PROCEEDINGS B

Research

Author for correspondence:
Charlotte Lindqvist
e-mail: cl243@buffalo.edu

THE ROYAL SOCIETY
PUBLISHING

Evolutionary history of enigmatic bears in the Tibetan Plateau–Himalaya region and the identity of the yeti

Tianying Lan[1], Stephanie Gill[1], Eva Bellemain[2], Richard Bischof[3], Muhammad Ali Nawaz[4,5] and Charlotte Lindqvist[1,6]

[1]Department of Biological Sciences, University at Buffalo (SUNY), Buffalo, NY 14260, USA
[2]SPYGEN, Savoie Technolac - BP 274, Le Bourget-du-Lac Cedex 73375, France
[3]Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, PO Box 5003, 1432 Ås, Norway
[4]Department of Animal Sciences, Quaid-i-Azam University, Islamabad, Pakistan
[5]Snow Leopard Trust, 4649 Sunnyside Ave N, Suite 325, Seattle, WA 98103, USA
[6]School of Biological Sciences, Nanyang Technological University, Singapore 637551

(iD) CL, 0000-0002-4190-727X

Although anecdotally associated with local bears (*Ursus arctos* and *U. thibetanus*), the exact identity of 'hominid'-like creatures important to folklore and mythology in the Tibetan Plateau–Himalaya region is still surrounded by mystery. Recently, two purported yeti samples from the Himalayas showed genetic affinity with an ancient polar bear, suggesting they may be from previously unrecognized, possibly hybrid, bear species, but this preliminary finding has been under question. We conducted a comprehensive genetic survey of field-collected and museum specimens to explore their identity and ultimately infer the evolutionary history of bears in the region. Phylogenetic analyses of mitochondrial DNA sequences determined clade affinities of the purported yeti samples in this study, strongly supporting the biological basis of the yeti legend to be local, extant bears. Complete mitochondrial genomes were assembled for Himalayan brown bear (*U. a. isabellinus*) and black bear (*U. t. laniger*) for the first time. Our results demonstrate that the Himalayan brown bear is one of the first-branching clades within the brown bear lineage, while Tibetan brown bears diverged much later. The estimated times of divergence of the Tibetan Plateau and Himalayan bear lineages overlap with Middle to Late Pleistocene glaciation events, suggesting that extant bears in the region are likely descendants of populations that survived in local refugia during the Pleistocene glaciations.

1. Introduction

The Tibetan Plateau, the most extensive and highest plateau in the world with an average altitude of 4500 m above sea level, is partly surrounded by the Himalayan range and many of Earth's highest mountains. Dramatic environmental changes caused by the uplift of the plateau and climatic oscillations during the Quaternary glaciations substantially impacted the evolution, diversification, and distribution of local plant and animal species [1]. Because of its heterogeneous habitat and topography, the region sustains a distinct biome with rich biological diversity and high level of endemism [2]. Extant plants and animals on the plateau are likely either descendants of relict colonists that migrated from other areas or recently derived endemic species [3–10]. However, the colonization and population expansion history of many species remains poorly understood, despite current and future impacts of climate change and anthropogenic threats to diversity loss.

Two brown bear subspecies, the Himalayan (*Ursus arctos isabellinus*) and the Tibetan (*U. a. pruinosus*) brown bear, inhabit the northwestern Himalayan
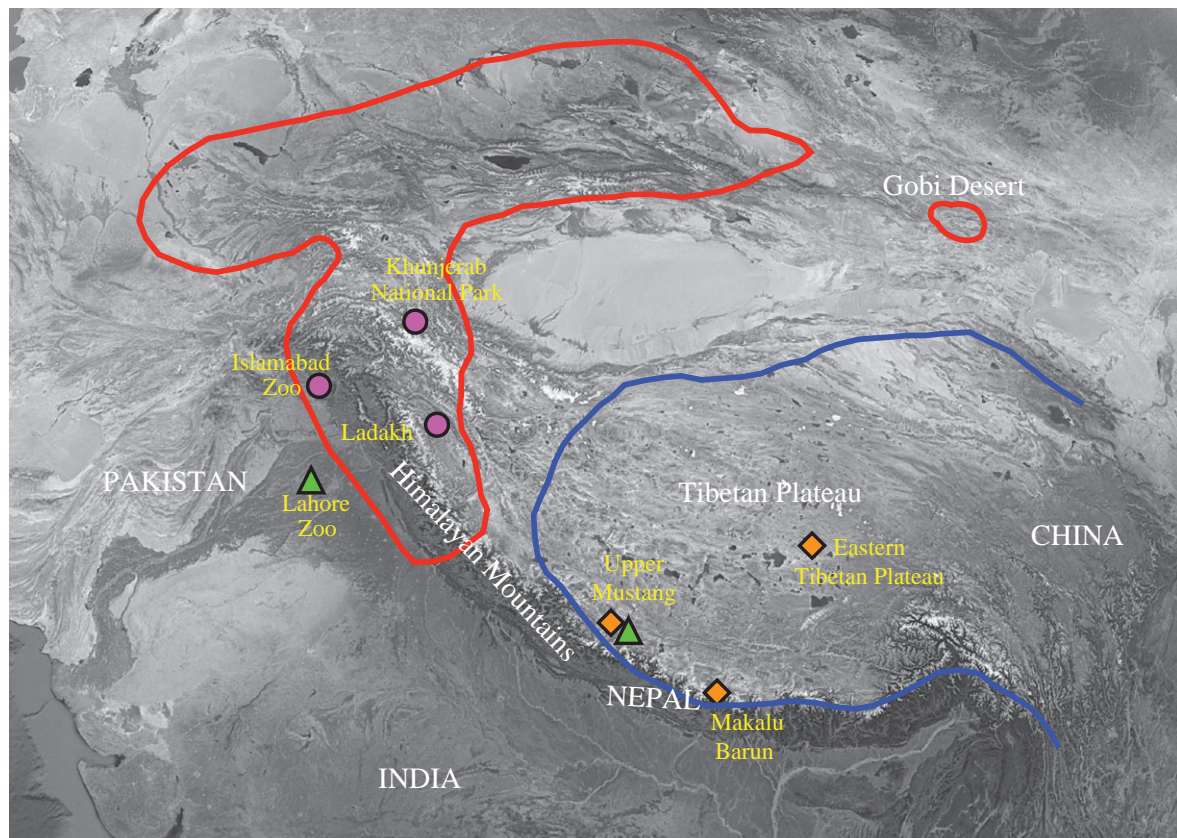
Figure 1. Distribution of Himalayan and Tibetan brown bear and localities of samples studied. Red and blue lines outline the approximate historical range of the Himalayan brown bear and the Tibetan brown bear, respectively (redrawn from Galbreath *et al.* [15]). The triangles, diamonds and circles, respectively, indicate the approximate collecting localities of the studied samples associated with Asian black bear, Tibetan brown bear and Himalayan brown bear.

region and southeastern Tibetan Plateau, respectively [11–14] (figure 1). These two subspecies have distinct skull features and the Himalayan brown bear is characterized by its paler and reddish-brown fur, while the Tibetan brown bear has generally darker fur with a developed, white 'collar' around the neck [11]. As the most widely distributed bear in the world, phylogeography of the brown bear has been well studied in North America, Europe and Japan [10,16–24]. However, due to limited sampling, very few studies have been conducted on these enigmatic subspecies. Two complete mitochondrial genomes (mitogenomes) from captive Tibetan brown bears are available, while only two short fragments of mitochondrial DNA (mtDNA) sequences from the Himalayan brown bear have been published [10,15]. Phylogenetic analyses based on these sequences suggested that the Tibetan brown bear might be a relict population of the Eurasian brown bear [10], and that the Himalayan brown bear, which is genetically distinct from the Tibetan brown bear, may represent a more ancient lineage [15]. However, phylogenetic relationships deduced from limited genetic data and number of individuals have put these preliminary findings into question. For example, the phylogenetic placement of a Gobi brown bear (*U. a. gobiensis*) sequence [25] was inconsistent with a later study also including sequences from Himalayan brown bear [15], and phylogenetic trees based on mtDNA control region and cytochrome *b* sequences, respectively, of the Tibetan brown bear were incongruent [26]. The other bear species found to inhabit the Tibetan Plateau–Himalaya region is the Asian black bear (*U. thibetanus*), which historically had a continuous distribution from southeastern Iran through Afghanistan and Pakistan to India, Nepal, China, Korea, Japan,

and south into Myanmar and the Malayan peninsula [12,27,28]. Today it occupies a patchy distribution throughout its historic range, including across a narrow band from Pakistan, Kashmir and to Bhutan, the home range of the Himalayan black bear (*U. t. laniger*) [27,29], which was described as distinguished from other black bear populations by its longer, thicker fur and smaller, whiter chest mark [11]. Although the range of Asian black bear overlaps with brown bear in the Tibetan Plateau–Himalaya region, it is mostly found at lower altitudes in forested hills ranging from 1200 to 3300 m [12,29]. So far, little is known about the evolutionary history of black bear in the region and no sequence data are available from the Himalayan black bear. To elucidate the evolutionary and migration history of the Himalayan and Tibetan bears, more genetic data from additional individuals are critically needed.

It has been reported that the brown bear populations in the Tibetan Plateau–Himalaya region have declined by more than half in the past century because of habitat loss, fragmentation, poaching and intense hunting by humans [12,29–31]. Facing the same threats as brown bears, Asian black bear populations have also decreased in the past few decades [29,32,33]. The Himalayan brown bear is listed in the IUCN (International Union for the Conservation of Nature) red list of threatened species as critically endangered [34], while the Asian black bear is listed as vulnerable [27]. Hence, clarifying population structure and genetic diversity for conservation management purposes is also urgently needed for these endangered bear species.

The Tibetan Plateau–Himalaya region is also known for the legend of purported 'hominid'-like creatures, referred to

as the 'yeti', 'chemo', 'mheti' or 'bharmando', among other regional monikers (for simplicity they are referred to in this paper as yeti). Despite decades of research and anecdotal association with bears and other mammals in the region [35,36], the species identity of the mysterious yeti is still debated, given the lack of conclusive evidence. A survey of hair samples attributed to yeti and other anomalous, supposed primates, was recently conducted to identify their genetic affinities [37]. Based on a short fragment of the mtDNA 12S rRNA gene from two samples collected in Ladakh, India and Bhutan, respectively, and a 100% match to a sequence recovered from a subfossil polar bear [38], Sykes *et al.* [37] speculated that an unclassified bear species or hybrid of polar bear and brown bear might be present in the Tibetan Plateau–Himalaya region. However, this speculation was critiqued by others [39,40], and their phylogenetic analyses using the sequences from Sykes *et al.* and other available Ursidae sequences did not rule out the possibility that the samples belonged to brown bear. Thus, to get accurate species identification, comprehensive phylogenetic analyses using genetic information from more variable and informative loci are needed.

Here, we report on new analyses of 24 field-collected and museum specimens, including hair, bone, skin and faecal samples, collected from bears or purported yetis in the Tibetan Plateau–Himalaya region. Based on both amplified mtDNA loci as well as complete mitogenomes, we reconstructed maternal phylogenies to increase knowledge about the phylogenetic relationships and evolutionary history of Himalayan and Tibetan bears.

2. Material and methods

(a) Samples

A total of 24 samples, including hair, tissue, bone and faeces, were analysed in this study (electronic supplementary material, table S1). Of these, 12 samples had been collected for a previous analysis of Himalayan brown bear in the Khunjerab National Park, Northern Pakistan [30], two samples were from purported Himalayan brown bears housed in the Lahore and Islamabad Zoos, one bone sample (M-70448) recorded as *U. a. pruinosus* was obtained from the American Museum of Natural History, and nine samples were provided to us by the Reinhold Messner Museum and the Icon Film Company.

(b) DNA extraction

Genomic DNA from 12 faecal samples collected in the Khunjerab National Park, Northern Pakistan [41], were previously extracted using the QIAmp DNA Stool Kit (Qiagen, USA) in a room dedicated to processing hairs and faeces [30]. DNA from two ethanol-preserved hair samples from Lahore and Islamabad Zoos were isolated in a room dedicated to nucleic acid extraction from modern samples. A DNeasy Blood & Tissue DNA Kit (Qiagen, USA) was used according to the manufacturer's protocol, except for the following modifications to optimize extraction of DNA from hair: 10 strands of hair from each sample were cut into fragments of approximately 0.5 cm with a sterile razor blade. Ethanol was allowed to evaporate (approx. 1 h), and hair fragments were transferred to a microcentrifuge tube. Three hundred microlitres of ATL buffer, 20 μl proteinase K, 20 μl 1M DTT (dithiothreitol) and 4 μl RNase A were added, and samples were incubated at 56°C overnight until completely lysed. A negative control was prepared alongside each hair sample. Following lysis, 300 μl AL buffer and 300 μl 100% EtOH were added to each sample, and the mixture

was pipetted into the DNeasy Mini Spin Column and centrifuged for 2 min. DNA was eluted twice with 50 μl AE buffer for a total elution volume of 100 μl. The remaining 10 samples, which had not been intentionally preserved for later extraction of DNA, were regarded as non-modern (ancient) samples, and thus DNA extractions and pre-amplifications were performed in a dedicated state-of-the-art cleanroom facility, physically separated from any modern DNA laboratory and appropriate for ancient DNA research. The following protocols designed for ancient DNA extraction were used: for bone samples, 50–100 mg fine bone powder was obtained from each sample by using a dental drill (HKM Surgical Handpiece, Pearson Dental, USA), and 50–100 mg skin samples were sliced into approximately 1 mm pieces with a sterile razor blade. DNA from the bone powder and the sliced skin samples was extracted using the protocol in Dabney *et al.* [42]. DNA from the hair samples were extracted using the protocol provided by Gilbert *et al.* [43] with the following modifications: 1 ml digestion buffer was used for each hair extraction. After purification with phenol and chloroform, additional purification was performed using Qiagen MinElute PCR Purification Kit (Qiagen, USA). Finally, a 12.5 μl EB buffer elution step was performed twice to obtain a total elution volume of 25 μl. DNA from approximately 100 mg faecal samples was extracted using the QIAmp DNA Stool Kit (Qiagen, USA). The final elution step was also performed twice to obtain a total volume of 100 μl. Negative controls were prepared alongside all extractions.

(c) PCR amplification

PCR amplifications from modern DNA were performed in a 25 μl reaction volume each containing 2.5 μl of $10 \times$ PCR buffer (Applied Biosystems, USA), 1.0 μl of dNTP mixture (2.5 mM each dNTP; Applied Biosystems), 2.5 μl of $MgCl_2$ (25 mM, Applied Biosystems), 0.1 μl of *Taq* DNA polymerase ($5–10 \text{ U } \mu l^{-1}$; Applied Biosystems, AmpliTaq Gold), 1 μl each of the forward and reverse primers (10 μM), 2 μl of the genomic DNA and 17.4 μl of H_2O. The PCR reaction mix for ancient DNAs was prepared in the cleanroom by adding 21 μl H_2O, 1 μl of each forward and reverse primer, and 2 μl genomic DNA to each GE illustra PuReTaq Ready-To-Go PCR bead (GE Healthcare, USA). A touchdown thermal cycling protocol was used as follows: 10 min at 94°C, 10 cycles of 30 s at 94°C, 30 s annealing with the temperature decreasing every cycle by 0.5°C from 55°C to 50°C, and 30 s extension at 72°C, followed by 25 cycles the annealing temperature set to 50°C and denaturation and extension phases as above. For samples of unknown identity, two sets of mtDNA 12S rRNA primers [44,45] were used to determine their approximate taxonomic affinity. Bear-specific primers targeting the mtDNA control region and cytochrome *b* ([46] and primers designed for this study; see electronic supplementary material, table S2) were used for samples identified as ursid bears. PCR products were Sanger sequenced directly using the same primers as in the PCR.

(d) Mitochondrial genome target enrichment and sequencing

Fifty microlitres of DNA extracts from four samples were sent to MYcroarray (http://www.mycroarray.com) for preparation of Ion Torrent sequencing libraries and mtDNA target enrichment and sequencing, using the following protocol. Sample libraries were quantified using spectrofluorometry, which indicated between 5 and 255 total nanograms ($0.2–8.5 \text{ ng } \mu l^{-1}$) of double-stranded DNA. Each library was then individually target enriched using a custom-designed ursid mitogenome bait set manufactured by MYcroarray. The standard MYbaits v. 3.0 protocol was applied with hybridization for 21 h at 60°C at all relevant steps. Following clean up, half of each bead-bound library was amplified in a 50 μl reaction with universal Ion Torrent adapter-primers for 10 cycles

using a KAPA HiFi premix (KAPA Biosystems) and the manufacturer's recommended thermal profile coupled with 62°C annealing temperature. After amplification, the beads were pelleted and the supernatant was purified using SPRI beads and eluted in Tris-HCl buffer containing 0.05% Tween-20. The enriched libraries were quantified with spectrofluorometry, which indicated between 1.12 and 4.21 total nanograms dsDNA per library $(0.03-0.12 \text{ ng } \mu l^{-1})$. Equal masses of each library were pooled, bead-templated and sequenced alongside other project libraries on the Ion Proton platform using the Ion PI Chip Kit v2 chemistry. Following sequencing, reads were de-multiplexed, quality trimmed and filtered using the default settings on the Ion Torrent Suite v. 4.4.3.

(e) Mitochondrial genome assembly

Assembly of mitochondrial genomes was performed using the following strategy: species-specific mitochondrial reference genomes were selected from initial species identification based on phylogenetic analyses of amplicon sequences (results not shown). All Ion Torrent reads were first aligned against the reference genome using BWA *aln* (v. 0.7.13) [47] using the default parameters, except for the parameter 'l 1024' to disable the seed and increase high-quality hits for the damaged ancient DNA reads [48]. The remaining unmapped reads were then aligned against the same reference using BWA *mem* with default parameters (see electronic supplementary material, table S3, for assembly statistics). We filtered for human contamination by applying an edit-distance based strategy [48]. All reads were mapped to a human mitochondrial genome reference (NCBI accession J01415.2) using the same BWA mapping method described above. Reads with a higher mapping edit-distance to human mtDNA than to bear mitochondrial genomes were considered of likely human origin and were removed from the bear mitogenome mapping results. PCR duplicates were removed with the MarkDuplicates tool in the Picard software suite v. 1.112 using lenient validation stringency (http://broadinstitute.github.io/picard/). Consensus calling was carried out using Samtools *mpileup* [49] with default settings.

(f) Phylogenetic analyses

Complete mitochondrial genomes, partial control region sequences, and cytochrome *b* sequences for 11 Asian black bears, 76 American black bears, two cave bears (*U. spelaeus*), 200 brown bears, and 52 polar bears were obtained from GenBank (electronic supplementary material, table S4). Two GenBank datasets were created: one dataset included only complete mitogenomes for the non-Tibetan/Himalayan bears and both partial (amplicon sequences) and complete mitogenomes for Tibetan and Himalayan bear lineages, while the other dataset included both amplicon sequences and complete mitogenomes for non-Tibetan/Himalayan bears. All new sequences produced in this study were added to these two GenBank datasets and used in the phylogenetic analyses. Sloth bear (*U. ursinus*) and sun bear (*U. malayanus*) sequences were included to root the trees (electronic supplementary material, table S4). Alignments were generated using MAFFT [50] followed by manual adjustment in BioEdit [51] to exclude the variable number tandem repeats of the D-loop. The total length of the final alignment was 16 412 bp. Maximum-likelihood (ML) phylogenetic analyses were performed using RAxML-HPC BlackBox v. 8.2.8 [52] in the CIPRES Science Gateway under the GTR substitution model, which was identified as the best-supported model by jmodeltest2 [53,54]. A total of 1000 bootstrap replicates were conducted to evaluate branch support. Bayesian inference (BI) phylogenetic analyses were carried out using MrBayes v. 3.2.6 [55] in two runs of 5 000 000 Markov chain Monte Carlo (MCMC) generations, with trees for estimation of the posterior probability distribution sampled every 100 generations. The best-fit substitution model was determined by the

program by setting Nst=mixed; 500 000 trees were discarded as burn-in.

(g) Divergence time estimation

Bayesian MCMC-based divergence time estimation was carried out using BEAST version 1.8.0 under the GTR substitution model. The dataset used for molecular dating analysis included only complete mitogenome, since shorter mtDNA regions (e.g. control region and cytochrome *b*) are generally associated with considerable uncertainty and may bias molecular dating analyses due to homoplasy [10,17]. The uncorrelated lognormal relaxed clock and the constant size coalescent prior were used. Radiocarbon dates and stratigraphically estimated dates for four ancient sequences were used to calibrate ages for terminal nodes, including three sequences from extinct bear species (*U. spelaeus* and *U. deningeri*) dated to 31.8 thousand years (ka) BP [56], 44.1 ka BP [57], and 409 ka BP [42], an approximately 120 ka BP polar bear subfossil [38], and seven European brown bears dated to approximately 4.1–37 ka BP [58]. Trees were sampled every 1000 generations from a total of 1 000 000 000 generations. The maximum clade credibility tree was generated using TreeAnnotator, implemented in the BEAST package [59], with 10% burn-in. Effective sampling size value greater than 200 for all parameters sampled from the MCMC and the posterior distributions were examined using Tracer v. 1.6 [60].

3. Results

(a) Identity and phylogenetic placement of the Tibetan Plateau–Himalayan samples

Except for one tooth sample collected from a stuffed exhibit at the Reinhold Messner Mountain Museum, which BLAST-matched dog (*Canis lupus familiaris*), all other samples were identified as ursid bears. ML tree reconstruction based on amplicon and mitogenome sequences (electronic supplementary material, figure S1) grouped the 23 samples within four bear lineages: Himalayan brown bear, Tibetan brown bear, Continental Eurasian brown bear and Asian black bear. Complete mitogenomes were assembled from one individual in each of the four identified bear lineages (electronic supplementary material, table S3). ML and BI phylogenetic trees were reconstructed using the newly obtained amplicon sequences, complete mitogenome sequences, and previously published bear mtDNA sequences, using sloth bear (*U. ursinus*) as an outgroup (figure 2 and electronic supplementary material, figures S2 and S3). In general, the ML and BI tree topologies are consistent and in agreement with previous studies [10,17,61], with all major polar, brown and black bear clades well-resolved and strongly supported. The two Tibetan–Himalayan black bear samples formed a well-supported sister lineage to all other Asian black bear subspecies. The polar and brown bears grouped into nine clades (clades 1, 2a, 2b, 3a1, 3a2, 3b, 4, 5 and a Himalayan clade, with numerical clade nomenclature following [10,17]). Fourteen samples collected in Pakistan and the Himalayas grouped with a previously reported Gobi brown bear (GOBI-1) and two Himalayan brown bears (DQ914409 and DQ914410), and formed a sister lineage to all other brown and polar bear clades with strong bootstrap support. Six samples collected from the Tibetan Plateau grouped with previously sequenced Tibetan brown bears, which together formed a sister clade to several other North American and Eurasian brown bear lineages (clade 3a1,
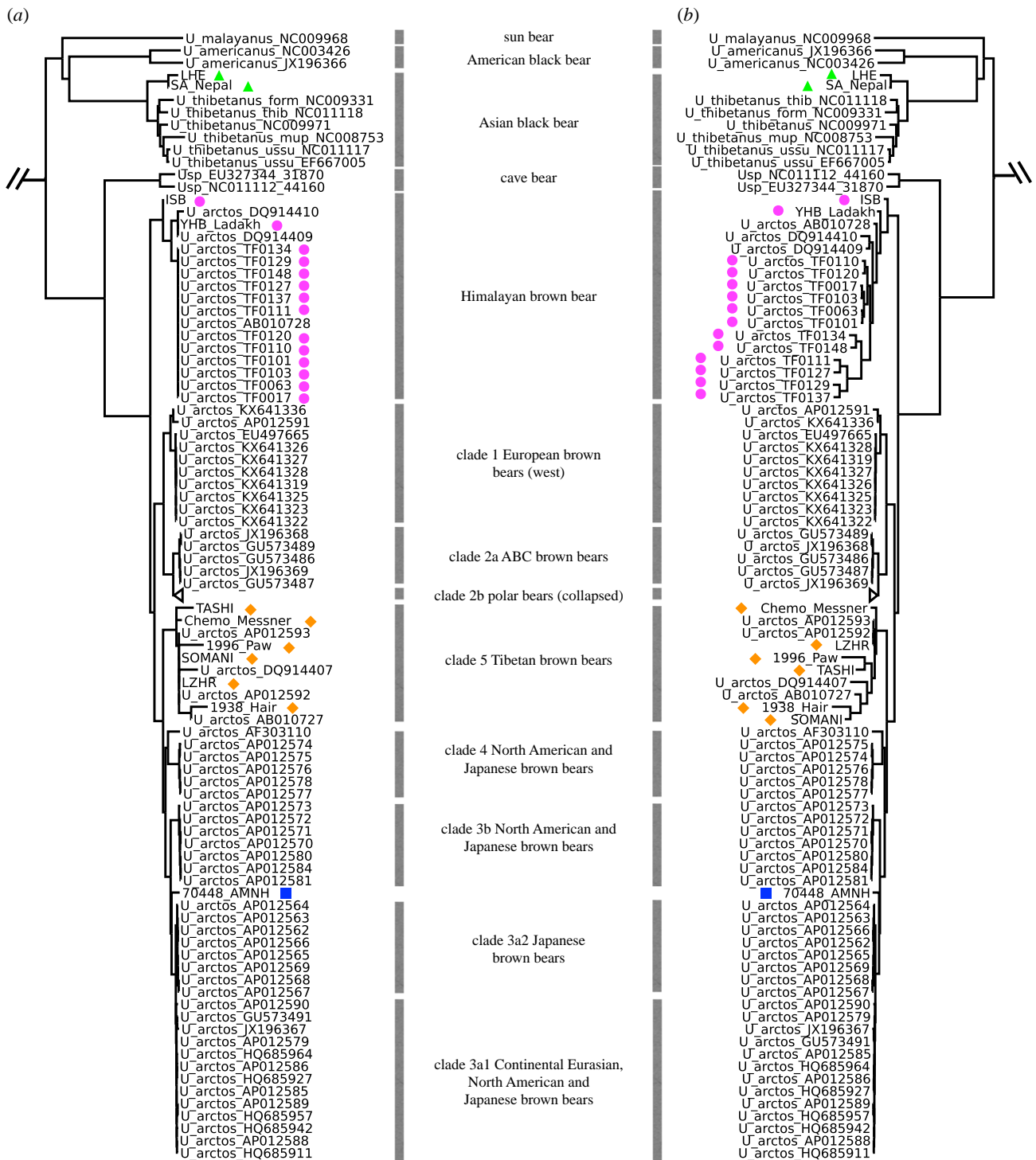
Figure 2. Phylogenetic trees based on (a) ML and (b) BI analyses of new mtDNA sequence data produced in this study and sequence data obtained from GenBank. New sequences are marked with triangles, diamonds, circles and a square, indicating the Asian black bear, Tibetan brown bear, Himalayan brown bear and the brown bear from the AMNH, respectively. GenBank data include complete mitogenomes of non-Tibetan–Himalayan bears, as well as amplicon and complete mitochondrial sequences of Tibetan and Himalayan bears. Major maternal clades and their geographic range are labelled following [10,17]. See electronic supplementary material, figures S2 and S3, for complete versions of the trees, shown with posterior probability and bootstrap values.

3a2, 3b and 4). One specimen (M-70448), which was sampled from the American Museum of Natural History's mammal collection and identified as a Tibetan brown bear, possibly of 'mixed breed', grouped in clade 3a with brown bears from Syria, Turkey, and animals held at Zoos in Europe [24] (electronic supplementary material, figure S1 and table S4).

(b) Divergence time estimations
MCMC-based divergence times discussed in the text are shown in figure 3 (see electronic supplementary material,

figure S4, for divergence times estimated for all nodes). For the brown bear clades, the divergence time between the Himalayan lineage and all other brown bear lineages was estimated to be 658 ka BP (95% HPD: 336–1258 ka BP). The divergence time between the Tibetan lineage and its sister North American and Eurasian lineages (clade 3 and 4) was estimated at 342 ka BP (95% HPD: 99–618 ka BP), and the split of the Continental Eurasian lineage (clade 3a) was estimated to be 146 ka BP (95% HPD: 14–799 ka BP). For the black bear clades, the ancestor of the Himalayan black bear
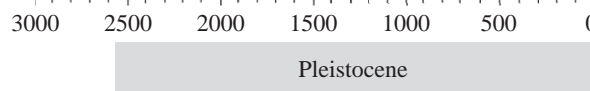
Figure 3. Maximum clade credibility tree from a BEAST analysis based on complete mitogenomes. The numbers at nodes indicate the median estimated divergence time in ka BP (HPD values are shown in brackets and the lower scale indicates time in ka BP). The coloured vertical bars indicate, from left to right, time spans of four Pleistocene glaciations: the Xixabangma, Nyanyaxungla, Guxiang and Baiyu. New mitogenomes sequenced in this study are indicated with symbols as in figure 2. See electronic supplementary material, figure S4, for a complete version of the tree and divergence times estimated for all nodes.

lineage diverged from other Asian black bear lineages at approximately 475 ka BP (95% HPD: 15–831 ka BP).

4. Discussion

(a) Phylogenetic placement and evolutionary history of Himalayan and Tibetan brown bears

Few genetic studies have been conducted of bears in the Tibetan Plateau and surrounding Himalaya region, and their evolutionary history remains enigmatic. Particularly little is known about the Himalayan brown bear (*U. a. isabellinus*). First, Masuda *et al.* [25] reported a 269 bp mtDNA control region sequence from a Gobi bear collected from the Great Gobi National Park in Mongolia, and suggested it was more closely related to Western European brown bears based on a neighbour-joining phylogenetic analysis. Later, Galbreath *et al.* [15] investigated homologous DNA fragments from two brown bears collected from the Deosai Plains of the western Himalayas. Their analyses demonstrated that the two Himalayan brown bears grouped together with the Gobi bear, confirming a close relationship between these two populations and a clear separation from European and Tibetan brown bears. Our results, providing more data and better resolution, demonstrate that the Himalayan brown bears, including the previously reported Gobi bear and Deosai bears, form a well-supported, sister lineage to all other extant brown bear clades included here. This result strongly supports Himalayan brown bears as a relict population that diverged early from other brown bear populations.

The phylogenetic position of Tibetan brown bears (*U. a. pruinosus*), which form a sister clade to North American and Eurasian brown bears consistent with previous reports [10,17–19,25], indicates that the Tibetan and other Eurasian brown bears, as well as North American brown bears, are all descendants of a common ancestral lineage. It was proposed that the Tibetan brown bears migrated to the Tibetan Plateau from its source population—ancestral Eurasian brown bears—approximately 343 ka BP, and that they remained geographically isolated from this source population thereafter [10]. Our phylogenetic analyses strongly support this migration scenario.

In our study, brown bear samples collected in the northwestern to western Himalayas were all identified as Himalayan brown bear, while the ones collected in the southeastern Himalayas and Tibetan Plateau were all identified as Tibetan brown bear (figure 1). The historical range of the Himalayan brown bear extends from the north and west of the Taklimakan Desert to the western Himalayas, while the historical range of the Tibetan brown bear lies in the Tibetan Plateau and the southeastern Himalayas [15]. While the Tibetan brown bears share a common ancestry with extant North American and Eurasian brown bears, the Himalayan brown bear appears to have originated from an ancient lineage that experienced long isolation in the mountains of central Asia, at least over the last 658 ka. Although the habitats of the two brown bear subspecies are geographically close, the high-altitude peaks of the Himalayan Mountains have likely impeded migration between these populations, and subsequently kept them as genetically distinct lineages.

(b) Phylogenetic placement and evolutionary history of the Himalayan black bear

The phylogenetic topology of Asian black bears is in agreement with a previous finding [61], except here we also include the rare Himalayan black bear (*U. t. laniger*), which forms a sister lineage to all other Asian black bears. Although sampling is limited, this result indicates that the Himalayan black bear originated from an ancient lineage and experienced long isolation in the Himalayan Mountains, a similar scenario to the divergence of the Himalayan brown bear lineage. However, the divergence time for the Himalayan black bear is younger, estimated at 475 ka BP, suggesting the isolation of Himalayan black bear occurred later than the isolation of the higher-altitude Himalayan brown bear. Reportedly, other described subspecies occur in the region, the Tibetan (*U. t. thibetanus*) and Indochinese (*U. t. mupinensis*) black bear, but whether these subspecies overlap is unclear given no modern revisionary work exists. Our phylogenetic relationships indicate that individuals from the Himalayas are genetically distant from other populations analysed, suggesting that little if any gene flow has occurred between this and other Asian black bear populations. Similar to the brown bear situation, the high mountains may also have separated the habitats of these black bear subspecies, possibly keeping *U. t. laniger* to the western Himalayas, and *U. t. mupinensis* and *U. t. thibetanus* to the east. Analyses of more individuals throughout the region and inclusion of nuclear DNA would be needed, however, to explore if this pattern is restricted to maternal gene flow only.

(c) Quaternary climatic oscillations and divergence of local bear lineages in the Tibetan Plateau–Himalaya region

The Tibetan Plateau is one of the youngest plateaus on Earth, created by the collision of the Indian subcontinent with the Eurasian continental plate in early Cenozoic times, followed by diachronous and extensive surface uplifts in the Miocene and even into the Pleistocene [62,63]. Although the dates and details of the uplifts have long been debated, many studies indicate they caused dramatic climatic changes and topographic variation, which facilitated the introduction and evolution of new plant and animal clades and greatly influenced the current spatial distribution of local species and their genetic diversity [64]. The Pleistocene glaciations of the Tibetan Plateau, which is closely related to the progressive uplift of the plateau and the surrounding Himalayan Mountains, have been suggested to have had a highly complex pattern, occurring asynchronously with the Northern Hemisphere glaciation events [65]. Four Pleistocene glaciations have been described in several geological and geographical studies [66–68]; the Xixabangma (Early Pleistocene, 1170–800 ka BP), Nyanyaxungla (Middle Pleistocene, 720–500 ka BP), Guxiang (Middle-Late Pleistocene, 300–130 ka BP) and Baiyu (Late Pleistocene, 70–10 ka BP) events. The most widespread Nyanyaxungla glaciation [64,69] was initiated by successive Kunlun-Huanghe tectonic movements. Interestingly, the divergence time of the Himalayan brown bear at around 658 ka BP overlaps with the Nyanyaxungla glaciation event, suggesting that this glaciation event may have caused the initial isolation of Himalayan brown bear. Glacial retreat occurred following the Nyanyaxungla glaciation, causing

changes in environmental conditions from cold and arid to warm and wet during the great interglacial period (500–300 ka BP) [68]. Both the divergence of the Himalayan black bear at around 475 ka BP and the Tibetan brown bear at around 342 ka BP overlap with this interglacial period, indicating that ancestors of these bear lineages migrated from lower altitudes to higher altitude locales after glaciers retreated. Subsequently, these populations may have diverged from lower altitude populations due to isolation in the high mountains and the following Guxiang glaciation event. Phylogeographic studies of many Tibetan plant and animal species indicate that local extant plant and animal populations, which mainly derived from colonists migrating from other areas or represent endemic species that diverged recently [3–10], experienced extensive oscillations and survived through glacial periods in multiple refugia or microrefugia on the plateau [1,65,70–75]. Similarly, we speculate that ancestral bear lineages on the Tibetan Plateau and Himalayan Mountains likely immigrated to the region from nearby Asian locales. These ancestral lineages then likely experienced extensive population oscillations caused by local climatic changes and diverged from other bear populations in refugia during the Pleistocene glaciations.

5. Conclusion

Samples collected in the field and archived in museum or private collections can significantly aid in our understanding of the genetic variation and phylogeographic patterns of rare and widespread species. To determine accurate species identification and clade affinity, however, phylogenetically informative genetic markers and appropriate phylogenetic analyses are critically needed. Based on a BLAST search using a 104 bp fragment of the mitochondrial 12S rRNA locus, which gave a 100% match to a complete mitogenome recovered from a subfossil polar bear [38], Sykes et al. [37] suggested that a previously unrecognized bear species or possibly a hybrid between brown bear and polar bear exists in the Himalayas. However, as also demonstrated by others [39,40], the short 12S rRNA gene fragment is insufficiently informative to determine precise taxonomic identity, particularly among closely related species, although it can be a useful screening marker to assess preliminary species affinities. We isolated DNA and assembled a complete mitogenome from a hair sample (collected in Ladakh, India, and named 'YHB' in this study), which based on their shared collection locality and other anecdotal evidence obtained from Icon Films, our sample source, may come from the same specimen that Sykes et al. [37] speculated represents an unknown or hybrid bear. Here, we unambiguously show that this sample is from a bear that groups with extant Himalayan brown bear. Similarly, we were able to determine the clade affinities of all other purported yeti samples in this study and infer their well-supported and resolved phylogenetic relationships among extant bears in the Tibetan Plateau and surrounding Himalayan Mountains. This study represents the most rigorous analysis to date of samples suspected to derive from anomalous or mythical 'hominid'-like creatures, strongly suggesting that the biological basis of the yeti legend is local brown and black bears.

References

1. Yang S, Dong H, Lei F. 2009 Phylogeography of regional fauna on the Tibetan Plateau: a review. Progress Nat. Sci. 19, 789–799. (doi:10.1016/j.pnsc.2008.10.006)

2. Xu J, Grumbine RE, Shrestha A, Eriksson M, Yang X, Wang YUN, Wilkes A. 2009 The melting Himalayas: cascading effects of climate change on water, biodiversity, and livelihoods. Conserv. Biol. 23, 520–530. (doi:10.1111/j.1523-1739.2009.01237.x)

3. Qiao C-Y, Ran J-H, Li Y, Wang X-Q. 2007 Phylogeny and biogeography of Cedrus (Pinaceae) inferred from sequences of seven paternal chloroplast and maternal mitochondrial DNA regions. Ann. Bot. 100, 573–580. (doi:10.1093/aob/mcm134)

4. Wu ZY. 1980 Vegetation of China. Beijing, China: Science Press.

5. Yang F-S, Wang X-Q, Hong D-Y. 2003 Unexpected high divergence in nrDNA ITS and extensive parallelism in floral morphology of Pedicularis (Orobanchaceae). Pl. Syst. Evol. 240, 91–105. (doi:10.1007/s00606-003-0005-2)

6. Wei XX, Wang XQ. 2004 Recolonization and radiation in Larix (Pinaceae): evidence from nuclear ribosomal DNA paralogues. Mol. Ecol. 13, 3115–3123. (doi:10.1111/j.1365-294X.2004.02299.x)

7. Liu J-Q, Wang Y-J, Wang A-L, Hideaki O, Abbott RJ. 2006 Radiation and diversification within the Ligularia–Cremanthodium–Parasenecio complex (Asteraceae) triggered by uplift of the Qinghai-Tibetan Plateau. Mol. Phyl. Evol. 38, 31–49. (doi:10.1016/j.ympev.2005.09.010)

8. Peng Z, Ho SYW, Zhang Y, He S. 2006 Uplift of the Tibetan Plateau: evidence from divergence times of glyptosternoid catfishes. Mol. Phyl. Evol. 39, 568–572. (doi:10.1016/j.ympev.2005.10.016)

9. Ran J-H, Wei X-X, Wang X-Q. 2006 Molecular phylogeny and biogeography of Picea (Pinaceae): Implications for phylogeographical studies using cytoplasmic haplotypes. Mol. Phyl. Evol. 41, 405–419. (doi:10.1016/j.ympev.2006.05.039)

10. Hirata D et al. 2013 Molecular phylogeography of the brown bear (Ursus arctos) in northeastern Asia based on analyses of complete mitochondrial DNA sequences. Mol. Biol. Evol. 30, 1644–1652. (doi:10.1093/molbev/mst077)

11. Pocock RI. 1941 The fauna of British India, including Ceylon and Burma. Volume II. Mammalia. London, UK: Taylor and Francis.

12. Servheen C, Herrero S, Peyton B. (compilers). 1999 Bears. Status survey and conservation action plan. IUCN/SSC bear and polar bear specialist groups, 32 p. Gland, Switzerland: IUCN.

13. Aryal A, Raubenheimer D, Sathyakumar S, Poudel BS, Ji W, Kunwar KJ, Kok J, Kohshima S, Brunton D. 2012 Conservation strategy for brown bear and its habitat in Nepal. Diversity 4, 301–317. (doi:10.3390/d4030301)

14. Nawaz MA. 2007 Status of the brown bear in Pakistan. Ursus 18, 89–100. (doi:10.2192/1537-6176(2007)18[89:SOTBBI]2.0.CO;2)

15. Galbreath GJ, Groves CP, Waits LP. 2007 Genetic resolution of composition and phylogenetic placement of the isabelline bear. *Ursus* **18**, 129–131. (doi:10.2192/1537-6176(2007)18[129:GROCAP]2.0.CO;2)

16. Barnes I, Matheus P, Shapiro B, Jensen D, Cooper A. 2002 Dynamics of Pleistocene extinctions in Beringian brown bears. *Science* **295**, 2267–2270. (doi:10.1126/science.1067814)

17. Davison J *et al.* 2011 Late-Quaternary biogeographic scenarios for the brown bear (*Ursus arctos*), a wild mammal model species. *Quat. Sci. Rev.* **30**, 418–430. (doi:10.1016/j.quascirev.2010.11.023)

18. Edwards CJ *et al.* 2011 Ancient hybridization and an Irish origin for the modern polar bear matriline. *Curr. Biol.* **21**, 1251–1258. (doi:10.1016/j.cub.2011.05.058)

19. Korsten M *et al.* 2009 Sudden expansion of a single brown bear maternal lineage across northern continental Eurasia after the last ice age: a general demographic model for mammals? *Mol. Ecol.* **18**, 1963–1979. (doi:10.1111/j.1365-294X.2009.04163.x)

20. Matheus P, Burns J, Weinstock J, Hofreiter M. 2004 Pleistocene brown bears in the mid-continent of North America. *Science* **306**, 1150. (doi:10.1126/science.1101495)

21. Talbot SL, Shields GF. 1996 Phylogeography of brown bears (*Ursus arctos*) of Alaska and paraphyly within the Ursidae. *Mol. Phyl. Evol.* **5**, 477–494. (doi:10.1006/mpev.1996.0044)

22. Waits LP, Talbot SL, Ward RH, Shields GF. 1998 Mitochondrial DNA phylogeography of the North American brown bear and implications for conservation. *Conserv. Biol.* **12**, 408–417. (doi:10.1046/j.1523-1739.1998.96351.x)

23. Bray SCE *et al.* 2013 Ancient DNA identifies post-glacial recolonisation, not recent bottlenecks, as the primary driver of contemporary mtDNA phylogeography and diversity in Scandinavian brown bears. *Div. Distr.* **19**, 245–256. (doi:10.1111/j.1472-4642.2012.00923.x)

24. Calvignac S, Hughes S, Hänni C. 2009 Genetic diversity of endangered brown bear (*Ursus arctos*) populations at the crossroads of Europe, Asia and Africa. *Div. Distr.* **15**, 742–750. (doi:10.1111/j.1472-4642.2009.00586.x)

25. Masuda R, Murata K, Aiurzaniin A, Yoshida MC. 1998 Phylogenetic status of brown bears *Ursus arctos* of Asia: a preliminary result inferred from mitochondrial DNA control region sequences. *Hereditas* **128**, 277–280. (doi:10.1111/j.1601-5223.1998.00277.x)

26. Matsuhashi T, Masuda R, Mano T, Murata K, Aiurzaniin A. 2001 Phylogenetic relationships among worldwide populations of the brown bear *Ursus arctos*. *Zool. Sci.* **18**, 1137–1143. (doi:10.2108/zsj.18.1137)

27. Garshelis DL, Steinmetz R. 2016 *Ursus thibetanus*. (errata version published in 2017) *The IUCN Red List of Threatened Species 2016*. e.T22824A114252336. See http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22824A45034242.e.

28. Sathyakumar S, Choudhury A. 2007 Distribution and status of the Asiatic black bear *Ursus thibetanus* in India. *J. Bombay Nat. Hist. Soc.* **104**, 316–323.

29. Abbas F, Bhatti ZI, Haider J, Mian A. 2015 Bears in Pakistan: distribution, population biology and human conflicts. *J. Biores. Manage.* **2**, 1–13.

30. Bellemain E, Nawaz MA, Valentini A, Swenson JE, Taberlet P. 2007 Genetic tracking of the brown bear in northern Pakistan and implications for conservation. *Biol. Conserv.* **134**, 537–547. (doi:10.1016/j.biocon.2006.09.004)

31. Nawaz MA. 2008 Ecology, genetics and conservation of Himalayan brown bears. PhD thesis, Norwegian University of Life Sciences, Norway.

32. Hwang D-S *et al.* 2008 A comprehensive analysis of three Asiatic black bear mitochondrial genomes (subspecies *ussuricus*, *formosanus* and *mupinensis*), with emphasis on the complete mtDNA sequence of *Ursus thibetanus ussuricus* (Ursidae). *DNA Seq* **19**, 418–429. (doi:10.1080/19401730802389525)

33. Ohnishi N, Uno R, Ishibashi Y, Tamate H, Oi T. 2009 The influence of climatic oscillations during the Quaternary Era on the genetic structure of Asian black bears in Japan. *Heredity* **102**, 579–589. (doi:10.1038/hdy.2009.28)

34. McLellan BN, Proctor MF, Huber D, Michel S. 2017 *Ursus arctos*. (amended version published in 2016) *The IUCN Red List of Threatened Species 2017*. e.T41688A114261661. See http://dx.doi.org/10.2305/IUCN.UK.2017–1.RLTS.T41688A114261661.en.

35. Messner R. 2001 *My quest for the yeti*. New York, NY: St. Martin's Press.

36. Milinkovitch MC, Caccone A, Amato G. 2004 Molecular phylogenetic analyses indicate extensive morphological convergence between the 'yeti' and primates. *Mol. Phyl. Evol.* **31**, 1–3. (doi:10.1016/S1055-7903(04)00045-4)

37. Sykes BC, Mullis RA, Hagenmuller C, Melton TW, Sartori M. 2014 Genetic analysis of hair samples attributed to yeti, bigfoot and other anomalous primates. *Proc. R. Soc. B* **281**, 20140161. (doi:10.1098/rspb.2014.0161)

38. Lindqvist C *et al.* 2010 Complete mitochondrial genome of a Pleistocene jawbone unveils the origin of polar bear. *Proc. Natl Acad. Sci. USA* **107**, 5053–5057. (doi:10.1073/pnas.0914266107)

39. Edwards CJ, Barnett R. 2015 Himalayan 'yeti' DNA: polar bear or DNA degradation? A comment on 'Genetic analysis of hair samples attributed to yeti' by Sykes *et al.* (2014). *Proc. R. Soc. B* **282**, 20141712. (doi:10.1098/rspb.2014.1712)

40. Gutiérrez E, Pine RH. 2015 No need to replace an 'anomalous' primate (Primates) with an 'anomalous' bear (Carnivora, Ursidae). *ZooKeys* **487**, 141–154. (doi:10.3897/zookeys.487.9176)

41. Bischof R, Hameed S, Ali H, Kabir M, Younas M, Shah KA, Din JU, Nawaz MA. 2014 Using time-to-event analysis to complement hierarchical methods when assessing determinants of photographic detectability during camera trapping. *Methods Ecol. Evol.* **5**, 44–53. (doi:10.1111/2041-210X.12115)

42. Dabney J *et al.* 2013 Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc.*

43. Gilbert MTP *et al.* 2007 Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. *Science* **317**, 1927–1930. (doi:10.1126/science.1146971)

44. Melton T, Holland C. 2007 Routine forensic use of the mitochondrial 12S ribosomal RNA gene for species identification. *J. Foren. Sci.* **52**, 1305–1307. (doi:10.1111/j.1556-4029.2007.00553.x)

45. Rohland N, Siedel H, Hofreiter M. 2004 Nondestructive DNA extraction method for mitochondrial DNA analyses of museum specimens. *Biotechniques* **36**, 814–821.

46. Taberlet P, Bouvet J. 1994 Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear *Ursus arctos* in Europe. *Proc. R. Soc. B* **255**, 195–200. (doi:10.1098/rspb.1994.0028)

47. Li H, Durbin R. 2010 Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595. (doi:10.1093/bioinformatics/btp698)

48. Schubert M, Ginolhac A, Lindgreen S, Thompson JF, Al-Rasheid KA, Willerslev E, Krogh A, Orlando L. 2012 Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics* **13**, 1. (doi:10.1186/1471-2164-13-178)

49. Li H. 2011 A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993. (doi:10.1093/bioinformatics/btr509)

50. Katoh K, Standley DM. 2013 MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780. (doi:10.1093/molbev/mst010)

51. Hall TA. 1999 BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* **41**, 95–98.

52. Stamatakis A. 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313. (doi:10.1093/bioinformatics/btu033)

53. Darriba D, Taboada GL, Doallo R, Posada D. 2012 jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**, 772. (doi:10.1038/nmeth.2109)

54. Guindon S, Gascuel O. 2003 A simple, fast, and accurate method to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704. (doi:10.1080/10635150390235520)

55. Ronquist F, Huelsenbeck JP. 2003 MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574. (doi:10.1093/bioinformatics/btg180)

56. Krause J *et al.* 2008 Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene–Pliocene boundary. *BMC Evol. Biol.* **8**, 220. (doi:10.1186/1471-2148-8-220)

Natl Acad. Sci. USA **110**, 15 758–15 763. (doi:10.1073/pnas.1314445110)

9

57. Bon C et al. 2008 Deciphering the complete mitochondrial genome and phylogeny of the extinct cave bear in the Paleolithic painted cave of Chauvet. Proc. Natl Acad. Sci. USA 105, 17 447 – 17 452. (doi:10.1073/pnas.0806143105)

58. Fortes GG et al. 2016 Ancient DNA reveals differences in behaviour and sociality between brown bears and extinct cave bears. Mol. Ecol. 25, 4907 – 4918. (doi:10.1111/mec.13800)

59. Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012 Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969 – 1973. (doi:10. 1093/molbev/mss075)

60. Drummond AJ, Rambaut A. 2007 BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 1. (doi:10.1186/1471-2148-7-214)

61. Wu J, Kohno N, Mano S, Fukumoto Y, Tanabe H, Hasegawa M, Yonezawa T. 2015 Phylogeographic and demographic analysis of the Asian Black Bear (Ursus thibetanus) based on mitochondrial DNA. PLoS ONE 10, e0136398.

62. Chung S-L, Lo C-H, Lee T-Y, Zhang Y, Xie Y, Li X, Wang K-L, Wang P-L. 1998 Diachronous uplift of the Tibetan Plateau starting 40 Myr ago. Nature 394, 769 – 773. (doi:10.1038/29511)

63. Sun J, Liu T. 2000 Stratigraphic evidence for the uplift of the Tibetan Plateau between ∼1.1 and ∼0.9 myr ago. Quat. Res. 54, 309 – 320. (doi:10. 1006/qres.2000.2170)

64. Lei F, Qu Y, Song G. 2014 Species diversification and phylogeographical patterns of birds in response to the uplift of the Qinghai – Tibet Plateau and Quaternary glaciations. Curr. Zool. 60, 149 – 161. (doi:10.1093/czoolo/60.2.149)

65. Zhang F, Jiang Z. 2006 Mitochondrial phylogeography and genetic diversity of Tibetan gazelle (Procapra picticaudata): implications for conservation. Mol. Phyl. Evol. 41, 313 – 321. (doi:10. 1016/j.ympev.2006.05.024)

66. Benxing Z, Rutter N. 1998 On the problem of Quaternary glaciations, and the extent and patterns of Pleistocene ice cover in the Qinghai – Xizang (Tibet) Plateau. Quat. Int. 45, 109 – 122. (doi:10. 1016/S1040-6182(97)00009-8)

67. Zhou S, Xu L, Colgan PM, Mickelson DM, Wang X, Wang J, Zhong W. 2007 Cosmogenic 10Be dating of Guxiang and Baiyu glaciations. Chin. Sci. Bull. 52, 1387 – 1393. (doi:10.1007/s11434-007-0208-y)

68. Zheng B, Xu Q, Shen Y. 2002 The relationship between climate change and Quaternary glacial cycles on the Qinghai – Tibetan Plateau: review and speculation. Quat. Int. 97, 93 – 101. (doi:10.1016/ S1040-6182(02)00054-X)

69. Li J, Fang X. 1999 Uplift of the Tibetan Plateau and environmental changes. Chin. Sci. Bull. 44, 2117 – 2124. (doi:10.1007/BF03182692)

70. Wang L, Abbott RJ, Zheng W, Chen P, Wang Y, Liu J. 2009 History and evolution of alpine plants endemic to the Qinghai – Tibetan Plateau: Aconitum gymnandrum (Ranunculaceae). Mol. Ecol. 18, 709 – 721. (doi:10.1111/j.1365-294X.2008.04055.x)

71. Wang GN, He XY, Miehe G, Mao KS. 2014 Phylogeography of the Qinghai – Tibet Plateau endemic alpine herb Pomatosace filicula (Primulaceae). J. Syst. Evol. 52, 289 – 302. (doi:10. 1111/jse.12089)

72. Opgenoorth L, Vendramin GG, Mao K, Miehe G, Miehe S, Liepelt S, Liu J, Ziegenhagen B. 2010 Tree endurance on the Tibetan Plateau marks the world's highest known tree line of the Last Glacial Maximum. New Phytologist 185, 332 – 342. (doi:10. 1111/j.1469-8137.2009.03007.x)

73. Ma YZ, Li ZH, Wang X, Shang BL, Wu GL, Wang YJ. 2014 Phylogeography of the genus Dasiphora (Rosaceae) in the Qinghai – Tibetan Plateau: divergence blurred by expansion. Biol. J. Linn. Soc. 111, 777 – 788. (doi:10.1111/bij. 12246)

74. Wan DS, Feng JJ, Jiang DC, Mao KS, Duan YW, Miehe G, Opgenoorth L. 2016 The Quaternary evolutionary history, potential distribution dynamics, and conservation implications for a Qinghai – Tibet Plateau endemic herbaceous perennial, Anisodus tanguticus (Solanaceae). Ecol. Evol. 6, 1977 – 1995. (doi:10.1002/ece3. 2019)

75. Yang SJ, Lei FM, Qu YH, Yin ZH. 2006 Intraspecific phylogeography of the white-rumped snowfinch (Onychostruthus taczanowskii) endemic to the Tibetan Plateau based on mtDNA sequences. J. Zool. 268, 187 – 192. (doi:10.1111/j.1469-7998.2005. 00009.x)