nature ecology & evolution

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41559-018-0654-8

In the format provided by the authors and unedited.

Partial genomic survival of cave bears in living brown bears

Axel Barlow[1,16]*, James A. Cahill[2,16], Stefanie Hartmann[1], Christoph Theunert[3,4], Georgios Xenikoudakis[1], Gloria G. Fortes[1,5], Johanna L. A. Paijmans[1], Gernot Rabeder[6], Christine Frischauf[6], Aurora Grandal-d'Anglade[7], Ana García-Vázquez[7], Marine Murtskhvaladze[8], Urmas Saarma[9], Peeter Anijalg[9], Tomaž Skrbinšek[10], Giorgio Bertorelle[5], Boris Gasparian[11], Guy Bar-Oz[12], Ron Pinhasi[13,14], Montgomery Slatkin[3], Love Dalén[15], Beth Shapiro[2] and Michael Hofreiter[1]

[1]Institute for Biochemistry and Biology, University of Potsdam, Potsdam, Germany. [2]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Santa Cruz, CA, USA. [3]Department of Integrative Biology, University of California, Berkeley, Berkeley, CA, USA. [4]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. [5]Department of Life Sciences and Biotechnology, University of Ferrara, Ferrara, Italy. [6]Institute of Palaeontology, University of Vienna, Vienna, Austria. [7]Instituto Universitario de Xeoloxía, Universidade da Coruña, A Coruña, Spain. [8]Ilia State University, Institute of Ecology, Tbilisi, Georgia. [9]Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia. [10]Department of Biology, Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia. [11]Institute of Archaeology and Ethnography, National Academy of Sciences of the Republic of Armenia, Yerevan, Armenia. [12]Zinman Institute of Archaeology, University of Haifa, Haifa, Israel. [13]Earth Institute, University College Dublin, Dublin, Ireland. [14]Department of Evolutionary Anthropology, University of Vienna, Vienna, Austria. [15]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden. [16]These authors contributed equally: Axel Barlow, James A. Cahill. *e-mail: axel.barlow.ab@gmail.com

Barlow et al. (2018) Partial genomic survival of cave bears in living brown bears

SUPPLEMENTARY TABLES AND FIGURES

Supplementary Table 1. Details of samples used in this study

Common name	Taxon	Sample code	Locality	Sex	Reference	Accession
Cave bear	*eremus*	WK01	Windischkopf, Austria	untested	This study	*
Cave bear	*ingressus*	GS136	Gamssulzen Cave, Austria	untested	This study	*
Cave bear	*spelaeus*	E-VD-1838	Eiros Cave, Spain	untested	This study	*
Cave bear	*kudarensis*	HV74	Hovk Cave, Armenia	untested	This study	*
Brown bear	*U. arctos*	Uap	Winden Cave, Austria	untested	This study	*
Brown bear	*U. arctos*	Ge	Georgia, Great Caucasus	untested	This study	*
Brown bear	*U. arctos*	191Y	Hrušica, Slovenia	untested	This study	*
Brown bear	*U. arctos*	235	Balakhtinsky District, Russia	untested	This study	*
Brown bear	*U. arctos*	Swe	Sweden	Female	[1]	SRX796442
Brown bear	*U. arctos*	LS039	Spain	Male	[2]	SRR5878347
Brown bear	*U. arctos*	Adm1	Alaska (Admiralty)	Female	[3]	SRX265457
Brown bear	*U. arctos*	Den	Alaska (Denali)	Female	[3]	SRX265456
Polar bear	*U. maritimus*	NB	North Beaufort Sea	Male	[3]	SRX265452
Polar bear	*U. maritimus*	SB	South Beaufort Sea	Male	[3]	SRX265435
Polar bear	*U. maritimus*	WH2	West Hudson Bay	Male	[3]	SRX265434
American black bear	*U. americanus*	Uam	Pennsylvania, USA	Female	[3]	SRX265459
Asiatic black bear	*U. thibetanus*	ERS781634	Zoo Madrid	Female	[4]	ERS781634
Spectacled bear	*T. ornatus*	ERR946788	Zoo Basel	Male	[4]	ERR946788

* European Nucleotide Archive run accessions ERR2678614–ERR2678640

Supplementary Table 2. Ancient sample radio-carbon ages and estimated endogenous contents

Common name	Taxon	Sample code	$_{14}$C age	Calibrated age BP	Lab No. MAMS	Endo %
Cave bear	*eremus*	WK01	>49000[5]	-	23137	66.1
Cave bear	*ingressus*	GS136	31026±500[5]	35062±966	23139	63.7
Cave bear	*spelaeus*	E-VD-1838	30737±500[5]	34806±931	23140	74.4
Cave bear	*kudarensis*	HV74	>49000[6]	-	23142	66.0
Brown bear	*U. arctos*	Uap	36680±500[5]	41201±895	23141	65.7

Endogenous % is an estimate calculated by dividing the proportion of reads mapping to the polar bear reference genome by the proportion mapping from a high quality modern brown bear sample (Ge), which we determined to be approximately 87.9%. This procedure provides a degree of correction for false negatives due to an incomplete or misassembled reference genome.

Supplementary Table 3. Mapping results. Outcome of mapping ancient DNA sequences to the reference genome assembly of the giant panda.

sample	unique mapped reads[A]	% uniquely mapping reads	Mean read depth[B]	Total mapped GB
WK01	148679922	44.6	3.49	6.26
GS136	69053691	23.5	3.55	3.82
E-VD-1838	116172796	50.5	2.85	4.69
HV74	81320041	43.7	2.39	3.82
Uap	61074228	43.6	1.93	2.71

A *unique mapped reads*: is the number of mapped reads with mapping quality Q30 or above, after removal of potential PCR duplicates using samtools[7] rmdup

B *Mean read depth*: is mean read depth per covered position of the reference

Supplementary Table 4. Evaluation of phylogenetic test of directional admixture using a well characterised system: gene flow from polar bears into brown bears. Shown are total counts of blocks returning all 15 possible rooted tree topologies for 2 different block sizes. The four taxa used for this investigation were: NB (Polar1), WH (Polar2), 191Y (EuroBrown), Adm1 (ABCbrown). Trees were rooted using the American black bear outgroup. The distribution of tree topologies is consistent with increased gene flow from polar bears into the ABC islands brown bears relative to European brown bears, and a lack of gene flow from brown bears into polar bears following the basal divergence of these polar bear populations.

Topology	Description	Tree	0.25MB	0.1MB	Prop*
1	species tree	((Polar1,Polar2)(EuroBrown,ABCbrown))	3898	8692	0.6599
2	symmetrical	((EuroBrown,Polar2)(Polar1,ABCbrown))	0	2	0.0002
3	symmetrical	((ABCbrown,Polar2)(Polar1,EuroBrown))	0	1	0.0001
4	Class 1	(((Polar1,Polar2),ABCbrown),EuroBrown)	855	2751	0.2089
5	Class 1	(((ABCbrown,Polar2),Polar1),EuroBrown)	8	49	0.0037
6	Class 1	(((Polar1,ABCbrown),Polar2),EuroBrown)	0	41	0.0031
7	Class 2	(((Polar1,Polar2),EuroBrown),ABCbrown)	354	1590	0.1207
8	Class 2	(((EuroBrown,Polar1),Polar2),ABCbrown)	0	5	0.0004
9	Class 2	(((EuroBrown,Polar2),Polar1),ABCbrown)	1	4	0.0003
10	Class 3	(((EuroBrown,ABCbrown),Polar2),Polar1)	2	20	0.0015
11	Class 3	(((EuroBrown,Polar2),ABCbrown),Polar1)	0	1	0.0001
12	Class 3	(((ABCbrown,Polar2),EuroBrown),Polar1)	0	2	0.0002
13	Class 4	(((EuroBrown,ABCbrown),Polar1),Polar2)	1	11	0.0008
14	Class 4	(((EuroBrown,Polar1),ABCbrown),Polar2)	0	0	0.0000
15	Class 4	(((Polar1,ABCbrown),EuroBrown),Polar2)	4	3	0.0002
	Totals		5123	13172	1

* “Prop” indicates proportion of blocks returning this topology at 0.1MB block size.

Supplementary Table 5. Observed length frequencies of topology Classes in brown bears, as measured by the number of contiguous 25Kbp blocks returning the same topology. Note that, although topologies are grouped here by topology Class, counts represent continuous genomic regions returning a single tree topology.

Sample	**Topology Class**	**25**	**50**	**75**	**100**	**125**	**150**	**175**
191Y	Class 1	6014	806	152	26	8	2	-
	Class 2	4931	620	89	14	3	2	-
	Class 3	3165	199	24	3	-	-	-
	Class 4	2862	173	22	1	1	-	-
235	Class 1	5717	808	159	26	6	3	-
	Class 2	4922	619	96	15	4	2	-
	Class 3	3171	226	21	1	1	-	-
	Class 4	2922	188	11	2	1	-	-
Adm1	Class 1	5304	702	116	18	6	1	-
	Class 2	4582	516	87	13	1	1	-
	Class 3	3203	247	27	5	1	-	-
	Class 4	2917	199	24	3	1	-	1
Ge	Class 1	6240	838	162	28	11	2	-
	Class 2	4956	610	102	15	2	2	-
	Class 3	3240	207	23	-	-	-	-
	Class 4	2961	186	10	2	1	-	-
Den	Class 1	5566	738	141	22	4	1	1
	Class 2	4691	570	88	16	1	1	-
	Class 3	3230	234	19	1	-	-	-
	Class 4	2969	172	13	4	2	-	-
Uap	Class 1	6756	987	201	52	8	1	-
	Class 2	5000	588	99	14	-	1	-
	Class 3	2798	170	18	4	-	-	-
	Class 4	2693	148	6	-	1	-	-
Swe	Class 1	6122	864	150	28	6	2	-
	Class 2	5032	621	99	14	4	2	-
	Class 3	3210	216	26	1	-	-	-
	Class 4	2877	174	10	4	1	-	-
LS039	Class 1	6301	851	146	20	10	1	-
	Class 2	5119	612	89	14	1	1	-
	Class 3	3185	215	21	3	-	-	-
	Class 4	2889	165	15	1	1	-	-

“-” indicates that this topology Class was not observed at this length, in this individual.

Supplementary Table 6. Summary of ancient DNA sequencing. The ancient data were generated over a series of experiments described elsewhere[6].

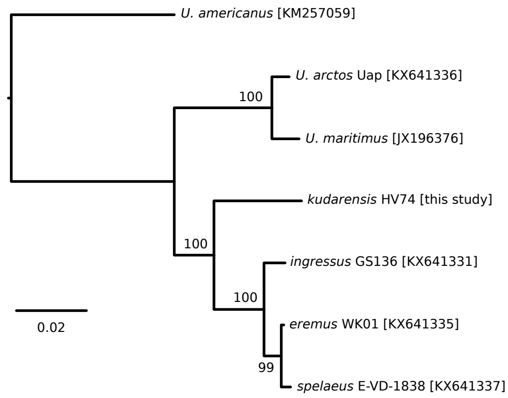
sample	library code	DNA extraction[A]	library[B]	primer[C]	platform	read length	read pairs
WK01	WK01_1	Dab	SS	CL72	MiSeq	75	5272087
					NextSeq	75	494438691
total							*499710778*
GS136	GS136_1	Dab	DS	standard	HiSeq	100	25064683
					HiSeq	100	303706660
	GS136_2	Roh	DS	standard	HiSeq	100	29871305
total							*358642648*
E-VD-1838	UD1838_1	Dab	SS	CL72	NextSeq	75	9201899
					NextSeq	75	399915983
total							*409117882*
HV74	HV74_1	Dab	DS	CL72*	MiSeq	75	2056535
	HV74_2	com	DS	CL72*	MiSeq	70	420279
	HV74_3	com	SS	CL72	MiSeq	70	209010
	HV74_4	Dab	DS	CL72*	MiSeq	70	493704
	HV74_5	Dab	SS	CL72	MiSeq	70	734342
					NextSeq	75	238422118
	HV74_6	Roh	DS	CL72*	MiSeq	70	412114
	HV74_7	Roh	SS	CL72	MiSeq	70	712976
total							*243461078*
Uap	Uap_1	Dab	DS	CL72*	MiSeq	75	848886
	Uap_2	com	DS	CL72*	MiSeq	70	687177
	Uap_3	com	SS	CL72	MiSeq	70	230543
	Uap_4	Dab	DS	CL72*	MiSeq	70	666349
	Uap_5	Dab	SS	CL72	MiSeq	70	297363
					NextSeq	75	154158768
					NextSeq	75	59016440
	Uap_6	Roh	DS	CL72*	MiSeq	70	639214
	Uap_7	Roh	SS	CL72	MiSeq	70	722978
total							*217267718*

[A] *DNA extraction*: "Dab" refers to the published method[8], "Roh" refers to the published method[9], and "com" refers to the combined protocol described in[6].
[B] *Library*: "DS" refers to the published method[10], "SS" refers to the published method[11]
[C] *primer*: describes the R1 sequencing primer used for sequencing, either standard Illumina or the custom primer CL72[11]
* DS libraries sequenced using the CL72 primer require removal of the first 5bp of R1 prior to data processing. This was accomplished using the software seqtk[12]

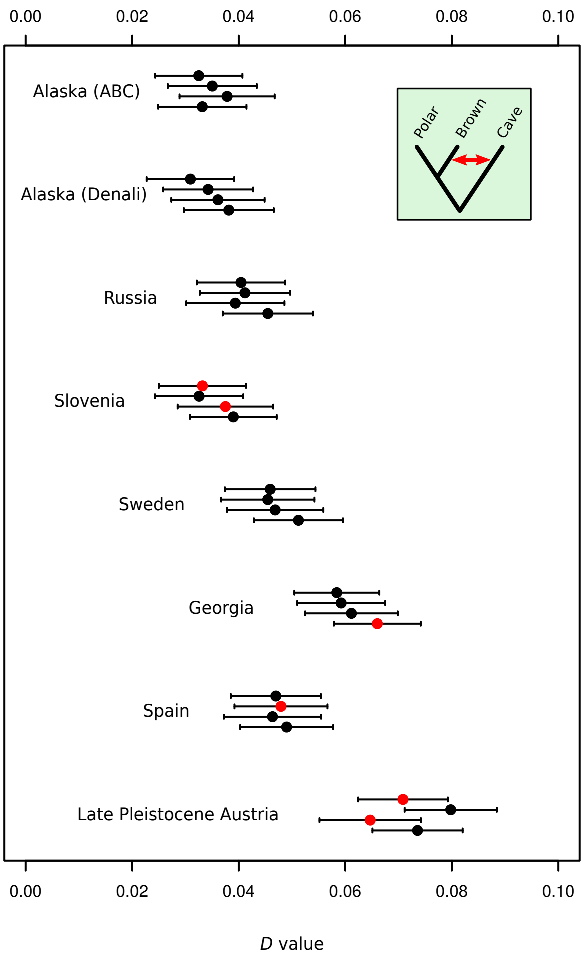
Supplementary Table 7. Block size and detection of alternative topologies. Shown are total counts of blocks returning all 15 possible rooted tree topologies for 6 different block sizes. 25Kbp was selected as an appropriate block size providing adequate sensitivity to detect incomplete lineage sorting and admixture. The four taxa used for this investigation were: NB (Polar), Uap (Brown), HV74 (Cave-), E-VD-1838 (Cave+). Trees were rooted using the black bear outgroup.

Topol.	**Description**	**Tree**	**1MB**	**0.5MB**	**0.25MB**	**100Kbp**	**50Kbp**	**25Kbp**	**Prop***
1	species tree	((Polar,Brown)(Cave-,Cave+))	1051	2269	4413	9518	15714	23817	0.4919
2	symmetrical	((Cave-,Brown)(Polar,Cave+))				11	77	298	0.0062
3	symmetrical	((Cave+,Brown)(Polar,Cave-))			1	16	75	347	0.0072
4	Class 1	(((Cave-,Cave+),Brown),Polar)	3	65	366	1667	4039	8979	0.1854
5	Class 1	(((Cave-,Brown),Cave+),Polar)				49	192	692	0.0143
6	Class 1	(((Cave+,Brown),Cave-),Polar)			1	36	208	1191	0.0246
7	Class 2	(((Cave-,Cave+),Polar),Brown)	3	19	146	918	2487	5330	0.1101
8	Class 2	(((Cave-,Polar),Cave+),Brown)			2	32	153	500	0.0103
9	Class 2	(((Polar,Cave+),Cave-),Brown)			2	24	129	512	0.0106
10	Class 3	(((Polar,Brown),Cave+),Cave-)			11	178	817	2839	0.0586
11	Class 3	(((Polar,Cave+),Brown),Cave-)				18	97	428	0.0088
12	Class 3	(((Cave+,Brown),Polar),Cave-)				19	117	532	0.0110
13	Class 4	(((Polar,Brown),Cave-),Cave+)		1	11	167	788	2123	0.0438
14	Class 4	(((Cave-,Polar),Brown),Cave+)				17	84	362	0.0075
15	Class 4	(((Cave-,Brown),Polar),Cave+)			1	12	82	473	0.0098
Totals			1057	2354	4954	12682	25059	48423	1

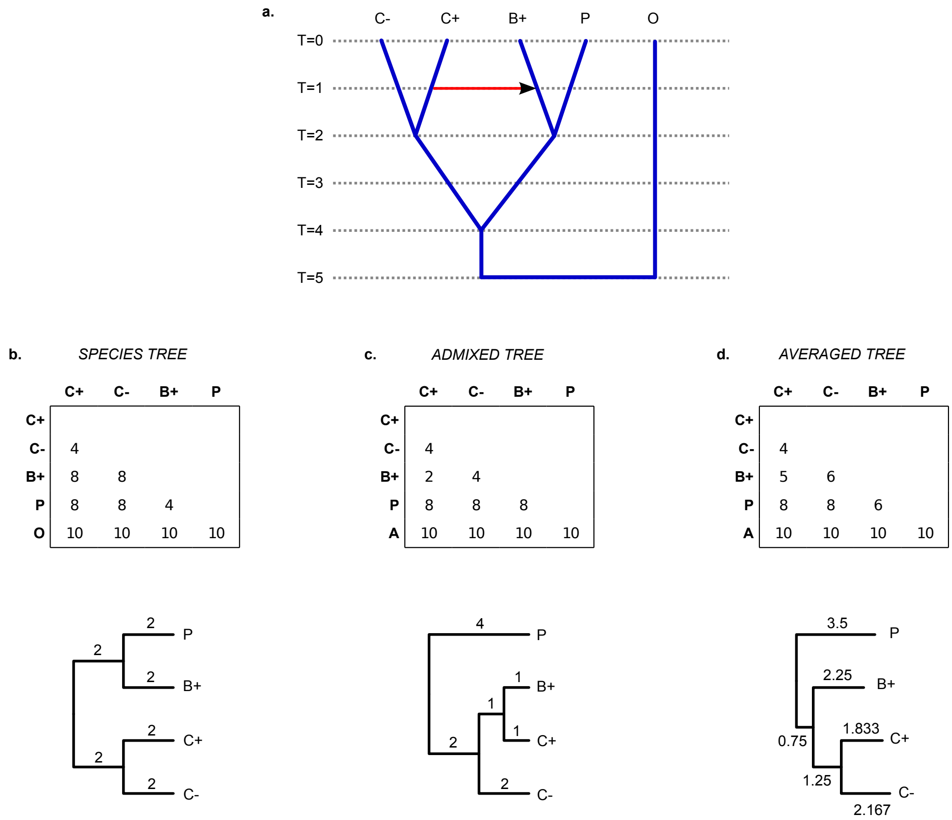
* “Prop” indicates proportion of blocks returning this topology at 25Kbp block size.



Supplementary Figure 1. Maximum-likelihood mitochondrial phylogeny of cave bears used in this study. The Pleistocene brown bear (*U. arctos* Uap) and a polar bear (*U. maritimus*) are included to show their respective phylogenetic positions, and the tree is rooted using the American black bear (*U. americanus*). Clade support values are bootstrap percentages, and the scale bar indicates substitutions per site. Genbank accession numbers for previously published sequences are shown inside square brackets.



Supplementary Figure 2. *D* statistic tests for admixture. Results are presented for eight brown bears (indicated by locality on the left) and a polar bear, with clusters of four points for each brown bear corresponding to four separate tests, each using a different cave bear as potential introgressor (ordered within each cluster from top to bottom: *eremus, spelaeus, ingressus*, *kudarensis*). X-axes indicate *D* values and bars indicate ± 1 weighted block jackknife standard error. Positive *D* values provide evidence of admixture (in either direction) between cave bears and brown bears following the divergence of brown bears and polar bears (inset top right). All *D* values are statistically significant ($Z > 3$ in all comparisons). Red points indicate comparisons involving geographically paired brown bears and cave bears. No obvious pattern of geographically localised admixture is apparent. Although tests of the Georgian brown bear returned a higher *D* value for the geographically paired Caucasus cave bear *kudarensis*, error estimates are widely overlapping with those for other cave bears precluding any conclusive interpretation of this pattern.

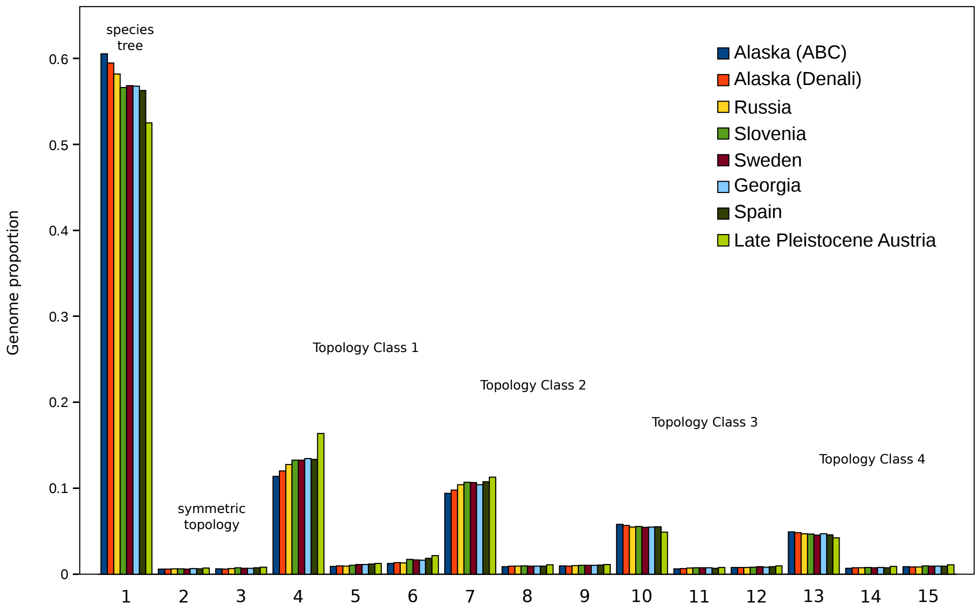


Supplementary Figure 3. Effect of pseudohaploid sequences, when an individual is heterozygous for admixed and unadmixed alleles. Consider an admixture scenario between cave bears and brown bears. We sample five individuals: an admixed cave bear (C+), an unadmixed cave bear (C-), an admixed brown bear (B+), a polar bear which is not admixed with cave bear (P), and an outgroup (O). a. shows the species tree with divergence times indicated by dashed lines. At T=1 geneflow from the admixed cave bear lineage to the brown bear lineage occurs. b.-c. Show distance matrices for alternative trees, proportional to divergence times indicated in a. Below each matrix is the calculated neighbour-joining tree, with branch lengths indicated. Pseudohaploidisation of a heterozygous region results in a matrix (d.) that is averaged across the species tree (b.) and an admixed tree (c.). The position of the admixed brown bear (B+) in the resulting averaged tree (d.) has shifted relative to the admixed tree (c.), but overall tree asymmetry has been preserved.

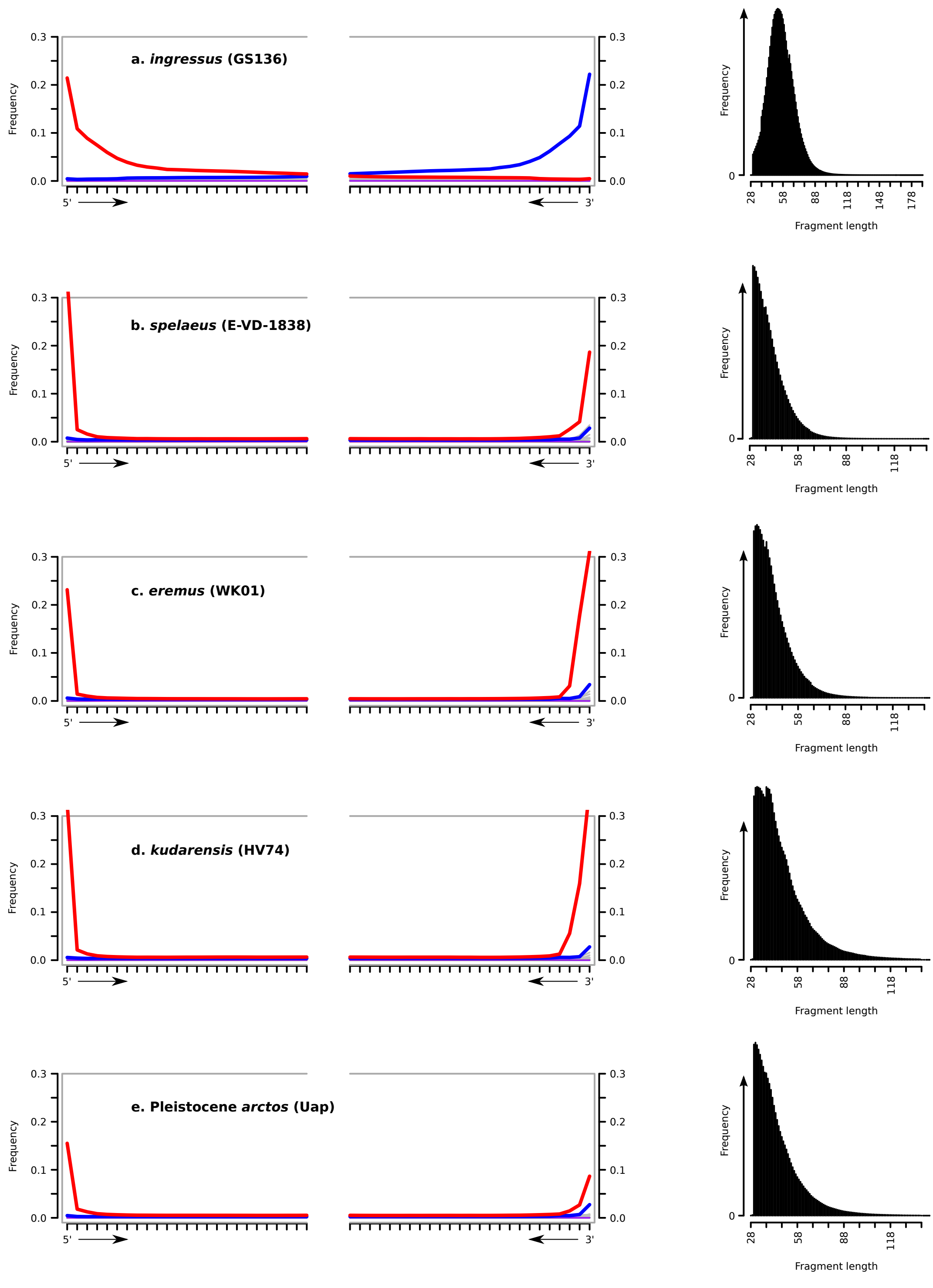
a

Topology	Description	Tree	Alaska (Admiralty)	Alaska (Denali)	Russia	Slovenia	Sweden	Caucasus	Spain	Austria Pleistocene
1	species tree	((Polar,Brown)(Cave-,Cave+))	32133	31551	30868	29205	30010	30128	29484	25495
2	symmetrical	((Cave-,Brown)(Polar,Cave+))	302	307	326	313	313	346	319	346
3	symmetrical	((Cave+,Brown)(Polar,Cave-))	319	316	340	375	364	361	383	385
4	Class 1	(((Cave-,Cave+),Brown),Polar)	6032	6370	6776	6835	6992	7131	7000	7940
5	Class 1	(((Cave-,Brown),Cave+),Polar)	4981	5186	5525	5510	5627	5526	5636	5475
6	Class 1	(((Cave+,Brown),Cave-),Polar)	455	490	490	491	496	493	487	526
7	Class 2	(((Cave-,Cave+),Polar),Brown)	364	393	400	397	394	406	388	432
8	Class 2	(((Cave-,Polar),Cave+),Brown)	450	435	448	488	486	482	486	529
9	Class 2	(((Polar,Cave+),Cave-),Brown)	479	504	496	530	590	596	614	602
10	Class 3	(((Polar,Brown),Cave+),Cave-)	502	494	525	520	536	545	554	534
11	Class 3	(((Polar,Cave+),Brown),Cave-)	318	339	378	378	385	394	360	374
12	Class 3	(((Cave+,Brown),Polar),Cave-)	2597	2550	2496	2398	2396	2488	2399	2051
13	Class 4	(((Polar,Brown),Cave-),Cave+)	3078	3009	2907	2858	2881	2910	2887	2367
14	Class 4	(((Cave-,Polar),Brown),Cave+)	407	411	410	411	458	419	443	467
15	Class 4	(((Cave-,Brown),Polar),Cave+)	653	712	690	873	872	854	963	1045
Total number of blocks			53070	53067	53075	51582	52800	53079	52403	48568
cave bear admixture proportion*			0.01155	0.01334	0.01340	0.01664	0.01700	0.01900	0.01813	0.03142

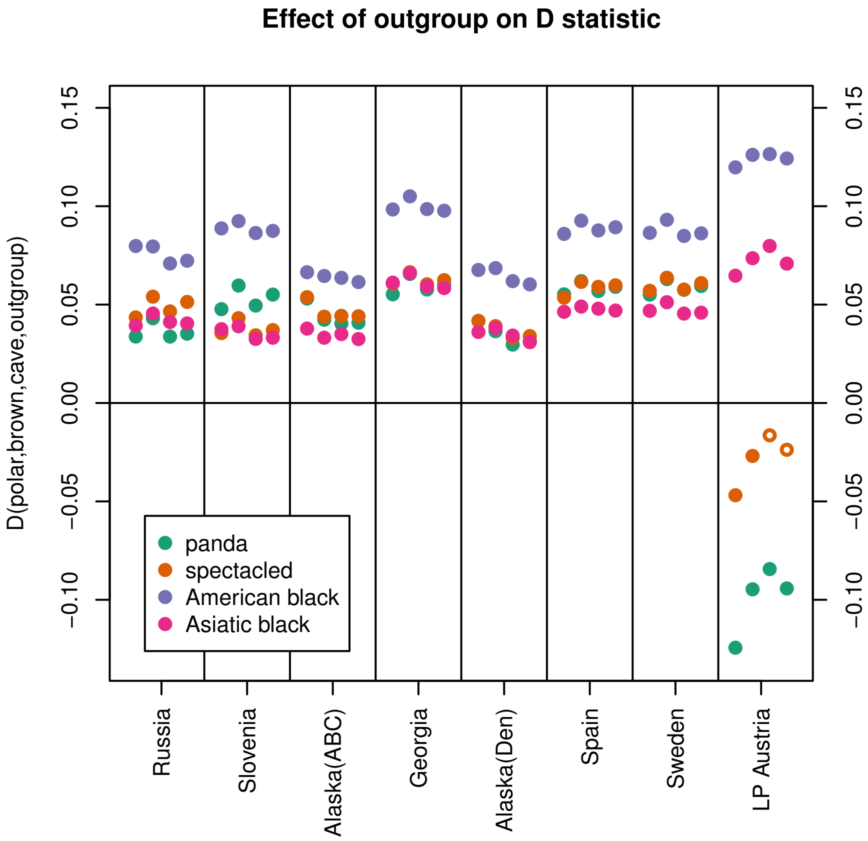
b



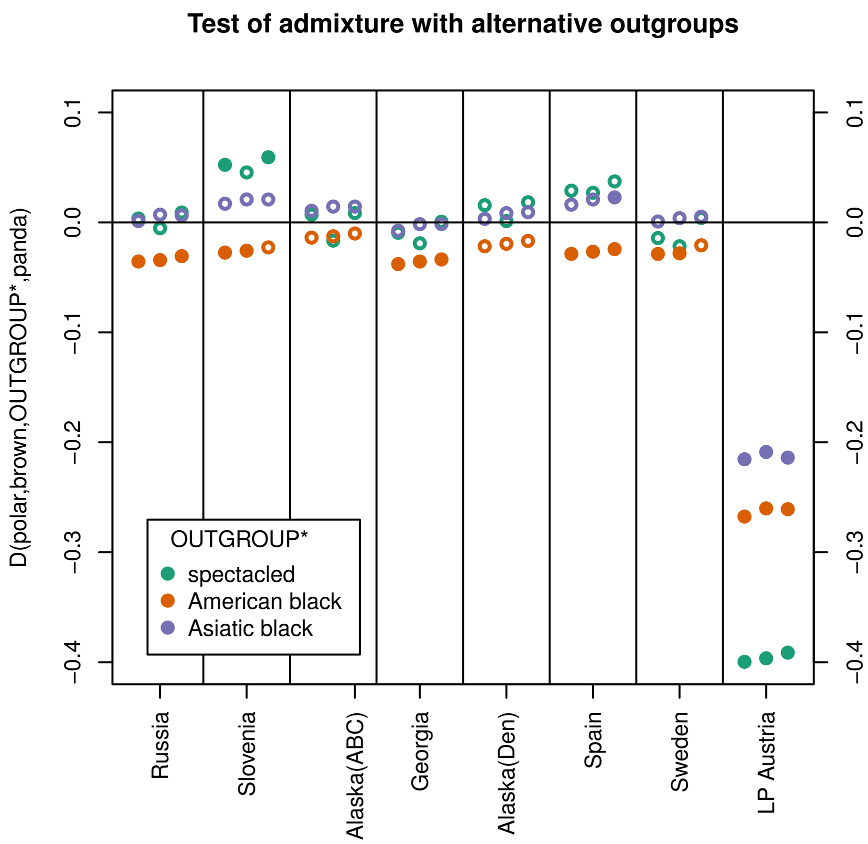
Supplementary Figure 4. a. Results of phylogenetic admixture tests, showing total counts of 25kbp non-overlapping genomic blocks returning 15 possible rooted tree topologies, for each brown bear investigated. The indicated tree topologies comprise polar bear (NB), *spelaeus* (Cave+), *kudarensis* (Cave-), and all brown bears investigated (right most eight columns, with locality indicated). The cave bear admixture proportion is estimated by subtracting the total number of 25Kbp blocks returning Class 2 from the total number returning Class 1, and dividing the resulting number by two times the total number of observed blocks. This method assumes that all introgressed blocks in brown bears are heterozygous for cave bear and brown bear haplotypes. It therefore represents an underestimate of the true admixture proportion, since some introgressed blocks appear to exist in a homozygous state (topologies 5 and 6). b. Bar charts showing the genomic representation of each topology for each investigated brown bear.



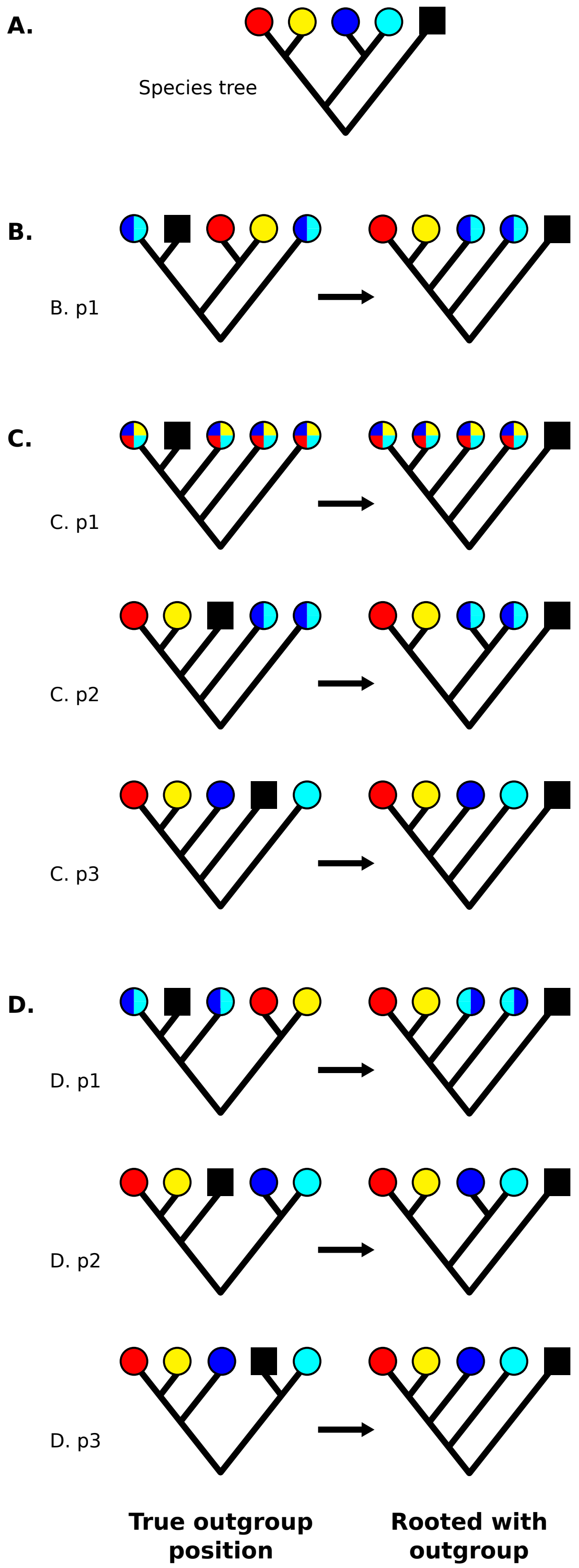
Supplementary Figure 5. Authentication of ancient DNA data from ancient bear samples (a.-e.). Plots on the left show cytosine deamination patterns. X axes indicate individual nucleotide positions of ancient DNA fragments, with 5' and 3' fragment ends indicated. Plots show the proportion of fragments possessing a T where the reference polar bear genome possesses a C (red plots), and proportion of fragments possessing an A where the reference possesses a G (blue plots). An overabundance of C->T substitutions at the fragment ends is indicative of cytosine deamination and ancient DNA damage. The overabundance in G->A substitutions at 3' read fragment ends in sample GS136 (a.) is an artefact of the double-stranded library procedure, and in reality represents C->T changes. Fragment length distributions for each sample are shown to the right, determined from the lengths of merged reads mapping to the polar bear reference. The minimum read length for mapping was 30bp, resulting in truncation of the distribution at this length. The double-stranded library preparation method used for sample GS136 (a.) is known to be associated with a loss of very short DNA fragments relative to the single-stranded methods (b.-e.), and likely explains the increase in fragment lengths observed in this dataset.



Supplementary Figure 6. Effect of outgroup selection on inferred patterns of admixture between brown bears and cave bears, following the divergence of brown bears and polar bears. The species tree is (((polar bear, brown bear),cave bear),outgroup), with American black bear, Asiatic black bear, spectacled bear and giant panda representing, respectively, increasingly divergent outgroups. We present *D* statistic (y-axis) results for 8 brown bears (x-axis) as points coloured according the outgroup utilised, with sets of 4 points per brown bear/outgroup combination representing values calculated each using one of four cave bears. Open points indicate non-significant *D* values (absolute Z score < 3). For modern brown bears, all four outgroups produce positive *D* values for all tests, but we observe elevated *D* values when using American black bear, relative to other outgroups. Positive values were also produced for the ancient brown bear (LP Austria) when using the two least divergent outgroups (American and Asiatic black bears). However, negative values are produced when the two most divergent outgroups (spectacled bear and giant panda) are used, which are significant in most cases. We attribute this effect to accumulated errors in the ancient pseudohaploid sequences as a result of both sequencing error and spurious read mapping, both of which tend to occur at higher rates in ancient relative to modern DNA datasets. Specifically, at sites where the outgroup has a private allele, accumulated errors in an ancient sample occupying the P2 position would convert a proportion of these BBBA sites to BABA sites, causing that individual to appear unadmixed relative to P1. This effect is amplified with more divergent outgroups since they will posses more private alleles, relative to the ingroup.



Supplementary Figure 7. Assessment of imbalance between polar bears and brown bears in derived allele sharing with three candidate outgroups. We present *D* statistic (y-axis) results for 8 brown bears (x-axis) as points coloured according the candidate outgroup utilised, with sets of 3 points per brown bear/outgroup combination representing values calculated each using a different polar bear. Open points indicate non-significant *D* values (absolute Z score < 3). The giant panda served as the P4 outgroup taxon for allele polarisation. We find a significant excess of derived alleles shared between polar bears and American black bear relative to modern brown bears in almost all comparisons. *D* values for Asiatic black bear are generally closest to zero and with the fewest significant values. All comparisons involving the ancient brown bear (LP Austria) produced large and significantly negative *D* values, which is attributable to accumulated errors in the ancient brown bear pseudohaploid sequence (see Supplementary Figure 6 for further explanation).

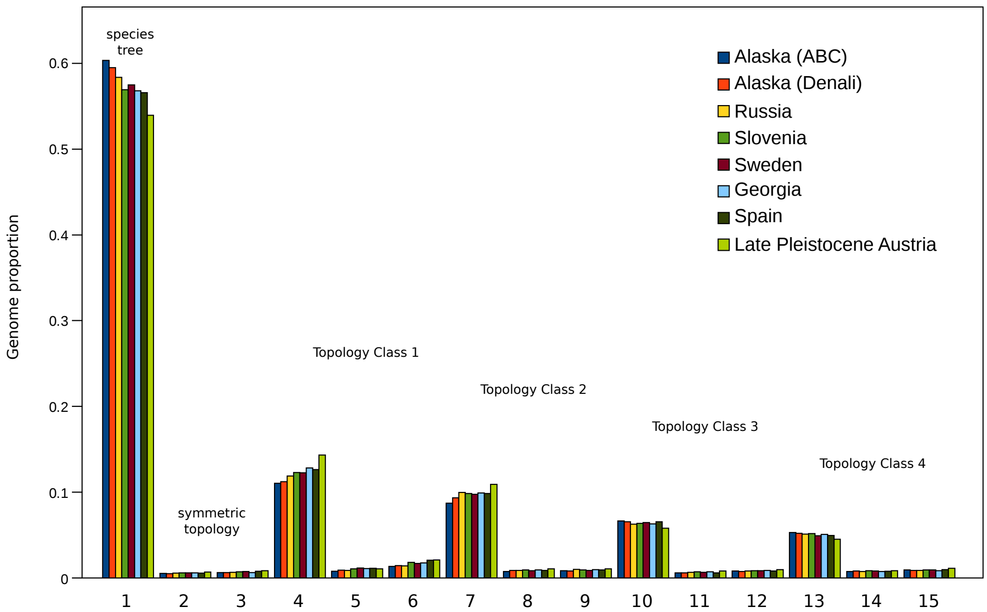


Supplementary Figure 8.
Effect of outgroup misspecification. We consider a rooted symmetrical four-species tree with a fifth outgroup taxon (A.) which is suited to admixture testing. We consider the effect of violating the assumed outgroup position. That is, the outgroup lineage in fact occupies an ingroup position for a given genomic block. A rooted five-taxon tree has 105 possible topologies representing three possible shapes (B., C., D.). For each of these tree shapes, the assumed outgroup taxon may occupy one (B. p1), three (C. p1–3) or three (D. p1–3) alternative ingroup positions, respectively. The three tree shapes with each possible outgroup position are shown on the left, and the trees resulting from rooting these trees using the assumed outgroup (black square) are shown on the right. For illustrative purposes, the four ingroups are represented by coloured circles, but it should be noted that this represents only a small subset of the total possible topologies. We can, however, consider the effect of outgroup misspecification across all topologies by reducing them to these seven combinations of tree shape and outgroup position. Where ingroup lineages have sorted with respect to the outgroup, these are shown as single-colour circles. Where lineages are incompletely sorted with respect the outgroup, they are shown as multi-coloured circles, reflecting that under a model of random lineage sorting these lineages will be represented equally at each of these positions if a large number of loci are sampled. In some cases, incomplete sorting of one (C. p3, D. p3) or two (B. p1, D. p1) ingroup lineages is correctly identified, albeit with an incorrect outgroup position, but in none of these cases do we predict an imbalance in frequency by which two incompletely sorted lineages occupy, respectively, basal ingroup positions. When no two lineages have sorted with respect to the outgroup, a topology which is informative on admixture is artifactually produced (C. p1), but under random lineage sorting all possible topologies will occur at equal frequencies and equilibrium of admixture informative topology classes is maintained. Outgroup misspecification can also artifactually group two lineages that have not yet sorted, but the resulting tree has a symmetrical ingroup and does not contribute to admixture inference (C. p2). Any decreases in admixture informative topology counts resulting from this effect will be balanced, respectively, between the two incompletely sorted lineages, preserving equilibrium of topology classes. When both ingroup clades have sorted with respect to each other, but not with respect to the outgroup (D. p2), ingroup monophyly is incorrectly inferred but the resulting tree is symmetrical and does not contribute to admixture inference.

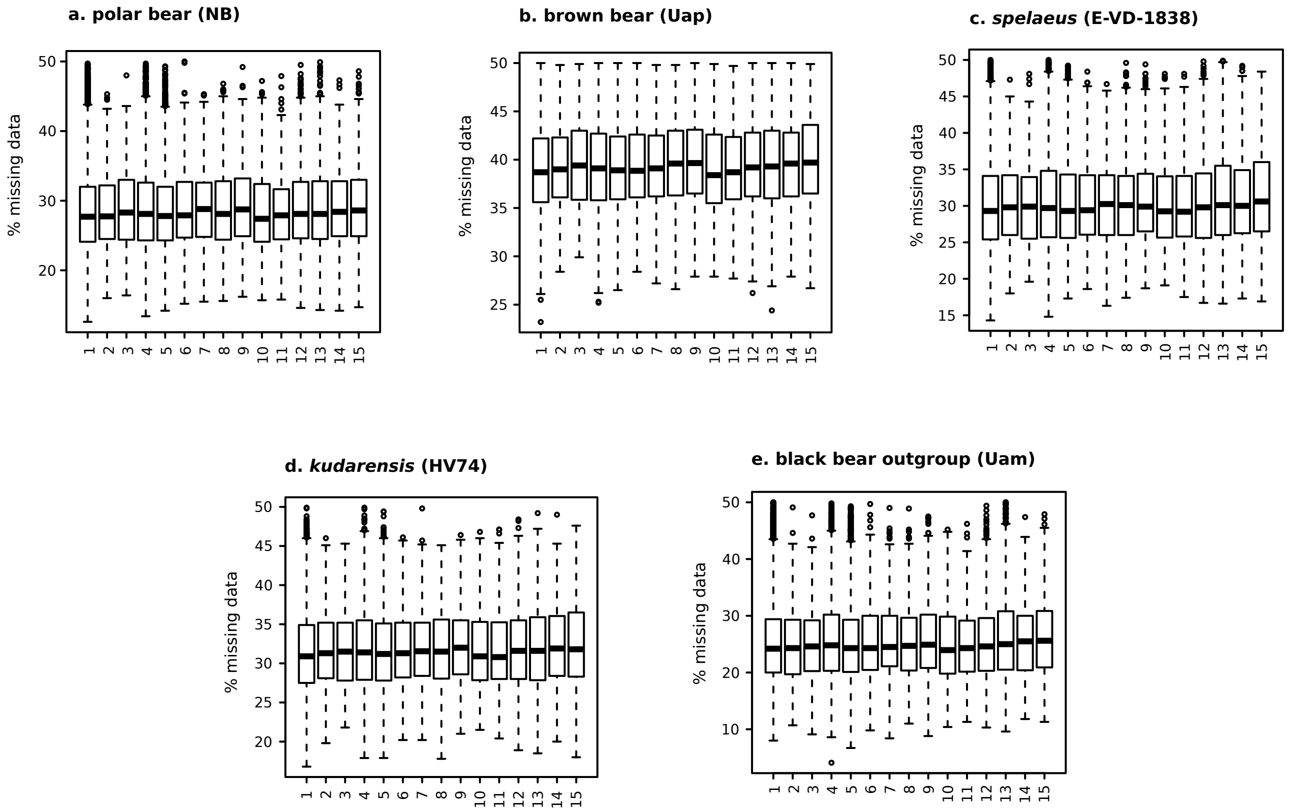
a

Topology	Description	Tree	Admiralty (Admiralty)	Alaska (Denali)	Russia	Slovenia	Sweden	Caucasus	Spain	Austria Pleistocene
1	species tree	((Polar,Brown)(Cave-,Cave+))	14625	14393	14073	13353	13810	13816	13428	11213
2	symmetrical	((Cave-,Brown)(Polar,Cave+))	130	119	135	139	139	139	129	142
3	symmetrical	((Cave+,Brown)(Polar,Cave-))	154	153	157	166	176	149	183	174
4	Class 1	(((Cave-,Cave+),Brown),Polar)	2672	2716	2862	2885	2946	3116	2994	2978
5	Class 1	(((Cave-,Brown),Cave+),Polar)	188	221	214	245	279	263	264	217
6	Class 1	(((Cave+,Brown),Cave-),Polar)	321	350	335	425	403	428	489	434
7	Class 2	(((Cave-,Cave+),Polar),Brown)	2110	2259	2406	2310	2342	2407	2335	2271
8	Class 2	(((Cave-,Polar),Cave+),Brown)	184	207	208	221	201	231	209	220
9	Class 2	(((Polar,Cave+),Cave-),Brown)	205	192	238	223	212	236	225	217
10	Class 3	(((Polar,Brown),Cave+),Cave-)	1610	1587	1513	1491	1550	1532	1551	1206
11	Class 3	(((Polar,Cave+),Brown),Cave-)	146	145	160	169	160	176	142	166
12	Class 3	(((Cave+,Brown),Polar),Cave-)	192	181	191	194	201	208	192	202
13	Class 4	(((Polar,Brown),Cave-),Cave+)	1284	1260	1232	1215	1182	1235	1176	936
14	Class 4	(((Cave-,Polar),Brown),Cave+)	183	194	180	199	194	183	186	175
15	Class 4	(((Cave-,Brown),Polar),Cave+)	225	212	209	215	225	206	227	235
Total number of blocks			24229	24189	24113	23450	24020	24325	23730	20786

b



Supplementary Figure 9. a. Results of phylogenetic admixture tests after filtering for only those blocks where the ingroup is monophyletic with respect to the assumed outgroup (Asiatic black bear). b. Bar charts showing the genomic representation of each topology for each investigated brown bear. Annotation is identical to Supplementary Figure 4, which shows results for the complete dataset. Although absolute topology counts are reduced in the filtered dataset compared to the complete dataset, relative abundances are highly consistent and support identical admixture inferences.



Supplementary Figure 10. Effect of missing data on the phylogenetic test for admixture. Shown are results for the test involving the Late Pleistocene brown bear (Uap). This dataset has far lower coverage than for any modern brown bear, and is therefore most sensitive to any potential effect of missing data. Panels a.-e. are boxplots showing the amounts of missing data across 25Kbp genomic blocks returning each of the 15 possible rooted tree topologies (x-axis, numbering corresponds with Supplementary Figure 4), for each of the five individuals used in the test (panels a.-e., respectively). Specific tree topologies are not associated with any obvious increase or decrease in missing data for any of the five individuals, indicating that the inferred admixture patterns are not driven by sequencing coverage.

SUPPLEMENTARY REFERENCES

1. Cahill, J. a *et al.* Genomic evidence of geographically widespread effect of gene flow from polar bears into brown bears. *Mol. Ecol.* **24,** 1205–1217 (2015).
2. Benazzo, A. *et al.* Survival and divergence in a small group: The extraordinary genomic history of the endangered Apennine brown bear stragglers. *Proc. Natl. Acad. Sci.* **7,** 9589–9597 (2017).
3. Cahill, J. A. *et al.* Genomic evidence for island population conversion resolves conflicting theories of polar bear evolution. *PLoS Genet.* **9,** e1003345 (2013).
4. Kumar, V. *et al.* The evolutionary history of bears is shaped by gene flow across species. *Sci. Rep.* **7,** 46487 (2017).
5. Fortes, G. G. *et al.* Ancient DNA reveals differences in behaviour and sociality between brown bears and extinct cave bears. *Mol. Ecol.* **25,** 4907–4918 (2016).
6. Barlow, A. et al. Massive influence of DNA isolation and library preparation approaches on palaeogenomic sequencing data. bioRxiv (2016).
7. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).
8. Dabney, J. *et al.* Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 15758–15763 (2013).
9. Rohland, N., Siedel, H. & Hofreiter, M. A rapid column-based ancient DNA extraction method for increased sample throughput. *Mol. Ecol. Resour.* **10,** 677–683 (2010).
10. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010,** pdb.prot5448 (2010).
11. Gansauge, M.-T. & Meyer, M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat. Protoc.* **8,** 737–748 (2013).
12. Li, H. "seqtk Toolkit for processing sequences in FASTA/Q formats." (2012).