

Specimens	Age (Ka)	Region	Site	Genotype	Element	Description	Dimension	Length (Mm)	Losey et al. 2014	BMe (kg)	Reference
Tirekhtyakh (CGG32)	52.5	Siberia	natural	CT	skull	Pleistocene wolf	TL (1)	253.5	$2.70 \cdot \log TL - 4.919$	38.1	Germonpré et al. (2017: table 4)
Eliseevichi (AL2657)	16.5	Western Russia	Epigravettian	TT	mandible	Pleistocene wolf	LPcC (4)	163.6	$2.945 \cdot \log LPcC - 5.104$	39.6	
Ulakhan Sular (CGG33)	16.9	Siberia	natural	CT	skull	Palaeolithic wolf	TL (1)	206.1	$2.70 \cdot \log TL - 4.919$	21.8	Germonpré et al. (2017: table 4)
Zhokhov (CGG6, ZH-03-97)	9.5	Siberia	prehistoric	CT	mandible	ancient dog				24.8	Pitulko and Kasparov (2017: table 9)

Figure S1. Body mass prediction using ancient canid fossils and *IGF1-AS* variant. Related to Figure 2 and STAR Methods. (A) Lateral view of mandible Eliseevichi AL2657. (B) Bivariate plot of Hp2p3 (mm) showing total length of mandible (mm). See table for specifics. Eliseevichi jaw only is shown in plot. Hp2p3: height of the mandible between p2 and p3. E: Eliseevichi AL2657; wolves from rNw: recent northern (n=39), PIW: Pleistocene (n=18). Dogs from rNd: recent northern (n=39), PalD: Palaeolithic (n=18). LPcC: Length from condyle process to border of the canine alveolus; TL: Total length of the skull as measured by ²⁰. (C) Body mass comparison between ancient dogs, modern dogs and coyotes. Diamonds = mean body mass; black horizontal bars = medians.

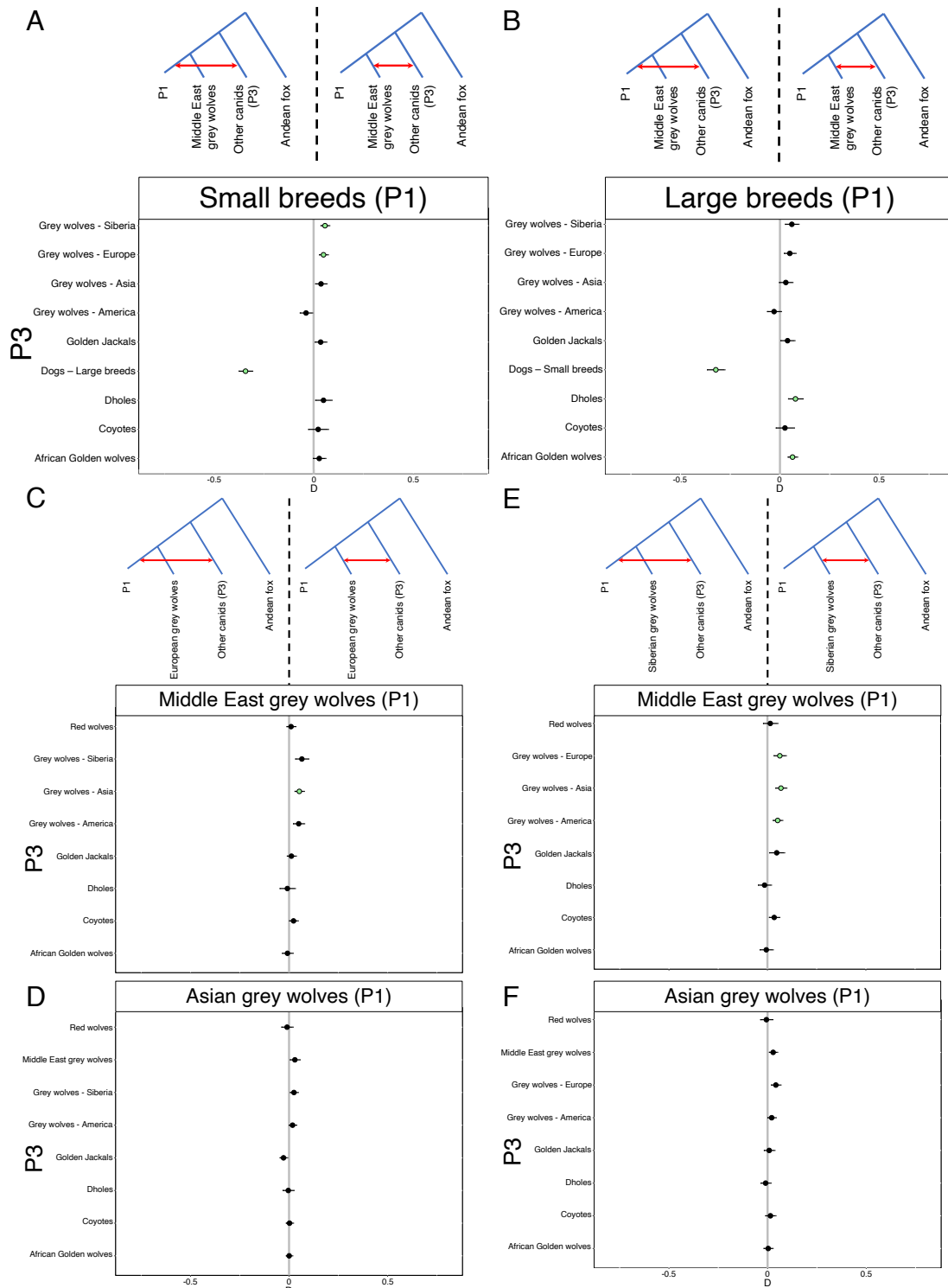


Figure S2. D-Statistics to test the hypothesis of gene flow existing (A-B) between dogs and other wild canids, and (C-F) between small grey wolves and other wild canids. Related to Figures 2, 3 and STAR Methods. (A-B) Small and large breeds are defined in STAR Methods. Results for analyses zoomed in on 2Mb locus centered on *IGF1-AS* variant. P1 population is indicated at the top of each graph. Significant values (P value<0.05) are highlighted in green. Error bars represent the standard error estimated via jackknife resampling with a block size of one MB.

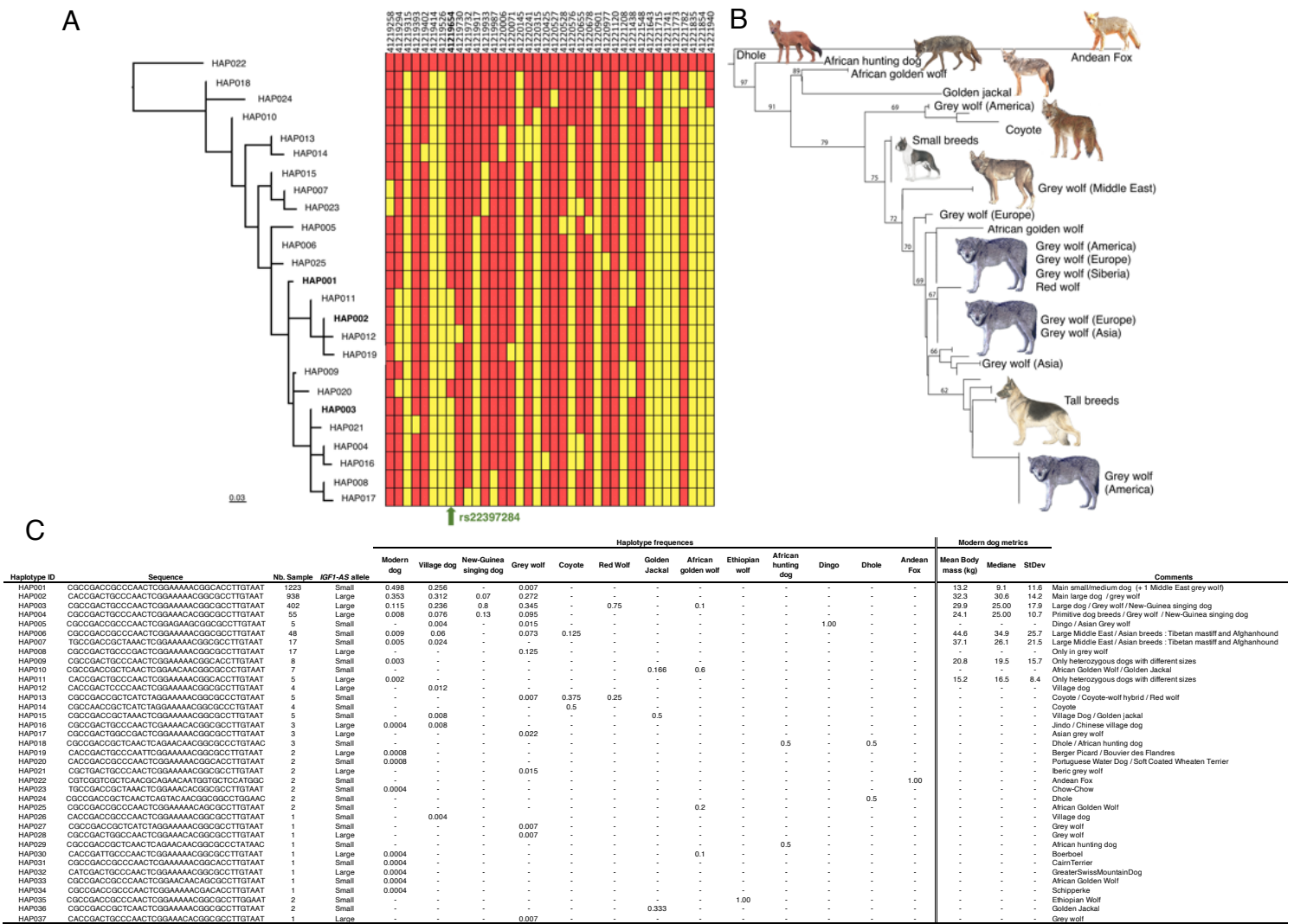


Figure S3. Haplotype analysis and IGF1-AS neighbor-joining tree. Related to Figures 2, 3 and STAR Methods (A) Haplotypes spanning IGF1-AS using 38 markers (see supplementary data) and WGS. IGF1-AS variant represented by green arrow. Ancestral alleles (red) defined from the Andean fox; yellow = derived alleles. (B) IGF1-AS neighbor-joining tree based on 2,682 bp of phased sequences rooted using the Andean Fox. Grey wolves are named using their geographic location; breed size assignments are defined in STAR Methods. (C) Haplotype frequencies among the 1,389 canid genomes analysed in this study.

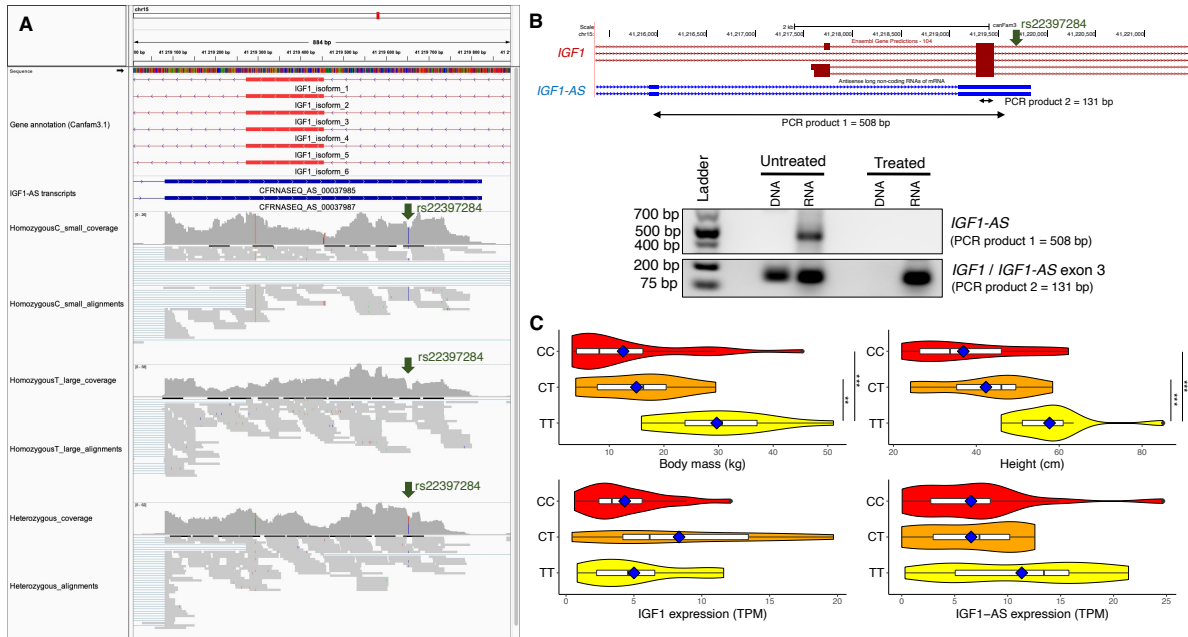


Figure S4. Functional analyses of *IGF1-AS*. Related to STAR Methods. A) Screenshot of RNA-SEQ data visualized IGV⁷⁰ with a zoom on the last exon of *IGF1-AS*. Red bar = *IGF1* third exon, blue bars = the last exon of *IGF1-AS* (two isoforms), green arrow = the SNP (rs22397284; chr15:41219654.g.T>C) and grey bars = read alignments. The first two RNA-Seq tracks show the homozygous C (small dog-ID:25363) in blue, lines 3-4 RNA-Seq indicate data for a homozygous T (large dog, ID:25460), and the last two tracks highlight data for a heterozygous dog. (B) Results of the RPA using DNase/RNase treatment. Top panel: Control indicating efficiency of the nuclease treatment on *IGF1-AS* detection. Bottom panel: Presence of *IGF1-AS*/*IGF1* mRNA duplex (exon 3) validated by RT-PCR in cDNA produced from nuclease treated RNA (C) RNA-seq analyses of testes from 26 small and 14 large dog breeds (STAR Methods). Upper boxplots: body mass and height distributions for the three genotypes. Diamonds = mean body mass; black vertical bars = medians. (**P < 0.001, ***P < 0.0001, Mann-Whitney-Wilcoxon tests). Bottom boxplots: *IGF1* and *IGF1-AS* expression levels.

GWAS Rank	Biallelic SNV	dbSNP number	Large allele (Dog reference)	Small allele	GWAS data					All WGS data							Sanger sequenced wild canids					Sanger sequenced specific breeds										
					All breeds population n=456		p _{Wald}	p _{Wald after correction} n=446	Small breeds (SBW <15 kg) n=162	Medium breeds (15<SBW<25) n=106	Large breeds (SBW >25 kg) n=188	Small breed (SBW <15 kg) n=438	Medium breed (15<SBW<25) n=191	Large breed (SBW >25 kg) n=426	Village dogs n=125	Grey wolves n=28	Coyote n=4	Red wolves n=2	Other canids n=18	Only wild canids body size GWAS p-values n=80	Grey wolves America/Europe n=29	Grey wolves MiddleEast/Asia n=17	Coyotes West-MidWest n=31	Coyotes East n=4	Other canids n=12	Miniature Schnauzer n=48	Standard Schnauzer n=42	Giant Schnauzer n=46	Toy Poodle n=48	Miniature Poodle n=48	Standard Poodle n=48	
					AI	Beta			Variance	Small allele frequency			Small allele frequency			Small allele frequency				Small allele frequency					Small allele frequency							
					1e-31	1.36e-30	0.82	0.42	0.19	0.8	0.43	0.21	0.19	0	0	0	0	1	1.68e-19	0.18	0.5	0.94	0.5	1	0.96	0.9	0	1	0.97	0.5		
1	chr15:41216368	NA	C	T	0.51	-7.25	0.11	1.36e-30	0.82	0.42	0.19	0.8	0.43	0.21	0.19	0	0	0	0	1	-	-	-	-	-	-	-	-	-	-		
2	chr15:41221438	rs22437444	G	A	0.52	-7.17	0.11	2.48e-31	0.83	0.41	0.19	0.8	0.42	0.21	0.18	0.01	0	0	0	0	0.32	0	0	0	0	0	-	-	-	-	-	
3	chr15:41216597	rs22362978	G	A	0.51	7.06	0.11	3.36e-30	0.82	0.44	0.19	0.79	0.45	0.22	0.23	0.09	0	0.25	0.06	0.82	-	-	-	-	-	-	-	-	-	-		
4	chr15:41219654	rs22397284	T	C	0.5	6.84	0.1	3.97e-29	0.83	0.42	0.21	0.8	0.44	0.23	0.28	0.12	1	0.25	0.97	1.68e-19	0.37	-	-	-	-	-	-	-	-	-		
5	chr15:41227725	NA	G	A	0.55	-6.92	0.1	4.56e-29	0.78	0.39	0.16	0.76	0.4	0.19	0.2	0.07	0	0	0.11	0.37	-	-	-	-	-	-	-	-	-	-		
6	chr15:41228068	NA	G	A	0.55	-6.92	0.1	4.56e-29	0.78	0.39	0.16	0.76	0.4	0.19	0.2	0.05	0	0	0.08	0.03	-	-	-	-	-	-	-	-	-	-		
7	chr15:41217964	rs22381307	C	A	0.5	6.82	0.1	3.99e-28	0.83	0.45	0.19	0.8	0.46	0.23	0.25	0	0	-	0.19	0.01	-	-	-	-	-	-	-	-	-	-		
8	chr15:41229597	rs22426115	A	G	0.54	-6.78	0.1	5.28e-28	0.78	0.43	0.15	0.76	0.43	0.18	0.27	0.01	0	0	0	0.32	-	-	-	-	-	-	-	-	-	-		
9	chr15:41217985	rs22381305	T	C	0.49	6.74	0.1	6.57e-28	0.83	0.46	0.21	0.8	0.45	0.24	0.28	0.14	1	-	0.97	2.82e-14	-	-	-	-	-	-	-	-	-	-		
10	chr15:41229800	rs22426118	T	C	0.55	-6.67	0.1	1.74e-27	0.78	0.42	0.15	0.76	0.42	0.18	0.27	0.01	0	0	0	0.32	-	-	-	-	-	-	-	-	-	-		
-	chr15:41220980	-	SINE	-	-	-	NA	NA	-	-	-	-	-	-	-	0	0	0	0	0	0	0	0	0	0	0	0.96	0.85	0	1	0.9	0.4

Table S1. Top 10 most associated variants with body size variations in dogs and wild canids. Related to Figures 1, 2 and 4.

The genetic variance attributable to each variant was estimated as $V=2^*A^*(1-A)^*(beta^2)/(variance(breed\ body\ mass\ average)-231.6)$. Beta is the estimated SNP effect (regression coefficient) given by GEMMA (Zhou et al. Nat. Genet 2012). For the SINE (207bp), we used IGV software to detect the presence of the SINE in WGS, and for Sanger sequenced samples, we only ran PCR product on an 2% agarose gel to detect presence/absence. * We removed three non-European breeds (four Chowchow, two Afghanhound, two Tibetan mastiff) that do not appear under selection for body size on this locus (these breeds carry their own haplotypes); AI: Small allele frequency ; - : Not checked ; NA: Not applicable ; Dog reference : Canfam3.1.

Individual ID	Species	Age	Location	Reference	Coverage	Lat	Long	IGF1-AS variant	BME (kg)	Name on the map (Fig. 2A)	Additional information for the three new samples								
											Dating Lab	NRCF Grant Code	OXA_ID	Extract_No	Uncal RC Age	Error +/-	From Cal BP (Int Cal13)	To Cal BP	Percentage
C32	Dog	100	Bulgunnyakhtakh lake, Russia	Bergstrom et al, <i>Science</i> 2020	7.20	61.2	128.4	CT	-	Bulgunnyakhtakh L. 0.1k									
F3781	Dog	100	East Siberian Sea coast, Russia	Bergstrom et al, <i>Science</i> 2020	0.80	69.6	164.3	TT	-	East Siberian Sea 0.1k									
UZA001	Dog	875	Uza, Israel	Bergstrom et al, <i>Science</i> 2020	5.80	31.6	34.8	CC	-	Uza 0.8-1.6k (x2)									
AL3223	Dog	1000	Weyanoke Old Town, Virginia, USA	Leathlobhair et al, <i>Science</i> 2018	0.40	NA	NA	-	-	-									
AL2022	Dog	1565	Marmara region, Turkey	Bergstrom et al, <i>Science</i> 2020	2.50	41.0	28.9	CT	-	Turkey 1.5k									
UZA002	Dog	1600	Uza, Israel	Bergstrom et al, <i>Science</i> 2020	5.50	31.6	34.8	TT	-	Uza 0.8-1.6k (x2)									
ASHQ01	Dog	2300	Ashkelon, Israel	Bergstrom et al, <i>Science</i> 2020	2.40	31.7	34.6	CC	~14.6	Ashkelon 2.3k (x2)									
ASHQ06	Dog	2300	Ashkelon, Israel	Bergstrom et al, <i>Science</i> 2020	1.80	31.7	34.6	CC	~14.6	Ashkelon 2.3k (x3)									
ASHQ08	Dog	2300	Ashkelon, Israel	Bergstrom et al, <i>Science</i> 2020	2.20	31.7	34.6	CC	~14.6	Ashkelon 2.3k (x3)									
TGEZ06	Dog	2300	Tel Gezer, Israel	Bergstrom et al, <i>Science</i> 2020	0.90	31.9	34.9	CC	-	Tel Gezer 2.3k									
C62	Dog	3100	Apalle, Sweden	Bergstrom et al, <i>Science</i> 2020	0.70	59.7	17.6	CC	-	Sweden 3.1k									
C5	Dog	3800	Krasnosamarskoe, Russia	Bergstrom et al, <i>Science</i> 2020	0.60	52.8	51.1	-	-	-									
AL2397	Dog	4000	Belverde di Cetona, Italy	Bergstrom et al, <i>Science</i> 2020	1.50	43.0	11.9	TT	-	Italy 4k									
AL3194	Dog	4000	Port au Choix, Canada	Leathlobhair et al, <i>Science</i> 2018	1.92	50.703	-57.352	TT	-	Port au Choix 4k									
G94	Dog	4000	Stora Fövar, Gotland, Sweden	Bergstrom et al, <i>Science</i> 2020	0.20	57.3	18.0	-	-	-									
ALP001	Dog	4500	Alymas-Podunavljje, Croatia	Bergstrom et al, <i>Science</i> 2020	1.60	45.5	19.0	CC	-	Croatia 4.5-4.9k (x2)									
C89	Dog	4800	Ajvide, Gotland, Sweden	Bergstrom et al, <i>Science</i> 2020	2.20	57.2	18.1	TT	-	Sweden 4.8k (x2)									
C90	Dog	4800	Ajvide, Gotland, Sweden	Bergstrom et al, <i>Science</i> 2020	0.60	57.2	18.1	TT	-	Sweden 4.8k (x2)									
CTC	Dog	4800	Cherry Tree Cave, Germany	Botigue et al, <i>Nat. Commun.</i> 2017	9.00	49.8	10.8	CC	-	Germany 4.8k									
NGDG	Dog	4800	Newgrange, Ireland	Frantz et al, <i>Science</i> 2016	30	53.694	-6.475	TT	-	Ireland 4.8k									
SOTN01	Dog	4900	Sotin, Croatia	Bergstrom et al, <i>Science</i> 2020	11.20	45.3	19.1	CC	-	Croatia 4.5-4.9k (x2)									
C88	Dog	5000	Frälsegården, Gökhem, Sweden	Bergstrom et al, <i>Science</i> 2020	0.70	58.1	13.2	TT	-	Sweden 5k									
AL2350	Wolf	5169	Botai, Kazakhstan	Unpublished	1.43	53.2	67.9	TT	-	Botai 5.1k	Oxford RLAHA	NF/2016/2/4	OxA-36898	AL2350	4492	29	5296	5041	95.4
AL2571	Dog	5826	Tepe Ghela Gap, Iran	Bergstrom et al, <i>Science</i> 2020	1.10	36.5	47.1	CC	-	Iran 5.8k									
OL4029	Dog	6230	Marizulo Cave, Gipuzkoa, Spain	Bergstrom et al, <i>Science</i> 2020	0.10	43.3	-2.1	-	-	-									
AL3185	Wolf	6307	Pietrele, Romania	Unpublished	4.21	43.9	24.9	CT	-	Pietrele 6.3k	Oxford RLAHA	NF/2016/2/4	OxA-36897	AL3185	5657	33	6505	6390	88
OL4222	Dog	6544	Skoteini cave, Tharrounia, Euboea Island, Greece	Bergstrom et al, <i>Science</i> 2020	4.50	38.5	24.0	CC	-	Greece 6.5k									
AL2946	Dog	6839	Pločnik, Serbia	Bergstrom et al, <i>Science</i> 2020	0.20	43.2	21.4	CC	-	Serbia 6.7k									
OL4223	Dog	6900	Pad' Kalashnikova, Russia	Bergstrom et al, <i>Science</i> 2020	2.20	52.7	103.7	-	-	-									
C26	Dog	7000	Pad' Kalashnikova, Russia	Bergstrom et al, <i>Science</i> 2020	0.20	53.1	103.4	-	-	-									
HXH	Dog	7000	Herxheim, Germany	Botigue et al, <i>Nat. Commun.</i> 2017	9.00	49.1	8.1	CT	-	Germany 7k									
THRZ02	Dog	7000	Tel Hreiz, Israel	Bergstrom et al, <i>Science</i> 2020	0.10	32.7	35.1	CC	-	Israel 7k									
C27	Dog	7400	Shamanka II, Russia	Bergstrom et al, <i>Science</i> 2020	0.30	51.7	103.7	-	-	-									
CGG6	Dog	9500	Zhokhov Island, Russia	Sinding et al, <i>Science</i> 2020	9.6	76.141	152.733	CT	24.8	Zhokhov Island 9.5k									
OL4061	Dog	10930	Vereye, Lake Lacha, Russia	Bergstrom et al, <i>Science</i> 2020	1.80	61.3	-8.9	TT	-	Karelia 10.9k									
Tumat2	Wolf	14122	Tumat, Russia	Ramos-Madrigal et al, <i>Curr. Biol.</i> 2020	5.162	70.72	139.23	TT	-	Tumat 14.1k									
CGG33	Wolf	16900	Ulakhan Sular, Russia	Ramos-Madrigal et al, <i>Curr. Biol.</i> 2020	15.279	56.71	131.85	CT	21.8	Ulakhan Sular 16.8k									
AL2657	Wolf	16500	Eliseevichi, W Russia	Unpublished	0.93	53.1	33.6	TT	39.6	Eliseevichi 17.6k	Oxford RLAHA	NF/2016/2/4	OxA-35646	AL2657	14445	65	17860	17405	95.4
CGG23	Wolf	33020	Yana site	Sinding et al, <i>Science</i> 2020	4.45	70.7	135.4	TT	-	Yana site 33k									
Taimyr-1	Wolf	34902	Taimyr Peninsula, C Siberia	Skoglund et al, <i>Curr. Biol.</i> 2015	2.42	73.3	104.3	TT	-	Taimyr Peninsula 35k									
CGG29	Wolf	48210	Bunge-Toll-1885 site, Yana river	Ramos-Madrigal et al, <i>Curr. Biol.</i> 2020	6.29	68.9	134.5	TT	-	Bunge-Toll-1885 site 48.2k									
CGG32	Wolf	52500	Tirekhtyakh	Ramos-Madrigal et al, <i>Curr. Biol.</i> 2020	15.10	68.9	147.2	CT	38.1	Tirekhtyakh 52.5k									

Table S3. Ancient genomes dataset and their IGF1-AS genotype. Related to Figure 2.