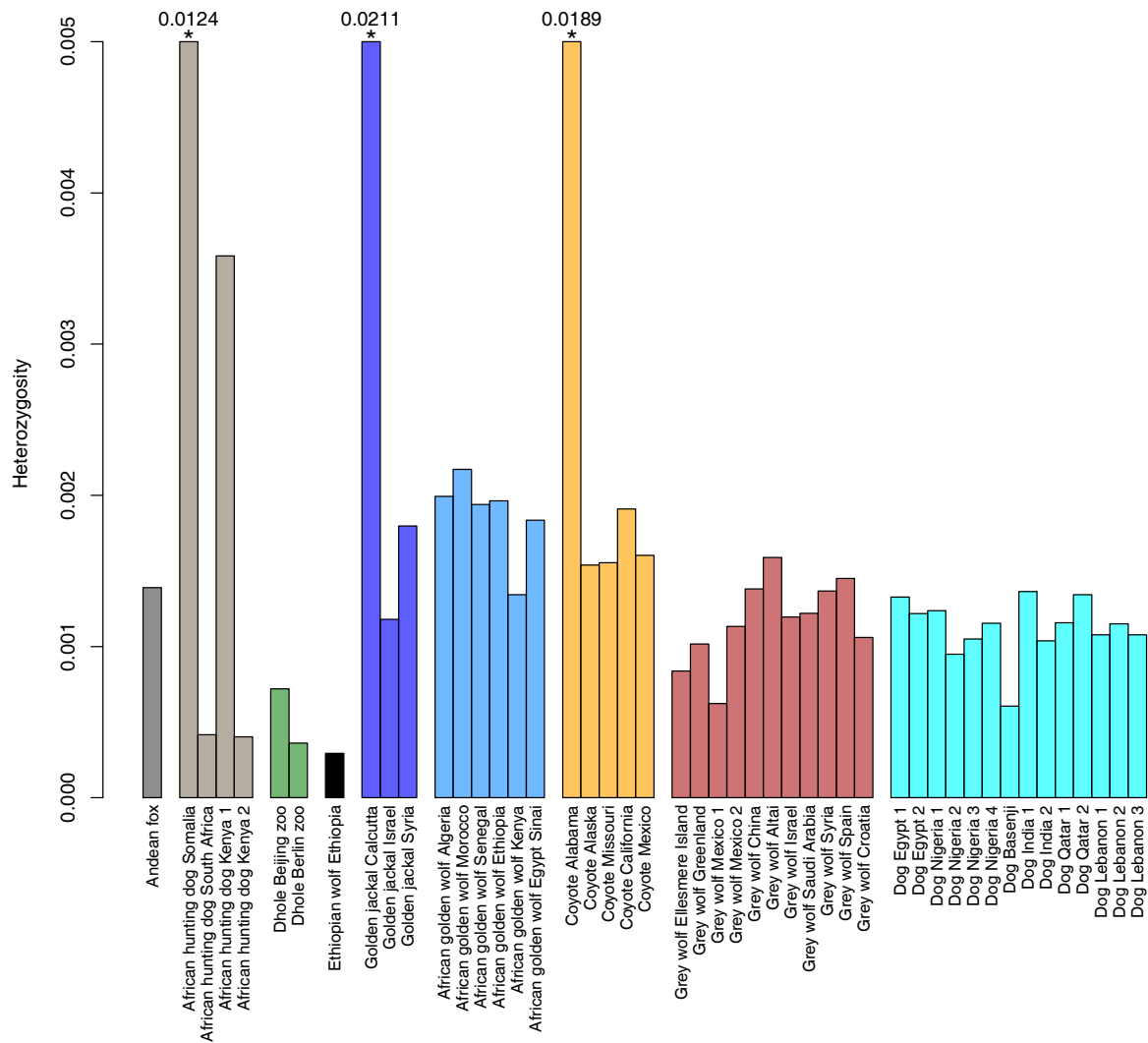


Current Biology, Volume 28

## Supplemental Information

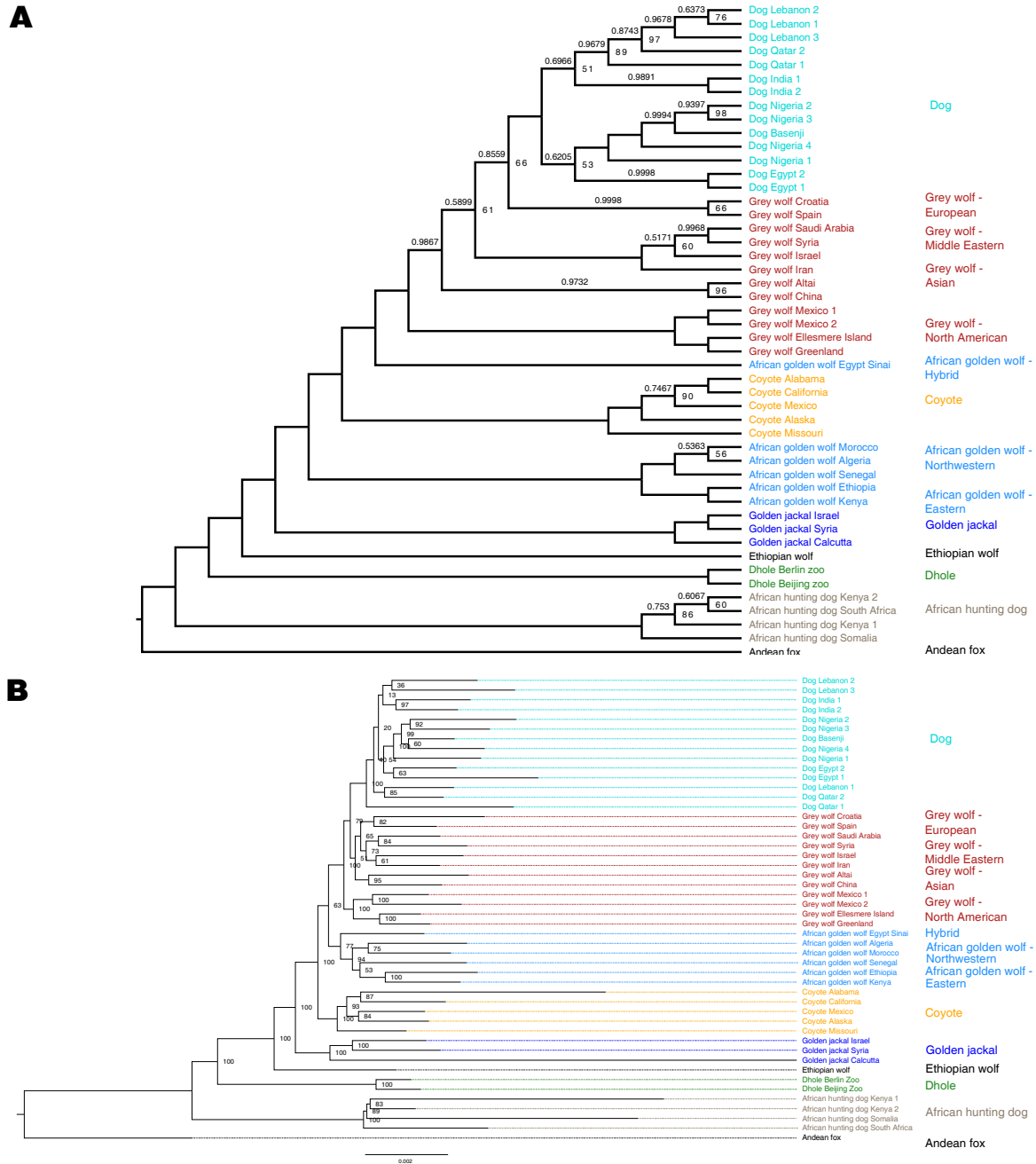
### Interspecific Gene Flow Shaped the Evolution of the Genus *Canis*

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



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

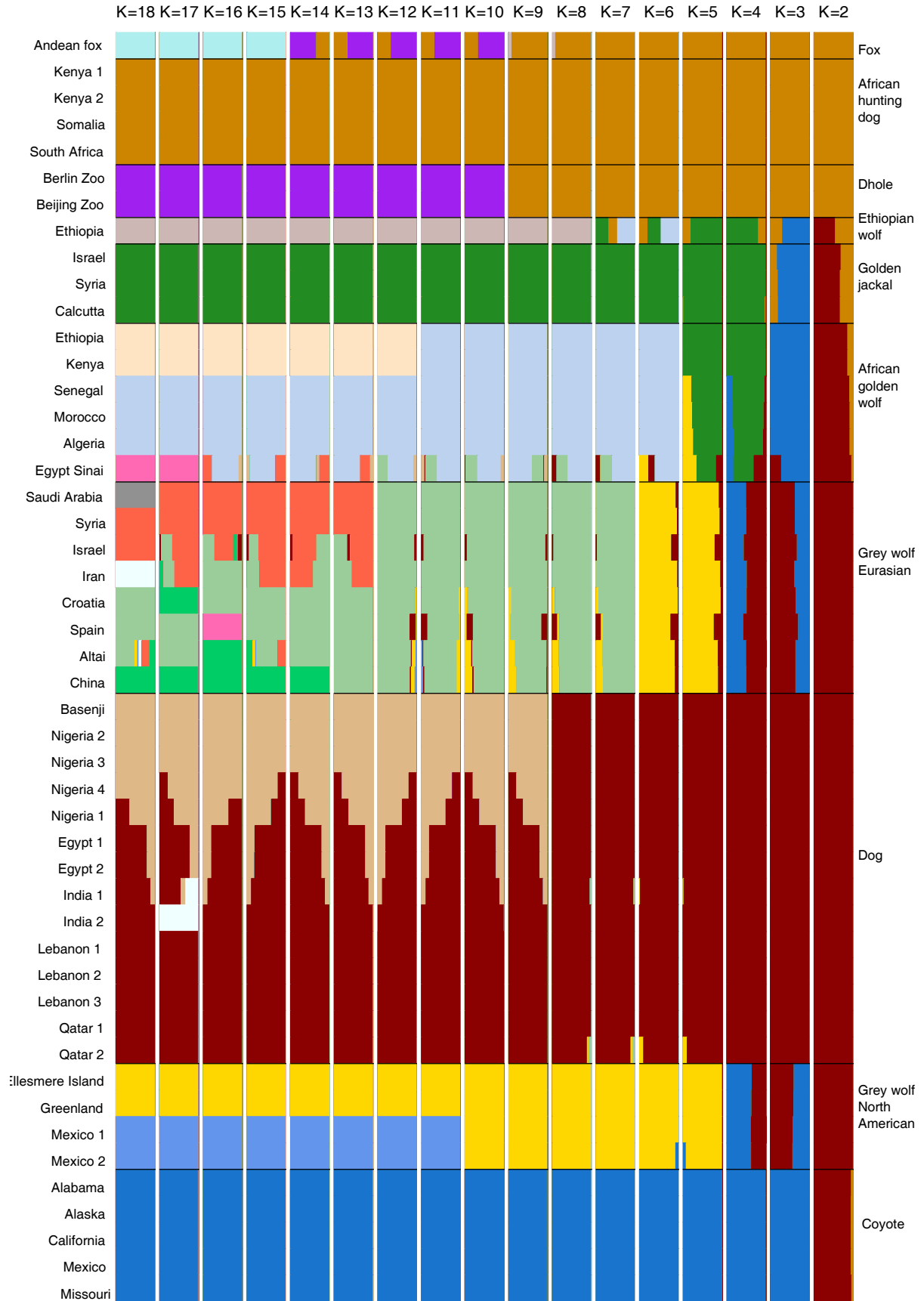
inbreeding. The Ethiopian wolf has the lowest heterozygosity at 0.0003, this extremely low figure being notable that it is comparable to the Mexican wolves, which descend from only 7-9 founders [S4, S5]. Thus, present-day Ethiopian wolf could potentially be subject to inbreeding depression as previously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

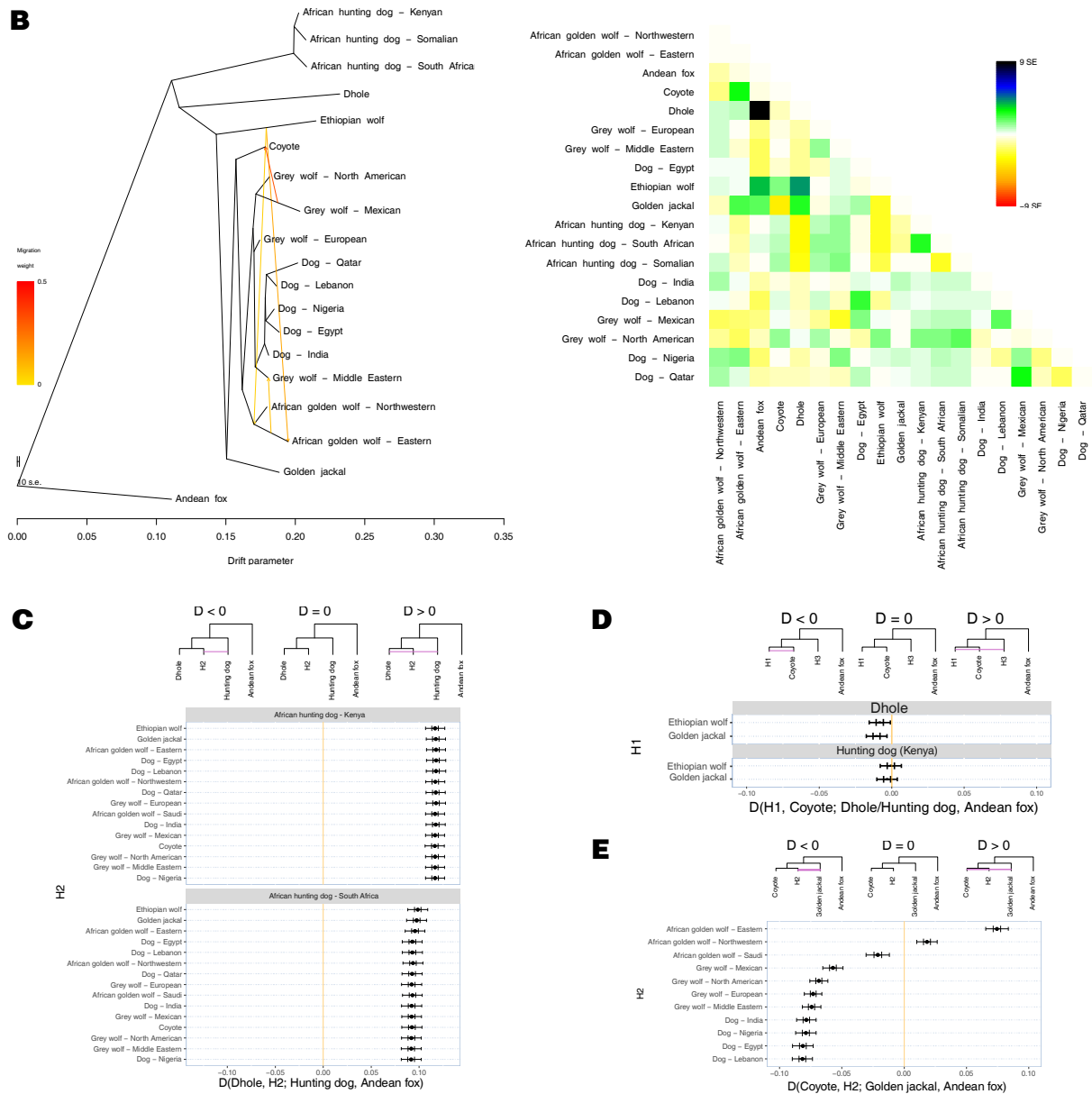


**Figure S2: Phylogeny of all samples in the study. A. Full nuclear phylogeny computed using ASTRAL II. Related to figure 2 in main article.** The phylogeny of all 48 samples, constructed as the consensus tree from 100 separate species trees estimated by Astral-II [S7] using gene trees obtained from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

values less than 100 and mean local posterior probabilities less than 1.0 are shown in the figure. The different groups of samples that were used for collapsing in Figure 1 of the main text, are shown on the right side, with the appropriate labels and colors. The branches and node labels are colored by the species of the sample. The nuclear phylogeny is consistent with previous findings where the Ethiopian wolf was basal to both jackal species, the golden jackal and the African golden wolf. The low average local posterior probabilities and bootstrap supports for the nodes and branches leading to the split of the dogs from the Eurasian wolves shows a lack of power to resolve the topology at these nodes. Further investigation of the resolution at various branches of the phylogeny using bipartition frequencies can be found in figure 2C. Note that the branch lengths in this phylogeny are related to the coalescent times, and do not directly correspond to actual time between clades.

**B. Full nuclear phylogeny using RaxML.** The phylogeny of all 48 samples, constructed using the concatenated analyses of 1000 randomly chosen 1kb regions strewn across the genome. The tree and the bootstrap values are obtained using RAxML [9], under the GTR-GAMMA model of sequence evolution. In this tree, the branch lengths are proportional to the actual split times between the different species.

**A**

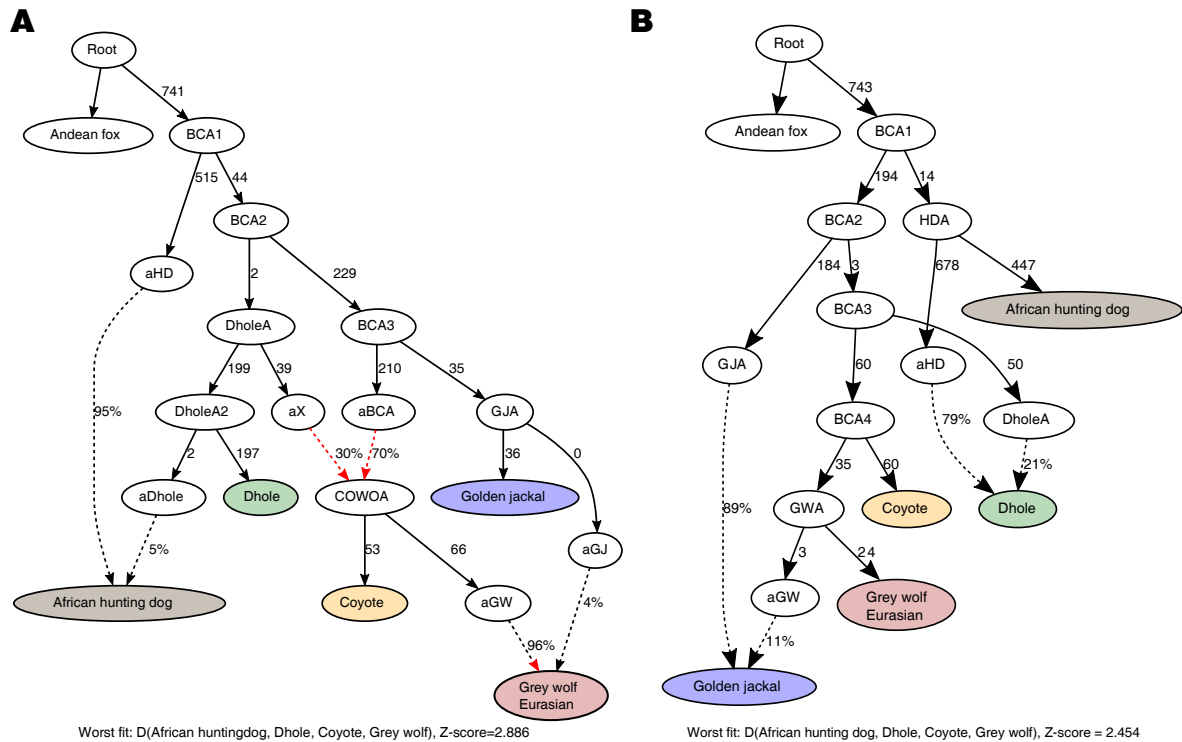


**Figure S3: Evidence for extensive gene flow in the genus *Canis*.** A. Related to figure 3 in main article. Admixture plots for  $K=2-18$ . Admixture analysis of all samples included in the manuscript. The plots show the ancestry proportions of the samples estimated using NGSAdmix, which performs an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

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

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





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

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

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

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

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

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

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

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

### Supplemental references

- [S1] Marino, J., and Sillero-Zubiri, C. (2011). *Canis simensis*. The IUCN Red List of Threatened Species 2011: e.T3748A10051312. <http://dx.doi.org/10.2305/IUCN.UK.2011-1.RLTS.T3748A10051312.en>. Downloaded on 05 April 2017.
- [S2] Woodroffe, R., and Sillero-Zubiri, C. (2012). *Lycaon pictus*. The IUCN Red List of Threatened Species 2012: e.T12436A16711116. <http://dx.doi.org/10.2305/IUCN.UK.2012.RLTS.T12436A16711116.en>. Downloaded on 05 April 2017.
- [S3] Kamler, J.F., Songsasen, N., Jenks, K., Srivathsa, A., Sheng, L., and Kunkel, K. (2015). *Cuon alpinus*. The IUCN Red List of Threatened Species 2015: e.T5953A72477893. <http://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T5953A72477893.en>. Downloaded on 05 April 2017.
- [S4] Garcia-Moreno, J., Matocq, M.D., Roy, M.S., Geffen, E., and Wayne, R.K. (1996). Relationships and Genetic Purity of the Endangered Mexican Wolf Based on Analysis of Microsatellite Loci. *Conserv. Biol.* 10, 376–389.
- [S5] Hedrick, P.W., and Fredrickson, R.J. (2008). Captive breeding and the reintroduction of Mexican and red wolves. *Mol. Ecol.* 17, 344–350.
- [S6] Gottelli, D., Sillero-Zubiri, C., Marino, J., Funk, S.M., and Wang, J. (2013). Genetic structure and patterns of gene flow among populations of the endangered Ethiopian wolf. *Anim. Conserv.* 16, 234–247.

- [S7] Mirarab, S., and Warnow, T. (2015). ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* 31, i44–52.
- [S8] Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 5, e9490.
- [S9] Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313.
- [S10] Auton, A., Rui Li, Y., Kidd, J., Oliveira, K., Nadel, J., Holloway, J.K., Hayward, J.J., Cohen, P.E., Gready, J.M., Wang, J., et al. (2013). Genetic recombination is targeted towards gene promoter regions in dogs. *PLoS Genet.* 9, e1003984.
- [S11] Fan, Z., Silva, P., Gronau, I., Wang, S., Armero, A.S., Schweizer, R.M., Ramirez, O., Pollinger, J., Galaverni, M., Ortega Del-Vecchio, D., et al. (2016). Worldwide patterns of genomic variation and admixture in gray wolves. *Genome Res.* 26, 163–173.
- [S12] Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchio, D., Han, E., Silva, P.M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., et al. (2014). Genome sequencing highlights the dynamic early history of dogs. *PLoS Genet.* 10, e1004016.
- [S13] Liu, Y.-H., Wang, L., Xu, T., Guo, X., Li, Y., Yin, T.-T., Yang, H.-C., Yang, H., Adeola, A.C., J Sanke, O., et al. (2017). Whole-genome sequencing of African dogs provides insights into adaptations against tropical parasites. *Mol. Biol. Evol.* DOI: 10.1093/molbev/msx258.
- [S14] Koepfli, K.-P., Pollinger, J., Godinho, R., Robinson, J., Lea, A., Hendricks, S., Schweizer, R.M., Thalmann, O., Silva, P., Fan, Z., et al. (2015). Genome-wide Evidence Reveals that African and Eurasian Golden Jackals Are Distinct Species. *Curr. Biol.* 25, 2158–2165.
- [S15] Schweizer, R.M., vonHoldt, B.M., Harrigan, R., Knowles, J.C., Musiani, M., Coltman, D., Novembre, J., and Wayne, R.K. (2016). Genetic subdivision and candidate genes under selection in North American grey wolves. *Mol. Ecol.* 25, 380–402.
- [S16] Schlebusch, C.M., Malmström, H., Günther, T., Sjödin, P., Coutinho, A., Edlund, H., Munters, A.R. et al. 2017. “Southern African Ancient Genomes Estimate Modern Human Divergence to 350,000 to 260,000 Years Ago.” *Science* 358 (6363): 652–55.
- [S17] Wang, G.-D., Zhai, W., Yang, H.-C., Fan, R.-X., Cao, X., Zhong, L., Wang, L., Liu, F., Wu, H., Cheng, L.-G., et al. (2013). The genomics of selection in dogs and the parallel evolution between dogs and humans. *Nat. Commun.* 4, 1860.
- [S18] Wang, G.-D., Zhai, W., Yang, H.-C., Wang, L., Zhong, L., Liu, Y.-H., Fan, R.-X., Yin, T.-T., Zhu, C.-L., Poyarkov, A.D., et al. (2016). Out of southern East Asia: the natural history of domestic dogs across the world. *Cell Res.* 26, 21–33.
- [S19] Campana, M.G., Parker, L.D, Hawkins, M.T.R, Young, H.S., Helgen K.M., Gunther, M.S., Woodroffe, R., Maldonado, J.E., Fleischer, R.C. (2016). Genome sequence, population history, and pelage genetics of the endangered African wild dog (*Lycaon pictus*). *BMC Genomics* 17:1013.