Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 47 (2008) 73–83
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Combined analysis of fourteen nuclear genes refines the Ursidae phylogeny

Marie Pagès [a,b], Sébastien Calvignac [a], Catherine Klein [a], Mathilde Paris [c], Sandrine Hughes [a], Catherine Hänni [a,*]

[a] *Paléogénétique et Evolution Moléculaire, Université de Lyon, Institut de Génomique Fonctionnelle, INRA, CNRS, Université Lyon 1, Ecole Normale Supérieure, 46 allée d'Italie, 69364 Lyon cedex 07, France*
[b] *Laboratoire de Paléontologie, Phylogénie et Paléobiologie, Institut des Sciences de l'Evolution de Montpellier, UMR CNRS 5554, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France*
[c] *Zoologie Moléculaire, Université de Lyon, Institut de Génomique Fonctionnelle, INRA, CNRS, Université Lyon 1, Ecole Normale Supérieure, 46 allée d'Italie, 69364 Lyon cedex 07, France*

Received 22 February 2007; revised 11 October 2007; accepted 17 October 2007
Available online 5 November 2007

Abstract

Despite numerous studies, questions remain about the evolutionary history of Ursidae and additional independent genetic markers were needed to elucidate these ambiguities. For this purpose, we sequenced ten nuclear genes for all the eight extant bear species. By combining these new sequences with those of four other recently published nuclear markers, we provide new insights into the phylogenetic relationships of the Ursidae family members. The hypothesis that the giant panda was the first species to diverge among ursids is definitively confirmed and the precise branching order within the *Ursus* genus is clarified for the first time. Moreover, our analyses indicate that the American and the Asiatic black bears do not cluster as sister taxa, as had been previously hypothesised. Sun and sloth bears clearly appear as the most basal ursine species but uncertainties about their exact relationships remain. Since our larger dataset did not enable us to clarify this last question, identifying rare genomic changes in bear genomes could be a promising solution for further studies.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Molecular phylogeny; Nuclear DNA; Hibernation; Ursidae

1. Introduction

Among Carnivores, the family of Ursidae comprises eight extant species, including the giant panda (*Ailuropoda melanoleuca*), the spectacled bear (*Tremarctos ornatus*) and six ursine species: the polar bear (*Ursus maritimus*), the brown bear (*Ursus arctos*), the American black bear (*Ursus americanus*), the Asiatic black bear (*Ursus thibetanus*), the sun bear (*Helarctos malayanus*) and the sloth bear (*Melursus ursinus*). Although the belonging of the giant panda to the Ursidae is not a riddle anymore (Davis, 1964; O'Brien et al., 1985; Van Valen, 1986; Mayr, 1986), the phylogenetic relationships among ursids still remain unclear. The fact that ursine species underwent a rapid diversification 5 Myr[1] ago (Kurtén, 1968; Wayne et al., 1991; Waits et al., 1999) could explain why resolving this phylogeny represents a tricky challenge. Despite previous molecular investigations, mainly based on mitochondrial DNA (Zhang and Ryder, 1994; Vrana et al., 1994; Talbot and Shields, 1996a; Waits et al., 1999; Yu et al., 2004; Fulton and Strobeck, 2006), three thorny questions are still pending: (Q1) what is the exact relationship between the giant panda and the spectacled bear at the base of the Ursidae

* Corresponding author. Address: Institut de Génomique Fonctionnelle, INRA, CNRS, Université Lyon 1, Ecole Normale Supérieure, F-69364, France. Fax: +33 0 4 72 72 80 80.
E-mail address: Catherine.Hanni@ens-lyon.fr (C. Hänni).

[1] *Abbreviations used:* Myr, million years; kb, kilobases.

1055-7903/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2007.10.019

radiation? (Q2) is the supposed sistership of the two black bears, *Ursus americanus* and *Ursus thibetanus* true?; (Q3) what is the branching order among the six ursine species?

A possible explanation of the inability to resolve these uncertainties is that highly variable molecular markers such as mitochondrial ones are too homoplasic (Talbot and Shields, 1996a; Waits et al., 1999). On the contrary, slowly evolving sequences, such as the only two nuclear markers available for all Ursidae species (transthyretin intron 1 and interphotoreceptor retinoid binding protein exon 1), are too poorly informative to allow robust phylogenetic inferences (Yu et al., 2004). However, concatenating numerous slowly evolving sequences is likely to overcome the lack of phylogenetic signal (Gadagkar et al., 2005). This strategy has already solved problematic relationships inside recently diversified families of other carnivores (see Bardeleben et al., 2005). Y-linked markers could in particular be valuable targets. Under the hypothesis of a male driven evolution, Y-linked genetic markers are supposed to be less variable than mitochondrial markers (and so, less homoplasic) but more variable than X-linked or autosomal genes (Haldane, 1947; Lessells, 1997; Pecon-Slattery and O'Brien, 1998). These markers already proved to be powerful to reconstruct phylogenetic trees concerning mammals (Pecon-Slattery and O'Brien, 1998; Nishida et al., 2003; Moreira, 2002; Pecon-Slattery et al., 2004).

In this paper, we provide the sequences of nine additional autosomal gene fragments and one Y-linked marker. Phylogenetic analyses performed on a concatenated dataset composed of these new sequences (nearly 6 kb) and four other recently published markers (nearly 4 kb) (two Y-linked and two autosomal ones; Yu et al., 2004; Pagés et al., submitted for publication) allow us to address each of the three still unresolved questions (Q1, Q2 and Q3) concerning the evolutionary history of the Ursidae.

2. Materials and methods

2.1. Samples

Samples were selected from a variety of sources including blood, tissues, and hairs from free ranging bears or individuals from zoos. A total number of 9 female and 13 male bears, representing all the eight species, were sampled (Table 1).

2.2. Choice of the nuclear genes for phylogenetic analyses

First, we selected three nuclear genetic markers that were recently used to resolve phylogenies of carnivore families that diversified at the same time as Ursidae (15–20 Myr, Waits et al., 1999): TRSP (selenocysteine tRNA) and FES (feline sarcoma protooncogene) were proved to be valuable for the resolution of the Canidae phylogeny (Canidae, 12 Myr, Bardeleben et al., 2005), and UBE1Y (ubiquitin activating enzyme E1 on Y) for the Felidae phylogeny (12–15 Myr, Pecon-Slattery et al., 2004). We also chose to sequence another marker, the vWF (von Willebrand Factor), recently used to infer the phylogeny of a rodent family (Echimyidae; Galewski et al., 2005).

Table 1
Ursid samples used in this study

Subfamily	Scientific name	Common name	Sex	Number of samples	Type	Sample source
Ailuropodinae	*Ailuropoda melanoleuca*	Giant panda	Male	1	Tissue	Collection of tissues conserved in alcohol, MNHN (France)
Tremarctinae	*Tremarctos ornatus*	Spectacled bear	Male	1	Hairs	Parc de la Tête d'Or (Lyon, France)
Ursinae			Female	1	Tissue	F. Catzeflis (University of Montpellier, France)
	Ursus thibethanus	Asiatic black bear	Male	1	Hairs and tissue	Réserve Africaine de Sigean (France)
	Melursus ursinus	Sloth bear	Male	1	Hairs	Moscow Zoo (Russia)
			Male	1	Hairs	Leipzig Zoo (Germany)
			Female	1	Hairs	Leipzig Zoo (Germany)
	Helarctos malayanus	Sun bear	Male	1	Hairs	Parc zoologique de Paris (France)
			Female	1	Hairs	Zoo de Saint Martin La Plaine (France)
	Ursus americanus	American black bear	Male	1	Tissue	F. Catzeflis (University of Montpellier, France)
			Male	1	DNA extract	I. Delisle (University of Alberta, Canada)
			Female	1	Hairs and blood	Zoo de Peaugres (France)
			Female	1	DNA extract	I. Delisle (University of Alberta, Canada)
	Ursus arctos	Brown bear	Male	1	DNA extract	I. Delisle (University of Alberta, Canada)
			Male	2	DNA extract	P. Taberlet (LECA, Grenoble, France)
			Unknown	1	Hairs and tissue	P. Taberlet (LECA, Grenoble, France)
			Female	3	DNA extract	P. Taberlet (LECA, Grenoble, France)
	Ursus maritimus	Polar bear	Male	1	Hairs	Zoo de la Palmyre (France)
			Male	1	DNA extract	I. Delisle (University of Alberta, Canada)
			Female	2	Hairs	Zoo de la Palmyre (France)
			Female	1	DNA extract	I. Delisle (University of Alberta, Canada)

Taxonomic denomination follows the classification of Wozencraft (1993).

Six supplemental genes, all involved into the thyroid hormone pathway, were targeted: DI1 (iodothyronine deiodinase type I), DI2 (iodothyronine deiodinase type II), TG (thyroglobulin), SIS (sodium iodide symporter), TRH (thyrotropin-releasing hormone), TSHβ (thyroid stimulating hormone beta-subunit). These markers were initially sequenced to determine whether hibernating and non-hibernating bear species had been submitted to different selective pressures on this pathway. As no positive selection was detected and since these markers appeared to be under neutral evolution (data not shown), we used these six nuclear markers for a phylogenetic purpose.

Finally, four nuclear markers already available for all ursid species were used in this study: SRY (sex determining region of the Y chromosome), ZFY (zinc finger protein on Y), TTR (transthyretin) and IRBP (interphotoreceptor retinoid binding protein) (Yu et al., 2004; Pagés et al., submitted for publication).

In this way, a total of 14 nuclear genes were considered for the phylogenetic analyses (see Table 2 for accession numbers).

2.3. DNA isolation, PCR amplification and sequencing

To avoid contaminations, pre-amplification procedures and post-amplification analyses were performed in independent rooms. DNA was extracted from blood and hair samples with QiAmp DNA Mini Kit (Qiagen) and from tissue with DNEasy Tissue Kit (Qiagen), both in accordance with the manufacturer's instructions.

Primer sets used to amplify the UBE1Y, TRSP, FES and vWF genes were found in the literature (Table 3). UBE1Y primers were tested on female bear samples in order to ensure their specificity for the Y copy. All amplifications were carried out in 25 μL containing about 30 ng of DNA extract, 0.2 mg/mL BSA (Roche, 1 mg/mL), 250 μM of each dNTP, 0.2 μM of each primer, 1 unit of Perkin Elmer Gold Taq polymerase (Applied Biosystems), 2.5 μL of 10× buffer, 1.5 mM of MgCl2. Cycling conditions were as follows: one activation step at 94 °C for 5 min followed by 40 cycles of denaturation at 94 °C for 30 s, annealing at 48–60 °C depending on the primers (Table 3) for 30 s, elongation at 72 °C for 45 s–1′30 min depending on the length of the target, and a final extension at 72 °C for 7 min.

To amplify the genes implicated in the thyroid pathway, we used a semi-nested PCR approach. For each fragment, two sets of degenerate primers were designed from conserved regions determined on the alignment of orthologous sequences of other mammals available in databank (Table 3). Reaction mixes were identical to that mentioned above. Two successive touch-down PCR were performed. To amplify each gene for each species, two different touch-down thermal cycle programs were used. It consisted of seven steps described in the Supplementary Data S1.

When PCR products proved to be difficult to sequence, they were cloned by using Topo TA Cloning for Sequencing kit (Invitrogen) and several clones were analysed. Clones and PCR products were sequenced by a service provider (Genome Express, France). All the sequences generated in this study were deposited in EMBL under the Accession Nos. AM748289 to AM748384 (Table 2).

2.4. Phylogenetic analyses

2.4.1. Sequence analyses

The nuclear sequences generated in this study and the already published ones (Table 2) were aligned by eye using

Table 2
Accession numbers of nuclear gene sequences used in this study

	Ursus maritimus	*Ursus arctos*	*Ursus americanus*	*Helarctos malayanus*	*Melursus ursinus*	*Ursus thibetanus*	*Tremarctos ornatus*	*Ailuropoda melanoleuca*	*Canis familiaris*
vWF	AM748289[d]	AM748290[d]	AM748291[d]	AM748292[d]	AM748293[d]	AM748294[d]	AM748295[d]	AM748296[d]	AF099154
ZFY	AM748297[c]	AM748298[c]	AM748299[c]	AM748300[c]	AM748301[c]	AM748302[c]	AM748303[c]	AM748304[c]	AF393756
SRY	AM748305[c]	AM748306[c]	AM748307[c]	AM748308[c]	AM748309[c]	AM748310[c]	AM748311[c]	AM748312[c]	AF107021[d]
IRBP	AY303843[b]	AY303842[b]	AY303837[b]	AY303839[b]	AY303838[b]	AY303841[b]	AY303840[b]	AY303836[b]	XM_546201
TTR	AY303848[b]	AF039741[a]	AY303844[b]	AY303846[b]	AY303845[b]	AY303847[b]	AF039740[a]	AF039738[a]	AY50579
TRSP	AM748313[d]	AM748314[d]	AM748315[d]	AM748316[d]	AM748317[d]	AM748318[d]	AM748319[d]	AM748320[d]	AY609084
FES	M748321[d]	AM748322[d]	AM748323[d]	AM748324[d]	AM748325[d]	AM748326[d]	AM748327[d]	AM748328[d]	AY885365
UBE1Y	AM748329[d]	AM748330[d]	AM748331[d]	AM748332[d]	AM748333[d]	AM748334[d]	AM748335[d]	AM748336[d]	
DI1	AM748337[d]	AM748338[d]	AM748339[d]	AM748340[d]	AM748341[d]	AM748342[d]	AM748343[d]	AM748344[d]	XM_536701
DI2	AM748345[d]	AM748346[d]	AM748347[d]	AM748348[d]	AM748349[d]	AM748350[d]	AM748351[d]	AM748352[d]	ENSCAFT00000027378
TG	AM748353[d]	AM748354[d]	AM748355[d]	AM748356[d]	AM748357[d]	AM748358[d]	AM748359[d]	AM748360[d]	ENSCAFT00000001727
SIS	AM748361[d]	AM748362[d]	AM748363[d]	AM748364[d]	AM748365[d]	AM748366[d]	AM748367[d]	AM748368[d]	XM_541946
TSH beta	AM748369[d]	AM748370[d]	AM748371[d]	AM748372[d]	AM748373[d]	AM748374[d]	AM748375[d]	AM748376[d]	U15644
TRH	AM748377[d]	AM748378[d]	AM748379[d]	AM748380[d]	AM748381[d]	AM748382[d]	AM748383[d]	AM748384[d]	XM_541757

Missing data: *Canis familiaris*, UBEIY, SIS INTRON 6, TSH*β* intron 1; *Ailuropoda melanoleuca*, UBE1Y intron 18.

[a] Flynn and Nedbal, 1998.
[b] Yu et al., 2004.
[c] Pagès et al., submitted for publication.
[d] This study.

Table 3
Abbreviation, name, sequence and original reference of primers used in this study

Designation	Gene name and amplified region	Nucleotide sequence 5′ → 3′	Annealing temperature	Original publication
vWF	von Willebrand Factor gene exon 28			
vWF-A		CTGTGATGGTGTCAACCTCACCTGTGAAGCCTG	60 °C	Porter et al., 1996
vWF-B		TCGGGGGAGCGTCTCAAAGTCCTGGATGA		
TRSP	Selenocyteine tRNA gene 5′ gene-flanking and tRNA coding regions			
TRSP-1F		GGGCTTCTGAAAGCCGACTT	50 °C	Bardeleben et al., 2005
TRSP-1R		CCGCCCGAAAGGTGGAATTG		
FES	Feline sarcoma protooncogene exon 13–intron 13–exon 14			
FESF		GGGGAACTTTGGCGAAGTGTT	50 °C	Venta et al., 1996
FESR		TCCATGACGATGTAGATGGG		
UBE1Y	Ubiquitin activating enzyme E1 on Y intron 17–exon 18–intron 18			
UBEF		GCTCTTCAAGCAGTCAGCTGAA	48 °C	Chang and Li, 1995
UBER		ATCCAGATGTAGGGG		
UBE-exon-forl[a]		AGACTCCAAKTTTGTGGAACG		This study
UBE-exon-revl[a]		TGKTCAGGTGGAAAGTTGTGC		
DI1[c]	Iodothyronine deiodinase type I exon 1			
DIl_forT2		CCCDGYYAGYGCTGTGG		
DIl_forl3		TCTGGGTSCTCTTKSAGGT		
Dllrevll		TCCCARATGYTGCACCTCTG	TC	This study
DIl_revl2		SYKGACCACVGGRCAGT		
DIlurs-Fl[b]		GGCCATGCACGTGGCCG		This study
DIlurs-Rl[b]		TCCTGAGAGCYGGACCAC		
DI2[c]	Iodothyronine deiodinase type II exon 2			
DI2-F21		GTGAAAYTGGGTGARGATGC		
DI2-F22		CCCAATTCCMGYGTGGTGC	TC	This study
DI2-R21		CTCTTGCTGAAATTCTTCTCC		
DI2-R22		GCCAAYGCCGGACTTCTTG		
TG[c]	Thyroglobulin exon 9			
TG-F91		STCMGARAGGMRGCAGGC		
TG-F92		GMCCTCVGGCTACTTCAG	TC	This study
TG-R91		SRCTTCGAGTTCCAGGAAT		
TG-R92		CACTCTGAGTTRAAGCACTG		
SIS[c]	Sodium iodide symporter exon 6–intron 6–exon 7			
SISF61		CYKGACCCGMGGMGCCG		
SIS_F62		GCCGCTAYACMTTCTGGAC	TC	This study
SISR71		BVARSAKRGGGTCRCAGTC		
SIS_R72		CTGGTCCGGGGCAGAGAT		
TRH[c]	Thyrotropin-releasing hormone exon 2			
TRH_for2		GTCAGCATCCWGGCAAAAG		
TRH_for22		GGYTCKCCAARCGTCAGC	TC	This study
TRH_rev2		CCTCCAGSGGYTCSCTG		
TRH_rev22		TGVCYYCKRCTCAGGTCA		
TSHB[c]	Thyroid stimulating hormone beta-subunit complete gene (exon 1–intron 1–exon 2)			
TSHb-for11		ACTGCTMYCTWYYTGATGTC		
TSHb-forl2		VTTTTTGGCCTDGCATGTGG	TC	This study
TSHb-rev21		CYAMCABATMRGACTTCTG		
TSHb-rev22		TTGYYTTRATGGCYTCATGT		

TC, touch down. (a) Used to amplify the exon 18 of UBE1Y of the taxa *Ailuropoda melanoleuca*; (b) the gene DI1 of the taxa *Helarctos malayanus*. (c) genes amplified by semi-nested PCR.

SEAVIEW (Galtier et al., 1996). Ambiguous regions due to nucleotide repetition were excluded from the alignment before performing the phylogenetic analyses.

To compare evolutionary rates between coding and non-coding sequences, pairwise divergence values were calculated for each part of the genes by computing uncorrected distances in PAUP v4.0b10 (Swofford, 1998) within the ursid species. GC-content at the third position for coding sequences and total GC-content for the non-coding sequences were computed using PHYLO_WIN (Galtier et al., 1996).

2.4.2. Datasets used

All the genes were concatenated and different datasets were used depending on the chosen outgroup. Outgroups were selected according to the phylogenetic issues to solve (Table 4). To address question (Q1), *Canis familiaris* was used as an outgroup for the analyses. Following this first analysis, the giant panda was then used as an outgroup to investigate other family member relationships (Q2 and Q3). Finally, since we did not succeed in amplifying the UBE1Y intron 18 for the giant panda, the spectacled bear was used as an outgroup in analyses where this intron was considered. Thus, three different datasets were generated (Table 4).

2.4.3. Phylogenetic analyses

The appropriate model of evolution was first determined for each region of each gene and then, for the three concatenated datasets, using MrModeltest 2.0 (Nylander et al., 2004), a Modeltest reduced version adapted to options available with Bayesian analysis softwares (Posada and Crandall, 1998). The selected models for each partition and dataset are shown in supplementary data S2. Maximum Likelihood (ML) analyses were performed with PHYML (Guindon and Gascuel, 2003) using the online interface http://atgc.lirmm.fr/phyml/(Guindon et al., 2005). For each analysis, the transition/transversion ratio, the proportion of invariable sites as well as the gamma distribution parameter (if necessary) were estimated and the starting tree was determined by BioNJ analysis of the datasets (default settings). Using optimization options, 1000

Table 4
Schematic representation of the three combined datasets used in this study to infer ursid phylogeny

Genes included in the datasets		Dataset 1 Ql *Out: Canis familiaris*	Dataset 2 Q2/Q3 *Out: Ailuropoda melanoleuca*	Dataset 3 Q2/Q3 *Out: Tremarctos ornatus*
vWF	exon 28	1076	1076	1076
ZFY	exon	350	397	397
SRY	5′ flanking region	417	438	438
	coding region	662	666	666
	3′ flanking region	152	167	168
IRBP	exon 1	1274	1274	1274
TTR	intron 1	946	984	987
TRSP	5′ flanking region	287	299	299
	tRNA	66	66	66
FES	exon 13	98	98	98
	intron 13	277	294	294
	exon 14	64	64	64
UBE1Y	intron 17		206	215
	exon 18		171	172
	intron 18			807
DI1	exon 1	192	192	192
DI2	exon 2	491	491	491
TG	exon 9	742	742	742
SIS	exon 6	117	117	117
	intron 6		80	80
	exon 7	111	111	111
TSHbeta	exon 1	115	115	115
	intron 1		389	389
	exon 2	166	166	167
TRH	exon 2	360	366	366
Total	in base pairs	7963	8969	9791
	% Coding regions	73.9	68.1	62.4
	% Non-coding regions	26.1	31.9	37.6
	% Y-linked regions	19.9	22.8	29.2

Each partition is associated to an outgoup (Out). Missing data for the outgroups are shaded. The ursid homologous sequences were consequently excluded from the partition. For each gene or part of gene, the number of aligned sites is given. Percentages of coding, non-coding and Y-linked regions are indicated for each dataset.

bootstrap (Bp) replicates were performed. ML analyses were first performed independently on each locus. Since the 14 genes yielded consistent, compatible topologies, analyses of the concatenated sequences were then carried out using the three different datasets described above.

Bayesian analyses (BA) were performed using MrBayes v3.1 (Ronquist and Huelsenbeck, 2003). Four independent runs of 5,000,000 generations each were performed applying appropriate independent models of evolution to each gene. A burn-in period of 100,000 generations was determined graphically using Tracer1.2 (Rambaud and Drumond, 2003), a software that allows an easy plotting of all parameters against the number of generations. For each dataset, all runs gave similar tree topologies and posterior probability (pp) values.

Alternative topologies were finally tested for significance using the Kishino-Hasegawa test (KH test) (Kishino and Hasegawa, 1989) and the Shimodaira-Hasegawa test (SH test) (Shimodaira and Hasegawa, 1999) (RELL option, 1000 Bp replicates) in PAUP* v4.0b10 (Swofford, 1998).

3. Results and discussion

3.1. Ursidae sequences analyses

The sequences of the ten targeted genes were generated for all the ursids, except one part of the UBE1Y fragment (intron 18) missing for the *A. melanoleuca* specimen. Information concerning these sequences is given in Table 5.

Proportionally to the sequence length, the number of variable sites within Ursidae ranged from 0.6% (1/166, TSHβ exon 2) to 5.1% (34/666, SRY coding sequence) among coding sequences, and from 3.3% (10/299, 5′ flanking region of TRSP; 13/389, TSHβ intron 1) to 6.3% (13/206, UBE1Y intron 17) among non-coding sequences. Y-chromosome fragments contained the highest number of variable sites either among coding or non-coding regions (5.1% SRY coding region; 6.3% UBE1Y intron 17).

For each gene used in this study, no significant difference in GC-contents was found between ursid sequences (data not shown).

Table 5
Characteristics of the nuclear Ursidae sequences used in this study

	Coding sequences				Non-coding sequences		
	NS	D	Vs	Pi	D	Vs	Pi
ZFY	397	2	8	3			
SRY							
5′untranslated region	438				2.5	16	3
Coding region	666	3.6	34	3			
3′untranslated region	167				4.1	8	1
UBE1Y							
Intron 17	206				5.7	13	2
Exon 18	171	1.7	3	0			
*Intron 18**	807				1.9	25	2
vWF	1076	3.1	45	9			
IRBP	1274	2	38	4			
TTR	984				3.3	48	13
TRSP							
5′ flanking region	299				3.3	10	2
tRNA coding region	66	3	2	0			
FES							
Exon 13	98	2	3	1			
Intron 13	294				2.7	11	1
Exon 14	64	3.1	2	0			
DI1	192	2.6	5	0			
DI2	491	1.2	6	2			
TG	742	1.7	19	3			
SIS							
Exon 6	117	2.6	3	1			
Intron 6	80				4.9	4	2
Exon 7	111	3.6	5	2			
TSHβ							
Exon 1	115	2.6	5	0			
Intron	389				1.8	13	1
Exon 2	166	0.6	1	0			
TRH	366	2.4	10	1			
Total	8969						
Coding regions	6112	2.4	189	29			
Non coding regions	2857				3.4	148	27

NS, number of aligned sites; D, highest uncorrected distance among Ursidae representatives (%); Vs, number of variable sites among Ursidae; Pi, number of parsimony informative sites. The asterisk indicates that all these parameters were computed without giant panda sequences.

3.2. New lights on the Ursidae evolutionary history

From the analysis of the 14 nuclear genes, we obtained the phylogenetic tree of the Ursidae family given in Fig. 1. BA and ML analyses of the three datasets gave identical topologies. Most relationships are well resolved (supports of 79–100% for Bp; 1.00 for pp), the only exceptions concern the phylogenetic positions of the sloth bear, *M. ursinus* and the sun bear, *H. malayanus*. This tree allows us to discuss the three last questions about the Ursidae phylogeny (Q1, Q2, Q3).

3.2.1. Q1: Basal ursid radiation

Three alternative hypotheses were initially proposed: (1) the giant panda, *Ailuropoda melanoleuca* may be the earliest species to diverge from other ursids, (2) alternatively, this may be the spectacled bear, *Tremarctos ornatus*, (3) or a common ancestor of the giant panda and the spectacled bear, i.e. these two bears may be a sister group. All these possibilities received strong to moderate supports in previous mitochondrial DNA analyses (Talbot and Shields, 1996a; Waits et al., 1999; Table 6). In the phylogenetic reconstructions inferred from our nuclear dataset 1 (*Canis familiaris* as outgroup), we observed that the giant panda unambiguously appears as the first ursid to diverge, followed by the spectacled bear (see supplementary data S3). This is supported by the highest values of bootstrap or posterior probabilities (ML 100% Bp; BA 1.00 pp; Table 6). In addition, topology tests performed on the dataset 1 reject the two other alternative hypotheses (KH test and SH test, $P < 0.03$). These findings are congruent with the recently published molecular phylogeny of the Arctoidea (Fulton and Strobeck, 2006) based on four nuclear

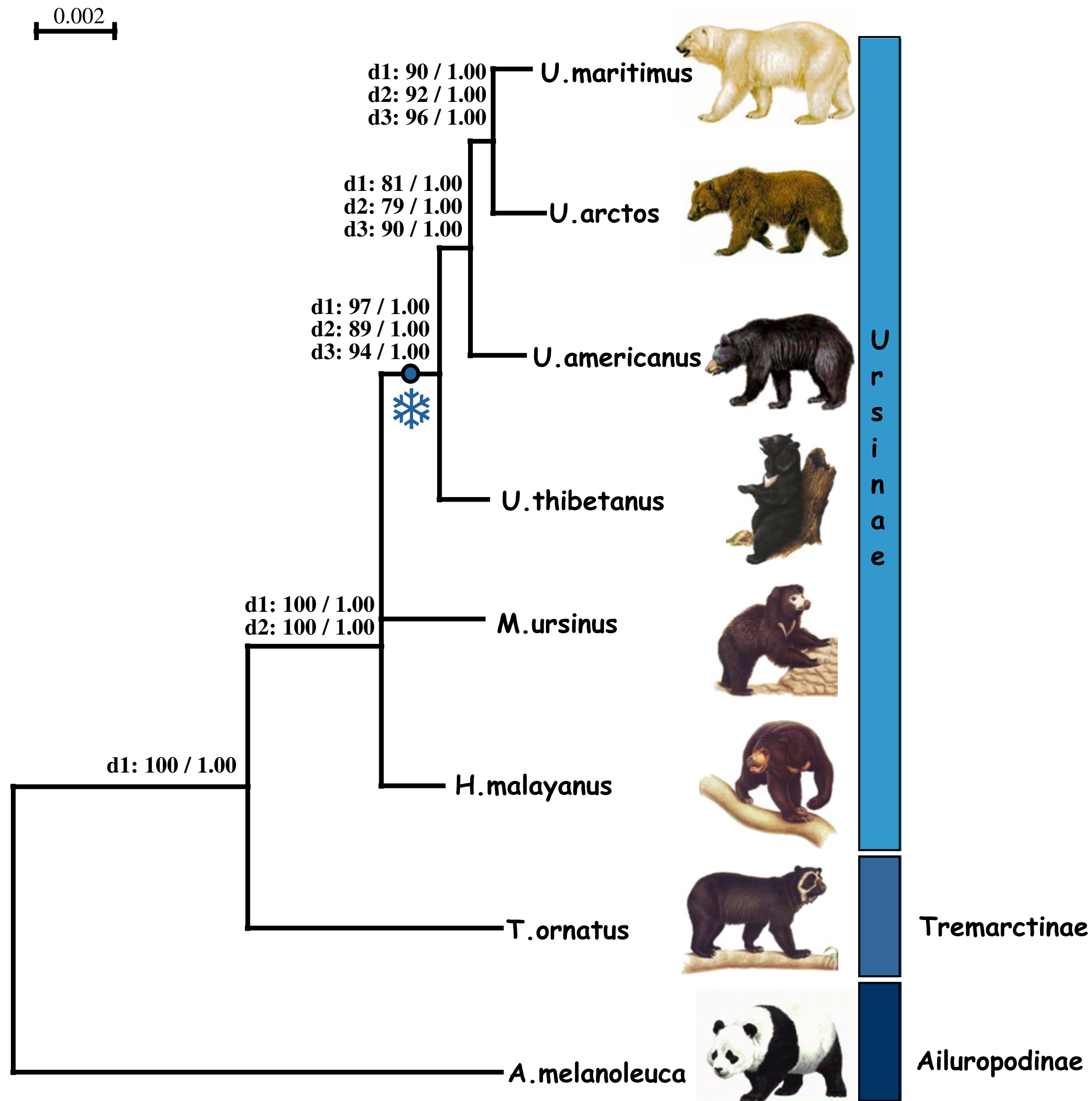

Fig. 1. Phylogenetic tree of the Ursidae family based on the analyses of 14 combined nuclear genes (dataset 2; see text) and reconstructed following Bayesian methods. BA and ML analyses of the 3 datasets gave an identical topology. Numbers above branches reflect supports (Bp/pp) obtained from the analysis of the three datasets (d1, d2, d3). The snow-flake stands for the putative acquisition of hibernating abilities by the common ancestor of the *Ursus* genus lineage.

sequence-tagged sites (STS) and IRBP exon 1 (68% Maximum Parsimony Bp, BA 0.98 pp).

3.2.2. Q2: Presumed sistership of the two black bears

The association of the two black bears, *Ursus americanus* and *Ursus thibetanus* as sister taxa was first proposed by Talbot and Shields using the entire cytochrome *b* sequences (1996a) (Table 6). However, it was shown that alternative hypotheses could not be excluded using larger mitochondrial DNA datasets (Waits et al., 1999). More recently, a sister branching of the black bears was once again suggested by the analysis of the TTR intron 1, but with a weak statistical support (40 and 52% Bp; Table 6). We show here, that topology tests performed on these datasets do not favour any of the hypotheses ($P > 0.05$; Table 7).

Our data are clearly not consistent with this sistership hypothesis. Indeed, all the three combined nuclear datasets gave identical results (Fig. 1): the American black bear clustered with the polar/brown bear clade and the Asiatic black bear is placed as the sister taxon to the polar/brown/American black bears. To test the reliability of this finding, we considered alternative hypotheses concerning Q2 (Table 7): (1) the Asiatic black bear groups with the polar/brown bear clade and the American black bear is placed as sister taxon to this former clade, and (2) the two black bears are clustered as sister taxa. KH and SH-tests show that trees where the two black bears are associated as sister taxa are significantly worse than trees where there are not clustered as sister taxa (e.g. $P < 0.05$ for KH and SH-tests based on dataset 3; Table 7). However, these tests failed to find significant differences between the alternative branching orders of the two black bears ($P > 0.05$ for KH and SH-tests based on datasets 2 and 3; Table 7). Nevertheless, the analyses of the three datasets

Table 6
Support values found in this study and in previous molecular analyses concerning the alternative hypotheses of the three unresolved ursid phylogenetic questions

	Previous published analysis					Multi nuclear genes analyses					
	Based on combined mitochondrial DNA		Based on exon (IRBP)	Based on intron (TTR)	Based on the 2 combined nuclear genes (IRBP/TTR)	Partition 1		Partition 2		Partition 3	
						Complete (7963 bp)	TTR excluded	Complete (8969 bp)	TTR excluded	Complete (9791 bp)	TTR excluded
Phylogenetic issue alternative topologies	Talbot and Shields, 1996	Waits et al., 1999	Yu et al., 2004	Yu et al., 2004	Yu et al., 2004	This study		This study		This study	
Q1 (O, ((A.m, T.o), U))	92	—				—	—				
(O, A.m, (T.o, U))	*71*	67/99	Not assessed	Not assessed	Not assessed	**100**/1.00	**100**/1.00	Not possible	Not possible	Not possible	Not possible
(O, T.o, (A.m, U))	54	—				—	—				
Q2 (O, (U.am, U.thib), U′)	*58/62/* 75	36/**38**		**40**/52	52/**58**/1.00	—	—	—		—	
(O, U.thib, (U.am, U′))	—	—	Unresolved	—	—	**81**/1.00	**82**/1.00	**79**/1.00	**75**/1.00	**90**/1.00	**85**/1.00
(O, U.am, (U.thib, U′))	—	—		—	—	—	—	—		—	
Q3 (O, (H.m, M.u)U″)			—	92/**98**	86/**86**/1.00	0.42	—	**50**		—	
(O, H.m, (M.u, U″))	Unresolved	Unresolved	63/**64**	—	—	**55**	**71**/0.78	0.96	**62**/0.99	**50**/0.98	**75**/0.99
(O, M.u, (H.m, U″))			—	—	—	—	—	—		—	

Support values correspond to the node highlighted in bold and represent the values observed for the best trees. Underlined numbers indicate posterior probabilities. NJ bootstrap values are in italics, ML bootstrap in bold and other numbers correspond to Maximum Parsimony bootstrap values.
Abbreviations: O, outgroup; A.m, *Ailuropoda melanoleuca*; T.o, *Tremarctos ornatus*; U, Ursinae; U.am, *Ursus americanus*; U.thib, *Ursus. thibetanus*, U′, (*Ursus maritimus, Ursus arctos*); H.m, *Helarctos malayanus*; M.u, *Melursus ursinus*; U″, *Ursus* genus representatives.
Q1: basal position of the Ursidae radiation?; Q2: sistership of the two black bears?; Q3: branching order among ursine group, i.e. positions of H.m and M.u?

Table 7
Likelihood ratio tests for the alternative hypotheses concerning phylogenetic relationships of the two black bears

Alternative topologies	IRBP Yu et al. (2004)			TTR Yu et al. (2004)			IRBP + TTR Yu et al. (2004)			Dataset 2 This study			Dataset 3 This study		
	Δ-lnL	KH-test	SH-test	Δ-lnL	KH-test	SH-test	Δ-lnL	KH-test	SH-test	Δ-lnL	KH-test	SH-test	Δ-lnL	KH-test	SH-test
(((U.mar, U.arct), U.am), U.thib), O1)	5.28	0.225	0.225	2.30	0.234	0.234	3.47	0.172	0.224	2.46	0.27	0.63	0.57	0.303	0.685
(((U.mar, U.arct), U.thib), U.am), O1)	5.28	0.225	0.225	2.30	0.234	0.234	3.47	0.172	0.224	15.63	0.095	0.117	14.05	0.088	0.092
(((U.mar, U.arct), (U.am, U.thib)), O1)	5.28	0.225	0.225	(1672)	—	—	1.29	0.275	0.577	20.28	0.035*	0.041*	16.76	0.039*	0.044*
(((U.mar, U.arct), U.am), U.thib), O2)	(1974)	—	—	2.30	0.234	0.234	2.39	0.227	0.333	(14803)	—	—	(15151)	—	—
(((U.mar, U.arct), U.thib), U.am), O2)	0	0.472	0.795	2.30	0.234	0.234	2.39	0.227	0.333	13.4	0.122	0.172	13.78	0.092	0.101
(((U.mar, U.arct), (U.am, U.thib)), O2)	0	0.347	0.790	0	0.506	0.804	(3705)	—	—	18.17	0.042*	0.067	16.57	0.042*	0.043*

Six topologies were tested representing all the possible topologies for the relationships between Ursinae. The association of brown and polar bears as sister taxa is considered as well-established. KH/SH-tests were performed on nuclear datasets provided by Yu et al., 2004 and on the two larger datasets out of the three generated in this study. Values in parenthesis correspond to the log-likelihood of the best tree; −ΔlnL, to the log-likelihood difference as compared to the best tree. An asterisk denotes that the test is significant at $P < 0.05$ (the topology is significantly worse than the best one). U.mar = *Ursus maritimus*; U.arct = *Ursus arctos*; U.am = *Ursus americanus*; U.thib = *Ursus thibetanus*; O1 = "*H. malayanus*), *M. ursinus*), *T. ornatus*), *A. melanoleuca*)"; O2 = "*M. ursinus*), *H. malayanus*), *T. ornatus*), *A. melanoleuca*)".

with two different methods of phylogenetic inferences sustain the hypothesis of the Asiatic black bear as the first black bear to diverge (Fig. 1) with strong supports (e.g. 90% Bp, 1.00 pp for the larger dataset 3; Table 6).

Consequently, with our larger nuclear datasets, the branching order inside the genus *Ursus* is clarified. Moreover, it should be noted that these analyses identified *Ursus* representatives as a monophyletic group and were in agreement on this point with phylogenetic reconstructions using nuclear TTR intron 1 (Yu et al., 2004). These findings are particularly interesting regarding the acquisition of hibernating capabilities by bears. As *Ursus* genus, including the extinct cave bear species (Hänni et al., 1994; Loreille et al., 2001), consists only in hibernating species (Stirling, 1993), their monophyly indicates that the hibernating capability likely appeared only once along the evolutionary history of the ursids (Fig. 1). Based on the current range distribution of the ursine species (Herreo et al., 1999), the acquisition of this peculiar feature by a common Asiatic ancestor could be the cause or the consequence of the colonisation by the *Ursus* representatives of the northern regions of the North hemisphere. Alternatively, hibernation could also represent a symplesiomorphic feature shared by the *Ursus* representatives, and lost by the other bear species.

3.2.3. Q3: The Ursinae radiation

The six ursine species underwent such a rapid radiation around 5 Myr ago (Kurtén, 1968; Wayne et al., 1991). Thus, determining their branching order is a difficult challenge. Molecular phylogenies based either on mitochondrial data (Zhang and Ryder, 1994; Vrana et al., 1994; Talbot and Shields, 1996a; Waits et al., 1999) or on nuclear data (Yu et al., 2004; Fulton and Strobeck, 2006) failed to clarify this problem. Up to now, the only well-established relationship within Ursinae was the sistership of brown and polar bears (Taberlet and Bouvet, 1992; Talbot and Shields, 1996a; Waits et al., 1999; Yu et al., 2004). More precisely, analyses of mitochondrial markers have concluded to closer relationships between polar bears and brown bear population of the ABC islands leading to the paraphyly of the brown bear species on mitochondrial data (Talbot and Shields, 1996b). Moreover, ancient DNA studies based on mitochondrial markers elucidated the sistership of the extinct cave bear with the brown-polar bear clade (Hänni et al., 1994; Loreille et al., 2001). We clarified above the relationships among the genus *Ursus* (Q2), so the last point to elucidate relates to the branching order of *M. ursinus* and *H. malayanus*.

Our data clearly placed these two species at the base of the ursine radiation and were excluded from the monophyletic group of *Ursus* representatives. However, ML reconstructions based on our 3 datasets do not discriminate between two of the three possible topologies given the weak Bp supports (Bp ⩽55%; Table 6). Nevertheless, BA performed on datasets 2 and 3 give stronger support to

M. ursinus clustered with the representatives of the genus *Ursus* (pp values of 0.96 and 0.98).

Combined genes could lead to an unresolved topology if the genes give contradictory results. In this case, Jeffroy et al. (2006) showed that it could be better to remove the problematic genes from the analyses. In their study combining IRBP and TTR, Yu et al. (2004) observed that both genes taken separately gave contradictory results: IRBP sustained the hypothesis of *M. ursinus* clustered with the representatives of the genus *Ursus*, whereas TTR supported the *M. ursinus* and *H. malayanus* sistership (Table 6). In order to assess whether TTR is responsible for the low supports observed, we performed the same phylogenetic analyses on our three datasets but excluding this marker. In this case, a single topology is well supported: *M. ursinus* clusters with the representatives of the genus *Ursus*. The exclusion of TTR increases all the support values for this hypothesis in both ML and BA analyses (e.g. Bp from 50–75% for dataset 3; Table 6). Nevertheless, further work is required to confirm the positions of the sloth and sun bears within the ursid tree.

3.2.4. Further investigations

In this study, we resolved the main ambiguities among the phylogeny of Ursidae (Q1 and Q2) and clarified the last one (Q3). Even the combination of 14 nuclear gene fragments was not informative enough to solve definitively the relationships between all the representatives of the Ursinae sub-family (Q3). This lack of resolution is certainly due to the rapid burst of ursine species during bear evolutionary history. A solution to circumvent this difficulty could be to identify Rare Genomic Changes (RGCs) in the Ursidae genomes (Springer et al., 2004). RGCs would be perfect arbiters and will certainly be needed, as there is little chance that increasing the nuclear sequence dataset will provide any further information. Searching RGCs could be directed by two theoretical and empirical points: (1) insertions, deletions, and retrotransposon integrations occur more likely in non-coding regions (Springer et al., 2004); (2) as other restricted regions of the genome with limited recombination, the non-recombining region of the Y chromosome is expected to accumulate more retrotransposon elements (Pecon-Slattery and O'Brien, 1998; Pecon-Slattery et al., 2004). So, scanning Y-linked chromosome intron sequences to find RGCs in bear genomes could prove to be useful to overcome the lack of phylogenetic signal among ursids. We are confident that these additional independent markers will consolidate the Ursidae tree.

Acknowledgments

We thank all the institutions that gave samples for this study: the MNHN (Paris), the Mammals Research Institute of Poland and Zoos of Moscow, Leipzig, Berlin, and in France: La Palmyre, Peaugres, Saint Martin La Plaine, Vincennes, Parc de la Tête d'Or de Lyon, Réserve Africaine de Sigean and more particularly I. Delisle (University of Alberta, Canada), G. Véron (MNHN, France), F. Catzeflis (CNRS, ISEM, France), and P. Taberlet (LECA, France). We are particularly grateful to M. Duffraisse for help in the lab, to P. Chevret, L. Orlando, A. Chaumont and V. Laudet for helpful discussions and critical reading, and to J.J. Jaeger for constant support. We wish to acknowledge J. Carew for English corrections. We finally thank the CNRS, Université Lyon 1, ENS de Lyon and MNRT for financial supports. M.P. holds a fellowship from the ARC Foundation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2007.10.019.

References

Bardeleben, C., Moore, R.L., Wayne, R.K., 2005. A molecular phylogeny of the Canidae based on six nuclear loci. Mol. Phylogenet. Evol. 37, 815–831.

Chang, B.H., Li, W.H., 1995. Estimating the intensity of male-driven evolution in rodents by using X-linked and Y-linked Ube 1 genes and pseudogenes. J. Mol. Evol. 40, 70–77.

Davis, D., 1964. The giant panda: a morphological study of evolutionary mechanisms. Fieldiana. Zool. Mem. 3, 1–339.

Flynn, J.J., Nedbal, M.A., 1998. Phylogeny of the Carnivora (Mammalia): congruence vs incompatibility among multiple data sets. Mol. Phylogenet. Evol. 9, 414–426.

Fulton, T.L., Strobeck, C., 2006. Molecular phylogeny of the Arctoidea (Carnivora): effect of missing data on supertree and supermatrix analyses of multiple gene data sets. Mol. Phylogenet. Evol. 41, 165–181.

Gadagkar, S.R., Rosenberg, M.S., Kumar, S., 2005. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. J. Exp. Zoolog. B Mol. Dev. Evol. 304, 64–74.

Galewski, T., Mauffrey, J.F., Leite, Y.L., Patton, J.L., Douzery, E.J., 2005. Ecomorphological diversification among South American spiny rats (Rodentia; Echimyidae): a phylogenetic and chronological approach. Mol. Phylogenet. Evol 34, 601–615.

Galtier, N., Gouy, M., Gautier, C., 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. 12, 543–548.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Guindon, S., Lethiec, F., Duroux, P., Gascuel, O., 2005. PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res. 33, W557–W559.

Haldane, J.B.S., 1947. The mutation rate of the gene for haemophilia, and its segregation ratios in males and females. Ann. Eugen. 13, 262–271.

Hänni, C., Laudet, V., Stehelin, D., Taberlet, P., 1994. Tracking the origins of the cave bear (*Ursus spelaeus*) by mitochondrial DNA sequencing. Proc. Natl. Acad. Sci. USA 91, 12336–12340.

Herreo, S. Compiled by Servheen, C., Herrero, H., Peyton, B. and the IUCN/SSC Bear and Polar Bear Specialist Groups (1999). Introduction, In: Bears: Status Survey and Conservation Action Plan, pp. 1–6. IUCN.

Jeffroy, O., Brinkmann, H., Delsuc, F., Philippe, H., 2006. Phylogenomics: the beginning of incongruence? Trends Genet. 22, 225–231.

Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. 29, 170–179.

Kurtén, B., 1968. The Pleistocene Mammals of Europe. Aldine, Chicago, III.

Lessells, K., 1997. More mutations in males. Nature 390, 236–237.

Loreille, O., Orlando, L., Patou-Mathis, M., et al., 2001. Ancient DNA analysis reveals divergence of the cave bear, *Ursus spelaeus*, and brown bear, *Ursus arctos*, lineages. Curr. Biol. 11, 200–203.

Mayr, E., 1986. Uncertainty in science: is the giant panda a bear or a raccoon? Nature 323, 769–771.

Moreira, M.A., 2002. SRY evolution in Cebidae (Platyrrhini: Primates). J. Mol. Evol. 55, 92–103.

Nishida, S., Pastene, A.L., Goto, M., Koike, H., 2003. SRY gene structure and phylogeny in the cetacean species. Mammal Study 28, 57–66.

Nylander, J.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53, 47–67.

O'Brien, S.J., Nash, W.G., Wildt, D.E., Bush, M.E., Benveniste, R.E., 1985. A molecular solution to the riddle of the giant panda's phylogeny. Nature 317, 140–144.

Pagès, M., Maudet, C., Bellemain, E., Taberlet, P., Hughes, S., Hänni, C., submitted for publication. A system for sex-determination from degraded DNA: a useful tool for paleogenetics and conservation genetics of Ursids.

Pecon-Slattery, J., O'Brien, S.J., 1998. Patterns of Y and X chromosome DNA sequence divergence during the Felidae radiation. Genetics 148, 1245–1255.

Pecon-Slattery, J., Pearks Wilkerson, A.J., Murphy, W.J., O'Brien, S.J., 2004. Phylogenetic assessment of introns and SINEs within the Y chromosome using the cat family felidae as a species tree. Mol. Biol. Evol. 21, 2299–2309.

Porter, C.A., Goodman, M., Stanhope, M.J., 1996. Evidence on mammalian phylogeny from sequences of exon 28 of the von Willebrand factor gene. Mol. Phylogenet. Evol. 5, 89–101.

Posada, D., Crandall, KA., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Rambaud A., Drumond A., 2003. Tracer, version v1.2.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Springer, M.S., Stanhope, M.J., Madsen, O., de Jong, W.W., 2004. Molecules consolidate the placental mammal tree. Trends Ecol. Evol. 19, 430–438.

Stirling, I., 1993. Bears-majestic creatures of the wild. Rodale Press Emmaus, Pennsylvania.

Swofford, D.L., 1998. PAUP*. Phylogenetic Analysis using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Taberlet, P., Bouvet, J., 1992. Génétique de l'Ours brun des Pyénées (*Ursus arctos*): premiers résultats. CR Acad. Sci. III. 314, 15–21.

Talbot, S.L., Shields, G.F., 1996a. A phylogeny of the bears (Ursidae) inferred from complete sequences of three mitochondrial genes. Mol. Phylogenet. Evol. 5, 567–575.

Talbot, S.L., Shields, G.F., 1996b. Phylogeography of brown bears (*Ursus arctos*) of Alaska and paraphyly within the Ursidae. Mol. Phylogenet. Evol. 5, 477–494.

Van Valen, L.M., 1986. Palaeontological and molecular views of panda phylogeny. Nature 319, 428.

Venta, P.J., Brouillette, J.A., Yuzbasiyan-Gurkan, V., Brewer, G.J., 1996. Gene-specific universal mammalian sequence-tagged sites: application to the canine genome. Biochem. Genet. 34, 321–341.

Vrana, P.B., Milinkovitch, M.C., Powell, J.R., Wheeler, W.C., 1994. Higher level relationships of the arctoid Carnivora based on sequence data and "total evidence". Mol. Phylogenet. Evol. 3, 47–58.

Waits, L.P., Sullivan, J., O'Brien, S.J., Ward, R.H., 1999. Rapid radiation events in the family Ursidae indicated by likelihood phylogenetic estimation from multiple fragments of mtDNA. Mol. Phylogenet. Evol. 13, 82–92.

Wayne, R.K., Van Valkenburgh, B., O'Brien, S.J., 1991. Molecular distance and divergence time in carnivores and primates. Mol. Biol. Evol. 8, 297–319.

Wozencraft, W.C., 1993. Order Carnivora. In: Wilson, D.E., Reeder, D.M. (Eds.), Mammal Species of the World: A Taxonomic and Geographic Reference. Smithsonian Institution Press, Washington, pp. 279–349.

Yu, L., Li, Q.W., Ryder, O.A., Zhang, Y.P., 2004. Phylogeny of the bears (Ursidae) based on nuclear and mitochondrial genes. Mol. Phylogenet. Evol. 32, 480–494.

Zhang, Y., Ryder, O., 1994. Phylogenetic relationships of bears (the Ursidae) inferred from mitochondrial DNA sequences. Mol. Phylogenet. Evol. 3, 351–359.