Report

Current Biology

Middle Pleistocene genome calibrates a revised evolutionary history of extinct cave bears

Highlights

- The oldest genome sequence from a non-permafrost environment.
- A revised phylogeny of cave bears based on nuclear genomes.
- Direct estimation of ursid nuclear and mitochondrial substitution rates.
- Shows importance of climatic changes on species evolution.

Authors

Axel Barlow, Johanna L.A. Paijmans, Federica Alberti, ..., Love Dalén, Gennady Baryshnikov, Michael Hofreiter

Correspondence

axel.barlow.ab@gmail.com

In Brief

Barlow et al. present the oldest published genome from a non-permafrost environment: from a 360,000-year-old cave bear that inhabited the Southern Caucasus during the Middle Pleistocene. Using this and other cave bear genomes, they determine nuclear and mitochondrial substitution rates and revise the evolutionary history of the extinct cave bear.

Barlow et al., 2021, Current Biology *31*, 1771–1779
April 26, 2021
https://doi.org/10.1016/j.cub.2021.01.073

CellPress

Current Biology

CellPress

Report

Middle Pleistocene genome calibrates a revised evolutionary history of extinct cave bears

Axel Barlow,[1,2,13,*] Johanna L.A. Paijmans,[3] Federica Alberti,[1] Boris Gasparyan,[4] Guy Bar-Oz,[5] Ron Pinhasi,[6] Irina Foronova,[7] Andrey Y. Puzachenko,[8] Martina Pacher,[9] Love Dalén,[10,11] Gennady Baryshnikov,[12] and Michael Hofreiter[1]

[1]School of Science and Technology, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS, UK
[2]Institute for Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Strasse 24–25, 14476 Potsdam, Germany
[3]School of Archaeology and Ancient History, University of Leicester, University Road, Leicester LE1 7RH, UK
[4]Institute of Archaeology and Ethnography, National Academy of Sciences of the Republic of Armenia, 0025, RA, Yerevan, 15 Charents st., Armenia
[5]The Zinman Institute of Archaeology, University of Haifa, 199 Aba-Hushi Avenue, Haifa, Israel 3498838
[6]Department of Evolutionary Anthropology, University of Vienna, Althanstraße 14, 1090 Vienna, Austria
[7]V.S. Sobolev Institute of Geology and Mineralogy, Siberian Branch of the Russian Academy of Sciences, 3, Ac. Koptyuga ave., Novosibirsk, Russia 630090
[8]Institute of Geography, Russian Academy of Sciences, Staromonetnyy Pereulok, 29, Moscow, Russia 119017
[9]Naturmuseum St. Gallen, Rorschacher Strasse 263, CH-9016 St. Gallen, Switzerland
[10]Centre for Palaeogenetics, Stockholm University, Svante Arrhenius väg 20C, 106 91 Stockholm, Sweden
[11]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Frescativägen 54, 114 18 Stockholm, Sweden
[12]Zoological Institute, Russian Academy of Sciences, Universitetskaya Naberezhnaya 1, 199034 St. Petersburg, Russia
[13]Lead contact
*Correspondence: axel.barlow.ab@gmail.com
https://doi.org/10.1016/j.cub.2021.01.073

SUMMARY

Palaeogenomes provide the potential to study evolutionary processes in real time, but this potential is limited by our ability to recover genetic data over extended timescales.[1] As a consequence, most studies so far have focused on samples of Late Pleistocene or Holocene age, which covers only a small part of the history of many clades and species. Here, we report the recovery of a low coverage palaeogenome from the petrous bone of a ~360,000 year old cave bear from Kudaro 1 cave in the Caucasus Mountains. Analysis of this genome alongside those of several Late Pleistocene cave bears reveals widespread mito-nuclear discordance in this group. Using the time interval between Middle and Late Pleistocene cave bear genomes, we directly estimate ursid nuclear and mitochondrial substitution rates to calibrate their respective phylogenies. This reveals post-divergence mitochondrial transfer as the dominant factor explaining their mito-nuclear discordance. Interestingly, these transfer events were not accompanied by large-scale nuclear introgression. However, we do detect additional instances of nuclear admixture among other cave bear lineages, and between cave bears and brown bears, which are not associated with mitochondrial exchange. Genomic data obtained from the Middle Pleistocene cave bear petrous bone has thus facilitated a revised evolutionary history of this extinct megafaunal group. Moreover, it suggests that petrous bones may provide a means of extending both the magnitude and time depth of palaeogenome retrieval over substantial portions of the evolutionary histories of many mammalian clades.

RESULTS AND DISCUSSION

Analyses of palaeogenomes have provided unparalleled insights into the evolution of numerous vertebrate lineages. Assembling these datasets represents a considerable technical challenge, however, due to postmortem degradation and the loss of endogenous DNA molecules over time.[1] This is especially true for warm temperate and tropical environments, where DNA degradation proceeds more rapidly than in colder boreal or arctic environments.[1] As a consequence, comparatively few ancient DNA studies have successfully retrieved genetic data from temperate zone samples of Middle Pleistocene age (Chibanian age, 129–774 ka[2]). Notable successes include high coverage genome datasets from samples dating around the Middle to Late Pleistocene boundary from Germany[3,4] and from the Altai Mountains.[5] Much older DNA sequences have been retrieved from samples dating to ~430 ka from Spain, but with a lower magnitude of data recovery, comprising of mitochondrial genome sequences[6,7] and 1–2 megabases of nuclear DNA.[8] Nonetheless, these achievements suggest that the retrieval of genome-scale datasets of this age is possible, provided samples of sufficient quality can be found.

Middle Pleistocene cave bear genome

One recent and notable advance in palaeogenome sequencing has been the discovery of the mammalian petrous bone as a

source of high purity ancient DNA.[9] One group where this approach has been successfully applied are extinct cave bears,[10,11] which form the sister lineage to the clade consisting of the extant brown (*Ursus arctos*) and polar (*Ursus maritimus*) bears. To investigate if petrous bones may provide a way of extending both the time depth and magnitude of Middle Pleistocene genome data retrieval, we investigated the petrous bone of a Middle Pleistocene cave bear from Kudaro 1 cave located in South Ossetia in the Southern Caucasus. The sedimentary layer (5c) from which this sample was recovered has been dated using radiothermoluminescence at 360,000 ± 90,000 years,[12] and multiple additional sources of evidence support a Middle Pleistocene age for the specimen (see STAR methods). The Kudaro 1 sample is assigned to the taxon *Ursus kudarensis praekudarensis*, which is thought to be ancestral to the Late Pleistocene Caucasian cave bear *U. k. kudarensis* based on morphological evidence.[13,14]

We extracted DNA from the *praekudarensis* petrous bone and sequenced it using Illumina technology. From a total of ~2.6 billion sequenced molecules, we were able to map 2.1 Gb of sequence with high confidence to the reference genome assembly of the polar bear (Table S1). This represents a low coverage genome dataset where the majority of sequenced positions are covered by a single sequencing read (Figure S1). The estimated proportion of endogenous DNA molecules in the *praekudarensis* extract is 3.6% (Table S1), which is remarkable given the age of the specimen, and exceeds the endogenous proportions of some previously studied temperate-zone Middle Pleistocene extracts[8] by several orders of magnitude.

Cave bear nuclear and mitochondrial relationships are highly incongruent

Phylogenetic relationships among cave bear taxa have been largely guided by analysis of their mitochondrial DNA.[15,16] Our previous study on nuclear genomes did reveal one instance of mito-nuclear discordance among three Late Pleistocene taxa,[10] but the limited sampling of this study precluded a broader assessment of cave bear nuclear relationships. To investigate this further, we analyzed the Middle Pleistocene *praekudarensis* genome dataset alongside novel genome datasets of Late Pleistocene cave bears generated from their petrous bones, representing the taxa *rossicus* (Kizel Cave, Ural Mountains, Russia), *kanivetz* (Medvezhiya Cave, Ural Mountains, Russia), and *kudarensis* (Hovk 1 Cave, Armenia). We also included published datasets from the taxa *spelaeus* (Eiros Cave, Spain),[10] *eremus* (Windischkopf Cave, Austia),[10] *ingressus* (Gamssulzen Cave, Austria),[11] and a second *kudarensis* individual from Hovk 1 Cave[10] in addition to modern Georgian and Late Pleistocene Austrian brown bears,[10] two modern polar bears,[17] and a modern Asiatic black bear[18] as outgroup (see Tables S1 and S2).

We investigated relationships among the cave bear nuclear genomes using Principal Components Analysis (PCA). This involved sampling a single mapped nucleotide from each individual at each position of the reference genome, which provided data from a total of 487,747 variable transversion sites after strict filtering (STAR methods). PCA suggested three major groups (Figure 1A) comprising: the Caucasus cave bears *praekudarensis* and *kudarensis,* which cluster together as predicted by morphology; a second, geographically widespread and broadly European group including *spelaeus*, *ingressus*, *eremus,* and *kanivetz* from the Urals; and finally, the Urals cave bear *rossicus*, which is distinct from all other cave bears and may represent a Urals-specific group. PCA further suggests a hierarchy of relationships within these groups, with each successive PC separating different taxa from one another (Figure 1B).

The three major groups identified by the PCA deviate from expectations based on mitochondrial DNA, which instead supports two major clades comprising the Caucasus cave bears and all European and Urals cave bears, respectively, with *rossicus* nested within the latter (Figure 1C). We further investigated these contrasting nuclear relationships using phylogenetic analysis, including representative brown bears, polar bears, and the Asiatic black bear outgroup. Palaeogenomic datasets are generally associated with high rates of error, which can distort estimates of phylogenetic branch lengths when only a single read is sampled.[11] We therefore applied a recently developed method, Consensify,[11] which calls the majority base from a random sample of three mapped nucleotides. This method provided a single high-quality allele from each individual for a total of 4,318,414 genomic sites, of which 39,122 were variable. Maximum likelihood phylogenetic analysis of this dataset (Figure 1D) supported the expected relationships between polar bears, brown bears, and the cave bear clade, as well as the position of the Caucasus cave bears within the latter. Among the sampled cave bears from Europe and the Urals, however, there is not a single sister-group relationship that agrees between the mitochondrial and nuclear phylogenies.

The relationships among cave bears inferred using nuclear DNA closely match estimates based on morphological characters,[19,20] in contrast to mitochondrial relationships, which are frequently incongruent with morphology. Mitochondrial DNA, which has provided the basis for our understanding of cave bear relationships for several decades, thus emerges as a phylogenetic outlier in contradiction with generally congruent evidence from morphology and nuclear genomes. Mitochondrial evolution in cave bears therefore appears to have been shaped to a large extent by incomplete lineage sorting and/or gene flow among cave bear lineages, as previously documented in brown bears and polar bears.[17]

The revised cave bear nuclear genome phylogeny provides several new insights into their evolution. Consistent with the results of the nuclear PCA, the Uralian cave bear *rossicus* represents a deeply divergent and isolated phylogenetic lineage (Figure 1D), supporting its recognition as a third major cave bear group, which has been obscured until now due to reliance on mitochondrial DNA. Within the European cave bear group, three large-bodied taxa, *spelaeus*, *ingressus* and *kanivetz,* form a clade that is sister to the smaller bodied cave bear *eremus*. This suggests a single increase in body size in their common ancestor, rather than two independent shifts as previously inferred from their mitochondrial relationships.[15] Moreover, the validity of the taxon *kanivetz* has itself been questioned by mitochondrial phylogeographic studies,[15] which show it to be nested within the wider *ingressus* mitochondrial clade. Nuclear genome analysis, however, supports its distinctiveness and position as sister to the European taxa *spelaeus* and *ingressus*. Finally, the first three primary phylogenetic divisions of cave bears involve Uralian or Asian lineages, potentially reflecting an eastern

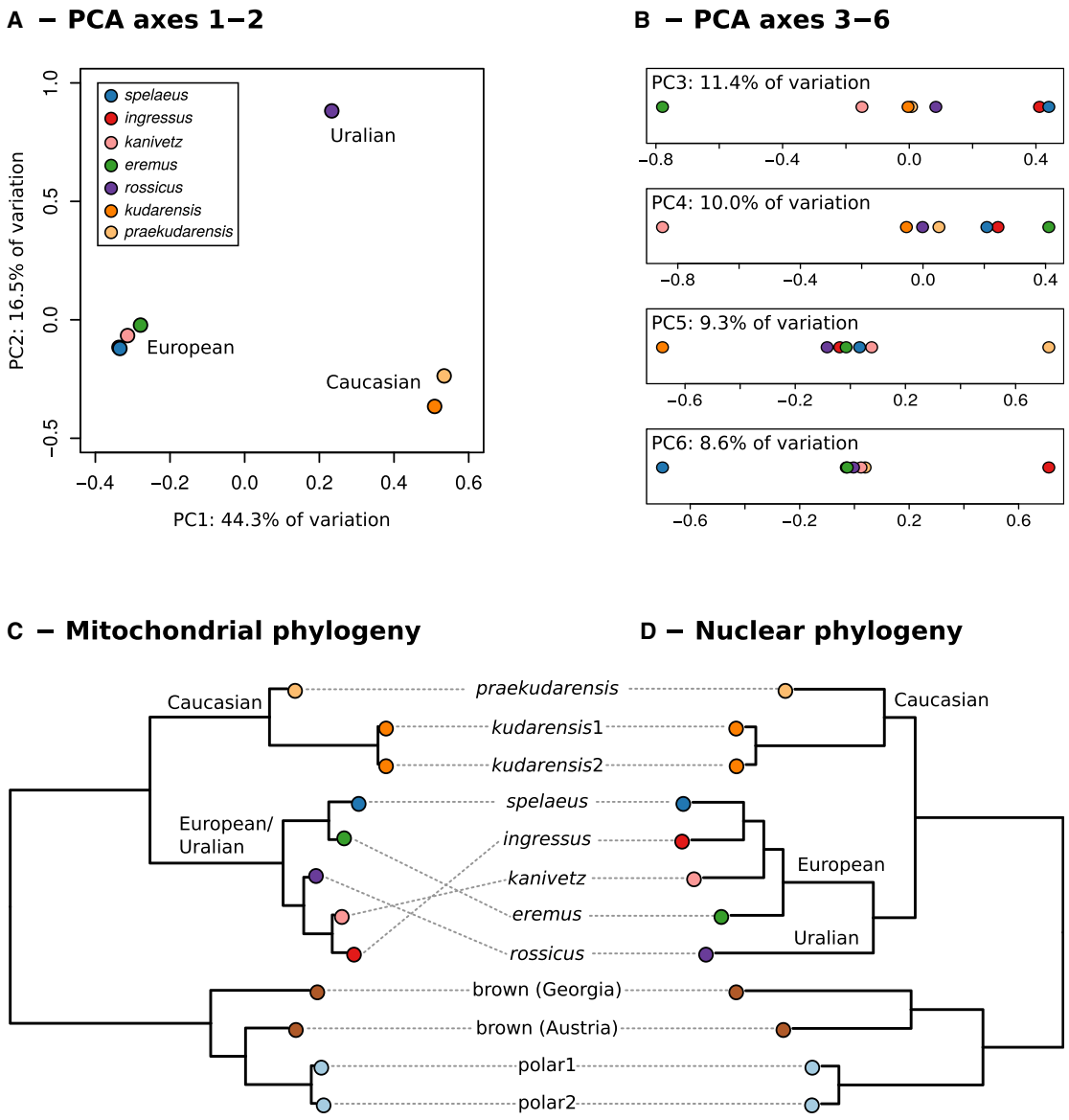


Figure 1. Cave bear relationships

(A) Ordination of individual cave bears along the first and second principal components of a PCA based on 487,747 filtered transversion sites supports three major groups.

(B) Ordination of the same individuals along PC3–PC6, which separate, respectively: *eremus, kanivetz, kudarensis* and *praekudarensis,* and *ingressus* and *spelaeus.*

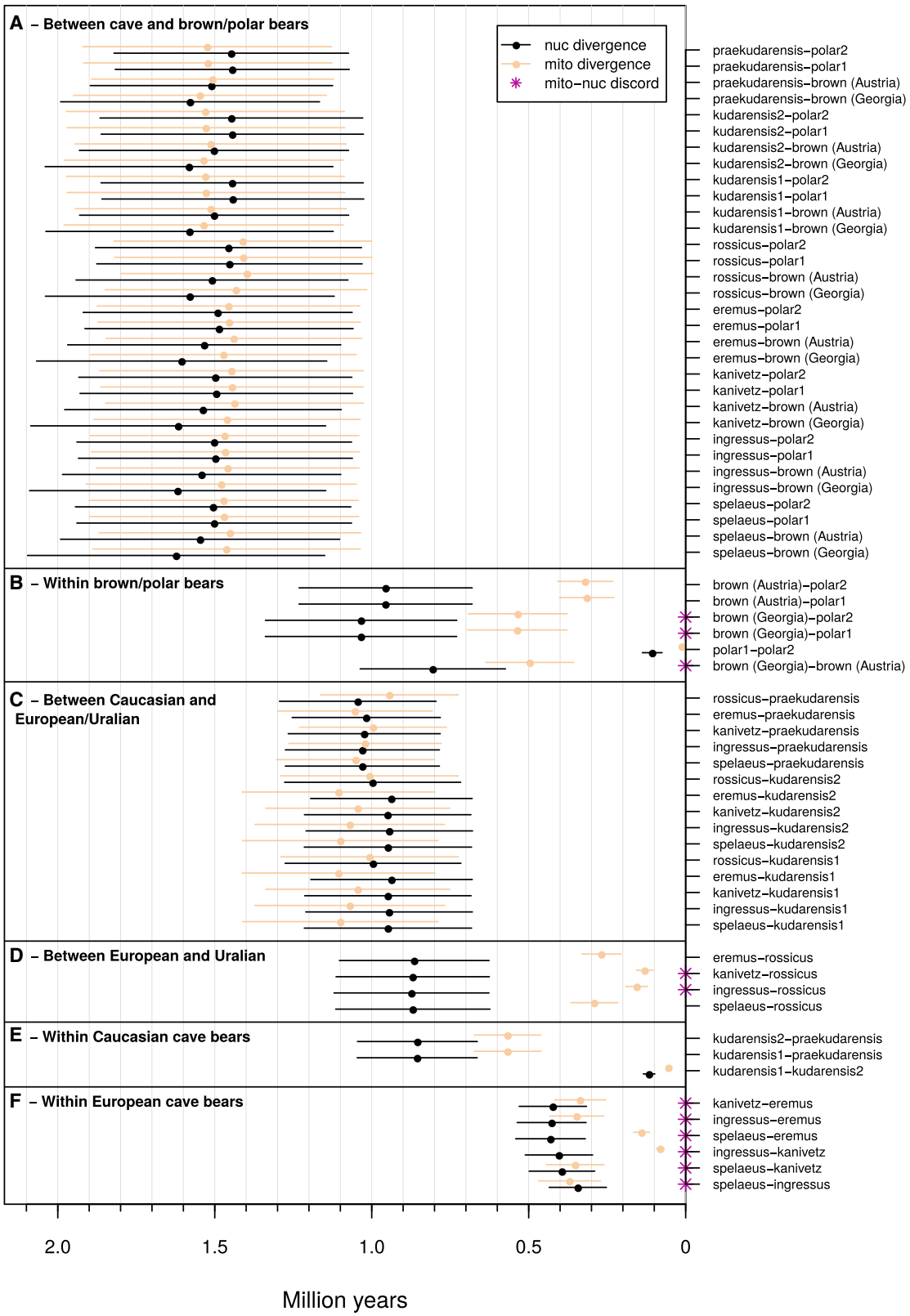
(C) Mitochondrial phylogeny based on 16,383 bp of aligned sequence and rooted using the Asiatic black bear outgroup (not shown). Bootstrap support is 100% for all nodes.

(D) Nuclear phylogeny based on 39,122 filtered, error-reduced variable sites. Rooting and bootstrap support are as described for C. There are multiple instances of mito-nuclear discordance within the European cave bears, among the European and Uralian cave bears, and among brown bears and polar bears.

origin for what has traditionally been regarded as a European radiation.[21]

Direct estimation of the genome-wide substitution rate

The temporal gap between *praekudarensis* and its Late Pleistocene sister lineage *kudarensis* provides an opportunity to estimate the *Ursus* substitution rate and calibrate the revised nuclear phylogeny of cave bears. We calculated the difference in their respective genetic divergences to modern brown bear and polar bear outgroups using the 4,318,414 Consensify error-reduced nuclear positions, which provides an estimate of the number of nucleotide substitutions occurring during their sampling interval. Based on the median *praekudarensis* radiothermoluminescence date and the age estimates of the *kudarensis* samples (Table S2), this equates to 305,400 years and yields a genome-wide nuclear substitution rate of 9.56×10^{-10} substitutions/site/year (range $7.39–13.56 \times 10^{-10}$ substitutions/site/year accommodating the radiothermoluminescence date uncertainty, Table S3). This



(legend on next page)

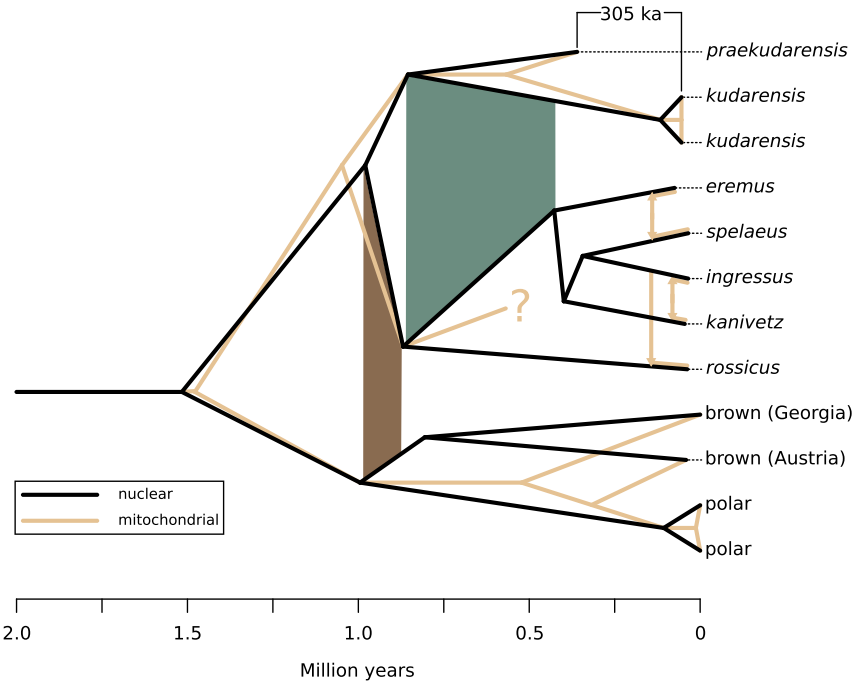


Figure 3. Calibrated nuclear and mitochondrial phylogenies

Branches terminate at the sample ages and nodes are centered on the mean of their respective pairwise estimates (see Figure 2 and Tables S3–S5). The complete mitochondrial evolutionary history of the European and Uralian cave bears is uncertain. The three recent mitochondrial transfer events that can be inferred from pairwise estimates (Figure 2) are indicated by vertical arrows. Shaded trapezoids connecting lineages indicate the two major episodes of nuclear gene flow identified by D-statistic analysis and are colored consistently with Figure 4.

estimate is substantially lower than ancient-DNA derived estimates for dogs and wolves (~1.2×10^{-8} substitutions/site/year[22]), but exceeds published estimates for humans and other great apes.[23] Notably, our estimated *Ursus* substitution rate is approximately double that estimated for humans (5×10^{-10} substitutions/site/year[24]), which aligns well with the difference in their respective generation times (brown and polar bears 11–12 years,[25,26] humans 20–25 years[27,28]), suggesting their underlying per-generation nuclear mutation rate is approximately equal. Applying the same methodology to mitochondrial DNA (excluding the control region) produced an estimated *Ursus* mitochondrial substitution rate of 1.81×10^{-8} substitutions/site/year (range $1.40–2.57 \times 10^{-8}$ substitutions/site/year, Table S3). This estimate falls within the lower range of mitochondrial substitution rate estimates for other vertebrates.[29] It also overlaps with estimated rates for human mitochondrial DNA,[30] deviating from the ratio predicted by generation times and suggesting some difference in the underlying rate of mitochondrial mutations between humans and bears.

We used these newly estimated substitution rates to calculate absolute times of nuclear and mitochondrial divergence for all pairs of individuals (Figure 2; Tables S4 and S5), and calibrate their respective phylogenies (Figure 3). We note that the divergence times of genetic lineages will be older than the divergence of their respective populations as they likely sample standing variation in the ancestral population. However, since the lineages under study span a large evolutionary timescale and show high levels of structuring (see D-statistic analysis below), the obtained times most likely provide reasonable approximations. Median nuclear divergence time estimates of cave bears and their sister clade, polar bears and brown bears, were found to be around 1.52 Ma (Figures 2A and 3). The mitochondrial divergence time is similar, around 1.48 Ma (Figures 2A and 3). These estimates are more recent than most previous estimates based on mitochondrial DNA, but highly consistent with the fossil record (see STAR methods). Median estimates for the polar bear and brown bear nuclear divergence are around 0.99 Ma (Figures 2B and 3), which is highly similar to previous phylogenetic estimates.[18] Median nuclear divergence times of the three major cave bear clades also fall around the same time, around 0.98 (Figures 2C and 3) and 0.87 Ma (Figures 2D and 3). Notably, these primary divergence events among cave bears, as well as between brown bears and polar bears, coincide with the Middle Pleistocene Transition, 1.2–0.8 Ma, when glacial cycles shifted from a ~40 to a ~100 ka periodicity causing extended glacial periods and more abrupt, intense interglacials,[31] which may have been a factor promoting their divergence. Within the Caucasian cave bears, we find a comparatively deep divergence time between *kudarensis* and *praekudarensis* (median estimate ~495 ka at the age of *praekudarensis*; Figures 2E and 3), potentially arguing against the direct ancestor descendant-relationship suggested by morphology, or, alternatively, for a genetically structured *praekudarensis* population. Among the European cave bears, we find a comparatively rapid sequence of divergence events among the four sampled taxa (median ages 427–344 ka, Figures 2F and 3), which are notably older than previous estimates based on mitochondrial DNA.[15,32]

Mitochondrial transfer explains mito-nuclear discordance in cave bears

Time calibration of the nuclear and mitochondrial cave bear phylogenies could reveal the underlying causes of their discordance.

Figure 2. Pairwise nuclear and mitochondrial divergence times

Points indicate ages estimated from substitution rates calculated using the median radiothermoluminescence age of the *praekudarensis* sample of 360 ka. Error bars represent the maximum and minimum ages reflecting the ± 90 ka uncertainty of the radiothermoluminescence age (see Tables S3–S5). Note that since the age uncertainties of the pairwise estimates are not independent, the rank order and relative separation of the point estimates will be maintained irrespective of the true *praekudarensis* age. Sample pairs showing mito-nuclear discordance are marked with asterisks. Results for specific clades discussed in the text are indicated (A–F).

Figure 4. D-statistics tests of cave bear admixture

Results are expressed as D(P1,P2,P3,P4) with significant positive values (filled circles) indicating admixture between P2 and P3 subsequent to the divergence of P1 and P2, colored consistently with the shaded trapezoids in Figure 3. To aid visualization, the assignment of the P1 and P2 individuals has been adjusted to make their respective D value positive. The results indicate two major episodes of nuclear gene flow among the sampled lineages: between brown bears and the ancestor of the European and Uralian cave bears, subsequent to their divergence from their respective sister lineages, polar bears D(polar,brown,cave,out) and the Caucasian cave bears D(Caucasian, European/Uralian,brown−polar,out); and between *kudarensis* and the ancestor of the European cave bears subsequent to their divergence from their respective sister lineages, *praekudarensis* D(praekudarensis,kudarensis, European/Uralian,out) and the Uralian cave bear *rossicus* D(Uralian,European, Caucasian,out).

Specifically, mitochondrial coalescence that considerably post-dates the respective nuclear coalescence of taxa implies the transfer of mitochondrial DNA through admixture. Examination of cave bear pairs exhibiting mito-nuclear discordance (defined as pairs with different taxa descending from the ancestral nodes of their respective nuclear and mitochondrial clades) reveals that all pairs but one (*spelaeus* and *ingressus*) show mitochondrial coalescence that is more recent than their respective nuclear coalescence (Figure 2; Tables S3 and S4), implicating mitochondrial transfer as the dominant factor shaping their mitochondrial relationships.

Multiple alternative transfer scenarios could explain the observed mitochondrial relationships, and selecting among them is challenging based on the available data. One notable set of comparisons are those involving the Uralian cave bear *rossicus* and the European cave bears *ingressus* and *kanivetz*, whose mitochondrial coalescence times post-date their respective nuclear coalescence by more than 700 ka (Figure 2D). Since none of the sampled European or Uralian mitochondrial lineage pairs approaches the nuclear coalescence of *rossicus* and the European cave bear group (median estimate ~0.87 Ma), it seems most likely that *ingressus* or *kanivetz* transferred its mitochondrial DNA to *rossicus*, although alternative scenarios such as multiple transfers or back transfers cannot be conclusively excluded. Also notable are mitochondrial coalescence times of the European cave bears *spelaeus* and *eremus*, and *ingressus* and *kanivetz*, which both considerably post-date their respective nuclear divergences (> 290 ka based on median substitution rate estimates; Figure 2F), implicating mitochondrial transfer also between lineages of these pairs. Determining the direction of these transfer events is challenging due the the rapid sequence of nuclear divergence events in European cave bears, around which many of their mitochondrial coalescence events also cluster (Figure 2F). The precise history of mitochondrial evolution is likely to remain uncertain until genetic data from samples pre-dating the inferred mitochondrial transfer events is obtained, which would require obtaining DNA sequences from multiple additional Middle Pleistocene specimens.

Mitochondrial transfer is not strongly associated with nuclear introgression

Since our analyses reveal numerous instances of mitochondrial transfer among cave bears, we investigated evidence of nuclear gene flow using D-statistic analysis of the error reduced genome sequences, which has been shown to represent a conservative approach that is more resistant to false positives than other methods.[11] Interestingly, we failed to detect nuclear gene flow among those cave bear taxa implicated in mitochondrial transfer events (Figure 4), suggesting that any residual nuclear introgression resulting from these gene flow events is below the detection

threshold of our low coverage genome datasets. In fact, only a single nuclear gene flow event is detected among the sampled cave bear lineages, between the lineage leading to *kudarensis* and the lineage leading to the European cave bears (Figure 4). Notably, this nuclear admixture event is not accompanied by any evidence of mitochondrial transfer. Similarly, we have previously shown that gene flow occurred between cave bears and brown bears,[10] again with no evidence of mitochondrial exchange. Replicating these tests including our novel cave bear genomes consistently supports this previous result, and moreover, our improved sampling pinpoints the common ancestor of the European and Uralian cave bears as the admixing cave bear lineage (Figure 4), with no evidence of gene flow following the divergence of its descendant clades.

This disparity between instances of nuclear and mitochondrial introgression within a recently diverging clade is puzzling. Enhanced mitochondrial introgression is documented in species with male-biased dispersal, such as bears, since this reduces the effective population size of maternally inherited loci in the incoming species at the contact zone.[33] However, this process does not explain the apparent absence of mitochondrial mixing among Caucasian and European cave bears, or among cave bears and brown bears. Here, some additional factor, such as a complete absence of females representing the incoming species, preferential mating of hybrids with their maternal species, or asymmetrical hybrid sterility,[34] must have operated.

Conclusions

We have shown that petrous bones provide a way to extend both the time depth and magnitude of Middle Pleistocene genome sequencing. The palaeogenome of the Middle Pleistocene *praekudarensis* cave bear sequenced in this study has provided important insights into the evolution of this iconic group of extinct animals. Critically, by providing a means of calculating the genome substitution rate, it has calibrated their evolutionary history, revealed numerous instances of mitochondrial transfer, and suggested a potential link between a profound change in climate dynamics and the divergence of major evolutionary lineages.

Palaeogenomes provide the opportunity to study the process of evolution in real time. For many temperate zone species, however, the current time depth for palaeogenome recovery represents a comparatively small part of their total evolutionary history. The age of the *praekudarensis* genome pre-dates the origin of some cave bear taxa and encompasses an estimated 24% of the total evolutionary history of cave bears. Petrous bones may thus provide a means of extending the time depth of palaeogenome recovery over substantial fractions of the evolutionary histories of many temperate zone species.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Laboratory methods
  - Data processing
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Assessment of ancient DNA authenticity
  - Nuclear genome PCA
  - Generation of error-reduced genome sequences
  - Phylogenetic analyses
  - Molecular dating
  - Comparison with other divergence estimates
  - Tests of nuclear admixture

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cub.2021.01.073.

ACKNOWLEDGMENTS

We pay tribute to our coauthor Irina Foronova for her contributions to the field of paleontology, who sadly passed away prior to the publication of this study. We thank Anthony J. Stuart for his help collecting radiocarbon dates of Ural cave bears. The authors would like to acknowledge support from the Science for Life Laboratory, the National Genomics Infrastructure (NGI), Sweden, the Knut and Alice Wallenberg Foundation and UPPMAX for providing assistance in massively parallel DNA sequencing and computational infrastructure. This work was funded by European Research Council (ERC) starting grant "gene flow" 310763 to M.H. The work was further carried out within a framework of the Federal theme of the Theriology laboratory of the Zoological Institute of the Russian Academy of Sciences no. АААА-А19-119032590102-7 "Phylogeny, morphology and systematics of placental mammals"; the project "The evolution of the organic world. Role and influence of planetary processes (subprogram I. Development of life and biosphere processes)" (G.B.); the Federal theme of the Laboratory of Biogeography of Institute of Geography of the RAS no. АААА-А19-119021990093-8 "Evaluation of physical and geographical, hydrological, and biotic environmental changes and their effects for the base of sustainable environmental management" (A.Y.P.); supported partly by an RFBR grant to A.Y.P. (17-01-00100-а); and on the state assignment of the V.S. Sobolev Institute of Geology and Mineralogy, Siberian Branch of the RAS (I.F.).

AUTHOR CONTRIBUTIONS

In alphabetical order: Conceptualization, A.B., G.B., and M.H.; Data curation, A.B.; Formal analysis, A.B. and J.L.A.P.; Funding acquisition, M.H.; Investigation, F.A. and M.P.; Methodology, A.B.; Project administration, A.B., L.D., and M.H.; Resources, B.G., E.B., G.B., G.B.-O., I.F., L.D., and R.P.; Software, A.B. and J.L.A.P.; Supervision, A.B., A.Y.P., G.B., and M.H.; Validation, A.B. and J.L.A.P.; Visualization, A.B.; Writing – original draft, A.B.; Writing – review & editing, all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 9, 2020
Revised: October 19, 2020
Accepted: January 21, 2021
Published: February 15, 2021

REFERENCES

1. Hofreiter, M., Paijmans, J.L.A., Goodchild, H., Speller, C.F., Barlow, A., Fortes, G.G., Thomas, J.A., Ludwig, A., and Collins, M.J. (2015). The future

of ancient DNA: Technical advances and conceptual shifts. BioEssays *37*, 284–293.

2. Cohen, K., Finney, S., Gibbard, P.L., and Fan, J.-X. (2020). The ICS International Chronostratigraphic Chart (updated). Episodes *36*, 199–204.
3. Palkopoulou, E., Lipson, M., Mallick, S., Nielsen, S., Rohland, N., Baleka, S., Karpinski, E., Ivancevic, A.M., To, T.-H., Kortschak, R.D., et al. (2018). A comprehensive genomic history of extinct and living elephants. Proc. Natl. Acad. Sci. USA *115*, E2566–E2574.
4. Meyer, M., Palkopoulou, E., Baleka, S., Stiller, M., Penkman, K.E.H., Alt, K.W., Ishida, Y., Mania, D., Mallick, S., Meijer, T., et al. (2017). Palaeogenomes of Eurasian straight-tusked elephants challenge the current view of elephant evolution. eLife *6*, 1–14.
5. Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A., Renaud, G., Sudmant, P.H., de Filippo, C., et al. (2014). The complete genome sequence of a Neanderthal from the Altai Mountains. Nature *505*, 43–49.
6. Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., and Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. USA *110*, 15758–15763.
7. Meyer, M., Fu, Q., Aximu-Petri, A., Glocke, I., Nickel, B., Arsuaga, J.-L., Martínez, I., Gracia, A., de Castro, J.M., Carbonell, E., and Pääbo, S. (2014). A mitochondrial genome sequence of a hominin from Sima de los Huesos. Nature *505*, 403–406.
8. Meyer, M., Arsuaga, J.-L., de Filippo, C., Nagel, S., Aximu-Petri, A., Nickel, B., Martínez, I., Gracia, A., Bermúdez de Castro, J.M., Carbonell, E., et al. (2016). Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. Nature *531*, 504–507.
9. Gamba, C., Jones, E.R., Teasdale, M.D., McLaughlin, R.L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., et al. (2014). Genome flux and stasis in a five millennium transect of European prehistory. Nat. Commun. *5*, 5257.
10. Barlow, A., Cahill, J.A., Hartmann, S., Theunert, C., Xenikoudakis, G., Fortes, G.G., Paijmans, J.L.A., Rabeder, G., Frischauf, C., Grandal-d'Anglade, A., et al. (2018). Partial genomic survival of cave bears in living brown bears. Nat. Ecol. Evol. *2*, 1563–1570.
11. Barlow, A., Hartmann, S., Gonzalez, J., Hofreiter, M., and Paijmans, J.L.A. (2020). Consensify: A method for generating pseudohaploid genome sequences from palaeogenomic datasets with reduced error rates. Genes (Basel) *11*, 11.
12. Lioubine, V.P. (1998). The Acheulian Epoch in the Caucasus (St. Petersburg: In Russian).
13. Baryshnikov, G.F. (1998). Cave bears from the Paleolithic of the Greater Caucasus. Illinois State Mus. Sci. Pap. *XXVII*, 69–118.
14. Baryshnikov, G.F., and Puzachenko, A.Y. (2019). Evolution and morphological variability of cheek teeth in the Kudaro cave bear (*Ursus kudarensis*, Carnivora, Ursidae). Zool. Zh. *98*, 1112–1136.
15. Stiller, M., Molak, M., Prost, S., Rabeder, G., Baryshnikov, G., Rosendahl, W., Münzel, S., Bocherens, H., Grandal-d'Anglade, A., Hilpert, B., et al. (2014). Mitochondrial DNA diversity and evolution of the Pleistocene cave bear complex. Quat. Int. *339–340*, 224–231.
16. Barlow, A., Hofreiter, M., and Knapp, M. (2019). Cave bears and ancient DNA: a mutually beneficial relationship. In Berichte der Geologischen Bundesanstalt, D. Nagel, and N. Kavcik-Graumann, eds. (Geologische Bundesanstalt), pp. 33–45.
17. Cahill, J.A., Green, R.E., Fulton, T.L., Stiller, M., Jay, F., Ovsyanikov, N., Salamzade, R., St John, J., Stirling, I., Slatkin, M., and Shapiro, B. (2013). Genomic evidence for island population conversion resolves conflicting theories of polar bear evolution. PLoS Genet. *9*, e1003345.
18. Kumar, V., Lammers, F., Bidon, T., Pfenninger, M., Kolter, L., Nilsson, M.A., and Janke, A. (2017). The evolutionary history of bears is characterized by gene flow across species. Sci. Rep. *7*, 46487.
19. Baryshnikov, G.F., and Puzachenko, A.Y. (2011). Craniometrical variability in the cave bears (Carnivora, Ursidae): Multivariate comparative analysis. Quat. Int. *245*, 350–368.
20. Baryshnikov, G.F., and Puzachenko, A.Y. (2019). Morphometry of upper cheek teeth of cave bears (Carnivora, Ursidae). Boreas *48*, 581–604.
21. Kurtén, B. (1976). The cave bear story. Life and death of a vanished animal (Columbia University Press).
22. Skoglund, P., Ersmark, E., Palkopoulou, E., and Dalén, L. (2015). Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. Curr. Biol. *25*, 1515–1519.
23. Besenbacher, S., Hvilsom, C., Marques-Bonet, T., Mailund, T., and Schierup, M.H. (2019). Direct estimation of mutations in great apes reconciles phylogenetic dating. Nat. Ecol. Evol. *3*, 286–292.
24. Scally, A. (2016). The mutation rate in human evolution and demographic inference. Curr. Opin. Genet. Dev. *41*, 36–43.
25. Cronin, M.A., Amstrup, S.C., Talbot, S.L., Sage, G.K., and Amstrup, K.S. (2009). Genetic variation, relatedness, and effective population size of polar bears (*Ursus maritimus*) in the southern Beaufort Sea, Alaska. J. Hered. *100*, 681–690.
26. De Barba, M., Waits, L.P., Garton, E.O., Genovesi, P., Randi, E., Mustoni, A., and Groff, C. (2010). The power of genetic monitoring for studying demography, ecology and genetics of a reintroduced brown bear population. Mol. Ecol. *19*, 3938–3951.
27. Eyre-Walker, A., and Keightley, P.D. (1999). High genomic deleterious mutation rates in hominids. Nature *397*, 344–347.
28. Nachman, M.W., and Crowell, S.L. (2000). Estimate of the mutation rate per nucleotide in humans. Genetics *156*, 297–304.
29. Allio, R., Donega, S., Galtier, N., and Nabholz, B. (2017). Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: Implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. Mol. Biol. Evol. *34*, 2762–2772.
30. Rieux, A., Eriksson, A., Li, M., Sobkowiak, B., Weinert, L.A., Warmuth, V., Ruiz-Linares, A., Manica, A., and Balloux, F. (2014). Improved calibration of the human mitochondrial clock using ancient genomes. Mol. Biol. Evol. *31*, 2780–2792.
31. Clark, P.U., Archer, D., Pollard, D., Blum, J.D., Rial, J.A., Brovkin, V., Mix, A.C., Pisias, N.G., and Roy, M. (2006). The Middle Pleistocene transition: characteristics, mechanisms, and implications for long-term changes in atmospheric pCO2. Quat. Sci. Rev. *25*, 3150–3184.
32. Gretzinger, J., Molak, M., Reiter, E., Pfrengle, S., Urban, C., Neukamm, J., Blant, M., Conard, N.J., Cupillard, C., Dimitrijević, V., et al. (2019). Large-scale mitogenomic analysis of the phylogeography of the Late Pleistocene cave bear. Sci. Rep. *9*, 10700.
33. Petit, R.J., and Excoffier, L. (2009). Gene flow and species delimitation. Trends Ecol. Evol. *24*, 386–393.
34. Konishi, M., and Takata, K. (2004). Impact of asymmetrical hybridization followed by sterile F 1 hybrids on species replacement in *Pseudorasbora*. Conserv. Genet. *5*, 463–474.
35. Gansauge, M.-T., and Meyer, M. (2013). Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nat. Protoc. *8*, 737–748.
36. Paijmans, J.L.A., Baleka, S., Henneberger, K., Taron, U.H., Trinks, A., Westbury, M.V., and Barlow, A. (2017). Sequencing single-stranded libraries on the Illumina NextSeq 500 platform. arXiv, 1–5.
37. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10.
38. Magoč, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics *27*, 2957–2963.
39. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.
40. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing

Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

41. Ginolhac, A., Rasmussen, M., Gilbert, M.T.P., Willerslev, E., and Orlando, L. (2011). mapDamage: testing for damage patterns in ancient DNA sequences. Bioinformatics *27*, 2153–2155.
42. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). Open Access ANGSD : Analysis of Next Generation Sequencing Data. BMC Bioinformatics *15*, 1–13.
43. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.
44. R Core Team (2014). R: A language and environment for statistical computing (R Found. Stat. Comput.), Available at. http://www.r-project.org/.
45. Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol. Biol. Evol. *35*, 1547–1549.
46. Lioubine, V.P., and Kulikov, O. (1991). About the age of the most ancient Paleolithic sites of the Caucasus. Sov. Archaeol. *4*, 4–6.
47. Lioubine, V.P. (2002). L'acheuléen du Caucase (Université de Liège, Service de Prehistoire).
48. Nesmeyanov, S.A. (1999). Geomorphological aspects of Paleolithic paleoecology of the Western Caucasus (Nauchniy Mir), [In Russian].
49. Baryshnikov, G.F. (2002). Local biochronology of Middle and Late Pleistocene mammals from the Caucasus. Russ. J. Theriology *1*, 61–67.
50. Guérin, C., and Baryshnikov, G.F. (1987). Le rhinocéros acheuléen de la grotte de Koudaro I(Géorgie, URSS) et le problème des espèces relictes du Pléistocène du Caucase. Geobios *20*, 389–396.
51. Barlow, A., Fortes, G.M.G., Dalen, L., Pinhasi, R., Gasparyan, B., Rabeder, G., Frischchauf, C., Paijmans, J.L.A., and Hofreiter, M. (2016). Massive influence of DNA isolation and library preparation approaches on palaeogenomic sequencing data. bioRxiv, 075911.
52. Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., Raczky, P., et al. (2015). Optimal ancient DNA yields from the inner ear part of the human petrous bone. PLoS ONE *10*, e0129102.
53. Basler, N., Xenikoudakis, G., Westbury, M.V., Song, L., Sheng, G., and Barlow, A. (2017). Reduction of the contaminant fraction of DNA obtained from an ancient giant panda bone. BMC Res. Notes *10*, 754.
54. Li, B., Zhang, G., Willerslev, E., and Wang, J. (2011). Genomic data from the Polar Bear (*Ursus maritimus*). Gigascience *157*, https://doi.org/10.5524/100008.
55. Hu, Y., Wu, Q., Ma, S., Ma, T., Shan, L., Wang, X., Nie, Y., Ning, Z., Yan, L., Xiu, Y., and Wei, F. (2017). Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. Proc. Natl. Acad. Sci. USA *114*, 1081–1086.
56. Krause, J., Unger, T., Noçon, A., Malaspinas, A.-S., Kolokotronis, S.-O., Stiller, M., Soibelzon, L., Spriggs, H., Dear, P.H., Briggs, A.W., et al. (2008). Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene-Pliocene boundary. BMC Evol. Biol. *8*, 220.
57. Abella, J., Alba, D.M., Robles, J.M., Valenciano, A., Rotgers, C., Carmona, R., Montoya, P., and Morales, J. (2012). *Kretzoiarctos* gen. nov., the oldest member of the giant panda clade. PLoS ONE *7*, e48985.
58. Sheng, G.-L., Basler, N., Ji, X.-P., Paijmans, J.L.A., Alberti, F., Preick, M., Hartmann, S., Westbury, M.V., Yuan, J.-X., Jablonski, N.G., et al. (2019). Paleogenome reveals genetic contribution of extinct giant panda to extant populations. Curr. Biol. *29*, 1695–1700.e6.
59. Günther, T., and Nettelblad, C. (2019). The presence and impact of reference bias on population genomic studies of prehistoric human populations. PLoS Genet. *15*, e1008302.
60. Fortes, G.G., Grandal-d'Anglade, A., Kolbe, B., Fernandes, D., Meleg, I.N., García-Vázquez, A., Pinto-Llona, A.C., Constantin, S., de Torres, T.J., Ortiz, J.E., et al. (2016). Ancient DNA reveals differences in behaviour and sociality between brown bears and extinct cave bears. Mol. Ecol. *25*, 4907–4918.
61. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.
62. Rustioni, M., and Mazza, P. (1992). The genus Ursus in Eurasia: Dispersal events and stratigraphical significance. Riv. Ital. di Paleontol. e Stratigr. *98*.
63. Rabeder, G., Pacher, M., and Withalm, G. (2010). Early Pleistocene bear remains from Deutsch-Altenburg (lower Austria). Geol. Carpath. *61*, 192.
64. Sala, B., and Masini, F. (2007). Late Pliocene and Pleistocene small mammal chronology in the Italian peninsula. Quat. Int. *160*, 4–16.
65. van Heteren, A.H., Arlegi, M., Santos, E., Arsuaga, J.L., and Gómez-Olivencia, A. (2019). Cranial and mandibular morphology of Middle Pleistocene cave bears (*Ursus deningeri*): implications for diet and evolution. Hist. Biol. *31*, 485–499.
66. Sheng, G.L., Barlow, A., Cooper, A., Hou, X.D., Ji, X.P., Jablonski, N.G., Zhong, B.J., Liu, H., Flynn, L.J., Yuan, J.X., et al. (2018). Ancient DNA from giant panda (*Ailuropoda melanoleuca*) of south-western China reveals genetic diversity loss during the Holocene. Genes (Basel) *9*, 9.
67. Loreille, O., Orlando, L., Patou-Mathis, M., Philippe, M., Taberlet, P., and Hänni, C. (2001). Ancient DNA analysis reveals divergence of the cave bear, *Ursus spelaeus*, and brown bear, *Ursus arctos*, lineages. Curr. Biol. *11*, 200–203.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Guanidine hydrochloride	Roth	Cat#0037.1
QIAGEN MinElute kit	QIAGEN	Cat#28004
Critical Commercial Assays		
D1000 Screen Tape (Tapestation2200)	Agilent	Cat#5067-5582
dsDNA HS Assay Kit (Qubit 2.0)	Thermofisher	Cat#Q32851
Deposited Data		
7t5-KU1_1 unprocessed data, fastq format	This paper	ENA: ERX 4890811
qgj-KU1_2 unprocessed data, fastq format	This paper	ENA: ERX 4890812
ucp-KU1_2 unprocessed data, fastq format	This paper	ENA: ERX 4890813
x54-KU1_2 unprocessed data, fastq format	This paper	ENA: ERX 4890814
4z6-KU1_3 unprocessed data, fastq format	This paper	ENA: ERX 4890815
85j-KU1_3 unprocessed data, fastq format	This paper	ENA: ERX 4890816
e5e-KU1_3 unprocessed data, fastq format	This paper	ENA: ERX 4890817
vup-KU1_3 unprocessed data, fastq format	This paper	ENA: ERX 4890818
2pq-KU1_4 unprocessed data, fastq format	This paper	ENA: ERX 4890819
4id-KU1_4 unprocessed data, fastq format	This paper	ENA: ERX 4890820
65v-KU1_4 unprocessed data, fastq format	This paper	ENA: ERX 4890821
6xw-KU1_4 unprocessed data, fastq format	This paper	ENA: ERX 4890822
9s1-KU1_4 unprocessed data, fastq format	This paper	ENA: ERX 4890823
j54-KU1_4 unprocessed data, fastq format	This paper	ENA: ERX 4890863
siw-KU1_4 unprocessed data, fastq format	This paper	ENA: ERX 4890824
009-KU1_5 unprocessed data, fastq format	This paper	ENA: ERX 4890825
6mn-KU1_5 unprocessed data, fastq format	This paper	ENA: ERX 4890826
9j7-KU1_5 unprocessed data, fastq format	This paper	ENA: ERX 4890827
b3b-KU1_5 unprocessed data, fastq format	This paper	ENA: ERX 4890828
4w5-KU1_6 unprocessed data, fastq format	This paper	ENA: ERX 4890829
6ux-KU1_6 unprocessed data, fastq format	This paper	ENA: ERX 4890864
8ux-KU1_6 unprocessed data, fastq format	This paper	ENA: ERX 4890830
dLd-KU1_6 unprocessed data, fastq format	This paper	ENA: ERX 4890831
mhh-KU1_6 unprocessed data, fastq format	This paper	ENA: ERX 4890832
n3p-KU1_6 unprocessed data, fastq format	This paper	ENA: ERX 4890833
we6-KU1_6 unprocessed data, fastq format	This paper	ENA: ERX 4890834
pk5-HV72_1 unprocessed data, fastq format	This paper	ENA: ERX 4890865
nxu-HV72_2 unprocessed data, fastq format	This paper	ENA: ERX 4890866
zfb-HV72_3 unprocessed data, fastq format	This paper	ENA: ERX 4890867
et7-HV72_4 unprocessed data, fastq format	This paper	ENA: ERX 4890868
1oa-HV72_5 unprocessed data, fastq format	This paper	ENA: ERX 4890869
0gw-HV72_6 unprocessed data, fastq format	This paper	ENA: ERX 4890870
xyw-HV72_7 unprocessed data, fastq format	This paper	ENA: ERX 4890871
ntw-HV72_8 unprocessed data, fastq format	This paper	ENA: ERX 4890810
w8o-HV72_8 unprocessed data, fastq format	This paper	ENA: ERX 4890872
2Ls-B05_1 unprocessed data, fastq format	This paper	ENA: ERX 4890807
vhf-B05_1 unprocessed data, fastq format	This paper	ENA: ERX 4890808

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
y3h-B05_1 unprocessed data, fastq format	This paper	ENA: ERX 4890809
9e1-B04_1 unprocessed data, fastq format	This paper	ENA: ERX 4890804
f9L-B04_1 unprocessed data, fastq format	This paper	ENA: ERX 4890806
vb5-B04_1 unprocessed data, fastq format	This paper	ENA: ERX 4890805
KU1 mapped reads, polar bear reference, bam format	This paper	ENA: ERX 4890915
KU1 mapped reads, panda reference, bam format	This paper	ENA: ERX 4890914
B05 mapped reads, polar bear reference, bam format	This paper	ENA: ERX 4890921
B05 mapped reads, panda reference, bam format	This paper	ENA: ERX 4890920
B04 mapped reads, polar bear reference, bam format	This paper	ENA: ERX 4890919
B04 mapped reads, panda reference, bam format	This paper	ENA: ERX 4890918
HV72 mapped reads, polar bear reference, bam format	This paper	ENA: ERX 4890917
HV72 mapped reads, panda reference, bam format	This paper	ENA: ERX 4890916
KU1 mitochondrial genome sequence	This paper	GenBank: MW491935
B05 mitochondrial genome sequence	This paper	GenBank: MW491933
B04 mitochondrial genome sequence	This paper	GenBank: MW491932
HV72 mitochondrial genome sequence	This paper	GenBank: MW491934
Oligonucleotides		
CL9 extension primer: GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	[35]	Sigma Aldrich
Double-stranded adaptor Strand 1 (CL53): CGACGCTCTTC-ddC (ddC = dideoxycytidine) Strand 2 (CL73): [Phosphate]GGAAGAGCGTCGTGTAGGGAAA GAG*T*G*T*A (* = phosphothioate linkage)	[35]	Sigma Aldrich
CL78: AGATCGGAAG[C3Spacer] $_{10}$ [TEG-biotin] (TEG = triethylene glycol spacer)	[35]	Sigma Aldrich
P5 indexing primer: AATGATACGGCGACCACCGAGATCTA CACnnnnnnnnACACTCTTTCCCTACACGACGCTCTT	[35]	Sigma Aldrich
P7 indexing primer: CAAGCAGAAGACGGCATACGAGATnnnn nnnnGTGACTGGAGTTCAGACGTGT	[35]	Sigma Aldrich
IS7 amplification primer: ACACTCTTTCCCTACACGAC	[35]	Sigma Aldrich
IS8 amplification primer: GTGACTGGAGTTCAGACGTGT	[35]	Sigma Aldrich
CL72 R1 sequencing primer: ACACTCTTTCCCTACACGAC GCTCTTCC	[35]	Sigma Aldrich
Gesaffelstein index 2 sequencing primer: GGAAGAGCGTCG TGTAGGGAAAGAGTGT	[36]	Sigma Aldrich
Software and Algorithms		
BEARCAVE ce78f40	N/A	https://github.com/nikolasbasler/BEARCAVE/
Cutadapt v1.12	[37]	https://cutadapt.readthedocs.io/en/stable/
Flash v1.2.11	[38]	https://ccb.jhu.edu/software/FLASH/
BWA v0.7.15 and v0.7.8	[39]	http://bio-bwa.sourceforge.net/
Samtools v1.3.1	[40]	https://sourceforge.net/projects/samtools/files/samtools/
PreSeq	N/A	http://smithlabresearch.org/software/preseq/
MapDamage v2.0.8	[41]	https://ginolhac.github.io/mapDamage/
ANGSD v0.916	[42]	http://www.popgen.dk/angsd
Consensify v0.1	[11]	https://github.com/jlapaijmans/Consensify
ReDuCToR v0.1	[11]	https://github.com/jlapaijmans/Consensify
RaxML v8.2.12	[43]	https://github.com/stamatak/standard-RAxML
R version 3.6.3	[44]	https://www.r-project.org/

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
MarkReadsByStartEnd.jar	N/A	https://github.com/dariober/Java-cafe/tree/master/MarkDupsByStartEnd
MEGA X v10.1.7	[45]	https://www.megasoftware.net/
D_stat.cpp	[10]	https://github.com/jacahill/Admixture
D-stat_parser.py	[10]	https://github.com/jacahill/Admixture
weighted_block_jackknife.py	[10]	https://github.com/jacahill/Admixture
Other		
Proteinase K	Promega	Cat#V3021
Zymo-spin V column extension reservoir	Zymo	Cat#C1016-50
Circligase II	Biozym	Cat#131402(CL9021K)
Endonuclease VIII	NEB	Cat#A0299S
Uracil-DNA glycosylase (Afu UDG)	NEB	Cat#M0279S
FastAP	Thermo Fisher	Cat#EF0651
MyOne C1 streptavidin beads	Thermo Fisher	Cat#65001
Bst 2.0 polymerase	NEB	Cat#M0537S
T4 DNA Polymerase	Thermo Fisher	Cat#EP0061
Buffer Tango (10x)	Thermo Fisher	Cat#BY5
T4 DNA ligase	Thermo Fisher	Cat#EL0011
Accuprime Pfx	Thermo Fisher	Cat#12344024
PEG-4000	Thermo Fisher	Cat#EP0061
Klenow fragment of DNA polymerase I	Thermo Fisher	Cat#EP0051
SYBR green PCR MasterMix	Thermo Fisher	Cat#4309155

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Axel Barlow (axel.barlow.ab@gmail.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The accession number for the raw, unprocessed sequencing reads in fastq file format, and the processed data mapped to each reference genome in bam file format, reported in this paper is European Nucleotide Archive (ENA) study accession PRJEB42442. The accession numbers for the novel cave bear consensus mitochondrial sequences reported in this paper are NCBI nucleotide database accessions MW491932-MW491935.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The novel cave bear datasets generated for this study are: *praekudarensis* (KU1), *kudarensis* (HV72), *rossicus* (B05), and *kanivetz* (B04). These were generated from subfossil petrous bone samples. Details of sample localities and ages are shown in Table S2.

Several sources of evidence support a Middle Pleistocene age of layer 5c in Kudaro 1 Cave, from which the *praekudarensis* petrous bone was recovered: layer 5c has been dated using radiothermoluminescence at 360,000 ± 90,000 yBp, and the overlaying layer 5b has been dated using the same method at 350,000 ± 70,000 yBp;[12,46,47] geomorphological studies of river terraces of the Caucasus[48] suggest the cave entrance of Kudaro 1 was opened 300–400 thousand years ago; archaeological material recovered from layer 5c belongs to the later Acheulean;[47] and other faunal material recovered from layer 5c is consistent with a Middle Pleistocene age,[49] including *Macaca sp., Canis mosbachensis, Panthera gombaszoegensis* and *Stephanorhinus hundsheimensis* (the rhinoceros was identified by[50] as *Dicerorhinus etruscus brachycephalus*). According to European markers, this faunal composition would suggest an even older age than indicated by the radiothermoluminescence dates. However, the Caucasus region is considered to represent a refugial area during the Pleistocene, which may explain the apparently more recent occurrence of these taxa in Kudaro 1 compared to Middle Pleistocene European deposits.

The *rossicus* and *kanivetz* samples have not been directly dated. For the purpose of molecular dating analyses, we used an indirectly estimated age for these samples based on the median age of radiocarbon dated cave bear bones from these sites (Table S6).

METHOD DETAILS

Laboratory methods

Six sequencing libraries (KU1_1–KU1_6) were prepared from the *praekudarensis* KU1 sample for this study. For the *kudarensis* sample HV72 we had previously prepared seven sequencing libraries (HV72_1–HV72_7) and carried out low level sequencing. Six of these libraries (HV72_2 – HV72_7) are described in.[51] We prepared one additional library (HV72_8) from this sample for this study, which was used for deeper sequencing. Single libraries were prepared from the *rossicus* B05 and *kanivetz* B04 samples.

For each library, bone was sampled from the otic capsule of the cave bear petrous bones[52] and ground to a fine powder using a RETSCH mixer mill mm 400 at a frequency of 30 Hz for 10 s. DNA was then extracted using a published protocol optimized for the recovery of short ancient DNA fragments,[6] with the modifications described in.[53] 50 mg of bone powder was digested in 1 mL of extraction buffer (0.45 M EDTA, 0.25 mg/mL Proteinase K) overnight at 37°C, with rotation. Centrifugation was used to pellet any undigested material. The supernatant was removed, combined with 13 mL of binding buffer (5 M guanidine hydrochloride, 40% (vol/vol) isopropanol, 0.05% Tween-20, and 90 mM sodium acetate), and then passed through a commercial silica spin column (QIAGEN MinElute) fitted with an extension reservoir (Zymo- spin V). Two wash steps were then carried out using PE buffer (QIAGEN). Purified DNA was then eluted in two steps each using 12.5 μL TET buffer (10mM Tris-HCl, 1 mM EDTA, 0.05% Tween-20).

Illumina sequencing libraries were prepared from the DNA extracts using a published protocol based on single stranded DNA, optimized for the recovery of short ancient DNA fragments,[35] with the modifications described in.[53] DNA was treated with the enzymes uracil-DNA glycosylase and Endonuclease VIII, to excise uracils resulting from the deamination of cytosine residues and to cleave DNA strands at abasic sites, in 44 μL reactions with the following reagent concentrations: 1.8x CircLigase buffer II, 4.5 mM MnCl2, 0.02 U/μL of uracil-DNA glycosylase, and 0.11 U/μL of endonuclease VIII. Residual phosphate groups were then removed from the template DNA fragment ends using 1 unit of FastAP. The DNA was then heat denatured, and oligo CL78 ligated to the 3' single-stranded fragment ends during overnight incubation in 80 μL reactions with the following reagent concentrations: 20% (vol/vol) PEG-4000, 0.125 μM CL78, and 2.5 units/μL Circligase II. Ligation products were then immobilised on streptavidin beads (MyOne C1) to allow the removal of reagent mixtures in subsequent library preparation steps. The CL9 extension primer was annealed to the complementary CL78 oligo sequence and the strand complementary to the template single-stranded molecules filled-in using Bst 2.0 polymerase in 50 μL reactions with the following reagent concentrations: 1x isothermal amplification buffer, 250 μM of each dNTP, 2 μM CL9 extension primer, and 0.48 U/μL Bst 2.0 polymerase. T4 DNA polymerase was then used to remove 3' overhangs, in 100 μL reactions with the following reagent concentrations: 1x Buffer Tango, 0.025% (vol/vol) Tween 20, 100 μM of each dNTP, and 0.05 U/μL T4 DNA polymerase. The double-stranded adaptor (CL53/CL73) was then ligated to the blunt-ended molecules using T4 DNA ligase in 100 μL reactions with the following reagent concentrations: 1x T4 DNA ligase buffer, 5% (vol/vol) PEG-4000, 0.025% (vol/vol) Tween 20, 2 μM double-stranded adaptor, and 0.1 U/μL T4 DNA ligase. The library strand complementary to the original single-stranded template molecule was then heat denatured and eluted in 25 μL TET buffer.

Template library molecules were PCR amplified using AccuPrime Pfx polymerase, incorporating unique 8 base-pair (bp) index sequences within both P5 and P7 adapters, in 80 μL reactions with the following reagent concentrations: 1x AccuPrime Pfx reaction mix, 10 μM each of P5 and P7 indexing primers, and 0.025 U/μL AccuPrime Pfx polymerase. Prior to library amplification, qPCR analysis of the unamplified library was used to identify the appropriate number of PCR cycles, corresponding to the cycle number at the point of inflection of the qPCR amplification curve, corrected for differing reaction volume and template amount in the subsequent library amplification PCR. The qPCR analysis involved 10 μL reactions with the following reagent concentrations: 1x SYBR green qPCR master mix, 0.2 μM each of IS7 and IS8 amplification primers, and 0.2% of the unamplified library. After amplification, the indexed libraries were quantified using a TapeStation 2200 instrument (Agilent) with D1000 screen tape and reagents, and a Qubit 2.0 instrument (Fisher) with the dsDNA HS Assay kit. Sequencing of the libraries was mostly performed on an Illumina NextSeq 500 sequencing platform, using the custom CL72 R1 sequencing primer[35] and the Gesaffelstein custom index 2 sequencing primer,[36] following the procedures described in.[36] Some *praekudarensis* libraries were additionally sequenced on an Illumina HiSeq 2500 sequencing platform using the same custom primers (see Table S1).

Data processing

Processing of the sequence reads was carried out within the BEARCAVE v.ce78f40 data analysis and storage environment (available at: https://github.com/nikolasbasler/BEARCAVE) and is reported in Table S1. BEARCAVE is freely available and can be used to obtain details of all software versions and parameter settings, as well as to replicate the described analyses. Data processing involved trimming adaptor sequences and removing reads < 30 bp using CutAdapt,[37] and merging overlapping paired-end reads using FLASH.[38] The specific BEARCAVE scripts used for these steps were: "*trim_merge_DS_PE_standard.sh*" for trimming and merging paired-end data generated from double stranded libraries (modern and some published ancient datasets); "*trim_merge_SS_PE_CL72.sh*" for trimming and merging paired-end data generated from single stranded libraries (ancient datasets); and "*trim_SE.sh*" for trimming single-end data. Reads were then mapped to the reference genome assemblies of the polar bear[54] and the giant panda[55] using the bwa[39] aln algorithm and samtools,[40] filtering for mapping quality (-q 30) and potential PCR duplicates (rmdup). For the *praekudarensis* data, duplicate removal was carried out separately for each library prior to merging them into a single

genomic dataset. The specific BEARCAVE scripts used for mapping datasets to the polar bear reference were: "*map_SE.sh*" for mapping ancient datasets with default bwa parameters, excluding unmerged read-pairs from paired-end datasets since they likely represent modern contamination, rendering these datasets effectively single-end and "*map_modern_PE.sh*" for mapping modern datasets with default bwa parameters including unmerged read pairs. The giant panda lineage is comparatively diverged from the investigated clade (around 12–19 Ma[56,57]), requiring relaxation of the number of allowed mismatches between read and reference (-n 0.01, implemented using the corresponding BEARCAVE scripts) in order to achieve acceptable mapping performance.[10] Nonetheless, we were consistently able to map more data to the polar bear than to the panda reference, reflecting its lower divergence from the investigated samples. For this reason, subsequent analyses investigating the broader scale patterns of divergence among the sampled genomes utilized the polar bear mapping reference in order to maximize the number of sampled genome positions. For analyses investigating patterns of admixture, the panda was used as mapping reference since these analyses may be biased by using the polar bear mapping reference, which represents an ingroup to the investigated clade.[58,59]

QUANTIFICATION AND STATISTICAL ANALYSIS

Assessment of ancient DNA authenticity

We assessed the authenticity of the cave bear datasets by estimating the endogenous fragment length distribution and extent of cytosine deamination for 10 million randomly sampled reads mapping successfully to the polar bear reference, using the program mapDamage v2.08[41] with Bayesian statistical estimation disabled and the merge reference sequences option enabled. All cave bear datasets showed evidence of DNA fragmentation and cytosine deamination consistent with the sample ages (Figure S1).

Nuclear genome PCA

We investigated the broad scale patterns of divergence among the seven sampled cave bear taxa, using principal components analysis (PCA) of a single representative genome of each taxon. The polar bear was used as mapping reference for this analysis. A covariance matrix was calculated by sampling a single nucleotide at random from the read stack at each position of the reference genome using single base identity by state (IBS) in ANGSD v0.916,[42] only considering reads with a minimum mapping quality score of 30 (-minMapQ 30) and nucleotides with a minimum base quality score of 30 (-minQ 30). We further only considered sites from scaffolds > 1 Mb in length, with no missing data (-minInd N, where N = number of individuals), and which were below the upper 95th percentile of global coverage (-setMaxDepth, determined in advance using the -doDepth function in angsd). Transition (identified using genotype likelihoods) and singleton (1/N < -minFreq < 2/N) sites were also excluded. PCA of the covariance matrix was then carried out using the ''eigen'' function in R.[44] The exclusion of singleton sites in this analysis is an effective way of reducing sequencing errors, which frequently occur at high abundance in ancient datasets. However, since private alleles are also removed, this approach is sensitive to unbalanced sampling of clades, with the tendency to underestimate divergence for undersampled lineages.[11] This effect was observed in preliminary analyses including both *kudarensis* individuals, since all other cave bear taxa were represented by single individuals. The sampling of the *kudarensis* lineage was therefore reduced to the individual with the higher coverage (HV74) in order to achieve a less biased assessment of cave bear relationships. The ordination of individuals along PC1 and PC2 of this analysis is shown in Figure 1A, and along PC3 to PC6 in Figure 1B.

Generation of error-reduced genome sequences

We prepared two error-reduced sets of genome sequences using Consensify: one for datasets mapped to the polar bear reference, which was used to estimate phylogeny and genetic distances; and one for datasets mapped to the panda reference for admixture tests using D-statistics.

For the first set of Consensify sequences, for each dataset we used angsd to count the observed frequency of mapped bases at each position of the polar bear reference (-doCounts function), filtering for mapping (-minMapQ 30) and base calling (-minQ 30) qualities, and only considering scaffold > 1 Mb in length. In order to achieve the most accurate estimates of genetic distances, we additionally excluded the terminal nucleotide of each mapped sequence to further reduce the probability of introducing errors resulting from cytosine deamination. The Consensify error-reduced sequence for each dataset was then generated from these base counts using the Consensify script (available from https://github.com/jlapaijmans/Consensify), applying a maximum depth filter of the integer < 95% depth for each individual dataset, calculated in advance using the -doDepth function in angsd. Generation of the second set of sequences used the same methodology applied to datasets mapped to the panda reference, except that the terminal nucleotides of the mapped sequences were included in base counts. The Consensify error reduced sequences are summarized in Table S1.

Phylogenetic analyses

We estimated phylogenetic relationships among the 12 sampled nuclear genomes of cave bears, brown bears and polar bears, using the Consensify error-reduced sequences generated from datasets mapped to the polar bear reference. We used the ReDuCToR script (included in the Consensify distribution) to combine the Consensify sequences into a single alignment, removing all invariant columns and any containing missing data. Maximum-likelihood phylogenetic analysis was then carried out using RaxML v8.2.12[43] under the GTR+GAMMA substitution model with 100 rapid bootstrap replicates and a thorough maximum likelihood search for the

final tree ("-f a" option), which was rooted using the Asiatic black bear outgroup. Bootstrap values > 80% were considered as statistically supported. The resulting phylogeny is shown in Figure 1D.

To estimate mitochondrial relationships, the *praekudarensis* mitochondrial genome sequence was generated from the datasets described in Table S1, except "6ux" and "j54," which were sequenced at a later date. Adaptor trimming was performed as described above (data processing section), except that reads < 28 bp were discarded. Subsequent manual inspection of the mapped reads indicated that this lower minimum read length threshold was appropriate for reconstruction of the *praekudarensis* mitochondrial genome. The reads were mapped to the published reference mitochondrial sequence of the *kudarensis* cave bear HV74[10] using the bwa aln algorithm, discarding reads with MapQuality score < 30 with samtools v1.3.1, and removing duplicate reads using MarkReadsByStartEnd.jar (https://github.com/dariober/Java-cafe/tree/master/MarkDupsByStartEnd). A consensus sequence was then generated from this alignment in Geneious v7.0, using a minimum sequence depth of 3x and a 75% majority rule for base calling. The consensus sequence was manually checked against the original alignment to exclude the possibility of erroneous or incorrect consensus base calls. Details of the mitochondrial genome reconstruction are shown in Table S7.

Mitochondrial genome sequences were also generated for the two brown bears (Ge, Uap), the two polar bears (SRS412584, SRS412585), the Asiatic black bear (ERS781634), and four of the Late Pleistocene cave bears (HV72, BO4, BO5, WK01), from the datasets described in Table S1. The mitochondrial genome sequences of the brown bear Uap and *eremus* cave bear WK01 have been published previously, based on much lower coverage.[60] We therefore recomputed them using the higher coverage datasets of[10] in order to achieve a more complete sequence. Mitochondrial reconstruction followed the methodology described above for *praekudarensis*, except that reads < 30 bp were discarded, reads from each dataset were mapped to a reference mitochondrial sequence selected as a close relative of the respective taxon, and consensus sequences were generated using a minimum sequence depth of 3x and a 90% majority rule for base calling. Details of the mitochondrial genome reconstruction are shown in Table S7.

The consensus sequences were aligned with published sequences of the *ingressus* (GS136), *spelaeus* (E-VD-1838), *eremus* (WK01) and *kudarensis* (HV74) individuals using the MUSCLE algorithm[61] implemented in MEGA X[45] with default parameters. The *Ursus* control region contains a microsatellite repeat which was removed as this cannot be reliably recovered using short read data. Maximum likelihood phylogenetic analysis was carried out as described above for the nuclear genome alignment. The resulting phylogeny is shown in Figure 1C.

Molecular dating

Genomic data from individuals sampled at different time points provides information on their genome-wide substitution rate. To estimate this rate for the cave, polar and brown bear clade, we compared the genomic divergences of the *praekudarensis* and *kudarensis* datasets to modern polar bears and brown bears, which are expected to be lower in the case of *praekudarensis* since its divergence time from the clade's common ancestor is considerably less. Thus, assuming a strict molecular clock, the difference in divergence divided by the median estimate of 305,400 years separating *praekudarensis* and *kudarensis* provides an estimate of the per-lineage substitution rate. The ReDuCToR alignment of Consensify sequences generated from datasets mapped to the polar bear reference was recomputed to include invariant positions. Pairwise genomic distances were then calculated from this alignment under the JC69 substitution model using the dist.dna function in the R package "ape," considering both transitions and transversions. Assuming the sites sampled using Consensify are a random and unbiased sample of the genome, these distances equate to whole genome divergences and can be used to estimate the genome-wide substitution rate. Six sets of substitution rate estimates were calculated using all combinations of the two *kudarensis* individuals, both polar bears and the modern Georgian brown bear (Table S3). The six estimates were highly consistent and their mean (9.56231×10^{-10} substitutions/site/year) was used for subsequent divergence time estimations. The consistency of the rate estimates supports both the validity of our method, and the assumption of a strict molecular clock. We additionally estimated the genome-wide substitution rate assuming ages of the *praekudarensis* sample ± 90 ka, representing the uncertainty in its radiothermoluminescence date. We also repeated this entire set of calculations using pairwise mitochondrial distances to estimate the mitochondrial substitution rate.

We applied substitution rate estimates to calculate nuclear (Table S4) and mitochondrial (S5) divergence times from pairwise genetic distances among all individuals. Since the estimates reflect the per-lineage substitution rate, they were multiplied by two to obtain the rate of genetic divergence between sister lineages, assuming a strict molecular clock. Genetic divergences between individuals were then divided by this estimated rate of genetic divergence to obtain the divergence time in years (Tables S4 and S5). The resulting pairwise divergence times represent the total time taken for lineages to achieve the observed genetic divergence, and do not take into account the non-contemporaneous ages of the individuals. The median age of each pair of individuals was therefore added to their respective pairwise divergence time, which, assuming a strict molecular clock, provides their absolute time of divergence before the present day (Tables S4 and S5). This procedure resulted in multiple absolute age estimates for most nodes of the phylogeny, with each estimate based on a different combination of individuals. As for the substitution rate estimates, the consistency of these absolute node age estimates supports both the validity of our method, and the assumption of a strict molecular clock. To provide the calibrated trees in Figure 3, nodes were centered on the mean of their calculated median age estimates.

Comparison with other divergence estimates

Based on a median radiothermoluminescence age of 360 ka for the *praekudarensis* sample, our nuclear genome analysis provided an estimated divergence time of cave bears and their sister clade, brown and polar bears, of 1.52 Ma. The estimate provided by mitochondrial DNA is remarkably similar, at 1.48 Ma. These estimates coincide with the last documented fossil occurrences (~1.6 Ma) of

their accepted common ancestor, *Ursus etruscus*.[62–64] They also moderately pre-date the earliest documented fossil occurrences of the accepted ancestral cave bear, *Ursus deningeri*, toward the end of the Early Pleistocene (a review of the literature is provided in[65]). Finally, they moderately pre-date the earliest documented fossil showing *arctos*-like characteristics, which have been assigned to the brown bear lineage. These fossils also date toward the end of the Early Pleistocene, around 1.2 Ma.[63]Our estimated divergence time is also considerably younger than a previous estimate based on complete mitochondrial genomes,[56] which reported a divergence estimate of 2.75 Ma (95% credibility interval (CI) 2.1–3.57 Ma) based on fossil calibration of the seal/bear divergence and of the *Ursus* lineage. A second study analyzing ~4kb of mitochondrial sequence[66] utilized four calibration points within the bear clade. Although the divergence time of cave bears from their sister clade was not reported in this study, as first author AB carried out this analysis we can confirm the estimate was 2.20 Ma (95% CI 1.58–3.00 Ma). In contrast, an early study[67] based on control region and cytochrome b sequences produced a more recent estimate than ours, around 1.2 Ma.

Tests of nuclear admixture

D-statistics were calculated from the Consensify error-reduced sequences generated from datasets mapped to the panda reference using the published C++ program D_stat.cpp,[10] and the results processed using the python scripts D-stat_parser.py and weighted_block_jackknife.py (available from https://github.com/jacahill/Admixture). Significance of the D-statistics was assessed by calculating the standard error using a weighted block jackknife analysis using 5 Mb genome windows, with D values deviating more than three standard-errors from zero (absolute Z-score > 3) considered as statistically significant. These tests used the Asiatic black bear as outgroup for allele polarization, which has previously been shown to be a suitable outgroup taxon for testing for admixture within the brown-polar-cave bear clade.[10] The two brown bears included in this study have previously been shown to exhibit high genomic proportions of admixture with cave bears,[10] and were chosen to maximize sensitivity in detecting the specific admixing cave bear lineage(s).

We calculated D-statistics for all possible combinations of individuals congruent with their nuclear phylogeny (Figure 1D). We found significant evidence of differential admixture between brown bears and cave bears subsequent to the divergence of brown bears and polar bears; between brown bears and European+Uralian cave bears subsequent to their divergence from the Caucasian cave bears; and between European cave bears and the *kudarensis* lineage, subsequent to their divergence from their respective sister clades, the Uralian cave bear *rossicus* and the *praekudarensis* lineage (Figure 4). All other comparisons were not significantly different from zero.