set to 70,000 with an AGC target of 3E6. MS2 spectra for up to 15 precursors were acquired with a resolution of 17,500 and an AGC target of 1E5 for up to 128ms and 28% normalized collision energy (higher-energy collision dissociation). The precursors were excluded for 27 seconds from re-selection.

The LC-MS/MS raw files were converted to MGF files using Proteowizard [43] and searched against a mammal collagen database which included common contaminants (http://www.thegpm.org/crap/) in PEAKS v7.5 [44,45]. Mass tolerances were set at 0.5Da for the fragment ions and 10ppm for precursor ions and up to 3 missed cleavages were permitted. Searches allowed various post-translational modifications (PTMs) including oxidation (MHW) and hydroxylation of proline (both +15.99), deamidation (NQ; +0.98) and pyro-glu from E (-18.01) and a fixed PTM of carbamidomethylation (+57.02) which occurs due to sample preparation. A maximum of 3 PTMs was allowed per peptide. Protein tolerances were set at 0.5% false discovery rate (FDR), >50% average local confidence (ALC; de novo only) and -10lgP score ≥ 20.

Sequences of both COL1A1 and COL1A2 were concatenated using previously published mammal collagen consensus sequences. Telopeptides very rarely survive in fossil samples and so these were removed from all sequences. Isoleucine and leucine cannot be differentiated using low energy tandem mass spectrometry and *de novo* sequencing as both amino acids are isobaric. Therefore, the identification of leucine/isoleucine was consistent with the consensus sequences. Once a potential collagen sequence was compiled for *Canis dirus*, the sequence was added to the collagen database and the sample was re-analysed using PEAKS to check for coverage and sequence substitutions. Any differences noted in the consensus were inspected manually. In order for a difference to be considered authentic, it had to occur in more than 1 product ion spectrum and be covered by both b and y ions.

Phylogenetic analysis

Here we built a phylogenetic tree using the COL1A1 and COL1A2 sequences generated by MS/MS analysis from a dire wolf as well as representatives of Carnivora species. We first obtained the amino acid sequences of COL1A1 and COL1A2 from a redfox (*Vulpes vulpes*; XP_025851655, XP_025859557), a dog (*Canis familiaris*; NP_001003090, NP_001003187), a dingo (*Canis familiaris dingo*; XP_025295726, XP_025327115), a grizzly bear (*Ursus arctos*; XP_026368913, XP_026361636), a northern seal (*Callorhinus ursinus*; XP_025715155, XP_025728689) and a cat (*Felis catus*; XP_003996748, XP_003982813) from GenBank. We also obtained amino acid sequences from the genomes of all modern species used in this study (gray fox, andean fox, black-backed jackal, side-striped jackal, african wild dog, dhole, ethiopian wolf, African wolf, Eurasian gray wolf, and Yellowstone gray wolf). To do so, we downloaded the GFF

file from Ensembl (v95) for canFam3.1. We then extracted and translated the coding sequence of COL1A1 and COL1A2 using gffread v0.10.5.

The amino acid sequence was then concatenated for each species and then aligned using MAFFT v7.123b [46]. We then masked all leucine (L) and isoleucine (I) from the alignment as these are isobaric and low-energy MS/MS sequencing is not capable of discriminating between them. A phylogenetic tree was generated using MrBayes version 3.2.1 [47] with the amino acid model estimated from the data (prset aamodellpr = mixed). We ran two runs of four chains each with 1,000,000 generations. Convergence was assessed by ensuring all ESS for all parameters were higher than 100. The resulting phylogeny indicates that gray wolves, African wolves, and coyotes, as well as dogs, all form a monophyletic clade (posterior >0.7) that excludes the dire wolf (Supplementary Fig. 7).

Genomics

Ancient DNA

Five samples (DireNTC, DireGB, DireSP, DireGWC), out of 46 specimens, possessed enough endogenous DNA for deeper sequencing (Supplementary Data 1). Here we describe the protocol used by the six laboratories which attempted DNA extraction and sequence from dire wolves remains (Oxford, Copenhaguen, UCSC, UCLA, ASU and Adelaide). The samples which were successful (selected for deep sequencing) and unsuccessful are mentioned in the title of each section. Additional metadata for these samples can be found in Supplementary Data 1 and site description for the five samples that were selected for deep sequencing can be found in the section above.

University of Oxford (PalaeoBarn)

Successful sample(s): DireAFR; unsuccessful sample(s): JH167-JH180/AJ62-AJ56.

DNA extraction

DNA was extracted from tooth or bone samples (see Supplementary Data 1) in a dedicated ancient DNA laboratory using appropriate sterile techniques and equipment. Extraction was carried out following the Dabney extraction protocol [48] but with the addition of a 30 minutes pre-digest stage [49].

Library preparation

Illumina libraries were built following [50], with the addition of a six base-pair barcode added to the IS1_adapter.P5 adapter. The libraries were then amplified on an Applied Biosystems StepOnePlus Real-Time PCR system to check that library building was successful, and to determine the optimum number of cycles to use during the indexing amplification PCR reaction. A six base-pair barcode was used during the indexing amplification reaction resulting in each library being double-barcoded with an “internal adapter” directly adjacent to the ancient DNA strand and which would form the first bases sequenced, and an external barcode that would be sequenced during Illumina barcode sequencing.

Capture

The lysates were sent to Arbor Biosciences. They used myBaits® Whole Genome Enrichment (WGE) (37) procedure following myBaits® manual version 3.0 using a probe set generated from a supplied *C. lupus* genomic DNA.The baits can be produced from any pure eukaryotic or prokaryotic

genomic DNA source, for example modern gray wolf (for capture of DireGWC and DireAFR). Arbor Biosciences' unique process converts this gDNA into a pool of biotinylated RNA baits. These baits can then enrich corresponding molecules from a user-supplied NGS library, for example from an ancient or extinct species such as the dire wolf, via the process of myBaits in-solution hybridization capture. This enriched library allows for orders-of-magnitude more efficient NGS than would be possible using the full, non-enriched library.The supplied DNA extracts were used to build Illumina TruSeq-style libraries, using blunt-end adapter ligation and a uracil non-stalling polymerase for 8 to 12 cycles of indexing amplification. Extracts of sample AJ66 were sonicated prior to library preparation, and size-selected to retain all fragments < 500nt (Fraction I-S). Fragments >500nt consequent this sonication were further sonicated to <300nt and converted to libraries separately (Fraction I-L). Arbor then employed the gray wolf WGE bait set to enrich between 270 and 700ng available library of Fraction I-S in two ~40hour rounds of hybridization capture. The first round used 55C for the hybridization and wash temperatures, and the entirety of the enriched product was amplified 8 cycles before being taken to a second round that used 60C temperatures. After the second round, half of each enriched library was amplified between 8 and 12 cycles and then sequenced. The libraries from Fraction I-L were sequenced without enrichment.

Sequencing

Libraries were screened on an Illumina HiSeq 2500 (Single End 80bp) sequencer at the Danish National High-Throughput Sequencing Centre and on a Illumina MiniSeq at the AMIS laboratory in Toulouse. Based on this data we selected **DireAFR** for deeper sequencing because of its preservation. Deeper sequencing was conducted on Illumina HiSeq 4000 (Paired End 150bp) at Novogene (Novogene Corporation Inc CA 91914, USA).

University of California Los Angeles (UCLA)

Successful sample(s): DireGWC; unsuccessful sample(s): None.

DNA extraction

DNA from sample DireGWC (Supplementary Data 1) was extracted from a lower right 4th premolar in a designated ultra-clean facility at UCLA using the appropriate sterile techniques and equipment in keeping with standard aDNA practice. Extraction was carried out following the protocol of [51] based on a silica-column based protocol.

Capture

The lysates were sent to Arbor Biosciences. They used myBaits® Whole Genome Enrichment (WGE) (37) procedure following myBaits® manual version 3.0 using a probe set generated from a supplied *C. lupus* genomic DNA. They used the DNA extracts to build Illumina TruSeq-style libraries, using blunt-end adapter ligation and a uracil non-stalling polymerase for 8 to 12 cycles of indexing amplification. Both extracts were taken to library prep without any treatment.

Library preparation

Arbor Biosciences employed the gray wolf WGE bait set to enrich between 270 and 700ng available library of the sample in two ~40hour rounds of hybridization capture. The first round used 55C for the hybridization and wash temperatures, and the entirety of the enriched product was amplified 8 cycles before being taken to a second round that used 60C temperatures. After the second round, half of each enriched library was amplified between 8 and 12 cycles.

Sequencing

Libraries were sent for sequencing on a HiSeq4000 (Paired-end 100bp) at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley.

Australian Centre for Ancient DNA, University of Adelaide

Successful sample(s): DireNTC, DireGB and, DireSP; unsuccessful sample(s): R46006, R30312, R13470, R21446, and P232517.

DNA extraction for successful sample(s): DireNTC, DireGB and, DireSP

Sample IMNH48001/52 (DireGB; Supplementary Data 1) was subjected to a silica-based DNA extraction method [52]. This protocol included decalcification in 4 mL 0.5 M EDTA for 1 hour at 37 °C under constant rotation, after which the EDTA was removed and replaced with 4 ml of fresh 0.5 M EDTA and incubated overnight under constant rotation at 37 °C. A final incubation with an additional 60 µl of Proteinase K (20 mg/mL) was then performed for 2 hours at 55°C, following which the supernatant (digestion buffer) was mixed with a modified QG buffer (15.5 mL QG buffer [Qiagen], 1.3% Triton X-100 [Sigma-Aldrich], 25 mM NaCl [Sigma-Aldrich], 0.17 M Sodium Acetate [Sigma-Aldrich]) and bound to silicon dioxide particles, which were then washed with 80% ethanol. Bound DNA was eluted in 200 uL of TE buffer (10 mM Tris, 1 mM EDTA, pH 8).

Sample KU48130 (DireNTC; Supplementary Data 1) was subjected to an alternative DNA extraction protocol optimised for recovery of short DNA fragments [48]. Bone samples were decalcified in 1 mL 0.5 M EDTA for 1 hour at 37 °C under constant rotation, after which the EDTA was removed and replaced with 980 µl of fresh 0.5 M EDTA and 20 uL of Proteinase K (20mg/mL) then incubated overnight under constant rotation at 55 °C. The digestion buffer was mixed with 13 mL of a modified PB buffer (12.6 mL PB buffer [Qiagen], 6.5 µL Tween-20, and 390 µL of 3M Sodium Acetate) and bound to silicon dioxide particles, which were then washed with 80% ethanol. Bound DNA was eluted in 100 µL of TE buffer.

Sample VP1737 (DireSP; Supplementary Data 1) was extracted following a third protocol, beginning with decalcification under rotation overnight in 10 mL of 0.5 M EDTA at room temperature. The decalcified material was then digested under rotation overnight in 3 mL of 100 mM Tris-HCl, 100 mM NaCl, 0.5 mg/mL proteinase K, 10 mg/mL dithiothreitol (DTT), and 1% sodium dodecyl sulphate (SDS) at 55 °C. Following digestion, 3 mL of Tris-saturated phenol was added and mixed under rotation for 10 min at room temperature, followed by centrifugation at 1500 g for 5 min. The aqueous phase was then transferred to a new tube. This process was repeated twice: once with 3 mL of Tris-saturated phenol, and once with 3 mL of chloroform. The final aqueous phase was de-salted with sequential additions of DNA-free water to an Amicon Ultra-4 Centrifugal Filter Unit (Millipore) and concentrated to a final volume of 200 µL.

DNA extraction for unsuccessful sample(s): R46006, R30312, R13470, R21446, and P232517.

Samples R46006, R30312, and R13470 were extracted following the same protocol as described above for DireGB. Sample R21446 and two subsamples of P2325171 were extracted following the same protocol as described above for DireNTC.

Library preparation for successful sample(s): DireNTC, DireGB and, DireSP

We used Gansauge *et al.* [53]'s single-stranded protocol to create Illumina sequencing libraries from the DNA extracts of IMNH48001/52, KU48130, and VP1737. After library preparation, we performed a real-time PCR assay to determine how many cycles of PCR were required to optimise

library quantity and complexity [54]. Duplicate real-time PCR assays were performed for each library in a final volume of 10 µL, each comprising: 1 µL of a 1:5 dilution of library, 1 x Platinum Taq DNA Polymerase High Fidelity buffer (ThermoFisher Scientific), 2 mM MgSO4 (ThermoFisher Scientific), 0.25 mM of each dNTP (ThermoFisher Scientific), 0.4 µM each of a P5 and P7 indexing primer (designed based on [50]), 0.004 x ROX (Life Tech), 0.2 x SYBR (Life Tech), 0.56 M DMSO (Sigma-Aldrich), and 0.2 U of Platinum Taq DNA Polymerase High Fidelity (ThermoFisher Scientific), in laboratory grade water. Real-time PCRs were performed on a LightCycler 96 (Roche) with the following cycling conditions: 94 °C for 6 min; 40 cycles of 94 °C for 30 s, 60 °C for 30 s, 68 °C for 40 s; followed by a high-resolution melt. The libraries were then amplified using conventional PCR. In order to maintain library complexity and minimise PCR bias, each library was amplified in eight separate 25 µL reactions, each comprising: 3 µL of undiluted library, 1 x Platinum Taq DNA Polymerase High Fidelity buffer (ThermoFisher Scientific), 2 mM MgSO4 (ThermoFisher Scientific), 0.25 mM of each dNTP (ThermoFisher Scientific), 0.4 µM each of a P5 and P7 indexing primer, and 0.2 U of Platinum Taq DNA Polymerase High Fidelity (ThermoFisher Scientific), in laboratory grade water. Cycling conditions for the PCR were as follows: 94 °C for 6 min; a number of cycles of 94 °C for 30 s, 60 °C for 30 s, 68 °C for 40 s as determined using rtPCR (8 for KU48130, 9 for VP1737, 9 for IMNH48001/52); and 68 °C for 10 min. PCR products from each library were pooled and purified using AMPure (Agencourt), and resuspended in 30 µL of buffer comprising 10 mM Tris, 0.1 mM EDTA, and 0.05% Tween-20.

Library preparation for unsuccessful sample(s): R46006, R30312, R13470, R21446, and P232517.

Extracted DNA was enzymatically repaired and blunt-ended, and had custom adapters ligated following the protocol of [50]. Adapter sequences featured unique barcodes in order to allow identification and exclusion of any downstream contamination. Libraries were subjected to a short round of PCR in order to increase the total quantity of DNA using primers complementary to the adapter sequences. Cycle number was kept low (exact number determined by rtPCR) and the template was split into eight separate PCRs per library in order to minimise PCR bias and maintain library complexity. Each individual PCR (25 µL) contained 1× PCR buffer, 2.5 mM MgCl2, 1 mM dNTPs, 0.5 mM each primer, 0.1 U AmpliTaq Gold and 2 µL DNA. Cycling conditions were as follows: 94 °C for 12 min; 12-13 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 40 s (plus 2 s/cycle); and 72 °C for 10 min. PCR products were purified using AMPure magnetic beads (Agencourt). Commercially synthesised biotinylated 80-mer RNA baits (MYcroarray, MI, USA) were used to enrich the libraries for placental mammal mitochondrial DNA.

Sequencing for successful sample(s): DireNTC, DireGB and, DireSP

The libraries from DireGB and DireSP were diluted to 1.5 nM and each was run on one lane of an Illumina HiSeq X Ten using 2 x 150 bp PE (300 cycle) chemistry. The library from DireNTC was diluted to 2 nM and ran on an Illumina NextSeq flow cell using the 2 x 75 PE (150 cycle) High Output chemistry. Due to the modified structure of the adapters used in (Gansuage *et al.*'s) [53] library protocol, these sequencing runs used a custom R1 sequencing primer (CL72; see [55]) instead of the default Illumina primer included in the kits.

Sequencing for unsuccessful sample(s): R46006, R30312, R13470, R21446, and P232517.

The enriched libraries were sequenced on Illumina high-throughput sequencing platforms. Only 21 sequencing reads in total (among all six libraries) could be mapped against the gray wolf mitochondrial reference, all of which could be excluded as common lab contaminants (e.g. *Homo*

sapiens, *Bos taurus*) mapping to conserved gene regions. Consequently, these samples were excluded from further analysis.

University of Copenhagen

Successful sample(s): None; unsuccessful sample(s): Bt2, Bb2, Bt3C, and Bb3C.

DNA extraction

Teeth and bone were subsampled from a lower jawbone from collections in UCLA (Bt1, Bb1, Bt2, Bb2, Bt3C, Bb3C, see Supplementary Data 1), originally found in the La Brea tar pits. Bone and teeth powder was digested in a EDTA, urea and proteinase K buffer as in [56] and purified as in [48] using a modified binding buffer as in [57]. Out of a total of 6 individual extractions (1 ml), 3 teeth and 3 bone sub-samples, 2 bone and 2 teeth were purified as above and the other bone and tooth involving an additional Phenol Chloroform purification step as in before purification as in [58].

Library preparation

Double stranded Illumina libraries were made on a tooth and a bone (without Phenol Chloroform treatment) extract using the commercial NEBNext DNA Sample Prep Master Mix Set 2 (E6070, New England Biolabs Inc., Beverly, Massachusetts, USA) and using MinElute (Qiagen) purifications in between reactions with 5x PB buffer [58]. Single stranded Illumina libraries were made on the remaining 4 purified extracts (including 2 Phenol Chloroform treated extracts) strictly according to [55]. qPCR was performed using IS7/8 primers [50] and Lightcycler 480 reagents using 1 uL 10x diluted library on an Mx3005 instrument (Agilent). This was done to check for successful library preparation and estimate the number of cycles for indexing amplification. Index PCR was performed in 100 uL reactions using Platinum Taq polymerase (Invitrogen) with 6-base indexed full length P5 primers and common non-indexed full length P7 primer. Samples were given 10-25 cycles in PCR and subsequently purified using MinElute columns (Qiagen). Quantity and quality of libraries was measured using a Qubit 2.0 fluorometer and a 2100 Bioanalyzer (Agilent).

Sequencing

While double strand libraries failed in amplification, single-strand libraries were screened using an Illumina HiSeq 2500 (SR 80 mode) platform at the Danish National High-Throughput Sequencing Centre in Copenhagen. After mapping, all reads were found to be non-mammalian and hence non-endogenous for the sample.

University of California Santa Cruz (UCSC)

Successful sample(s): None; unsuccessful sample(s): JK376.

DNA Extractions

DNA from sample JK376 (Supplementary Data 1) was extracted and libraries were built in a dedicated aDNA lab at the University of California, Santa Cruz. The coprolite was first crushed to expose undigested bone and then the sample was divided into two separate tubes for extraction, one with bone and one with all other material. The bone fraction was powdered using a Retsch MM 400 ball mill. DNA extraction was performed following the method outlined in Dabney et al. (2013)[48] on 120mg of bone fraction and 250mg of non-bone material was extracted with MoBio's Powerlyzer kit.

Library Preparation

We prepared Illumina libraries from both extractions using the double stranded DNA library preparation protocol outlined in Meyer and Kircher using 20uL of extract [50]. Libraries were amplified for 25 cycles using Amplitaq Gold hot start polymerase (2U Polymerase, 0.2uM each primer, 0.25mM each dNTP, 2.5mM MgCl2, 1X Amplitaq Buffer) and were SPRI purified using a factor of 1.75x.

Capture

Target enrichment was performed on the post amplified library from the non-bone fraction using Arbor Biosciences myBaits® Custom. The RNA probe set includes 75 mammalian mitochondria and was designed after [59].

Sequencing

Shotgun and enriched libraries were sequenced separately on an Illumina MiSeq (2x75bp) at the University of California, Santa Cruz.

Arizona State University (ASU)

Successful sample(s): None; unsuccessful sample(s): DW01-DW19.

DNA extraction

DNA was extracted from teeth and bones (samples DW01-DW19, Supplementary Data 1) in a dedicated ancient DNA laboratory at Arizona State University using previously established protocols [60].

Library preparation

Libraries were constructed by following a modified version of a protocol created by Meyer and Kircher (2010). Double stranded libraries were constructed from DNA extracts and amplified in duplicate using AmpltiTaq Gold Polymerase (Applied Biosystems). Samples were purified using a MinElute PCR Purification kit (Qiagen) and quantified using an Agilent 2100 Bioanalyzer (Agilent).

Capture

Mitochondrial DNA was captured using previously established protocols. Modern DNA was extracted from cheek swabs of two dogs using a standard phenol-chloroform extraction [61]. The mitochondrial genome was amplified in two separate PCRs using primers and methods from previously published sources [62]. In brief, the PCR products were pooled into equimolar amounts for primer sets A and B, which were constructed based on previous work on the mitochondrial genome of *Canis familiaris* [62,63]. The bait DNA was then sheared to 200-300bp using a Covaris sonicator and attached to beads that were thoroughly washed [64].

Ancient samples were captured using these modern baits in accordance with previously published protocols [65]. In summary, a hybridization mixture was created which included the indexed libraries, an Agilent blocking agent, and an Agilent hybridization buffer among other reagents. The beads were then incubated in this mixture for two nights before being washed, removing everything that was not bound to the beads, and eluted. This enriched library elution was purified using a MinElute column and quantified using a KAPA Library Quantification Kit by Illumina Platforms. Based on these results, some samples were amplified or diluted to ensure they reached a concentration of at least 4 nM, as requested for sequencing.

Sequencing

Amplified libraries were pooled in equimolar amounts and sequenced on an Illumina MiSeq (paired-end 150bp).

Modern DNA

Black-backed jackal and gray wolf from Montana - UCLA

DNA from a black-backed jackal (*C. mesomelas*) (Supplementary Data 13) was extracted from a blood sample collected in 1987 (Soysambu, Eastern Africa) using the Qiagen DNeasy Blood and Tissue Kit. The DNA extract was sent to ArborSciences for library preparation. The DNA was sonicated and size-selected to 300nt modal fragment lengths and converted to two identical Illumina TruSeq-style libraries using standard a-tail chemistry and 6 indexing cycles, and subsequently sequenced at Novogene (Novogene Corporation Inc CA 91914, USA) on an Illumina Hiseq4000 (150bp paired-end). DNA from a captive gray wolf (*Canis lupus*) from the wolf haven in Montana was extracted using the same method and sequenced on a NovaSeq S4 at the California Institute for Quantitative Biosciences (QB3).

Side-striped jackal - Copenhagen

A ~37.5x nuclear genome (Supplementary Data 13) was sequenced from a tissue sample (T-1252) from a Side-striped jackal (*C. adustus*) collected in 2002 from Guinea, (mtDNA-sequenced was previously sequenced in [66]). The DNA was extracted using a Kingfisher Duo extraction robot, using Cell and Tissue DNA Kit from ThermoFisher Scientific using manufacturer's protocol (ThermoFisher Scientific, Waltham, MA). The DNA was fragmented into 400-600bp molecules using a Bioruptor NGS device (Diagenode, Liège, Belgium). A next generation sequencing library was made using the commercial NEBNext DNA Sample Prep Master Mix Set 2 (E6070, New England Biolabs Inc., Beverly, Massachusetts, USA) in combination with BGIseq adaptors [67], and was sequenced on a lane of PE100 on a BGIseq500 platform by BGI-Europe.

Data processing

Modern data

Modern samples sequenced in this study as well as those downloaded from public repositories (Supplementary Data 13) were aligned to canFam3.1 (dog reference genome) and VulVul2.2 (red fox assembly; accession: *GCA_001887905.1)* using BWA mem [68], with a realignment step as implemented in GATK [69]. For each sample, we then computed depth of coverage using the following command in ANGSD [70]:

angsd -doCounts 1 -i input.bam -doDepth 1 -out sample_cov -minQ 0 -minMapQ 0

The output was used to build a cumulative distribution of per sample depth of coverage. All regions, within each sample, that fell within the 5% highest and lowest coverage were excluded from further analysis. This procedure ensured that no abnormally covered (e.g. repetitive regions or copy number variation) regions were included in the analysis. The likelihood of each possible genotype, in single sample for every base of the reference genome (excluding those in abnormal coverage regions) was then computed using the GATK genotype likelihood function as implemented in ANGSD:

angsd -GL 2 -out file_name -doCounts 1 -setMinDepth min -setMaxDepth max -i input_bam_file -doGlf 4

At each base the genotype was encoded as missing (N) unless the likelihood of the highest genotype was ten fold higher than the next best possible genotype. To accommodate the difference in coverage among modern and ancient samples in our D-statistics analysis (see below), we also called genotypes by randomly sampling a single read of 20 base pair minimum and with a mapping quality (MAQ) and base quality (BQ) of at least 30 at each covered position in the genome [71–73].

Additional modern genomes for D-statistics

To explore the potential of admixture between dire wolf and North American canids, we obtained 24 additional genomes from North American gray wolves (16), coyotes (2), a red wolf (1), a Great Lakes wolf (1), an ancient Eurasian wolf (1), and three high coverage ancient dogs including one from pre-Columbian America (Supplementary Data 13). To accommodate the difference in coverage among modern and ancient samples in our D-statistics analysis (see below), we also genotypes in all genomes by randomly sampling a single read of 20 base pair minimum and with a mapping quality (MAQ) and base quality (BQ) of at least 30 in each genome at sites that were ascertained as transversion (see below) [71–73].

Ancient data

Raw reads were filtered allowing one mismatch to the indices used in library preparation. Adapter sequences were removed using AdapterRemoval [74]. Reads were aligned using Burrows-Wheeler Aligner (BWA) version 0.7.17 [68] to canFam3.1 (dog reference genome) and VulVul2.2 (red fox assembly; accession: *GCA_001887905.1)* with the following parameters ("--l 1024,-n 0.01, -o 2") [75]. FilterUniqueSAMCons [76] was then used to remove duplicates. BAM files from different sequencing lanes were merged using the MergeSamFiles tool from Picard v1.137 (http://broadinstitute.github.io/picard/). To accommodate the low coverage of the nuclear genome of the dire wolf samples, genotypes were called by randomly sampling a single read of 20 base pair minimum and with a mapping quality (MAQ) and base quality (BQ) of at least 30 at each covered position in the genome, excluding bases within 5bp of the start and end of a read [71–73]. Molecular damage was assessed using MapDamage2.0 using default parameters [77] (Supplementary Fig. 8-9). The damage plots from DireAFR and DireGWC are expected when building double-stranded libraries from ancient DNA, as C to T misincorporations are observed on both strands (with damage on the complementary strand manifesting as G to A misincorporations). In contrast, single-stranded libraries [55] were created from the remaining dire wolf samples, meaning that damage from the complementary strand is not observed, resulting in only C to T misincorporations.

Ascertainment

Specific analyses, such as the supermatrix phylogeny based on SNPs (using ascertainment correction implemented in RAxML) as well as for the D-statistics, necessitate a list of pre-defined SNPs. Here, we used the genome consensus (see genotype calling section above) of all high coverage modern genomes (see column used-for-ascertainment in Supplementary Data 13) to obtain a list of SNPs for these analyses. All variable positions, with a minor allele found at least in two high coverage modern genomes (as heterozygous or homozygous) were kept for further analyses. On canFam3.1, this resulted in ~46M SNPs, ~13M of which were transversions and ~38M SNPs on VulVul2.2, ~11M of which were transversions.

Phylogenetic analyses

Mitochondrial genomes

We used htsbox (https://github.com/lh3/htsbox) to generate mitochondrial majority consensus sequences from the bam files using BQ>=30 and MAPQ>=30 while excluding bases within 5bp of the start and end of a read to limit the incorporation of deamination in the analysis. To reconstruct the mitochondrial phylogeny, we retrieved the modern mitochondrial genomes from several extant canids from the BAM files aligned to the dog reference genome, including the genome of the gray wolf (*C. lupus*), the North American endemic coyote (*C. latrans)*, the African wolf (*C. anthus*), Ethiopian wolf (*C. simensis*), and wild dog (*Lycaon pictus*); and the dhole (*Cuon alpinus*), the black-backed (*C. mesomelas*) and the side-striped African jackals (*C. adustus*) (Supplementary Data 13). Additional mitochondrial genomes were obtained from NCBI for the Arabian wolf (*C. lupus arabs*), the dog (*C. lupus familiaris*), the red wolf (*C. rufus*) and the Great Lakes wolf (*C. lupus lycaon*) were added as well in the analyses. The gray fox (*Urocyon cinereoargenteus*) and Andean fox (*Lycalopex culpaeus*) were used as outgroups. These were aligned together using MAFFT. Accession numbers are listed in Supplementary Table 2.

Partitioning of the mitochondrial genome into 13-protein coding and two rRNA genes (12S,16S) was performed using AMAS [78]. For this analysis, we removed any genes which were not covered in at least one dire wolf (Supplementary Table 3). PartitionFinder2 [79] implemented in the Cipres web server [80] was subsequently used to determine the optimal substitution model for the 9 remaining genes (Supplementary Table 4). Bayesian phylogeny reconstructions were done in MrBayes 3.2.6 [47] as implemented in the Cipres web server [80]. Markov chain Monte Carlo (MCMC) sampling was performed with 4 chains run for 5 x 10^6 generations with one tree sampled every 1000 generations. A strict majority rule consensus tree was built from the combined chains, excluding the 25% first iterations as burn-in.

The amount of missing data ranged between 49% to 85% for the dire wolf specimens (see Supplementary Table 3). DireAFR and DireNTC have the highest amount of missing data with 79% and 85% respectively. To address potential bias introduced by the missing data [81,82], we performed several tree analyses. Firstly with the data covered in 9 genes (10587bp) in three dire wolves (Supplementary Table 3), excluding the two dire wolves with the most missing data (DireAFR and DireNTC). The phylogeny recovered was similar to those recently obtained by [83] and [84] (Supplementary Fig. 10 A), including a monophyletic clade that contains Eurasian and American wolves (*C. lupus*), the dog (*C. familiaris*) and the African wolf (*C. anthus*). The Great Lakes wolf (*C. lupus lycaon*) clusters with the coyote (*C. latrans*) and the red wolf (*C. rufus*), most likely due to admixture events between gray wolves and coyotes [85], [86]. While we obtained relatively lower for the node (posterior probability ~0.5) leading to the Ethiopian wolf (*C.simensis*) and the golden jackal (*C. aureus*) yet their position is consistent with the tree obtained by [84]. The dire wolves form a strong monophyletic cluster, basal to all extant *Canis*, except for the *C. adustu*s (the side-striped jackal) and *L. pictus* (African wild dogs) (Supplementary Fig. 10 A).

The second phylogenetic reconstruction included all five dire wolves (9 genes, 10587bp; Supplementary Table 4; Supplementary Fig. 10 B). We retrieved similar relationships except for the deep nodes in the phylogeny, which could not be resolved. The five dire wolf specimens, however, still formed a highly supported monophyletic cluster. The third phylogenetic reconstruction excluded all the missing data from the alignment (566bp; Supplementary Fig. 10 C). *C.adustus* (side-striped jackal) is now basal to all canids and the Andean fox. This dramatically reduced our power to resolve deeper nodes within the phylogeny, yet the five dire wolves and the

wolf like candis (wolves, coyote, and dog) clustered in two separate, highly supported clades (posterior probability > 0.9).

Nuclear DNA (ascertainment free pipeline)

Filtering

We used bedtools [87] to obtain all regions of the canFam3.1 and VulVul2.2 assemblies that were covered by at least 1 read, in each dire wolf, excluding repetitive elements and CpG islands (see Supplementary Table 5). We then extracted the sequence of these loci from the consensus sequences obtained of each modern species (see genotyping). We then merged the modern data with each dire wolf separately, filtering out loci that 1) had >20% missing data in any species 2) were shorter than 30 bp. DireNTC was excluded from this analysis due to its low coverage (Supplementary Data 1).

Supermatrix

For each data set (multiple, and single dire wolf data sets) we built a maximum likelihood tree, with 100 bootstrap replicates using the GTR+G model as implemented in RAxML [88] by concatenating all loci into a "supermatrix". Analyses of single dire wolf samples, as well as combined samples (DireSP and DireGB; Supplementary Table 5) provided enough power to retrieve a well supported topology (Supplementary Fig. 10). In all cases, the dire wolf samples were basal to all *Canis*, *Lycaon* and *Cuon* species.

We assessed whether the basal placement of the dire wolf could be due to the excessive long external branch induced by singleton DNA damage (Supplementary Fig. 8). To do so, we first estimated the proportion of "true" singletons in each dire wolf sequence by comparing the length of branches in other species to the root of *Caninae* based on the RAxML trees. We then randomly edited the direwolf sequences back to the ancestral state with a probability equal to the excess branch length (see Supplementary Table 5) and built a maximum likelihood tree with the resulting alignment. This correction dramatically reduced the external branch length of both samples (Supplementary Fig. 11). This correction dramatically reduced the external branch length of both samples (Supplementary Fig. 11). In fact after applying this correction, the external branch of each dire wolf was roughly the same length as that of other taxa in the tree suggesting that it removed close to 100% of the excess singletons. This correction, however, did not affect the topology (Supplementary Fig. 11) and was also applied to alleviate potential issues of deamination affecting coalescent-based analysis and molecular dating (see below).

Species tree

We first built species trees based on the multispecies coalescent model using both SNAPP [89] and BPP [90]. These methods have the advantage of taking either very short loci (BPP) or SNPs (SNAPP) rather than well resolved gene tree topologies (e.g. such as produced by ASTRAL [91]) as an input. The latter are difficult to obtain for this study due to the highly fragmented nature of the dataset (see Supplementary Table 5).

BPP

BPP uses the multispecies coalescent, to jointly estimate the species tree topology, divergence time (τ, in coalescent unit) and nucleotide diversity (θ). Here we used BPP to obtain a tree topology (A01 analysis). We used a uniform species tree prior, a diffuse inverse Gamma distribution (3, 0.015) for τ (roughly corresponding to ~10Myr root age) and a diffuse inverse Gamma distribution (3, 0.001) for θ.

We ran BPP on the two highest coverage direwolf samples (DireSP and DireGB). To do so we used every contiguous region of the genome of >30bp in which both direwolf were covered (see filtering section above). We used a burnin of 2,000 samples, sample frequency of 2, and collected 20,000 samples. The topology inferred placed the dire wolves as basal to *Caninae* (all species except the gray fox; Supplementary Fig. 12 A.). To assess whether basal placement could be due to an excess of singletons (see section *Supermatrix* above) we ran the same analysis (with the same priors) on the corrected data (see above). When removing the excess of singleton the resulting tree was identical to the one inferred by RAxML (Supplementary Fig. 12 B.).

SNAPP

We also analysed our SNP data using a multispecies coalescent approach implemented in SNAPP [89]. As for BPP, this analysis was restricted to the two highest coverage direwolf samples (DireSP and DireGB). For each locus, we extracted biallelic SNPs, converting these to 0 (ancestral allele; using the gray fox to polarise allele), 1 (heterozygote) and 2 (derived allele), excluding any site that was missing in a single species. This resulted in 2,433 SNPs without singleton correction and 1,711 SNPs with singleton correction (see section *Supermatrix* above). We sampled *u* and *v* mutation rates parameters from a normal distribution (mean and sigma of 1.0). A uniform distribution (0, 1) was used for the species tree prior, and a Gamma distribution (3, 1000) for θ. This analysis was repeated on the corrected data-set (with correction for singleton excess in dire wolves), with the same priors.

As for BPP, the topology inferred by SNAPP put the dire wolf as basal to *Caninae* (all species except the gray fox; Supplementary Fig. 13 A&B). But similarly, this basal placement was most likely based on excess of singleton as demonstrated by the tree obtained after correcting for excess singletons (Supplementary Fig. 13 C&D).

Discordance visualisation using Discovista

We then inferred maximum likelihood trees, using RAxML, in 500kb, 1Mb and 5Mb sliding windows across the genome (100kb overlap). This analysis was conducted for each dire wolf sequence separately. For each bin we concatenated all sequences that were covered in the dire wolf, using the same filters as for the supermatrix analysis. Only bins with at least 2kb of coverage in a dire wolf were considered (Supplementary Table 6). The frequency, and support, of different topologies were visualised using discovista [92]. Overall this analysis strongly rejects a topology in which the dire wolf is sister species to wolves (dire/can in Supplementary Fig. 14). Topologies in which the dire wolves (dire/out) or the black jackal/side-striped jackal (dire/in) were the most basal group in the phylogeny were almost equally supported - although the dire/in topology obtained slightly higher support in this analysis (Supplementary Fig. 14).

Nuclear DNA (pipeline based on SNPs ascertained in modern genomes)

SNP-calling and variant filtering

We created indexed VCF files for each BAM file using samtools v0.1.18 (mpileup; part of the SAMtools package, [93]) and BCFtools v0.1.17 (call, index; part of the SAMtools package). We used Parallel v20170822 [94] to process each BAM file in parallel. We then converted the autosomal biallelic variants from the VCF files to random “pseudo-haploid” eigenstrat formatted variants using vcf2eig (part of eig-utils; https://github.com/grahamgower/eig-utils) with the following options: -m (include monomorphic sites), -s (include singleton sites), and -t (exclude transitions). The eigenstrat formatted files were then filtered to contain only variants from the list ascertained in the high-coverage modern samples (see Ascertainment section above; see Supplementary Data 13 for a list of genomes that were used for the ascertainment) using eigreduce (part of eig-utils;

https://github.com/grahamgower/eig-utils) with the following options: -m (include monomorphic sites) and -s (include singleton sites). We then used eigreduce (-i, -m, -s) to create eight sets of samples: all samples, modern (non-dire wolf) samples only, modern samples plus DireSP, modern samples plus DireGB, modern samples plus DireSP and DireGB, modern samples plus DireNTC, modern samples plus DireGWC, and modern samples plus DireAFR. Focusing only on transversion ascertained in modern genomes alleviates issues arising from ancient DNA bias as it reduces the incorporation of DNA damage in the analyses (i.e. remove transition singletons found only in ancient genomes).

Supermatrix

The filtered eigenstrat formatted files for each of the eight sets of samples (see above) were then transposed to a PHYLIP file using eig2phylip (part of eig-utils; https://github.com/grahamgower/eig-utils) (Supplementary Table 7). For each set of samples we created supermatrix trees (i.e. all variants concatenated) using the rapid bootstrapping algorithm in RAxML v8.2.4 [95](-f a, -m ASC_GTRCAT) with 100 bootstrap replicates (-# 100) and the Felsenstein ascertainment correction based on the number of invariant sites (--asc-corr=felsenstein), which was calculated as the total ungapped length of the canFam3.1 autosomes minus the length of the alignment.

Discordance visualisation using Discovista

For each of the eight sets of samples (see above) we further broke the genome down into 422 non-overlapping 5Mb windows using eigreduce (-R) (Supplementary Table 8). For each of the 422 windows we then created a tree (i.e. all variants concatenated) using the rapid bootstrapping algorithm in RAxML v8.2.4 (-f a, -m ASC_GTRCAT) with 100 bootstrap replicates (-# 100) and the Felsenstein ascertainment correction based on the number of invariant sites (--asc-corr=felsenstein), which was calculated as the length of the window (5 million bases) minus the length of the alignment. As for the ascertainment free pipeline (see above) we then summarised and visualised the frequency and support of different tree topologies using discovista [92]. Ultimately, we obtained comparable results to those from the ascertainment free pipeline (Supplementary Fig. 15-16).

Molecular dating - MCMCtree

Fossil calibration

Two fossil calibrations were used to calibrate the canid phylogeny so the absolute divergence times could be estimated with MCMCtree (part of the PAML suite v4.9 [96]):

1. We calibrated the root of the tree (*Urocyon* versus the other samples) using a uniform distribution with a minimum of 10.3 Ma and a maximum of 20 Ma. The uniform distribution had soft bounds, implemented as described by [97], such that the true age is between 10.3 Ma and 20 Ma with the left and right tail probabilities being 0.025. The minimum bound was based on the occurrence of *Metalopex macconnelli* at the end of the Clarendonian NALMA (dated to 10.3 Ma; [98]), which appears to be closer to *Urocyon* than to *Canis* [63]. The maximum bound was set close to the end of the Harrisonian NALMA (dated to 20.6 Ma; [98]), which allows for the possibility that some *Leptocyon* species could have been early stem members of an extant canid lineage.

2. We calibrated the divergence between Andean fox (Cerdocyonina) and Canina using a soft-bounded uniform distribution (implemented as above) with a minimum of 4.9 Ma and a maximum

of 10.3 Ma. The minimum bound of 4.9 Ma was based on the occurrence of *Canis ferox*, which is likely a stem member of Canina [63]. The maximum bound was set at the end of the Clarendonian NALMA (dated to 10.3 Ma; [98]), based on the absence of recognised crown members of Canini prior to this point in time.

Molecular data

We used the same data as for the BPP analysis above (corrected and uncorrected). The data was concatenated and used to construct the molecular alignment. For this analysis we used only one representative of each species, randomly selecting one African wild dog and using the Eurasian wolf (ptw, SRS661492 see Supplementary Data 13) instead of the Yellowstone wolf genome (based on overall coverage). As for the species tree, this analysis was restricted to the two highest coverage direwolf samples (DireSP and DireGB). We ran the analysis on both corrected and uncorrected data, i) without direwolf sequences (10 species), ii) direwolf sequence with the highest coverage (DireGB; 11 species), and (iii) with both direwolf sequences (DireGB and DireSP; 12 species).

Simulated data

We carried out a simulation study to investigate the impact that incorporating deamination (from aDNA) can have on divergence time estimates. To do so, we simulated molecular alignments based on a set tree topology, and gradually increased the number of external substitutions on one or two (i.e., if either one or two ancient sequences were simulated to be included in the alignment) branches in the tree.

We first ran BASEML (part of the PAML suite v4.9 [96]) to estimate the (i) best-scoring maximum-likelihood (ML) tree for the molecular data under the HKY+Γ5 substitution model, the (ii) transition/transversion ratio, the (iii) base frequencies, and the (iv) shape parameter α for the discrete-gamma model for rate heterogeneity. These parameters were estimated for the corrected molecular alignments (alignments with one and two direwolf sequences, respectively) under both tree topologies ("A" and "B"), being then used to simulate the corresponding molecular alignments under each topology to which substitutions were subsequently added.

Alignments were simulated using seqgen [99] as implemented in the phyclust R package [100]. We used the parameters estimated by BASEML for the corrected molecular alignments under each tree topology evaluated. This means that the simulated alignments had the same sequence length and biological properties than the corrected alignments. We then added an increasing proportion of "errors", err_{added} = 0%, 0.5%, 1%, 3%, and 5%; on the simulated direwolf sequence(s). Altogether, this resulted in 10 simulated alignments (5 error rates x 2 topologies) with 12 species (two dire wolves) and 10 simulated alignments with 11 species (one dire wolf).

Bayesian model selection for molecular clock and tree topology

We used the autocorrelated-rates, also known as geometric Brownian diffusion (GBM [97,101]), and the independent log-normal rates (ILN [102,103]) models. The analysis was run on different topologies for both alignments: "A" (dire/out in Supplementary Fig. 6) and "B", (dire/in in Supplementary Fig. 14). Altogether, we evaluated the fit of 4 models clock/topology combinations for both corrected and uncorrected data sets, prior to the dating analysis: (i) model 1: GBM + topology A, (ii) model 2: GBM + topology B, (iii) model 3: ILN + topology A, and (iv) model 4: ILN + topology B.

For each model, the likelihood values collected during the MCMC for each model were used to estimate marginal likelihoods with the mcmc3r R package [104]. Marginal likelihoods were used to compute Bayes factors (BFs) and posterior probabilities, which were then used to select the best-fitting model for each alignment. This analysis supported the combination of topology A + clock GBM (dire/out in Supplementary Fig. 6) as the model that best fit the data (Bayes Factor=5.53; Posterior probability= 0.941) and was used for the subsequent Bayesian inference analyses. The results obtained for each model tested are shown in Supplementary Table 9.

Divergence time estimation

Priors and substitution model selection

Estimation of divergence times for the molecular alignments was carried out using MCMCtree under the GBM model (see above) and the HKY+Γ substitution model with 5 discrete-gamma categories. We used a uniform prior for node ages using the birth-death (BD) process [97] with the following parameter values: λ_{BD} = 1 (birth-rate), μ_{BD} = 1 (death-rate), and ρ_{BD} = 0.1 (sampling fraction for extant species).

We used the gamma-Dirichlet distribution for the rate (r) prior as implemented in MCMCtree [105]. The shape parameter was set to α = 2, which corresponds to a diffuse prior. The scale parameter β was chosen based on the estimated molecular branch lengths of the phylogeny. To do so, we first ran RAxML v8.2.10 [95] on the concatenated molecular alignment. The resulting best-scoring ML tree was used to estimate the distance from the tips to the root (i.e., the number of substitutions from the tips to the root), $b_{root\text{-}tips}$. Given that $b_{root\text{-}tips} = r \times t_{root}$ and that the mean rate of the gamma-Dirichlet distribution is defined as $r = \alpha / \beta$, then, $\alpha / \beta = b_{root\text{-}tips} / t_{root}$. We can therefore use $b_{root\text{-}tips}$ to estimate the scale parameter as $\beta = (\alpha \times time_{root}) / b_{root\text{-}tips}$. Supplementary Table 10 lists the priors used for each analysis. Lastly, the prior on σ_i^2 was defined using a diffuse gamma-Dirichlet distribution, $\sigma_i^2 \sim \Gamma(2,2)$.

Results - simulated and real data sets

The divergence times were estimated for both real and simulated data sets using the preferred model according to BFs: autocorrelated-rates + topology “A”. The results for the estimated mean age of the (i) root, the (ii) Vulpini-Canini divergence, the (iii) direwolf divergence from the jackals and the rest of canids, the (iv) jackals split from the rest of canids, and (vi) the divergence between the two dire wolves are shown in Supplementary Table 11. Note that the posterior divergence times for alignments without dire wolf sequences were only estimated for the corrected and uncorrected real data sets. This is because the simulations required alignments in which at least one direwolf sequence was present, so aDNA deamination could be simulated.

Altogether, the results with the simulated data sets indicate that the presence of errors slightly affected divergence times, although the observed effect was minimal (e.g., the time of divergence between dire wolves and other canids increased by 3% at 0.5% error and 6% at 1% error). Nevertheless, introducing errors dramatically affected divergence time between the two ancient sequences on which we added errors; i.e., the time of divergence between the two dire wolves increased by ~two fold (Supplementary Fig. 17 and Supplementary Table 11). Last, we also ran the analysis without the dire wolf sequences, which yielded highly similar results (see column “Estimated divergence times for the alignment without dire wolf specimens” in Supplementary Table 11). The results presented in Fig. 2A are based on the alignment with only one dire wolf specimen (DireGB).

D-statistics

We used D-statistics, as implemented in Admixtools [106], to detect gene-flow from the dire wolf into other canid lineages and to further explore taxonomic relationships. We only used the two highest coverage dire wolf genomes for this analysis (DireSP and DireGB). We computed D of the form D(gray_fox, (dire wolf, (P1, P2)), where P1 or P2 can be any extant canid genome. We did not compute these statistics using the black-backed or side-striped jackal given the uncertainty in the topology at the root of *Canis* (see above). We used randomly sampled reads per site instead of genotype calls for all genomes, including in high coverage modern genomes, in order to account for different depth of coverage (same procedure as for the dire wolf; see above). The SNPs used in this analysis were ascertained in modern high coverage sequences (see section ascertainment above). We used only transversions to reduce potential biases arising from ancient DNA damage in the analysis as the inclusion of transitions have been shown to significantly affect this type of analyses [107]. We used a weighted block jackknife procedure over 5Mb blocks to assess significance (|Z|>3) and repeated the analysis on datasets aligned to both the Red fox (VulVul2.2) and dog (Canfam3.1) references genomes to avoid potential issues arising from a reference bias. We only used scaffolds longer than 5Mb.

We found no evidence of an excess of shared derived allele between dire wolves and North American wolves or coyotes , since the most recent common ancestor of African wolves, and gray wolves using both canFam3.1 and VulVul2.2 as reference (Supplementary Fig. 18; Fig. 2B;Supplementary Data 14). These results indicate that our data does not support the existence of gene flow between dire wolf and extant populations of North American canids (coyotes and gray wolves) taking place since they diverge from their most closely related Eurasian gray wolves and African wolves counterparts.

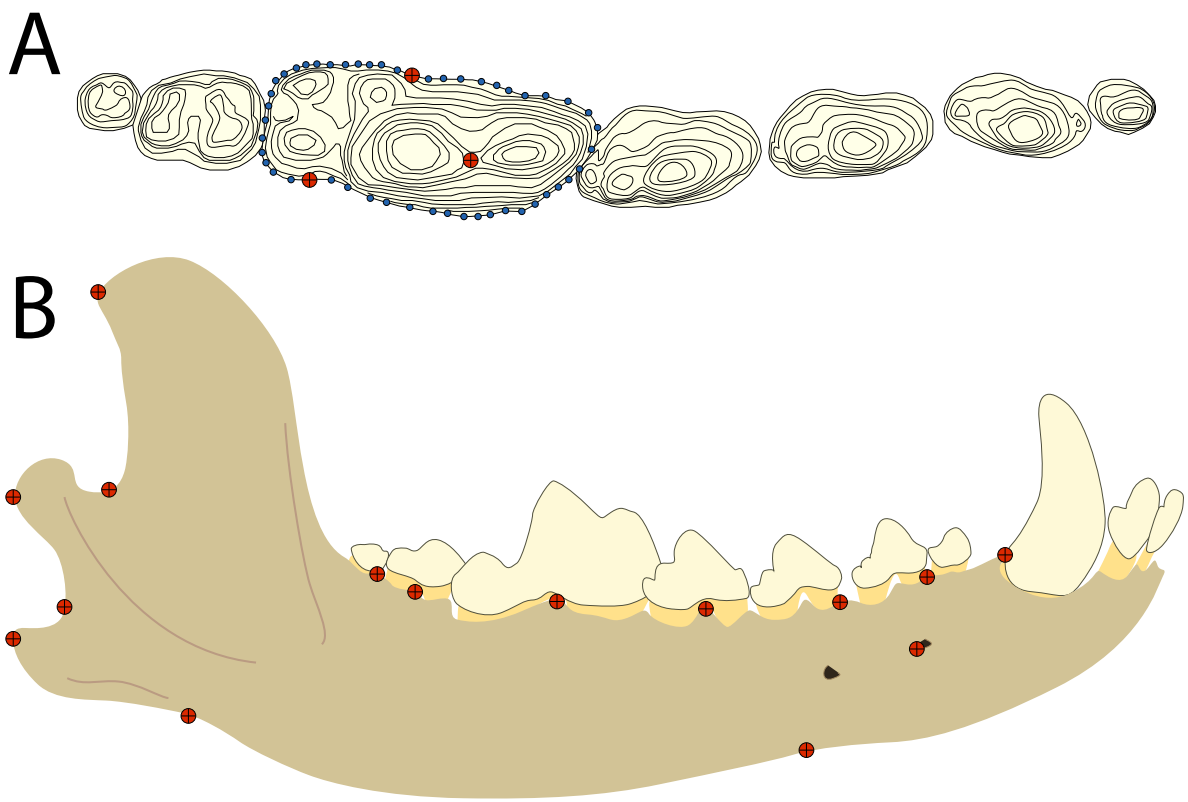
Next we used D-statistics to further assess taxonomic relationships. We first evaluated whether the wolf-like canids (coyotes, wolves etc.) and dhole shared more derived alleles with African wild dogs or with dire wolves. We computed all possible combinations involving both dire wolves and african wild dog genomes. Our results show that wolf-like canids and dhole share more derived alleles with the African wild dog than with dire wolves (Supplementary Fig. 19). This finding is consistent with our phylogenetic analysis indicating that the dire wolf represents an outgroup to these lineages. Our results, however, are also consistent with admixture between dire wolf lineage and the ancestor of the dhole, wolves and coyotes (Supplementary Fig. 20). This is demonstrated by the fact that the dire wolf genome shares significantly more derived alleles with the genome of the gray wolves, coyote, African wolf, ethiopian wolf, and dhole, than with the genome of the African wild dog (Supplementary Fig. 20). This result was replicated using both the dog reference genome (Supplementary Data 15).This admixture between ancestral lineages could partly explain why our phylogenetic analyses could not resolve the root of this phylogeny.

Using the same approach we were also able to confirm that the dire wolf represents an outgroup lineage to wolf-like canids and dhole (Supplementary Fig. 21). In fact, all combinations in which the dire wolf was set as P3 yielded D-values of ~0, indicating that no gene flow took place between the dire wolf lineage and wolf-like canids since their common ancestor. Lastly, we repeated this analysis with the Ethiopian wolf instead of the dhole. This analysis also yielded values of ~0 when the dire wolf was set as P3 (Supplementary Figure 22). These D-statistics, however, also the existence of gene-flow between the dire wolf and the ancestor of all wolves (including Eurasian and African wolves) and coyotes, using DireGB yet this signal became non-significant when using either DireSP and when using the VulVul2.2 assembly (Supplementary Data 16) This suggests that this signal might be due to reference bias toward CanFam3.1.

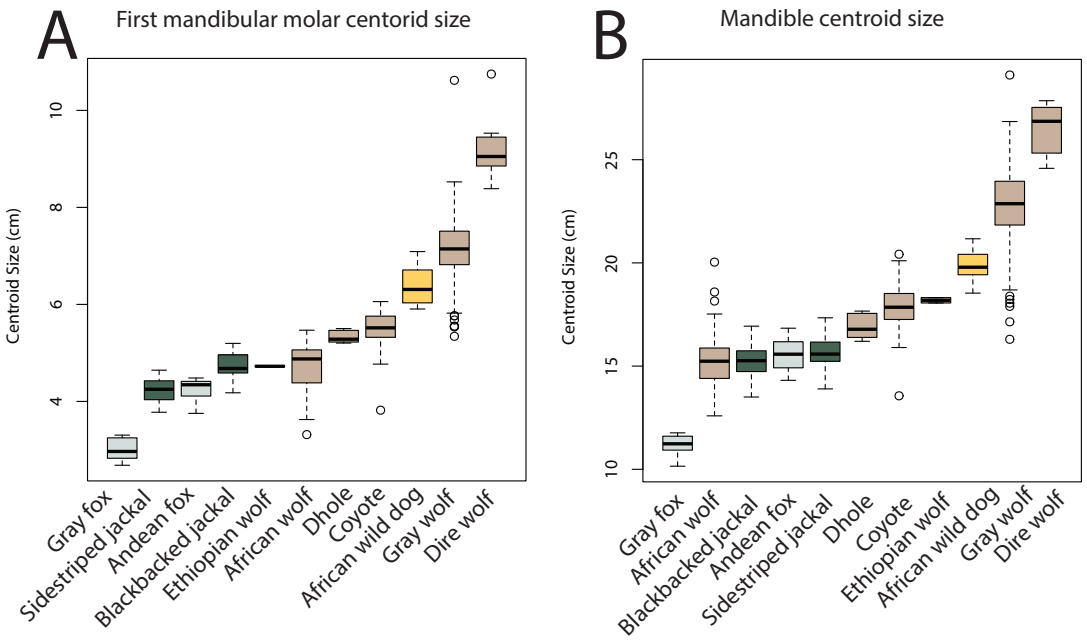
Our results are also consistent with admixture signal described in [83]; namely admixture between wolf and golden jackal (*e.g.* D(gray fox (golden jackal, (Eurasian wolf, coyote))=-0.104, Z<-3 using canFam3.1 and D=-0.25, Z<-3 using VulVul2.2) and between African wolf and Ethiopian wolf (*e.g.* D(gray fox(ethiopian wolf, (eurasian wolf gray wolf))=-0.35, Z<-3 using canFam3.1 and D=-0.33, Z<-3, using VulVul2.2).

We also found evidence of gene flow between side-striped jackals and the ancestor of the dhole, wolves and coyotes (Supplementary Fig. 23). At first, this signal seems to involve the ancestor of the dhole, wolves and coyotes as well as african wild dogs (see result on canFam3.1 in Supplementary Fig. 23). This signal, however, disappeared when the analysis using the VulVul2.2 reference genome (Supplementary Fig. 23), suggesting the existence of a reference bias in the African wild dog genomes toward the dog (canFam3.1) reference genome.

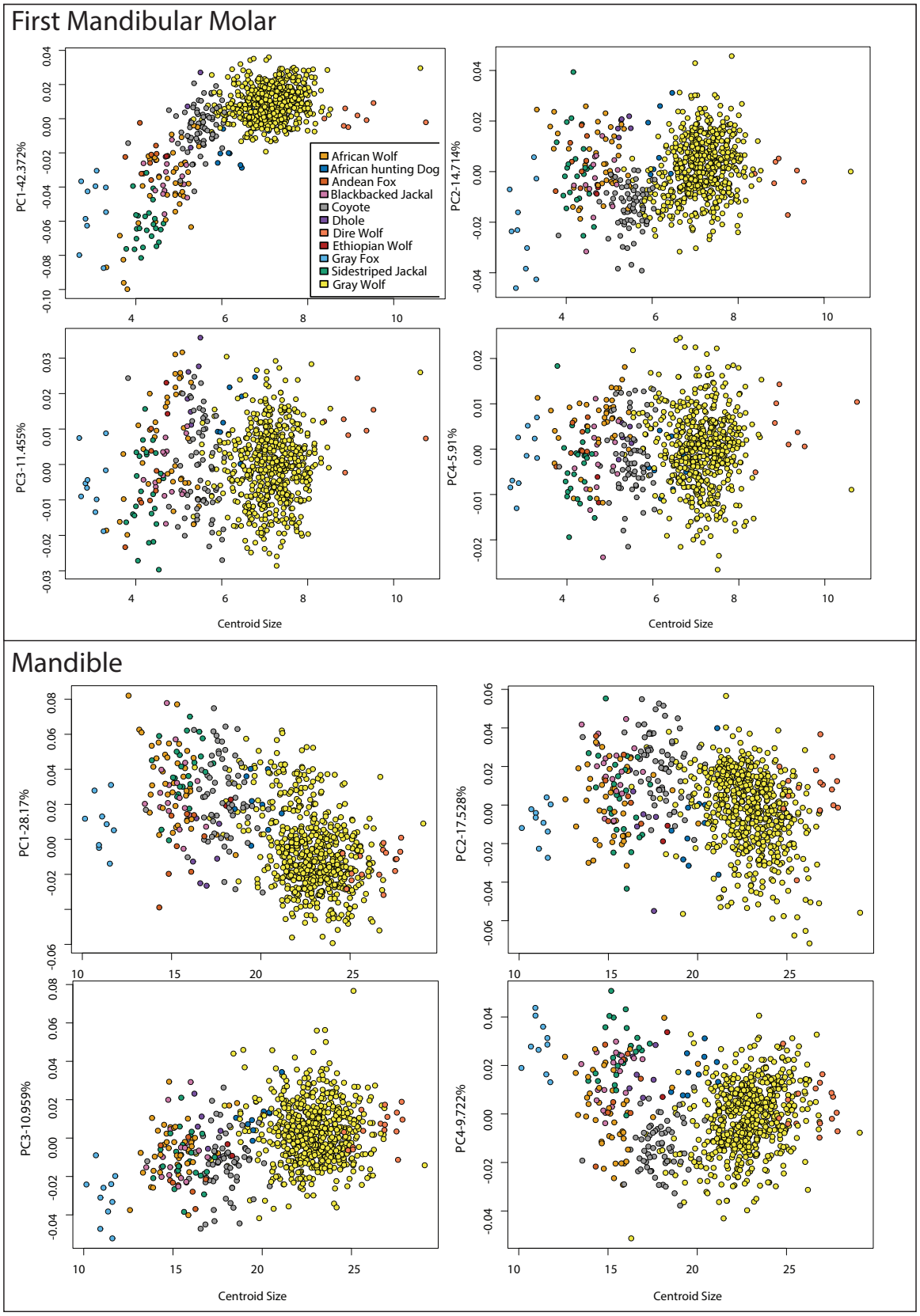
Supplementary Figures



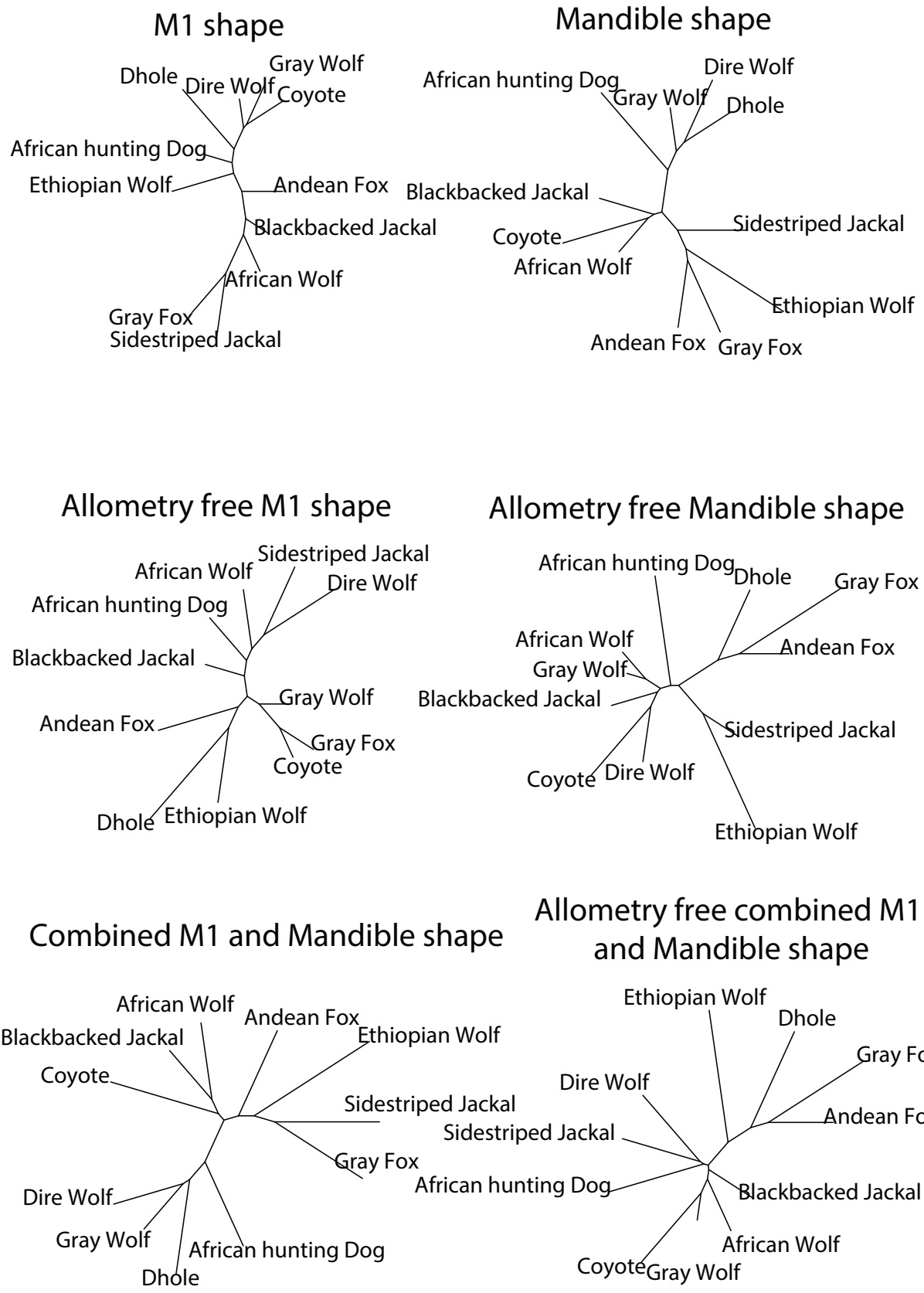
Supplementary Figure 1. Landmark configuration for geometric morphometric protocol. Landmarks are marked in red while sliding semi landmarks are marked in blue. A. First mandibular molar with both landmarks and sliding semilandmarks configurations; B. Mandible with landmark configuration



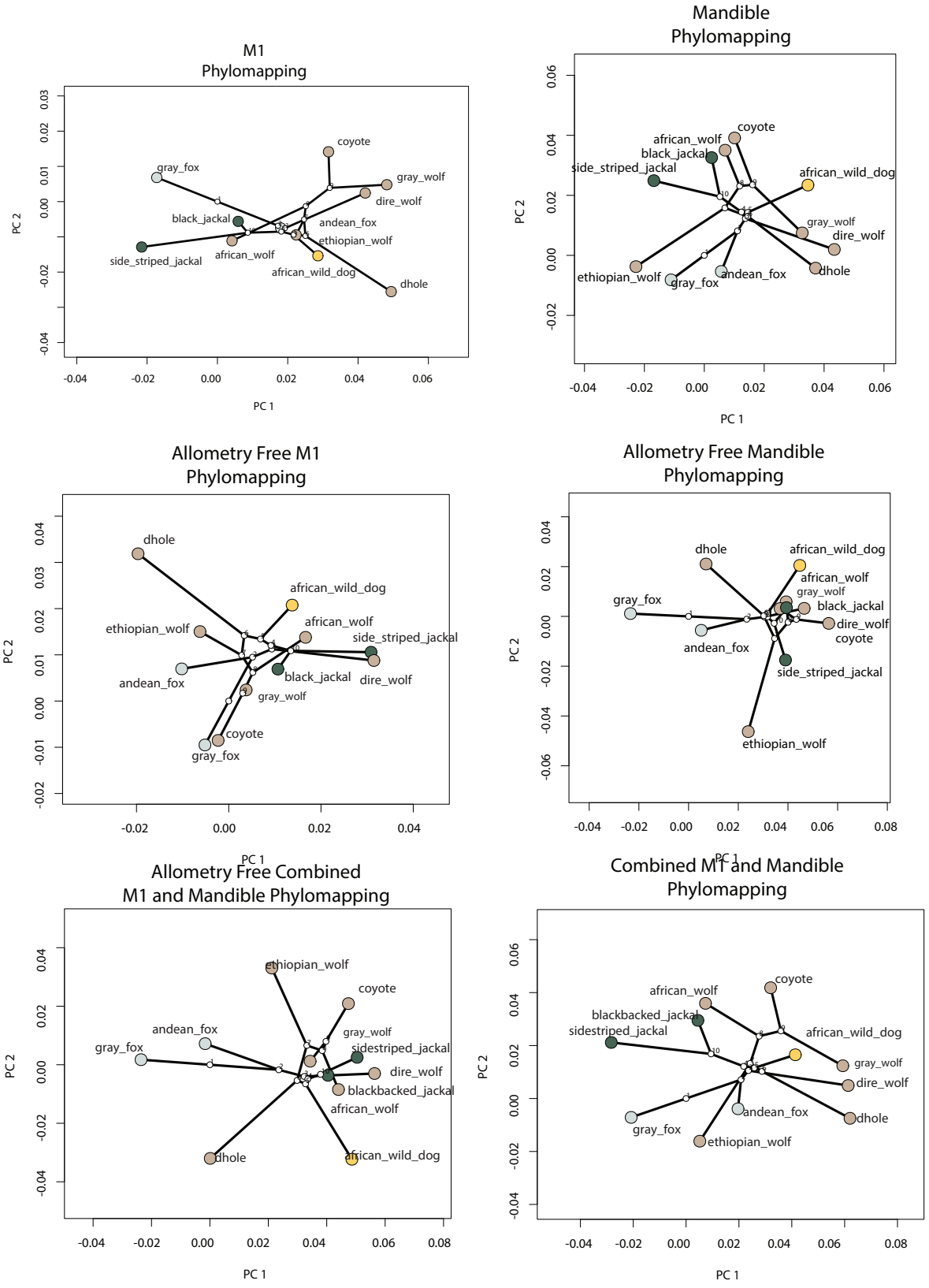
Supplementary Figure 2. Boxplot of centroid size comparisons. Colors correspond to: Gray=outgroups; Green=jackals; Brown=wolf-like canids; Yellow=African hunting dog. A. Mandible centroid size; B. First molar centroid size.



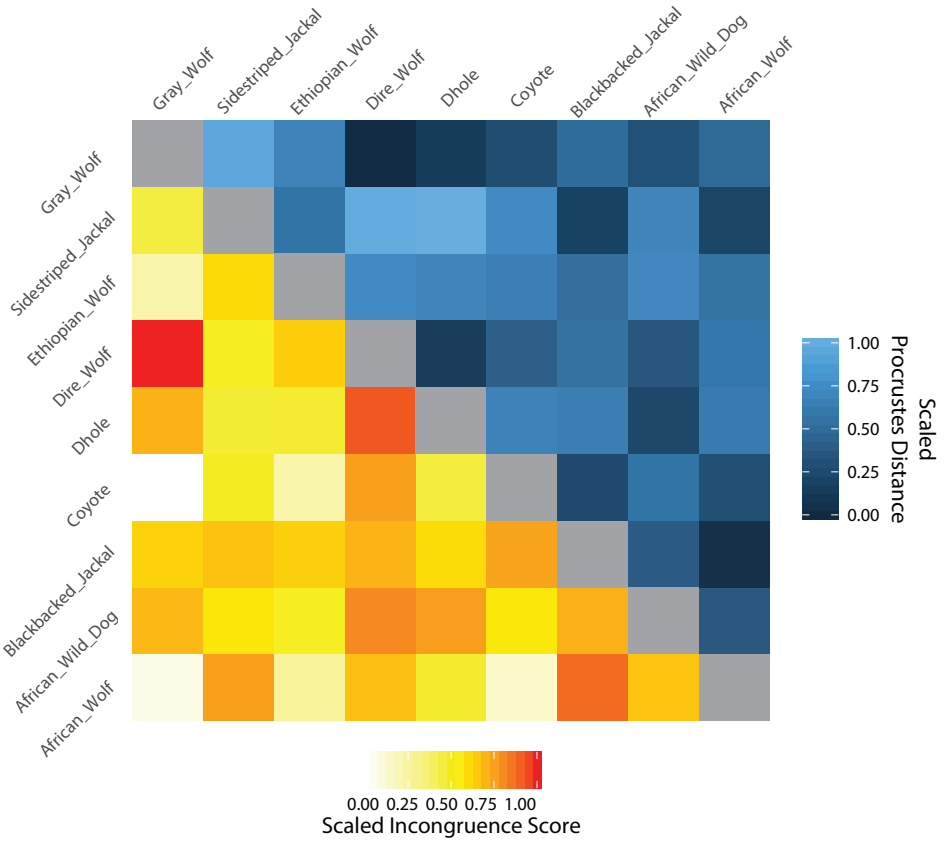
Supplementary Figure 3. Panel of first two PCs versus CS in both mandibular and first molar datasets. CS is on the x axis for all plots. Plots show PCs 1–4 for each morphological data set versus CS. PC variance is listed on the y axis of each plot.



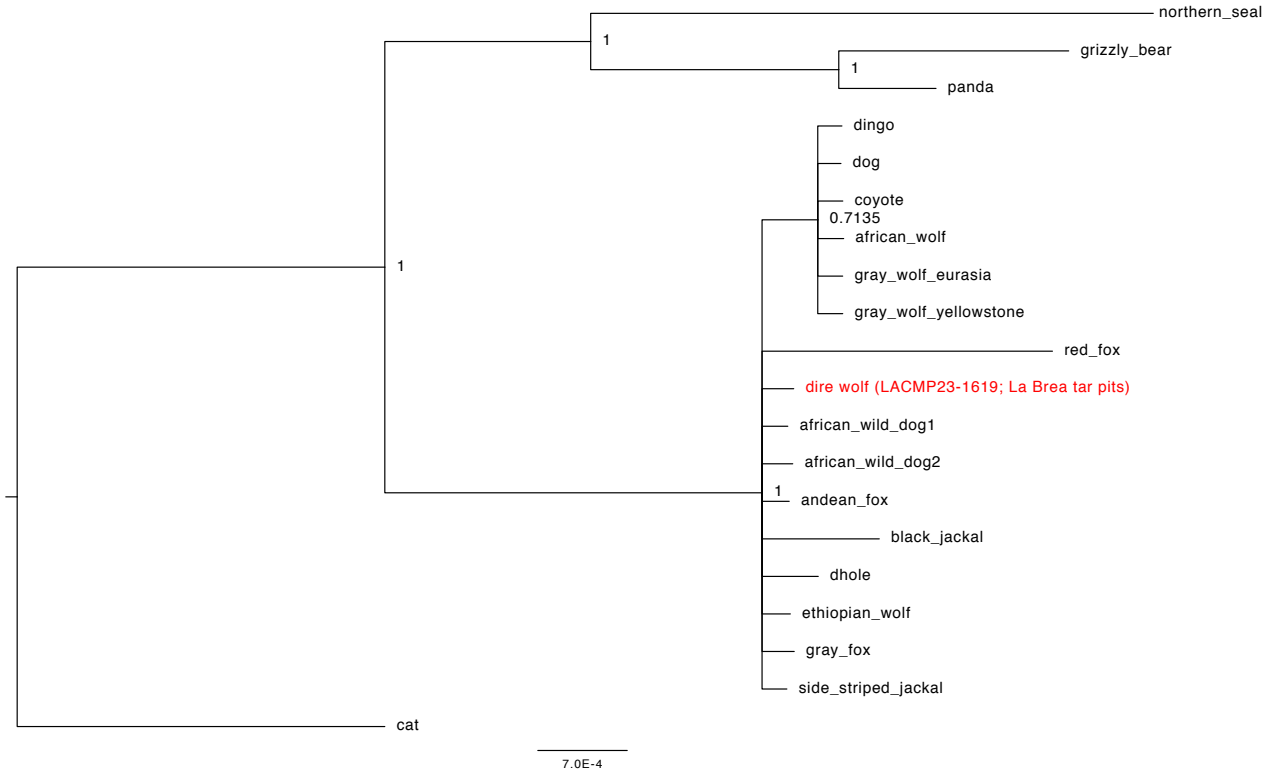
Supplementary Figure 4. Panel of unrooted neighbor joining trees based on different metrics.



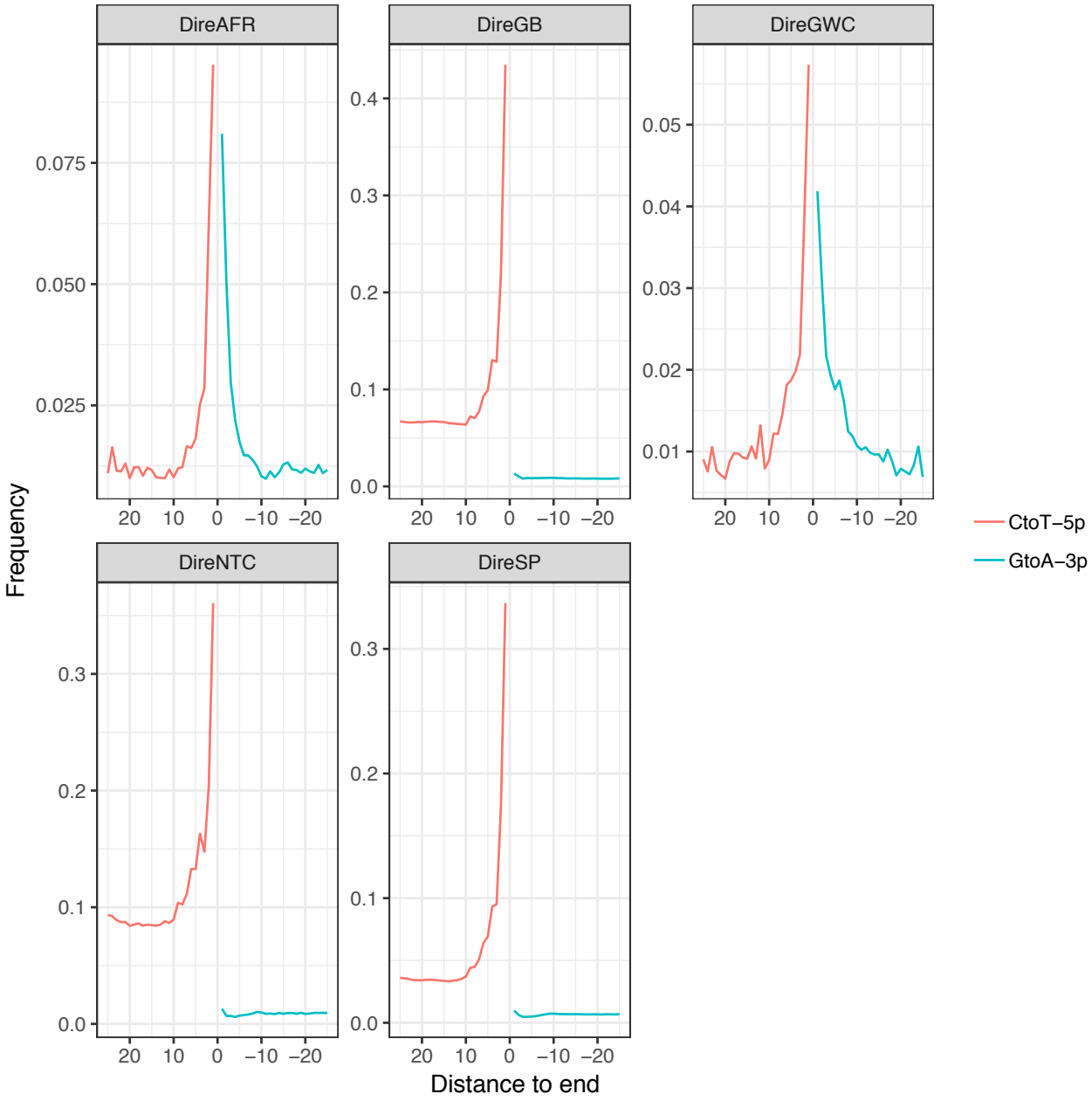
Supplementary Figure 5. Panel of PCAs with phylogenies mapped to mean shapes. Colors correspond to: Gray=outgroups; Green=jackals; Brown=wolf-like canids; Yellow=African hunting dog.



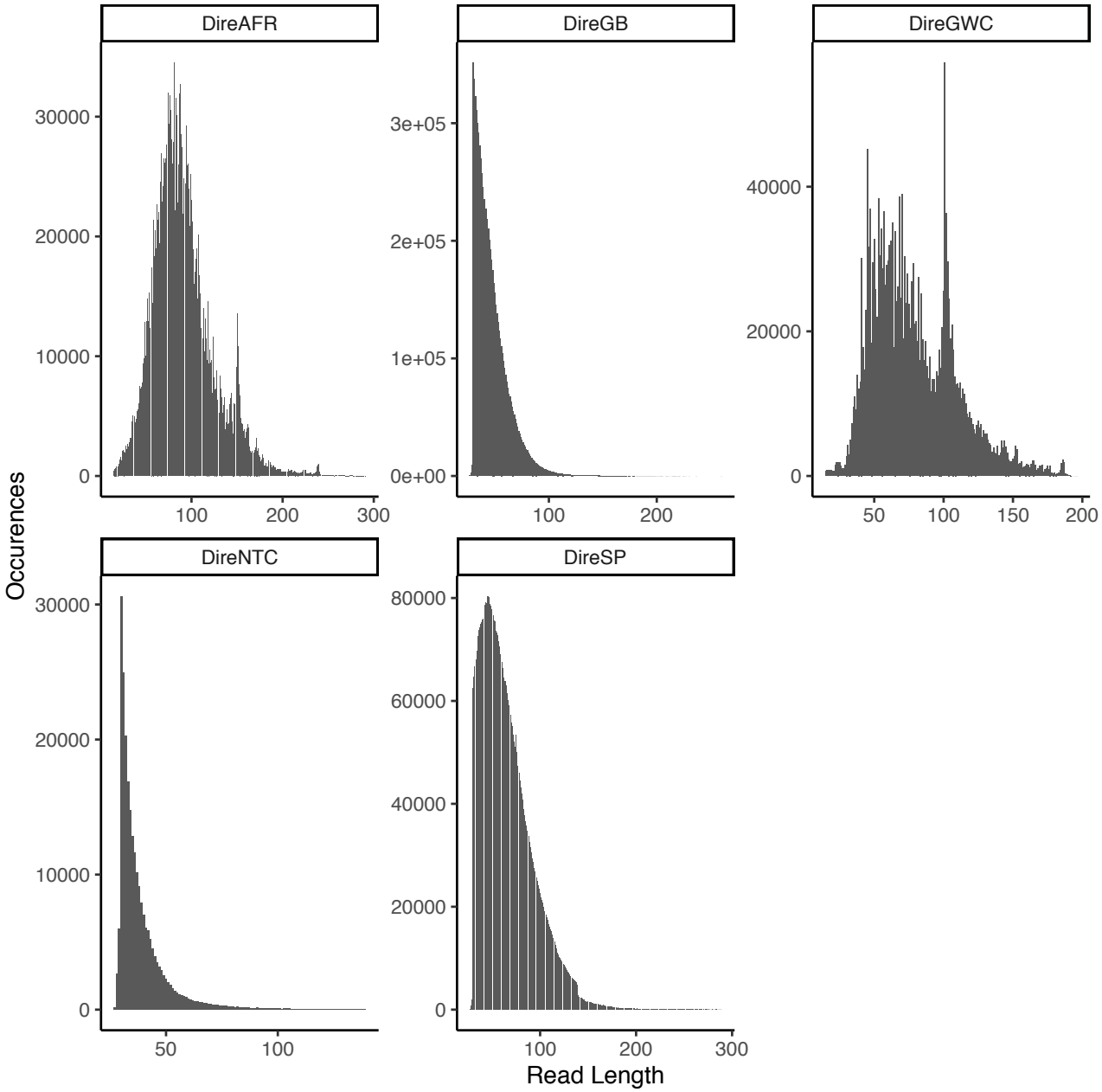
Supplementary Figure 6. Heatmap visualisation of incongruence scores and full Procrustes distances. The lower triangle, scaled in red to white consists of incongruence scores constructed from combined morphological datasets versus genetic distances and transformed and scaled for ease of visualisation. Red scores represent highly similar morphologies with distant genetic divergence, whereas whiter scores represent distant morphologies with recent genetic divergence. The upper triangle represents normalised Procrustes distances (i.e. the distances have been normalised to fall between 0 and 1) from blue to black, where darker blue/blacks colors represent shorter distances and lighter blues represent greater distances.



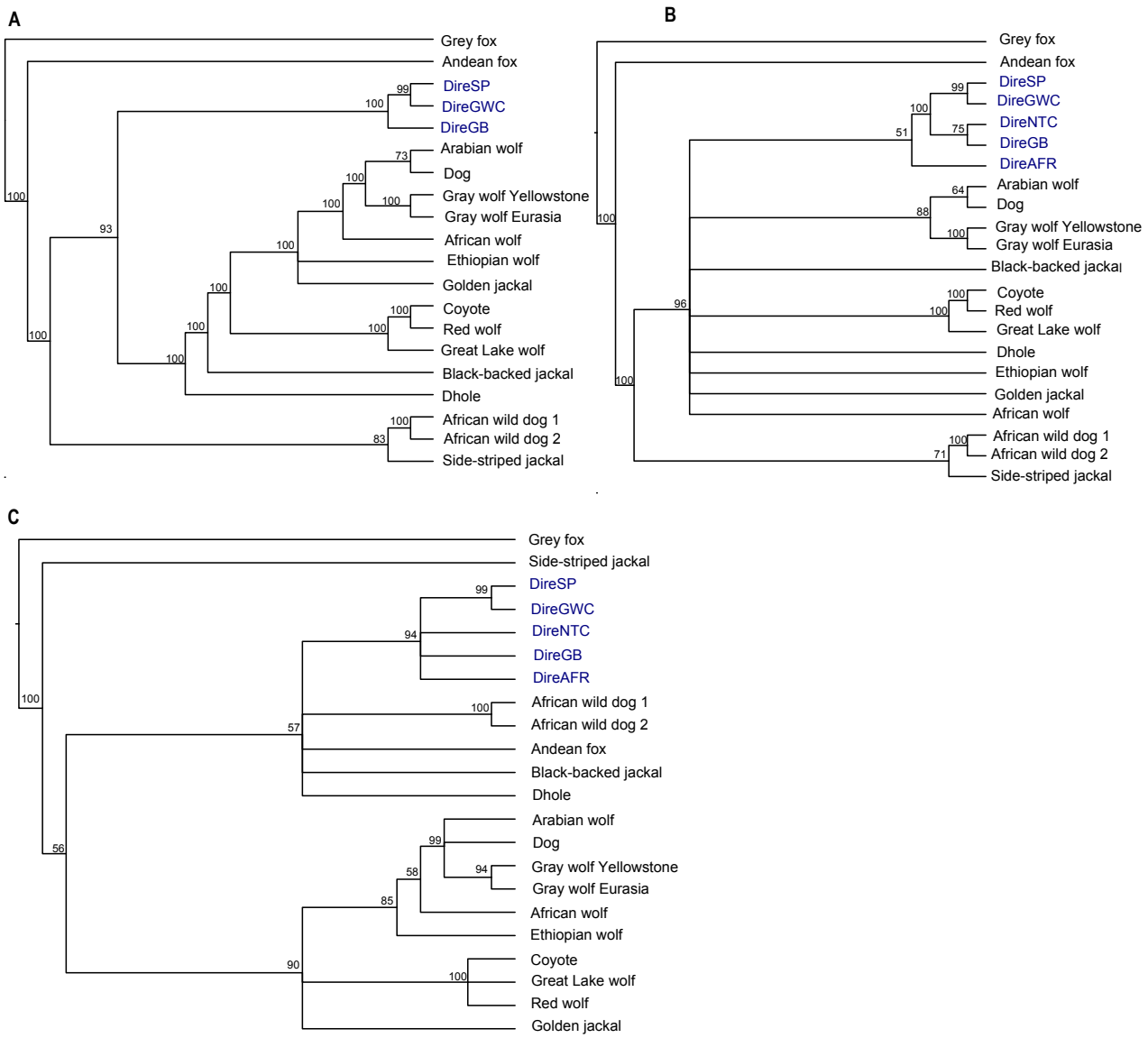
Supplementary Figure 7. Bayesian phylogeny (MrBayes) of the concatenated COL1A1 and COL1A2 amino acid sequence.



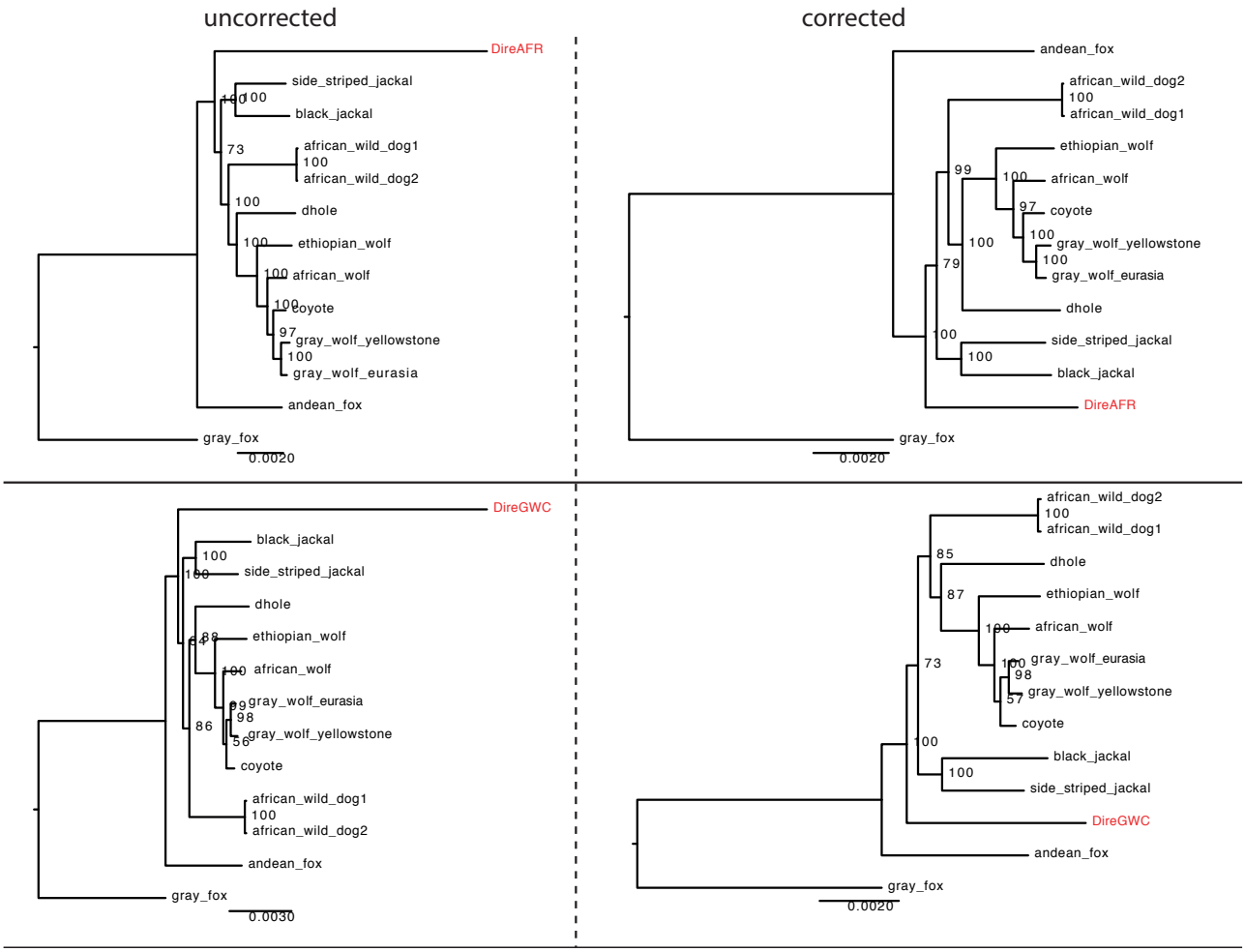
Supplementary Figure 8. Per library C to T (red) and G to A (blue) frequency of mis-incorporation at 3' and 5' end of read for samples used in nuclear and mitochondrial genome analyses.



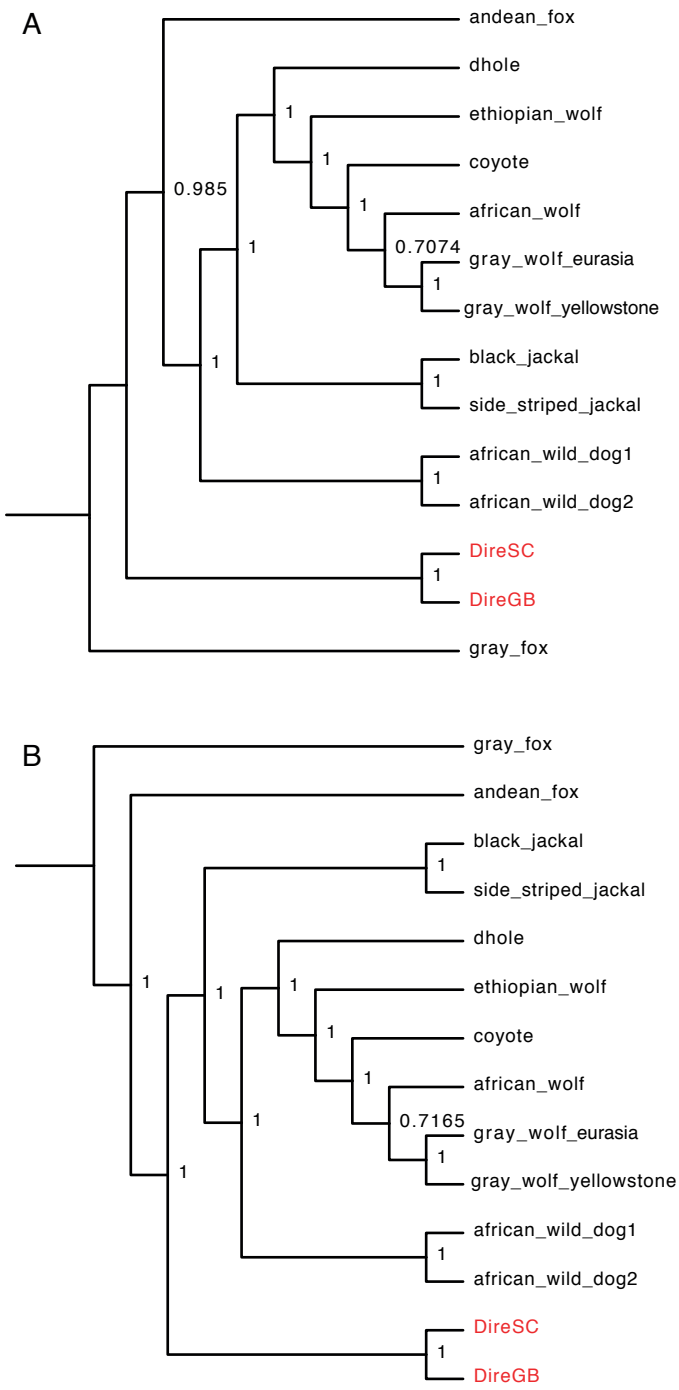
Supplementary Figure 9. Per library read length distribution.



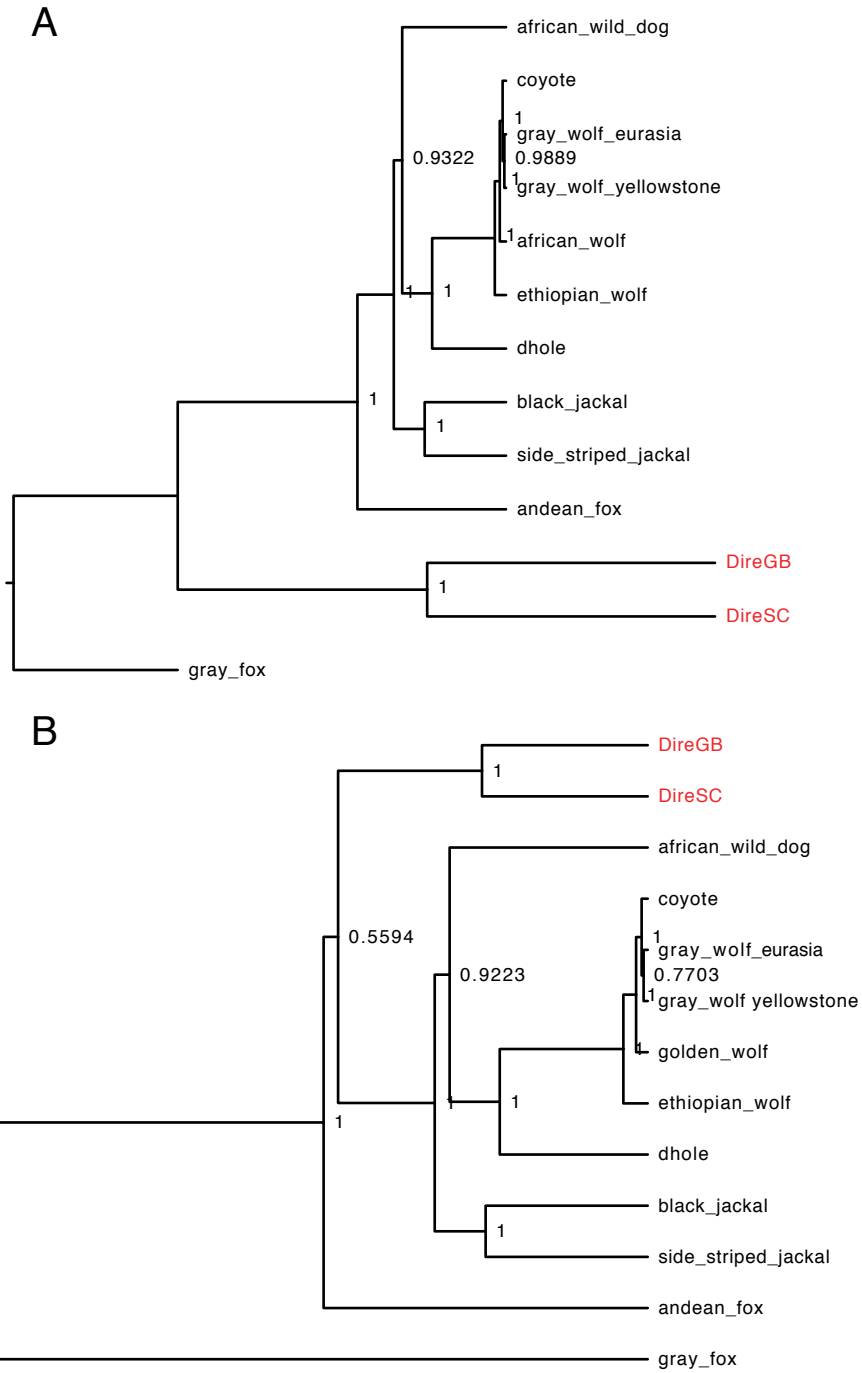
Supplementary Figure 10. Bayesian phylogeny (MrBayes) of mitochondrial DNA. A. Three dire wolves were included - based on 9 mitochondrial genes (10587bp) (DireNTC, DireAFR were excluded) B. Same as A but with the five dire wolves. C. All five direwolves specimens without any missing data (566bp).



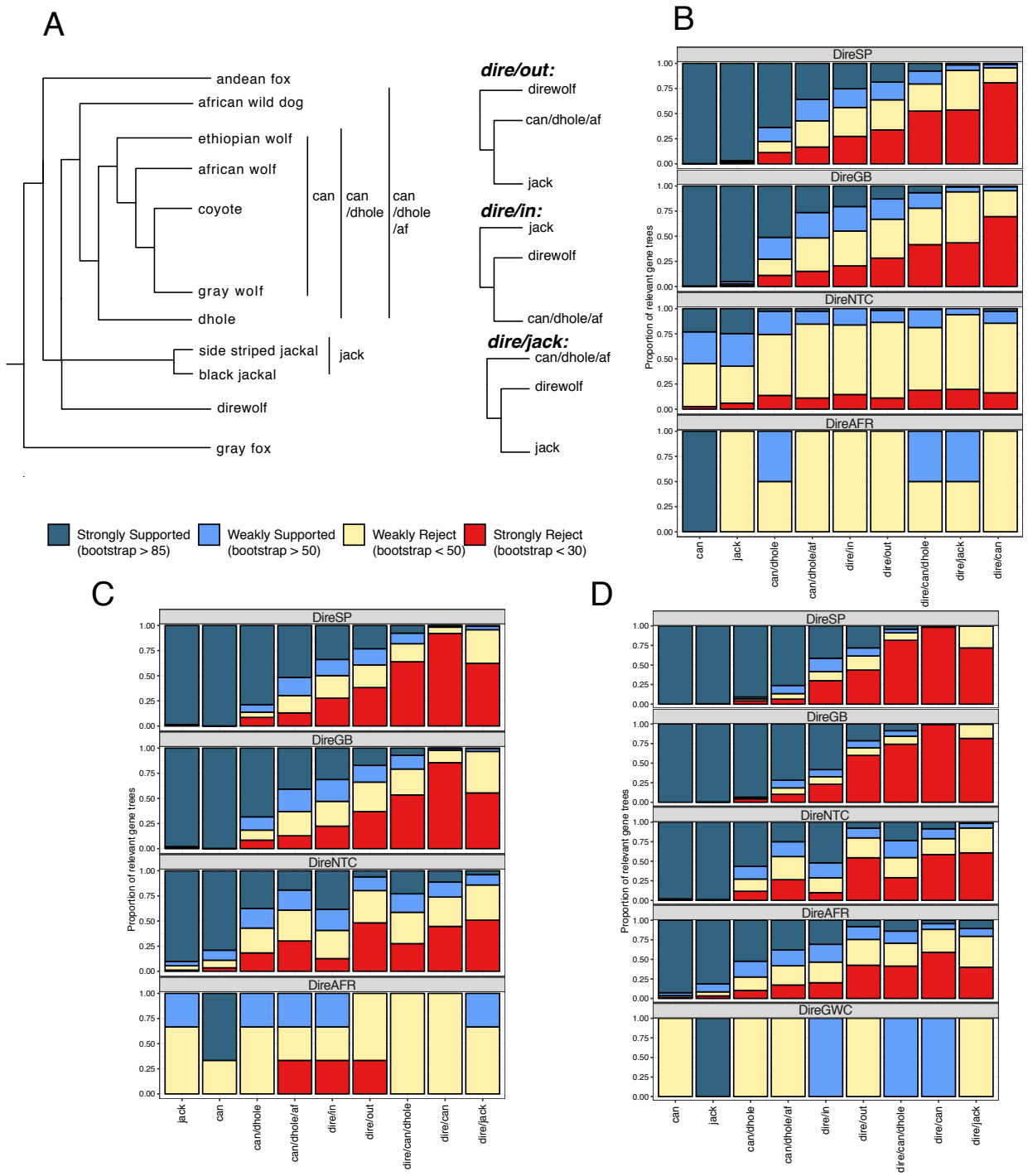
Supplementary Figure 11. Maximum likelihood (RAxML) trees based on concatenation of sequence covered in each dire wolf, with and without branch correction scheme (removing singleton in dire wolf; see section supermatrix).



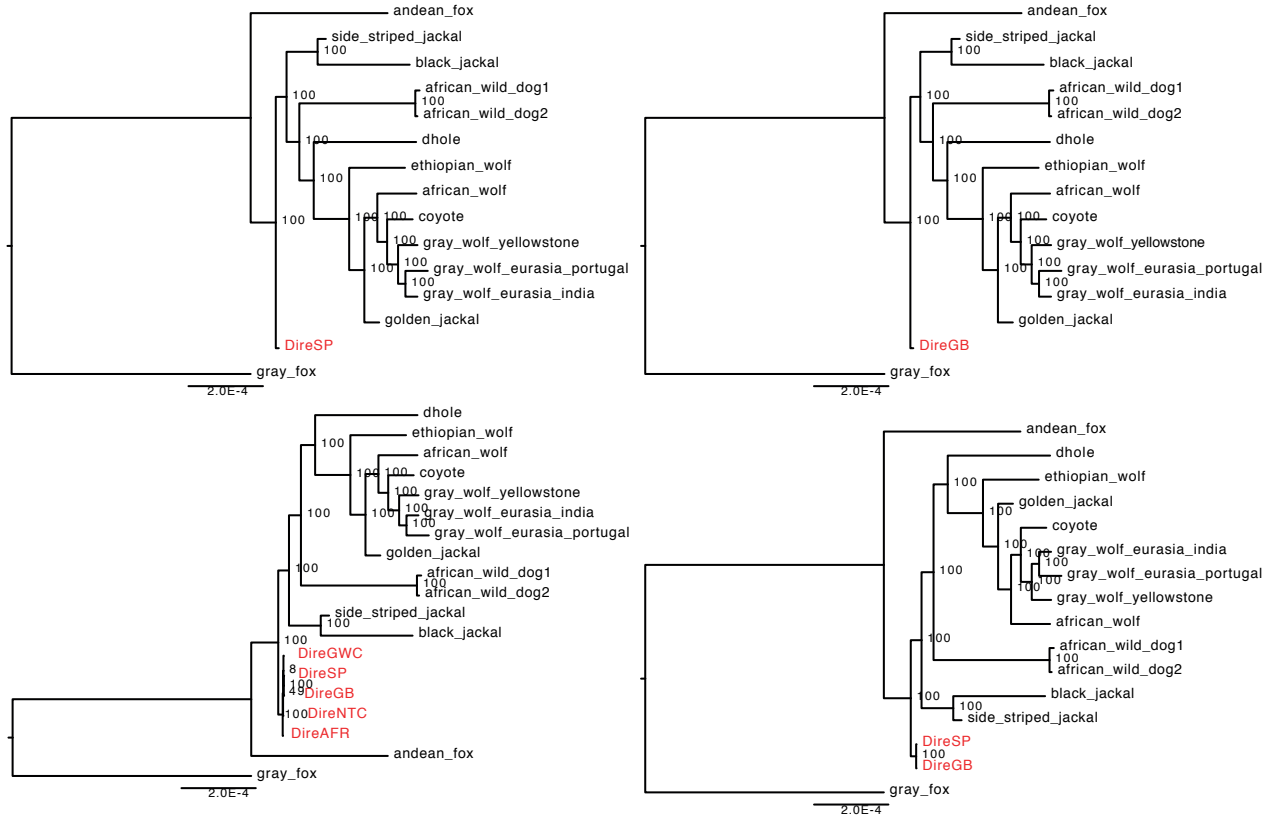
Supplementary Figure 12. Phylogeny tree built using BPP. A. Without branch correction scheme B. With branch correction scheme.



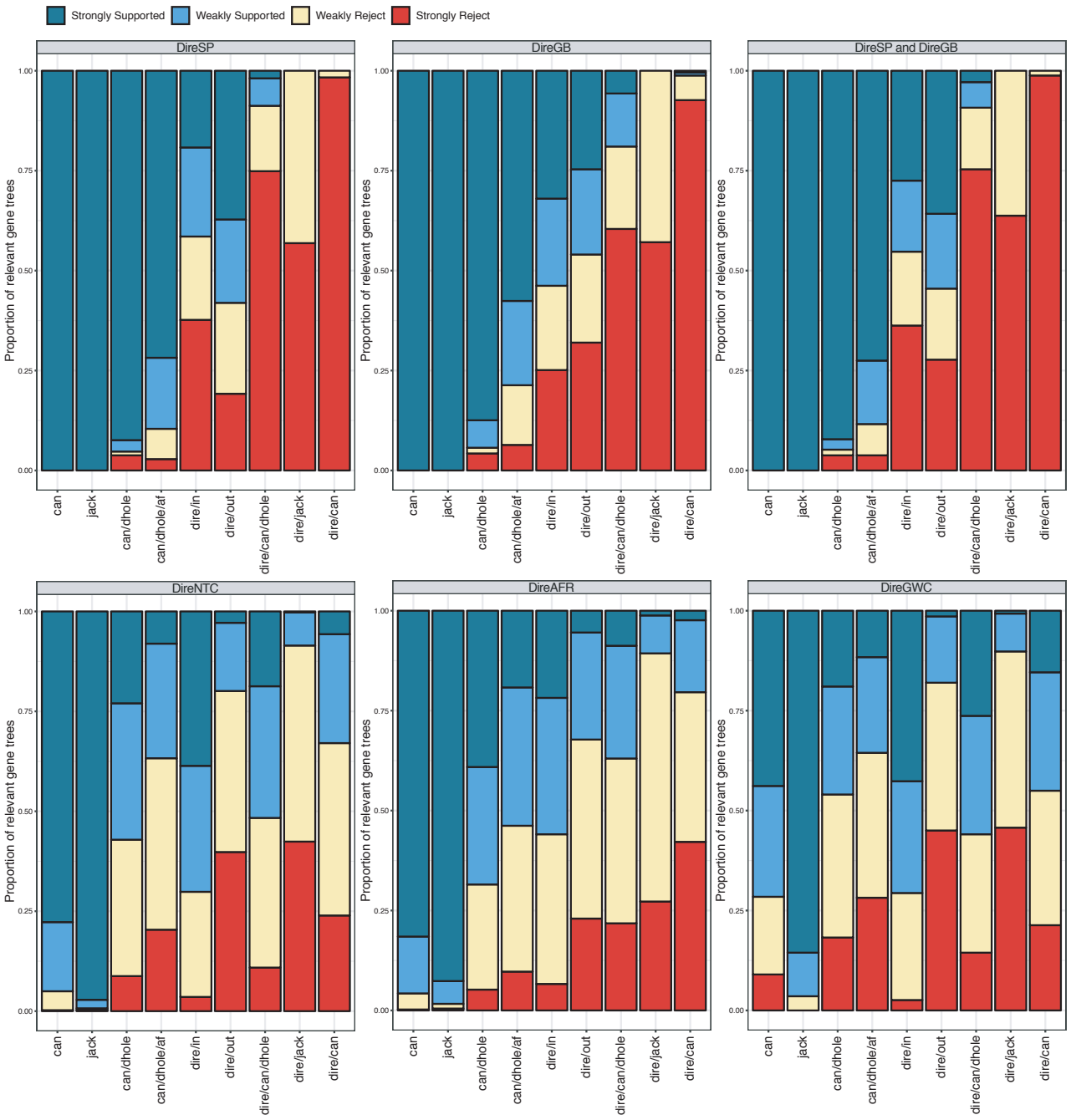
Supplementary Figure 13. Phylogeny based on SNPs, built using SNAPP. A. Without branch correction scheme B. With branch correction scheme. Node labels correspond to posterior probabilities.



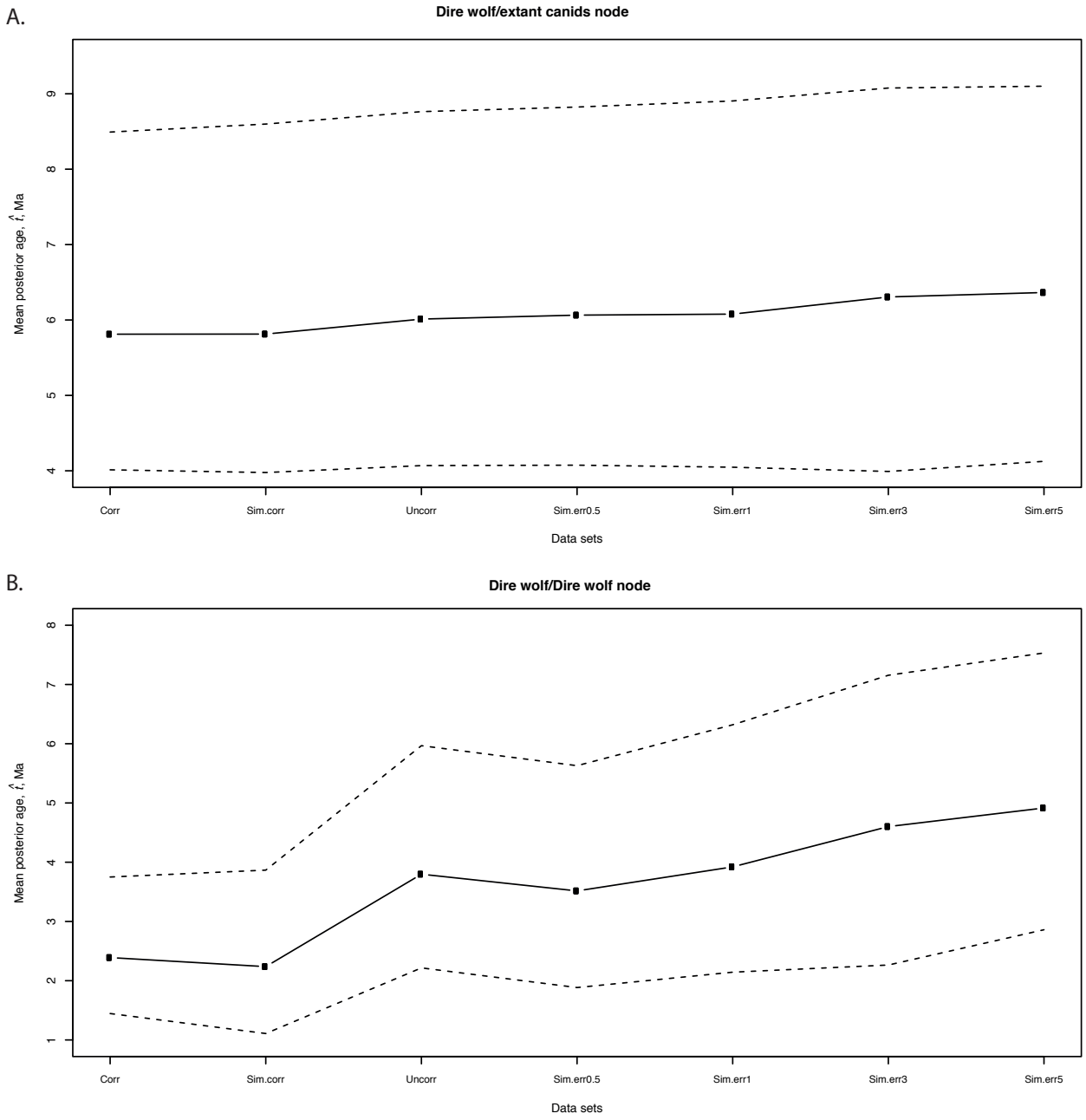
Supplementary Figure 14. Discordance visualisation of maximum likelihood trees build from different bins size. A. Tree defining clades (jack, can, cand/dhole, can/dhole/af) for which we measured support. In addition, the three possible arrangements of the dire wolf, African jackal, and wolf-like canid lineages are displayed (dire/out, dire/in, dire/jack). B. Clades and/or alternative topologies displayed in panel A are listed on the x-axis. The y-axis depicts the proportion of individual trees (made from 500kb bins along the genome) that strongly support (>85% bootstrap support; dark blue), weakly support (□85% bootstrap support; light blue), weakly reject (best tree favours alternative arrangement with □85% bootstrap support; orange), or strongly reject that clade/topology (best tree favours alternative arrangement with >85% bootstrap support; red). C. As for panel B, but using 1Mb bins. D. As for panel B, but using 5Mb bins.



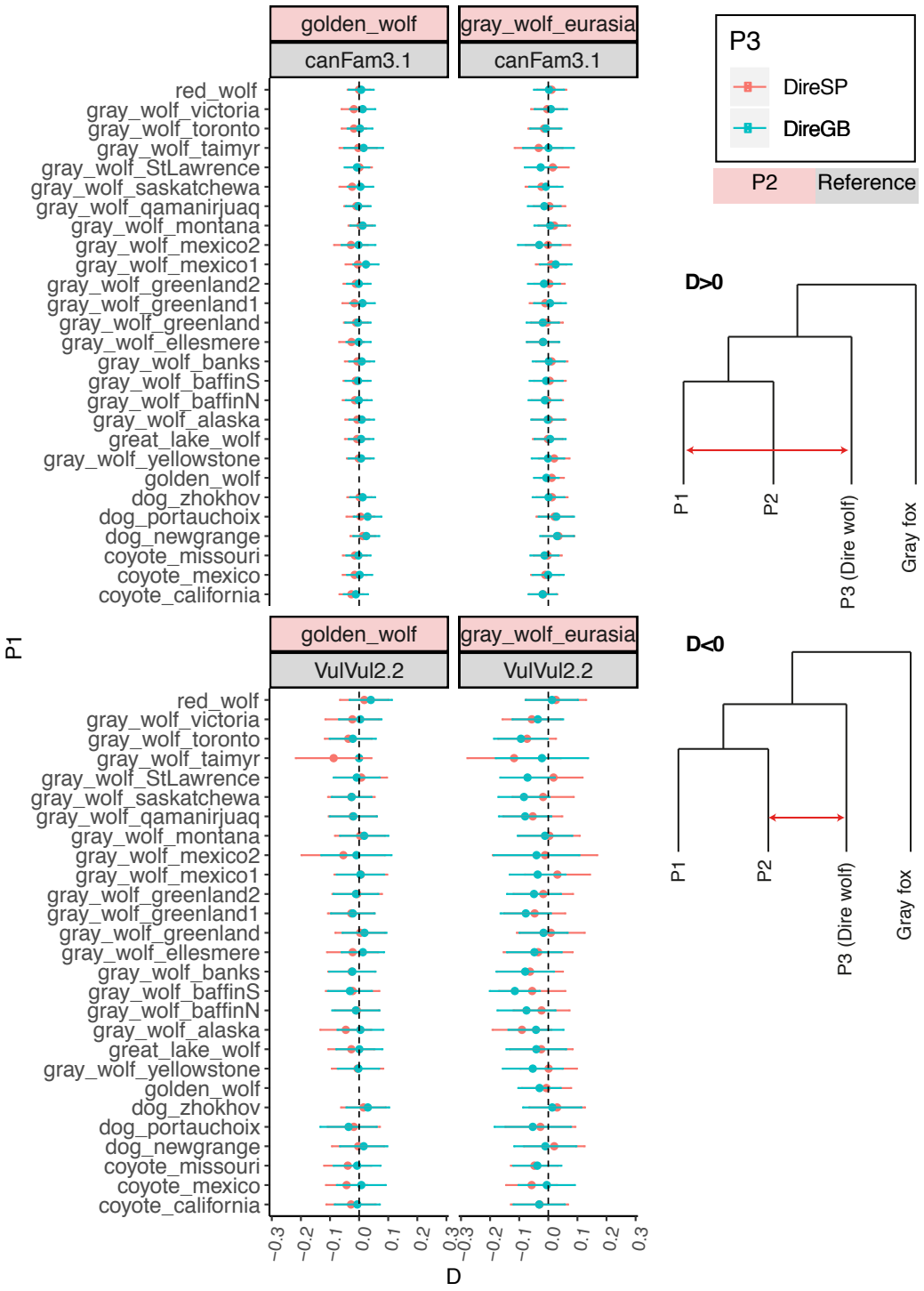
1216
1217 **Supplementary Figure 15. Maximum likelihood trees built using ascertainment correction**
1218 **as implemented in RAxML (based on pre-ascertained SNPs in modern canids; see section**
1219 ***supermatrix*).** Samples in red represent dire wolves.
1220



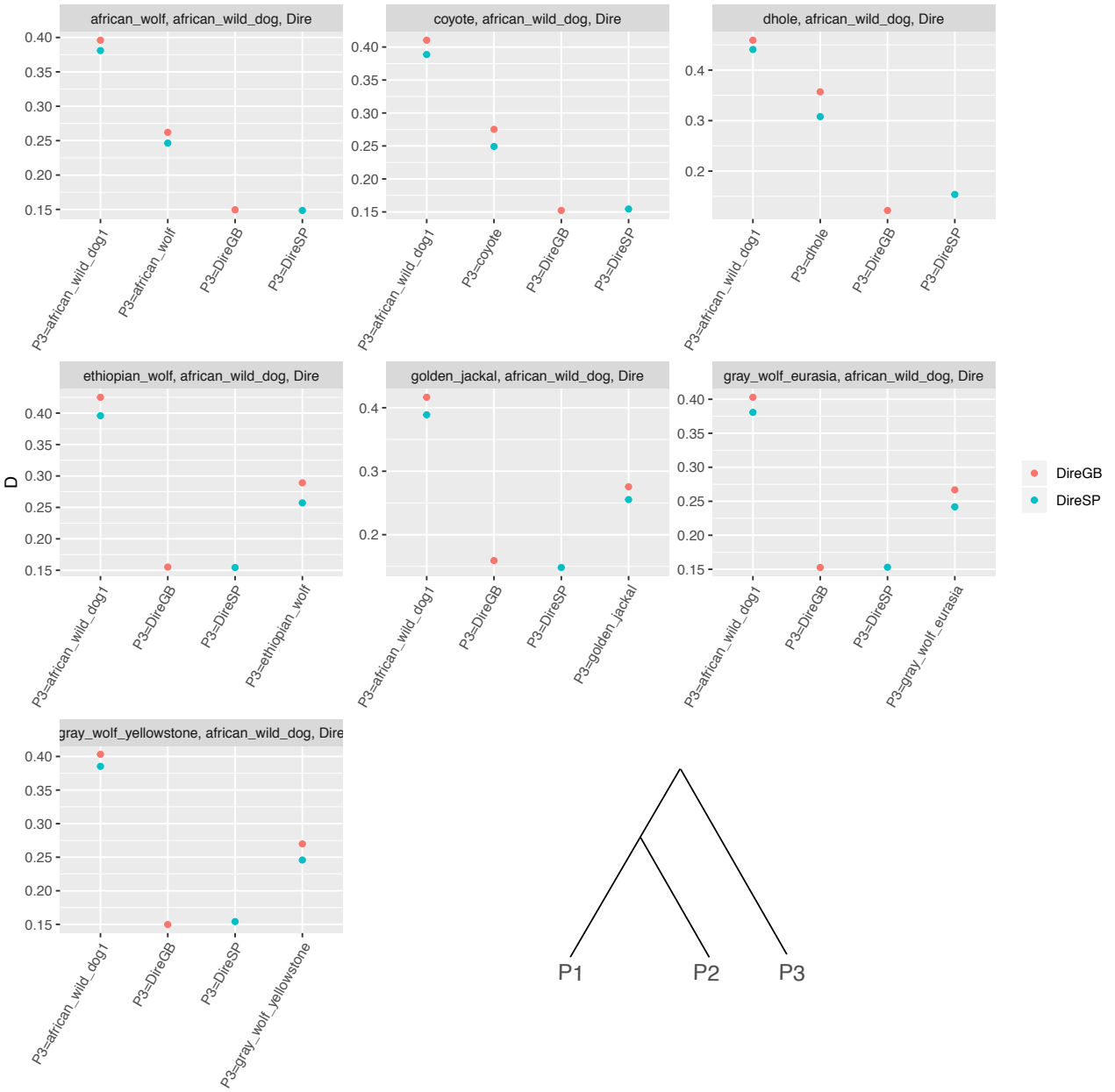
1221
1222 **Supplementary Figure 16. Discordance visualisation of maximum likelihood trees built in**
1223 **5Mb window size, and using the ascertainment ascertainment correction as implemented in**
1224 **RAxML.** Clades and/or alternative topologies displayed in Supplementary Fig. 14 are listed on the x-
1225 axis. The y-axis depicts the proportion of individual trees that strongly support (>85% bootstrap support;
1226 dark blue), weakly support (□85% bootstrap support; light blue), weakly reject (best tree favours alternative
1227 arrangement with □85% bootstrap support; orange), or strongly reject that clade/topology (best tree favours
1228 alternative arrangement with >85% bootstrap support; red).
1229



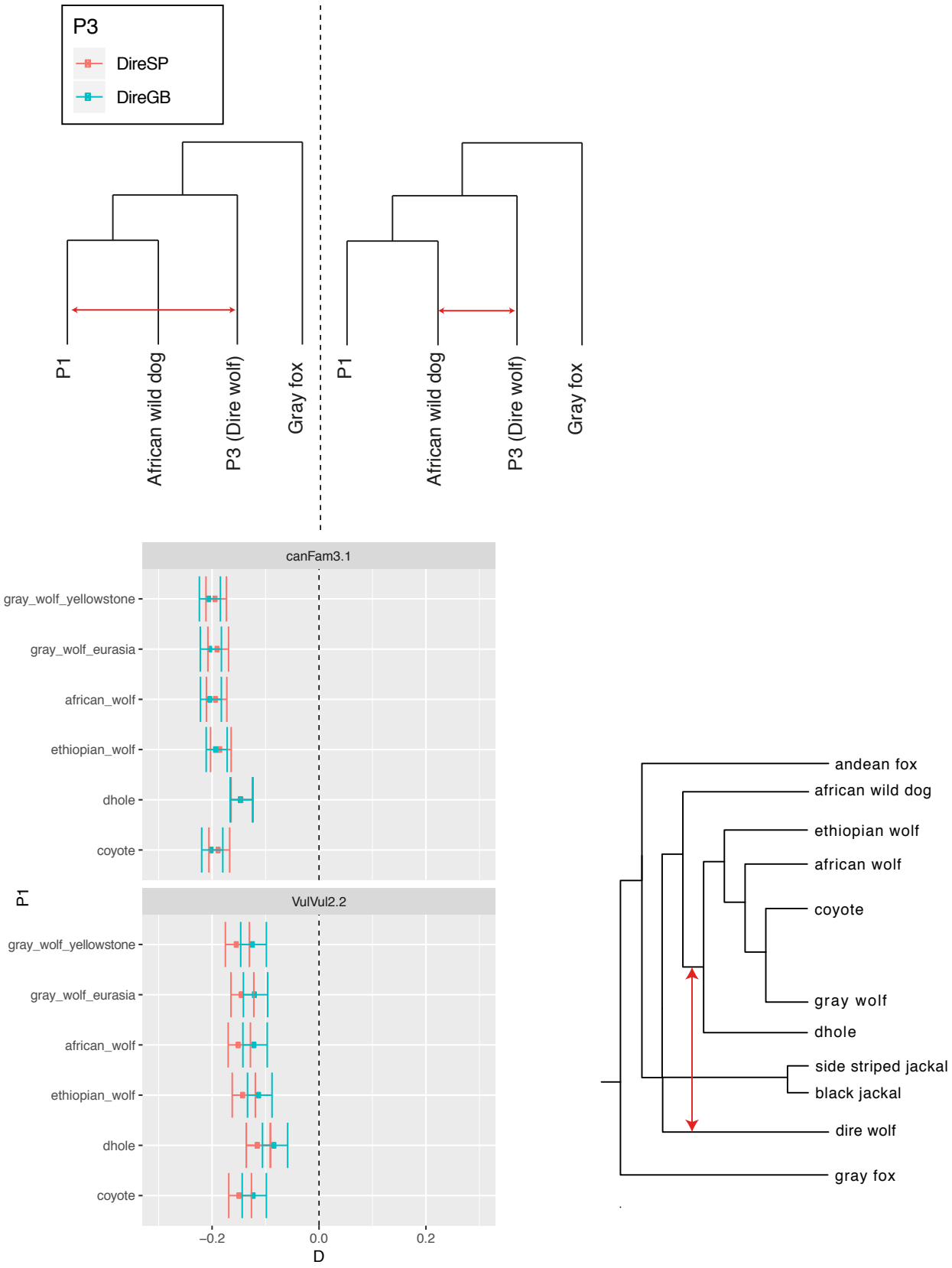
Supplementary Figure 17. Different of age node across different error rate. Age of node (y-axis), under different data sets (x-axis; Corr=corrected data; Sim.corr=simulated data based on corrected tree; Uncorr=uncorrected data; Sim.err0.5-5=simulated with 0.5%-5% error on the dire wolf branch), representing the ancestor of A. dire wolf and extant canids B. two dire wolves. Dashed line represents 95% HPDI. This figure demonstrates that the error either introduced via simulation (Sim.errX) or due to deamination (Uncorr) on the branch of the dire wolf 1) does not affect the age of the node representing the ancestor of the dire wolf and extant canids (A.) 2) inflates the age of the node representing the ancestor of two dire wolves (B.). We note that the real data (Uncorr) is most similar to simulated data with 0.5% singleton in the dire wolf sequence (Sim.err0.5).



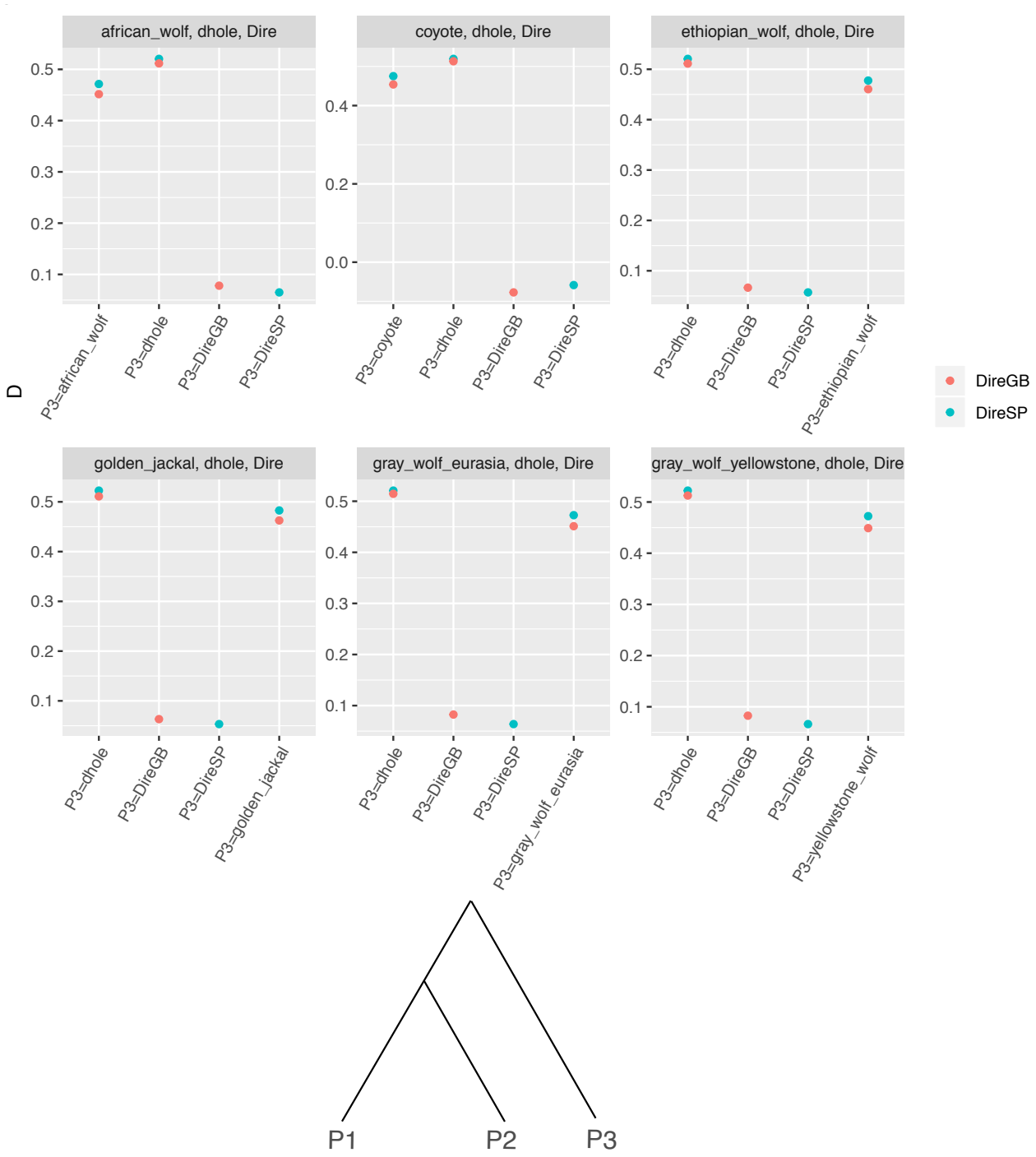
Supplementary Figure 18. Results of D statistics used to assess the possibility of gene flow between the dire wolf and extant North American canids. Each dot represents the mean D calculated along the genome and the error bar represents 3 standard deviations. D-statistics were computed using two reference genomes (dog: canFam3.1 and redfox: VulVul2.2), as well as two different P2 taxa: a Eurasian wolf (from Portugal; see Supplementary Data 1). We computed each possible combination of D(gray_fox; dire_wolf (P1, P2)) where P2 is either a Eurasian wolf (from Portugal; see Supplementary Data 1) or an African wolf (*Canis anthus* from Kenya; see Supplementary Data 1) and P3 a North American canid genome (see Supplementary Data 13). These plots show that the dire wolf genomes do not share significantly more derived alleles with extant North American canids compared to African wolves or Eurasian wolves (values of D not significantly different to zero), suggesting that no hybridization occurred between the dire wolf and the ancestor of extant North American canids (Supplementary Data 14).



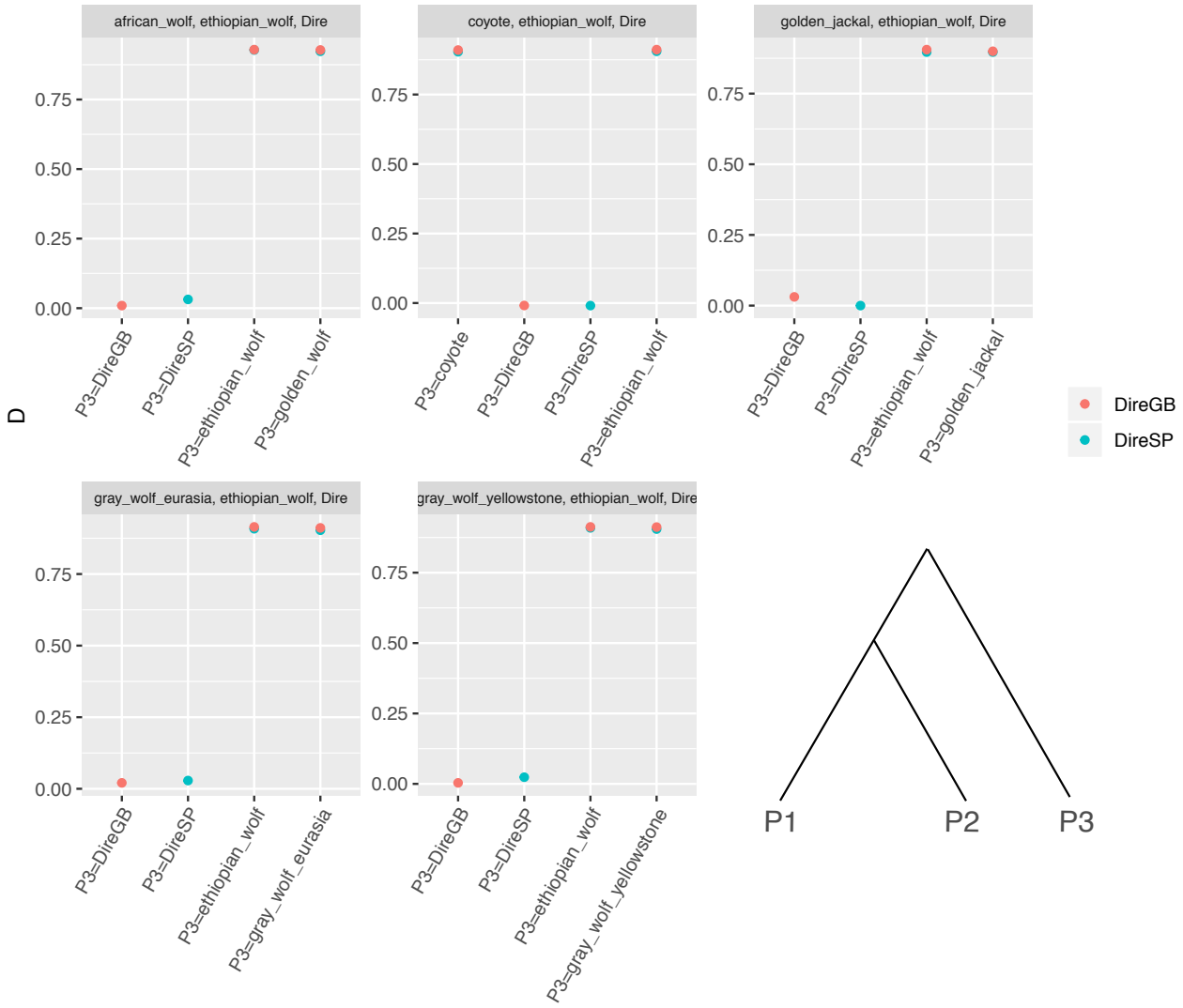
Supplementary Figure 19. Results of D-statistics to assess the taxonomic relationship between dire wolves, african wild dogs, dhole and wolf like canids. Each panel represents a different combination of three genomes (see tree on the bottom right corner). Genomes used as P3 are labelled on the x axis. Values close to 0 indicate that the species on the x axis (P3) share roughly the same number of derived alleles with P1 and P2, while positive values indicate that P3 shares more derived alleles with P1 or P2. A value close to 0 thus indicates that the genome in P3 represents an outgroup to P1 and P2. These results show that wolf-like canids and dhole share more derived alleles with the African wild dog than with dire wolves. This is consistent with our phylogenetic analysis indicating that the dire wolf represents an outgroup to these lineages.



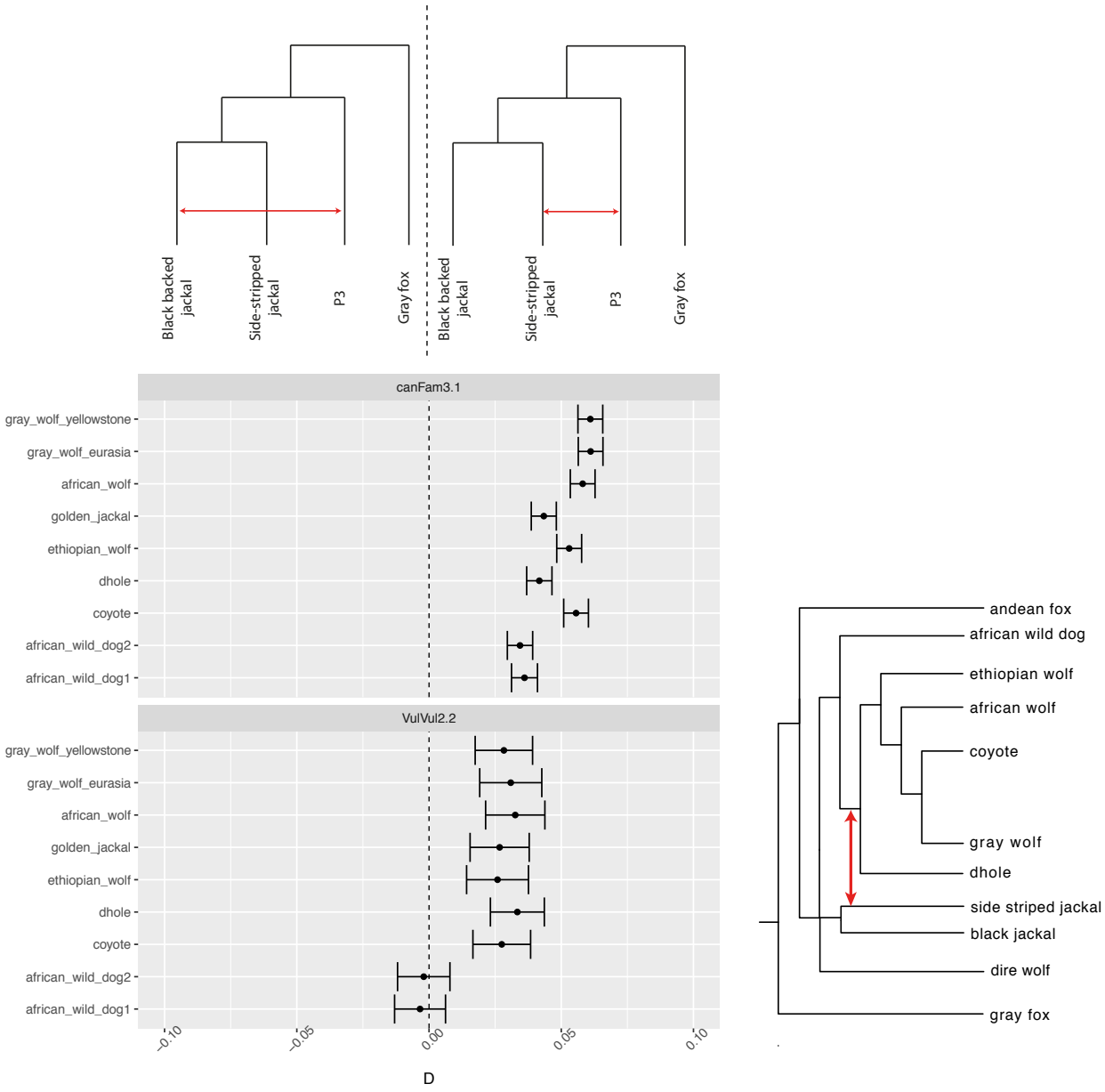
Supplementary Figure 20. Results of the D-statistics of the form D(gray_fox; dire_wolf (African wild dog, P1)). P1 on the y-axis represents any wolf like canids (coyotes, wolves etc.), or the dhole. Each dot represents the mean D calculated along the genome and the error bar represents 3 standard deviations. The analysis was run on both Dog (canFam3.1) and Red fox assembly (VulVul2.2). This analysis suggests that there was an admixture event between the dire wolf lineage and the ancestor of the dhole, wolves and coyotes.



pplementary Figure 21. Results of D-statistics to assess the taxonomic relationship between dire wolves, dhole and wolf like canids. Each panel represents a different combination of three genomes (see tree on the bottom right corner). Genomes used as P3 are labelled on the x axis. Values close to 0 indicate that the species on the x axis (P3) share roughly the same number of derived alleles with P1 and P2, while positive values indicate that P3 shares more derived alleles with P1 or P2. A value close to 0 thus indicates that the genome in P3 represents an outgroup to P1 and P2. These results show that wolf-like canids share more derived alleles with the dhole than with dire wolves. This is consistent with our phylogenetic analysis indicating that the dire wolf represents an outgroup to these lineages.



Supplementary Figure 22. Results of D-statistics to assess the taxonomic relationship between dire wolves and wolf like canids. Each panel represents a different combination of three genomes (see tree on the bottom right corner). Genomes used as P3 are labelled on the x axis. Values close to 0 indicate that the species on the x axis (P3) share roughly the same number of derived alleles with P1 and P2, while positive values indicate that P3 shares more derived alleles with P1 or P2. A value close to 0 thus indicates that the genome in P3 represents an outgroup to P1 and P2. These results show that wolf-like canids share more derived allele with the Ethiopian wolf than with dire wolves. This is consistent with our phylogenetic analysis indicating that the dire wolf represents an outgroup to these lineages.



Supplementary Figure 23. Results of the D-statistics of the form D(gray_fox; P3(side-striped jackal, black-backed jackal)). P3 label can be found on the x-axis. Each dot represents the mean D calculated along the genome and the error bar represents 3 standard deviations. The analysis was run on both Dog (canFam3.1) and Red fox assembly (VulVul2.2). This analysis suggests that there was an admixture event between the ancestor of the wolf like canids (dhole, wolves, ethiopian wolves etc) and the side-striped jackal.

Supplementary Data

Supplementary Data 1. Information about ancient samples sequenced in this study including provenance, age (radiocarbon or stratigraphic information), and sequencing statistics (endogeneous content etc.). * represents duplicate samples.

Supplementary Data 2. Locations and age of dire wolf remains in the Americas, obtained from [10,63,108–114]

Supplementary Data 3. Pairwise CS testing . All p values have been adjusted for multiple comparisons using the FDR method. The upper triangles are the first molar results. Lower triangles are mandibular results.

Supplementary Data 4. Pairwise morphological variance comparisons by species. All p values have been adjusted for multiple comparisons using the FDR method. The upper triangles are the first molar results. Lower triangles are mandibular results. The Procrustes variance scores are listed at the side for the first molar and below for the mandible.

Supplementary Data 5. Pairwise Procrustes ANOVA results by species. All p values have been adjusted for multiple comparisons using the FDR method. The upper triangles are the results from the first molar dataset. The lower triangles are results from the mandibular dataset. The Ethiopian wolf sample size (2) was too small for testing.

Supplementary Data 6. First molar pairwise CCV results from LDA optimised with reduced dimensionality. Upper triangle are numbers of PCs used to achieve maximum discrimination of balanced groups. Lower triangle contains the maximum CCV result achieved. Ethiopian wolf sample was too small for LDA procedures.

Supplementary Data 7. Mandibular pairwise CCV results from LDA optimised with reduced dimensionality. Upper triangle are numbers of PCs used to achieve maximum discrimination of balanced groups. Lower triangle contains the maximum CCV result achieved. Ethiopian wolf sample was too small for LDA procedures.

Supplementary Data 8. Allometry corrected first molar pairwise CCV results from LDA optimised with reduced dimensionality. Upper triangle are numbers of PCs used to achieve maximum discrimination of balanced groups. Lower triangle contains the maximum CCV result achieved. Ethiopian wolf sample was too small for LDA procedures.

Supplementary Data 9. Allometry corrected mandibular pairwise CCV results from LDA optimised with reduced dimensionality. Upper triangle are numbers of PCs used to achieve maximum discrimination of balanced groups. Lower triangle contains the maximum CCV result achieved. Ethiopian wolf sample was too small for LDA procedures.

Supplementary Data 10. Incongruence scores from full shape. Upper triangles are scores calculated from first molar morphological data. Lower triangle contains incongruence scores calculated from mandibular data.

Supplementary Data 11. Incongruence scores from allometry corrected shape. Upper triangles are scores calculated from first molar morphological data. Lower triangle contains incongruence scores calculated from mandibular data.

Supplementary Data 12. Incongruence scores from combined mandibular and first molar datasets. Upper triangles are scores calculated from allometry corrected morphological data. Lower triangle incongruence scores are calculated from full shape.

Supplementary Data 13: Tables containing information (coverage, accession etc.) of modern whole genomes used in this study and the additional genomes used for the D-statistics.

Supplementary Data 14. D-statistics results of the form D(Gray fox, Dire wolf [P3]; North American wolf-like canid [P1], Eurasian/African wolf-like canid [P2]). These results are plotted in Supplementary Fig. 18.

Supplementary Data 15. D-statistics results of the form D(gray_fox, dire_wolf [P3]; wolf-like canid [P1], African wild dog [P2]). These results are plotted in Supplementary Fig. 20.

Supplementary Data 16. D-statistics results of the form D(gray_fox, dire_wolf [P3]; wolves/coyotes [P1], Ethiopian wolf [P2]).

Supplementary Tables

Supplementary Table 1. Sample sizes for geometric morphometric datasets

Species	Mandible	Mandible M1
African wolf	44	43
African hunting dog	10	10
Andean fox	8	8
Black-backed jackal	18	16
Coyote	65	67
Dhole	6	6
Dire wolf	16	8
Ethiopian wolf	2	2
Golden jackal	27	26
Gray fox	10	10
Side-striped jackal	24	24
Gray wolf	607	541

Supplementary Table 2. Species used for the mitochondrial phylogenies and their accession numbers. Species with accession numbers in bold were used as well in the nuclear analyses.

	Accession number
Gray fox (*Urocyon cinereoargenteus*)	**SRS1937014**
Andean fox (*Lycalopex culpaeus*)	**SRS523207**
Yellowstone wolf (*Canis lupus*)	**SRS661496**
Eurasia wolf (*Canis lupus*)	**SRS661492**
Dog (*Canis lupus familiaris*)	canFam3.1

Arabian wolf (*Canis lupus arabs*)	DQ480507
Coyote (*Canis latrans*)	**SRS661477**
Red wolf (*Canis rufus*)	SRS661493
Great Lakes wolf (*Canis lupus lycaon*)	SRS661486
African wolf (*Canis anthus*)	**ERS3334821**
Ethiopian wolf (*Canis simensis*)	**SRS3929738**
Black-backed jackal (*Canis mesomelas*)	**ERS3216353**
African wild dog 1 (*Lycaon pictus*)	**SRR2971425**
African wild dog 2 (*Lycaon pictus*)	**SRR2971441**
Side-striped jackal (*Canis adustu*s)	**ERS3216352**
Golden jackal (*Canis aureus*)	**SRS1025419**
Dhole (*Cuon alpinus*)	**SRS3929739**

Supplementary Table 3. Amount of missing data (in %) for each direwolf specimen. Genes with 100% of missing data in at least one specimen are in bold.

	DireGB	DireSP	DireNTC	DireGWC	DireAFR
ATP6 (681bp)	77%	62%	97%	63%	**100%**
ATP8 (204bp)	72%	42%	**100%**	42%	75%
COX1 (1545bp)	55%	65%	91%	65%	67%
COX2 (684bp)	37%	24%	68%	4%	68%
COX3 (784bp)	95%	91%	97%	83%	98%
Cytb(1140bp)	76%	83%	91%	60%	97%
ND1(951bp)	70%	78%	91%	71%	96%
ND2(1042bp)	62%	60%	88%	63%	**100%**
ND3 (346bp)	86%	77%	**100%**	74%	71%
ND4 (1378bp)	68%	60%	87%	49%	64%
ND4L (297bp)	74%	59%	87%	37%	**100%**
ND5 (1821bp)	66%	70%	87%	60%	78%
ND6 (528bp)	65%	50%	73%	55%	88%

S12 (954bp)	33%	34%	69%	7%	66%
S16 (1580bp)	38%	35%	74%	11%	64%
% missing data	61.8%	60.2%	85.5%	49.3%	79.7%

Supplementary Table 4. Best model for each partition in the mitochondrial genome as inferred using partitionFinder2 [79].

Best Model	Sites	Partition names
TRN+I+G	4692	ATP8_pos1, ND2_pos1, CYTB_pos1, ATP8_pos2, S16_pos1, S12_pos1, ND5_pos1, ND4_pos1, ATP6_pos1
HKY+I+G	3733	COX1_pos2, COX2_pos2, RCND6_pos2, COX3_pos2, ATP6_pos2, ND1_pos2, ND4_pos2, CYTB_pos2, ND5_pos2, ND2_pos2, ND4L_pos2, ND3_pos2
TIM+I+G	3110	ND1_pos3, CYTB_pos3, ND2_pos3, COX3_pos3, ATP6_pos3, COX2_pos3, ND4_pos3, ND5_pos3, ATP8_pos3, ND4L_pos3, ND3_pos3
TRNEF+I+G	1715	RCND6_pos1, ND3_pos1, COX1_pos1, ND4L_pos1, COX2_pos1, COX3_pos1, ND1_pos1
TRN+G	515	COX1_pos3
HKY+G	176	RCND6_pos3

Supplementary Table 5. Total size of alignment, number of loci and correction factor used in the supermatrix analysis using RAxML[95] (Supplementary Fig. 3).

Sample	Total Size	Number of loci	Correction factor
DireAFR (AJ66)	139817	1395	0.77
DireGWC (RW001)	67487	815	0.803
DireSP (ACAD1735)	28646071	477919	0.728
DireGB (ACAD18742)	28593900	595622	0.78
Combined (DireSP+DireGB)	412476	7295	NA

Supplementary Table 6. Number of bins, and average number of coverage for different bin sizes used in the DISCOVISTA analysis presented in Supplementary Fig. 6.

Sample	Average number of base covered per bin	Total coverage	Total number of bins with coverage	Total number of bins with coverage >2000bp	Bin size
DireAFR	154	901979	5839	2	500kb
DireAFR	363	946090	2606	4	1Mb
DireAFR	2063	1238090	600	305	5Mb
DireGWC	47	272110	5839	0	500kb
DireGWC	123	321804	2606	0	1Mb
DireGWC	824	494514	600	4	5Mb
DireSP	40367	235704849	5839	5517	500kb
DireSP	80380	209469949	2606	2463	1Mb
DireSP	390380	234228275	600	569	5Mb
DireGB	42599	248734410	5839	5516	500kb
DireGB	84820	221040684	2606	2465	1Mb
DireGB	411956	247173801	600	569	5Mb
DireNTC	1159	6770271	5839	134	500kb
DireNTC	2335	6084926	2606	1909	1Mb
DireNTC	11440	6864245	600	558	5Mb

Supplementary Table 7. **Alignment length and dire wolf sample data missingness for RAxML supermatrix analyses.**

Set	Total alignment	Dire wolf	Percentage dire wolf

	length		data missingness %
All samples	12,994,077	DireSP Dire GB DireNTC DireGWC DireAFR	79.3 78.5 99.4 99.7 99.5
Modern only	12,993,672	N/A	N/A
Modern plus DireSP	**12,993,899**	**DireSP**	**79.3**
Modern plus DireGB	**12,993,861**	**DireGB**	**78.5**
Modern plus DireSP and DireGB	12,994,060	DireSP Dire GB	79.3 78.5
Modern plus DireNTC	12,993,683	DireNTC	99.4
Modern plus DireGWC	12,993,674	DireGWC	99.7
Modern plus DireAFR	12,993,679	DireAFR	99.5

Supplementary Table 8. **Number of variable sites and dire wolf sample data missingness for the 422 5mb windows analysed using RAxML for visualisation of phylogenetic discordance with Discovista.**

Set	**Variable sites per window (mean ± standard deviation**	**Dire wolf**	**Percentage dire wolf data missingness % (mean ± standard deviation)**
All samples	29,438.9 ± 3,062.3	DireSP Dire GB DireNTC DireGWC DireAFR	79.3 ± 2.5 78.4 ± 1.7 99.4 ± 0.1 99.7 ± 0.1 99.5 ± 0.2
Modern only	29,438.0 ± 3,062.1	N/A	N/A
Modern plus DireSP	**29,438.5 ± 3,062.2**	**DireSP**	**79.4 ± 2.5**
Modern plus DireGB	**29,438.4 ± 3,062.2**	**DireGB**	**78.4 ± 1.7**
Modern plus DireSP and DireGB	29,438.8 ± 3,062.3	DireSP Dire GB	79.3 ± 2.5 78.4 ± 1.7
Modern plus DireNTC	29,438.0 ± 3,062.1	DireNTC	99.4 ± 0.1
Modern plus DireGWC	29,438.0 ± 3,062.1	DireGWC	99.7 ± 0.1
Modern plus DireAFR	29,438.0 ± 3,062.1	DireAFR	99.5 ± 0.2

Supplementary Table 9. Bayesian model selection of clock and tree topology for the molecular data sets.

Data[a]	Model[b]	log m*L* ± S.E[c]	Pr (CI-2.5%, CI-97.5%)[d]	Bayes factor
ALN1 (1 dw)	**Topo A + GBM + BD prior**	**-176267.70 ± 0.04**	**0.942 (0.935, 0.948)**	5.58
	Topo A + ILN + BD prior	-176270.50 ± 0.04	0.058 (0.052, 0.065)	
	Topo B + GBM + BD prior	-176282.10 ± 0.04	$5.42 \cdot 10^{-7}$ ($4.85 \cdot 10^{-7}$, $6.03 \cdot 10^{-7}$)	
	Topo B + ILN + BD prior	-176284.30 ± 0.05	$5.55 \cdot 10^{-8}$ ($4.92 \cdot 10^{-8}$, $6.22 \cdot 10^{-8}$)	
ALN2 (2 dws)	**Topo A + GBM + BD (priorD)**	**-178325.40 ± 0.05**	**0.969 (0.965, 0.972)**	6.87
	Topo A + ILN + BD prior	-178328.90 ± 0.05	0.031 (0.028, 0.035)	
	Topo B + GBM + BD prior	-178341.30 ± 0.05	$1.22 \cdot 10^{-7}$ ($1.08 \cdot 10^{-7}$, $1.37 \cdot 10^{-7}$)	
	Topo B + ILN + BD prior	-178344.70 ± 0.05	$3.99 \cdot 10^{-9}$ ($3.51 \cdot 10^{-9}$, $4.50 \cdot 10^{-9}$)	

[a] ALN*X*: molecular alignments with one direwolf (ALN1) or two direwolf specimens (ALN2).

[b] Topo A: tree topology that supports “dire/out” (Supplementary Fig. 6), Topo B: tree topology that supports “dire/in” (Supplementary Fig. 6), GBM: autocorrelated-rates model, ILN: independent log-normal rates model, BD: birth-death prior.

[c] log mL ± S.E : log marginal likelihood ± standard error; calculated with the mcmc3r R package.

[d] Pr (CI-2.5%, CI-97.5%): posterior probabilities for the model and the corresponding equal trail bootstrap confidence intervals; calculated with the mcmc3r R package.

Supplementary Table 10. Priors used for the Bayesian divergence times estimation and Bayesian model selection analyses.

Analysis	Prior on rates[a]	Prior on root age[b]	Prior on Andean fox – Canini split
Divergence times	Γ(2,3040)	U(10.3, 20.0)	U(4.9, 10.3)
Bayes factors	Γ(2,200)	U(0.999, 1.001)	U(4.9, 10.3)

[a] The prior on the rates used is a diffuse gamma distribution regardless of the analysis and the data. Therefore, the shape parameter is always α = 2. The scale parameter β changes according to the analysis performed. For divergence times estimation analyses with both corrected and uncorrected molecular alignments, $b_{root\text{-}tips}$ = 0.0093 ~ 0.01 and t_{root} = 15.2 Ma, thus $\beta = (\alpha \times t_{root}) / b_{root\text{-}tips}$ = (2 x 15.2) / 0.01 = 3040. For Bayes factors analyses, we fix t_{root}= 1 Ma because we are not interested in estimating absolute divergence times, only in selecting the best fitting-model. Therefore, β = (2 x 1) / 0.01 = 200.

[b] The prior on the root age is required when using MCMCtree. For divergence time estimation, we use a fossil calibration on the root ranging from the earlier part of the Miocene until the end of the Clarendonian, U(10.3,20.0). For the Bayes factors analyses, we cannot use a fossil calibration on the root as we are not interested in obtaining absolute divergence times. Therefore, as t_{root} = 1 Ma, we construct a narrow uniform distribution with soft bounds U(0.999,1.001).

Supplementary Table 11. Estimated mean divergence times of four external nodes for the molecular alignments using the model selected according to the Bayes factors analysis: autocorrelated-rates model and topology “A”. Results for both real and simulated molecular data have been included. Note that, to simulate aDNA deamination, alignments required to have at least one direwolf sequence present; thus posterior divergence times for alignments without direwolf sequences were only estimated for the corrected and uncorrected real data sets.

Tree node + Data type	**Estimated divergence times for the alignment without direwolf specimens**	**Estimated divergence times for the alignment with 1 direwolf specimen**	**Estimated divergence times for the alignment with 2 direwolf specimens**
Root.corr	**16.203 (10.415,22.777)**	**15.979 (10.369,22.740)**	**15.916 (10.366,22.722)**
Root.simcorr	**-**	16.008 (10.407,22.739)	15.813 (10.286,22.738)
Root.uncorr	16.161 (10.413,22.764)	15.816 (10.255,22.742)	15.762 (10.274,22.751)
Root.err0.5	**-**	15.921 (10.304,22.793)	15.650 (10.227,22.737)
Root.err1	-	15.835 (10.245,22.803)	15.579 (10.231,22.721)
Root.err3	-	15.633 (10.240,22.735)	15.470 (10.198,22.672)
Root.err5	-	15.578 (10.210,22.726)	15.274 (10.180,22.602)
Vulpini-Canini.corr	**6.835 (4.825,10.012)**	**6.771 (4.821,9.972)**	**6.790 (4.819,9.969)**
Vulpini-Canini.simcorr	**-**	6.724 (4.809,9.936)	6.921 (4.826,10.077)
Vulpini-Canini.uncorr	6.824 (4.828,9.984)	6.993 (4.834,10.133)	6.986 (4.826,10.103)
Vulpini-Canini.err0.5	**-**	6.928 (4.828,10.078)	7.053 (4.832,10.135)
Vulpini-Canini.err1	-	7.048 (4.840,10.147)	7.094 (4.846,10.163)
Vulpini-Canini.err3	-	7.196 (4.855,10.182)	7.212 (4.863,10.199)
Vulpini-Canini.err5	-	7.326 (4.882,10.228)	7.363 (4.889,10.235)
Dw-(jk+rest).corr	**-**	**5.694 (3.963,8.452)**	**5.700 (3.967,8.419)**
Dw-(jk+rest).simcorr	-	5.609 (3.938,8.329)	5.882 (4.003,8.649)
Dw-(jk+rest).uncorr	-	6.063 (4.075,8.919)	5.988 (4.026,8.798)
Dw-(jk+rest).err0.5	-	5.828 (3.939,8.585)	6.042 (4.029,8.833)

Dw-(jk+rest).err1	-	5.950 (3.950,8.745)	6.096 (4.034,8.914)
Dw-(jk+rest).err3	-	6.126 (3.983,8.929)	6.255 (4.087,9.065)
Dw-(jk+rest).err5	-	6.280 (4.032,9.041)	6.414 (4.150,9.178)
jk-rest.corr	**5.207 (3.551,7.731)**	**5.090 (3.509,7.604)**	**5.101 (3.519,7.567)**
jk-rest.simcorr	**-**	5.024 (3.502,7.487)	5.267 (3.537,7.813)
jk-rest.uncorr	5.201 (3.555,7.714)	5.558 (3.672,8.239)	5.430 (3.592,8.051)
jk-rest.err0.5	**-**	5.285 (3.528,7.839)	5.465 (3.578,8.081)
jk-rest.err1	-	5.425 (3.547,8.046)	5.540 (3.599,8.191)
jk-rest.err3	-	5.623 (3.583,8.299)	5.700 (3.623,8.402)
jk-rest.err5	-	5.760 (3.611,8.422)	5.850 (3.678,8.543)
dw1-dw2.corr	-	-	**2.339 (1.409,3.713)**
dw1-dw2.simcorr	-	-	2.265 (1.143,3.936)
dw1-dw2.uncorr	-	-	3.804 (2.201,6.030)
dw1-dw2.err0.5	-	-	3.519 (1.897,5.725)
dw1-dw2.err1	-	-	3.931 (2.135,6.305)
dw1-dw2.err3	-	-	4.582 (2.529,7.175)
dw1-dw2.err5	-	-	4.970 (2.758,7.639)

*- Abbreviations: **corr**: corrected alignment (real data), **uncorr**: uncorrected alignment (real data), **simcorr**: simulated alignment using the corrected alignment with real data, **err0.5**: simulated alignment with 0.5% added error, **err1**: simulated alignment with 1% added error, **err3**: simulated alignment with 3% added error, **err5**: simulated alignment with 5% added error, Vulpini-Canini: Divergence between Vulpini and Canini, **Dw-(jk+rest)**: Divergence between the direwolf and the clade with the split between jackals and the rest of canids, **jk-rest**: Divergence between the jackals and the rest of canids. **dw1-dw2**: Divergence between the two dire wolf specimens.

References

1. Martin, L. D. & Miles Gilbert, B. Excavations at Natural Trap Cave. (1978).
2. Kohn, M. J. & McKay, M. P. Paleoecology of late Pleistocene–Holocene faunas of eastern and central Wyoming, USA, with implications for LGM climate models. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **326-328**, 42–53 (2012).
3. Meachen, J. A., Brannick, A. L. & Fry, T. J. Extinct Beringian wolf morphotype found in the continental U.S. has implications for wolf migration and evolution. *Ecol. Evol.* **6**, 3430–3438 (2016).
4. Scott, W. E., Pierce, K. L., Bradbury, J. P. & Forester, R. M. Revised Quaternary stratigraphy and chronology in the American Falls area, southeastern Idaho. *Cenozoic Geology of Idaho: Idaho Bureau of Mines and Geology Bulletin* **26**, 581–595 (1982).
5. Tankersley, K. B. Sheriden: A Clovis cave site in eastern North America. *Geoarchaeology* **12**, 713–724 (1997).
6. Ford, K. M., Bair, A. R. & Holman, J. A. Late Pleistocene Fishes from Sheriden Pit, Northwestern Ohio. *Mich. Acad.* **28**, 135–145 (1996).
7. McDonald, H. G. The late Pleistocene vertebrate fauna in Ohio: coinhabitants with Ohio's Paleoindians. *The First Discovery of America: Archaeological Evidence of the Early Inhabitants of the Ohio Area, The Ohio Archaeological Council, Columbus* 23–39 (1994).
8. Holman, J. A. Amphibians and reptiles from the Pleistocene (late Wisconsinan) of Sheriden Pit Cave, northwestern Ohio. *Mich. Acad.* **29**, 1–20 (1997).
9. Bills, T. M. & McDonald, H. G. Fauna from Late-Pleistocene Sediments of the Sheriden Cave Site (33WY252), Wyandot County, Ohio. *Current Research in the Pleistocene* **15**, 101–103 (1998).
10. Schubert, B. W. & Wallace, S. C. Late Pleistocene giant short-faced bears, mammoths, and large carcass scavenging in the Saltville Valley of Virginia, USA. *Boreas* **38**, 482–492 (2009).
11. Nye, A. S. Pleistocene Peccaries from Guy Wilson Cave, Sullivan County, Tennessee. (East Tennessee State University, 2007).
12. Stock, C. & Harris, J. M. Rancho La Brea: A Record of Pleistocene Life in California. Science Series No. 37. *Natural History Museum of Los Angeles County, Los Angeles, CA* (1992).

13. Rohlf, J. PSDIG, version 1.31. SUNY Morphometrics, Stony Brook, New York. *Book PSDIG, version* **1**, (2001).
14. Core Team, R. & Others. R: A language and environment for statistical computing. *R Foundation for statistical computing, Vienna* (2013).
15. Dryden, I. L. & Mardia, K. V. *Statistical Shape Analysis*. (Wiley, 1998).
16. Goodall, C. Procrustes Methods in the Statistical Analysis of Shape. *J. R. Stat. Soc. Series B Stat. Methodol.* **53**, 285–321 (1991).
17. BOOKSTEIN & L, F. Applying landmark methods to biological outline data. *Image Fusion and Shape Variability Techniques* (1996).
18. Schlager, S. Morpho and Rvcg--Shape Analysis in R: R-Packages for geometric morphometrics, shape analysis and surface manipulations. in *Statistical shape and deformation analysis* 217–256 (Elsevier, 2017).
19. Dryden, I. L. & Dryden, M. I. Shapes package. *Vienna, Austria: R Foundation for Statistical Computing* (2012).
20. Foote, M. Contributions of individual taxa to overall morphological disparity. *Paleobiology* **19**, 403–419 (1993).
21. Anderson, M. J. Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* **62**, 245–253 (2006).
22. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* vol. 57 289–300 (1995).
23. Adams, D. C., Collyer, M. L., Kaliontzopoulou, A. & Sherratt, E. Geomorph: software for geometric morphometric analyses. R package version 3.0. 6. (2016).
24. Evin, A. *et al.* The long and winding road: identifying pig domestication through molar size and shape. *J. Archaeol. Sci.* **40**, 735–743 (2013).
25. Baylac, M. & Frieß, M. Fourier Descriptors, Procrustes Superimposition, and Data Dimensionality: An Example of Cranial Shape Analysis in Modern Human Populations. in *Modern Morphometrics in Physical Anthropology* (ed. Slice, D. E.) 145–165 (Springer US, 2005).

26. Meloro, C. & Raia, P. Cats and Dogs Down the Tree: The Tempo and Mode of Evolution in the Lower Carnassial of Fossil and Living Carnivora. *Evol. Biol.* **37**, 177–186 (2010).
27. Piras, P. *et al.* Ecological and functional correlates of molar shape variation in European populations of Arvicola (Arvicolinae, Rodentia). *Zoologischer Anzeiger - A Journal of Comparative Zoology* **251**, 335–343 (2012).
28. Felsenstein, J. Maximum-likelihood estimation of evolutionary trees from continuous characters. *Am. J. Hum. Genet.* **25**, 471–492 (1973).
29. Butler, M. A. & King, A. A. Phylogenetic Comparative Analysis: A Modeling Approach for Adaptive Evolution. *Am. Nat.* **164**, 683–695 (2004).
30. Harmon, L. J. *et al.* EARLY BURSTS OF BODY SIZE AND SHAPE EVOLUTION ARE RARE IN COMPARATIVE DATA. *Evolution* no–no (2010) doi:10.1111/j.1558-5646.2010.01025.x.
31. Pagel, M. Inferring the historical patterns of biological evolution. *Nature* **401**, 877–884 (1999).
32. Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E. & Challenger, W. GEIGER: investigating evolutionary radiations. *Bioinformatics* **24**, 129–131 (2008).
33. Blomberg, S. P., Garland, T., Jr & Ives, A. R. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* **57**, 717–745 (2003).
34. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* (2012).
35. Adams, D. C. A generalized K statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data. *Syst. Biol.* **63**, 685–697 (2014).
36. Klingenberg, C. P. & Gidaszewski, N. A. Testing and quantifying phylogenetic signals and homoplasy in morphometric data. *Syst. Biol.* **59**, 245–261 (2010).
37. Rueness, E. K. *et al.* The cryptic African wolf: Canis aureus lupaster is not a golden jackal and is not endemic to Egypt. *PLoS One* **6**, e16385 (2011).
38. Fuller, B. T., Harris, J. M., Farrell, A. B., Takeuchi, G. & Southon, J. Sample Preparation for Radiocarbon Dating and Isotopic Analysis of Bone from Rancho La Brea. *La Brea and beyond: The paleontology of asphalt-preserved biotas, ed. JM Harris. Natural History Museum of Los Angeles County, Science Series* 151–167 (2015).
39. Welker, F., Soressi, M., Rendu, W., Hublin, J.-J. & Collins, M. Using ZooMS to identify

fragmentary bone from the late Middle/Early Upper Palaeolithic sequence of Les Cottes, France. *J. Archaeol. Sci.* **54**, 279–286 (2015).

40. Welker, F. *et al.* Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. *Nature* **522**, 81–84 (2015).
41. Demarchi, B. *et al.* Protein sequences bound to mineral surfaces persist into deep time. *Elife* **5**, (2016).
42. van Doorn, N. L., Hollund, H. & Collins, M. J. A novel and non-destructive approach for ZooMS analysis: ammonium bicarbonate buffer extraction. *Archaeol. Anthropol. Sci.* **3**, 281–289 (2011).
43. Chambers, M. C. *et al.* A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920 (2012).
44. Ma, B. *et al.* PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **17**, 2337–2342 (2003).
45. Zhang, J. *et al.* PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol. Cell. Proteomics* **11**, M111.010587 (2012).
46. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
47. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
48. Dabney, J. *et al.* Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15758–15763 (2013).
49. Damgaard, P. B. *et al.* Improving access to endogenous DNA in ancient bones and teeth. *Sci. Rep.* **5**, 11184 (2015).
50. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, db.prot5448 (2010).
51. Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 5257 (2014).
52. Brotherton, P. *et al.* Neolithic mitochondrial haplogroup H genomes and the genetic origins of

Europeans. *Nat. Commun.* **4**, 1764 (2013).

53. Gansauge, M.-T. *et al.* Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic Acids Res.* **45**, e79 (2017).
54. Gamba, C. *et al.* Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. *Mol. Ecol. Resour.* **16**, 459–469 (2016).
55. Gansauge, M.-T. & Meyer, M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat. Protoc.* **8**, 737–748 (2013).
56. Ersmark, E., Orlando, L. & Sandoval-Castellanos, E. Population demography and genetic diversity in the Pleistocene cave lion. *Open* (2015).
57. Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).
58. Carøe, C. *et al.* Single-tube library preparation for degraded DNA. *Methods Ecol. Evol.* (2017).
59. Slon, V. *et al.* Mammalian mitochondrial capture, a tool for rapid screening of DNA preservation in faunal and undiagnostic remains, and its application to Middle Pleistocene specimens from Qesem Cave (Israel). *Quat. Int.* **398**, 210–218 (2016).
60. Nieves-Colón, M. A. *et al.* Comparison of two ancient DNA extraction protocols for skeletal remains from tropical environments. *Am. J. Phys. Anthropol.* (2018) doi:10.1002/ajpa.23472.
61. Sambrook, J. & Russell, D. W. Purification of nucleic acids by extraction with phenol:chloroform. *CSH Protoc.* **2006**, (2006).
62. Verscheure, S., Backeljau, T. & Desmyter, S. Dog mitochondrial genome sequencing to enhance dog mtDNA discrimination power in forensic casework. *Forensic Sci. Int. Genet.* **12**, 60–68 (2014).
63. Tedford, R. H., Wang, X. & Taylor, B. E. Phylogenetic Systematics of the North American Fossil Caninae (Carnivora: Canidae). *Bull. Am. Mus. Nat. Hist.* **325**, 1–218 (2009).
64. Maricic, T., Whitten, M. & Pääbo, S. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS One* **5**, e14004 (2010).
65. Maricic, T., Whitten, M. & Pääbo, S. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS One* **5**, e14004 (2010).

66. Gaubert, P. *et al.* 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 (2012).
67. Mak, S. S. T. *et al.* Erratum to: Comparative performance of the BGISEQ-500 vs Illumina HiSeq2500 sequencing platforms for palaeogenomic sequencing. *Gigascience* **7**, (2018).
68. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
69. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
70. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356 (2014).
71. Green, R. E. *et al.* A draft sequence of the Neandertal genome. *Science* **328**, 710–722 (2010).
72. Skoglund, P. *et al.* Genomic insights into the peopling of the Southwest Pacific. *Nature* **538**, 510–513 (2016).
73. Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
74. Lindgreen, S. AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res. Notes* **5**, 337 (2012).
75. Schubert, M. *et al.* Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics* **13**, 178 (2012).
76. Kircher, M. Analysis of high-throughput ancient DNA sequencing data. *Methods Mol. Biol.* **840**, 197–228 (2012).
77. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
78. Borowiec, M. L. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ* **4**, e1660 (2016).
79. Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T. & Calcott, B. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological

Phylogenetic Analyses. *Mol. Biol. Evol.* **34**, 772–773 (2017).

80. Miller, M. A., Pfeiffer, W. & Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. in *2010 Gateway Computing Environments Workshop (GCE)* 1–8 (2010).

81. Kvist, S. & Siddall, M. E. Phylogenomics of A nnelida revisited: a cladistic approach using genome-wide expressed sequence tag data mining and examining the effects of missing data. *Cladistics* **29**, 435–448 (2013).

82. Xi, Z., Liu, L. & Davis, C. C. The Impact of Missing Data on Species Tree Estimation. *Mol. Biol. Evol.* **33**, 838–860 (2016).

83. Gopalakrishnan, S. *et al.* Interspecific Gene Flow Shaped the Evolution of the Genus Canis. *Curr. Biol.* **28**, 3441–3449.e5 (2018).

84. Koepfli, K.-P. *et al.* Genome-wide Evidence Reveals that African and Eurasian Golden Jackals Are Distinct Species. *Curr. Biol.* **25**, 2158–2165 (2015).

85. vonHoldt, B. M. *et al.* 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 (2016).

86. Sinding, M.-H. S. *et al.* Population genomics of grey wolves and wolf-like canids in North America. *PLoS Genet.* **14**, e1007745 (2018).

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

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

89. Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A. & RoyChoudhury, A. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Mol. Biol. Evol.* **29**, 1917–1932 (2012).

90. Yang, Z. The BPP program for species tree estimation and species delimitation. *Current Zoology* vol. 61 854–865 (2015).

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

92. Sayyari, E., Whitfield, J. B. & Mirarab, S. DiscoVista: Interpretable visualizations of gene tree

discordance. *Mol. Phylogenet. Evol.* **122**, 110–115 (2018).

93. Li, H. *et al.* The Sequence Alignment / Map (SAM) Format and SAMtools 1000 Genome Project Data Processing Subgroup. *Data Processing* 1–2 (2009).

94. Tange, O. *GNU Parallel 2018*. (Lulu.com, 2018).

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

96. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).

97. Yang, Z. & Rannala, B. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* **23**, 212–226 (2006).

98. Alroy, J. New methods for quantifying macroevolutionary patterns and processes. *Paleobiology* **26**, 707–733 (2000).

99. Rambaut, A. & Grassly, N. C. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Comput. Appl. Biosci.* **13**, 235–238 (1997).

100. Chen, W.-C. Overlapping codon model, phylogenetic clustering, and alternative partial expectation conditional maximization algorithm. (Iowa State University, 2011).

101. Thorne, J. L., Kishino, H. & Painter, I. S. Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* **15**, 1647–1657 (1998).

102. Rannala, B. & Yang, Z. Inferring speciation times under an episodic molecular clock. *Syst. Biol.* **56**, 453–466 (2007).

103. Lemey, P., Rambaut, A., Welch, J. J. & Suchard, M. A. Phylogeography takes a relaxed random walk in continuous space and time. *Mol. Biol. Evol.* **27**, 1877–1885 (2010).

104. Reis, M. D. *et al.* Using Phylogenomic Data to Explore the Effects of Relaxed Clocks and Calibration Strategies on Divergence Time Estimation: Primates as a Test Case. *Syst. Biol.* **67**, 594–615 (2018).

105. Dos Reis, M., Zhu, T. & Yang, Z. The impact of the rate prior on Bayesian estimation of divergence times with multiple Loci. *Syst. Biol.* **63**, 555–565 (2014).

106. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).

107. Soraggi, S., Wiuf, C. & Albrechtsen, A. Powerful Inference with the D-Statistic on Low-

Coverage Whole-Genome Data. *G3* **8**, 551–566 (2018).

108. Bravo-Cuevas, V. M., Priego-Vargas, J., Cabral-Perdomo, M. Á. & Pineda Maldonado, M. A. First occurrence of Panthera atrox (Felidae, Pantherinae) in the Mexican state of Hidalgo and a review of the record of felids from the Pleistocene of Mexico. *Fossil Record* **19**, 131–141 (2016).

109. Dundas, R. G. Quaternary records of the dire wolf, Canis dirus, in North and South America. *Boreas* **28**, 375–385 (1999).

110. Czaplewski, N. J., Rogers, K. J. & Russell, C. LATE PLEISTOCENE VERTEBRATES FROM THREE-FORKS CAVE, ADAIR COUNTY, OKLAHOMA OZARK HIGHLAND. *J. Cave Karst Stud.* **80**, (2018).

111. Prevosti, F. J. & Rincón, A. D. A new fossil canid assemblage from the late Pleistocene of northern South America: the canids of the Inciarte asphalt pit (Zulia, Venezuela), fossil record and biogeography. *J. Paleontol.* **81**, 1053–1065 (2007).

112. Holen, S. R. *et al.* A 130,000-year-old archaeological site in southern California, USA. *Nature* **544**, 479–483 (2017).

113. Scott, E. & Springer, K. B. First records of Canis dirus and Smilodon fatalis from the late Pleistocene Tule Springs local fauna, upper Las Vegas Wash, Nevada. *PeerJ* **4**, e2151 (2016).

114. Smith, M. R. & Polly, P. D. A reevaluation of the Harrodsburg Crevice fauna (late Pleistocene of Indiana, U.S.A.) and the climatic implications of its mammals. *J. Vert. Paleontol.* **33**, 410–420 (2013).