Ancient genomes reveal hybridisation between extinct short-faced bears and the extant spectacled bear (*Tremarctos ornatus*)

Alexander T Salis[1#*], Graham Gower[1,2], Blaine W. Schubert[3], Leopoldo H. Soibelzon[4], Holly Heiniger[1], Alfredo Prieto[5], Francisco J. Prevosti[6,7], Julie Meachen[8], Alan Cooper[9], Kieren J. Mitchell[1*]

[1]Australian Centre for Ancient DNA (ACAD), School of Biological Sciences, University of Adelaide, South Australia 5005, Australia

[2]Lundbeck GeoGenetics Centre, GLOBE Institute, University of Copenhagen, Copenhagen 1350, Denmark

[3]Center of Excellence in Paleontology and Department of Geosciences, East Tennessee State University (ETSU), Johnson City, Tennessee 37614, USA

[4]División Paleontología de Vertebrados, Museo de La Plata, 1900 La Plata, Argentina

[5]Centro de Estudios del Hombre Austral: Instituto de la Patagonia, Universidad de Magallanes, Punta Arenas, Chile

[6]Museo de Ciencias Antropológicas y Naturales, Universidad Nacional de La Rioja (UNLaR), La Rioja, Argentina

[7]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

[8]Anatomy Department, Des Moines University, Des Moines, IA, USA

[9]South Australian Museum, Adelaide, South Australia 5000, Australia

Lead contact

*Corresponding author(s): A.T.S. (alexander.t.salis@gmail.com), and K.J.M. (kieren.mitchell@adelaide.edu.au)

Summary:

Two genera and multiple species of short-faced bear from the Americas went extinct during or toward the end of the Pleistocene, and all belonged to the endemic New World subfamily Tremarctinae [1-7]. Two of these species were giants, growing in excess of 1,000 kg [6, 8, 9], but it remains uncertain how these extinct bears were related to the sole surviving short-faced bear: the spectacled bear (*Tremarctos ornatus*). Ancient mitochondrial DNA has recently suggested phylogenetic relationships among these lineages that conflict with interpretations based on morphology [1, 10-12]. However, widespread hybridisation and incomplete lineage sorting among extant bears mean that the mitochondrial phylogeny frequently does not reflect the true species tree [13, 14]. Here we present ancient nuclear genome sequences from representatives of the two extinct short-faced bear genera, *Arctotherium* and *Arctodus*. Our new data support a third hypothesis for the relationships among short-faced bears, which conflicts with existing mitochondrial and morphological data. Based on genome-wide D-statistics, we suggest that the extant spectacled bear derives substantial ancestry from Pleistocene hybridisation with an extinct short-faced bear lineage, resulting in a discordant phylogenetic signal between the mitochondrion and portions of the nuclear genome.

Results and Discussion:

The spectacled bear (*Tremarctos ornatus)* is the only extant species of short-faced bear (Tremarctinae), a once diverse subfamily endemic to the Americas. This subfamily also includes many species that became extinct during the Pleistocene, including the Florida cave bear (*Tremarctos floridanus*), two species of North American short-faced bears (*Arctodus* spp. [3, 4]), and as many as five species of South American short-faced bears (*Arctotherium* spp. [2, 6]), one of which (*Arctotherium wingei*) has recently been discovered as far north as the Yucatan of Mexico [5]. Notably, the genera *Arctodus* and *Arctotherium* both included giant (>1,000kg) forms [8, 9] — *Arctodus simus* and *Arctotherium angustidens*, respectively — and based on morphology it was hypothesised that these genera were closely related [1, 6, 10, 11]. However, recently published mitochondrial DNA data suggested that *Arctotherium* was most closely related to the extant spectacled bear, to the exclusion of North American *Arctodus* [12]. While this result supported the convergent evolution of giant bears in North and South America, the mitochondrial genome does not always reflect the true relationships among species [e.g. 15, 16-19]. Importantly, discordance between mitochondrial and nuclear loci has been previously noted in bears, and has been attributed to a combination of stochastic processes and the rapid evolution of bears [13], as well as hybridisation between species [13, 14, 20-25]. To further resolve the evolutionary history of short-faced bears, we applied ancient DNA techniques to retrieve and analyse whole genome data from both *Arctodus* and *Arctotherium*.

Ancient DNA (aDNA) was extracted and sequenced from three *Arctodus simus* specimens: one each from placer mines at Sixty Mile Creek (ACAD 438; Canadian Museum of Nature; CMN 42388) and Hester Creek (ACAD 344; Yukon Government; YG 76.4) in the Yukon Territory, Canada; and one from Natural Trap Cave in Wyoming, USA (ACAD 5177; University of Kansas; KU 31956). We also analysed one specimen of *Arctotherium* sp. from Cueva del Puma, Patagonia, Chile (ACAD 3599; complete right femur, no. 32104, Centro de Estudios del Hombre Austral, Instituto de la Patagonia, Universidad de Magallanes). The *Arctotherium* specimen was previously dated to 12,105 ± 175 cal yBP (Ua-21033) [26], while two of the *Arctodus* specimens have been dated: ACAD 438 at 47,621 ± 984 cal yBP (TO-2699) [27] and ACAD 5177 at 24,300 ± 208 cal yBP (OxA-37990) (Table S1). The *Arctotherium* specimen has yielded mitochondrial aDNA in a previous studies [12], however, here we shotgun sequenced this specimen, along with the three *A. simus* specimens, at much

greater depth in order to reconstruct nuclear genome sequences. Mapping our new sequencing data from these specimens to the giant panda (*Ailuropoda melanoleuca*) reference genome (LATN01) yielded average depths of coverage between 0.12x to 5.9x (Table S3). We compared these new genomic data to previously published genomes from all extant species of bear (Table S2): spectacled bear, giant panda, brown bear (*Ursus arctos*), American black bear (*U. americanus*), Asian black bear (*U. thibetanus*), polar bear (*U. maritimus*), sloth bear (*U. ursinus*), and sun bear (*U. malayanus*).

Phylogenetic analyses on a concatenated dataset of genome-wide SNPs revealed relationships within Ursinae that were consistent with previous genomic studies: *U. americanus*, *U. maritimus*, and *U. arctos* formed a monophyletic clade sister to a clade consisting of *U. thibetanus*, *U. malayanus*, and *U. ursinus* [13, 14]. In contrast, within short-faced bears (Tremarctinae) we recovered strong support for a close relationship between the spectacled bear and the North American short-faced bear (*Arctodus simus*) to the exclusion of the South American *Arctotherium* (Figure 1A, Figure S2). This result conflicts with the mitochondrial tree, which instead supports a clade comprising *Arctotherium* and *Tremarctos ornatus* to the exclusion of *Arctodus simus* [12] (Figure 1B). As the radiation of bears is thought to have occurred rapidly during the Miocene - Pliocene transition, it is possible that this discordance could be explained by incomplete lineage sorting (ILS) [28], a process whereby pre-existing genetic variation in an ancestral species is randomly inherited and fixed in descendant species [29, 30]. Alternatively, given the observed propensity of bears for hybridisation [e.g. 13, 14, 20-22, 25, 31], mitochondrial/nuclear discordance within short-faced bears may instead result from gene flow between *Tremarctos* and either *Arctodus* or *Arctotherium*.

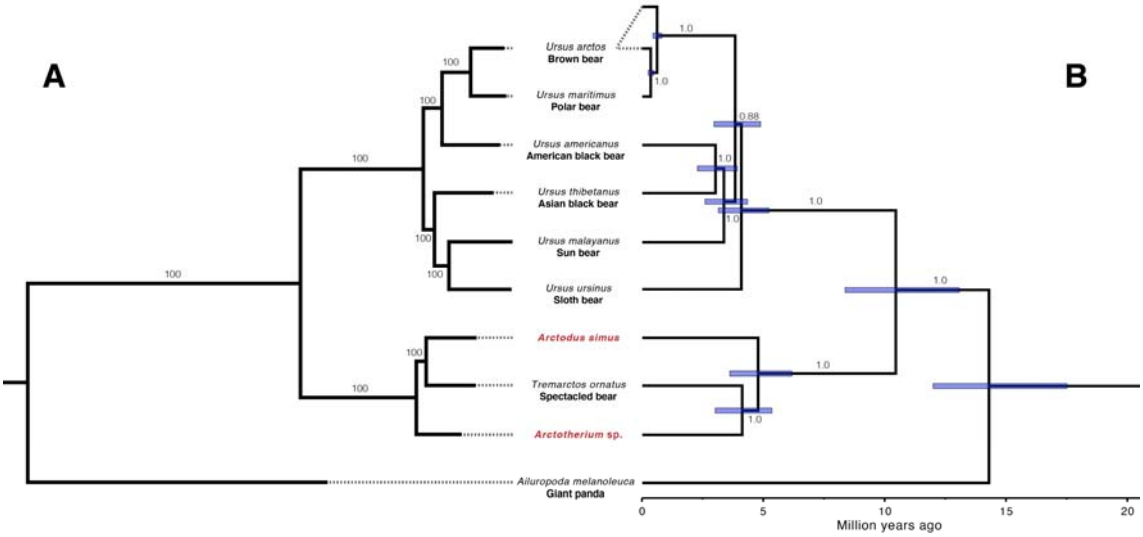

Figure 1: Phylogenetic relationships among ursids **A.** Maximum likelihood tree based on nuclear SNPs constructed in RAxML. Branch labels represent bootstrap support percentages. For RAxML tree with all individuals analysed see Figure S2. **B.** Bayesian phylogeny based on full mitochondrial genomes adapted from Mitchell, et al. [12]. Blue bars represent 95% highest posterior density interval on node ages. Branch labels represent BEAST posterior support values.

To test for potential phylogenetic discordance across our short-faced bear genomes, we constructed phylogenetic trees from 500 kb non-overlapping windows (n = 2622) across the 85 largest autosomal scaffolds of the giant panda reference genome (LATN01). Trees created from roughly 70% of windows agreed with the results from our genome-wide concatenated dataset (Topology 1; *i.e. Tremarctos* + *Arctodus*; Figure 2B & S3). However, approximately 30% of windows instead supported the mitochondrial tree topology (Topology 2; *i.e. Tremarctos* + *Arctotherium*; Figure 2B & S3), while the third possible topology where the two extinct genera form a clade — *Arctodus* + *Arctotherium* — was rejected for over 95% of windows. The frequencies of the three possible tree topologies are difficult to explain as a result of ILS, which we would expect to result in a more even representation of the two "minority" topologies (*i.e.* Topologies 2 and 3). Our results therefore suggest that introgression may be the most likely explanation for the observed phylogenetic discordance. Consequently, we calculated D-statistics [32, 33] using our concatenated genome-wide SNPs in order to identify signals of hybridisation between the bear species in our dataset.

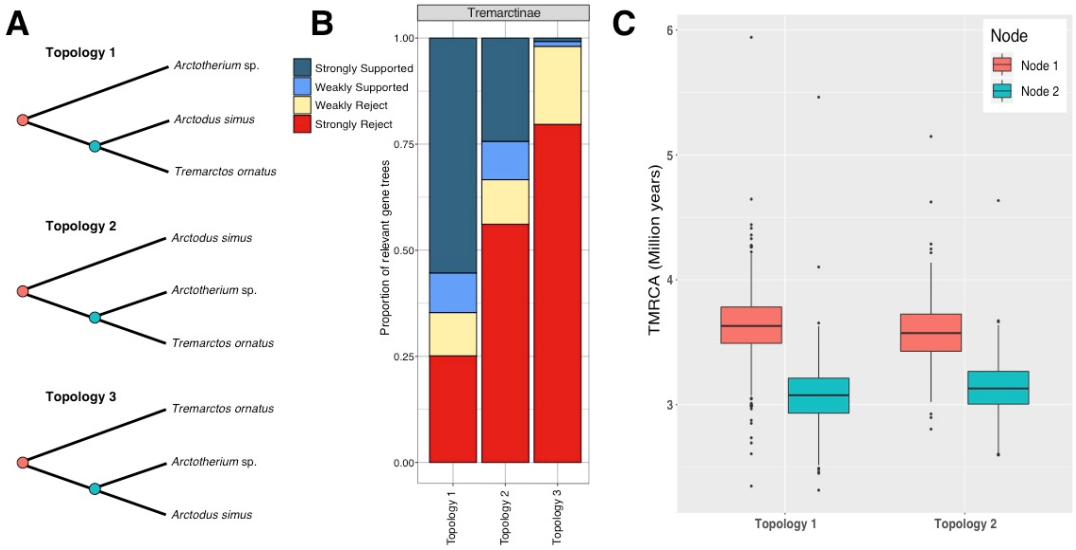

Figure 2: A. The three possible short-faced bear (Tremarctinae) tree topologies. **B.** Discordance visualisation using DiscoVista from 2622 500 kb genomic fragments. The x-axis represents topologies tested and the y-axis the proportion of fragments that support the topology, with >80% bootstrap support used to define strong support. For more comprehensive tests of phylogenetic placements see Figure S3. **C.** Divergence time estimates (TMRCA) of *Tremarctos, Arctodus, and Arctotherium* (Node 2) and for sister species (Node 1) of the two most common short-faced bear topologies.

Consistent with previous studies [i.e. 14], our D-statistics revealed compelling evidence for hybridisation between: Asian black bears and all North American ursine bears (including the polar bear); sun bears and North American ursine bears; and Asian black bear and sun bear (Table S4). In contrast, we did not obtain any significantly non-zero values for D-statistics calculated using our two extinct short-faced bear genomes, any member of Ursinae, and the panda outgroup (Table 1). This result suggests that no gene flow occurred between *Arctodus* or *Arctotherium* and the ancestors of any modern ursine bear, and also demonstrates a lack of any discernible reference bias in the ancient genomic data (which would result in asymmetrical allele sharing with the reference). Thus, it appears *Arctodus* and *Arctotherium* did not hybridise with brown and black bears in the Americas during the late Pleistocene, even though the distribution of *Arctodus* overlapped with both ursines, and *Arctotherium* may have encountered them in Mexico or Central America [5]

Contrary to previous studies, our D-statistics revealed signals consistent with gene flow between the spectacled bear and members of Ursinae (Table 1 & S5), suggesting the possibility that *Tremarctos* hybridised with ancestors of either the brown bear or American

black bear during the Pleistocene. This signal is surprising given the deep divergence between ursine and short-faced bears, having split approximately 10 million years ago (mya) [12, 14, 28]. However, in support of this hypothesis, offspring between spectacled bear and American black bear have resulted from hybridisation in zoos, although whether these hybrids were fertile remains unknown [34]. Importantly, members of *Tremarctos* and the ancestors of modern American black bears had overlapping distributions throughout the Pleistocene in North America [4, 10], meaning that hybridisation may have occurred when the two lineages were less divergent and reproductive barriers had had less time to evolve.

In addition to evidence for hybridisation between *Tremarctos* and ursine bears, we also recovered convincing evidence for hybridisation between *Arctotherium* and *Tremarctos* (Table 1). These results are consistent with a model where the divergence between *Arctodus* and *Tremarctos* occurred in North America after the ancestors of *Arctotherium* dispersed southwards into South America, with subsequent hybridisation between *Tremarctos* and *Arctotherium*. This interpretation is supported by the presence of *Arctodus* and *Tremarctos* (and absence of *Arctotherium*) in the late Pliocene fossil record of North America [3, 4, 7, 10]. The fossil record further suggests that contact between *Tremarctos* and *Arctotherium* occurred during the late Pleistocene, when representatives of *Arctotherium* were distributed as far north as the Yucatan of Mexico [5], providing an opportunity for hybridisation.

Table 1: D-statistics for short-faced bears (Tremarctinae). D-statistics (D), standard error, and Z-Score (significant if > |3|) are displayed, with ABBA-BABA counts and the number of SNPs considered in the analysis. It is clear that there is an excess of allele sharing between the spectacled bears (*T. ornatus*) and *Arctotherium*. However, neither of the spectacled bear individuals show elevated D-statistics in relation to each other meaning gene flow likely occurred in the ancestor of both individuals, or they carry similar proportions of hybridised DNA.

D-statistic: D(H1, H2, H3, Giant Panda)	D	Stderr	Z-Score	BABA	ABBA	nSNPs
D(*T. ornatus* (Chaparri), *Arctodus*, *Arctotherium*)	0.3116	0.009559	32.6*	41219	21633	6071021
D(*T. ornatus* (Nobody), *Arctodus*, *Arctotherium*)	0.3112	0.009462	32.889*	41187	21637	6070446
D(*T. ornatus* (Chaparri), *T. ornatus* (Nobody), *Arctodus*)	0.0206	0.022575	0.911	1375	1320	6393667
D(*T. ornatus* (Chaparri), *T. ornatus* (Nobody), *Arctotherium*)	0.0172	0.021533	0.801	1227	1186	6258343
D(*T. ornatus* (Nobody), *Arctotherium* sp., *Ursus arctos*)	0.2079	0.008878	23.418*	11177	7329	6184195
D(*T. ornatus* (Chaparri), *Arctotherium* sp., *Ursus arctos*)	0.2074	0.009242	22.442*	11185	7342	6185696
D(*T. ornatus* (Nobody), *Arctodus simus*, *Ursus arctos*)	0.2152	0.009651	22.302*	11172	7214	6317724
D(*T. ornatus* (Chaparri), *Arctodus simus*, *Ursus arctos*)	0.2131	0.010012	21.281*	11191	7260	6319318
D(*Arctodus simus*, *Arctotherium* sp., *Ursus malayanus* (Klaus))	0.0215	0.00854	2.523	9662	9254	6010428
D(*Arctodus simus*, *Arctotherium* sp., *Ursus americanus*)	0.0225	0.009168	2.453	9661	9236	6025070
D(*Arctodus simus*, *Arctotherium* sp., *Ursus malayanus* (Anabell))	0.0211	0.009128	2.308	9691	9291	6009103
D(*Arctodus simus*, *Arctotherium* sp., *Ursus maritimus* (PB1))	0.0187	0.009374	1.99	9677	9323	6019285
D(*Arctodus simus*, *Arctotherium* sp., *Ursus arctos*)	0.0167	0.009247	1.802	9717	9398	6033406
D(*Arctodus simus*, *Arctotherium* sp., *Ursus maritimus* (PB9))	0.016	0.009483	1.683	9752	9446	6044167
D(*Arctodus simus*, *Arctotherium* sp., *Ursus ursinus*)	0.016	0.009584	1.666	9608	9306	6012634
D(*Arctodus simus*, *Arctotherium* sp., *Ursus thibetanus*)	0.0157	0.009487	1.652	9644	9346	6019681

*Significantly positive d-statistic, denote deviations from a typical bifurcating tree with H1 and H3 being closer than expected

If the ancestors of the spectacled bear hybridised with *Arctotherium* somewhere in the American mid-latitudes during the migration of *Tremarctos* into South America, then signals of gene flow between members of these two genera could date to the latest Pleistocene or earliest Holocene, when spectacled bears are thought to have migrated into South America [6, 35, 36]. To test this hypothesis, we estimated divergence times among the three short-faced bear lineages for all 500 kb windows from the largest 40 scaffolds corresponding to either Topology 1 (n = 980) or Topology 2 (n = 413) and summarised the results (Figure 2c). The age of the most recent common ancestor (TMRCA) of *Tremarctos, Arctodus, and Arctotherium* was similar irrespective of topology (Topology 1: 3.6 mya; Topology 2: 3.6 mya), as was the subsequent divergence between the remaining two lineages (Topology 1: 3.1 mya; Topology 2: 3.1 mya). Assuming that members of *Tremarctos* migrated southward no earlier than the latest Pleistocene, our results superficially appear to be incompatible with late

Pleistocene/Holocene hybridisation between *Tremarctos* and *Arctotherium*. The fossil record suggests two ways these observations may be explained.

Late Pleistocene fossil data indicate that the ancestors of the spectacled bear are likely to have encountered *Arctotherium* individuals from Mexico, Central America, and/or northern South America, which were comparable in size and diet to the spectacled bear [1, 5, 37] and which may have represented a different *Arctotherium* species from the Chilean specimen sequenced in the present study [1, 6, 12, 26]. Indeed, throughout the Pleistocene a number of *Arctotherium* species have been described across South and Central America, with putative species ranging from gigantic in the early-mid Pleistocene to relatively small in the late Pleistocene [1, 2, 6, 9]. If the ancestors of our sampled Patagonian *Arctotherium* specimen diverged from those of more northerly *Arctotherium* species during the Pliocene or early Pleistocene, then our molecular dating results remain consistent with hybridisation being the primary driver of phylogenetic discordance in our genomic data. Alternatively, hybridisation between *Tremarctos* and *Arctotherium* could have occurred in Central America during the Pleistocene. *Tremarctos* and *Arctotherium* have both been recorded in Central American cave deposits [5, 38], however, the extent of occupation by both genera in the region is unknown, and conceivably Central America represents a contact zone between the genera throughout the Pleistocene where hybridisation may have occurred.

An alternative interpretation of our phylogenetic results is that Topology 2 (*Tremarctos* + *Arctotherium*), which is supported by the mitochondrion and ~30% of our nuclear genome windows, is the pre-hybridisation tree. Recently, Li, et al. [39] suggested that under scenarios involving substantial gene flow the predominant phylogenetic signal across the genome may not reflect the pre-hybridisation tree. If this were the case for short-faced bears, the majority of support for Topology 1 would actually result from extensive hybridisation between *Arctodus* and *Tremarctos* in North America. Li, et al. [39] contend that the phylogenetic signal of the pre-hybridisation tree may be enriched in regions of low recombination, especially on the X-chromosome. In order to test this hypothesis, we identified panda scaffolds corresponding to the ~40 Mb recombination cold-spot on the X-chromosome highlighted by Li, et al. [39] and produced phylogenetic trees for each 500 kb window along this region (Figure S4). Interestingly, the majority of these fragments supported Topology 2 (*Tremarctos* + *Arctotherium*), the same topology as the mitochondrial phylogeny but contrasting with the majority of autosomal scaffolds.

Unlike felids [e.g. 40, 41, 42], a high-quality reference assembly and linkage map does not exist for any bear species, meaning scaffolds pertaining to high and low recombination areas of the genome could not be identified. Unfortunately, this currently makes it impossible to further explore the possibility that Topology 2 (*Tremarctos* + *Arctotherium*) may reflect the pre-hybridisation short-faced bear tree, rather than Topology 1 (*Tremarctos* + *Arctodus*). In the absence of a linkage map, sequencing aDNA from either the extinct *Tremarctos floridanus* or more northerly *Arctotherium* populations will be key to further resolving the evolutionary history of short-faced bears, though this will be challenging given that the core range of these species lies in the lower-latitudes where aDNA preservation is less reliable. For now we conclude that the weight of evidence supports a closer relationship between the spectacled bear and the extinct short-faced bears from North America (*Arctodus*) rather than South America (*Arctotherium*). In any case, our genomic data imply extensive hybridisation occurred between the spectacled bear and one of the extinct short-faced bear lineages. These results contribute to the growing consensus that hybridisation is widespread among carnivoran groups generally [13, 14, 39, 43].

Acknowledgements:

We would like to thank the following institutions for allowing access to specimens: Canadian Museum of Nature, University of Kansas Natural History Museum, Yukon Government, Centro de Estudios del Hombre Austral, Instituto de la Patagonia, Universidad de Magallanes. In addition, we are grateful to the following individuals who helped in the collection and identification of specimens and/or provided laboratory support: Grant Zazula (Yukon Territorial Government, Palaeontology Program, Canada), Fabiana Martin (Universidad de Magallanes, Chile), Jeremy Austin (University of Adelaide, Australia) and Sarah Bray (University of Adelaide, Australia). We would like to thank the Wyoming BLM and permit number PA-13-WY-207. This research was funded by an Australian Research Council Laureate Fellowship awarded to AC (FL140100260), U.S. National Science Foundation grant (EAR/SGP#1425059) awarded to JM and AC, and Agencia Nacional de Promoción Científica y Técnica' (ANPCyT, Argentina) (PICT 2015–966) awarded to FJP.

Author Contributions:

Conceptualization, A.T.S., A.C. and K.J.M.; Methodology, A.T.S., G.G., A.C., and K.J.M.; Investigation, A.T.S., H.H., and K.J.M.; Writing – Original Draft, A.T.S. and K.J.M.; Writing – Review & Editing, G.G., B.W.S., L.H.S., H.H., A.P., F.J.P., J.M., and A.C.; Funding

259 Acquisition, F.J.P., J.M., and A.C.; Resources, J.M., A.P., and F.J.P.; Supervision, A.C., and
260 K.J.M.

Materials and Methods:

Sampling

Analyses were performed on three bone samples identified as *Arctodus simus* and one sample identified as *Arctotherium* sp. (Table S1). The *Arctotherium* specimen ACAD 3599, had previously be radiocarbon dated, as well as one of the *A. simus* specimens (ACAD 438), a further *A. simus* specimen was radiocarbon dated at the Oxford Radiocarbon Accelerator Unit of the University of Oxford. All radiocarbon dates were calibrated with the either the IntCal13 curve [44] or the SHCal13 curve [45] using OxCal 4.4 [46] (Table S1).

Sample preparation and extraction

All pre-PCR steps (*i.e.*, extraction, library preparation) were conducted in purpose-built ancient DNA clean-room facilities at the University of Adelaide's Australian Centre for Ancient DNA (ACAD). Potential surface contamination on each sample was reduced by UV irradiation for 15 min each side, followed by abrasion of the exterior surface (c. 1 mm) using a Dremel tool and a disposable carborundum disk. The sample was then pulverised using a metallic mallet. Approximately 100 mg of powder was extracted using an in-house silica-based extraction protocol adapted from Dabney, et al. [47] optimised for the recovery of small fragments. the powder was digested first in 1 mL 0.5 M EDTA for 60 min, followed by an overnight incubation in 970 µL fresh 0.5 M EDTA and 30 µL proteinase K (20 mg/ml) at 55°C. The samples were centrifuged and the supernatant mixed with 13 mL of a modified PB buffer (12.6 mL PB buffer (Qiagen), 6.5 µL Tween-20, and 390 µL of 3M Sodium Acetate) and bound to silicon dioxide particles, which were then washed two times with 80% ethanol. The DNA was eluted from silica particles with 100 µL TE buffer.

Library preparation

Double-stranded Illumina libraries were constructed following the protocol of Meyer, et al. [48] from 25 µL of DNA extract. In addition, all samples underwent partial uracil-DNA glycosylase (UDG) treatment [49] to restrict cytosine deamination, characteristic of ancient DNA, to terminal nucleotides. A short round of PCR using PCR primers complementary to the library adapter sequences was performed to increase the total amount of DNA and add full-length Illumina sequencing adapters. Cycle number was determined via rtPCR and each library split into 8 separate PCR reactions to minimise PCR bias and maintain library complexity. Each PCR of 25 µL contained 1× HiFi buffer, 2.5 mM MgSO4, 1 mM dNTPs,

0.5 mM each primer, 0.1 U Platinum Taq Hi-Fi polymerase and 3 μL DNA. The cycling conditions were 94 °C for 6 min, 8–10 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 40 s, followed by 72 °C for 10 min. Following PCR, replicates were pooled and purified using AxyPrep™ magnetic beads, eluted in 30 μL H2O quantified on TapeStation (Agilent Technologies).

Sequencing

Libraries were initially pooled and sequenced on an Illumina NextSeq using 2 x 75 bp PE (150 cycle) High Output chemistry. For deeper sequencing, libraries were diluted to 1.5 nM and each was run on one lane of an Illumina HiSeq X Ten using 2 x 150 bp PE (300 cycle) chemistry, except for ACAD 438 which was run on two lanes of an Illumina HiSeq X Ten.

Data processing

Demultiplexed sequencing reads were processed through Paleomix v1.2.12 [50]. Within Paleomix, raw reads were filtered, adapter sequences removed, and pair-end reads merged using ADAPTER REMOVAL v2.1.7 [51], trimming low quality bases (<Phred20 --minquality 4) and discarding merged reads shorter than 25 bp (--minlength 25). Read quality was visualised before and after adapter trimming using fastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to ensure efficient adapter removal. Reads were mapped to the Panda ASM200744v1 genome [52] with BWA v0.7.15 using the mem algorithm [53]. Reads with mapping Phred scores less than 25 were removed using SAMtools 1.5 [54] and PCR duplicates were removed using "paleomix rmdup_collapsed" and MARKDUPLICATES from the Picard package (http://broadinstitute.github.io/picard/). Indel realignment was performed using GATK [55] and damage profiles assessed using MapDamage v2.0.8 [56] (Figure S1).

Sequencing reads were downloaded from the European Nucleotide Archive for all extant bear species (Table S2) [14, 21, 24, 57, 58] and processed using the same pipeline as for the ancient samples.

Phylogenetic analysis

Indexed VCF files were created for each BAM file using mpileup, part of the SAMtools package v0.1.19 [54], and the call and index functions as a part of the BCFtools package v0.1.19. Parallel v2010622 [59] was used to process each BAM file in parallel. BCFtools was

then used to filter SNPs within 3 bp of an indel (--SnpGap 3). The 85 largest scaffolds of the Panda reference genome were renamed as chromosomes (chr1–85) in each VCF file using BCFtools annotate. Biallelic variants in VCF files were converted to random pseudohaploid variants in eigenstrat format for the 85 largest scaffolds using vcf2eig (part of eig-utils; https://github.com/grahamgower/eig-utils) including monomorphic (-m) and singleton (-s) sites, and excluding transitions (-t). Eigenstrat formatted files were then converted to PHYLIP files using eig2phylip (part of eig-utils; https://github.com/grahamgower/eig-utils). A supermatrix tree was then created in RAxML v8.2.4 [60] using the rapid bootstrapping algorithm (-f a) and using the GTRCAT model of substitution with ascertainment correction (-m ASC_GTRCAT) with 100 bootstrap replicates (-#100) and using the Felsenstein ascertainment correction (--asc-corr=felsenstein) based on the number of invariant sites (calculated from the total ungapped length of the largest 85 scaffolds of the Panda reference genome minus the length of the alignment).

Discordance analysis using DiscoVista

The eigenstrat files were broken down into non-overlapping 500kb sliding windows using eigreduce (part of eig-utils; https://github.com/grahamgower/eig-utils). For each window a PHYLIP file and tree were created as described above. The frequency and support of different tree topologies was then summarised and visualised using DiscoVista [61], using bootstrap values of 80 as the cutoff for strong support. Topologies tested included: 1) the inclusion of *Arctotherium* with ursine bears; 2) the inclusion of *Arctodus* with ursine bears; 3) the inclusion of *Tremarctos* with ursine bears; 4) any combination of tremarctine bears included with ursine bears; 5) the monophyly of Tremarctinae; 6) monophyly of *Tremarctos* and *Arctodus*; 7) monophyly of *Tremarctos* and *Arctotherium*; and 8) monophyly of *Arctotherium* and *Arctodus*.

D-statistics

To test for signals of gene flow within Tremarctinae and between tremarctine and ursine lineages we used D-statistics as implemented by Admixtools [62] in admixr [63]. We only used the higher coverage *A. simus* sample (ACAD 344) in this analysis. The giant panda was used as outgroup and block jack-knife procedure used to test for significant departures from zero (|Z|>3). D-statistics within Tremarctinae were calculated in the form D(*Arctodus*, *Tremarctos*, *Arctotherium*, panda) and for detecting gene flow between Tremarctinae and Ursinae in the form D(U1, U2, T1, panda), where T1 is any short-faced bear and U1 and U2

any ursine individual. D-statistics were also performed to detect gene flow within Ursinae (as per Kumar, et al. 2017), using either the giant panda or spectacled bear as outgroup. To account for the possibility of a reference bias in ancient samples, within Tremarctinae D-statistics were recalculated using the Asiatic black bear as outgroup (Table S5).

Molecular dating

Divergence times were estimated for each 500kb fragment from the discordance analysis using MCMCtree, part of the PAML package v4.8a [64], using the topology from the ML tree produced in the discordance analysis as the input tree. Four calibrations were used to calibrate the phylogeny:

1. The crown-age of Ursidae (*i.e.* the divergence of the giant panda lineage) was constrained to between 11.6 and 23 million years ago (mya) based on the presence of *Kretzoiarctos* [65], a putative ailuropodine, in the middle Miocene and the assumption that early Miocene *Ursavus* representatives are likely ancestral to modern ursids [10].
2. The divergence of Tremarctinae and Ursinae was constrained to between 7 and 13 mya based on the presence of putative early tremarctine bears (*e.g. Plionarctos*) in the Late Miocene/Early Pliocene [66].
3. The common ancestor of all sampled ursine bears was constrained to between 4.3 and 6 mya based on the occurrence of *Ursus minimus* [14, 67].
4. The divergence of polar and brown bears was constrained to between 0.48 and 1.1 mya based on previous nuclear estimates [21, 24, 25].

The JC +G substitution model with 5 discrete gamma categories was used with autocorrelated-rates, also known as the geometric Brownian diffusion clock model. Uniform priors for node ages using the birth-death (BD) process were used [λ_{BD} = 1 (birth-rate), μ_{BD} = 1 (death-rate), and ρ_{BD} = 0.1 (sampling fraction for extant species)]. A gamma-Dirichlet distribution was used for the prior on rate with an α shape parameter of 2 (diffuse prior). The σ_i^2 prior was defined as a diffuse gamma-Dirichlet distribution (2,2). MCMC tree runs were performed with a burn-in of 10000, and a sample size of 10000, sampling every ten iterations. Median node ages were then averaged for each tree topology.

Low-recombining region of X-chromosome

Scaffolds of the panda ASM200744v1 reference genome [52] corresponding to low recombination regions of the X chromosome were identified by mapping all scaffolds to the recombination cold-spot of the X chromosome of the domestic cat (FelCat5) using minimap2. Default parameters were used, meaning the alignment lacked base-level precision (to account to phylogenetic distance between giant panda and the domestic cat). Only scaffolds larger than 500kb and with greater than 100 kb of segments mapping to the low recombining region of the domestic cat X-chromosome were retained, resulting in 15 scaffolds linked to the low recombination region of the X-chromosome. A maximum-likelihood phylogenetic tree and gene-tree discordance analysis were performed on these 15 scaffolds as described above for the genome-wide dataset.

References:

1. Soibelzon, L.H., Tonni, E.P., and Bond, M. (2005). The fossil record of South American short-faced bears (Ursidae, Tremarctinae). J. South Am. Earth Sci. *20*, 105-113.

2. Soibelzon, L.H. (2004). Revisión sistemática de los Tremarctinae (Carnivora, Ursidae) fósiles de América del Sur. Rev. Mus. Argent. Cienc. Nat *6*, 107-133.

3. Kurtén, B. (1966). Pleistocene bears of North America. 1. Genus *Trematctos*, spectacled bears. Acta Zool. Fenn. *115*, 1-120.

4. Kurtén, B., and Anderson, E. (1980). Pleistocene Mammals of North America, (New York: Columbia University Press).

5. Schubert, B.W., Chatters, J.C., Arroyo-Cabrales, J., Samuels, J.X., Soibelzon, L.H., Prevosti, F.J., Widga, C., Nava, A., Rissolo, D., and Erreguerena, P.L. (2019). Yucatan carnivorans shed light on the Great American Biotic Interchange. Biol Lett *15*, 20190148.

6. Prevosti, F.J. (2018). Evolution of South American Mammalian Predators During the Cenozoic: Paleobiogeographic and Paleoenvironmental Contingencies, 1st ed. 2018. Edition, (Cham: Springer International Publishing).

7. Kurtén, B. (1967). Pleistocene bears of North America. 2. Genus *Arctodus*, short-faced bears. Acta Zool. Fenn. *117*, 1-60.

8. Christiansen, P. (1999). What size were *Arctodus simus* and *Ursus spelaeus* (Carnivora: Ursidae)? Ann. Zool. Fenn. *36*, 93-102.

9. Soibelzon, L.H., and Schubert, B.W. (2011). The largest known bear, *Arctotherium angustidens*, from the early Pleistocene Pampean region of Argentina: with a discussion of size and diet trends in bears. J. Paleontol. *85*, 69-75, 67.

10. McLellan, B., and Reiner, D.C. (1994). A review of bear evolution. Bears Their Biol. Manag. *9*, 85-96.

11. Trajano, E., and Ferrarezzi, H. (1995). A Fossil Bear from Northeastern Brazil, with a Phylogenetic Analysis of the South American Extinct Tremarctinae (Ursidae). J. Vertebr. Paleontol. *14*, 552-561.

12. Mitchell, K.J., Bray, S.C., Bover, P., Soibelzon, L., Schubert, B.W., Prevosti, F., Prieto, A., Martin, F., Austin, J.J., and Cooper, A. (2016). Ancient mitochondrial DNA reveals convergent evolution of giant short-faced bears (Tremarctinae) in North and South America. Biol. Lett. *12*, 20160062.

13. Kutschera, V.E., Bidon, T., Hailer, F., Rodi, J.L., Fain, S.R., and Janke, A. (2014). Bears in a forest of gene trees: Phylogenetic inference is complicated by incomplete lineage sorting and gene flow. Mol. Biol. Evol. *31*, 2004-2017.

14. 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.

15. Funk, D.J., and Omland, K.E. (2003). Species-Level Paraphyly and Polyphyly: Frequency, Causes, and Consequences, with Insights from Animal Mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. *34*, 397-423.

16. Leache, A.D., and McGuire, J.A. (2006). Phylogenetic relationships of horned lizards (*Phrynosoma*) based on nuclear and mitochondrial data: evidence for a misleading mitochondrial gene tree. Mol. Phylogenet. Evol. *39*, 628-644.

17. Wallis, G.P., Cameron-Christie, S.R., Kennedy, H.L., Palmer, G., Sanders, T.R., and Winter, D.J. (2017). Interspecific hybridization causes long-term phylogenetic discordance between nuclear and mitochondrial genomes in freshwater fishes. Mol. Ecol. *26*, 3116-3127.

18. Toews, D.P., and Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. *21*, 3907-3930.

19. Gompert, Z., Forister, M.L., Fordyce, J.A., and Nice, C.C. (2008). Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in *Lycaeides*. Mol. Ecol. *17*, 5231-5244.

20. 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.

21. Cahill, J.A., Green, R.E., Fulton, T.L., Stiller, M., Jay, F., Ovsyanikov, N., Salamzade, R., St. John, J., Stirling, I., Slatkin, M., et al. (2013). Genomic evidence for island

population conversion resolves conflicting Theories of polar bear evolution. PLoS Genet. *9*, e1003345.

22. Cahill, J.A., Stirling, I., Kistler, L., Salamzade, R., Ersmark, E., Fulton, T.L., Stiller, M., Green, R.E., and Shapiro, B. (2015). Genomic evidence of geographically widespread effect of gene flow from polar bears into brown bears. Mol. Ecol. *24*, 1205-1217.

23. Hailer, F. (2015). Introgressive hybridization: brown bears as vectors for polar bear alleles. Mol. Ecol. *24*, 1161-1163.

24. Liu, S.P., Lorenzen, E.D., Fumagalli, M., Li, B., Harris, K., Xiong, Z.J., Zhou, L., Korneliussen, T.S., Somel, M., Babbitt, C., et al. (2014). Population genomics reveal recent speciation and rapid evolutionary adaptation in polar bears. Cell *157*, 785-794.

25. Hailer, F., Kutschera, V.E., Hallstrom, B.M., Klassert, D., Fain, S.R., Leonard, J.A., Arnason, U., and Janke, A. (2012). Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. Science *336*, 344-347.

26. Martin, F.M., Prieto, A., Morello, F., Prevosti, F., and Borrero, L. (2004). Late Pleistocene megafauna at Cueva del Puma, Pali-Aike Lava Field, Chile. Curr. Res. Pleistocene *21*, 101-103.

27. Harington, C.R., Naughton, D., Dalby, A., Rose, M., and Dawson, J. (2003). Annotated Bibliography of Quaternary Vertebrates of Northern North America, (Toronto: University of Toronto Press).

28. Krause, J., Unger, T., Nocon, 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.

29. Nichols, R. (2001). Gene trees and species trees are not the same. Trends Ecol. Evol. *16*, 358-364.

30. Maddison, W.P. (1997). Gene trees in species trees. Syst. Biol. *46*, 523-536.

31. Cahill, J.A., Heintzman, P.D., Harris, K., Teasdale, M.D., Kapp, J., Soares, A.E.R., Stirling, I., Bradley, D., Edwards, C.J., Graim, K., et al. (2018). Genomic evidence of widespread admixture from polar bears into brown bears during the last ice age. Mol. Biol. Evol. *35*, 1120-1129.

32. Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W.W., Fritz, M.H.Y., et al. (2010). A Draft Sequence of the Neandertal Genome. Science *328*, 710-722.

33. Durand, E.Y., Patterson, N., Reich, D., and Slatkin, M. (2011). Testing for ancient admixture between closely related populations. Mol. Biol. Evol. *28*, 2239-2252.

34. Mondolfi, E., and Boede, E. (1981). A hybrid of a spectacled bear (*Tremarctos ornatus*) and an asiatic black bear (*Selenarctos thibetanus*) born at the Maracay zoological park, Venezuela. Mem. Soc. Cienc. Nat. La Salle *41*, 143-148.

35. Stucchi, M., Salas-Gismondi, R., Baby, P., Guyot, J.-L., and Shockey, B.J. (2009). A 6,000 year-old specimen of a spectacled bear from an Andean cave in Peru. Ursus *20*, 63-68.

36. García-Rangel, S. (2012). Andean bear *Tremarctos ornatus* natural history and conservation. Mammal Rev. *42*, 85-119.

37. Figueirido, B., and Soibelzon, L.H. (2010). Inferring palaeoecology in extinct tremarctine bears (Carnivora, Ursidae) using geometric morphometrics. Lethaia *43*, 209-222.

38. Czaplewski, N.J., Krejca, J., and Miller, T.E. (2003). Late quaternary bats from Cebada Cave, Chiquibul cave system, Belize. Caribb. J. Sci. *39*, 23-33.

39. Li, G., Figueiro, H.V., Eizirik, E., and Murphy, W.J. (2019). Recombination-aware phylogenomics reveals the structured genomic landscape of hybridizing cat species. Mol. Biol. Evol. *36*, 2111-2126.

40. Menotti-Raymond, M., David, V.A., Roelke, M.E., Chen, Z.Q., Menotti, K.A., Sun, S., Schaffer, A.A., Tomlin, J.F., Agarwala, R., O'Brien, S.J., et al. (2003). Second-generation integrated genetic linkage/radiation hybrid maps of the domestic cat (*Felis catus*). J. Hered. *94*, 95-106.

41. Menotti-Raymond, M., David, V.A., Lyons, L.A., Schaffer, A.A., Tomlin, J.F., Hutton, M.K., and O'Brien, S.J. (1999). A genetic linkage map of microsatellites in the domestic cat (*Felis catus*). Genomics *57*, 9-23.

42. Li, G., Hillier, L.W., Grahn, R.A., Zimin, A.V., David, V.A., Menotti-Raymond, M., Middleton, R., Hannah, S., Hendrickson, S., Makunin, A., et al. (2016). A high-resolution SNP array-based linkage map anchors a new domestic cat draft genome assembly and provides detailed patterns of recombination. G3 (Bethesda) *6*, 1607-1616.

43. Li, G., Davis, B.W., Eizirik, E., and Murphy, W.J. (2016). Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). Genome Res. *26*, 1-11.

44. Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Ramsey, C.B., Buck, C.E., Cheng, H., Edwards, R.L., Friedrich, M., et al. (2013). Intcal13 and Marine13 radiocarbon age calibration curves 0-50,000 years cal BP. Radiocarbon *55*, 1869-1887.

45. Hogg, A.G., Hua, Q., Blackwell, P.G., Niu, M., Buck, C.E., Guilderson, T.P., Heaton, T.J., Palmer, J.G., Reimer, P.J., Reimer, R.W., et al. (2013). Shcal13 Southern Hemisphere Calibration, 0-50,000 Years Cal Bp. Radiocarbon *55*, 1889-1903.

46. Ramsey, C.B. (2009). Bayesian analysis of radiocarbon dates. Radiocarbon *51*, 337-360.

47. Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., et al. (2013). 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.

48. Meyer, M., Kircher, M., Gansauge, M.T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay, F., Prufer, K., de Filippo, C., et al. (2012). A high-coverage genome sequence from an archaic Denisovan individual. Science *338*, 222-226.

49. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., and Reich, D. (2015). Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. Philos. Trans. R. Soc. Lond. B. Biol. Sci *370*, 20130624.

50. Schubert, M., Ermini, L., Sarkissian, C.D., Jonsson, H., Ginolhac, A., Schaefer, R., Martin, M.D., Fernandez, R., Kircher, M., McCue, M., et al. (2014). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. Nat. Protoc. *9*, 1056-1082.

51. Schubert, M., Lindgreen, S., and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res. Notes *9*, 88.

52. Hu, Y., Wu, Q., Ma, S., Ma, T., Shan, L., Wang, X., Nie, Y., Ning, Z., Yan, L., Xiu, Y., et al. (2017). Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. Proc. Natl. Acad. Sci. U. S. A. *114*, 1081-1086.

53. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

54. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078-2079.

55. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297-1303.

56. Jonsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics *29*, 1682-1684.

57. Zhao, S.C., Zheng, P.P., Dong, S.S., Zhan, X.J., Wu, Q., Guo, X.S., Hu, Y.B., He, W.M., Zhang, S.N., Fan, W., et al. (2013). Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation. Nat. Genet. *45*, 67-U99.

58. Miller, W., Schuster, S.C., Welch, A.J., Ratan, A., Bedoya-Reina, O.C., Zhao, F.Q., Kim, H.L., Burhans, R.C., Drautz, D.I., Wittekindt, N.E., et al. (2012). Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. Proc. Natl. Acad. Sci. U. S. A. *109*, E2382-E2390.

59. Tange, O. (2011). Gnu parallel-the command-line power tool. The USENIX Magazine *36*, 42-47.

60. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312-1313.

584 61. Sayyari, E., Whitfield, J.B., and Mirarab, S. (2018). DiscoVista: Interpretable
585 visualizations of gene tree discordance. Mol. Phylogenet. Evol. *122*, 110-115.

586 62. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck,
587 T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics
588 *192*, 1065-1093.

589 63. Petr, M., Vernot, B., and Kelso, J. (2019). admixr-R package for reproducible analyses
590 using ADMIXTOOLS. Bioinformatics *35*, 3194-3195.

591 64. Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol.
592 Evol. *24*, 1586-1591.

593 65. Abella, J., Alba, D.M., Robles, J.M., Valenciano, A., Rotgers, C., Carmona, R.,
594 Montoya, P., and Morales, J. (2012). *Kretzoiarctos* gen. nov., the oldest member of the
595 giant panda clade. PLoS ONE *7*, e48985.

596 66. Tedford, R.H., and Martin, J. (2001). *Plionarctos*, a tremarctine bear (Ursidae:
597 Carnivora) from western North America. J. Vertebr. Paleontol. *21*, 311-321.

598 67. Vangengeim, E.A., Vislobokova, I., and Sotnikova, M. (1998). Large Ruscinian
599 Mammalia in the territory of the former Soviet Union. Stratigr. Geol. Correl *6*, 368-382.
600