Complete mitochondrial genome of a sun bear from Malaysia and its position in the phylogeny of Ursidae

Authors: Lai, Wai Ling, Ratnayeke, Shyamala, Austin, Christopher, Rahman, Sadequr, Ayub, Qasim, et al.

Source: Ursus, 2021(32e19) : 1-10

Published By: International Association for Bear Research and Management

URL: https://doi.org/10.2192/URSUS-D-20-00032.1

Complete mitochondrial genome of a sun bear from Malaysia and its position in the phylogeny of Ursidae

Wai Ling Lai[1], Shyamala Ratnayeke[1], Christopher Austin[2], Sadequr Rahman[2], Qasim Ayub[2,3], Noor Azleen Mohd Kulaimi[4], Sagathevan Kuppusamy[5], and Jactty Chew[1,6]

[1]*Department of Biological Sciences, School of Science and Technology, Sunway University, No. 5, Jalan Universiti, Bandar Sunway, 47500 Selangor Darul Ehsan, Malaysia*
[2]*Monash University Malaysia Genomics Facility, Tropical Medicine and Biology Multidisplinary Platform, Jalan Lagoon Selatan, Bandar Sunway, 47500 Selangor Darul Ehsan, Malaysia*
[3]*School of Science, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, 47500 Selangor Darul Ehsan, Malaysia*
[4]*National Wildlife Forensic Laboratory (NWFL), Ex-situ Conservation Division, Department of Wildlife and National Parks (DWNP), KM 10 Jalan Cheras, 56100 Kuala Lumpur, Malaysia*
[5]*Science and Engineering Resources, Sunway University, No. 5, Jalan Universiti, Bandar Sunway, 47500 Selangor Darul Ehsan, Malaysia*

Abstract: Whole mitochondrial genome sequences have important applications for phylogenetic inference, population evolution, and population structure. In this study, we sequenced the entire mitochondrial genome of a sun bear (*Helarctos malayanus*) from Peninsular (West) Malaysia using Illumina Miseq technology and used 26 additional mitochondrial genomes from the Ursidae, including 5 sun bears, to generate a phylogeny. The complete mitochondrial genome of the sun bear consisted of 16,770 base pairs (bp), including 13 protein-coding genes, 2 ribosomal subunit genes, 22 transfer RNAs, and a noncoding, adenine–thymine (AT) -rich control region. Maximum likelihood and Bayesian inference phylogenetic trees revealed topologies identical to trees previously published using whole mitochondrial genomes. Sun bears formed 2 distinct mitochondrial lineages, with the peninsular genome occupying a clade separate from the clade including a sun bear from Yunnan, China. Within the control region, all 5 sun bear genomes differed at a microsatellite repeat region and all 5 genomes consistently lacked a 6-bp imperfect repeat, which is found in some bear species. Ursine phylogenies constructed with whole mitochondrial genomes conflict with recent well-resolved coalescent trees employing whole genome data. However, both phylogenies suggest a historical split in the sun bear lineage. Furthermore, the inclusion of the peninsular sun bear mitochondrial genome suggests that this split does not conform to the current subspecies delineation in sun bears. Genomic data from multiple individuals of known geographic origin will help to resolve this question.

Key words: bear phylogeny, complete mitochondrial genome, genetic sequencing, *Helarctos malayanus,* microsatellite region, Sundaic region, Ursidae, Ursinae

DOI: 10.2192/URSUS-D-20-00032.1 ***Ursus 32:article e19 (2021)***

The sun bear (*Helarctos malayanus*) is the smallest of the 8 bear species. The species' historical distribution ranged from northeastern India to southern Yunnan Province in China to Malaysia and Indonesia (Servheen 1999). The 2 currently recognized subspecies of sun bear are *H. m. malayanus,* which occurs on the Southeast Asian mainland, including Peninsular (West) Malaysia and Sumatra, and *H. m. euryspilus,* which is limited to Borneo (Meijaard 2004). Currently, sun bears occur patchily throughout the distribution range owing to rapid economic development and land clearing for agricultural purposes (Scotson et al. 2017).

The evolution of the Ursidae is linked to a sequence of rapid evolutionary radiation events dating back to the early Miocene about 20 million years ago (mya; Waits et al. 1999, Krause et al. 2008). The lineage leading to the giant panda (*Ailuropoda melanoleuca*) is widely accepted as the most basal offshoot in the bear family (O'Brien et al. 1985, Hashimoto et al. 1993, Zhang and Ryder 1994). The subsequent divergence of the Tremarctine bears 7–13 mya is represented now by one extant

[6]email: jacttyc@sunway.edu.my

species, the Andean bear (*Tremarctos ornatus*) of South America (Kurten 2017). The Ursine lineage radiated <1–6 mya (Kumar et al. 2017), leading to the remaining 6 extant species: the brown bear (*Ursus arctos*), the polar bear (*U. maritimus*), the North American black bear (*U. americanus*), the Asiatic black bear (*U. thibetanus*), the sun bear, and the sloth bear (*Melursus ursinus*).

Resolving phylogenetic relationships within the Ursidae has been challenging. Apart from the recent divergence and close evolutionary relationship between the brown bear and polar bear, which is now widely accepted, phylogenetic relationships among the more recently evolved ursine bears have shown conflicting relationships (e.g., Zhang and Ryder 1994; Talbot and Shields 1996; Waits et al. 1998; Yu et al. 2004, 2007; Pages et al. 2008). Mitochondrial DNA has proved to be extremely useful for inferring evolutionary history and resolving taxonomic uncertainties at the level of genera and species (DeSalle et al. 2017). Previously, partial regions of the mitogenome, including combinations of a few genes, resulted in inconclusive branching patterns in the ursine phylogeny (Zhang and Ryder 1994, Talbot and Shields 1996, Waits et al. 1999, Yu et al. 2004). With advances in molecular sequencing techniques, complete mitochondrial genome data has generated consistent and well-supported ursid phylogenies (Yu et al. 2007, Krause et al. 2008). Nevertheless, ursine phylogenies using mitochondrial DNA have conflicted with those produced using nuclear genes (e.g., Yu et al. 2004, Pages et al. 2008).

Kutschera et al. (2014) compared Ursidae phylogenies constructed from multiple autosomal introns, Y-chromosome loci, and whole mitochondrial genomes. Using coalescent-based gene-flow analyses, they concluded that incomplete lineage sorting and interspecific gene flow in the ursine genealogical history explained the discordance between mitochondrial and nuclear phylogenies. Using a large molecular data set sourced from mitochondrial, autosomal, and Y-chromosome DNA, Kumar et al. (2017) generated a well-supported coalescent species tree for the Ursidae. This study further supported the hypothesis that early and episodic gene flow during ursine radiation had been largely responsible for the difficulties in resolving the phylogeny of the more recently evolved Ursinae. Kumar et al. (2017) highlighted a potential split in the sun bear lineage about 100 kya, including 2 distinct mitochondrial lineages, which they thought might have corresponded to the 2 described subspecies. However, the geographic origin of only one of those sun bears was known. Compared with other bear species, few complete mitochondrial genomes from confirmed geographic origins are available for sun bears.

In this study, we recovered one complete sun bear mitochondrial genome from Peninsular Malaysia using next generation sequencing (NGS) technology. We then constructed phylogenetic trees using 27 complete mitochondrial genomes of extant bear species, including those of 4 other sun bears, with the aim of contributing a genome of known geographic origin. This newly sequenced mitochondrial genome will also provide the foundation for developing markers to advance population genetic studies, and potentially, to identify unique mitochondrial lineages among sun bear populations in the Sundaic region.

Materials and methods

Sampling, DNA extraction, next generation sequencing

We conducted this study under permit No. JPHL&TN(IP):100-34/1.24 Jld. 2(15) provided by the Department of Wildlife and National Parks (DWNP) Malaysia. We recovered total genomic DNA from a sample of muscle tissue provided by DWNP, May 2017. The sample was obtained from a road-killed wild sun bear from Peninsular Malaysia. We extracted DNA from 40 mg of muscle tissue using the DNeasy Blood and Tissue Kit (Qiagen, Haldane, Germany), following the manufacturer's protocol with minor modification: overnight incubation of muscle tissue at 56°C to ensure total lysis of the long-frozen sample. We sheared genomic DNA to 550 bp using a M220 focused-ultrasonicator (Covaris, Woburn, Massachusetts, USA). We prepared a library using the NEBNext® Ultra™ DNA Library Prep Kit (New England Biolabs Inc., Ipswich, Massachusetts, USA) and subsequently sequenced on an Illumina MiSeq platform (Illumina, San Diego, California, USA) using 250-bp paired-end reads.

Reference guided de novo assembly of complete sun bear mitochondrial genome

We used Geneious v11.1.5 software (Biomatters Ltd., Auckland, New Zealand; Kearse et al. 2012) to view, process, and analyze sequence data. We downloaded 3 published mitochondrial genomes of sun bears—EF196664 (Yu et al. 2007), NC009968 (revised version of EF196664), and FM177765 (Krause et al. 2008)—from the National Center for Biotechnology Information (NCBI). We obtained forward and reverse sequence reads (length 2 × 250 bp) from the Illumina Miseq Benchtop Sequencer. These reads were "set-paired" using Geneious. We then mapped paired sequencing reads to all 3 downloaded reference mitochondrial genomes

to obtain pairwise percentage identity and to confirm that the DNA from the peninsular bear was of the same species.

The raw reads contained sequences from both nuclear and mitochondrial genomes. During mapping, mitochondrial sequences were filtered out from raw reads and saved as the input file for de novo assembly of the complete sun bear mitochondrial genome to produce contigs. We aligned this raw de novo consensus sequence from mapping with the published reference genomes using Geneious' ClustalW aligner to generate a raw mitochondrial genome.

Upon checking the raw assembled mitochondrial genome, we found ambiguous nucleotides within a 100-bp length within the mitochondrial microsatellite region, located roughly at the mitochondrial position of 500–600 bp referred to as the mitochondrial control region of EF196664. To clarify the ambiguous nucleotides, we designed one set of primers to amplify 57–900 bp of the mitochondrial genome: MtSB2F (5'-ACTTGCTATGACTCAGCTAT-3') and MtSB2R (3'-CTACATGGACGTAGATGGTT-5'); amplicons were sent for Sanger sequencing. We then aligned and trimmed the sequences, and compared them with the raw mitochondrial genome in Geneious. We then replaced ambiguous nucleotides according to the aligned sequencing results. We recorded the length and the guanine–cytosine (GC) content of the full sun bear mitochondrial genome. We did not try to detect nuclear DNA sequences of mitochondrial origin (numts). We considered numt coverage to be much lower than that of mtDNA, which typically has hundreds to thousands of copies per cell in muscle tissue. We performed annotation of the mitochondrial genome using MITOS (Bernt et al. 2013). We generated the visualization of the complete circular mitochondrial genome with detailed gene annotation using Geneious.

Comparisons of the newly sequenced sun bear mitochondrial genome with published sun bear genomes

We performed ClustalW alignment of 5 sun bear mitochondrial genomes, including our newly sequenced sample (MN807949), 2 published genomes (EF196664 and FM177765), and 2 genomes from Kumar et al. (2017; Anabell and Klaus). We then carefully examined the alignment for nucleotide differences, after which we calculated pairwise nucleotide distances among the genomes using MEGA7. For a better understanding of genetic relatedness, we then performed phylogenetic analysis.

Phylogenetic analysis of 8 bear species

We downloaded 24 complete mitochondrial genomes of 8 different bear species from NCBI, consisting of 2 American black bears, 4 brown bears, 5 polar bears, 3 Asiatic black bears, 2 Andean bears, 4 giant pandas, 2 sloth bears, and 2 sun bears. We used one gray wolf (*Canis lupus*) mitochondrial genome (KF857179) as an outgroup for the following phylogenetic analysis. Together with 2 additional sun bear mitochondrial genomes (Kumar et al. 2017) and the newly sequenced mitochondrial genome from Malaysia, we checked 28 complete mitochondrial genomes in BioEdit 7.0.9.0 (Hall 1999) and aligned them using Clustal X (Thompson et al. 1997). Maximum likelihood (ML) analysis was performed in MEGA 7 (Kumar et al. 2016) using the best-fitting model (GTR + G + I) and 1,000 bootstrap replicates to evaluate nodal support. For Bayesian inference (BI), we used MrBayes v3.2 (Ronquist et al. 2011) incorporating the best-fitting model, GTR + G + I, with Markov chain Monte Carlo calculations of 2,000,000 generations and sample frequency of 500. We used Tracer v1.7.1 (Rambaut et al. 2018) to ensure split frequencies were below 0.05 and effective sample sizes ≥200. We viewed both ML- and BI-constructed phylogenetic trees in FigTree v.1.4.0 (Rambaut 2012). We computed nucleotide p-distances (proportion of nucleotide sites that were different) within and among each bear species in MEGA 7, with rate variation among sites modeled with a gamma distribution (shape parameter = 1) and 1,000 bootstrap replicates for variance estimation.

Results

The complete mitochondrial genome of the Peninsular Malaysia sun bear was 16,770 bp with a GC percentage of 41% (Fig. 1). All genes encoded in the mitochondrial genome, including 13 protein-coding genes, 22 transfer RNA genes, and 2 ribosomal RNA genes, were recovered (Table 1) and matched to reference genomes (EF196664, FM177765) with the same gene order. The overall nucleotide composition of the sequenced sun bear was A = 31.2%, C = 25.6%, G = 15.4%, and T = 27.8%. The control region (1,326 bp) contained a microsatellite region with 23 10-bp repeat sequences of 5'-ACGCACGTGT-3' and 8 10-bp repeat sequences of 5'-ACGCATGTGT-3'.

From the alignment of all 5 sun bear mitochondrial genomes, we noticed 233 nucleotide variation sites, resulting most probably from point mutations. The control region (~1,300 bp) contained 49 variable sites, comprising 21% of all variable sites in the genome. A complex of short tandem repeats (microsatellite) was observed within

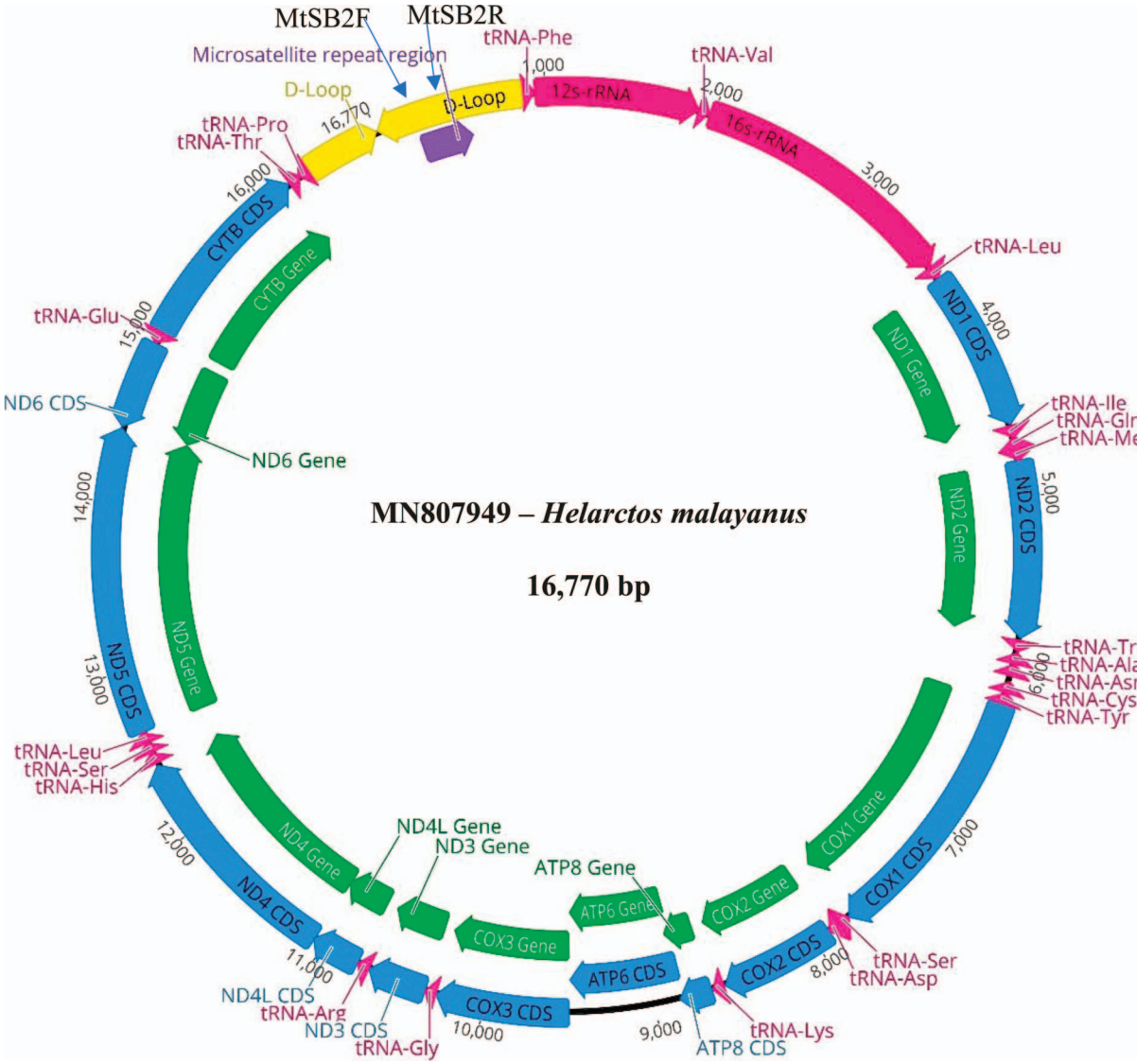

Fig. 1. Map of complete mitochondrial genome of a sun bear (*Helarctos malayanus;* 16,770 base pairs), sequenced using Illumina Miseq technology and using 26 additional mitochondrial genomes from the Ursidae, including 5 sun bears, to generate a phylogeny. We recovered total genomic DNA from a sample of muscle tissue of a road-killed bear in Peninsular Malaysia, as provided by Department of Wildlife and National Parks. COX1-3 indicates cytochrome c oxidase subunits; CYTB, cytochrome b; ATP6-8, ATPase subunits 6 and 8; ND1-6/4L, NADH dehydrogenase subunits 1-6/4L. Color code: Pink—tRNA, Fuchsia—rRNA, Blue—Coding sequence site for coding genes, Green—Name of coding genes, Yellow—Mitochondrial control region (D-Loop), and Purple arrow—Microsatellite repeat region. Arrow heads indicate direction of transcription. Mitochondrial markers (MtSB2F and MtSB2R) designed for mitochondrial genome clean-up are shown above.

the control region. This has been reported in other bear species (Choi et al. 2010) and can be generally divided into a 10-bp perfect repeat (PR) and a 6-bp imperfect repeat (IR). The 2 major types of perfect repeats seen in bear species are PR1 (5'-ACGCACGTGT-3') and PR2 (5'-ACGCATGTGT-3'). The imperfect repeats, however, are a 5'-end truncated form of PR or its derivatives. For instance, the imperfect repeat "ACGTGT" may be formed

Table 1. Annotation of sun bear (*Helarctos malayanus;* MN807949) mitochondrial genome using MITOS. 'bp' is base pair.

Type		Name	Start	Stop	Strand	Length (bp)
tRNA	1	tRNA-Phe	872	939	+	68
	2	tRNA-Val	1,904	1,969	+	66
	3	tRNA-Leu	3,550	3,624	+	75
	4	tRNA-Ile	4,583	4,651	+	69
	5	tRNA-Gln	4,649	4,721	−	73
	6	tRNA-Met	4,723	4,791	+	69
	7	tRNA-Trp	5,834	5,900	+	67
	8	tRNA-Ala	5,909	5,977	−	69
	9	tRNA-Asn	5,978	6,050	−	73
	10	tRNA-Cys	6,084	6,150	−	67
	11	tRNA-Tyr	6,151	6,217	−	67
	12	tRNA-Ser	7,761	7,829	−	69
	13	tRNA-Asp	7,836	7,902	+	67
	14	tRNA-Lys	8,590	8,657	+	68
	15	tRNA-Gly	10,285	10,353	+	69
	16	tRNA-Arg	10,701	10,769	+	69
	17	tRNA-His	12,438	12,506	+	69
	18	tRNA-Ser	12,507	12,565	+	59
	19	tRNA-Leu	12,566	12,635	+	70
	20	tRNA-Glu	14,968	15,036	−	69
	21	tRNA-Thr	16,181	16,250	+	70
	22	tRNA-Pro	16,250	16,315	−	66
rRNA	1	s-rRNA	940	1,903	+	964
	2	l-rRNA	1,969	3,548	+	1,580
Gene	1	ND1	3,627	4,577	+	951
	2	ND2	4,792	5,820	+	1,029
	3	COX1	6,219	7,757	+	1,539
	4	COX2	7,903	8,583	+	681
	5	ATP8	8,660	8,854	+	195
	6	ATP6	8,821	9,495	+	675
	7	COX3	9,501	10,283	+	783
	8	ND3	10,354	10,698	+	345
	9	ND4L	10,770	11,063	+	294
	10	ND4	11,060	12,427	+	1,368
	11	ND5	12,636	14,438	+	1,803
	12	ND6	14,446	14,964	−	519
	13	CYTB	15,041	16,171	+	1,131

from truncation of PR1, removing 4 bp at the 5' end (Choi et al. 2010). None of the 5 sun bears possessed an imperfect repeat, but all possessed PR1 and PR2. The number of repeats for sun bears MN807949, EF196664, FM177765, Anabell, and Klaus were 23, 16, 26, 17, and 18 for PR1 and 8, 16, 7, 15, and 14 for PR2. Variation in the number of repeats in the microsatellite region was mostly responsible for the differences in genome lengths among the 5 sun bears.

Mean within-species pairwise nucleotide p-distances were greatest in the Asiatic black bear, followed by the American black bear, and least in polar bears (Table 2). Among sun bears, pairwise nucleotide p-distances ranged from 0.089% to 1.21% with FM177765 and Anabell possessing the smallest pairwise nucleotide difference (Table 3). Within-group pairwise p-distance for sun bears in the Yunnan clade (EF19664, FM 177765, Anabell) was 0.163% (SE = 0.032) and 0.227% (SE = 0.039) for bears in the Malaysian clade (MN807949 and Klaus). Nucleotide p-distance between the 2 clades of sun bears was 1.17% (SE = 0.083).

Maximum likelihood (ML) and Bayesian inference (BI) phylogenetic trees produced identical tree topologies (Fig. 2). The giant panda lineage formed the basal branch of the family Ursidae, followed by 2 distinct clusters, one corresponding to the Andean bear and the other including all 6 ursine species. Both trees displayed the sloth bear as the basal branch of the subfamily Ursinae with strong support. Bayesian probabilities were high for all branch nodes, but ML support was weak for the remaining 5 ursine species, which clustered in 2 major clades. One clade was composed of the polar bear and brown bear with robust support; a second well-supported clade consisted of the sun bear as a sister taxon to the cluster comprising the American black bear and Asiatic black bear. Sun bears formed 2 well-supported clades: the Malaysian sun bear and another sun bear of unidentified origin occupied one clade; a sun bear from Yunnan Province, China, and 2 others of identified origin occupied the other.

Discussion

The complete mitochondrial genome (16,770 bp) of a sun bear from Peninsular Malaysia was sequenced and characterized in detail. This is the first mitochondrial genome of a sun bear successfully sequenced in Southeast Asia using NGS technology. The overall characteristics, including mitochondrial genome organization and gene arrangement pattern, were identical to those of other ursids (Delisle and Strobeck 2002, Hou et al. 2007, Peng et al. 2007, Yu et al. 2007, Krause et al. 2008). The control region is the most variable stretch in the sun bear mitochondrial genome. The microsatellite repeat region in sun bears is variable. This region, with different types and numbers of repeats among bear species, has been described as promising for species and subspecies differentiation in the Ursidae (Hoelzel et al. 1994, Hou et al. 2007, Choi et al. 2010). Although all 5 genomes consistently lacked a 6-bp imperfect repeat, the 10-bp PR regions (PR1 and PR2) in the sun bear mitochondrial genomes were variable among all individuals. The considerable variability in this microsatellite region suggests that it may have potential for phylogeographic studies in this species.

Table 2. Pairwise p-distances (d) within Ursidae species estimated using whole mitochondrial genome sequences. Pairwise distances are expressed as the percentage of base differences per site, averaging over all sequence pairs within each group.

Species		No. of individuals	Pairwise p-distance	SE
Giant panda	*Ailuropoda melanoleuca*	4	0.421	0.042
Polar bear	*Ursus maritimus*	5	0.219	0.024
Brown bear	*Ursus arctos*	4	0.835	0.038
Asiatic black bear	*Ursus thibetanus*	2	2.997	0.155
American black bear	*Ursus americanus*	3	1.259	0.066
Sun bear	*Helarctos malayanus*	5	0.741	0.049
Sloth bear	*Melursus ursinus*	2	0.937	0.078
Andean bear	*Tremarctos ornatus*	2	0.271	0.047

Most branching events in the phylogenetic trees we generated were supported by high bootstrap values and posterior probabilities, and were identical in topology to trees generated by Yu et al. (2007) and Krause et al. (2008). Yu et al. (2007) used one mitochondrial genome for each extant bear species, but excluded the tandem repeats region of the genome; Krause et al. (2008) used one mitochondrial genome for each extant and 2 extinct bear species, but omitted the control region of the genome. Similar to Kumar et al.'s 2017 observation, we observed 2 distinct mitochondrial lineages in sun bears. Kumar et al. (2017) believed the 2 lineages might correspond to the subspecies in Borneo (*U. m. euryspilus*) and *U. m. malayanus* from the Southeast Asian mainland, Sumatra, and the Malay Peninsula. However, the newly sequenced sun bear mitochondrial genome from Peninsular Malaysia occupies a clade separate to that containing a sun bear from Yunnan, China, suggesting a possible historical separation between sun bears in the Sundaic region (Malay Peninsula, Borneo, and Sumatra) and those in other parts of the Southeast Asian mainland. Sequence divergence (p-distance) values were also smaller within the 2 clades than between them. Frequent fluctuation in sea levels in Sundaland during the late Pleistocene has been hypothesized to have led to strong faunal disjunctions between Indochina and the Sundaic region, including diversification within Sundaland (Woodruff and Turner 2009). More genomic data from known geographic locations in sun bear range, particularly from the Sundaic region, is needed to clarify the evolutionary history and may identify potentially significant conservation units in this species.

Trees based on whole mitochondrial genomes have consistently placed the sloth bear and sun bear as the 2 most basal ursine species and the American and Asiatic black bears as sister taxa (Fig. 3; Yu et al. 2007, Krause et al. 2008, this study). Trees employing nuclear genes have not reflected these relationships (e.g., Pagés et al. 2008, Kutschera et al. 2014). Discrepancies between nuclear and mitochondrial phylogenies are well-documented in the literature and do not necessarily imply that one or both forms of data are inaccurate; such conflicts can often lead to an enhanced understanding of the evolutionary history of the taxa (Rubinoff and Holland 2005). Conflicting trees generated from mitochondrial, autosomal, and Y-chromosome data led Kutschera et al. (2014) to track the history of individual alleles and conclude that differential sorting and introgression events had influenced evolutionary history in the Ursidae. Employing larger sets of heritable characters further strengthens phylogenetic inference (Grandcolas et al. 2001). A recent coalescent species tree including >18,000 genome

Table 3. Pairwise p-distance (d) among 5 sun bears (*Helarctos malayanus*) estimated from whole mitochondrial genome sequences. Standard error estimates are shown above the diagonal. Pairwise distances (below diagonal) are expressed as the percentage of base differences per site between sequences.

Sun bear mitochondrial genomes	EF196664	FM177765	MN807949	Anabell	Klaus
EF196664 (Yu et al. 2004)		0.040	0.087	0.034	0.089
FM177765 (Krause et al. 2008)	0.233		0.077	0.023	0.089
MN807949 (This study)	1.163	1.181		0.077	0.040
Anabell (Kumar et al. 2017)	0.167	0.089	1.211		0.081
Klaus (Kumar et al. 2017)	1.103	1.205	0.227	1.175	

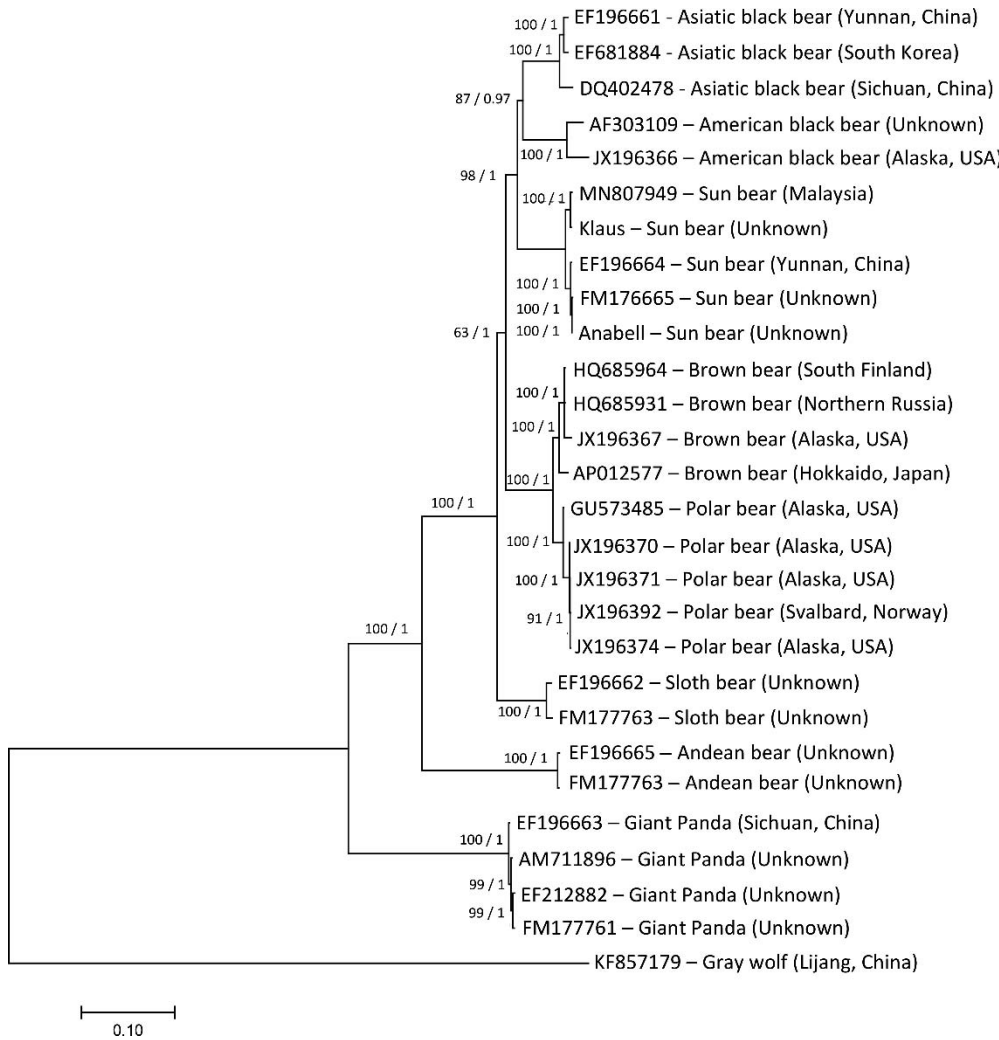

Fig. 2. Maximum likelihood tree of the phylogeny of the Ursidae based on whole mitochondrial genomes and the General Time Reversible model. The gray wolf (*Canis lupus*) was used as the outgroup. Support values for nodes represent % bootstraps obtained from maximum likelihood analyses and Bayesian posterior probabilities. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 28 nucleotide sequences.

fragments produced a well-resolved ursid phylogeny (Fig. 3; Kumar et al. 2017).

We did not detect greater pairwise nucleotide differences in older bear species or in bear species with larger geographic ranges (e.g., brown bears). Older lineages, such as the giant panda and Andean bear, would be expected to have accumulated more variation in the genome. The current distribution ranges of the Andean bear and giant panda are relatively small and may represent only a small part of their historical range, compared with the American and Asiatic black bears, which possessed the greatest within-clade sequence divergence. We caution, however, that these results are based on a very limited number of genomes.

Mitochondrial DNA still remains invaluable for resolving taxonomic uncertainties at the level of family, genus, and species, and is a well-established tool for detecting hybridization events, direction of gene flow, and species limits (Rubinoff and Holland 2005). The noncoding control region is especially useful for

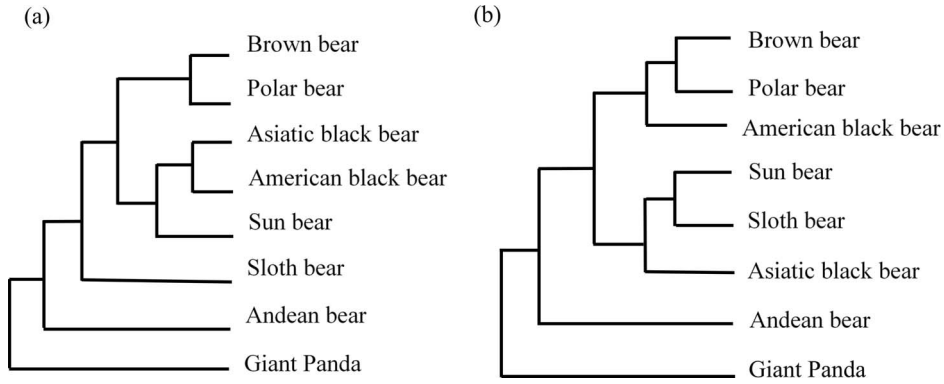

Fig. 3. (a) Ursidae phylogeny based on whole mitochondrial genomes (Yu et al. 2007, Krause et al. 2008, Kutschera et al. 2014, this study) and (b) Coalescent species tree for the Ursidae based on nonoverlapping whole genome fragments (Kumar et al. 2017).

investigating intraspecific genetic variation and elucidating the structure of contemporary populations (Wan et al. 2004, Gupta et al. 2015). Both partial and whole mitochondrial genomes have proved valuable for inferring evolutionary history and identifying significant conservation units within species (e.g., Patel et al. 2017). Mitochondria exist in such large copy numbers in animal cells that well-chosen markers may be amplified from even partially degraded samples. For species such as sun bears, whose cubs and body parts are heavily trafficked (Shepherd and Shepherd 2010, Gomez et al. 2020), mitochondrial markers can provide the baseline data to identify the origins of those individuals. Using mitochondrial markers developed from this newly sequenced mitochondrial genome, we successfully amplified DNA from noninvasively recovered root bulbs from sun bear hair (Tee et al. 2020). For population genetic studies employing partially degraded or trace amounts of genetic material, mitochondrial markers may lead to better success than nuclear markers and will thus continue to play a pivotal role in galvanizing population studies and improving forensic efforts for this imperiled species.

Acknowledgments

This work was supported by the Sunway Internal Grant (INT-FST-BIOS-0115-02) funded by Sunway University, Malaysia, and the International Association of Bear Research and Management (IBA). We thank A. Kannan for help with phylogenetic analysis. Also, we are grateful for all the valuable comments from the Associate Editor and reviewers, which helped to improve this manuscript.

Literature cited

BERNT, M., A. DONATH, F. JÜHLING, F. EXTERNBRINK, C. FLORENTZ, G. FRITZSCH, J. PÜTZ, M. MIDDENDORF, AND P.F. STADLER. 2013. MITOS: Improved de novo metazoan mitochondrial genome annotation. Molecular Phylogenetics and Evolution 69:313–319.

CHOI, E.H., S.K. KIM, S.H. RYU, K.H. JANG, AND U.W. HWANG. 2010. Mitochondrial genome phylogeny among Asiatic black bear *Ursus thibetanus* subspecies and comprehensive analysis of their control regions. Mitochondrial DNA 21:105–114.

DELISLE, I., AND C. STROBECK. 2002. Conserved primers for rapid sequencing of the complete mitochondrial genome from carnivores, applied to three species of bears. Molecular Biology and Evolution 19:357–361.

DESALLE, R., B. SCHIERWATER, AND H. HADRYS. 2017. MtDNA: The small workhorse of evolutionary studies. Frontiers in Bioscience 22:873–887.

GOMEZ, L., C.R. SHEPHERD, AND M.S. KHOO. 2020. Illegal trade of sun bear parts in the Malaysian states of Sabah and Sarawak. Endangered Species Research 41: 279–287.

GRANDCOLAS, P., P. DELEPORTE, L. DESUTTER-GRANDCOLAS, AND C. DAUGERON. 2001. Phylogenetics and ecology: As many characters as possible should be included in the cladistic analysis. Cladistics 17:104–110.

GUPTA, A., A. BHARDWAJ, P. SHARMA, AND Y. PAL. 2015. Mitochondrial DNA—A tool for phylogenetic and biodiversity search in equines. Journal of Biodiversity & Endangered Species S1:S1.006. doi:10.4172/2332-2543.S1-006.

HALL, T.A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95–98.

HASHIMOTO, T., E. OTAKA, J. ADACHI, K. MIZUTA, AND M. HASEGAWA. 1993. The giant panda is closer to a bear, judged by α-and β-hemoglobin sequences. Journal of Molecular Evolution 36:282–289.

HOELZEL, A.R., J.V. LOPEZ, G.A. DOVER, AND S.J. O'BRIEN. 1994. Rapid evolution of a heteroplasmic repetitive sequence in the mitochondrial DNA control region of carnivores. Journal of Molecular Evolution 39:191–199.

HOU, W.-r., Y. CHEN, X. WU, J.-c. HU, Z.-s. PENG, J. YANG, Z.-x. TANG, C.-Q. ZHOU, Y.-m. LI, AND S.-k. YANG. 2007. A complete mitochondrial genome sequence of Asian black bear Sichuan subspecies (*Ursus thibetanus mupinensis*). International Journal of Biological Sciences 3: 85.

KEARSE, M., R. MOIR, A. WILSON, S. STONES-HAVAS, M. CHEUNG, S. STURROCK, S. BUXTON, A. COOPER, S. MARKOWITZ, AND C. DURAN. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.

KRAUSE, J., T. UNGER, A. NOÇON, A. MALASPINAS, S. KOLOKOTRONIS, M. STILLER, L. SOIBELZON, H. SPRIGGS, P.H. DEAR, A.W. BRIGGS, S.C. BRAY, S.J. O'BRIEN, G. RABEDER, P. MATHEUS, A. COOPER, M. SLATKIN, S. PAABO, AND M. HOFREITER. 2008. Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene–Pliocene boundary. BMC Evolutionary Biology 8:220. doi: 10.1186/1471-2148-8-220.

KUMAR S., G. STECHER, AND K. TAMURA. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. Molecular Biology and Evolution 33:1870–1874.

KUMAR, V., F. LAMMERS, T. BIDON, M. PFENNINGER, L. KOLTER, M.A. NILSSON, AND A. JANKE. 2017. The evolutionary history of bears is characterized by gene flow across species. Scientific Reports 7:46487. doi: 10.1038/srep46487.

KURTEN, B. 2017. Pleistocene mammals of Europe. Routledge, New York, New York, USA.

KUTSCHERA, V.E., T. BIDON, F. HAILER, J.L. RODI, S.R. FAIN, AND A. JANKE. 2014. Bears in a forest of gene trees: Phylogenetic inference is complicated by incomplete lineage sorting and gene flow. Molecular Biology and Evolution 31: 2004–2017.

MEIJAARD, E. 2004. Craniometric differences among Malayan sun bears (*Ursus malayanus*); evolutionary and taxonomic implications. Raffles Bulletin of Zoology 52: 665–672.

O'BRIEN, S.J., W.G. NASH, D.E. WILDT, M.E. BUSH, AND R.E. BENVENISTE. 1985. A molecular solution to the riddle of the giant panda's phylogeny. Nature 317:140–144.

PAGES, M., S. CALVIGNAC, C. KLEIN, M. PARIS, S. HUGHES, AND C. HÄNNI. 2008. Combined analysis of fourteen nuclear genes refines the Ursidae phylogeny. Molecular Phylogenetics and Evolution 47:73–83.

PATEL, R.P., S. WUTKE, D. LENZ, S. MUKHERJEE, U. RAMAKRISHNAN, G. VERON, J. FICKEL, A. WILTING, AND D.W. FÖRSTER. 2017. Genetic structure and phylogeography of the leopard cat (*Prionailurus bengalensis*) inferred from mitochondrial genomes. Journal of Heredity 108(4):349–360. doi:10.1093/jhered/esx017.

PENG, R., B. ZENG, X. MENG, B. YUE, Z. ZHANG, AND F. ZOU. 2007. The complete mitochondrial genome and phylogenetic analysis of the giant panda (*Ailuropoda melanoleuca*). Gene 397:76–83.

RAMBAUT, A. 2012. Rambaut A. FigTree v.1.4.2: Tree figure drawing tool. http://tree.bio.ed.ac.uk/software/figtree. Accessed 30 Sep 2020.

———, A.J. DRUMMOND, D. XIE, G. BAELE, AND M.A. SUCHARD. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic Biology 67:901–904. doi: 10.1093/sysbio/syy032.

RONQUIST, F., J. HUELSENBECK, AND M. TESLENKO. 2011. Draft MrBayes version 3.2 manual: Tutorials and model summaries. Software available from http://brahms.biology.rochester.edu/software.html,1-105.

RUBINOFF, D., AND B.S. HOLLAND. 2005. Between two extremes: Mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Systematic Biology 54:952–961.

SCOTSON, L., G. FREDRIKSSON, D. AUGERI, C. CHEAH, D. NGOPRASERT, AND W. WAI-MING. 2017. *Helarctos malayanus* (errata version published in 2018). The IUCN Red List of Threatened Species 2017:e. T9760A123798233.

SERVHEEN, C. 1999. Bears: Status survey and conservation action plan. International Union for Conservation of Nature, Gland, Switzerland, and Cambridge, England, UK.

SHEPHERD, C.R., AND L. SHEPHERD. 2010. The poaching and trade of Malayan sun bears in Peninsular Malaysia. Traffic Bulletin 23:49–52.

TALBOT, S.L., AND G.F. SHIELDS. 1996. A phylogeny of the bears (Ursidae) inferred from complete sequences of three mitochondrial genes. Molecular Phylogenetics and Evolution 5:567–575.

TEE, T.T., W.L. LAI, T.K.J. KOK, Z.H. OOI, F.T. VAN MANEN, S.P. SHARP, S.T. WONG, J. CHEW, AND S. RATNAYEKE. 2020. An evaluation of non-invasive sampling techniques for Malayan sun bears. Ursus 31:e16. doi: 10.2192/URSUS-S-20-00004.1.

THOMPSON, J.D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN, D.G. HIGGINS. 1997. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25:4876-82. doi: 10.1093/nar/25.24.4876. PMID: 9396791; PMCID: PMC147148.

WAITS, L.P., J. SULLIVAN, S.J. O'BRIEN, AND R. WARD. 1999. Rapid radiation events in the family Ursidae indicated by likelihood phylogenetic estimation from multiple fragments of mtDNA. Molecular Phylogenetics and Evolution 13: 82–92.

———, S.L. TALBOT, R. WARD, AND G. SHIELDS. 1998. Mitochondrial DNA phylogeography of the North American brown bear and implications for conservation. Conservation Biology 12:408–417.

WAN, Q.H., H. WU, T. FUJIHARA, AND S.G. FANG. 2004. Which genetic marker for which conservation genetics issue? Electrophoresis 25:2165–2176.

WOODRUFF, D.S., AND L.M. TURNER. 2009. The Indochinese–Sundaic zoogeographic transition: A description and analysis of terrestrial mammal species distributions. Journal of Biogeography. doi: 10.1111/j.1365-2699.2008.02071.x.

YU, L., Q-W. LI, O. RYDER, AND Y-P. ZHANG. 2004. Phylogeny of the bears (Ursidae) based on nuclear and mitochondrial genes. Molecular Phylogenetics and Evolution 32:480–494.

———, ———, ———, AND ———. 2007. Analysis of complete mitochondrial genome sequences increases phylogenetic resolution of bears (Ursidae), a mammalian family that experienced rapid speciation. BMC Evolutionary Biology 7:1. doi: 10.1186/1471-2148-7-198.

ZHANG, Y-P., AND O.A. RYDER. 1994. Phylogenetic relationships of bears (the Ursidae) inferred from mitochondrial DNA sequences. Molecular Phylogenetics and Evolution 3:351–359.

Received: December 6, 2019
Accepted: December 16, 2020
Associate Editor: E. Bellemain