Current Biology

Genome-wide Evidence Reveals that African and Eurasian Golden Jackals Are Distinct Species

Graphical Abstract



Authors

Klaus-Peter Koepfli, John Pollinger, Raquel Godinho, ..., Stephen J. O'Brien, Blaire Van Valkenburgh, Robert K. Wayne

Correspondence

koepflik@si.edu (K.-P.K.), rwayne@ucla.edu (R.K.W.)

In Brief

Koepfli et al. assess divergence between golden jackals (Canis aureus) from Africa and Eurasia using data from the mitochondrial and nuclear genomes. They show that African and Eurasian golden jackals are genetically distinct and independent lineages, and that African golden jackals likely represent a separate species.

Highlights

- African and Eurasian golden jackals are genetically distinct lineages
- Divergence between lineages is concordant across multiple molecular markers
- Morphologic convergence is observed between African and Eurasian golden jackals
- African golden jackals merit recognition as a distinct species





Genome-wide Evidence Reveals that African and Eurasian Golden Jackals Are Distinct Species

Klaus-Peter Koepfli,^{1,2,17,*} John Pollinger,^{3,17} Raquel Godinho,^{4,5} Jacqueline Robinson,³ Amanda Lea,⁶ Sarah Hendricks,⁷ Rena M. Schweizer,³ Olaf Thalmann,^{8,9} Pedro Silva,⁴ Zhenxin Fan,¹⁰ Andrey A. Yurchenko,² Pavel Dobrynin,² Alexey Makunin,² James A. Cahill,¹¹ Beth Shapiro,¹¹ Francisco Álvares,⁴ José C. Brito,⁴ Eli Geffen,¹² Jennifer A. Leonard,¹³ Kristofer M. Helgen,¹⁴ Warren E. Johnson,¹⁵ Stephen J. O'Brien,^{2,16} Blaire Van Valkenburgh,³ and Robert K. Wayne^{3,*}

¹Smithsonian Conservation Biology Institute, National Zoological Park, 3001 Connecticut Avenue NW, Washington, DC 20008, USA ²Theodosius Dobzhansky Center for Genome Bioinformatics, St. Petersburg State University, 41A Sredniy Prospekt, St. Petersburg 199034, Russia

³Department of Ecology and Evolutionary Biology, University of California, Los Angeles, 610 Charles Young Drive East, Los Angeles, CA 90095-1606, USA

⁴CIBIO/InBIO - Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, and Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s⁄n, 4169-007 Porto, Portugal

⁵Department of Zoology, University of Johannesburg, PO Box 534, Auckland Park 2006, South Africa

⁶Department of Biology, Duke University, PO Box 90388, Durham, NC 27708, USA

⁷Institute for Bioinformatics and Evolutionary Studies, Department of Biological Sciences, University of Idaho, 875 Perimeter MS 3051, Moscow, ID 83844, USA

⁸Department of Biological Sciences, Division of Genetics and Physiology, University of Turku, Itäinen Pitkäkatu 4, 20014 Turku, Finland ⁹Department of Biology, University of Oulu, PO Box 3000, 90014 Oulu, Finland

¹⁰Sichuan Key Laboratory of Conservation Biology on Endangered Wildlife, College of Life Sciences, Sichuan University, Chengdu 610064, China
 ¹¹Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA
 ¹²Department of Zoology, Tel Aviv University, Tel Aviv 69978, Israel

¹³Estación Biológica de Doñana, Conservation and Evolutionary Genetics Group (EBD-CSIC), Avenida Américo Vespucio s/n, 41092 Sevilla, Spain

¹⁴Division of Mammals, National Museum of Natural History, MRC 108, Smithsonian Institution, PO Box 37012, Washington, DC 20013-7012, USA

¹⁵Smithsonian Conservation Biology Institute, National Zoological Park, 1500 Remount Road, Front Royal, VA 22630, USA

¹⁶Nova Southeastern University, Oceanographic Center, 8000 North Ocean Drive, Dania Beach, FL 33004 USA

17Co-first author

*Correspondence: koepflik@si.edu (K.-P.K.), rwayne@ucla.edu (R.K.W.)

http://dx.doi.org/10.1016/j.cub.2015.06.060

SUMMARY

The golden jackal of Africa (Canis aureus) has long been considered a conspecific of jackals distributed throughout Eurasia, with the nearest source populations in the Middle East. However, two recent reports found that mitochondrial haplotypes of some African golden jackals aligned more closely to gray wolves (Canis lupus) [1, 2], which is surprising given the absence of gray wolves in Africa and the phenotypic divergence between the two species. Moreover, these results imply the existence of a previously unrecognized phylogenetically distinct species despite a long history of taxonomic work on African canids. To test the distinct-species hypothesis and understand the evolutionary history that would account for this puzzling result, we analyzed extensive genomic data including mitochondrial genome sequences, sequences from 20 autosomal loci (17 introns and 3 exon segments), microsatellite loci, X- and Y-linked zinc-finger protein gene (ZFX and ZFY) sequences, and whole-genome nuclear sequences in African and Eurasian golden jackals and gray wolves. Our results provide consistent and robust evidence that populations of golden jackals from Africa and Eurasia represent distinct monophyletic lineages separated for more than one million years, sufficient to merit formal recognition as different species: C. anthus (African golden wolf) and C. aureus (Eurasian golden jackal). Using morphologic data, we demonstrate a striking morphologic similarity between East African and Eurasian golden jackals, suggesting parallelism, which may have misled taxonomists and likely reflects uniquely intense interspecific competition in the East African carnivore guild. Our study shows how ecology can confound taxonomy if interspecific competition constrains size diversification.

RESULTS

Two recent studies based on mtDNA reported that the larger-sized golden jackals from Ethiopia and North and West Africa were more closely related to gray wolves than to other



populations of golden jackals, suggesting that some populations of African golden jackals represent a cryptic subspecies of gray wolf, designated *Canis lupus lupaster*, the African wolf [1, 2]. These results were consistent with earlier findings based on morphological and zoogeographic evidence that had suggested the large jackals of Egypt (*C. aureus lupaster*) were actually a small-sized subspecies of gray wolf [3]. However, this conclusion leaves the position of golden jackal populations in East Africa problematic, as they were never considered distinct from conspecifics in Eurasia. Consequently, either both golden jackal and African wolf occur in Africa, as has been suggested [2, 3], or these represent a single polytypic species. The former scenario suggests separate invasions of wolf- and jackal-like forms into North and East Africa, whereas the latter scenario suggests stable coexistence of distinct morphs within the same species that evolved in situ.

Evolutionary history is best verified through concordance among different molecular markers, which can provide a genome-wide history of divergence, and along with ecological and morphological data can be used to understand the context of evolutionary divergence [4–7]. Here, we present detailed analyses of the genome history of golden jackals employing a comprehensive set of molecular markers that include (1) mitochondrial genome sequences, (2) 20 autosomal DNA segments, (3) microsatellites, (4) sequences from the X- and Y-linked zincfinger protein gene (*ZFX* and *ZFY*), and (5) \sim 7.6 million SNPs derived from whole-genome sequences. We compare the data generated from golden jackals to that from gray wolves and other wolf-like canids (see Supplemental Experimental Procedures).

Phylogenetic Analyses of Mitochondrial and Nuclear Sequences

Phylogenies estimated from sequences of the mitochondrial cytochrome b gene, 13 protein-coding and two rRNA genes from complete mitochondrial genomes (13,890 bp), and 17 intron plus 3 exon segments (13,727 bp) were all consistent in showing that golden jackals are separated into two well-supported clades. The cytochrome b phylogeny (Figure 1A) includes both published and novel sequences from golden jackals sampled in Africa and Eurasia (Figure 1B), which are assorted into two clades. Golden jackal haplotypes from Kenya, Mauritania, and Morocco are included in a clade containing haplotypes of canids from Algeria, Egypt, Mali, and Senegal referred to as C. lupus *lupaster* [1, 2]. This African golden jackal clade is closely related to Eurasian gray wolves with strong nodal support and up to 6.7% divergence from Eurasian golden jackals. The only exceptions to this geographic pattern are haplotypes of golden jackals from Israel, which are grouped into both Eurasian and African clades, and three canids originating from Egypt that are classified as African wolf, gray wolf, and golden jackal (the latter indicated by arrows in Figure 1B). The phylogeny of sequences derived from complete mitochondrial genomes also shows that African golden jackals from Kenya group strongly with gray wolves, and not with a Eurasian golden jackal from Israel (Figure S1).

Phylogenies estimated from nuclear data likewise suggest a close relationship between representative African golden jackals and gray wolves, but in these analyses, the gray wolf clade is sister to coyotes as found previously [8], suggesting that the divergence between golden jackals and gray wolves preceded that of gray wolves and coyotes (Figure 2). Phlyogenies estimated using both concatenation and multispecies coalescent approaches were identical, except for the relative placement of the Ethiopian wolf (*C. simensis*) (Figure S2). Divergence times estimated using the concatenated nuclear dataset show that gray wolf, coyote, Ethiopian wolf, and the two lineages of golden jackals diversified during the Pleistocene, beginning about 1.9 million years ago (mya) (95% highest posterior density [HPD] = 1.5-2.4 mya) with the divergence of the Eurasian lineage of golden jackals (Figure 2). The divergence between the African lineage of golden jackals and the gray wolf + coyote clade was estimated at 1.3 mya (95% HPD = 1.0-1.7 mya). These estimates are slightly earlier than the corresponding values from the mitochondrial genome analysis (Figure S1).

The mitochondrial gene trees and nuclear species trees differ significantly in topology, which may be due to differences in lineage sorting (see <u>Supplemental Experimental Procedures</u>). Nonetheless, topologies in which the two jackal lineages were constrained to be monophyletic were less significantly supported compared to their optimal topologies (Table S1).

Sex Chromosome Sequences

Genetic distinctness between African and Eurasian golden jackals is further supported by analyses of sequences from the final intron of the zinc-finger X-chromosomal (*ZFX*) and Y-chromosomal (*ZFY*) genes. Eurasian golden jackals, including most individuals from Israel, carry *ZFX* or *ZFY* haplotypes distinct from those seen in gray wolves, coyotes, and African golden jackals (Figure 3A; Table S2). Notably, African golden jackals lack a 210 bp SINE II insertion, a 9 bp insertion, and a 2 bp insertion observed in Eurasian golden jackals (Table S2) [9]. A PCR assay of a larger panel of 31 male golden jackals from Eurasia and Africa confirmed that, with two exceptions (both from Israel), all male golden jackals from Eurasia had the *ZFY* SINE II element insertion, though this insert was absent in African golden jackals (Table S3).

Whole-Genome Sequences and Admixture

Whole-genome sequence analysis of three Eurasian wolves and one African (Kenya) and one Eurasian (Israel) golden jackal yielded 7,675,363 SNPs, and pairwise comparisons among these taxa confirm their distinctiveness across the genome (Figure 3B). We found relatively low diversity among gray wolf sequences (\sim 42-45 ± 22 sites in 50,000 bp) despite the sampled wolves originating from geographically distant populations across Eurasia. The African golden jackal was equally divergent from all three gray wolves, differing at \sim 72 ± 30 sites in 50,000 bp. The Eurasian golden jackal showed a higher level of divergence from the gray wolves of \sim 87 ± 29 sites in 50,000 bp. Most strikingly, the African and Eurasian golden jackals were the most divergent, differing by \sim 94 ± 31 sites in 50,000 bp. Principal-component analysis (PCA) and historical trajectories of effective population size from the five canid genomes further reinforce the distinction between the two golden jackals relative to gray wolves (Figure S3).

We found evidence confirming historical gene flow among the canid lineages in *D* statistic analyses of the genome-wide SNP data (Figure 3C; Table S4) [10]. Low *D* values (D = 0.0



Figure 1. Phylogenetic Tree Based on Mitochondrial Cytochrome *b* Sequences and Sampling Localities of Golden Jackals Used in This Study (A) Maximum-likelihood phylogram of 104 cytochrome *b* sequences (1,140 bp). Haplotype number is shown next to taxon name and locality. Accession numbers indicate sequences downloaded from GenBank. Haplotypes without accession numbers are novel sequences generated for the present study. Asterisks at nodes indicate bootstrap support \geq 80% based on maximum-likelihood analyses (500 pseudoreplicates) and \geq 0.95 posterior probability from Bayesian inference. *Canis* spp. from Egypt are indicated by thick arrows. Haplotypes labeled as *Canis lupus lupaster* refer to the African wolf. The tree was rooted using Sechuran fox (*Lycalopex sechurae*) as outgroup. Scale bar indicates the number of substitutions per site. Photo credits: left, golden jackal from Senegal (© CIBIO/Monia Nakamura); center, Mexican gray wolf (© Tom and Pat Leeson); right, golden jackal from lsrael (© Eyal Cohen). (B) Map of geographic localities showing where golden jackals were sampled. Relative number of animals sampled from each locality is shown. Hatched lines indicates geographic range of golden jackal based on IUCN distribution (http://www.iucnredlist.org/details/3744/0). See also Figure S1 and Table S1.

to -0.04) indicate only infrequent gene flow between the Kenyan golden jackal and gray wolf lineages, comparable to comparisons between anatomically modern humans and Neanderthals [10]. In contrast, higher *D* statistic values (0.16 to 0.18) suggest significant gene flow has occurred between Eurasian golden jackals and the gray wolf/dog group after the latter's divergence from the African golden jackal lineage (Figure 3C; Table S4).

Interestingly, the signal of gene flow is greatest between the Eurasian golden jackal and the basenji and dingo. The closeness of dog breeds and golden jackals may indicate recent admixture. Alternatively, some dog genome component may derive from admixture with gray wolves that have admixed with dogs in the past. However, previous genome analysis suggests that the dog component in Middle Eastern gray wolves is <9% [11]. Additional evidence for genetic admixture in Israeli golden jackals comes from comparisons of our cytochrome *b*, nuclear DNA, microsatellite, and *ZFX/ZFY* sequence results (see above and Supplemental Experimental Procedures).

Microsatellite Analysis of Population Structure

Bayesian clustering analysis of 128 individuals genotyped at 38 microsatellite loci corroborates our findings above (Figure 3D). Our results showed that K = 3 had the highest likelihood (see Supplemental Experimental Procedures), with Eurasian golden jackals; golden jackals from Kenya; and a group containing North African golden jackals, gray wolves, and dogs resolved as distinct genetic clusters. Notably, at K = 2 all African golden



Figure 2. Chronogram Estimated from Concatenated Analysis of Twenty Nuclear Gene Segments Using a Relaxed Molecular Clock Tree is based on analysis of 13,727 bp of sequence collected from 17 intron- and 3 exon-containing segments. Values shown at nodes are, respectively: Shimodaira-Hasegawa-like approximate likelihood ratio test (SH-aLRT, PhyML), bootstrap support with 1,000 pseudoreplicates (BS, RAxML), and posterior probability from Bayesian inference (PP, BEAST). Asterisks indicate SH-aLRT = 100%, BS = 100%, and PP = 1.0. Node bars show 95% highest posterior density (HPD) for divergence times. Four individuals were used each for gray wolf, golden jackal (Africa), and golden jackal (Eurasia), and two individuals were used for coyote. Letters correspond to list of estimated divergence times and 95% HPD for internodes (inset). The tree was rooted using red fox (*Vulpes vulpes*) and gray fox (*Urocyon cinereoargenteus*) as outgroups. Scale bar indicates the number of substitutions per site. Timescale at bottom is in million years ago (mya), and geological timescale (epochs) are shown at top. Photo credits: top, Mexican gray wolf (© Tom and Pat Leeson); middle, golden jackal from Senegal (© CIBIO/ Monia Nakamura); bottom, golden jackal from Israel (© Eyal Cohen). See also Table S1 and Figure S2.

jackals are grouped together with gray wolves and dogs in a single cluster, while at K = 4 North African golden jackals are resolved as a cluster distinct from gray wolves and dogs (Figure 3D). Critically, our results suggest that the presence of two mtDNA clades in golden jackals from Israel (see cytochrome *b* results above) does not reflect the occurrence of two reproductively distinct entities in this region, as the microsatellite results suggest that haplogroups do not form distinct genetic clusters (Figure 3D).

Size and Morphological Parallelism

We tested whether the patterns revealed by the genetic and genomic data were also manifested in morphology. PCA of 45 cranial and dental measurements taken from 140 golden jackals sampled from throughout the range of the species [12] revealed that golden jackals from North Africa are distinct from golden jackals from Eurasia and East Africa on PC1 (58.3% variation explained), which reflects the larger body size of North African golden jackals and is consistent with the equal loading across measurements on this PC (Figure 4A). PC2 (7.0% variation explained) does not segregate these three populations further,

but PC3 (4.4% variation explained) suggests that there are some differences in relative tooth size and skull shape between Eurasian and Middle Eastern golden jackals and all other African golden jackals (North, East, West, and Central) (Figure S4). To explore this further, we conducted PCA on the arcsine-transformed values of nine shape ratios for three groups: North African, East African, and Middle Eastern golden jackals (see Supplemental Experimental Procedures). The first PC accounted for 33% of the variance and separated East African from Middle Eastern golden jackals (Figure 4B). Compared with East African golden jackals, Eurasian golden jackals had high values on this axis, reflecting their broader muzzles, shorter molars, and the rounder cross-sections of their premolars and upper canines (see Supplemental Experimental Procedures). North African golden jackals overlap with the other two populations on the first PC, perhaps because this sample includes both larger "African wolf" individuals and others that are more closely related to Middle Eastern golden jackals. Notably, the North African golden jackals have more negative or near-zero values on the first PC and thus are more similar to East African than Middle Eastern golden jackals in shape. The North and East



Figure 3. Patterns of Genetic Differentiation and Admixture of African and Eurasian Golden Jackals Based on Sex Chromosome Sequences, Genome-wide SNP Data, and Microsatellite Multilocus Genotypes

(A) Haplotype networks showing relationships among ZFX and ZFY final intron sequences among golden jackals from Africa and Eurasia, gray wolves, and coyotes. Circle size is proportional to haplotype frequency (see scale). Small dots on internodes indicate number of indels and nucleotide substitutions between haplotypes. Internodes without dots indicate single substitutions between haplotypes. The 210 bp SINE II insertion in the ZFY sequences separating Eurasian golden jackals from African golden jackals, gray wolves, and coyotes is indicated. See Table S2 for specific sequence features of each haplotype. See also Table S3.

(B) Comparison of genome-wide divergence between golden jackals and gray wolves. Histograms of genome-wide pairwise distance estimates were calculated from 50 kb non-overlapping windows (41,999 windows total) for all ten possible pairwise comparisons between the three gray wolf genomes and two golden jackal genomes. Gray wolves are from China, Croatia, and Israel. Pairwise differences are the number of differences per 50 kb. See also Figure S3.

(C) Diagram showing the phylogenetic relationships among dogs, gray wolves, African golden jackals (Kenya), and Eurasian golden jackals (Israel) used in the D statistic comparisons. The phylogeny was rooted using the Channel Island fox (not shown). D statistic values above double-headed arrows indicate detectable admixture (gene flow) between lineages. Gray wolves are from China, Croatia, and Israel, and domestic dogs represent the dingo and basenji breeds. See also Table S4.

(D) Estimated population structure of 128 individuals genotyped for 38 microsatellite loci. Analysis and posterior probability assignments to each cluster assuming two (K = 2) to five (K = 5) genetic clusters were estimated using STRUCTURE (see Supplemental Experimental Procedures). ΔK likelihood was highest for K = 3 (see Supplemental Information). The origin of individuals in each cluster is indicated at the bottom of the figure.

African golden jackals are similar in having narrower, more blade-like upper canines, as well as more slender premolars and muzzles, all of which are gray wolf-like features. These results suggests that parallelism in size and body conformation between Eurasian and East African jackals is accompanied by more subtle differences that support common ancestry of the latter with North African jackals.

DISCUSSION

Our results from mtDNA, nuclear loci, and whole genomes provide consistent, compelling evidence that golden jackals from Africa and Eurasia constitute largely distinct gene pools with independent evolutionary histories. We estimate that the African lineage has been on an independent trajectory for at least one million years. Our results extend and contrast with the findings of previous genetic studies, based exclusively on mitochondrial DNA, which suggested that some golden jackal populations in Africa constitute a subspecies of gray wolf [1, 2]. Specifically, we show that, given our current sampling, there are no golden jackals of Eurasian affinity in Africa. Instead, African golden jackals define a distinct lineage, which includes those from East Africa showing phenotypic similarity to Eurasian golden jackals. African golden jackals are distinct by all genetic measures in this study, showing diagnostic differences across a range of markers and with levels of genome divergence similar



in magnitude to those found between other recognized species. Thus, our results suggest that African golden jackals merit recognition as a full species, as they meet the primary defining criterion of a separate and independently evolving metapopulation lineage under the unified species concept [13]. Accordingly, we propose that African golden jackals be designated as Canis anthus (Cuvier, 1820) based on the earliest description of golden jackals from Senegal [14] (see Supplemental Experimental Procedures). Furthermore, we suggest that the common names "African golden wolf" (C. anthus) and "Eurasian golden jackal" (C. aureus) be applied to distinguish these taxa, and to distinguish the former from the Ethiopian wolf (C. simensis). We propose that the African golden wolf is distributed across Africa and includes individuals that have been referred to as C. lupus lupaster [1-3] or C. aureus, sensu lato. Morphologic parallelism of African golden wolves and Eurasian golden jackals may have resulted in their mistaken attribution to a single species

Figure 4. Principal Component Analyses of the Morphometric Data for African and Eurasian Golden Jackals

(A) Plot of principal component 2 (PC2) against PC1 based on 45 linear measurements of teeth and skulls of 140 African and Eurasian golden jackals from five different geographic regions. See [12] for details of geographic sampling of golden jackals.
(B) Plot of PC2 against PC1 based on nine ratio variables that describe dental and cranial shape for three populations: North Africa (Egypt, Libya, Tunisia, Algeria, Morocco, Senegal, Western Sahara), East Africa (Kenya, Ethiopia) and the Middle East (Iran, Turkey, Jordan, Israel, Greece).
Numbers in parentheses indicate percent variance explained on each axis. See also Figure S4.

in most taxonomic treatments since *C. aureus* was first formally described by Linnaeus [15]. *C. anthus* merits conservation concern and assessment independent of *Canis aureus*, as it represents a unique legacy of adaptation and divergence within the extant Canidae.

Our nuclear DNA analyses indicate that the African golden wolf lineage split from the gray wolf + coyote clade about 1.0– 1.7 mya during the Pleistocene. More broadly, our phylogenetic analyses suggest that extant wolf-like canids have colonized Africa from Eurasia at least five times throughout the Pliocene and Pleistocene, which is consistent with fossil evidence suggesting that much of African canid fauna diversity resulted from the immigration of Eurasian ancestors [16, 17], likely coincident with Plio-Pleistocene climatic oscillations between arid and humid conditions [18, 19].

Our analyses of genome-wide SNP data revealed evidence of admixture in the histories of Eurasian golden jackals

and African golden wolves. Eurasian golden jackals from Israel show signals of hybridization with gray wolves, dogs, and the African golden wolf based on D statistic analyses and comparisons of cytochrome b, microsatellite, and ZFX/ZFY sequence results. The close geographic proximity and connectivity between the Levant and Northeastern Africa (e.g., Egypt) may have facilitated admixture and mitochondrial capture of African golden wolf haplotypes by Eurasian golden jackals. Furthermore, Eurasian golden jackals have only recently recolonized parts of Israel following a large-scale eradication program begun in the 1960s to control rabies [20], and the greater amount of hybridization detected in Eurasian golden jackals from Israel may be related to colonization of migrants from elsewhere. Interestingly, microsatellites revealed no evidence of admixture, suggesting that the admixture we detected in the genome-wide SNP data was relatively ancient. Previous analysis of complete genome sequences of gray wolves and Israeli golden jackals also supported ancient hybridization between the two species, suggesting that as much as 15% of the current Israeli wolf genome is derived from ancient admixture with golden jackals [11]. Our results suggest a dynamic genetic history among these canids in the Middle East and North Africa, similar to that observed in North American wolf-like canids and other carnivoran taxa such as brown and polar bears [21–23]. Increased sampling of gray wolves, African golden wolves, and Eurasian golden jackals from throughout the Middle East and North Africa will be required to fully resolve the details of this history.

Despite their distinct genetic ancestries, African golden wolves and Eurasian golden jackals are phenotypically similar in craniodental anatomy, and African golden wolves from East Africa and Eurasian golden jackals are similar in body size. This striking example of parallel evolution highlights the importance of natural selection in constraining morphologic divergence in sympatric carnivores [24-26]. However, there are subtle shape similarities in craniodental form that unite African golden wolves and distinguish them from Eurasian golden jackals. The phylogenetic affinities of the African golden wolves to gray wolves or gray wolves + coyotes, the canine fossil record, and macroevolutionary dynamics of canine body-size evolution suggest that they were derived from ancestors of larger body size [16, 27]. The convergent evolution of a smaller, more omnivorous jackallike form, especially in East Africa, from larger, more carnivorous wolf-like forms is uncommon in canids [28, 29] and may have been facilitated by intense competition from a uniquely diverse carnivoran community including species larger and smaller than jackals, thus inhibiting size divergence [12].

ACCESSION NUMBERS

GenBank accession numbers for the sequences reported here were not yet available from GenBank as of the date this article was finalized for press; please contact the corresponding authors for the GenBank accession numbers. Additional files associated with the Supplemental Information have been deposited at the Dryad Digital Repository at http://dx.doi.org/10.5061/dryad.3b77f.

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, four tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx. doi.org/10.1016/j.cub.2015.06.060.

AUTHOR CONTRIBUTIONS

K.-P.K. designed the study, performed experiments, analyzed data, and drafted the manuscript. J.P. designed the study, performed experiments, and analyzed the sex chromosome and SNP data. R.G. collected and analyzed the microsatellite data. J.R., A.L., S.H., and R.M.S. performed experiments and collected data. O.T. performed experiments and collected the wholemitochondrial-genome data. P.S., Z.F., J.A.C., and B.S. analyzed the wholegenome SNP data. P.D. and A.M. analyzed the mitochondrial genome and nuclear DNA data. A.A.Y. and B.V.V. analyzed the morphological data. F.A., J.C.B., E.G., J.A.L., and K.M.H. provided materials and reagents to the study. W.E.J. and S.J.O. contributed to the scientific strategy and assisted with the manuscript. R.K.W. co-designed and supervised the study and co-drafted the manuscript. All authors contributed to and approved the final manuscript.

ACKNOWLEDGMENTS

K.-P.K., A.A.Y., P.D., A.M., and S.J.O. were supported by Russian Ministry of Science Mega-grant 11.G34.31.0068. R.G. and J.C.B. were supported by FCT

contracts (IF/00564/2012 and IF/00459/2013, respectively). Fieldwork of J.C.B. and F.A. in North Africa was supported by the National Geographic Society (CRE 7629-04 and CRE 8412-08) and CIBIO, respectively. Microsatellite lab work was partially funded by Project "Genomics Applied to Genetic Resources" cofinanced by North Portugal Regional Operational Programme 2007/2013 (ON.2 - O Novo Norte), under the National Strategic Reference Framework, through the European Regional Development Fund. R.M.S. was supported by a National Science Foundation Graduate Research Fellowship. O.T. is financed by a Marie Curie Intra-European Fellowship within the 7th European Community Framework Program and is grateful to M. Webster. We thank the Tel Aviv University Zoological Museum for providing samples of golden jackals and gray wolves used in this study. We gratefully acknowledge Frank Zachos (Naturhistorisches Museum Wien) for providing samples of golden jackals from Serbia and for constructive comments on the manuscript. We also thank N. Ferrand for helpful comments on the manuscript. We thank Michael Campana (Smithsonian Conservation Biology Institute) for conducting additional phylogenetic analyses on the mitochondrial genome dataset. We are grateful to Pauline Charruau-Dau and rev.com for providing translations of Frédéric Cuvier's description of Canis anthus. We also thank D. Gordon E. Robertson for permission to use the photograph of a golden jackal from Serengeti National Park, Tanzania, for the graphical abstract. Finally, we thank four anonymous reviewers for providing excellent comments that improved the manuscript.

Received: November 11, 2014 Revised: April 15, 2015 Accepted: June 22, 2015 Published: July 30, 2015

REFERENCES

- Rueness, E.K., Asmyhr, M.G., Sillero-Zubiri, C., Macdonald, D.W., Bekele, A., Atickem, A., and Stenseth, N.C. (2011). The cryptic African wolf: *Canis aureus lupaster* is not a golden jackal and is not endemic to Egypt. PLoS ONE 6, e16385.
- Gaubert, P., Bloch, C., Benyacoub, S., Abdelhamid, A., Pagani, P., Djagoun, C.A.M.S., Couloux, A., and Dufour, S. (2012). Reviving the African wolf *Canis lupus lupaster* in North and West Africa: a mitochondrial lineage ranging more than 6,000 km wide. PLoS ONE 7, e42740.
- Ferguson, W.V. (1981). The systematic position of *Canis aureus lupaster* (Carnivora: Canidae) and the occurrence of *Canis lupus* in North Africa, Egypt and Sinai. Mammalia 45, 459–465.
- Dupuis, J.R., Roe, A.D., and Sperling, F.A.H. (2012). Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. Mol. Ecol. 21, 4422–4436.
- Buckley-Beason, V.A., Johnson, W.E., Nash, W.G., Stanyon, R., Menninger, J.C., Driscoll, C.A., Howard, J., Bush, M., Page, J.E., Roelke, M.E., et al. (2006). Molecular evidence for species-level distinctions in clouded leopards. Curr. Biol. *16*, 2371–2376.
- Kitchener, A.C., Beaumont, M.A., and Richardson, D. (2006). Geographical variation in the clouded leopard, Neofelis nebulosa, reveals two species. Curr. Biol. 16, 2377–2383.
- Christiansen, P. (2008). Species distinction and evolutionary differences in the clouded leopard (Neofelis nebulosa) and Diard's clouded leopard (Neofelis diardi). J. Mammal. 89, 1435–1446.
- Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas, E.J., 3rd, Zody, M.C., et al. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438, 803–819.
- Tsubouchi, A., Fukui, D., Ueda, M., Tada, K., Toyoshima, S., Takami, K., Tsujimoto, T., Uraguchi, K., Raichev, E., Kaneko, Y., et al. (2012). Comparative molecular phylogeny and evolution of sex chromosome DNA sequences in the family Canidae (Mammalia: Carnivora). Zoolog. Sci. 29, 151–161.

- Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.-Y., et al. (2010). A draft sequence of the Neandertal genome. Science 328, 710–722.
- Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E., Silva, P.M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., et al. (2014). Genome sequencing highlights the dynamic early history of dogs. PLoS Genet. *10*, e1004016.
- Van Valkenburgh, B., and Wayne, R.K. (1994). Shape divergence associated with size convergence in sympatric East African jackals. Ecology 75, 1567–1581.
- De Queiroz, K. (2007). Species concepts and species delimitation. Syst. Biol. 56, 879–886.
- Geoffroy Saint-Hilaire, E., and Cuvier, F. (1824). Histoire Naturelle des Mammifères, Tome Deuxième (Chez A. Belin).
- Linnaeus, C. (1758). Systema Naturæ per Regna Tria Naturæ, secundum Classes, Ordines, Genera, Species, cum Characteribus, Differentiis, Synonymis, Locis, Tomus I (Laurentius Salvius).
- Tedford, R.H., Wang, X., and Taylor, B.E. (2009). Phylogenetic systematics of the North American fossil Caninae (Carnivora: Canidae). Bull. Am. Mus. Nat. Hist. 325, 1–218.
- Wang, X., and Tedford, R.H. (2008). Dogs: Their Fossil Relatives and Evolutionary History (Columbia University Press).
- deMenocal, P.B. (2004). African climate change and faunal evolution during the Pliocene–Pleistocene. Earth Planet. Sci. Lett. 220, 3–24.
- Le Houerou, H.N. (1997). Climate, flora and fauna changes in the Sahara over the past 500 million years. J. Arid Environ. 37, 619–647.
- Cohen, T.M., King, R., Dolev, A., Boldo, A., Lichter-Peled, A., and Bar-Gal, G.K. (2013). Genetic characterization of populations of the golden jackal and the red fox in Israel. Conserv. Genet. *14*, 55–63.

- vonHoldt, B.M., Pollinger, J.P., Earl, D.A., Knowles, J.C., Boyko, A.R., Parker, H., Geffen, E., Pilot, M., Jedrzejewski, W., Jedrzejewska, B., et al. (2011). A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. Genome Res. *21*, 1294–1305.
- Miller, W., Schuster, S.C., Welch, A.J., Ratan, A., Bedoya-Reina, O.C., Zhao, F., Kim, H.L., Burhans, R.C., Drautz, D.I., Wittekindt, N.E., et al. (2012). Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. Proc. Natl. Acad. Sci. USA 109, E2382–E2390.
- Cahill, J.A., Green, R.E., Fulton, T.L., Stiller, M., Jay, F., Ovsyanikov, N., Salamzade, R., St John, J., Stirling, I., Slatkin, M., and Shapiro, B. (2013). Genomic evidence for island population conversion resolves conflicting theories of polar bear evolution. PLoS Genet. 9, e1003345.
- Wayne, R.K., Van Valkenburgh, B., Kat, P.W., Fuller, T.K., Johnson, W.E., and O'Brien, S.J. (1989). Genetic and morphological divergence among sympatric canids. J. Hered. 80, 447–454.
- Davies, T.J., Meiri, S., Barraclough, T.G., and Gittleman, J.L. (2007). Species co-existence and character divergence across carnivores. Ecol. Lett. 10, 146–152.
- Van Valkenburgh, B. (2007). Deja vu: the evolution of feeding morphologies in the Carnivora. Integr. Comp. Biol. 47, 147–163.
- Finarelli, J.A. (2007). Mechanisms behind active trends in body size evolution of the Canidae (Carnivora: Mammalia). Am. Nat. 170, 876–885.
- Van Valkenburgh, B., Wang, X., and Damuth, J. (2004). Cope's rule, hypercarnivory, and extinction in North American canids. Science 306, 101–104.
- Slater, G.J. (2015). Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. Proc. Natl. Acad. Sci. USA *112*, 4897–4902.

Current Biology Supplemental Information

Genome-wide Evidence Reveals that African

and Eurasian Golden Jackals Are Distinct Species

Klaus-Peter Koepfli, John Pollinger, Raquel Godinho, Jacqueline Robinson, Amanda Lea, Sarah Hendricks, Rena M. Schweizer, Olaf Thalmann, Pedro Silva, Zhenxin Fan, Andrey A. Yurchenko, Pavel Dobrynin, Alexey Makunin, James A. Cahill, Beth Shapiro, Francisco Álvares, José C. Brito, Eli Geffen, Jennifer A. Leonard, Kristofer M. Helgen, Warren E. Johnson, Stephen J. O'Brien, Blaire Van Valkenburgh, and Robert K. Wayne



Supplemental Figure S1 (related to Figure 1). Chronogram estimated from the 13 protein-coding and two rRNA genes of the mitochondrial genome (13,890bp) using a relaxed molecular clock.

Values of bootstrap support (BS) based on maximum likelihood analyses (RAxML) with 1000 pseudoreplicates and posterior probability (PP) from Bayesian inference (BEAST) are shown at nodes, respectively. Asterisks indicate BS = 100% and PP = 1.0. Node bars show 95% highest posterior density (HPD) for divergence times. Letters correspond to list of estimated divergence times and 95% HPD for internodes (inset). Numbers in parentheses indicates number of individuals used for that taxon. The tree was rooted using red fox (*Vulpes vulpes*) and arctic fox (*V. lagopus*) as outgroups. Time scale at bottom in millions of years before present (MYBP) and geological time scale (epochs) shown at top.



Supplemental Figure S2 (related to Figure 2). Phylogenies estimated from 20 nuclear gene segments (three exons, 17 introns; 13,727 bp).

(A) Phylogram estimated from concatenated analysis of 20 nuclear gene segments using maximum likelihood (RAxML). This tree has the same topology as shown in Figure 2 but taxa for which more than one individual was used are presented here along with their localities. Values shown at nodes are, respectively: Shimodaira-Hasegawa-like approximate likelihood ratio test (SH-aLRT, PhyML); bootstrap support with 1000 pseudoreplicates (RAxML); and posterior probability from Bayesian inference (MrBayes). Scale bar = number of substitutions per site. (B) Densitree plot of species tree from phased sequences of 20 autosomal gene segments estimated using Bayesian multispecies coalescent analysis (*BEAST; see Supplemental Experimental Procedures). Posterior probability values shown at nodes. For the following taxa, more than one individual was used in the multispecies coalescent analysis, as shown in Figure S2 (A): *Canis lupus*, n =4; *Canis latrans*, n = 2; *Canis aureus* (Africa), n =4; *Canis aureus* (Eurasia), n =4. Both trees were rooted with red fox (*Vulpes vulpes*) and gray fox (*Urocyon cinereoargenteus*).



Supplemental Figure S3 (related to Figure 3B). PCA results and trajectory of effective population size based on genome-wide SNPs from three gray wolves and two golden jackals.

(A) Principal component analysis (PCA) plot of the three gray wolves (*Canis lupus* = CLU) and two golden jackals (*C. aureus* = CAU) for a linkage disequilibrium-pruned subset of 264,937 SNPs for PC1 (X-axis) and PC2 (Y-axis). PC1 explains 36.5% of the overall variation while PC2 explains 24.4% of the variation. The inset presents eigenvalues for all four principal components. (B) Reconstruction of historical patterns of effective population size (*Ne*) for individual genome sequences of golden jackals (Kenya and Israel) and gray wolves (China, Croatia, and Israel) based on the genomic distribution of heterozygous sites using the pairwise sequential Markovian coalescent (PSMC) method [S91]. Data from gray wolves and the golden jackal from Israel were used from [S17]. The inferred pattern of the African golden jackal indicates a significant spike in *Ne* from ancient levels of ~50,000 to ~62,000 approximately 200 kya, followed by a continuous >10-fold decline towards the present. In comparison, the Eurasian golden jackal has an ancient level of *Ne* of ~30,000, with a spike to *Ne*~55,000 approximately 100 kya.



Supplemental Figure S4 (related to Figure 4). Principal component analyses (PCA) of the morphometric data.

(A) Plot of PC3 against PC1, and (B) PC3 against PC2, based on 45 linear measurements of the teeth and skulls of 140 African and Eurasian golden jackals from five different geographic regions. Note that both Eurasian and Middle Eastern jackals tend to have similar and more negative values on PC3 than all African jackals. Numbers in parentheses indicate percent variance explained on each axis.

Table S1. Results of Shimodaira-Hasegawa test comparing log-likelihoods (-ln L) of alternative topologies in which African and Eurasian golden jackals are unconstrained or constrained to be monophyletic across the cytochrome *b*, mitochondrial genome and nuclear sequence data sets. Trees with the highest likelihood across comparisons are in bold. * = P < 0.05.

Data set	Length	-In L unconstrained	-In L constrained	Difference -In L	p-value
Cytochrome b	1,140	4880.12241	4939.84854	59.72613	0.0001*
Mitochondrial genome	13,875	60828.28524	60902.1965	73.91126	<0.0001*
Nuclear - concatenated	13,727	25801.96978	25853.54305	51.57327	0.0094*

Supplemental Table S2 (related to Figure 3A). Informative positions found in the final intron sequence of the canid *ZFX* (X chromosome) and *ZFY* (Y chromosome) genes for male and female individuals. Positions based on *ZFX* and *ZFY* final intron alignments [S75]. Informative positions and features displayed for golden jackals (*Canis aureus*) from localities in Eurasia and Africa, along with a reference set of coyotes (*C. latrans*), gray wolves (*C. lupus*). Sequences from gray wolves included samples from Eurasia and North America. Domestic dogs (*C. l. familiaris*) were also included in the reference set and had the same feature states as gray wolves. Red represents the gray wolf feature state; green represents the Eurasian golden jackal feature state; and yellow the African golden jackal feature state where unique. A dash (-) indicates feature state is absent. Corresponding cytochrome *b* haplotype clades (based on phylogeny shown in Figure 1A) are presented for reference.

			ZFX fir	nal int	tron	ZFY final intron									
	POSITION		670	721	765	23	130	202-208	243-453	941-970	1072	1098-99	1118	1149	
	FEATURE		1 bp insertion	T/A	A/G	A/G	1 bp insertion	9 bp insertion	210 bp SINE II insertion	30 bp deletion	T/G	2 bp insertion	G/A	т/с	Cytochrome <i>b</i> haplotype clade
SPECIES	SAMPLE	SEX													
Canis latrans	North America	М		Т	Α	Α				30 bp	Т		G	Т	Canis latrans
Canis lupus	Holarctic	М		Т	Α	Α				30 bp	G		G	С	Canis lupus
	Afghanistan 1	F	G	Α	Α										
	Serbia DS3	F	G	Α	Α										
	Serbia DS5	М	G	Α	Α	G	Α	9 bp	210 bp		Т	ТА	Α	С	Canis aureus
	Serbia VP2	М	G	Α	Α	G	Α	9 bp	210 bp		Т	TA	Α	С	Eurasia
Canis aureus Eurasia	Serbia VP3	М	G	Α	Α	G	Α	9 bp	210 bp		т	TA	Α	С	
	Serbia VP4	F	G	Α	Α										
	Israel 214	М	G	Α	Α	G	Α	9 bp	210 bp	-	Т	TA	Α	С	Canis aureus
	Israel 31	м	G	Т	G	G	Α	9 bp	210 bp	-	Т	TA	Α	С	Eurasia
	Israel 24	М	G	Α	Α	G	Α	9 bp	210 bp	-	Т	TA	Α	С	
	Israel 27	М	G	Α	Α	G	Α	9 bp	210 bp		Т	TA	Α	С	
	Israel 37	М	G	Α	Α	G	Α	9 bp	210 bp		Т	TA	Α	С	Canis aureus
	Israel 39	М	G	Т	G	G	Α	9 bp	210 bp		Т	TA	Α	С	Africa
	Israel 34	М	G	Α	Α	G	Α	9 bp	210 bp		Т	TA	Α	С	
	Morocco 1592	М	G	Т	G	Α					Т		G	С	
Canis aureus Africa	Mauritania 2646	F	G	Т	G										
	Mauritania 3054	F	G	Т	G										
	Morocco 3544	F	G	Т	G										Canis aureus
	Kenya 445	М	G	Т	G	Α					Т		G	С	Africa
	Kenya 51	М	G	Т	G	Α					Т		G	С	
	Kenya 623	F	G	Т	G										
	Kenya 641	М	G	Т	G	Α					Т		G	С	

Supplemental Table S3 (related to Figure 3A). PCR assay results for *ZFY* final intron SINE-II element. Presence or absence of the 210bp SINE-II element within the final intron of the *ZFY* gene of male golden jackals (*Canis aureus*) from Eurasia and Africa. Corresponding cytochrome *b* haplotype clade for tree shown in Figure 1A of main text is also presented for comparison for each individual. A dash (-) indicates feature is absent.

		Y	
		Chromosome ZFY SINE-II	
		210bp	
Sample ID	Origin	Element	Cytochrome <i>b</i> haplotype clade
DS1	Serbia	Present	<i>C. aureus</i> - Eurasia
DS4	Serbia	Present	<i>C. aureus</i> - Eurasia
DS5	Serbia	Present	<i>C. aureus</i> - Eurasia
VP2	Serbia	Present	<i>C. aureus</i> - Eurasia
VP3	Serbia	Present	<i>C. aureus</i> - Eurasia
VP5	Serbia	Present	<i>C. aureus</i> - Eurasia
21	Israel	Present	<i>C. aureus</i> - Eurasia
24	Israel	Present	<i>C. aureus</i> - Eurasia
27	Israel	Present	C. aureus - Africa
31	Israel	Present	<i>C. aureus</i> - Eurasia
34	Israel	Present	C. aureus - Africa
37	Israel	Present	C. aureus - Africa
39	Israel	Present	C. aureus - Africa
54	Israel	-	<i>C. aureus</i> - Eurasia
214	Israel	Present	<i>C. aureus</i> - Eurasia
215	Israel	Present	<i>C. aureus</i> - Eurasia
216	Israel	Present	<i>C. aureus</i> - Eurasia
259	Israel	-	<i>C. aureus</i> - Eurasia
270	Israel	Present	<i>C. aureus</i> - Eurasia
51	Kenya	-	C. aureus - Africa
441	Kenya	-	C. aureus - Africa
445	Kenya	-	C. aureus - Africa
615	Kenya	-	C. aureus - Africa
631	Kenya	-	C. aureus - Africa
641	Kenya	-	C. aureus - Africa
645	Kenya	-	C. aureus - Africa
669	Kenya	-	C. aureus - Africa
671	Kenya	-	C. aureus - Africa
695	Kenya	-	C. aureus - Africa
2441	Kenya	-	C. aureus - Africa
1592	Morocco	-	C. aureus - Africa

Supplemental Table S4 (related to Figure 3C). *D*-statistic results for comparisons among three gray wolves and two golden jackals. The results are sorted by Z score from greatest to least. Z scores > 3 are considered significant evidence of a nonzero *D*-statistic value consistent with the presence of admixutre. Positive *D*-statistic values are indicative of gene flow between P2 and P3. Negative values are indicative of gene flow between P1 and P3. Gray wolf (*Canis lupus*) = Chinese, Croatian, Israeli; dog (*C. lupus familiaris*) = basenji and dingo; golden jackal (*C. aureus*) = Israeli and Kenyan. O = outgroup, Fox = Channel Island fox (*Urocyon littoralis*).

				Standard			
P1	P2	P3	0	D	Error	Z score	
Croatian wolf	Kenyan jackal	Israeli jackal	Fox	-0.167121	0.003649	45.801216	
Chinese wolf	Kenyan jackal	Israeli jackal	Fox	-0.164595	0.003728	44.149772	
Israeli wolf	Kenyan jackal	Israeli jackal	Fox	-0.166891	0.003813	43.771383	
Dingo	Kenyan jackal	Israeli jackal	Fox	-0.170839	0.004077	41.903663	
Basenji	Kenyan jackal	Israeli jackal	Fox	-0.178066	0.004286	41.550637	
Israeli wolf	Chinese wolf	Kenyan jackal	Fox	-0.028628	0.002929	9.775284	
Dingo	Israeli wolf	Kenyan jackal	Fox	0.029512	0.003253	9.070969	
Israeli wolf	Croatian wolf	Kenyan jackal	Fox	-0.019824	0.003003	6.602334	
Basenji	Chinese wolf	Israeli jackal	Fox	-0.021852	0.003315	6.592149	
Basenji	Israeli wolf	Israeli jackal	Fox	-0.021925	0.00339	6.467152	
Basenji	Croatian wolf	Israeli jackal	Fox	-0.019795	0.003265	6.063218	
Basenji	Israeli wolf	Kenyan jackal	Fox	0.019235	0.003296	5.836858	
Dingo	Croatian wolf	Kenyan jackal	Fox	0.011046	0.002802	3.942871	
Basenji	Dingo	Israeli jackal	Fox	-0.01347	0.003474	3.877272	
Basenji	Dingo	Kenyan jackal	Fox	-0.013683	0.003706	3.691705	
Croatian wolf	Chinese wolf	Kenyan jackal	Fox	-0.00984	0.002667	3.689018	
Dingo	Chinese wolf	Israeli jackal	Fox	-0.011024	0.003073	3.58731	
Basenji	Chinese wolf	Kenyan jackal	Fox	-0.010126	0.003037	3.333776	
Dingo	Israeli wolf	Israeli jackal	Fox	-0.009477	0.002855	3.319387	
Dingo	Croatian wolf	Israeli jackal	Fox	-0.007037	0.002968	2.371088	
Croatian wolf	Chinese wolf	Israeli jackal	Fox	-0.003197	0.002703	1.182682	
Israeli wolf	Croatian wolf	Israeli jackal	Fox	0.001646	0.002762	0.595872	
Israeli wolf	Chinese wolf	Israeli jackal	Fox	-0.001541	0.002614	0.58969	
Basenji	Croatian wolf	Kenyan jackal	Fox	-0.001354	0.003026	0.447323	
Dingo	Chinese wolf	Kenyan jackal	Fox	0.000836	0.003141	0.266251	

Supplemental Experimental Procedures

Sample collection and DNA extraction

We collected or obtained blood and tissue samples of golden jackals from the following countries: Afghanistan (n = 1), Algeria (n = 5), Israel (n = 25), Kenya (n = 15), Mauritania (n = 14), Morocco (n = 6), Serbia (n = 10), and Western Sahara (n=1). The sample from Afghanistan was from a female individual kindly provided by B.C. Yates at the US Fish and Wildlife Service's National Fish and Wildlife Forensic Laboratory (USFWS accession N2234). Samples from Israel were obtained from road-killed or legally culled animals. Golden jackals from Kenya were collected from live-captured animals in 1990 and 1991. Samples from Serbia were from legally culled animals collected in Donji Srem and Velika Plana and were part of a previous study [S1].

Our analyses also included samples of gray wolves selected from throughout their Holarctic range, including subspecies such as the Mexican wolf (*Canis lupus baileyi*) and Eurasian wolf (*C. lupus lupus*). Importantly, we included samples from Israel, one of the locations in Eurasia where gray wolves and golden jackals are sympatric and have the potential to hybridize. Sequences of *Canis* species from previously published studies [S2-S11] were downloaded from Genbank and included in the cytochrome *b* and/or mitochondrial genome analyses (http://dx.doi.org/10.5061/dryad.3b77f/1). In order to place the divergence of golden jackals and gray wolves in a wider phylogenetic context, we also collected specific types of data from other species of wolf-like canids such as coyote (*C. latrans*), Ethiopian wolf (*C. simensis*), black-backed jackal (*C. mesomelas*), side-striped jackal (*C. adustus*), dhole (*Cuon alpinus*), and African wild dog (*Lycaon pictus*). To root phylogenetic trees, Sechuran fox (*Lycalopex sechurae*), gray fox (*Urocyon cinereoargenteus*), red fox (*Vulpes vulpes*) and/or arctic fox (*Vulpes lagopus*) were used as outgroups.

We extracted genomic DNA from blood and tissue samples using the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA) following the manufacturer's protocol. For the canid samples used in the whole genome analyses (see below), we used phenol-chloroform and ethanol precipitation to extract and isolate genomic DNA [S12].

Mitochondrial genome amplification and sequencing

We amplified the complete mitochondrial genome in two overlapping segments by means of long range PCRs using the Expand Long Range dNTPack (Roche, Indianapolis, IN). 150 ng of genomic DNA was used in a 50 μ I PCR consisting of 1x PCR buffer including magnesium, 0.125 μ M each dNTP, 3.5 U Expand Long Range Enzyme mix and 0.3 mM each primer (sequences available upon request from the authors). Each individual mitochondrial segment was amplified under the following conditions: initial denaturation at 92°C for 2 min; followed by 10 cycles each consisting of denaturation at 92°C for 10 sec, primer annealing at 57°C for 15 sec and elongation at 68°C for 10 min; 27 cycles of denaturation at 92°C for 10 sec, annealing at 57°C for 15 sec and elongation at 68°C for 10 min with an additional 20 sec per cycle. Following a final elongation at 68°C for 7 min and a cooling step at 8°C, PCR products were electrophoresed and then assessed under UV-light using an ethidium-bromide stained 1% agarose gel.

We prepared individually barcoded-sequencing libraries following published procedures [S13]. PCR products were purified with magnetic Agencourt AMPure SPRI beads (Beckman Coulter, Indianapolis, IN), quantified (NanoDrop, Thermo Scientific, Waltham, MA) and the two long-range PCR segments pooled in equimolar ratios. The pool was sheared to the desired fragment length by nebulization according to the 454 GS FLX+ library preparation manual. Since we prepared the samples for several libraries to be sequenced on a 454 GS FLX and a 454 GS FLX Titanium (Roche, Indianapolis, IN), the protocols were adjusted to accommodate the increased fragment size of the latter platform. In general, we added an individual barcode to each canid sample by performing the following steps: blunt-end repair, ligation of adapters, which contain the barcode and the sequences necessary for successful sequencing on the 454 GS FLX+ machine, adapter fill-in reaction, single library quantification using a Pico-green assay, pooling of the respective barcoded libraries, dephosphorylation and restriction digestion followed by a final small fragment removal. One μ I of pooled libraries was quantified [S13] on a LightCycler 480

(Roche, Indianapolis, IN) using the following High Resolution Melting Master (Roche, Indianapolis, IN) protocol for a 20 μ l reaction: 1x MasterMix (dNTPs, polymerase, reaction buffer and HRM dye), 0.2 μ M of each primer, 4.375 mM MgCl₂, and additional ddH₂O. The real time PCR was run as follows: preincubation at 95°C for 10 min, 45 amplification cycles each with denaturation at 94°C for 10 sec, annealing at 60°C for 15 sec and elongation at 72°C for 25 sec. The melting reaction had the following steps: 95°C for 1 min, 40°C for 1 min, 65-95°C for 1 sec and cooling at 40°C. The quantified libraries were subsequently processed according to 454 GS FLX+ manuals and sequenced on fractions (1/16th) of full picotiter sequencing plates. Sequencing was carried out at the UCLA Genotyping and Sequencing Core facility. Raw sequencing reads were adapter trimmed and filtered on the machine according to default 454 GS FLX+ parameters, de-tagged as described [S13] and prepared for subsequent data processing.

Filtered and de-tagged reads were assembled into complete mt-genomes using two strategies. First, we employed the iterative mapping approach using a modified version of *mia* [S14]. This version was adjusted to handle long 454 reads of modern DNA origin and yielded a consensus sequence of all reads mapped against the reference mt-genome. We used a complete mitochondrial genome of a wolf (Genbank accession no. AM711902) [S15] as a reference and applied the *D*-flag in *mia*, indicating a distant relationship of the reference to the assembly whenever non-canids were assembled. In cases where the iterations in *mia* did not yield a consensus, we exchanged the reference to a more closely related species according to the known canid phylogeny [S16]. The second strategy used 454's default programs *gsAssembler* for a *de-novo* assembly of the mt-genome and *gsMapper*, which was used to map reads against the reference mt-genome of the wolf and respective substitutes. Both programs were used with default parameters.

Only consensus sequences that had a minimum average coverage of 10fold were further used and for consensus sites with ambiguities the respective assembly files were inspected by eye and a majority rule adopted. The primer binding sites of the PCR-amplified mt-genomes were also evaluated manually and consensus nucleotides were retained from the respective reverse strands. For Ethiopian wolf (*Canis simensis*) and side-striped jackal (*C. adustus*), we were only able to obtain 8092bp and 10277bp out of a total 13890bp used in the final alignments that contained the 13 protein-coding and two rRNA genes. Complete sequences were obtained from the following loci for the Ethiopian wolf: *ATP8*, *CYTb*, *NADH5*, *NADH6*, 12S rRNA, and 16S rRNA. For the side striped jackal, we obtained complete sequences from: *ATP8*, *COX2*, *COX3*, *CYTb*, *NADH1*, *NADH2*, *NADH3*, *NADH4*L, 12S rRNA, and 16S rRNA.

Mitochondrial genomes of an Israeli golden jackal, three gray wolves from China, Croatia and Israel, and two domestic dogs (basenji and dingo) were obtained from the study by Freedman et al. [S17]. The reads of the mitochondrial genome for the Kenyan golden jackal were mapped to the dog genome [S16] with Bowtie2 [S18].

Cytochrome *b* gene amplification and sequencing

Two primer pairs were used to amplify overlapping segments of the complete mitochondrial cytochome b gene (1,140bp) – L14724/H15513 and L15162/H15915 [S19]. PCR reactions (total volume of 25µl) consisted of 1.5µl DNA, 0.1µl AmpliTag Gold (Life Technologies, Grand Island, NY), 2.5µl 10X buffer, 2.0µl dNTPs (2.5mM), 1.5µl MgCl₂ (25mM), 1.5µl F primer and 1.5µl R primer (10µM), 1.0µl BSA (10mg/ml) and 13.4µl ddH₂O. PCR cycle conditions were: one cycle of 95°C for 15 min, followed by 45 cycles of 94°C for 45sec, 55°C for 45sec, and 72°C for 90sec, and then a final extension at 72°C for 10min. PCR products were electrophoresed, checked under UV-light using an ethidiumbromide stained 1% agarose gel, and purified using Exonuclease I, Shrimp Alkaline Phosphatase (Affymetrix, Santa Clara, CA). Sequencing reactions were performed with both forward and reverse primers using BigDye 3.1 and products were sequenced on a 3730xl DNA Analyzer (Life technologies, Grand Island, NY). Chromatograms from forward and reverse sequences were assembled and edited in Geneious Pro 5.5.6 [S20]. We aligned sequences using MAFFT v7.017 [S21] with the following settings: scoring matrix = 200PAM/k=2, gap open penalty = 1.53, offset value = 0.123. The sequences were translated as well as compared to cytochrome b sequences generated from the whole mitochondrial genomes by

long-range PCR (see above) to assure that the sequence products did not include nuclear mitochondrial inserts (NUMTs).

For the historic specimen labeled as *Canis lupus lupaster* [Z.D.1901.3.17.3, Natural History Museum (London)], DNA was extracted by phenol-chloroform with negative controls in an isolated ancient DNA laboratory, as in [S22]. The cytochrome *b* gene was amplified with the three sets of primers from Leonard et al. [S23] in 25 μ l with AmpliTaq Gold (Life Technologies). All PCRs were performed at least twice, and included amplification negative controls (no DNA). All cleaned PCR products were Sanger sequenced in both directions on an ABI sequencer. The final cytochrome *b* data matrix consisted of 104 sequences, 92 (88.5%) of which were 1140bp in length.

Nuclear gene segment amplification and sequencing

We amplified 20 nuclear gene segments from a subset of golden jackals (from Africa and Eurasia), gray wolves, coyotes, and other species of wolf-like canids. Primers for these gene segments (http://dx.doi.org/10.5061/dryad.3b77f/2) were obtained from previous phylogenetic studies or studies reporting sequence-tagged sites [S24-S34]. Seventeen of the segments contained complete introns with partial flanking exons while the remaining three segments were from exon sequences.

We amplified the nuclear gene segments with either one of two (I and II) polymerase chain reaction (PCR) reagent and touchdown protocols in a GeneAmp PCR System 9700 thermal cycler: I) A 15μ I reaction contained 6.98 μ I of sterile double-distilled water, 1.5µl of 10X PCR Gold Buffer, 1.2µl 25mM MqCl₂, 1.2 μ l of 10mM dNTP mix, 1.5 μ l of both 2 μ M forward and reverse primers, 0.12 μ l of AmpliTag Gold Tag polymerase (Life Technologies, Grand Island, NY), and 1.0μ I of $0.1-1.0\mu$ g genomic DNA. This mix was then subjected to one cycle of 95°C for 3 min; followed by 6 cycles of 94°C for 15 sec, 60°C to 50°C for 30 sec, with a decrease in annealing temperature of 2°C per cycle, and 72°C for 45 sec; followed by 30 cycles of 94°C for 15 sec, 50°C for 30 sec, and 72°C for 45 sec; and one cycle of 72°C for 30 min. II) A 25μ l reaction contained 14.9μ l of sterile double-distilled water, 2.5μ of 10X PCR Gold Buffer, 1.5μ 25mM MgCl₂, 1.0μ DMSO, 2.0 μ I of 10mM dNTP mix, 1.0 μ I of both 2 μ M forward and reverse primers, 0.1μ I of AmpliTaq Gold Taq polymerase, and 1.0μ I of $0.1-1.0\mu$ g genomic DNA. Thermal amplification included one cycle of 95°C for 10 min; followed by 16 cycles of 94°C for 60 sec, 63°C to 50.2°C for 60 sec, with a decrease in annealing temperature of 0.8°C per cycle, and 72°C for 90 sec; followed by 30 cycles of 94°C for 60 sec, 50°C for 60 sec, and 72°C for 90 sec; and one cycle of 72°C for 5 min. All PCRs included a negative control (no DNA) to check for contamination. Following electrophoresis in 1% agarose/Tris-acetic acid-EDTA agarose gels stained with ethidium bromide to verify amplification success, we purified PCR products with Exonuclease I, Shrimp Alkaline Phosphatase

(Affymetrix, Santa Clara, CA). The products were then cycle sequenced in both directions using the original amplification primers and the BigDye Terminator v3.1 Cycle Sequencing Kit (Life Technologies, Grand Island, NY). Dye terminators were removed from cycle sequencing products using Agencourt CleanSEQ (Beckman Coulter, Indianapolis, IN) and then run on a 96-capillary 3730xl DNA Analyzer (Life Technologies, Grand Island, NY). Chromatograms from forward and reverse sequences were assembled, checked and edited using the Geneious Pro v5.5.6 software package [S20].

Sequences generated from gray wolves and golden jackals were used in BLAST searches to extract the orthologous regions of the 20 nuclear gene segments from the *de novo* sequenced genomes of three gray wolves and two golden jackals (see below) and were included in the nuclear gene phylogenetic analyses.

Phylogenetic analyses

Mitochondrial genomes: We extracted the 13 protein-coding and two rRNA (12S and 16S) genes from the mitochondrial genomes for use in phylogenetic analyses. Individual alignments for each region were estimated using MAFFT v7.017 [S21] with the following settings: scoring matrix = 200PAM/k=2, gap open penalty = 1.53, offset value = 0.123. Alignments were then checked by eye and verified using the orthologous regions from the domestic dog mitochondrial genome [S16]. The 15 regions were then concatenated into a supermatrix with

13,890 sites (bp), including insertions and deletions. The best-fitting partitioning scheme and nucleotide substitution models were estimated using PartitionFinder v1.1.1 [S35]. We tested among several partitioning schemes including division of protein-coding genes into 1st, 2nd and 3rd codon positions and the rRNA genes as separate partitions. Models were selected by the Bayesian information criterion (BIC). We found the optimal partitioning scheme includes three partitions (optimal models are indicated in brackets) - 1st+2nd codons (HKY+G), 3rd codon (HKY+G+I) and rRNA genes (TrN+G+I).

We estimated the mitochondrial gene tree using maximum likelihood (ML) and Bayesian inference (BI) methods. ML phylogenetic analysis of the partitioned dataset was performed using raxmlGUI 1.3.1 [36, 37] with the GTR+G substitution model specified for each partition. Nodal support was measured with 1000 bootstrap replicates using the ML + thorough bootstrap setting and branch lengths saved in the bootstrap trees (BS brL enabled). For BI, the BEAST 1.7.5 program package [S38] was employed to co-estimate topology and divergence times using a strict molecular clock model. We used the BEAUti program to set up the MCMC run using the following parameters: unlinked substitution models using the partitioning scheme and corresponding models as indicated by PartitionFinder (see above), link trees enabled (given that the mitochondrial genome is non-recombining), strict clock, and tree prior specified as Yule process of speciation. The auto optimize setting was enabled in the Operators window. We calibrated the molecular clock using two fossil priors as soft bounds: 1) a root

age spanning 9-12 mya based on the approximate split between the tribes Canini and Vulpini as suggested by the earliest known fossils of *Eucyon* (Canini) [S39]; and 2) a prior spanning 1-3 mya, which encompasses the earliest known fossils of Canis lupus (~1 mva) and Canis edwardii (~3 mva), the latter of which is sister to a clade containing C. aureus and C. latrans based on phylogenetic analyses of morphological data of extant and fossil taxa [S39]. Time calibration priors were approximated with lognormal distributions: mean=0.0, stdev=0.7, offset=8.7 (95% HPD 9.0-11.9 Mya) for prior 1 and mean=0.0, stdev=0.5, offset=0.7 (95% HPD 1.1-3.0 Mya) for prior 2. Three independent MCMC analyses were run for 10 million generations with sampling each 1,000 generations. The first 25% trees were discarded as burn-in. We also performed an analysis with sampling from the priors only (without sequence data), with an MCMC run for 30 million generations and sampling each 3,000 generations. Analysis of the posterior distributions of tree likelihood and other parameters using Tracer v1.6 [S40] showed ESS values >500 and trace plots indicating rapid mixing. The maximum clade credibility tree with mean node heights was visualized using FigTree v1.4.0 [S41].

Cytochrome *b*: For the cytochrome *b* dataset, the best-fitting model of DNA substitution was HKY+I+G, chosen among 24 models and a base tree estimated with BIONJ [S42], using the corrected Bayesian information criterion (BIC) implemented in jModelTest v2.1.4 [S43]. ML phylogenetic trees were estimated

using RAxML v7.4.2 [S36] as implemented in raxmIGUI 1.3.1 [S37]. As RAxML only applies the GTR model, we ran our analyses under the GTR+G model, with BSbrL enabled and 500 thorough bootstrap replicates to evaluate branch support. Bayesian inference was also used to estimate trees using MrBayes v3.2.1 [S44]. Under the HKY+I+G model, two independent analyses of Metropolis-coupled MCMC (one cold and three heated chains) were run simultaneously for 10 million generations and sampled every 1000th generation. Average standard deviation of split frequencies of ~0.005 between independent runs as well as trace plots and effective sample sizes >>1000 for each run measured using Tracer v1.6 [S40] indicated rapid convergence and mixing of chains. The first 1000 (10%) samples of the posterior distribution were discarded as burn-in from each run.

Nuclear DNA sequences: Sequences from each of the 20 unlinked nuclear loci were aligned using MAFFT v7.017 [S21] with the same settings as used for the mitochondrial genome data (see above). We estimated phylogenies from phased and unphased nuclear data using both multispecies coalescent and concatenation methods. The alignments of each gene segment were phased into haplotypes using SeqPHASE [S45]. The best-fitting model of DNA substitution for each locus and the concatenated data set

(http://dx.doi.org/10.5061/dryad.3b77f/3) was chosen among 24 models and a base tree estimated with BIONJ [S42], using the corrected Bayesian information criterion (BIC), implemented in jModelTest v2.1.4 [S43].

Preliminary ML + bootstrapping analyses of individual loci using RAxML v7.4.2 [S36] resulted in gene trees with heterogeneous topologies, albeit with generally low nodal support values (results not shown). To account for heterogeneity in coalescence among gene trees, we used the Markov-chain Monte Carlo approach implemented in *BEAST [S46] to co-estimate gene trees and the species tree among the sampled taxa. The BEAUti program from the BEAST 1.7.5 program package [S38] was used to set up the MCMC run using the following parameters: *BEAST species tree ancestral reconstruction enabled prior to importation of sequence alignments; unlinked substitution models using the corresponding best-fitting models as indicated by ModelTest; clock models for each locus unlinked; trees for each locus unlinked; strict clock for each locus; unlinked clock rates; species tree prior specified as Yule process of speciation; and the piecewise linear and constant root prior was used for the population size model. The auto optimize setting was enabled in the Operators window. Three independent MCMC analyses were run for 200 million generations, sampling trees every 20,000 generations for each run. The first 10% of sampled trees were discarded as burn-in. The posterior outputs from the three runs were combined using LogCombiner [S38]. Analysis of the posterior distributions of the gene tree and species tree likelihoods and other parameters using Tracer v1.6 [S40] showed ESS values >200. The posterior probability species tree was visualized using Densitree [S47].

We also concatenated sequences of the 20 loci into a supermatrix and then used ML and BI to estimate phylogenetic trees. ML phylogenetic trees were estimated using RAxML v7.4.2 and PhyML 3.0 [S36, S48]. As PhyML cannot perform partitioned analyses, the HKY+I+G model was selected as the bestfitting model of substitution for the supermatrix, based on the corrected BIC implemented in ModelTest v2.1.4 [S43]. Only the substitution model was specified and all other parameters were estimated as a part of the PhyML analyses. Tree search settings included: starting trees estimated with BIONJ: SPR+NNI branch swapping; random starting trees =5; branch length and tree topology optimization enabled. Branch support was estimated using the approximate likelihood ratio test, Shimodaira-Hasegawa-like (aLRT, SH-like) method [S48]. For the RAXML analyses we used the partitioning scheme estimated with PartitionFinder v1.1.1 [S35] and the GTRCAT model for all concatenated loci. Branch support was assessed by bootstrap resampling with 1000 pseudo-replicates. The command line settings used were: raxmIHPC-PTHREADS -T 50 -f a -m GTRCAT -x 127897 -p 12345 -# 1000 -s Concatenated_alignmentsPhased.phy -q partition.txt -n

Concatenated_alignmentsPhased_out). We used MrBayes 3.2 [S44] for BI of the supermatrix phylogeny. We used the partitioning scheme and substitution models selected with PartitionFinder v1.1.1 [S35]. Three independent analyses of Metropolis-coupled MCMC (one cold and three heated chains) were run simultaneously for 10 million generations and sampled every 1000th generation.

Average standard deviation of split frequencies of ~0.005 between independent runs as well as trace plots and effective sample sizes >>1000 for each run measured using Tracer v1.6 [S40] indicated rapid convergence and mixing of chains. The first 1000 (10%) samples of the posterior distribution were discarded as burn-in from each run.

We used BEAST 1.7.5 [S38] to estimate divergence times from the concatenated nuclear data set using a partitioned model scheme and an uncorrelated lognormal relaxed molecular clock model. BEAUti was used to set up the MCMC run using the following parameters and settings: unlinked substitution models using the partition scheme and corresponding models as indicated by PartitionFinder v1.1.1 [S35], link clock models enabled, link trees enabled, and tree prior = Yule process of speciation. The auto optimize setting was enabled in the Operators window. The molecular clock was calibrated using two fossil priors as soft bounds: 1) a root height age spanning 9-12 mya based on the approximate split between the tribes Canini and Vulpini as suggested by the earliest known fossils of *Eucyon* (Canini) [S39]; and 2) a prior spanning 1-3 mya, which encompasses the earliest known fossils of *Canis lupus* (~1 mya) and Canis edwardii (~3 mya), the latter of which is sister to a clade containing C. aureus and C. latrans based on phylogenetic analyses of morphological data of extant and fossil taxa [S39]. Calibration priors were approximated with lognormal distributions: mean=0.0, stdev=0.7, offset=8.7 (95% HPD 9.0-11.9 Mya) for prior 1 and mean=0.0, stdev=0.5, offset=0.7 (95% HPD 1.1-3.0 Mya) for prior 2. Three
independent MCMC analyses were run for 20 million generations, with parameters sampled every 2,000 generations. The first 25% trees were discarded as burn-in from each MCMC run. Analysis of the posterior distributions of tree likelihood, substitution, clock and other parameters using Tracer v1.6 [S40] showed ESS values >500 for each of the three runs. The post-burn-in posterior samples from each run were combined using LogCombiner [S38]. The maximum clade credibility tree with mean node heights was visualized using FigTree v1.4.0 [S41]. Lastly, we ran an analysis with sampling from the prior distribution only (without sequence data), with an MCMC run for 30 million generations and sampling each 3,000 generations.

Sequence characteristics and tests of alternative topologies

We calculated the number of variable and informative sites, as well as the mean empirical base frequencies, for the cytochrome *b*, mitochondrial genome (13,890bp) and nuclear DNA (concatenation of the 20 autosomal gene segments) data sets using PAUP* 4.0a145 [S49]. The table reporting these sequence characteristics has been deposited in the Dryad Digital Repository http://dx.doi.org/10.5061/dryad.3b77f/4).

We compared the topologies derived from maximum likelihood analyses of the cytochrome *b*, mitochondrial genome (13,890bp) and concatenated nuclear (13,727bp) data sets against topologies in which sequences of golden jackals from Africa and Eurasia were topologically constrained to be monophyletic. For each data set, trees files in Newick format were revised so that golden jackals from Africa and Eurasia were enforced to be monophyletic, while all other taxa were unconstrained. These tree files were then analyzed using RAxML [S36] and the original substitution models and data partitioning schemes to implement the non-parametric Shimodaira-Hasegawa test [S50, S51] ('-f h' option) to compare the likelihood values of the best tree (from the unconstrained analyses) and the constrained tree (all golden jackals monophyletic). P-values of \leq 0.05% were used as threshold to test for significant differences between likelihoods of the unconstrained and constrained topologies (see Table S1).

Microsatellite samples, genotyping and analyses

Microsatellite analysis was performed with the following tissue or blood samples of golden jackals: Greece (n = 6), Croatia (n = 5), Slovenia (n = 1), Serbia (n = 1), Iran (n = 3), Israel (n = 23), Kenya (n = 23), Mauritania (n = 14), Morocco (n = 7), Algeria (n = 1) and Senegal (n = 2). Microsatellite analysis also included tissue samples of gray wolves from Eurasia: Russia (Siberia; n = 5), Romania (n = 6), Slovenia (n = 4) and Portugal (n = 5). European samples from Romania and Slovenia represent locations where gray wolves and golden jackals are sympatric and have the potential to hybridize. Dog samples included stray dogs (Portugal n = 7; Morocco n = 1 and Kenya n = 1), herding dogs (n = 9) and hound dogs (n = 4). All samples were extracted using the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA) and amplified for 38 microsatellite loci in four multiplex reactions following Godinho *et al.* [S52, S53]: AHT111 [S54], AHT132 (by N. Holmes), AHTh171, FH2848 and REN64E19 [S55], AHTk211 and AHTk253 [S56] C04.140, C09.173, C22.279 and C20.253 [S57], C08.410, C09.474, C20.446 and CXX.459 [S58], C13.758 and C14.866 [S59], FH2010, FH2054, FH2079 and FH2161 [S60], INRA21 [S61], PEZ1, PEZ3 and PEZ5 [S62], REN162C04, REN169D01, REN169O18, REN247M23 and REN54P11 [S63], VWF [S64], INU005, INU030 and INU055 (Finnzymes, Fisher Scientific, Loures, Portugal) and CPH02, CPH05, CPH09 and CPH14 [S65]. PCR products were separated by size on an ABI 3130xl capillary sequencer. Alleles were determined using GENEMAPPER v4.1 (Life Technologies, Grand Island, NY) and checked manually.

Standard genetic diversity indices were evaluated for each of the 38 loci at each population (four jackal populations: Europe, Middle East, Kenya and North Africa, wolves and dogs) using GenAlEx v 6.5 [S66]. The number of alleles (Na), observed heterozygosity (Ho), expected heterozygosity (He) and fixation index (Fis) were computed as in Weir and Cockerham [S67] (see:

http://dx.doi.org/10.5061/dryad.3b77f/5 and

http://dx.doi.org/10.5061/dryad.3b77f/6). Guo and Thompson's exact test [S68] was implemented in GENEPOP v. 3.4 [S69] to statistically evaluate deviations from Hardy–Weinberg expectations for all locus-population combinations (228 comparisons) and to test significance of association between genotypes at pairs of loci in each population (LD; 703 comparisons). Statistical significance levels were adjusted using sequential Bonferroni corrections for multiple comparisons [S70].

Individual genotypes of 128 golden jackal, gray wolf and dog samples were analysed using the Bayesian clustering procedure implemented in STRUCTURE 2.3.3 [S74] assuming one to eight clusters (K = 1 to K = 8). This analysis was performed independently 10 times for 10⁶ permutations after a burn-in period of 10⁵ permutations using the admixture model with correlated allele frequencies among populations and no priors for individual identification (POPFLAG=0). We inferred mean likelihood values for the different number of genetic clusters and estimated the delta K between successive values of K. (http://dx.doi.org/10.5061/dryad.3b77f/7). The Nei's Da genetic distance [S71] was calculated between all pairs of individuals and converted into a dendrogram using the Neighbour-Joining algorithm [S72] provided with the Populations v 1.2 software [S73] and graphically displayed with FigTree v1.3.1 [S41]. This dendrogram has been deposited in the Dryad Digital Respository (http://dx.doi.org/10.5061/dryad.3b77f/8).

ZFX and ZFY final intron amplification and sequencing

We sequenced the complete final intron for the zinc-finger X-chromosomal gene (*ZFX*, 834-839 bp) for a subset of our samples: eight gray wolves, two domestic

dogs (Basenji and Dingo), and 21 golden jackals (Eurasia, n = 13, Africa, n = 8). We also sequenced the complete final intron for the zinc-finger Y-chromosomal gene (*ZFY*, 924-1146 bp) for the subset of individuals that were male: both domestic dogs, gray wolves (n =5), and golden jackals (n= 13). We used two different sets of primers to amplify each of the *ZFX* and *ZFY* final introns using the previously reported primers and amplification conditions [S75]. Amplification products were cycle sequenced with both forward and reverse primers using BigDye 3.1 and then sequenced on a 3730xl DNA Analyzer (Life Technologies, Grand Island, NY). Sequences were assembled, edited and aligned with Geneious Pro 5.5.6 [S20] and aligned to previously published sequences [S75]. Haplotype trees were generated with Geneious Pro 5.5.6 [S20] and networks were constructed using Haploviewer [S76]. All deletions and insertions were considered as a single evolutionary event regardless of size.

ZFY SINE-II element presence/absence assay

The final intron of the zinc-finger Y-chromosomal gene (*ZFY*) contains two shortinterspersed elements (SINEs), one of which ("SINE-II") is present only in golden jackals, as compared to other assayed canid species [S75]. We designed a PCR assay to amplify the portion of the canid *ZFY* final intron that contains the diagnostic SINE-II element, using the forward primer U-ZF-2F [S77] and a novel reverse primer designed from the *ZFY* final intron sequences of 12 canid species [S75]: – Canidpost-*ZFY*R (5'-AAATTTCTTCACTCAGATGAAATAACA-3'). The PCR product generated is 327 bp size for reference gray wolf samples (Asian, European, North American) and 537 bp for reference Eurasian golden jackal samples (Serbian). A total of 31 male golden jackals were assayed for the presence/absence of the *ZFY* SINE-II element. PCR reactions (total volume of 25μ I) consisted of 3.75μ I DNA, 12.5μ I 2X Qiagen Multiplex Mix (Qiagen, Valencia, CA), 2.5μ I Qiagen Q solution, 2.5μ I primer mix (2μ M), 1.0μ I BSA (10mg/mI) and 2.75μ I H₂O. PCR cycle conditions were: one cycle at 95°C for 15min., 14 cycles of 94°C for 30sec, 60°C (-0.5°C/cycle) for 90sec, 72°C for 60sec, followed by 33 cycles at 89°C for 30sec, 53°C for 90sec, and 72°C for 60sec, and a final cycle at 60°C for 30min. PCR product sizes were visualized on Sybr-safe (Invitrogen, Carlsbad, CA) stained agarose gels.

Golden jackal and gray wolf genome SNP calling

Genome sequences of three gray wolves from China (24.6X coverage), Croatia (25.3X) and Israel (21.6X), two domestic dog breeds, Basenji (12.6X) and Dingo (25.8X), and one golden jackal from Israel (23.8X) were individually aligned to the dog genome (CanFam 3.0) using BioScope v1.2 (Life Technologies, Grand Island, NY) for SOLiD data, or Novalign v2.07.11 (www.novocraft.com) for Illumina HiSeq2000 data as part of the study by Freedman et al. [S17]. The resulting aligned sequences for each individual (as SAM files) were sorted, PCR duplicates removed, and local realignment performed with samtools and picard-tools, then formatted as .bam files [S78]. For the set of six aligned genome

sequences from the Freedman et al. study, we utilized the genotypes for each genome that had already been called; see details of sequence alignment and genotyping pipeline in Text S3 of Freedman et al. [S17]. We converted the coordinates of the six genomes from the Freedman et al. study to CanFam3.1 based on our custom python script.

The genome of a golden jackal from Kenya was also sequenced (25.77X coverage) using the Illumina HiSeq2000, and the sequence reads aligned to CanFam 3.1 using Bowtie2 [S18] under very-sensitive mode and saved as a .bam format file, following the pipeline described in Freedman et al. [S17]. We separately analyzed the aligned sequences and called genotypes in the Kenyan golden jackal using GATK [S79] per the protocol of Freedman et al. [S17]. After merging the seven genomes into a single .vcf file, we applied several filters to remove false positive and low quality SNPs [see S17]. The result was a set of 7,675,363 SNPs for the seven-canid genome dataset.

Whole genome pairwise distance comparisons using non-overlapping window approach

The level of diversity and divergence within and between the gray wolf and golden jackal genomes was estimated by pairwise comparison of nonoverlapping 50kb windows along the whole aligned genomes (41,999 windows total). Pairwise distance for each window for all 10 pairwise genome combinations was calculated following the method of Gronau et al. [S80]. Histograms of divergence were generated by comparing numbers of windows for number of pairwise distances per 50kb window.

Pruning of SNPs due to linkage disequilibrium and principal component analysis (PCA)

The 7,675,363 SNP dataset was pruned to remove SNPs with high pairwise genotypic association (r^2) (i.e. high linkage disequilibrium or LD) prior to principal component analysis of the three gray wolves and two golden jackals. Highly linked SNPs with an $r^2 > 0.2$ were removed from the variant calls of the seven-canid dataset using PLINK [S81] with the setting "indep-pairwise 50 5 0.2" (50 SNP window, 5 SNP step shift, and $r^2>0.2$ prune of any SNP pair within the window) resulting in a subset of 264,937 SNPs after pruning. We then used the smartpca program distributed in the EIGENSTRAT package [S82] for PCA to visualize the dominant relationships among the 264,937 SNP genotype data.

Estimation of *D*-statistic between golden jackals and gray wolves

We used a Channel Island fox (*Urocyon littoralis*) as an outgroup to define the ancestral allele in our *D*-statistic analyses. The genome of a single Channel Island fox was sequenced (R. Wayne et al. manuscript in prep.) and SNPs were annotated as described above. We merged reads with overlapping sequences, where a portion of the template DNA molecule was observed in both the forward and reverse read, using MergeReadsFastQ_cc provided by M. Kircher,

combining overlapping paired-end reads into a single read and reevaluating base quality scores [S83]. For paired-end reads that did not merge, we trimmed bases from the 3' tail until we encountered a base of greater than 95% confidence using Trimmomatic [S84]. If either the forward or reverse read was trimmed to less than 36bp we discarded the pair, to minimize multiple mappings. We mapped the remaining high quality paired end and merged reads to the reference dog genome [S16] using bwa [S85] and removed sequence duplicates from library preparation with samtools [S78]. We then called the genotype of the Channel Island fox at all sites where the dogs, wolves and jackals in our panel varied using the GATK UnifiedGenotyper [S86].

We tested for admixture between canids in our study using the *D*-statistic test [S87, S88] (see: http://dx.doi.org/10.5061/dryad.3b77f/9). The *D*-statistic is the excess of shared-derived alleles between two conspecific or closely related individuals (P1 and P2), and a more distantly-related candidate admixing individual (P3). An outgroup to all of the other taxa in the comparison is used to define the ancestral state. When the outgroup is heterozygous, the ancestral state is considered ambiguous and the site is excluded from analysis. Specifically, the *D*-statistic compares the number of ABBA sites where P2 and P3 share a derived allele but P1 has the ancestral allele and BABA sites where P1 and P3 share a derived allele but P2 has the ancestral allele.

Equation: D=(ABBA-BABA)/(ABBA+BABA) (1)

In the absence of admixture *D* is expected to equal zero. Negative values indicate gene flow between P1 and P3 while positive values indicate gene flow between P2 and P3. We calculated *D*-statistic values for all combinations of individuals consistent with the consensus phylogeny (Figure 3C and Table S4) from our genotype calls. To test for results significantly different from zero we used a weighted block jackknife with 5Mb nonoverlapping blocks [S87, S89, S90]. *D*-statistic results differing from zero by more than 3 times the standard error (Z>3) are considered significant evidence of gene flow. Although P3 is commonly termed a candidate admixing individual, the *D*-statistic does not assign a direction to gene flow. An additional limitation of the *D*-statistic test is that it is a relative test and so if P1 and P2 share the same amount of admixture with P3 no admixture will be detected.

Additional evidence for genetic admixture in Israeli golden jackals comes from comparisons of our cytochrome *b*, nuclear DNA, microsatellite, and *ZFX/ZFY* sequence results. Five golden jackals from Israel show a discordant relationship between their *ZFX/ZFY* haplotypes and their phylogenetic position in the cytochrome *b* gene tree. For example, the "Israel 39" golden jackal in Table S2 has the same *ZFY* haplotype as other Eurasian golden jackals but has a *ZFX* and cytochrome *b* haplotype that also suggests a genetic ancestry with golden jackals found in Africa. Moreover, the seven Israeli golden jackal haplotypes that group with golden jackals from Africa in the cytochrome *b* tree (Figure 1A) cluster with other golden jackals from Israel in the microsatellite results, suggesting historical but no recent admixture. Lastly, one of the golden jackals used in the nuclear DNA sequence analyses (sample ID = Israel 37) was grouped with the other three golden jackals from Eurasia in both the concatenated and multispecies coalescent trees (Figures 2 and S2), had a *ZFX* genotype and a *ZFY* haplotype characteristic of Eurasian golden jackals (Table S2), but had a cytochrome *b* haplotype (*Canis aureus* Israel 6) that was grouped in the clade containing golden jackals (and *Canis lupus lupaster*) sequences from Africa, again suggesting past admixture between the two golden jackal lineages.

Reconstruction of historical patterns of effective population size

To evaluate and contrast historical patterns of effective population size in golden jackals and gray wolves, we employed the pairwise sequential Markovian coalescent (PSMC) method [S91]. This method infers ancestral effective population sizes (*Ne*) over time, based on a probabilistic model of coalescence with recombination and changes in heterozygosity rates along a single diploid genome. We applied PSMC to each of the two golden jackal genomes and three gray wolf genomes and converted the mutation-scaled estimates of time (to years) and population size (to numbers of individuals) by assuming an average mutation rate per generation of 1.0 x 10^{-8} and an average generation time of three years [S16, S17, S92].

Morphometric analysis of golden jackals in Africa and Eurasia

To assess whether golden jackals showed significant differences in morphology across their geographic range, especially between localities in Africa versus those in Eurasia, we re-analyzed skull and teeth measurements originally reported in the study by Van Valkenburgh and Wayne [S93] on shape divergence among sympatric jackal species. The original data included 62 measurements of the cranium, mandibles, and teeth taken from museum specimens of 171 golden jackals from throughout their geographic range in Africa and Eurasia. Concatenation of all variables reduced the sample size to 149 individuals (22 excluded due to missing data) and six dental measurements were excluded because of significant differences due to sexual dimorphism. Morphometric analysis of 45 log-transformed variables with 140 individuals was conducted with PAST 2.17 software [S94] using principal component analysis (PCA) (data deposited in Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.3b77f/10). A PCA plot was constructed using the variance-covariance matrix and iterative imputation of missing values. Significance of the principal components was assessed by using 1000 bootstrap replications. Discriminant analysis of two predefined groups (African and Eurasian jackals) was conducted using the same

software and its significance was assessed using a two-group permutation test with 3000 permutations. The raw data analysis revealed that populations differed in various dental and skull proportions. Consequently, we computed nine ratios describing relative tooth size and shape, as well as skull shape. All ratios were arcsine-transformed and used in the PCA as described above (Table S7). The nine ratios were:

RelC1Ci = maximum mediolateral breadth of the rostrum across the upper canines divided by skull length;

P4SH = shape of the upper fourth premolar as estimated by maximum width across the protocone divided by maximum mesiodistal length;

LP3SH = shape of the lower third premolar as estimated by maximum width divided by maximum mesiodistal length;

RBL = relative blade length of the lower first molar as estimated by trigonid length divided by maximum mesiodistal length;

URGA = relative grinding area of the upper tooth row as estimated by the square root of the sum of areas of the first and second upper molars divided by the mesiodistal length of the upper fourth premolar;

LRGA = relative grinding area of the lower tooth row as estimated by the square root of the sum of areas of the talonid of the first lower molar and entire second lower molar divided by the mesiodistal length of the trigonid of the first lower molar;

SKLSH = skull shape as estimated by maximum width acoross the zygomatic arches divided by skull length;

IOBMCW = minimum mediolateral width between the orbits (interorbital breadth) divided by maximum mediolateral width of the braincase.

The table showing the loadings for nine arcsine-transformed craniodental ratios included in PCA analysis has been deposited in the Dryad Digital Repository (http://dx.doi.org/10.5061/dryad.3b77f/11).

Taxonomy

Following the first formal description and classification of the golden jackal of Eurasia as *Canis aureus* by Linnaeus in 1758 [S95; type locality later restricted to the Benna Mountains, Laristan, Southern Persia (Iran), (S96)], various additional taxonomic names long included in the synonymy for this species [S97] were applied to jackals collected in other localities across Asia and Africa that were judged to differ from typical *aureus* in certain features such as pelage coloration or body size. In Africa, Cuvier (1820) [S98] described a female golden jackal from Senegal that differed in several respects from golden jackals from Eurasia and other regions of Africa. He applied the name *Canis anthus* to this taxon, which is the earliest name for a specimen of a "golden jackal" from Africa. We provide the translation of the original description by Cuvier (1820):

"The Senegalese Jackal, Female"

To us, this animal appears to belong to a species that is essentially distinct from that which is correctly known as a Jackal, the Canid that is found in central and southern Asia, and perhaps all across Africa, which gathers in a pack and eats carcasses, and which we believe to have pictured under the name of "male Jackal."

The name, *Chacal du Sénégal*¹ [Senegalese Jackal], is doubtless incorrect, even more so because the true Jackal is also apparently found in Senegal; but since we received this animal under this name, and not knowing its name in its own country, we have not found it inappropriate to keep it while we wait until we are able to give it a more accurate name.

Nothing is as obscure as this branch of the family of Canids of the old continent, which is made up of the Jackal, the *Adive*, the *Corsac*, the *Mésomélas*, etc., even though it is natural, and the structure of the species making it up is uniform. It is to be assumed that the animal that is the subject of this article was never distinguished from the Jackal by the Europeans living in Senegal; and we ourselves might not have appreciated the differences that characterize it, if we had not had them alive at the same time, and if we had not been able to compare one directly with the other; to be convinced that our Senegalese Jackal definitely does not belong to one of the species we named above, we first had to rigorously determine their defining characteristics.

The *Corsac* and the *Adive* are in no way different from each other, as M. G. Cuvier presumed, if the *Adive* is this small species of the *Chien de l'Inde*,¹ called Nougi-Hari in Malabar.² Actually, the Museum collection holds several of these Canids, which were sent to it by M. Léchenault. When they are compared to the description that Guldenstaet made of the *Corsac*, and the description Buffon

made of the animal he calls *Isatis* (Suppl., Volume III, pages 113 and 114), which had been sent to him under its true name of *Corsac*, it is not possible to discern any difference between them.

The *Corsac* is no larger in size than the Marten, and its tail, which is very long in proportion to its body, falls, when fully hanging, to three inches below its feet. The entire upper portion of its body and its tail are of a uniform fawn-grey color with a very soft tint: this coloring is the result of fawn and white rings that cover most of the visible part of the hair; however, some of these rings are black. The legs are entirely fawn, the tip of its tail is black, and on the upper side of the tail, about three inches from its base, one can see a small black spot. The entire lower part of its body is yellowish white.

The true Jackal is twice as large as the *Corsac*, and its tail hangs only as far down as its heels. The pelt on the upper parts of its body is also made up of hair covered by rings alternating between black, fawn, and white; but the result is not the uniform soft grey tint that the *Corsac* has; in certain places it is darker, and in others, less so; however there is nothing regular in the distribution of these different hues; in this regard, the Jackal rather resembles the Wolf. This grey color becomes paler on the neck, shoulders, and thighs, and the legs are fawn. The tail is grey-brown, with its lower third being black. The underside of its body and the inside surface of the legs are a dirty white.

The *Mésomélas* is also gray and tan; it is similar in size to the Jackal, and its tail almost reaches the ground. The hair on its back, like that of the above

species, is covered with fawn, black, and white rings; but since these rings are generally very wide, the resulting color is even less uniform than that of the Jackal's upper body parts. These are irregular patches of white and black, and they contrast strongly with each other and with the brilliant [reddish] tan of the other body parts. Its fawn-colored tail also has a black tip; the color of the back, which in the front descends as far as the shoulders, becomes narrower toward the back, so that above the rump it is no more than two inches wide. The ears of the *Mésomélas* are twice as large as those of the Jackal.

These three species having been clearly described, I can now describe the one whose picture I publish here today. One can see that it cannot be confused with the Jackal: it is much larger than the *Corsac* and lacks the dorsal triangle of the *Mésomélas*. Now this new animal differs essentially from the Jackal, as can be seen when their two pictures are compared, and as the following description will make even more evident.

Its proportions are more elegant than those of the Jackal, and it is lighter in form. Its height at the middle of its back is 15 inches, its length from the base of the tail to the neck is 14 inches; the length of its head from the occiput to the tip of the nose is 7 inches, and the length of its tail is 10 inches. Its back and sides are covered with dark grey fur, marked with some yellowish coloring; the hair is covered with black and white rings, as a well as a few tan rings; this grey color is not uniformly dispersed, because of the length of the hair, which is separated into locks, revealing sometimes the white, and sometimes the dark, hair. Its neck is a greyish fawn color that becomes greyer still on the head, especially on the jowls under the ears. The top of the muzzle, the front and rear legs, the backs of the ears, and the tail are of a quite pure fawn color; only one sees a longitudinal black mark running down the upper third of the tail, and a few black hairs, very few indeed, at its tip. The area under its lower jaw, and its throat, chest, belly and the inner sides of its legs are whitish. The hair on the neck and tail is very long, a little less so on its flanks and rump, and the hair on its head and legs is short. In general, it runs from front to back, except between the front legs, where it emerges from back to front.

This animal resembles a Dog in every way; it usually holds its tail as shown in our picture; however, when it experiences fear, it quickly puts it between its legs while baring its teeth. However, this threatening stance is not an indication of anger, once one reassures it with a few words, it approaches, whines, and licks ones hand. Its voice is fairly soft, and its sound is prolonged instead of our Jackal's burst of barks. When it wants something, its cry is soft, like that of a young Dog; and if it hears other animals cry, it also cries. It gives off an odor that, although quite strong, is infinitely less marked than that of a Jackal.

All the other aspects of their structure resemble those of Canids in general; therefore I will not repeat what I already said about them in the article on the Jackal, in the understanding that I am commenting on the generic characteristics of these animals. Travelers have undoubtedly spoken of this animal, which they appear to have confused with the Jackal; however, it is a point on which they do not agree, and this could be because some of them may have heard mention of our new species. Some report that the Jackal gives off a very strong odor, which would indicate an actual Jackal, others affirm instead that there is almost no odor, which could relate to the animal being described here.

Since this new species of Canid does not yet have its own name, I would suggest that it be named *Anthus* in the scientific catalogues. This was the name of a family in Arcadia,³ a member of which was transformed every year into a Wolf, according to the beliefs of the inhabitants of this land (Pliny, Book VIIII).

June 1820"

¹[Translator's note: literally, "Dog of India"]

²[Translator's note: Malabar is a region in south-west India]

³[Translator's note: A region in Greece]

Other specimens of "golden jackals" from Africa were given additional taxonomic names (species or subspecies) that postdate Cuvier's naming of *C. anthus*. These include the following, some of which have previously been regarded as subspecies of *C. aureus* [S97].

Synonymy (unique names as originally described, with type localities). Details from Allen [S99].

[Canis] anthus F. Cuvier, 1820. Senegal.

Thos senegalensis Hamilton Smith, 1839. Senegal.

Canis variegatus Cretzschmar, 1826. Nubia and upper Egypt (name preoccupied

by *C. familiaris variegatus* Gmelin, 1788).

Canis lupaster Hemprich and Ehrenberg, 1832. Fayum, Egypt.

Canis sacer Hemprich and Ehrenberg, 1832. Fayum, Egypt.

Canis riparius Hemprich and Ehrenberg, 1832. Coast of Abyssinia, near Arkiko [Eritrea].

Canis aureus algirensis Wagner, 1841. "Algeria."

Canis aureus tripolitanus Wagner, 1841. No locality (Tripoli implied) [Libya].

Canis hagenbecki Noack, 1886. Somaliland [Somalia].

Canis mengesi Noack, 1886. Coast of Somaliland [Somalia].

Canis anthus soudanicus Thomas, 1903. El Obeid, Kordofan [Sudan].

Canis somalicus Lorenz, 1906. Ireso, near Agada, Somaliland [Somalia].

Canis gallaënsis Lorenz, 1906. Ginea, Arussi, Ethiopia.

Canis doederleini Hilzheimer, 1906. Upper Egypt [Egypt].

Canis thooides Hilzheimer, 1906. Sennaar [Sudan].

Canis mengesi lamperti Hilzheimer, 1906. ?Somaliland.

Thos aureus bea Heller, 1914. Loita Plains [Kenya].

Thos lupaster maroccanus Cabrera, 1921. Mogador, Morocco.

Thos aureus nubianus Cabrera, 1921. Replacement name for *Canis variegatus* Cretzschmar, 1826.

Therefore, *Canis anthus* (Cuvier, 1820) is the earliest name applied to a golden jackal specimen from Africa, which pre-dates the next available name, *lupaster* (Hemprich and Ehrenberg, 1833) by thirteen years. Consequently, we recommend that golden jackals originating from localities in Africa be formally recognized as *Canis anthus*. Additional samples of "golden jackals" from localities in Africa not sampled in our or previous genetic studies can be used to verify our phylogenetic and taxonomic hypotheses.

Incongruence between mitochondrial and nuclear DNA phylogenies

Although the phylogenies from the cytochrome *b* (Figure 1), mitochondrial genome (Figure S1) and nuclear sequence (Figures 2 and S2) data sets are consistent in showing that golden jackals from Africa and Eurasia are not monophyletic, the topologies revealed by each of these data sets are markedly incongruent with one another, with the majority of nodes in the mitochondrial genome and nuclear sequence topologies showing strong bootstrap and/or posterior probability support. For example, (*Canis adustus, C. mesomelas*) and (*C. lupus, C. latrans*) are each grouped as sister taxa in the nuclear DNA phylogenies whereas two and three nodes separate these groupings respectively, in the mitochondrial genome tree. Furthermore, the positions of *Cuon* and *Lycaon* differ between the trees inferred from the different data types and the monophyly of the two species of South American canids (*Lycalopex* and *Speothos*) is not recovered in the mitochondrial genome phylogeny. Such discordance between

nuclear and mitochondrial histories is not uncommon and has been documented, for example, in elephants [S100], phyllostomid bats [S101] bears [S102-103], gibbons [S104], hares [S105] and many other groups [S106]. Differences in ploidy, mode of inheritance, effective population size, mutation rate, and recombination, along with processes such as incomplete lineage sorting, mitochondrial capture via ancient hybridization [e.g., S107], and selection may produce different mitochondrial and nuclear genealogical [S108-109]. Moreover, under some conditions mitochondrial DNA is more vulnerable to being distorted by these processes or under certain conditions [S110-111].

With regards to the discordance between the topologies based on the mitochondrial genome and nuclear data sets, our analyses employed models of DNA substitution and data partitioning schemes that should help to correct for mutational saturation (especially 3rd codon positions in mitochondrial protein loci), heterogeneity in base composition, and among-site rate variation (see Phylogenetic Analyses section above). Furthermore, the higher substitution rate of the mitochondrial genes combined with the short internodes, especially among the more recent species (*Canis lupus, C. latrans*, the two lineages of *C. aureus*, and *C. simensis*) can quickly obscure phylogenetic signal due to homoplasy related to mutiple substituions at the same sites, leading to a phylogeny that can differ from the topology based on nuclear sequence data [e.g., S112]. We ran additional analyses with the mitochondrial genome data (with BEAST 1.8.2, using the same data partitioning scheme and models as described above) set to

mitigate the influence of homoplasy on phylogenetic signal by excluding 3rd codon positions entirely or coding them as RY [S113]. Both analyses resulted in identical topologies, which matched the topology shown in Figure S1 with one exception: The Ethiopian wolf (*C. simensis*) and the golden jackal (*C. aureus*) from Israel were joined together as sister taxa, albeit with posterior probability <0.95, in analyses in which 3rd codon positions were excluded or RY coded. Excluding distant outgroups (i.e., the two *Vulpes* species or *Urocyon* + *Vulpes*) or taxa with long branches (*Lycalopex* and *Speothos*) from the phylogenetic analyses also does not change the topologies recovered by the mitochondrial genome and nuclear DNA data sets (results not shown).

We also considered the possibility of nuclear-mitochondrial copies (paralogues or NUMTs) contaminating our mitochondrial genome data. However, we used long-range PCR to amplify two overlapping pieces of the total mitochondrial genome, which can minimize the amplification of NUMTs [S114]. In addition, the gray wolf mitochondrial genome sequences we amplified were very similar (small distance) to gray wolf sequences published in previous studies. Also, one of the golden jackals from Kenya sequenced from long-range PCR products was 99.99% identical to the mitochondrial genome extracted from the Kenyan golden jackal that was whole-genome sequenced. Lastly, we translated the 13 protein coding genes of the mitochondrial genome into amino acids and observed no unusual signs suggesting NUMTs such as premature stop codons or other nonsense mutations. Regarding the different topologies found for the cytochrome *b* (Figure 1) and mitochondrial genome (Figure S1) data sets, since the mitochondrial genome is a single non-recombining linkage group, it might be expected that the cytochrome *b* gene phylogeny (and all other loci containing a sufficient number of informative sites) should have the same topology as that based on the larger mitochondrial genome data. However, this has been shown to be not the case, even at low taxonomic levels [S115-116]. Unlike the mitochondrial genome phylogeny, most of the deeper nodes in the cytochrome *b* phylogeny are not well supported (<50% bootstrap, <0.95 posterior probability support). However, the two topologies each show high support for the clade that includes *Canis lupus, C. latrans*, the two lineages of *C. aureus* and *C. simensis* (which is also consistent with the nuclear DNA topology – Figure 2) and with the placement of African golden jackals sister to gray wolves rather than Eurasian golden jackals.

We argue that the topologies inferred from the nuclear gene segments (Figures 2 and S2) provide a more accurate history of the species tree compared to the mitochondrial genome phylogeny for several reasons. First, with the exception of the non-monophyly of African and Eurasian golden jackals, the relationships among species of *Canis, Cuon, Lycaon, Lycalopex* and *Speothos* in the multispecies coalescent species tree and the tree based on concatenated analyses of the nuclear DNA data are almost perfectly congruent with trees based on a largely independent set of 22 concatenated nuclear gene segments (only two loci shared between the two datasets) [S16] containing a broader taxon

sample of the Canidae. Therefore, nearly non-overlapping samplings of the nuclear genome recover similar topologies among the wolf-like canids (Canis, *Cuon*, *Lycaon*). Such replication with independent evidence is not possible with the mitochondrial genome, as it effectively represents only one non-recombining locus. Second, we recovered almost identical nuclear DNA phylogenies using either a multispecies coalescent approach (Bayesian) that co-estimated the gene trees and species tree or a supermatrix approach (likelihood and Bayesian) in which the gene segments were concatenated. These trees differed only in the relative placement of *C. simensis* and the clade containing Eurasian golden jackals. This suggests that despite the different assumptions and analytical approaches underlying these methods, the nuclear gene data yield nearly similar evolutionary histories. Third, during the revision of our manuscript an unpublished preprint came to our attention, which was based on a study that examined the phylogenetic position of the African wolf (Canis lupus lupaster, which we would consider to represent *Canis anthus*) using genome-wide polymorphic sites derived from restriction-site associated DNA sequencing (RAD-seq) [S117]. Phylogenetic analyses of seven individuals of African wolf from Ethiopia, along with gray wolf, domestic dog, Ethiopian wolf, Eurasian golden jackal, and rooted with side-striped jackal, were consistent with our findings based on the 20 nuclear DNA segments. The seven African wolves and two Eurasian golden jackals defined separate clades and the authors concluded that the former represents a distinct taxon sister to a clade containing gray wolves and domestic

dog (coyote was not included in this study) [S117]. We suggest that the mitochondrial genome topology represents an accurate depiction of the gene tree for this single, non-recombining locus that differs fundamentally from the concatenated and coalescent-based species tree(s).

Supplemental References

- S1. Zachos, F. E., Cirovic, D., Kirschning, J., Otto, M., Hartl, G. B., Petersen,
 B., and Honnen, A.-C. (2009). Genetic variability, differentiation, and
 founder effect in golden jackals (Canis aureus) from Serbia as revealed by
 mitochondrial DNA and nuclear microsatellite loci. Biochem. Genet. 47,
 241–50.
- S2. Aggarwal, R. K., Kivisild, T., Ramadevi, J., and Singh, L. (2007).
 Mitochondrial DNA coding region sequences support the phylogenetic distinction of two Indian wolf species. J. Zool. Syst. Evol. Res. 45, 163–172.
- S3. Matsumura, S., Inoshima, Y., and Ishiguro, N. (2014). Reconstructing the colonization history of lost wolf lineages by the analysis of the mitochondrial genome. Mol. Phylogenet. Evol. *80*, 105–12.
- S4. Sharma, D. K., Maldonado, J. E., Jhala, Y. V, and Fleischer, R. C. (2004).
 Ancient wolf lineages in India. Proc. Biol. Sci. *271 Suppl*, S1–4.
- S5. Björnerfeldt, S., Webster, M. T., and Vilà, C. (2006). Relaxation of selective constraint on dog mitochondrial DNA following domestication.
 Genome Res. *16*, 990–994.
- S6. Meng, C., Zhang, H., and Chen, Y.-C. (2008). Sequencing and analysis of

mitochondrial genome of Chinese gray wolf (Canis lupus chanco). Chinese J. Biochem. Mol. Biol. *24*, 1170-1176.

- S7. Meng, C., Zhang, H., and Meng, Q. (2009). Mitochondrial genome of the Tibetan wolf. Mitochondrial DNA *20*, 61–3.
- S8. Rueness, E. K., Asmyhr, M. G., Sillero-Zubiri, C., Macdonald, D. W., Bekele, A., Atickem, A., and Stenseth, N. C. (2011). The cryptic African wolf: Canis aureus lupaster is not a golden jackal and is not endemic to Egypt. PLoS One *6*, e16385.
- S9. Gaubert, P., Bloch, C., Benyacoub, S., Abdelhamid, A., Pagani, P., Djagoun, C. A. M. S., Couloux, A., and Dufour, S. (2012). Reviving the African wolf Canis lupus lupaster in North and West Africa: a mitochondrial lineage ranging more than 6,000 km wide. PLoS One *7*, e42740.
- S10. Zhang, H., Zhang, J., Chen, L., and Liu, G. (2014). The complete mitochondrial genome of Chinese Shinjang wolf. Mitochondrial DNA 25, 106–8.
- S11. Zhang, H., Zhang, J., Zhao, C., Chen, L., Sha, W., and Liu, G. (2015).
 Complete mitochondrial genome of Canis lupus campestris. Mitochondrial DNA *26*, 255–6.
- S12. Sambrook J., Fritsch E.F., and Maniatus T. (1989). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor: Cold Spring Harbor Laboratory Press).
- S13. Meyer, M., Briggs, A. W., Maricic, T., Höber, B., Höffner, B., Krause, J., Weihmann, A., Pääbo, S., and Hofreiter, M. (2008). From micrograms to

picograms: quantitative PCR reduces the material demands of highthroughput sequencing. Nucleic Acids Res. *36*, e5.

- S14. Green, R. E., Malaspinas, A.-S., Krause, J., Briggs, A. W., Johnson, P. L.
 F., Uhler, C., Meyer, M., Good, J. M., Maricic, T., Stenzel, U., et al. (2008).
 A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. Cell *134*, 416–26.
- S15. Arnason, U., Gullberg, A., Janke, A., and Kullberg, M. (2007).
 Mitogenomic analyses of caniform relationships. Mol. Phylogenet. Evol. 45, 863–74.
- S16. Lindblad-Toh, K., Wade, C. M., Mikkelsen, T. S., Karlsson, E. K., Jaffe, D. B., Kamal, M., Clamp, M., Chang, J. L., Kulbokas, E. J., Zody, M. C., et al. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature *438*, 803–819.
- S17. Freedman, A. H., Gronau, I., Schweizer, R. M., Ortega-Del Vecchyo, D.,
 Han, E., Silva, P. M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B.,
 et al. (2014). Genome sequencing highlights the dynamic early history of
 dogs. PLoS Genet. *10*, e1004016.
- S18. Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–9.
- S19. Irwin, D. M., Kocher, T. D., and Wilson, A. C. (1991). Evolution of the cytochrome b gene of mammals. J. Mol. Evol. 32, 128–44.
- S20. Geneious version 5.5.6 created by Biomatters. Available from http://www.geneious.com/
- S21. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. *30*, 3059–66.

- S22. Byrd, B.F., Cornellas, A., Eerkens, J.W., Rosenthal, J., Carpenter, T.R., Leventhal, A., and Leonard, J.A. (2013). The role of canids in ritual and domestic contexts: new ancient DNA insights from complex huntergatherer sites in prehistoric central California. J. Archaeol. Sci. 40, 2176-2189.
- S23. Leonard, J.A., Wayne, R.K., and Cooper, A. (2000). Population genetics of Ice Age brown bears. Proc. Nat. Acad. Sci. USA *97*, 1651-1654.
- S24. Väli, U., Einarsson, A., Waits, L., and Ellegren, H. (2008). To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? Mol. Ecol. *17*, 3808–17.
- S25. Amrine-Madsen, H., Koepfli, K. P., Wayne, R. K., and Springer, M. S.
 (2003). A new phylogenetic marker, apolipoprotein B, provides compelling evidence for eutherian relationships. Mol. Phylogenet. Evol. *28*, 225–240.
- S26. Lyons L.A., Laughlin T.F., Copeland N.G., Jenkins N.A., Womack J.E., and O'Brien S.J. (1997). Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. Nat. Genet. 15, 47–56.
- S27. Janecka, J. E., Miller, W., Pringle, T. H., Wiens, F., Zitzmann, A., Helgen,
 K. M., Springer, M. S., and Murphy, W. J. (2007). Molecular and genomic data identify the closest living relative of primates. Science *318*, 792–4.
- S28. Venta, P. J., Brouillette, J. a., Yuzbasiyan-Gurkan, V., and Brewer, G. J. (1996). Gene-specific universal mammalian sequence-tagged sites:
 Application to the canine genome. Biochem. Genet. *34*, 321–341.
- Stanhope, M.J., Czelusniak, J., Si, J.-S., Nickerson, J., and Goodman, M.,
 (1992). A molecular perspective on mammalian evolution from the gene encoding interphotoreceptor retinoid binding protein, with convincing

evidence for bat monophyly. Mol. Phylogenet. Evol. 1, 148–160.

- S30. Sato, J. J., Wolsan, M., Minami, S., Hosoda, T., Sinaga, M. H., Hiyama, K.,Yamaguchi, Y., and Suzuki, H. (2009). Deciphering and dating the red panda's ancestry and early adaptive radiation of Musteloidea. Mol. Phylogenet. Evol. *53*, 907–22.
- S31. Housley, D. J. E., Zalewski, Z. A, Beckett, S. E., and Venta, P. J. (2006). Design factors that influence PCR amplification success of cross-species primers among 1147 mammalian primer pairs. BMC Genomics 7, 253.
- S32. Murphy, W. J., Eizirik, E., Johnson, W. E., Zhang, Y. P., Ryder, O. A., and O'Brien, S. J. (2001). Molecular phylogenetics and the origins of placental mammals. 409, 614-618.
- S33. Johnson, W. E., Eizirik, E., Pecon-Slattery, J., Murphy, W. J., Antunes, A., Teeling, E., and O'Brien, S. J. (2006). The late Miocene radiation of modern Felidae: a genetic assessment. Science *311*, 73–7.
- S34. Flynn, J. J., and Nedbal, M. A. (1998). Phylogeny of the Carnivora (Mammalia): congruence vs incompatibility among multiple data sets. 9, 414–426.
- S35. Lanfear, R., Calcott, B., Ho, S. Y. W., and Guindon, S. (2012).
 Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. *29*, 1695– 701.
- S36. Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models.
 Bioinformatics *22*, 2688–90.
- S37. Silvestro, D., and Michalak, I. (2011). raxmlGUI: a graphical front end for RAxML. Org. Divers. Evol. *12*, 335–337.
- S38. Drummond, A. J., Suchard, M. A, Xie, D., and Rambaut, A. (2012).

Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. *29*, 1969–73.

- S39. Tedford, R.H., Wang, X., and Taylor, B.E. (2009). Phylogenetic
 Systematics of the North American Fossil Caninae (Carnivora: Canidae).
 Bull. Am. Museum Nat. Hist. 325, 1–218.
- S40. Rambaut A., Suchard M. A., Xie D., and Drummond A. J. (2014). Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer.
- S41. Rambaut, A. (2006–2009). FigTree v1.3.1. Available from http://tree.bio.ed.ac.uk/software/figtree/
- S42. Gascuel, O. (1997). BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol. Biol. Evol. 14, 685–95.
- S43. Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest
 2: more models, new heuristics and parallel computing. Nat. Methods *9*, 772.
- S44. Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A.,
 Höhna, S., Larget, B., Liu, L., Suchard, M. A, and Huelsenbeck, J. P.
 (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. *61*, 539–42.
- S45. Flot, J.-F. (2010). Seqphase: a Web Tool for Interconverting Phase
 Input/Output Files and Fasta Sequence Alignments. Mol. Ecol. Resour. 10, 162–6.
- S46. Heled, J., and Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. *27*, 570–80.
- S47. Bouckaert, R. R. (2010). DensiTree: making sense of sets of phylogenetic trees. Bioinformatics 26, 1372–3.
- S48. Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W., and Gascuel O. (2010). New algorithms and methods to estimate maximum

likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307-21.

- S49. Swofford, D.L. (2002). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.
- S50. Shimodaira, H., and Hasegawa, M. (1989). Multiple comparisons of loglikelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.
- S51. Goldman, N., Anderson, J.P., and Rodrigo, A.G. (2000). Likelihood-based tests of topologies in phylogenetics. Syst. Biol. *49*, 652–670.
- S52. Godinho, R., Llaneza, L., Blanco, J. C., Lopes, Álvares F., García, E. J., Palacios, V., Cortés, Y., Talegón, J. and Ferrand, N. (2011). Genetic evidence for multiple events of hybridization between wolves and domestic dogs in the Iberian Peninsula. Mol. Ecol. *20*, 5154–5166.
- S53. Godinho, R., López-Bao, J. V., Castro, D., Llaneza, L., Lopes, S., Silva, P., and Ferrand, N. (2014). Real-time assessment of hybridization between wolves and dogs: combining non-invasive samples with ancestry informative markers. Mol. Ecol. Resour. doi: 10.1111/1755-0998.12313.
- S54. Holmes, N.G., Mellersh, C. S., Humphreys, S. J., Binns, M. M., Holliman,
 A., Curtis, R., Sampson, J. (1993). Isolation and characterization of
 microsattelites from the canine genome. Anim. Genet. *24*, 289-292.
- S55. Breen, M., Jouquand, S., Renier, C., Mellersh, C. S., Hitte, C., Holmes, N. G., Chéron, a, Suter, N., Vignaux, F., Bristow, A. E., et al. (2001).
 Chromosome-specific single-locus FISH probes allow anchorage of an 1800-marker integrated radiation-hybrid/linkage map of the domestic dog genome to all chromosomes. Genome Res. *11*, 1784–95.

- S56. Lingaas, F., Sorensen, A., Juneja, R. K., Johansson, S., Fredholm, M.,
 Winterø, A. K., Sampson, J., et al. (1996). Towards construction of a canine linkage map: Establishment of 16 linkage groups. Mamm. Genom. *8*, 218-223.
- S57. Ostrander, E.A., Sprague, G.F. and Rine, J. (1993). Identification and characterization of dinucleotide repeat (ca)n markers for genetic-mapping in dog. Genomics *16*, 207-213.
- S58. Ostrander, E.A., Mapa, F.A., Yee, M., and Rine, J. (1995). One hundred and one new simple sequence repeat-based markers for the canine genome. Mamm. Genom. *6*, 192-195.
- S59. Mellersh, C.S., Langston, A.A., Acland, G.M., Fleming, M.A., Ray, K.,
 Wiegand, N.A., Fransisco, L.V., Gibbs, M., Aguirre, G.D., and Ostrander,
 E.A. (1997). A linkage map of the canine genome. Genomics *46*, 326–336.
- S60. Francisco, L. V., Langston, A. A., Mellersh, C. S., Neal, C. L., and Ostrander, E.A. (1996). A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. Mamm. Genom. *7*, 359-362.
- S61. Mariat, D., Kessler, J. L., Vaiman, D., and Panthier, J. J. (1996).
 Polymorphism characterization of five canine microsatellites. Anim. Genet. 27, 434–435.
- S62. Neff, M.W., Broman, K. W., Mellersh, C. S., Ray, K., Acland, G. M., Aguirre, G. D., Ziegle, J. S., Ostrander, E. A., and Rine, J. (1999). A second-generation genetic linkage map of the domestic dog, *Canis familiaris*. Genetics *151*, 803–820.
- S63. Guyon, R., Lorentzen, T. D., Hitte, C., Kim, L., Cadieu, E., Parker, H. G., Quignon, P., Lowe, J. K., Renier, C., Gelfenbeyn, B., et al. (2003). A 1-Mb resolution radiation hybrid map of the canine genome. Proc. Natl. Acad. Sci. U. S. A. *100*, 5296–301.

- S64. Shibuya, H., Collins, B. K., Huang, T. H., and Johnson, G.S. (1994). A polymorphic (AGGAAT)ⁿ tandem repeat in an intron of the canine von Willebrand factor gene. Anim. Genet. *25*, 122.
- S65. Fredholm, M., and Winterø, A.K. (1995). Variation of short tandem repeats within and between species belonging to the Canidae family. Mamm. Genom. *6*, 11-18.
- S66. Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics, *28*, 2537-2539.
- S67. Weir, B. S., and Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population structure. Evolution *38*, 1358-1370.
- S68. Guo S. W., and Thompson, E. A. (1992). Performing the exact test ofHardy-Weinberg proportions for multiple alleles. Biometrics *48*, 361–372.
- S69. Raymond, M., and Rousset, F. (1995). GENEPOP (Version 1.2):
 population genetics software for exact tests and ecumenicism. J. Hered.
 86, 248–249.
- S70. Rice, W. R. (1989). Analyzing tables of statistical tests. Evolution 43, 223–225.
- S71. Nei, M., Tajima, F., and Tateno, Y. (1983), Accuracy of estimated phylogenetic trees from molecular data. J. Mol. Evol. *19*, 153-170.
- S72. Saitou, N., and Nei, M. (1987). The Neighbor-Joining method A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406-425.
- S73. Langella, O. (1999). POPULATIONS 1.2.30 Population genetic software.CNRS UPR9034. Available at bioinformatics.org/~ tryphon/populations.
- S74. Falush, D., Stephens, M., and Pritchard, J. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics *164*, 1567–1587.

- S75. Tsubouchi, A., Fukui, D., Ueda, M., Tada, K., Toyoshima, S., Takami, K., Tsujimoto, T., Uraguchi, K., Raichev, E., Kaneko, Y., et al. (2012).
 Comparative molecular phylogeny and evolution of sex chromosome DNA sequences in the family Canidae (Mammalia: Carnivora). Zoolog. Sci. 29, 151–61.
- S76. Ewing, G. Haploviewer. Available from http://www.cibiv.at/~greg/haploviewer
- S77. Nakagome, S., Pecon-Slattery, J., and Masuda, R. (2008). Unequal rates of Y chromosome gene divergence during speciation of the family Ursidae.
 Mol. Biol. Evol. *25*, 1344–56.
- S78. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N.,
 Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence
 Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–9.
- S79. DePristo, M. a, Banks, E., Poplin, R., Garimella, K. V, Maguire, J. R.,
 Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., et al.
 (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. *43*, 491–8.
- S80. Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G., and Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. Nat. Genet. 43, 1031–4.
- S81. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. a R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., et al. (2007). PLINK: a tool set for whole-genome association and populationbased linkage analyses. Am. J. Hum. Genet. *81*, 559–75.
- S82. Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–9.

- S83. Kircher, M. (2012). Analysis of high-throughput ancient DNA sequencing data. Methods Mol. Biol. 840, 197-228.
- S84. Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-2120.
- S85. Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589–95.
- S86. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297– 1303.
- S87. Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher,
 M., Patterson, N., Li, H., Zhai, W., Fritz, M. H.-Y., et al. (2010). A draft
 sequence of the Neandertal genome. Science *328*, 710–22.
- S88. Durand, E. Y., Patterson, N., Reich, D., and Slatkin, M. (2011). Testing for ancient admixture between closely related populations. Mol. Biol. Evol. 28, 2239–52.
- S89. Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. Ann. Stat. 17, 1217-1241.
- S90. Busing, F. M. T. A., Meijer, E., and Ven Der Leeden, R. (1999). Delete-m jackknife for unequal m. Stat. Comput. 9, 3-8.
- S91. Li, H., and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. Nature 475, 493–6.
- S92. Skoglund, P., Götherström, A., and Jakobsson, M. (2011). Estimation of population divergence times from non-overlapping genomic sequences: examples from dogs and wolves. Mol. Biol. Evol. 28, 1505–17.
- S93. Van Valkenburgh B., and Wayne R.K. (1994). Shape divergence
associated with size convergence in sympatric East African jackals. Ecology *75*, 1567–1581.

- S94. Hammer, Ø., Harper, D.A.T., Ryan, P.D. 2001. PAST: Paleontological statistics software package for education and data analysis.
 Palaeontologia Electronica 4, 1-9.
- S95. Linnaeus, C. (1758). Systema Naturæ Per Regna Tria Naturæ, Secundum Classes, Ordines, Genera, Species, cum Characteribus,
 Differentiis, Synonymis, Locis Tomus I (Stockhom: Laurentius Salvius).
- S96. Thomas, O. (1911). The mammals of the tenth edition of Linnaeus; an attempt to fix the types of the genera and the exact bases and localities of the species. Proc. Zool. Soc. Lond. *81*, 120-158.
- S97. Wozencraft W.C. (2005). Order Carnivora. In Mammal Species of the World: A Taxonomic and Geographic Reference, D.E. Wilson D. E. and D.M. Reeder, eds. (Baltimore: The Johns Hopkins University Press), pp. 279-348.
- S98. Geoffroy-Saint-Hilaire, E., and Cuvier, F. (1824). Histoire Naturelle desMammifères , Tome Deuxième (Paris: Chez A. Belin).
- S99. Allen, G.M. (1939). A checklist of African mammals. Bull. Mus. Comp. Zool. Harvard *83*, 1-763.
- S100. Roca, A. L., Georgiadis, N., and O'Brien, S. J. (2005). Cytonuclear genomic dissociation in African elephant species. Nat. Genet. *37*, 96–100.

- S101. Dávalos, L. M., Cirranello, A. L., Geisler, J. H., and Simmons, N. B. (2012).
 Understanding phylogenetic incongruence: lessons from phyllostomid bats.
 Biol. Rev. *87*, 991-1024.
- S102. Hailer, F., Kutschera, V. E., Hallström, B. M., Klassert, D., Fain, S. R., Leonard, J. A., Arnason, U., and Janke, A. (2012). Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. Science *336*, 344–7.
- S103. Kutschera, V. E., Bidon, T., Hailer, F., Rodi, J. L., Fain, S. R., and Janke,
 A. (2014). Bears in a Forest of Gene Trees: Phylogenetic Inference Is
 Complicated by Incomplete Lineage Sorting and Gene Flow. Mol. Biol.
 Evol. *31*, 2004–2017.
- S104. Carbone, L., Alan Harris, R., Gnerre, S., Veeramah, K. R., Lorente-Galdos, B., Huddleston, J., Meyer, T. J., Herrero, J., Roos, C., Aken, B., et al. (2014). Gibbon genome and the fast karyotype evolution of small apes. Nature *513*, 195–201.
- S105. Melo-Ferreira, J., Boursot, P., Suchentrunk, F., Ferrand, N., and Alves, P.
 C. (2005). Invasion from the cold past: extensive introgression of mountain hare (Lepus timidus) mitochondrial DNA into three other hare species in northern Iberia. Mol. Ecol. *14*, 2459–64.
- S106. Toews, D. P. L., and Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. *21*, 3907–30.

- S107. Good, J. M., Hird, S., Reid, N., Demboski, J. R., Steppan, S. J., Martin-Nims, T. R., and Sullivan, J. (2008). Ancient hybridization and mitochondrial capture between two species of chipmunks. Mol. Ecol. *17*, 1313–27.
- S108. Funk, D. J., and Omland, K. E. (2003). Species-level paraphyly and polyphyly : Frequency, Causes, and Consequences, with Insights from Animal Mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34, 397–423.
- S109. Ballard, J. W. O., and Whitlock, M. C. (2004). The incomplete natural history of mitochondria. Mol. Ecol. *13*, 729–744.
- S110. Chan, K. M. A., and Levin, S. A. (2005). Leaky prezygotic isolation and porous genomes : rapid introgression of maternally inherited DNA. Evolution *59*, 720–729.
- S111. Currat, M., Ruedi, M., Petit, R. J., and Excoffier, L. (2008). The hidden side of invasions: massive introgression by local genes. Evolution *62*, 1908–20.
- S112. McCracken, K., and Sorenson, M. (2005). Is homoplasy or lineage sorting the source of incongruent mtDNA and nuclear gene trees in the stiff-tailed ducks (Nomonyx-Oxyura)? Syst. Biol. *54*, 35–55.
- S113. Phillips, M. J., and Penny, D. (2003). The root of the mammalian tree inferred from whole mitochondrial genomes. Mol. Phylogenet. Evol. 28, 171–185.

- S114. Triant, D., and DeWoody, J. (2007). The occurrence, detection, and avoidance of mitochondrial DNA translocations in mammalian systematics and phylogeography. J. Mammal. *88*, 908–920.
- S115. Duchêne, S., Archer, F. I., Vilstrup, J., Caballero, S., and Morin, P. a (2011). Mitogenome phylogenetics: the impact of using single regions and partitioning schemes on topology, substitution rate and divergence time estimation. PLoS One *6*, e27138.
- S116. Meiklejohn, K. A., Danielson, M. J., Faircloth, B. C., Glenn, T. C., Braun,
 E. L., and Kimball, R. T. (2014). Incongruence among different
 mitochondrial regions: a case study using complete mitogenomes. Mol.
 Phylogenet. Evol. *78*, 314–23.
- S117. Rueness, E. K., Trosvik, P., Atickem, A., Sillero-Zubiri, C., and Trucchi, E. (2015). The African wolf is a missing link in the wolf-like canid phylogeny. bioRxiv doi: http://dx.doi.org/10.1101/017996