

1 **Modern wolves trace their origin to a late Pleistocene expansion from Beringia**

2 Liisa Loog<sup>1,2,3\*</sup>, Olaf Thalmann<sup>4†</sup>, Mikkel-Holger S. Sinding<sup>5,6,7†</sup>, Verena J. Schuenemann<sup>8,9†</sup>,  
3 Angela Perri<sup>10</sup>, Mietje Germonpré<sup>11</sup>, Herve Bocherens<sup>9,12</sup>, Kelsey E. Witt<sup>13</sup>, Jose A.  
4 Samaniego Castruita<sup>5</sup>, Marcela S. Velasco<sup>5</sup>, Inge K. C. Lundstrøm<sup>5</sup>, Nathan Wales<sup>5</sup>, Gontran  
5 Sonet<sup>15</sup>, Laurent Frantz<sup>2</sup>, Hannes Schroeder<sup>5,15</sup>, Jane Budd<sup>16</sup>, Elodie-Laure Jimenez<sup>11</sup>, Sergey  
6 Fedorov<sup>17</sup>, Boris Gasparyan<sup>18</sup>, Andrew W. Kandel<sup>19</sup>, Martina Lázničková-Galetová<sup>20,21,22</sup>,  
7 Hannes Napierala<sup>23</sup>, Hans-Peter Uerpmann<sup>8</sup>, Pavel A. Nikolskiy<sup>24,25</sup>, Elena Y. Pavlova<sup>26,25</sup>,  
8 Vladimir V. Pitulko<sup>25</sup>, Karl-Heinz Herzig<sup>4,27</sup>, Ripan S. Malhi<sup>26</sup>, Eske Willerslev<sup>2,5,29</sup>, Anders J.  
9 Hansen<sup>5,7</sup>, Keith Dobney<sup>30,31,32</sup>, M. Thomas P. Gilbert<sup>5,33</sup>, Johannes Krause<sup>8,34</sup>, Greger  
10 Larson<sup>1\*</sup>, Anders Eriksson<sup>35,2\*</sup>, Andrea Manica<sup>2\*</sup>

11

12 \*Corresponding Authors: L.L. (liisaloog@gmail.com), G.L. (greger.larson@arch.ox.ac.uk),  
13 A.E. (anders.eriksson@kcl.ac.uk), A.M. (am315@cam.ac.uk)

14

15 †These authors contributed equally to this work

16

17 1 Palaeogenomics & Bio-Archaeology Research Network Research Laboratory for  
18 Archaeology and History of Art, University of Oxford, Dyson Perrins Building, South Parks  
19 Road, Oxford OX1 3QY, UK

20 2 Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ,  
21 UK

22 3 Manchester Institute of Biotechnology, School of Earth and Environmental Sciences,  
23 University of Manchester, Manchester, M1 7DN, UK

24 4 Department of Pediatric Gastroenterology and Metabolic Diseases, Poznan University of  
25 Medical Sciences, Szpitalna 27/33, 60-572 Poznan, Poland

26 5 Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen,  
27 Øster Voldgade 5-7, DK-1350 Copenhagen, Denmark

28 6 Natural History Museum, University of Oslo, P.O. Box 1172 Blindern, NO-0318 Oslo,  
29 Norway

30 7 The Qimmeq project, University of Greenland, Manutooq 1, PO Box 1061, 3905 Nuussuaq,  
31 Greenland

32 8 Institute for Archaeological Sciences, University of Tübingen, Rümelinstr. 23, 72070  
33 Tübingen, Germany

34 9 Senckenberg Centre for Human Evolution and Palaeoenvironment, University of Tübingen,  
35 72070 Tübingen, Germany

- 1 10 Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology,  
2 Deutscher Platz 6, 04103 Leipzig, Germany
- 3 11 OD Earth and History of Life, Royal Belgian Institute of Natural Sciences, Vautierstraat  
4 29, 1000 Brussels, Belgium
- 5 12 Department of Geosciences, Palaeobiology, University of Tübingen, Tübingen, Germany
- 6 13 School of Integrative Biology, University of Illinois at Urbana-Champaign, 109A  
7 Davenport Hall, 607 S. Mathews Avenue, Urbana IL 61801, USA
- 8 14 OD Taxonomy and Phylogeny, Royal Belgian Institute of Natural Sciences, Vautierstraat  
9 29, 1000 Brussels, Belgium
- 10 15 Faculty of Archaeology, Leiden University, Postbus 9514, 2300 RA Leiden, The  
11 Netherlands
- 12 16 Breeding Centre for Endangered Arabian Wildlife, PO Box 29922 Sharjah, United Arab  
13 Emirates
- 14 17 Mammoth Museum, Institute of Applied Ecology of the North of the North-Eastern  
15 Federal University, ul. Kulakovskogo 48, 677980 Yakutsk, Russia
- 16 18 National Academy of Sciences, Institute of Archaeology and Ethnography, Charents St.  
17 15, Yerevan 0025, Armenia
- 18 19 Heidelberg Academy of Sciences and Humanities: The Role of Culture in Early  
19 Expansions of Humans, Rümelinstr. 23, 72070 Tübingen, Germany
- 20 20 Departement of Anthropology, University of West Bohemia, Sedláčková 15, 306 14  
21 Pilzen, Czech republic
- 22 21 Moravian museum, Zelný trh 6, 659 37 Brno, Czech republic
- 23 22 Hrdlička Museum of Man, Faculty of Science, Charles University, Viničná 1594/7,128 00  
24 Praha, Czech republic
- 25 23 Institute of Palaeoanatomy, Domestication Research and History of Veterinary Medicine,  
26 Ludwig-Maximilians-University Munich, Kaulbachstraße 37 III/313, D-80539 Munich,  
27 Germany
- 28 24 Geological Institute, Russian Academy of Sciences, 7 Pyzhevsky per., 119017 Moscow,  
29 Russia
- 30 25 Institute for Material Culture History, Russian Academy of Sciences, 18 Dvortsovaya nab.,  
31 St Petersburg 191186, Russia
- 32 26 Arctic and Antarctic Research Institute, 38 Bering St., St Petersburg 199397, Russia
- 33 27 Insitute of Biomedicine and Biocenter of Oulu, Medical Research Center and University  
34 Hospital, University of Oulu, Aapistie 5, 90220 Oulu University, Finland
- 35 28 Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign,  
36 1206 W Gregory Dr., Urbana, Illinois 61820, USA

- 1 29 Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK
- 2 30 Department of Archaeology, Classics and Egyptology, University of Liverpool, 12-14
- 3 Abercromby Square, Liverpool L69 7WZ, UK
- 4 31 Department of Archaeology, University of Aberdeen, St Mary's, Elphinstone Road,
- 5 Aberdeen AB24 3UF, UK
- 6 32 Department of Archaeology, Simon Fraser University, Burnaby, B.C. V5A 1S6, 778-782-
- 7 419, Canada
- 8 33 Norwegian University of Science and Technology, University Museum, N-7491
- 9 Trondheim, Norway
- 10 34 Max Planck Institute for the Science of Human History, Khalaische Straße 10, 07745 Jena,
- 11 Germany
- 12 35 Department of Medical & Molecular Genetics, King's College London, Guys Hospital,
- 13 London SE1 9RT, UK

1 **ABSTRACT**

2

3 **Grey wolves (*Canis lupus*) are one of the few large terrestrial carnivores that maintained**  
4 **a wide geographic distribution across the Northern Hemisphere throughout the**  
5 **Pleistocene and Holocene. Recent genetic studies have suggested that, despite this**  
6 **continuous presence, major demographic changes occurred in wolf populations between**  
7 **the late Pleistocene and early Holocene, and that extant wolves trace their ancestry to a**  
8 **single late Pleistocene population. Both the geographic origin of this ancestral**  
9 **population and how it became widespread remain a mystery. Here we analyzed a large**  
10 **dataset of novel modern and ancient mitochondrial wolf genomes, spanning the last**  
11 **50,000 years, using a spatially and temporally explicit modeling framework to show that**  
12 **contemporary wolf populations across the globe trace their ancestry to an expansion**  
13 **from Beringia at the end of the Last Glacial Maximum - a process most likely driven by**  
14 **the significant ecological changes that occurred across the Northern Hemisphere during**  
15 **this period. This study provides direct ancient genetic evidence that long-range**  
16 **migration has played an important role in the population history of a large carnivore**  
17 **and provides an insight into how wolves survived the wave of megafaunal extinctions at**  
18 **the end of the last glaciation. Moreover, because late Pleistocene grey wolves were the**  
19 **likely source from which all modern dogs trace their origins, the demographic history**  
20 **described in this study has fundamental implications for understanding the**  
21 **geographical origin of the dog.**

22

1 The Pleistocene epoch harbored a large diversity of top predators though most became extinct  
2 during or soon after the Last Glacial Maximum (LGM) ~24 thousand years ago. The grey  
3 wolf (*Canis lupus*) was one of the few large carnivores that survived and maintained a wide  
4 geographical range throughout the period (1), and both the paleontological and archaeological  
5 records attest to the continuous presence of grey wolves across the Northern Hemisphere for  
6 at least the last 300,000 years (2) (reviewed in Supplementary Information 1). This  
7 geographical and temporal continuity across the Northern Hemisphere contrasts with analyses  
8 of complete modern genomes which have suggested that all contemporary wolves and dogs  
9 descend from a common ancestral population that existed as recently as ~20,000 years ago  
10 (3–5). These analyses point to a bottleneck followed by a rapid radiation from an ancestral  
11 population around or just after the LGM, but the geographic origin and dynamics of this  
12 radiation remain unknown. Resolving these demographic changes is necessary for  
13 understanding the ecological circumstances that allowed wolves to survive the late  
14 Pleistocene megafaunal extinctions. Furthermore, because dogs were domesticated from late  
15 Pleistocene grey wolves (6), a detailed insight into wolf demography during this time period  
16 would provide an essential context for reconstructing the history of dog domestication.

17 Reconstructing past demographic events solely from modern genomes is challenging since  
18 multiple demographic histories can lead to similar genetic patterns in present-day samples (7).  
19 Analyses that incorporate ancient DNA sequences can eliminate some of these alternative  
20 histories by quantifying changes in population genetic differences through time. While  
21 nuclear markers provide greater power relative to mitochondrial DNA (mtDNA), the latter is  
22 more easily retrievable and better preserved in ancient samples due to its higher copy number  
23 compared to the nuclear DNA, thus allowing for the generation of datasets with greater  
24 geographical and temporal coverage. Furthermore, the nuclear mutation rate in canids is  
25 poorly understood, leading to wide date ranges for past demographic events reconstructed  
26 from panels of modern whole genomes (e.g. (3, 5)). Having samples from a broad time period  
27 can reduce mutation rate uncertainty by calibrating the evolutionary rate with directly dated  
28 samples (8–10).

29 Although mtDNA can be retrieved from a wider range of ancient samples, the sparseness of  
30 samples across space and time, compounded by the stochasticity of a single non-recombining  
31 genetic marker, can lead to patterns that are difficult to interpret intuitively (7). Population  
32 genetic models that explicitly capture the expected temporal (e.g. (11)) and spatial (e.g. (12,  
33 13)) differences between samples can be used to overcome this problem.

34 In order to reconstruct the demographic history of wolves, we assembled a substantial dataset  
35 (Fig. 1, Table S1) consisting of 90 modern and 45 ancient wolf whole mitochondrial genomes

1 (55 of which are newly sequenced), spanning the last 50,000 years and the geographic breadth  
2 of the Northern Hemisphere. We first used the ancient mitogenomes to estimate a wolf  
3 mutation rate using BEAST. We then designed a spatially and temporally explicit population  
4 genetic (coalescent) model that accounts for the stochasticity of the mitochondrial  
5 phylogenetic tree, as well as the uneven spatial and temporal distribution of the samples. We  
6 used our spatial model to investigate the origin and population dynamics of the expansion of  
7 grey wolves during the LGM.

## 8 **Bayesian Phylogenetic Analysis**

9

10 All sequences included in the study were subjected to stringent quality criteria with respect to  
11 coverage and damage patterns. We used the 38 ancient samples for which we had direct  
12 radiocarbon dates to estimate mitochondrial mutation rates using BEAST (14), and  
13 molecularly dated the remaining seven ancient samples (Supplementary Information 3.1).

14

15 Our Bayesian phylogenetic analysis suggests that the most recent common ancestor of all  
16 North Eurasian and American wolf samples dates to ca. 90,000 (95% CIs: 82,000 – 99,000)  
17 years ago (Fig. 2, see Figs. S11 and S12 for node support values and credibility intervals). At  
18 the root of this tree, we find a divergent clade consisting exclusively of ancient samples from  
19 Europe and the Middle East that has not contributed to present day mitochondrial diversity in  
20 our data (see also (15)). The rest of the tree consists of a monophyletic clade made up of  
21 ancient and modern samples from across the Northern Hemisphere and shows a pattern of  
22 rapid bifurcations of genetic lineages centered on 25,000 years ago. A Bayesian skyline  
23 analysis (Fig. S13, see Supplementary Information 3 for details) also shows a recent reduction  
24 in effective population size. This pattern is compatible with a scenario of rapid radiation that  
25 has also been suggested by whole genome studies (e.g. (3, 5)). At the root of this clade we find  
26 predominantly samples from Beringia, pointing to a possible expansion out Northeast Eurasia  
27 or the Americas. However, given the uneven temporal and geographic distribution of our  
28 samples and the stochasticity of a single genetic marker (16), it is important to explicitly test  
29 the extent to which this pattern can occur by chance under other plausible demographic  
30 scenarios, taking the geographic and temporal distribution of our samples into account.

31

## 32 **Spatial Modelling of Past Wolf Demography**

33 Motivated by the population structure observed in whole genome studies of modern wolves  
34 (5), we tested the degree of spatial genetic structure among the modern wolf samples in our  
35 dataset, and found a strong pattern of genetic isolation by distance across Eurasia ( $\rho=0.3$ ,

1  $p < 0.0001$ ; see Fig. S8). To take this structure into account in our spatial framework, we  
2 represented the wolf distribution in the Northern Hemisphere as seven demes (Fig. 1), each of  
3 which is defined by major geographic barriers including mountain ranges, seas, oceans and  
4 deserts (see Materials and Methods). We then used coalescent simulations (see Materials and  
5 Methods) to test a range of different explicit demographic scenarios (illustrated in Fig. 3a),  
6 with sampling matching the empirical spatial and temporal distribution of our samples.

7 The first scenario consisted of a constant population size and uniform movement between  
8 neighboring demes. This allowed us to test the null hypothesis that drift within a structured  
9 population alone can explain all the patterns observed in the mitochondrial tree. We then  
10 considered two additional demographic processes that could explain the observed patterns: 1)  
11 a temporal sequence of two population size changes that affected all demes simultaneously  
12 (thus allowing for a bottleneck); and 2) an expansion out of one of the demes, which had a  
13 continuous population through time and sequentially replaced the populations in the other  
14 demes (repeated for all seven possible expansion origins). We considered each demographic  
15 event in isolation as well as their combined effect (resulting in a total of 16 scenarios) and  
16 used Approximate Bayesian Computation (ABC) to calculate the likelihood of each scenario  
17 and estimate parameter values.

18 Both the null scenario and the scenario of only population size change in all demes were  
19 strongly rejected (Bayes Factor (BF)  $\leq 0.1$ , Fig. 3b and Table S6), illustrating the power of  
20 combining a large dataset of ancient samples with statistical modeling. Scenarios that  
21 combined an expansion with a change in population size (bottleneck) were better supported  
22 than the corresponding scenarios (i.e. with the same origin) with constant population size  
23 (Fig. 3b). The best-supported scenario (BF 1, Fig. 4) was characterized by the combination of  
24 a rapid expansion of wolves out of the Beringian deme ~25,000 years ago (95% CI: 33,000-  
25 14,000 years ago) with a population bottleneck between 15,000 and 40,000 years ago, and  
26 limited gene flow between neighboring demes (see Table S7 and Fig. S15 for posterior  
27 distributions of all model parameters). We also found relatively strong support for a scenario  
28 that describes a wolf expansion out of the East Eurasian deme (BF 0.7) with nearly identical  
29 parameters to the best-supported scenario (Table S8 & Fig. S16). This can be explained by  
30 geographic proximity of East Eurasia to Beringia and the genetic similarity of wolves from  
31 these areas.

## 32 **Discussion**

33 Recent whole-genome studies (3–5) found that modern grey wolves (*Canis lupus*) across  
34 Eurasia are descended from a single source population. The results of our analyses using both

1 ancient and modern grey wolf samples (Fig. 1) within a spatially and temporally explicit  
2 modeling framework (Fig. 3), suggest that this process began ~25,000 (95% CI:13,000-  
3 33,000) years ago when a population of wolves from Beringia (or a Northeast Asian region in  
4 close geographic proximity) expanded outwards and replaced indigenous Pleistocene wolf  
5 populations across Eurasia (Fig. 4). The star-like topology of modern wolves observed in  
6 these whole genome studies is also consistent with our inferred scenario (Fig. 4) in which the  
7 wave of expansion is divided by geographic barriers, leading to divergence of subpopulations  
8 within the Northern Hemisphere due to subsequent limited gene flow.

9 In the Americas, the Beringian expansion was delayed due to the presence of ice sheets  
10 extending from Greenland to the northern Pacific Ocean (Fig. 4) (13). A recent study by (17)  
11 suggested that wolf populations that were extant south of these ice sheets were replaced by  
12 Eurasian wolves crossing the Beringian land bridge. Our analyses support the replacement of  
13 North American wolves (following the retreat of the ice sheets), and our more extensive  
14 ancient DNA sampling combined with a spatially explicit model has allowed us to narrow  
15 down the geographic origin of this expansion.

16 Were the wolves before and after this replacement ecologically equivalent? Analyses of wolf  
17 specimens have noted morphological differences between Late-Pleistocene and Holocene  
18 wolves: late Pleistocene specimens have been described as cranio-dentally more robust than  
19 the present-day grey wolves, as well as having specialized adaptations for carcass and bone  
20 processing (18–20) associated with megafaunal hunting and scavenging (21, 22). In contrast,  
21 the early Holocene archaeological record has only yielded a single sample with the  
22 Pleistocene wolf morphotype (in Alaska) (19), suggesting that this ecomorph had largely  
23 disappeared from the Northern Hemisphere by the Pleistocene-Holocene transition. This  
24 change in wolf morphology coincides with a shift in wolf isotope composition (23) and the  
25 disappearance of many megafaunal herbivores as well as other large predators, such as cave  
26 hyenas and cave lions, suggesting a possible change in the ecological niche of wolves.

27 It has been unclear whether the morphological change was the result of population  
28 replacement (genetic turnover), a plastic response to a dietary shift, or both. Our results  
29 suggest that the Pleistocene-Holocene transition was accompanied by a genetic turnover in  
30 most of the Northern Hemisphere wolf populations since most indigenous wolf populations  
31 experienced a large-scale replacement resulting in the loss of all native Pleistocene genetic  
32 lineages (Fig. 4). Similar population dynamics of discontinuity and replacement by  
33 conspecifics have been observed in several other large Pleistocene mammals in Europe  
34 including cave bears, woolly mammoths (24, 25), giant deer (24) and even humans (11, 26).



1 The geographic exception to this pattern of widespread replacement is Beringia, where we  
2 infer demographic continuity between late Pleistocene and Holocene wolf populations (Fig.  
3 4). This finding is consistent with a recent study using the mtDNA control region by (27) that  
4 failed to reject continuity in this region, but at odds with a previous suggestion of genetic  
5 turnover in Beringia (19). This contradiction is likely the result of both the amount of data  
6 available and the analytical methodology: (19) used a short segment (427 bases long) of the  
7 mitochondrial control region and employed a descriptive phylogeographic approach, whereas  
8 our conclusions are based on an expanded dataset both in terms of sequence length, sample  
9 number, and geographic and temporal range (Fig. 1) and a formal hypothesis testing within a  
10 Bayesian framework (Figs. 3 and 4). As a consequence, the morphological and dietary shift  
11 observed in Beringian wolves between the late Pleistocene and Holocene (19) cannot be  
12 explained by a population turnover, but instead requires an alternative explanation such as  
13 adaptation or plastic responses to the substantial environmental and ecological changes that  
14 took place during this period. Indeed, grey wolves are a highly adaptable species. Studies of  
15 modern grey wolves have found that differences in habitat - specifically precipitation,  
16 temperature, vegetation, and prey specialization, can strongly affect their cranio-dental  
17 morphology (28–32).

18 The specific causal factors for the replacement of indigenous Eurasian wolves during the  
19 LGM by their Beringian conspecifics (and American wolves following the disappearance of  
20 the Cordilleran and Laurentide ice sheets) are beyond the scope of this study. However, one  
21 possible explanation may be related to the relatively stable climate of Beringia compared to  
22 the substantial climatic fluctuations that impacted the rest of Eurasia and Northern America  
23 during the late Pleistocene (33). These fluctuations have been associated with dramatic  
24 changes in food webs, leading to the loss of most of the large Pleistocene predators in the  
25 region (23, 34–36). In addition, the hunting of large Pleistocene predators by Upper  
26 Palaeolithic people (e.g. (37–39) may have also negatively impacted large carnivore  
27 populations (5). An interdisciplinary approach involving morphological, isotopic as well as  
28 genetic data is necessary to better understand the relationship between wolf population  
29 dynamics and dietary adaptations in the late Pleistocene and early Holocene period.

30 In summary, we have found that that, despite a continuous fossil record through the late  
31 Pleistocene, wolves experienced a complex demographic history involving population  
32 bottlenecks and replacements (Fig. 4). Our analysis suggests that long-range migration played  
33 an important role in the survival of wolves through the wave of megafaunal extinctions at the  
34 end of the last glaciation. These results will enable future studies to examine specific local  
35 climatic and ecological factors that enabled the Beringian wolf population to survive and  
36 expand across the Northern Hemisphere.

1 Lastly, the complex demographic history of Eurasian grey wolves reported here (Fig. 4) also  
2 has significant implications for identifying the geographic origin(s) of wolf domestication and  
3 the subsequent spread of dogs. For example, the limited understanding of the underlying wolf  
4 population structure may explain why previous studies have produced conflicting geographic  
5 and temporal scenarios. Numerous previous studies have focused on the patterns of genetic  
6 variation in modern domestic dogs, but have failed to consider potential genetic variation  
7 present in late Pleistocene wolf population, thereby implicitly assuming a homogenous wolf  
8 population source. As a result, both the domestication and the subsequent human-mediated  
9 movements of dogs were the only processes considered to have affected the observed genetic  
10 patterns in dog populations. However, both domestication from and admixture with a  
11 structured wolf population will have consequences for patterns of genetic variation within  
12 dogs.

13 In light of the complex demographic history of wolves (and the resulting population genetic  
14 structure) reconstructed by our analysis, several of the geographic patterns of haplotype  
15 distribution observed in previous studies, including differences in levels of diversity found  
16 within local dog populations (40), and the deep phylogenetic split between Eastern and  
17 Western Eurasian dogs (41), could have resulted from known admixture between domestic  
18 dogs and grey wolves (3, 5, 42, 43). Future analyses should therefore explicitly include the  
19 demographic history of wolves and demonstrate that the patterns of variation observed within  
20 dogs fall outside expectations that take admixture with geographically structured wolf  
21 populations into account.

## 1 MATERIALS & METHODS

### 2 Data preparation

3 We sequenced whole mitochondrial genomes of 40 ancient and 22 modern wolf samples.  
4 Sample information, including geographic locations, estimated ages and archaeological  
5 context information for the ancient samples, is provided in the Table S1 and Supplementary  
6 Information (SI) 1.2. Of the 40 ancient samples, 24 were directly radiocarbon dated for this  
7 study and calibrated using the IntCal13 calibration curve (see Table S1 for radiocarbon dates,  
8 calibrated age ranges and AMS laboratory reference numbers). DNA extraction, sequencing  
9 and quality filtering, and mapping protocols used are described in SI 2.

10

11 We included 16 previously published ancient mitochondrial wolf genomes (Table S1 and SI  
12 2). In order to achieve a uniform dataset, we re-processed the raw reads from previously  
13 published samples using the same bioinformatics pipeline as for the newly generated  
14 sequences.

15

16 We subjected the aligned ancient sequences to strict quality criteria in terms of damage  
17 patterns and missing data (Figs. S3 – S5). First, we excluded all whole mitochondrial  
18 sequences that had more than 1/3 of the whole mitochondrial genome missing (excluding the  
19 mitochondrial control region – see below) at minimum three-fold coverage. Secondly, we  
20 excluded all ancient whole mitochondrial sequences that contained more than 0.1% of  
21 singletons showing signs of deamination damage typical for ancient DNA (C to T or A to G  
22 singletons). After quality filtering, we were left with 32 newly sequenced and 13 published  
23 ancient whole mitochondrial sequences (Table S1).

24

25 We also excluded sequences from archaeological specimens that postdate the end of  
26 Pleistocene and that have been identified as dogs (Table S1), since any significant population  
27 structure resulting from a lack of gene flow between dogs and wolves could violate the  
28 assumption of a single, randomly mating canid population. Some of the Pleistocene  
29 specimens used in the demographic analyses (TH5, TH12, TH14) have been argued to show  
30 features commonly found in modern dogs and have therefore been suggested to represent  
31 Paleolithic dogs (e.g. (22, 44–48). Here, we disregard such status calls because of the  
32 controversy that surrounds them ( (49–52), and because early dogs would have been  
33 genetically similar to the local wolf populations from which they derived. This reasoning is  
34 supported by the close proximity of these samples to other wolf specimens confidently  
35 described as wolves in the phylogenetic tree (see Fig. S10).

36

1 Finally, we added 66 modern published wolf sequences from NCBI and two sequences from  
2 (3) (Table S1) resulting in a final dataset of 135 complete wolf mitochondrial genome  
3 sequences, of which 45 were ancient and 90 were modern. We used ClustalW alignment tool  
4 (version 2.1) (53) to generate a joint alignment of all genomes. In order to avoid the  
5 potentially confounding effect of recurrent mutations in the mitochondrial control region (54)  
6 in pairwise difference calculations, we removed this region from all subsequent analyses. This  
7 resulted in an alignment of sequences 15,466bp in length, of which 1301 sites (8.4%) were  
8 variable. The aligned dataset is located in Supplementary File S1.

9

## 10 **Phylogenetic analysis**

11 We calculated the number of pairwise differences between all samples (Fig. S6) and  
12 generated a neighbor-joining tree based on pairwise differences (Fig. S7). This tree shows a  
13 clade consisting of samples exclusively from the Tibetan region and the Indian sub-continent  
14 that are deeply diverged from all ancient and other modern wolf samples (see also (55, 56)). A  
15 recent study of whole genome data showed a complex history of South Eurasian wolves (5)  
16 that is beyond the scope of our study. While their neighbor-joining phylogeny grouped South  
17 Eurasian wolves with East and North East Asian wolves (Fig. 3 in Fan et al. (2016)), they  
18 cluster outside of all other grey wolves in a Principal Component Analysis (Fig. 4 in (5)), and  
19 also show a separate demographic history within a PSMC analysis (Fig. 5 in (5)). Because our  
20 study did not possess sufficient samples from the Himalayas and the Indian subcontinent to  
21 unravel their complex demography, we excluded samples from these regions and focused on  
22 the history of North Eurasian and North American wolves, for which we have good coverage  
23 through time and space.

24

25 We used PartitionFinder (57) and BEAST (v.1.8.0) (14) to build a tip calibrated wolf  
26 mitochondrial tree (with a strict global clock, see SI 3.2 for full details) from modern and  
27 directly dated ancient samples, and to estimate mutation rates for four different partitions of  
28 the wolf mitochondrial genome (see Tables S3 and S4 for results).

29

30 We used BEAST to molecularly date seven sequences from samples that were not directly  
31 radiocarbon dated (TH4, TH6, TH14, TU15) or that had been dated to a period beyond the  
32 limit of reliable radiocarbon dating (>48,000 years ago) (CGG12, CGG29, CGG32). We  
33 estimated the ages of the samples by performing a BEAST run where the mutation rate was  
34 fixed to the mean estimates from the previous BEAST analysis and all other parameter  
35 settings were set as described in the SI 3.2. We cross-validated this approach through a leave-

1 one-out analysis where we sequentially removed a directly dated sample and estimated its  
2 date as described above. We find a close fit ( $R^2=0.86$ ) between radiocarbon and molecular  
3 dates (Fig. S9). We combined the seven undated samples with the 110 ancient and modern  
4 samples from the previous run and used a uniform prior ranging from 0 to 100,000 years to  
5 estimate the ages of the seven undated samples (see Table S5 for results).

6

7 Finally, in order to estimate the mitochondrial divergence time between the South Eurasian  
8 (Tibetan and Indian) and the rest of our wolf samples, we performed an additional BEAST  
9 run in which we included all modern and ancient grey wolves ( $N = 129$ ) as well as five  
10 Tibetan and one Indian wolf, and used parameters identical to the ones described above. The  
11 age of the ancient samples was set as the mean of the calibrated radiocarbon date distribution  
12 (for radiocarbon dated samples) or as the mean of the age distribution from the BEAST  
13 analyses (for molecularly dated samples).

14

#### 15 **Isolation by distance analysis**

16 We performed isolation by distance (IBD) analyses to see the extent to which wolf  
17 mitochondrial genetic variation shows population structure. To this end, we regressed the  
18 pairwise geographic distances between 84 modern wolf samples (Table S1) against their  
19 pairwise genetic (mitochondrial) distances. The geographic distance between all sample pairs  
20 was calculated in kilometres as the great circle distance from geographic coordinates, using  
21 the Haversine Formula (58) to account for the curvature of the Earth as follows:

$$22 \quad G_{ij} = 2r \arcsin \left( \sqrt{\sin^2((\phi_j - \phi_i)/2) + \cos(\phi_i) \cos(\phi_j) \sin^2((\lambda_i - \lambda_j)/2)} \right)$$

23 Where  $G$  is the distance in kilometers between individuals  $i$  and  $j$ ;  $\phi_i$  and  $\phi_j$  are the latitude  
24 coordinates of individuals  $i$  and  $j$ , respectively;  $\lambda_i$  and  $\lambda_j$  are the longitude coordinates of  
25 individuals  $i$  and  $j$ , respectively; and  $r$  is the radius of the earth in kilometers. The pairwise  
26 genetic distances were calculated as the proportion of sites that differ between each pair of  
27 sequences (excluding the missing bases), using *dist.dna* function in the R package APE (59).

28

#### 29 **Geographical deme definitions**

30 We represented the wolf geographic range as seven demes, defined by major geographic  
31 barriers through time.

- 1        1. The *European* deme is bordered by open water from the North and the West (the  
2            Arctic and the Atlantic oceans, respectively); the Ural Mountains from the East; and  
3            the Mediterranean, the Black and the Caspian Sea and the Caucasus mountains from  
4            the South.
- 5        2. The *Middle-Eastern* deme consists of the Arabian Peninsula, Anatolia and  
6            Mesopotamia and is bordered by the Black Sea, the Caspian Sea and the Aral Sea in  
7            the North; the Indian Ocean in the South; the Tien Shen mountain range, the Tibetan  
8            Plateau and the Himalayas from the East; and the Mediterranean Sea in the West.
- 9        3. The *Central North Eurasian* deme consist of the Siberian Plateau and is bordered by  
10           the Arctic Ocean from the North; the Ural Mountains from the West; the Lena River  
11           and mountain ranges of North Eastern Siberia (Chersky and Verkhoyansk ranges)  
12           from the East; and the Tien Shen mountain range, the Tibetan Plateau and the Gobi  
13           Desert from South-East.
- 14       4. The *East Eurasian deme* is bordered by the Tien Shen mountain range, the Tibetan  
15           Plateau and Gobi desert from the West; the Pacific Ocean from the East; and the Lena  
16           river and the mountain ranges of North Eastern Siberia (Chersky and Verkhoyansk  
17           ranges) from the North.
- 18       5. The *Beringia* deme spans the Bering Strait, which was a land bridge during large  
19           parts of the Late Pleistocene and the Early Holocene. It is bordered to the West by the  
20           Lena River and mountain ranges of North Eastern Siberia (Chersky and Verkhoyansk  
21           ranges), and to the South and East by the extent of the Cordillerian and Laurentide ice  
22           sheets during the Last Glacial Maximum.
- 23       6. The *Arctic North America* deme consists of an area of the North American continent  
24           east of the Rocky Mountains and west of Greenland, that was covered by ice during  
25           the last Glaciation and is at present known as the Canadian Arctic Archipelago.
- 26       7. The *North America* deme consists of an area in the Northern American sub-continent  
27           that was south of the Cordillerian and Laurentide ice sheets during the last glaciation  
28           (13).

29

### 30 **Demographic scenarios**

31 We tested a total of 16 demographic scenario combinations, from four different kinds of  
32 demographic scenarios (illustrated in Fig. 3a in the main text):

- 33        1) Static model (the null hypothesis) – neighboring demes exchange migrants, no  
34           demographic changes.

- 1        2) Bottleneck scenarios – demes exchange migrants as in the static model but
- 2            populations have different size in different time periods. We consider three time
- 3            periods: 0-15k years ago, 15k-40k years ago, and >40k years ago.
- 4        3) Expansion scenarios - demes exchange migrants like in the static model but a single
- 5            deme experiences an expansion starting between 5k and 40k years ago (at a minimum
- 6            rate of 1,000 years per deme, so the whole world could be colonized within 3,000
- 7            years or faster).
- 8        4) Combinations of scenarios 2 & 3.

9

## 10 **Population genetic coalescent framework**

11 We implemented coalescent population genetic models for the different demographic  
12 scenarios to sample gene genealogies.

13 In the static scenario, we simulated local coalescent processes (60) within each deme (scaled  
14 to rate  $1/K$  per pair of lineages, where  $K$  is the mean time to most recent common ancestor in  
15 a deme and is thus proportional to the effective population size). In addition, we moved  
16 lineages between demes according to a Poisson process with rate  $m$  per lineage. To match the  
17 geographic and temporal distribution of the data, we represented each sample with a lineage  
18 from the corresponding deme and date.

19 The bottleneck scenario was implemented as the static one but with piecewise constant values  
20 for  $K$  as a function of time. We considered three time periods, each with its own value of  $K$   
21 ( $K_1$ ,  $K_2$  and  $K_3$ ), motivated by the archaeological and genetic evidence of wolf population  
22 changes described in the main text. The first time period was from present to early Holocene,  
23 0-15k years ago. The second time period extended from early Holocene to late Pleistocene  
24 and covered the last glacial maximum, 15-40k years ago. Finally, the third time period  
25 covered the late Pleistocene and beyond, i.e. 40k years ago and older.

26 The population expansion scenarios were based on the static model but with an added  
27 population expansion model with founder effects and replacement of local populations (we  
28 refer to populations not yet replaced by the expansion as "indigenous"). Starting at time  $T$ , the  
29 population expanded from the initial deme and replaced its neighboring populations. After the  
30 start of the expansion, the expansion proceeded in fixed steps of  $\Delta T$  (in time). At each step,  
31 colonized populations replaced neighboring indigenous populations (if an indigenous deme  
32 bordered to more than one colonized deme, these demes contributed equally to the  
33 colonization of the indigenous deme). In the coalescent framework (that simulates gene  
34 genealogies backwards in time) the colonization events corresponds to forced migrations from  
35 the indigenous deme to the source deme. If there were more than one source deme, the source

1 of each lineage was chosen randomly with equal probability. Finally, founder effects during  
2 the colonization of an indigenous deme were implemented as a local, instantaneous  
3 population bottleneck in the deme (after the expansion), with a severity scaled to give a fixed  
4 probability  $x$  of a coalescent event for each pair of lineages in the deme during the bottleneck  
5 (61). ( $x=1$  correspond to a complete loss of genetic diversity in the bottleneck, and  $x=0$   
6 corresponds to no reduction in genetic diversity.)

7 Finally, the combined scenario of population expansion and bottlenecks was implemented by  
8 making the population size parameter  $K$  in the population expansion model time dependent as  
9 in the population bottleneck model.

10

### 11 **Approximate Bayesian Computation analysis**

12 We used Approximate Bayesian Computation (ABC) analysis (62) with ABCtoolbox (63) to  
13 formally test the fit of our different demographic models. This approach allows formal  
14 hypothesis testing using likelihood ratios in the cases where the demographic scenarios are  
15 too complex for a direct calculation of the likelihoods given the models. We used the most  
16 likely tree from BEAST (see SI 3.2 for details) as data, and simulated trees using the  
17 coalescent simulations described above.

18 To match the assumption of random mixing within each deme in the population genetic  
19 model, we removed closely related sequences if they came from the same geographic location  
20 and time period, by randomly retaining one of the closely related sequences to be included in  
21 the analysis (Table S1, column “Samples\_used\_in\_Simulation\_Analysis”).

22 To robustly measure differences between simulated and observed trees we use the matrix of  
23 time to most recent common ancestor (TMRCA) for all pairs of samples. This matrix also  
24 captures other allele frequency based quantities frequently used as summary statistics with  
25 ABC, such as  $F_{ST}$ , as they can be calculated from the components of this matrix.

26 In principle the full matrix could be used, but in practice it is necessary to use a small number  
27 of summary statistics for ABC to work properly (63). To this end, we grouped our seven  
28 demes into four super demes (Fig. S14), based on geographic proximity and genetic similarity  
29 in the dataset, and used mean TMRCA within each super deme and mean TMRCAs between  
30 all super demes as summary statistics in the ABC analysis. The four super demes are 1)  
31 Europe; 2) Middle East; 3) North East Eurasia, Beringia and East Eurasia combined; and 4)  
32 Artic and Continental North America combined. This resulted in 10 summary statistics in  
33 total.



1 An initial round of fitting the model showed that all scenarios underestimate the within-super  
2 deme TMRCA for the Middle East, while the rest of the summary statistics were well  
3 captured by the best fitting demographic scenarios. This could be explained by a scenario  
4 where the Middle East was less affected by the reduction in population size during the last  
5 glacial maximum. However, we currently lack sufficient number of samples from this area to  
6 explicitly test a more complex scenario such as this hypothesis. To avoid outliers biasing the  
7 likelihood calculations in ABC (63) we removed this summary statistic, resulting in nine  
8 summary statistics in total.

9 For each of the 16 scenarios we performed 1 billion simulations with randomly chosen  
10 parameter combinations, chosen from the following parameter intervals for the different  
11 scenarios:

- 12 • The static scenario:  $m$  in  $[0.001,20]$  and  $K$  in  $[0.01,100]$ .
- 13 • The bottleneck scenarios:  $m$  in  $[0.001,20]$  and  $K_1, K_2, K_3$  in  $[0.01,100]$ .
- 14 • The expansion scenarios:  $m$  in  $[0.001,20]$ ,  $K$  in  $[0.01,100]$ ,  $x$  in  $[0,1]$ ,  $T$  in  $[5,40]$  and  
15  $\Delta T$  in  $[0.001,1]$ . For expansion out of the North American scenario and the expansion  
16 out of the Arctic North American scenario, the glaciation and during the LGM in  
17 North American and sea level rise during the de-glaciation mean that  $T$  must be in the  
18 range  $[9,16]$
- 19 • The combined bottleneck and expansion scenarios:  $m$  in  $[0.001,20]$ ,  $K_1, K_2, K_3$  in  
20  $[0.01,100]$ ,  $x$  in  $[0,1]$ ,  $T$  in  $[5,40]$  and  $\Delta T$  in  $[0.001,1]$ .

21 The parameter  $m$  is measured in units of 1/1,000 years, and  $T$ ,  $\Delta T$ ,  $K$ ,  $K_1$ ,  $K_2$  and  $K_3$  are  
22 measured in units of 1,000 years. The parameters  $x$ ,  $T$  and  $\Delta T$  were sampled according to a  
23 uniform distribution over the interval, while all other parameters were sampled from a  
24 uniform distribution of their log-transformed values. To identify good parameter  
25 combinations for ABC, we first calculated the Euclidian square distances between predicted  
26 and observed statistics and restricted analysis to parameter combinations within the lowest  
27 tenth distance percentile. We then ran the ABCtoolbox (63) on the accepted parameter  
28 combinations to estimate posterior distributions of the model parameters, and to calculate the  
29 likelihood of each scenario as described in the ABCtoolbox manual.

30 See Table S6 for ABC likelihoods and Bayes factors for all demographic scenarios tested.

31 See Tables S7 and S8 for posterior probability estimates and Figs. S.15 and S16 for posterior  
32 density distributions for estimated parameters ( $\Delta T$ ,  $T$ ,  $\log_{10} K_1$ ,  $\log_{10} K_2$ ,  $\log_{10} K_3$ ,  $\log_{10} m$ ,  $x$ ) in  
33 the two most likely models (An expansion out of Beringia with a population size change and  
34 an expansion out of East Eurasia with a population size change).

1 **Data availability**

2 New sequences are available to download from GenBank database (accession numbers XX-  
3 XX). The raw sequence reads are available from ENA database (accession numbers XX-XX).  
4 The scripts used in the analyses are available up on request from L.L. and A.E.

5

## 1    **References**

- 2    1.    Puzachenko AY, Markova AK (2016) Diversity dynamics of large- and medium-sized  
3        mammals in the Late Pleistocene and the Holocene on the East European Plain:  
4        Systems approach. *Quat Int*. doi:10.1016/j.quaint.2015.07.031.
- 5    2.    Sotnikova M, Rook L (2010) Dispersal of the Canini (Mammalia, Canidae: Caninae) across  
6        Eurasia during the Late Miocene to Early Pleistocene. *Quat Int* 212(2):86–97.
- 7    3.    Freedman AH, et al. (2014) Genome Sequencing Highlights the Dynamic Early History of  
8        Dogs. *PLOS Genet* 10(1):e1004016.
- 9    4.    Skoglund P, Ersmark E, Palkopoulou E, Dalén L (2015) Ancient Wolf Genome Reveals an  
10        Early Divergence of Domestic Dog Ancestors and Admixture into High-Latitude Breeds.  
11        *Curr Biol* 25(11):1515–1519.
- 12    5.    Fan Z, et al. (2016) Worldwide patterns of genomic variation and admixture in gray  
13        wolves. *Genome Res* 26(2):163–173.
- 14    6.    Larson G, et al. (2012) Rethinking dog domestication by integrating genetics, archeology,  
15        and biogeography. *Proc Natl Acad Sci* 109(23):8878–8883.
- 16    7.    Groucutt HS, et al. (2015) Rethinking the dispersal of Homo sapiens out of Africa. *Evol*  
17        *Anthropol Issues News Rev* 24(4):149–164.
- 18    8.    Rambaut A (2000) Estimating the rate of molecular evolution: incorporating non-  
19        contemporaneous sequences into maximum likelihood phylogenies. *Bioinformatics*  
20        16(4):395–399.
- 21    9.    Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) Estimating Mutation  
22        Parameters, Population History and Genealogy Simultaneously From Temporally  
23        Spaced Sequence Data. *Genetics* 161(3):1307–1320.
- 24    10.    Rieux A, et al. (2014) Improved calibration of the human mitochondrial clock using  
25        ancient genomes. *Mol Biol Evol*:msu222.
- 26    11.    Posth C, et al. (2016) Pleistocene Mitochondrial Genomes Suggest a Single Major  
27        Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. *Curr Biol*  
28        26(6):827–833.
- 29    12.    Warmuth V, et al. (2012) Reconstructing the origin and spread of horse domestication in  
30        the Eurasian steppe. *Proc Natl Acad Sci* 109(21):8202–8206.
- 31    13.    Raghavan M, et al. (2015) Genomic evidence for the Pleistocene and recent population  
32        history of Native Americans. *Science*:aab3884.
- 33    14.    Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian Phylogenetics with  
34        BEAUti and the BEAST 1.7. *Mol Biol Evol* 29(8):1969–1973.
- 35    15.    Thalmann O, et al. (2013) Complete Mitochondrial Genomes of Ancient Canids Suggest a  
36        European Origin of Domestic Dogs. *Science* 342(6160):871–874.

- 1 16. Nielsen R, Beaumont MA (2009) Statistical inferences in phylogeography. *Mol Ecol*  
2 18(6):1034–1047.
- 3 17. Koblmüller S, et al. (2016) Whole mitochondrial genomes illuminate ancient  
4 intercontinental dispersals of grey wolves (*Canis lupus*). *J Biogeogr* 43(9):1728–1738.
- 5 18. Kuzmina I., Sablin MV (1993) Pozdnepleistotsenovy pesets verhnei Desny. *Materiali po*  
6 *mezozoickoi i kainozoickoi istorii nazemnykh pozvonochnykh* (Trudy 17  
7 Zoologicheskogo Instituta RAN 249), pp 93–104.
- 8 19. Leonard JA, et al. (2007) Megafaunal Extinctions and the Disappearance of a Specialized  
9 Wolf Ecomorph. *Curr Biol* 17(13):1146–1150.
- 10 20. Baryshnikov GF, Mol D, Tikhonov AN (2009) Finding of the Late Pleistocene carnivores in  
11 Taimyr Peninsula (Russia, Siberia) with paleoecological context. *Russ J Theriol* 8(2):107–  
12 113.
- 13 21. Fox-Dobbs K, Leonard JA, Koch PL (2008) Pleistocene megafauna from eastern Beringia:  
14 Paleoecological and paleoenvironmental interpretations of stable carbon and nitrogen  
15 isotope and radiocarbon records. *Palaeogeogr Palaeoclimatol Palaeoecol* 261(1):30–  
16 46.
- 17 22. Germonpré M, et al. (2017) Palaeolithic and prehistoric dogs and Pleistocene wolves  
18 from Yakutia: Identification of isolated skulls. *J Archaeol Sci* 78:1–19.
- 19 23. Bocherens H (2015) Isotopic tracking of large carnivore palaeoecology in the mammoth  
20 steppe. *Quat Sci Rev* 117:42–71.
- 21 24. Stuart AJ, Kosintsev PA, Higham TFG, Lister AM (2004) Pleistocene to Holocene  
22 extinction dynamics in giant deer and woolly mammoth. *Nature* 431(7009):684–689.
- 23 25. Palkopoulou E, et al. (2013) Holarctic genetic structure and range dynamics in the woolly  
24 mammoth. *Proc R Soc B* 280(1770):20131910.
- 25 26. Fu Q, et al. (2016) The genetic history of Ice Age Europe. *Nature* 534(7606):200–205.
- 26 27. Ersmark E, et al. (2016) From the Past to the Present: Wolf Phylogeography and  
27 Demographic History Based on the Mitochondrial Control Region. *Front Ecol Evol* 4.  
28 doi:10.3389/fevo.2016.00134.
- 29 28. Geffen E, Anderson MJ, Wayne RK (2004) Climate and habitat barriers to dispersal in the  
30 highly mobile grey wolf. *Mol Ecol* 13(8):2481–2490.
- 31 29. Pilot M, et al. (2006) Ecological factors influence population genetic structure of  
32 European grey wolves. *Mol Ecol* 15(14):4533–4553.
- 33 30. O’Keefe FR, Meachen J, Fet EV, Brannick A (2013) Ecological determinants of clinal  
34 morphological variation in the cranium of the North American gray wolf. *J Mammal*  
35 94(6):1223–1236.
- 36 31. Flower LOH, Schreve DC (2014) An investigation of palaeodietary variability in European  
37 Pleistocene canids. *Quat Sci Rev* 96(Supplement C):188–203.

- 1 32. Leonard JA (2015) Ecology drives evolution in grey wolves. *Evol Ecol Res* 16(6):461–473.
- 2 33. Clark PU, et al. (2012) Global climate evolution during the last deglaciation. *Proc Natl*  
3 *Acad Sci* 109(19):E1134–E1142.
- 4 34. Lister AM, Stuart AJ (2008) The impact of climate change on large mammal distribution  
5 and extinction: Evidence from the last glacial/interglacial transition. *Comptes Rendus*  
6 *Geosci* 340(9–10):615–620.
- 7 35. Hofreiter M, Stewart J (2009) Ecological Change, Range Fluctuations and Population  
8 Dynamics during the Pleistocene. *Curr Biol* 19(14):R584–R594.
- 9 36. Lorenzen ED, et al. (2011) Species-specific responses of Late Quaternary megafauna to  
10 climate and humans. *Nature* 479(7373):359–364.
- 11 37. Münzel SC, Conard NJ (2004) Change and continuity in subsistence during the Middle  
12 and Upper Palaeolithic in the Ach Valley of Swabia (south-west Germany). *Int J*  
13 *Osteoarchaeol* 14(3–4):225–243.
- 14 38. Germonpré M, Hämäläinen R (2007) Fossil Bear Bones in the Belgian Upper Paleolithic:  
15 The Possibility of a Proto Bear-Ceremonialism. *Arct Anthropol* 44(2):1–30.
- 16 39. Cueto M, Camarós E, Castañón P, Ontañón R, Arias P (2016) Under the Skin of a Lion:  
17 Unique Evidence of Upper Paleolithic Exploitation and Use of Cave Lion (*Panthera*  
18 *spelaea*) from the Lower Gallery of La Garma (Spain). *PLOS ONE* 11(10):e0163591.
- 19 40. Wang G-D, et al. (2016) Out of southern East Asia: the natural history of domestic dogs  
20 across the world. *Cell Res* 26(1):21–33.
- 21 41. Frantz LAF, et al. (2016) Genomic and archaeological evidence suggest a dual origin of  
22 domestic dogs. *Science* 352(6290):1228.
- 23 42. Verardi A, Lucchini V, Randi E (2006) Detecting introgressive hybridization between free-  
24 ranging domestic dogs and wild wolves (*Canis lupus*) by admixture linkage  
25 disequilibrium analysis. *Mol Ecol* 15(10):2845–2855.
- 26 43. Godinho R, et al. (2011) Genetic evidence for multiple events of hybridization between  
27 wolves and domestic dogs in the Iberian Peninsula. *Mol Ecol* 20(24):5154–5166.
- 28 44. Sablin M, Khlopachev G (2002) The Earliest Ice Age Dogs: Evidence from Eliseevichi 1.  
29 *Curr Anthropol* 43(5):795–799.
- 30 45. Germonpré M, et al. (2009) Fossil dogs and wolves from Palaeolithic sites in Belgium,  
31 the Ukraine and Russia: osteometry, ancient DNA and stable isotopes. *J Archaeol Sci*  
32 36(2):473–490.
- 33 46. Druzhkova AS, et al. (2013) Ancient DNA Analysis Affirms the Canid from Altai as a  
34 Primitive Dog. *PLOS ONE* 8(3):e57754.
- 35 47. Germonpré M, Lázničková-Galetová M, Sablin MV (2012) Palaeolithic dog skulls at the  
36 Gravettian Předmostí site, the Czech Republic. *J Archaeol Sci* 39(1):184–202.

- 1 48. Germonpré M, Lázničková-Galetová M, Losey RJ, Rääkkönen J, Sablin MV (2015) Large  
2 canids at the Gravettian Předmostí site, the Czech Republic: The mandible. *Quat Int*  
3 359–360:261–279.
- 4 49. Crockford SJ, Kuzmin YV (2012) Comments on Germonpré et al., *Journal of*  
5 *Archaeological Science* 36, 2009 “Fossil dogs and wolves from Palaeolithic sites in  
6 Belgium, the Ukraine and Russia: osteometry, ancient DNA and stable isotopes”, and  
7 Germonpré, Lázničková-Galetová, and Sablin, *Journal of Archaeological Science* 39,  
8 2012 “Palaeolithic dog skulls at the Gravettian Předmostí site, the Czech Republic.” *J*  
9 *Archaeol Sci* 39(8):2797–2801.
- 10 50. Morey DF (2014) In search of Paleolithic dogs: a quest with mixed results. *J Archaeol Sci*  
11 52:300–307.
- 12 51. Drake AG, Coquerelle M, Colombeau G (2015) 3D morphometric analysis of fossil canid  
13 skulls contradicts the suggested domestication of dogs during the late Paleolithic. *Sci*  
14 *Rep* 5:8299.
- 15 52. Perri A (2016) A wolf in dog’s clothing: Initial dog domestication and Pleistocene wolf  
16 variation. *J Archaeol Sci* 68:1–4.
- 17 53. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–  
18 2948.
- 19 54. Excoffier L, Yang Z (1999) Substitution rate variation among sites in mitochondrial  
20 hypervariable region I of humans and chimpanzees. *Mol Biol Evol* 16(10):1357–1368.
- 21 55. Sharma DK, Maldonado JE, Jhala YV, Fleischer RC (2004) Ancient wolf lineages in India.  
22 *Proc R Soc Lond B Biol Sci* 271(Suppl 3):S1–S4.
- 23 56. Aggarwal RK, Kivisild T, Ramadevi J, Singh L (2007) Mitochondrial DNA coding region  
24 sequences support the phylogenetic distinction of two Indian wolf species. *J Zool Syst*  
25 *Evol Res* 45(2):163–172.
- 26 57. Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PartitionFinder: Combined Selection of  
27 Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Mol Biol Evol*  
28 29(6):1695–1701.
- 29 58. Sinnott R (1984) Virtues of the Haversine. *Sky Telesc* 68(2):159.
- 30 59. Paradis E, Claude J, Strimmer K (2004) APE: Analyses of Phylogenetics and Evolution in R  
31 language. *Bioinformatics* 20(2):289–290.
- 32 60. Kingman JFC (1982) The coalescent. *Stoch Process Their Appl* 13(3):235–248.
- 33 61. Eriksson A, Mehlig B (2004) Gene-history correlation and population structure. *Phys Biol*  
34 1(4):220.
- 35 62. Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian Computation in  
36 Population Genetics. *Genetics* 162(4):2025–2035.

- 1 63. Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L (2010) ABCtoolbox: a
- 2 versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics*
- 3 11(1):116.
- 4
- 5

## 1 **Acknowledgements**

2 The authors are grateful to Daniel Klingberg Johansson & Kristian Murphy Gregersen from  
3 the Natural History Museum of Denmark; Gabriella Hürlimann from the Zurich Zoo; Jane  
4 Hopper from the Howlett's & the Port Lympne Wild Animal Parks; Cyrintha Barwise-Joubert  
5 & Paul Vercammen from the Breeding Centre for Endangered Arabian Wildlife; Link Olson  
6 from the University of Alaska Museum of the North; Joseph Cook & Mariel Campbell from  
7 the Museum of Southwestern Biology; Lindsey Carmichael & David Coltman from the  
8 University of Alberta; North American Fur Auctions; Department of Environment Nunavut  
9 and Environment and Natural Resources Northwest Territories for DNA samples from the  
10 modern wolves.

11 The authors are also grateful to the staff at the Danish National High-Throughput Sequencing  
12 Centre for technical assistance in the data generation; the Qimmeq project, funded by The  
13 Velux Foundations and Aage og Johanne Louis-Hansens Fond, for providing financial  
14 support for sequencing ancient Siberian wolf samples; the Rock Foundation (New York,  
15 USA) for supporting radiocarbon dating of ancient samples from the Yana site; to Stephan  
16 Nylinder from the Swedish Museum of Natural History for advice on phylogenetic analyses  
17 and Terry Brown from the University of Manchester for comments on this manuscript.

18 L.L., K.D. & G.L. were supported by Natural Environment Research Council, UK (grant  
19 numbers NE/K005243/1, NE/K003259/1); LL. was also supported by the European Research  
20 Council grant (339941-ADAPT); A.M. & A.E. were supported by the European Research  
21 Council Consolidator grant (grant number 647787-LocalAdaptation); L.F. & G.L. were  
22 supported by the European Research Council grant (ERC-2013-StG 337574-UNDEAD); T.G  
23 was supported by European Research Council Consolidator grant (681396-Extinction  
24 Genomics) & Lundbeck Foundation grant (R52-5062); O.T. was supported by the National  
25 Science Center, Poland (2015/19/P/NZ7/03971) with funding from EU's Horizon 2020  
26 program under the Marie Skłodowska-Curie grant agreement (665778) and Synthesys Project  
27 (BETAF 3062); V.P., E.P. & P.N. were supported by the Russian Science Foundation grant  
28 (N16-18-10265 RNF); A.P. was supported by the Max Planck Society; M.L-G. was supported  
29 by Czech Science Foundation grant (GAČR15-06446S).

30

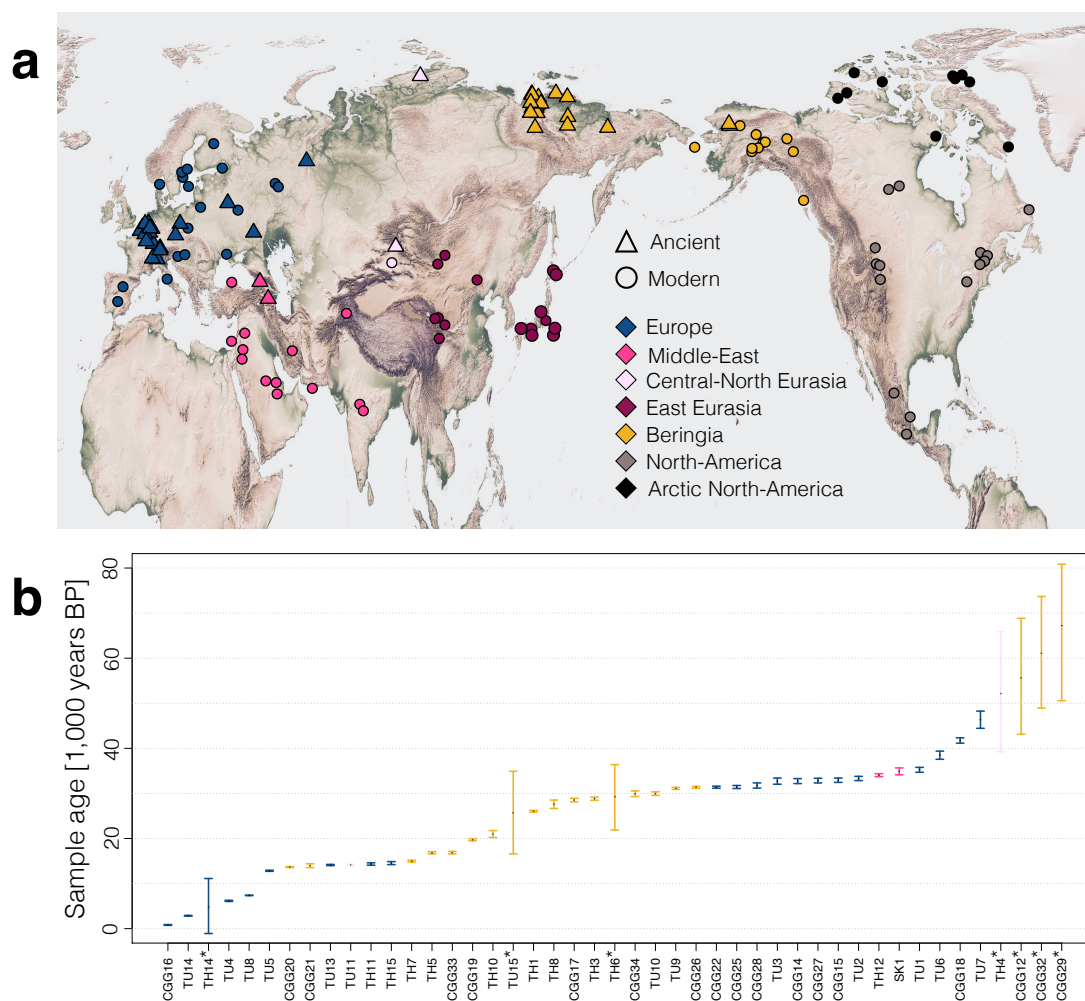
## 31 **Author contributions**

32 L.L., O.T., M.T.P.G., J.K., G.L., A.E. and A.M. designed the research; O.T., M-H.S.S.,  
33 V.J.S., K.E.W., M.S.V., I.K.C.L., N.W. and G.S. performed ancient DNA laboratory work  
34 with input from J.K., M.T.P.G., H.S., K-H.H., R.S.M. and K-H.H.; M-H.S.S. performed  
35 modern DNA laboratory work with input from M.T.P.G; O.T., J.A.S.C. and L.L. performed  
36 bioinformatic analyses; L.L., A.E. and A.M. designed the population genetic analyses; L.L.  
37 Performed phylogenetic analyses; A.E. implemented the spatial analyses framework; L.L. and



1 A.E. performed spatial analyses; M.G., J.B., V.V.P., E.Y.P., P.A.N., S.E.F., J.E-L., A.W.K.,  
2 B.G., H.N., H-P.U. and M.L-G. provided samples; V.V.P., M.G., M. L-G., H.B., H.N.,  
3 A.W.K., E.Y.P. and P.A.N. provided context for archaeological samples; A.P., M.G., H.B.  
4 and K.D. Helped setting the results of genetic analyses into an archaeological context; A.M.,  
5 M.T.P.G., A.J.H., G.L., J.K., E.W. and K.D. secured funding for the project; L.L., O.T. and  
6 A.E. wrote the initial draft of the manuscript with input from A.M.; L.L., O.T. and A.E wrote  
7 the manuscript and the supplementary information with input from A.P., M.G., H.B., M-  
8 H.S.S., M.T.P.G., K.E.W., A.M., G.L and K.D.; V.J.S., L.F., A.W.K., K-H.H., A.J.H.,  
9 R.S.M., H.S., G.S., V.V.P., E.Y.P., P.A.N. and J.E-L. provided comments to the manuscript  
10 and/or to the supplementary information.

## 1 Figures

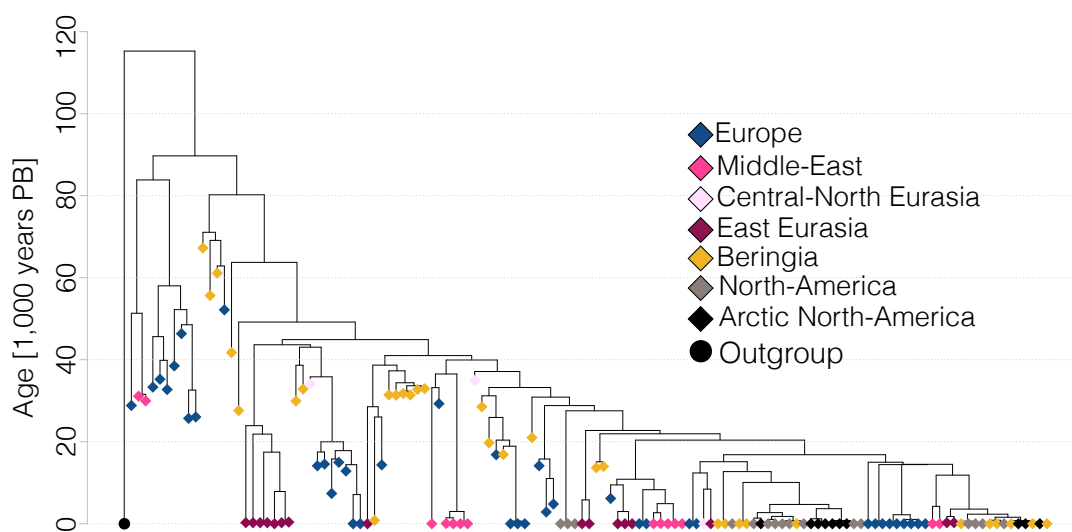


2

3 Fig. 1. Geographic distribution of modern (circles) and ancient (triangles) samples, grouped  
4 into seven geographic regions (demes, colour coded) (a) and temporal distribution of ancient  
5 samples (b) used in the analyses. \* Samples dated by molecular dating.

6

1



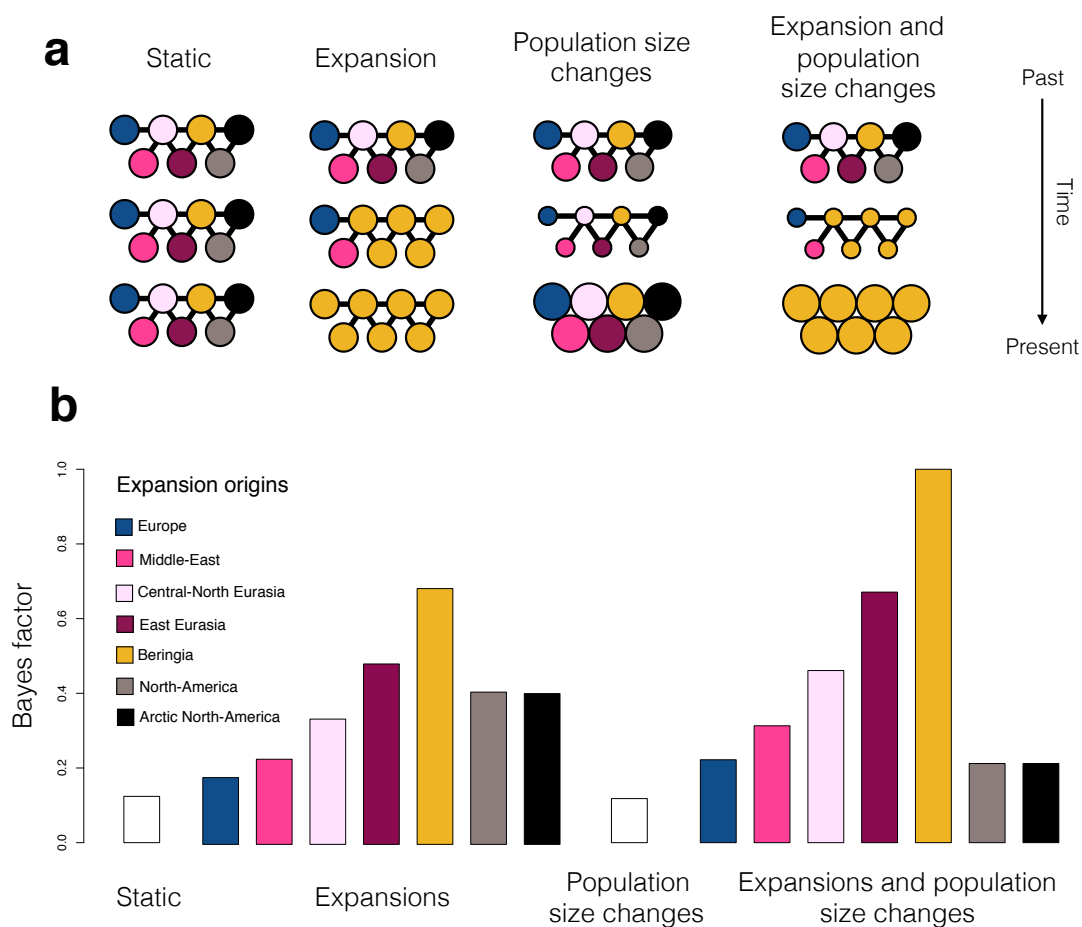
2

3

4 Fig. 2. Tip calibrated BEAST tree of all samples used in the spatial analyses (diamonds),  
5 coloured by geographic region. The circle represents an outgroup (modern Indian wolf, not  
6 used in the analyses).

7

1

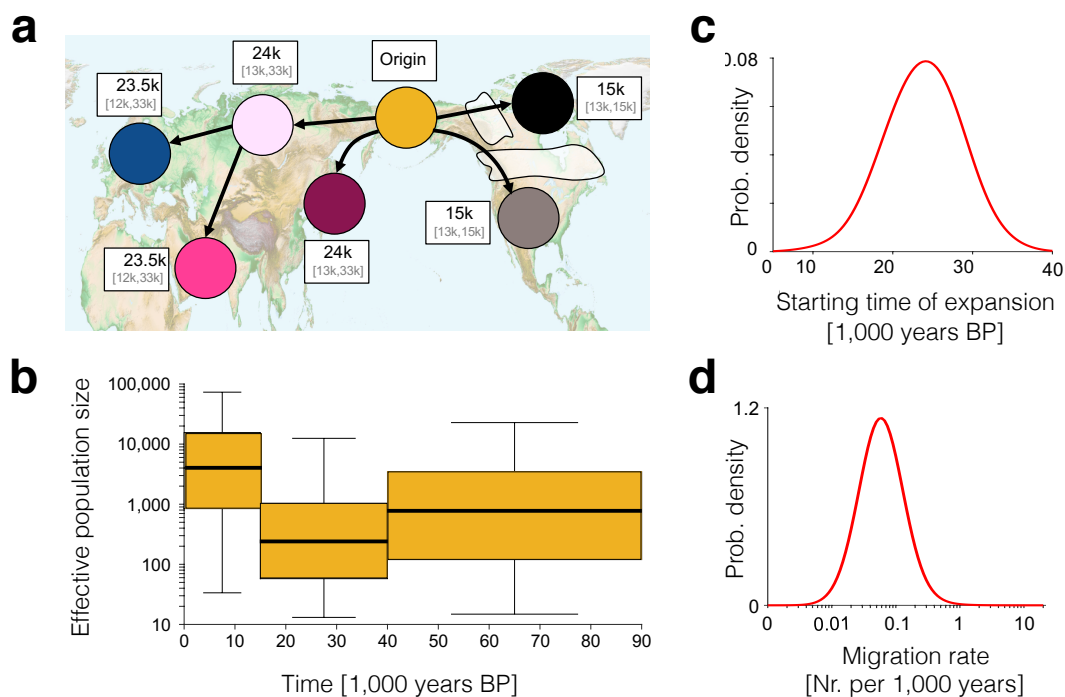


2

3

4 Fig. 3. Spatially and temporally explicit analysis. (a) Illustration of the different scenarios,  
5 with circles representing one deme each for the seven different geographic regions (see panel  
6 b for colour legend and text for full description of the scenarios). Solid lines represent  
7 population connectivity. The *static* scenario (far left) shows stable populations through time.  
8 The *expansion* scenarios (middle left) shows how one deme (here yellow, corresponding to  
9 Beringia) expands and sequentially replaces the populations in all other demes (from top to  
10 bottom). The *population size change scenario* (middle right) illustrates how population size in  
11 the demes can change through time (large or small population size shown as large or small  
12 circles, respectively). We also show a combined scenario (far right) of both expansion and  
13 population size change. (b) Likelihood of each demographic scenario relative to the most  
14 likely scenario, shown as Bayes factors, estimated using Approximate Bayesian Computation  
15 analyses (see text for details). For expansion scenarios (including the combined expansion  
16 and population size changes), we colour code each bar according to the origin of the  
17 expansion (see colour legend).

1



2

3 Fig. 4. The inferred scenario of wolf demography from the Bayesian analysis using our  
4 spatially and temporally explicit model (see Fig. 3 and the main text). (a) Geographic  
5 representation of the expansion scenario (out of Beringia) with median and 95% CI for the  
6 date of the population replacement in each deme given in white boxes next to each deme. (b)  
7 Effective population size (thick line, boxes and whiskers show the median, interquartile range  
8 and 95% CI, respectively, for each time period). (c) Posterior distribution of migration rate  
9 and (d) starting time of expansion.