Conserv Genet (2009) 10:897–907
DOI 10.1007/s10592-008-9650-x

RESEARCH ARTICLE

A system for sex determination from degraded DNA: a useful tool for palaeogenetics and conservation genetics of ursids

Marie Pagès · Célia Maudet · Eva Bellemain · Pierre Taberlet · Sandrine Hughes · Catherine Hänni

Received: 6 June 2007 / Accepted: 24 July 2008 / Published online: 18 December 2008

Abstract In this paper, we characterise three sex-specific genes (ZFX/Y, SRY, AMLX/Y) for all eight extant bear species and propose a new, robust and accurate molecular procedure to identify the sex of bears from non-invasive samples and fossil remains. These materials contain tiny amounts of poorly preserved deoxyribonucleic acid (DNA), leaving Polymerase Chain Reaction (PCR) amplification very prone to contamination and difficult to analyse. By taking into account the ancient DNA requirements, the duplex procedures that we developed are efficient not only on DNA extracted from bear faeces but also on ancient DNA extracted from a brown bear fossil 7,500 years old. Defined specifically for ursids, the procedure for faecal samples (co-amplification of ZFX/Y and SRY markers) appears more accurate than other published procedures, as it prevents cross-amplification of potential ingested prey and contamination (19 non-ursid species tested). This system can be applied to threatened bear populations to improve the reliability of sex-ratio and population-size estimates based on non-invasive samples.

Keywords Molecular sex identification · DNA degradation · Non-invasive sampling · Ancient DNA · Ursidae

Introduction

The giant panda, as the peerless emblem of the world's endangered species, sums up the situation of its whole family: most ursid species are faced with extinction and persist only as remnant populations. To better preserve our last bear representatives, numerous studies have been performed to assess population genetic parameters (e.g. of brown bear: Taberlet et al. 1997; Köhn et al. 1995; polar bear: Carmichael et al. 2005; Paetkau et al. 1999; Asiatic black bear: Yamamoto et al. 2002; spectacled bear: Ruiz-Garcia 2003; Ruiz-Garcia et al. 2005; giant panda: Zhan et al. 2006). Among these parameters, sex ratio provides fundamental information for managing populations in order to avoid their extinction (Taberlet et al. 1993). For endangered bears, non-invasive sampling represents an easy way to obtain deoxyribonucleic acid (DNA) without any danger or disturbance (Höss et al. 1992; Constable et al. 1995). However, DNA extracted from these samples (mostly faeces or hair) is usually present in low quality and quantity (Taberlet et al. 1999). It is thus subject to contamination by exogenous DNA and often co-extracted with some Taq polymerase inhibitors (Köhn and Wayne 1997; Taberlet et al. 1999). Because of such difficulties, determining the sex of bear individuals from non-invasive sampling constitutes to be a real challenge.

Electronic supplementary material The online version of this article (doi:10.1007/s10592-008-9650-x) contains supplementary material, which is available to authorised users.

M. Pagès · C. Maudet · S. Hughes · C. Hänni (✉)
Paléogénétique et Evolution Moléculaire, Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, UMR 5242 CNRS, INRA, Université Claude Bernard Lyon 1, Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France
e-mail: Catherine.Hanni@ens-lyon.fr

M. Pagès
Institut des Sciences de l'Evolution, CNRS UMR 5554, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France

E. Bellemain · P. Taberlet
Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, B.P. 53, 38041 Grenoble Cedex 09, France

In the literature, three classic molecular markers have been used for sex determination of Ursidae (Fig. 1): (1) genes encoding for the zinc finger proteins ZFX and ZFY, (2) genes encoding for the amelogenin proteins AMLX and AMLY and (3) the sex-determining region of the Y chromosome SRY. The absence or presence of a single marker localised on the Y chromosome cannot be used for faecal extracts (Taberlet et al. 1993). Indeed, Y-specific amplifications could fail for a male individual for which the DNA extracted is not optimal and may thus mislead the sex assignation. A solution to prevent such errors consists in co-amplifying an internal nuclear Polymerase Chain Reaction (PCR)-positive control (duplex approach; Taberlet et al. 1997).

The sequences of these three markers are not available in databanks for all eight extant ursid species, and no simple and specific method has been yet defined for the molecular sexing of Ursidae from degraded substrates (Fig. 1). The previous procedures proposed are suboptimal for four main reasons: (1) heteroduplex formation between the homologous X and Y copies hampers accurate analysis of PCR results (Yamamoto et al. 2002; Xu et al. 2007), (2) faint SRY-like band for SRY female diagnosis leads to incorrect sex assignation (Taberlet et al. 1993; Woods et al. 1999), (3) lack of bear specificity of the method when applied to degraded bear substrates does not rule out the possibility of exogenous contaminations or cross-amplification of consumed prey (Murphy et al. 2003), and (4) long fragments targeted though maximum length of amplifiable DNA from faecal DNA extract is evaluated to be less than 200–250 bp (Taberlet et al. 1997; Murphy et al. 2000; Teletchea et al. 2005).

In this study, we sequenced three sex markers, SRY, ZF and AML, for all extant bear species. It allowed us to assess the genetic variability of these markers among the whole family but also to develop a new method especially devised for bear DNA extracted from faecal and fossil samples (the latters being referred to as delicate samples). We tested different molecular strategies on non-invasive samples as well on one of the most difficult substrates, a fossil bone of a brown bear, *Ursus arctos*, dating back 7,500 years. Our sex identification systems proved to be robust and reliable on degraded DNA templates from bears and opens up new possibilities for conservation biology and palaeontology.

Materials and methods

Samples

Fresh samples

A variety of DNA sources was selected, including blood, tissues, and plucked hairs from free-ranging bears or bears

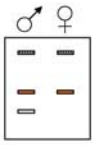
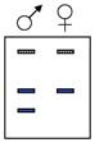
Molecular sexing strategy	Y marker	Internal PCR control	References	Critiques
Y-chromosome specificity → *SRY*  **4 primers**	SRY	CR *mt*DNA	Taberlet et al. 1993 and using the same primers, Köhn et al. 1995	• faint SRY-like band for SRY female diagnosis
	SRY	ZF	Wasser et al. 1997	• lack of bear specificity (primers designed for cattle by Aasen and Medrano (1990)) • length of targeted fragment unsuitable for degraded DNA
			Woods et al. 1999	• lack of bear specificity
Sequence polymorphism → *ZF*  **3 primers**	ZFY	ZFX	Durnin et al. 2007 *One combination of primers amplified a same size fragment on both X and Y copies, the second amplified a shorter fragment on the Y copy (one primer is specific to the Y because of its sequence)*	• lack of bear specificity • faint amplification of the ZFY-fragment for male diagnosis
Length polymorphism → *AML*  **2 primers**	AMLY	AMLX	Yamamoto et al. 2002	• heteroduplex formation between the homologous X- and Y-copies • non-bear specificity when using primers developed for cattle by Ennis and Gallagher (1994)
			Carmichael et al. 2005	• lack of bear-specificity
			Xu et al. 2007	• lack of bear-specificity • heteroduplex formation between the homologous X- and Y-copies

Fig. 1 Available molecular sex determination methods for ursids and pitfalls of these methods for analysing degraded DNA. *SRY*: sex-determining region of the Y chromosome; *ZF*: gene encoding for the zinc finger protein; *AML*: gene encoding for the amelogenin protein; CR mtDNA : control region of the mitochondrial DNA

from zoos (Table 1). To assess the intraspecific variability of the SRY gene, we sampled two male specimens for each extant species coming from different locations when possible. A total of ten females and 15 males was sampled, including all the eight bear species. Furthermore, 22 non-bear DNA extracts were used to test the bear specificity of our molecular sexing procedure (Table 2; see also Discussion). The different species were chosen either to test a possible DNA cross-amplification of ingested prey (four bovid, one cervid and one lagomorph species) or a contamination by exogenous DNA (12 carnivore species and three human individuals).

Faecal samples

Thirty DNA extracts from faeces were used to ensure performance of our molecular sexing procedure on degraded substrates of different bear species. Twenty-eight brown bear DNA samples came from previous studies (Bellemain et al. 2005, 2007). Two giant panda DNA samples were extracted from scats of individuals in the Berlin Zoo using DNeasy Tissue Kit (Qiagen) and following the manufacturer's instructions.

Fossil samples

Ancient DNA was extracted from a fossil brown bear (*U. arctos*) from Takouatz cave (Algeria) following the protocol described in Loreille et al. 2001. This bone, radiocarbon-dated to 7,345 ± 40 years BP, was chosen because of its peculiar mitochondrial haplotype (Calvignac et al. 2008). Until now, this haplotype was indeed only described for extinct brown bear specimens from North

Table 1 Fresh samples of extant ursid species analysed

Subfamily	Scientific name	Common name	Sex	Number of samples	Type	Sample source	Sequenced genes		
							SRY	ZF	AML
Ailuropodinae	*Ailuropoda melanoleuca*	Giant panda	Male	1	Tissue	MNHN (France)	X	X	X
Tremarctinae	*Tremarctos ornatus*	Spectacled bear	Male	2	Hairs	Parc de la Tête d'Or (Lyon, France)	XX	XX	X
			Female	1	Tissue	University of Montpellier II (France)		X	X
Ursinae	*Ursus thibetanus*	Asiatic black bear	Male	2	Hairs and tissue	Réserve Africaine de Sigean (France)	X–	X–	X–
	Melursus ursinus	Sloth bear	Male	1	Hairs	Moscow Zoo (Russia)	X	X	X
			Male	1	Hairs	Leipzig Zoo (Germany)	X	X	X
			Female	2	Hairs	Leipzig Zoo (Germany)		XX	XX
	Helarctos malayanus	Sun bear	Male	1	Hairs	Parc zoologique de Vincennes (MNHN, France)	X	X	X
			Female	1	Hairs	Zoo de Saint Martin La Plaine (France)		X	X
	Ursus americanus	American black bear	Male	1	Tissue	University of Montpellier II (France)	X		
			Male	1	DNA extract	University of Alberta (Canada)	X	X	X
			Female	1	Hairs and blood	Zoo de Peaugres (France)		X	
			Female	1	DNA extract	University of Alberta (Canada)			X
	Ursus arctos	Brown bear	Male	1	DNA extract	University of Alberta (Canada)	X	X	X
			Male	1	DNA extract	LECA (Grenoble, France)	X		
			Female	1	DNA extract	University of Alberta (Canada)			X
			Male	1	DNA extract	LECA (Grenoble, France)		X	
			Female	1	DNA extract	LECA (Grenoble, France)		X	
			Female	1	DNA extract	University of Alberta (Canada)			
	Ursus maritimus	Polar bear	Male	1	Hairs	Zoo de la Palmyre (France)	X	X	X
			Male	1	DNA extract	University of Alberta (Canada)	X		
			Female	1	Hairs	Zoo de la Palmyre (France)		X	X

Taxonomic denomination follows the classification of Wozencraft (1993);

"X": Polymerase Chain Reaction (PCR) positive amplifications; "–": PCR failures

Table 2 Non-ursid deoxyribonucleic acid (DNA) extracts used to test specificity of the ursid duplex molecular sexing procedure

Family	Scientific name	Common name	Sex	ZF amplification
Mustelidae	*Lutra longicaudis*	Neotropical river otter	X	+
	Eira barbara	Tayra	X	+
	Galictis vittata	Greater grison	X	+
	Mustela herminea	Short-tailed weasel	X	+
	Mustela vison	American mink	X	+
Procyonidae	*Procyon cancrivorus*	Crab-eating racoon	X	−
Canidae	*Canis familiaris*	Domestic dog	M	−
			F	−
Felidae	*Felis catus*	Domestic cat	F	−
	Herpailurus yagouaroundi	Jaguarundi	X	−
	Leopardus pardalis	Ocelot	X	−
	Panthera onca	Jaguar	X	−
	Puma concolor	Puma	X	−
Hominidae	*Homo sapiens*	Human	M	−
			M	−
			F	−
Bovidae	*Capra pyrenaica*	Spanish ibex	M	−
	Bos taurus	Cow	X	−
	Bison bison	American bison	X	−
	Bison bonasus	European bison	X	−
Cervidae	*Cervus elaphus*	Red deer	M	−
Lagomorph	*Lepus europaeus*	European hare	X	−

X: unknown sex; M: male sample; F: female sample; "+": ZF-positive PCR amplification; "−"; no PCR amplification

Africa and thus could be easily distinguished from other possible bear contamination.

Precautions to avoid contaminations

Preamplification procedures and postamplification analyses were systematically performed in independent rooms. In the case of ancient DNA work, analyses were performed following guidelines for ancient DNA in a specific laboratory devoted to ancient DNA analysis and using specific equipment and personal protection (Cooper and Poinar 2000). Ancient DNA facilities are physically separated from the two rooms where fresh and faecal substrates, respectively, were handled.

DNA extraction of fresh samples, amplification and sequencing of ursid sex markers

DNA was extracted from blood samples with QiAmp DNA Blood Mini Kit (Qiagen) and from tissue and hair samples with DNeasy Tissue Kit (Qiagen) in accordance with the manufacturer's instructions.

The three molecular markers (SRY, ZF and AML genes) were amplified with the primer sets previously described for mammal or carnivore analyses (Table 3). Single pairs of primers were used to amplify ZF [447 base pairs (bp)] and AML (244 bp for AMLX and/or 190 bp for AMLY) exons. However, for SRY, two overlapping fragments (846 and 1,094 bp) were amplified to cover the entire gene (around 1,300 bp). All amplifications were carried out in 25 µl containing about 30 ng of DNA extract, 0.2 mg/ml bovine serum albumin (BSA) (Roche, 1 mg/ml), 250 µM of each deoxynucleotide triphosphate (dNTP), 0.2 µM of each primer, 1 U of Amplitaq Gold DNA polymerase (Applied Biosystems), 2.5 µl of 10× buffer, and 1.5 mM of magnesium chloride ($MgCl_2$). Fresh DNA was amplified with an average of 40 cycles of denaturation (94°C, 30 s), annealing (for temperature, see Table 3, 30 s) and elongation [72°C, 1 min/kilobase (kb) according to target length]. When the sequencing of PCR products proved to be difficult or when PCR products consisted of an admixture of homologous fragments shared by the X and Y chromosomes in the case of a male sample, they were cloned with Topo TA Cloning for Sequencing Kit (Invitrogen), and several clones were sequenced. All sequences were performed by a service provider (Genome Express, France) in both strand directions.

Sequence analysis of ursid SRY, ZF and AML genes

To assess the variability of the three markers among ursids, we used PAUP software (Swofford 1998) to compute the uncorrected pairwise divergence between the orthologous sequences of the eight bear species. As the SRY gene is composed of many domains with different properties and

Table 3 Primers used in this study

Designation	Nucleotide sequence 5′ → 3′	Annealing temperature		Fragment length	Original publication
Primers used for the molecular sexing of carnivores or mammals and to obtain the ursid sequences					
SRY gene					
SRYW	Unpublished	55°C		846 bp	Pers. com. M. Kinnear
SRYM	GCCTTCCGACGRGGTCGATA				This study
SRYC-COR	GGAAGTTTTGCTTGAGAATGC	50°C		1,094 bp	This study
SRYK	Unpublished				Pers. com. M. Kinnear
ZFY/X fragment					
P1-5EZ	ATAATCACATGGAGAGCCACAAGCT	53°C		447 bp	Aasen and Medrano 1990
P2-3EZ	GCACTTCTTTGGTATCTGAGAAAGT				
AMLX/Y fragment					
SE47F	CAGCCAAACCTCCCTCTGC	60°C	X-copy	244 bp	Ennis and Gallagher 1994
SE48R	CCCGCTTGGTCTTGTCTGTTGC		Y-copy	190 bp	
Primers used for sexing bears from degraded DNA (excremental DNA and ancient DNA)					
SRY system					
MP-SRY-Ursus	TGGTCTCGTGATCAAAGGCGC	55°C		115 bp	This study
MP-SRY-1R-Ursus	GCCATTTTTCGGCTTCCGTAAG				This study
ZFY/X system					
MP-ZF-F	GACAGCTGAACAAGGGTTG	55°C		144 bp	This study
MP-ZF-R	GCTTCTCGCCGGTATGGATG				
Mitochondrial system					
CB2670MP	CTACTTAAACTATTCCCTGGTACATAC	55°C		97 pb	Modified from Hofreiter et al. 2004
CBH45	GGACATACTATGATGGTACAGTACAT				Hofreiter et al. 2004

When different sizes of fragments were obtained for the X and Y copy (for the AML gene only), each specific size is indicated

functionalities, we partitioned the analysis as proposed in Tucker and Lundrigan 1993. The divergence was thus computed for the five following regions of the gene: the 5′ untranslated region, the N-terminal coding region, the HMG box, the C-terminal coding region and the 3′ untranslated region.

Development of the ursid sexing procedure based on bear faecal DNA

Requirements for faeces analysis

At the first glance, a system based on the AML gene appears advantageous: (1) a visual discrimination of the PCR amplification product on agarose gel between a male (two bands) and a female (one band), as AML is shared by the X and the Y chromosomes but the AMLY has a deletion compared with the X copy; (2) an internal PCR control (AMLX) to monitor PCR failure. However, because of the strong conservation of the sequence among mammals, the design of bear-specific primers was not possible (see the alignment of the eight ursid sequences obtained for the AML with other mammalian ones in Supplementary Data S1). Consequently, we were prompted to develop a new system offering the same advantages as the AML system: a co-amplification of a bear-specific Y marker (SRY gene) with a bear-specific internal PCR control (ZF gene). The two primer sets used for the new duplex procedure were visually designed to target short fragments (less than 200 bp; Table 3). To co-amplify a robust internal positive control, the ZF fragment was chosen to be longer than the SRY fragment.

Experimental protocol for the new duplex procedure using SRY and ZF markers

Appropriate conditions of the duplex PCR were determined by testing the primer sets on non-degraded DNA extracted from blood/tissue bear samples. To validate the procedure on degraded bear substrates, this sexing system was applied to the faecal samples, and this experiment was repeated five times independently. The sex of giant panda individuals was known before molecular assignment, whereas the real sex of the defecator (determined independently by Bellemain et al. 2005, 2007 using another molecular sexing method) was known only after the molecular sexing assignation was blindly performed on brown bear DNA. PCR amplifications were carried out in 25 µl reaction

volume containing 1.25 U of Amplitaq Gold DNA polymerase (Applied Biosystems), 0.5 mg/ml BSA (Roche, 1 mg/ml), 1.5 mM $MgCl_2$, 250 μM each dNTP, 0.4 μM of ZF primers, 0.12 μM of SRY primers and 2 μl of faecal DNA. PCR conditions were the same as the procedure based on fresh samples, except for the time of elongation (45 s). The number of cycles was also increased to 48. To test for potential contamination and the reliability of this procedure, two PCR amplification products from both a male and a female brown bear sample were cloned using Topo TA Cloning for Sequencing Kit (Invitrogen) as a control. Species identification of clone sequences was then performed using a BLAST program.

Development of the ursid sexing procedure based on ancient bear DNA

Considering the requirements for ancient DNA studies (see Discussion), we defined two independent duplex procedures based on the co-amplification of one sexing marker (ZF, 144 bp or SRY, 115 bp) and a highly variable mitochondrial (mt) fragment (97 bp) (Table 3). At least two independent PCR amplifications were performed in 25 μl reaction volume containing 2.5 U of Amplitaq Gold DNA polymerase (Applied Biosystems), 1 mg/ml BSA (Roche, 1 mg/ml), 2 mM $MgCl_2$, 250 μM each dNTP, 0.125 μM of mitochondrial primers and 0.5 μM of each nuclear primer. For each PCR attempt, a range of dilutions was performed to find the best compromise between inhibitor concentration and target DNA molecule concentration (Hänni et al. 1995). PCR conditions were the same as those mentioned above, except that the number of cycles was increased to 55. All products of amplification obtained were cloned and sequenced in accordance with ancient DNA studies.

Results

Characterisation of the ursid SRY, ZF and AML genes

All sequences obtained for the SRY, ZF and AML genes were deposited in the European Molecular Biology Laboratory (EMBL) under the accession numbers from AM748297 to AM748312 and from AM941047 to AM941070.

SRY gene

Both overlapping fragments were successfully amplified for 13 of 14 male bears tested, permitting us to determine the entire sequences of the SRY gene (1,303 bp) for the eight extant ursid species (Supplementary Data S2). No polymorphism was observed between individuals within a species (except one substitution between the two males of *U. americanus*; Supplementary Data S2). The highest sequence difference for the entire SRY sequence was observed between *Tremarctos ornatus* and *Ailuropoda melanoleuca* (3%) and the lowest difference between *U. americanus* and *U. maritimus* (0.08%) (see Supplementary Data S3). The most conserved domain for the ursid SRY corresponded to the HMG box (from 0% to 3% sequence divergence among Ursidae) and the most variable to the C-terminal region (from 1% to 5.3% sequence divergence among Ursidae) (see Fig. 2 and Supplementary Data S3). We observed that the C-terminal region was heterogeneous in length. A difference in the stop codon position was detected for *U. arctos* compared with other ursids, leading to a longer C-terminal region of the encoded protein. Sequences from *U. arctos* individuals comprise a mutation at position 664 of the coding region, which converts the stop codon of the other ursid species in a sense codon and thus extends the protein to 11 amino acids (S2).

ZF gene

A 447-bp homologous fragment was successfully amplified for all 11 males and all seven females tested, furnishing the ZF sequences for all extant bear species (Supplementary Data S4). Low levels of variability were observed among ursid species (from 0% to 1.2% for ZFX and from 0% to 2% for ZFY; Supplementary Data S3). However, 11 dimorphic sites could be distinguished by scanning the whole set of the ursid ZFX and ZFY sequences (Table 4). Four sex diagnostic sites were specific to only one bear species, the seven others were shared by three or more bear species. Only one X/Y substitution was shared by all eight bear species (position 64).

AML gene

We obtained a single band of 244 bp for all seven tested females and two bands of 244 bp (AMLX) and 190 bp (AMLY) for nine males (out of the ten tested). An additional nonspecific band (heteroduplex AMLX/AMLY, 280 bp approximately) was detected for males (Yamamoto et al. 2002). Very low levels of variability were observed among ursid species (from 0% to 0.95% for AMLX and from 0% to 2.7% for AMLY; Supplementary Data S3). Alignment of the sequences of all eight extant bear species is provided in Supplementary Data S5. Most ursid species have a 54-nucleotide deletion on the Y-chromosome-specific gene compared with the X copy, which includes, instead, a motif of 9 bp repeated three times (Table 5). The AMLX fragment of the sun bear and the giant panda comprises an additional repetition of this motif, leading to a deletion of 63 bp on the AMLY fragment. Two polymorphic positions between X and Y (positions 49 and 130 of the alignment S5) were observed for all bear species. Other detected dimorphic sites were

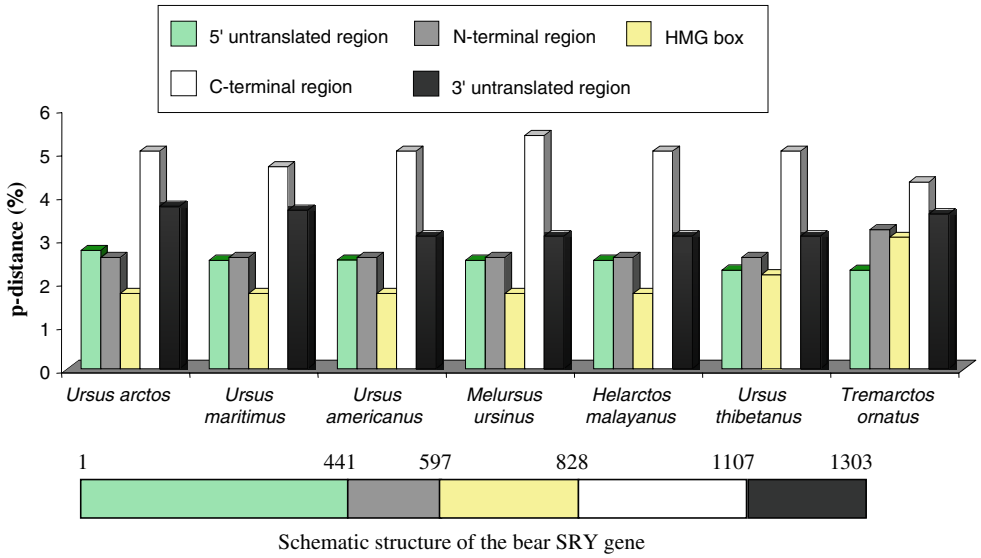


Fig. 2 Percentage of observed divergence between sequences of the most basal ursid representative—the giant panda, *Ailuropoda melanoleuca*—and the seven other bear species. According to its properties and functionalities, the SRY gene was partitioned in the five following regions: the 5′ untranslated region, the N-terminal coding region, the HMG box, the C-terminal coding region, and the 3′ untranslated region. Numbers on the schematic representation of the SRY gene indicate the positions delimiting each of these regions according to the giant panda sequence (AM748312). The most conservative domain for the ursid SRY gene corresponds to the HMG box, and the most variable region corresponds to the C-terminal region

Table 4 Mutational point diagnosis of sex between the X and Y copies of the ZF fragment among the Ursidae family

Ursids species	Copies	Positions of variable sites X/Y										
		16	58	64	68	97	160	178	187	241	252	277
U. arctos	X/Y	G/A		G/A						G/A		
U. maritimus	X/Y	G/A		G/A						G/A		
U. americanus	X/Y	G/A	C/T	G/A						G/A		
U. thibetanus	X/Y	G/A		G/A	T/C		G/A		C/T	G/A		T/C
H. malayanus	X/Y	G/A		G/A	T/C		G/A	A/G	C/T	G/A		T/C
M. ursinus	X/Y	G/A		G/A	T/C		G/A		C/T	G/A		T/C
T. ornatus	X/Y	G/A		G/A	T/C		G/A		C/T			
A. melanoleuca	X/Y			G/A		G/A					A/G	

Positions are numbered according to the sequence of brown bear (*U. arctos*) deposited in European Molecular Biology Laboratory (EMBL) under the accession number AM941047

species specific or shared by only a few of the eight bear species. Table 5 displays all the sex-diagnosis sites between the X and Y copies of the AML fragment among the Ursidae family.

Efficiency of the new molecular sex identification systems

Application to faeces: duplex PCR amplification of ZF and SRY fragments

The five independent PCR attempts yielded consistent amplification for all 30 bear excremental DNA tested. Only two male samples were misidentified as females (only one band obtained instead of the two expected) in one out of five attempts. Molecular sex assignment from excremental DNA perfectly matched with the real sex of the bear defecator when usual procedures of replication were performed (i.e. to detect allelic dropout; Taberlet et al. 1996).

Despite several amplification attempts in optimal conditions on non-bear DNA extracts, the internal positive control (ZF) was not amplified on most species tested, including humans (Table 2). Only ZF of representatives of the Mustelidae family could be amplified with our system. Moreover, of the 69 clones analysed for the two independent amplicons cloned for two brown bear faecal samples collected in the field, no sequence of human origin or of other potential ingested prey was detected. All sequences were of brown bear origin.

Table 5 Mutational point diagnosis of sex between the X and Y copies of the AML fragment among the Ursidae family

Ursids species	Copies	Positions of variable sites X/Y									Bp deletion
		22	28	49	130	137	142	151	170	199	
U. arctos	X/Y			G/C	G/A					G/A	54
U. maritimus	X/Y			G/C	G/A					G/A	54
U. americanus	X/Y			G/C	G/A	C/T				G/A	54
U. thibetanus	X/Y			G/C	G/A					G/A	54
H. malayanus	X/Y			G/C	G/A						63
M. ursinus	X/Y			G/C	G/A					G/A	54
T. ornatus	X/Y			G/C	G/A			G/A		G/A	54
A. melanoleuca	X/Y	C/A	T/C	G/C	G/A		T/A		C/T	G/A	63

Positions are numbered according to the longest sequence of the sun bear, *H.malayanus*, deposited in European Molecular Biology Laboratory (EMBL) under the accession number AM941061

Application to ancient DNA: duplex PCR amplification of ZF (or SRY) and CR fragments

MtDNA (CR) and nuclear (ZF or SRY) sequences were obtained in the same tube by a duplex PCR procedure on a brown bear fossil from North Africa (Takouatz cave, Calvignac et al. 2008). We are confident that the sequences we obtained were genuine ancient brown bear sequences for the following reasons: (1) All clones analysed (80) were identified as brown bear, and all mtDNA copies belonged to the peculiar haplotype of the ancient brown bear described elsewhere (Calvignac et al. 2008). (2) Errors induced by DNA damage were perfectly consistent with the pattern generally observed for ancient DNA sequences (strong bias toward type 2 transitions caused by deamination of cytosine; Hofreiter et al. 2001; Gilbert et al. 2003). (3) For each amplification, all three PCR blanks were negative (Loreille et al. 2001). (4) No PCR product was obtained when using ursid primers on the non-bear species extract co-extracted or co-amplified with the Takouatz sample, suggesting that no cross-contamination and no carrier effect occurred during the extraction and the amplification steps (Loreille et al. 2001). (5) Independent PCRs were performed to ensure the reliability of this result and furnished the same conclusions. These points satisfy criteria of authentication for the ancient DNA work (Gilbert et al. 2005). As amplicons of the ZF fragment contained equally X and Y copies, and as the SRY fragment was amplified, both independent systems of molecular sex determination were congruent: this ancient brown bear, that lived 7,500 years ago in North Africa, was a male.

Discussion

In this study, we determined the sequences of three sex markers (AML, ZF and SRY) for all eight bear species. In this large data set, little genetic variability was detected among the ursid family and even less among the Ursinae subfamily (less than 1% sequence divergence for the three genes). These findings are particularly noteworthy regarding the SRY gene. A growing body of evidence suggests that the flanking regions of the HMG box domain evolve rapidly and are poorly conserved between mammals (Whitfield et al. 1993; Tucker and Lundrigan 1993). In this study, in contrast to this rapid sequence evolution of SRY gene among mammals, a low interspecific divergence was found between ursid SRY sequences. Among Ursinae, this value is below 0.5%. Such a very low level of variability between ursine sequences is probably due to the recent and rapid burst of ursine species during bear evolutionary history (Waits et al. 1999; Yu et al. 2004; Pagès et al. 2008). Little sequence divergence for SRY has been described among closely related species [between 0% and 3.7% among sequences of rock wallabie species (Petrogale, Macropodidae) (O'Neill et al. 1997)], and no polymorphism has been observed between closely related species of Cebidae (Moreira 2002). This also suggests that SRY is not suitable for phylogeographical purposes to track male migrations and to better define conservation units. In the same way, the SRY sequences, as well as those of AML or ZF, are not ideal for assessing the unresolved phylogeny of the Ursidae family (Pagès et al. 2008). Nevertheless, this set of sequences enabled us to develop a straightforward, effective and specific sexing method from degraded remains of all extant and extinct bear species.

A new and efficient tool for non-invasive genetic studies in bears

Classic molecular sexing procedures use trans-specific primers (e.g. Aasen and Medrano 1990; Griffiths and Tiwari 1993; Shaw et al. 2003) that can amplify many species. For example, molecular sex identification methods

based on the AML system usually use the AML primer set SE47/SE48 (Table 3), originally described for cattle (Ennis and Gallagher 1994) and now used for sexing various mammals (Yamamoto et al. 2002; Pfeiffer and Brenig 2005; Weikard et al. 2006). Transspecific primers are efficient when used on good-quality DNA samples, such as tissue or blood, but are not reliable when a low amount of highly degraded DNA is the primary source of DNA template, such as DNA extracted from non-invasive samples (Taberlet et al. 1999). As potential contaminants of higher quality may be preferentially amplified when working with highly degraded templates, sex assignation mistakes would be inevitable when using nonspecific primers (Murphy et al. 2003).

Molecular sexing is even more problematic when dealing with faecal DNA that could be a mix of DNA from different sources. Bears are consummate opportunists and can eat hoofed mammals, such as red deer [up to 48.5% of volume diet for the Yellowstone grizzly bears (Mattson et al. 1991)], moose calves (e.g. Dahle et al. 1998) or small mammals. Murphy and colleagues (2003) showed that of the ten brown bear females exclusively fed with male deer meat, 40% of sex assignment performed on excremental DNA led to a false identification of sex. This warns against the use of universal primers with non-invasive samples and emphasises the importance of developing group- or species-specific primers to analyse such delicate DNA. The exhaustive data set of sex markers generated here allowed us to design bear-specific primers. Tested on different mammalian species, it should be noticed that ZF amplification could occur for species that are phylogenetically close to bears, as observed here with the Mustelids (Table 2), probably because of low sequence variability (Supplementary Data S4). However, as it seems highly unlikely that bear and mustelid faeces would be misidentified in the field, a species determination based on molecular methods should be performed first for any questionable specimen.

The bear specificity of this system is suitable for non-invasive studies, as proven by the following observations: (1) No amplification was obtained from hoofed mammals and lagomorph samples, suggesting that these primers avoid cross-amplification of most potential bear food items (Table 2). (2) The duplex procedure comprises an internal control (ZF) that prevents contamination by most other species tested, including humans (Table 2). (3) as a control, all 69 clone sequences obtained from faecal DNA samples of two free-ranging bears were of brown bear origin only, confirming again that when usual precautions are respected, these primers prevent exogenous contaminations.

By taking into account all requirements when working on degraded DNA, this bear-specific system represents an accurate tool for sex assignment of ursids, from the brown bear to the most basal ursid representative, the giant panda. However, as this system cannot prevent a ZF or SRY dropout due to sampling stochasticity when pipetting template DNA in a very dilute DNA extract, we recommend using a multiple-tube procedure to obtain reliable sex assignation (Taberlet et al. 1996).

We are optimistic that this method will improve some conservation studies. The procedure could be particularly suitable to analyse the sex ratio of free-ranging bears in tropical regions, for which the droppings are badly preserved and data are sorely missing [e.g. the sun bear, data deficient for the International Union for Conservation of Nature (IUCN)].

Molecular sexing of bear fossils

Problems encountered in ancient DNA studies are similar to those faced when dealing with DNA extracted from non-invasive samples. However, they are clearly exacerbated, particularly when nuclear sequences are targeted. Ancient DNA is indeed very prone to exogenous contamination (Pääbo et al. 2004). Therefore, a required criterion is to clone and sequence all positive PCR amplifications to authenticate the sequence obtained. The use of a highly variable sex marker (or a combination of a sex marker with a variable autosomal sequence) is thus interesting because it permits one to discriminate between eventual contamination, including those from extant or extinct bears (e.g. cave bear, *Ursus spelaeus*; or Atlas brown bear). However, no such sequences are yet available in databanks for ursids, even from recent palaeogenomics data (Noonan et al. 2005). Instead, a short, highly variable fragment of CR of the mtDNA offers the following advantages: (1) numerous diagnosis sites allow identification of the bear species, and (2) the geographical origin of individuals can be assigned for ursid species with well-known phylogeography (e.g. brown bear: Taberlet and Bouvet 1994; Waits et al. 1998; Calvignac et al. 2008; black bear: Wooding and Ward 1997).

Our molecular sex determination system, which proved reliable on a brown bear fossil aged 7,500 years, represents a new and original tool for palaeontological research. Although only a single specimen was tested in this study, preliminary results demonstrate that the procedure was efficient on other ancient samples (Pagès et al. in prep.). Some confused nomenclature for fossil ursids based on body size, and potentially non-detected sexual dimorphism, could be clarified by this method, as evidenced for the extinct moas (Bunce et al. 2003).

Acknowledgments We thank I. Delisle (University of Alberta, Canada), F. Catzeflis (CNRS, ISEM, France), N. Valière (BBE, UMR 5558), G. Queney (ANTAGENE, Lyon, France) and all the

institutions that gave samples for this study: the Natural History Museum of Aix-en-Provence (France), the MNHN (Paris, France), 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 (Lyon), Réserve Africaine de Sigean. Faecal DNA extracts were supplied by the LECA (Grenoble, France) with the help of C. Miquel and the international Scandinavian Brown Bear Research Project. Furthermore, we express gratitude to M. Kinnear who offered us judicious advice and gave us sequences of his unpublished SRY primers, to J.P. Noonan who transmitted to us his ursid nuclear sequences data set and to V. Laudet for his interesting comments. We thank anonymous referees for comments and suggestions. Finally, we are particularly grateful to other members of the palaeogenetics group: A. Seguin for initial work, M. Duffraisse for help in the lab, F. Télétchéa and L. Orlando for helpful discussions and S. Calvignac for critical reading. We acknowledge J. Carew for English language corrections. Finally, we thank the CNRS (ECLIPSE program), Université Lyon 1, ENS Lyon and MNRT for financial support.

References

Aasen E, Medrano JF (1990) Amplification of the ZFY and ZFX genes for sex identification in humans, cattle, sheep and goats. Biotechnology (N Y) 8(12):1279–1281. doi:10.1038/nbt1290-1279

Bellemain E, Swenson JE, Tallmon D, Brunberg S, Taberlet P (2005) Estimating population size of elusive animals with DNA from hunter-collected feces: comparing four methods for brown bears. Conserv Biol 19(1):150–161. doi:10.1111/j.1523-1739.2005.00549.x

Bellemain E, Ali Nawaz M, Valentini A, Swenson JE, Taberlet P (2007) Genetic tracking of the brown bear in northern Pakistan and implications for conservation. Biol Conserv 134(4):537–547. doi:10.1016/j.biocon.2006.09.004

Bunce M, Worthy TH, Ford T, Hoppitt W, Willerslev E, Drummond A et al (2003) Extreme reversed sexual size dimorphism in the extinct New Zealand moa *Dinornis*. Nature 425(6954):172–175. doi:10.1038/nature01871

Calvignac S, Hughes S, Tougard C, Michaux J, Thevenot M, Philippe M et al (2008) Ancient DNA evidence for the loss of a highly divergent brown bear clade during historical times. Mol Ecol (in press)

Carmichael L, Krizan P, Polischuk Blum S, Strobeck C (2005) Genotyping of pseudohermaphrodite polar bears in Nunavut and advances in DNA sexing techniques. J Mammal 86(1):160–169. doi:10.1644/1545-1542(2005)086<0160:GOPPBI>2.0.CO;2

Constable JJ, Packer C, Collins DA, Pusey AE (1995) Nuclear DNA from primate dung. Nature 373(6513):393. doi:10.1038/373393a0

Cooper A, Poinar HN (2000) Ancient DNA: do it right or not at all. Science 289(5482):1139. doi:10.1126/science.289.5482.1139b

Dahle B, Sørensen OJ, Wedul H, Swenson E, Sandegren F (1998) The diet of brown bears in central Scandinavia: effect of access to free-ranging domestic sheep. Wildl Biol 3:147–158

Durnin M, Palsbøll P, Ryder O et al (2007) A reliable genetic technique for sex determination of giant panda (*Ailuropoda melanoleuca*) from non-invasively collected hair samples. Conserv Genet 8:715–720. doi:10.1007/s10592-006-9196-8

Ennis S, Gallagher TF (1994) A PCR-based sex-determination assay in cattle based on the bovine amelogenin locus. Anim Genet 25(6):425–427

Gilbert MT, Hansen A, Willerslev E, Rudbeck L, Barnes I, Lynnerup N et al (2003) Characterization of genetic miscoding lesions caused by postmortem damage. Am J Hum Genet 72:48–61. doi:10.1086/345379

Gilbert MT, Bandelt HJ, Hofreiter M, Barnes I (2005) Assessing ancient DNA studies. Trends Ecol Evol 20(10):541–544. doi:10.1016/j.tree.2005.07.005

Griffiths R, Tiwari B (1993) Primers for the differential amplification of the sex-determining region Y gene in a range of mammal species. Mol Ecol 2(6):405–406. doi:10.1111/j.1365-294X.1993.tb00034.x

Hänni C, Brousseau T, Laudet V, Stehelin D (1995) Isopropanol precipitation removes PCR inhibitors from ancient bone extracts. Nucleic Acids Res 23(5):881–882. doi:10.1093/nar/23.5.881

Hofreiter M, Jaenicke V, Serre D, von Haeseler A, Pääbo S (2001) DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res 29(23):4793–4799. doi:10.1093/nar/29.23.4793

Hofreiter M, Rabeder G, Jaenicke-Despres V, Withalm G, Nagel D, Paunovic M, Jambresić G, Pääbo S (2004) Evidence for reproductive isolation between cave bear populations. Curr Biol 14(1):40–43. doi:10.1016/j.cub.2003.12.035

Höss M, Köhn M, Pääbo S, Knauer F, Schröder W (1992) Excrement analysis by PCR. Nature 359(6392):199. doi:10.1038/359199a0

Köhn M, Wayne R (1997) Factes from feces revisited. Trends Ecol Evol 12:223–227. doi:10.1016/S0169-5347(97)01050-1

Köhn M, Knauer F, Stoffella A, Schroder W, Pääbo S (1995) Conservation genetics of the European brown bear–a study using excremental PCR of nuclear and mitochondrial sequences. Mol Ecol 4(1):95–103. doi:10.1111/j.1365-294X.1995.tb00196.x

Loreille O, Orlando L, Patou-Mathis M, Philippe M, Taberlet P, Hänni C (2001) Ancient DNA analysis reveals divergence of the cave bear, *Ursus spelaeus*, and brown bear, *Ursus arctos*, lineages. Curr Biol 11(3):200–203. doi:10.1016/S0960-9822(01)00046-X

Mattson D, Blanchard B, Knight R (1991) Food habits of Yellowstone grizzly bears, 1977–87. Can J Zool 69:1619–1629

Moreira MA (2002) SRY evolution in Cebidae (Platyrrhini: Primates). J Mol Evol 55(1):92–103. doi:10.1007/s00239-001-2308-7

Murphy M, Waits L, Kendall KC (2000) Quantitative evaluation of fecal drying methods for brown bear DNA analysis. Wildl Soc Bull 28:951–957

Murphy MA, Waits LP, Kendall KC (2003) The influence of diet on faecal DNA amplification and sex identification in brown bears (*Ursus arctos*). Mol Ecol 12(8):2261–2265. doi:10.1046/j.1365-294X.2003.01863.x

Noonan JP, Hofreiter M, Smith D, Priest JR, Rohland N, Rabeder G et al (2005) Genomic sequencing of Pleistocene cave bears. Science 309(5734):597–599. doi:10.1126/science.1113485

O'Neill RJ, Eldridge MD, Crozier RH, Graves JA (1997) Low levels of sequence divergence in rock wallabies (Petrogale) suggest a lack of positive directional selection in SRY. Mol Biol Evol 14(3):350–353

Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N et al (2004) Genetic analyses from ancient DNA. Annu Rev Genet 38:645–679. doi:10.1146/annurev.genet.37.110801.143214

Paetkau D, Amstrup SC, Born EW, Calvert W, Derocher AE, Garner GW et al (1999) Genetic structure of the world's polar bear populations. Mol Ecol 10:1571–1584. doi:10.1046/j.1365-294x.1999.00733.x

Pagès M, Calvignac S, Klein C, Paris M, Hughes S, Hänni C (2008) Combined analysis of fourteen nuclear genes refines the Ursidae family. Mol Phyl Evol (in press)

Pfeiffer I, Brenig B (2005) X- and Y-chromosome specific variants of the amelogenin gene allow sex determination in sheep (*Ovis aries*) and European red deer (*Cervus elaphus*). BMC Genet 6(1):16. doi:10.1186/1471-2156-6-16

Ruiz-Garcia M (2003) Molecular population genetic analysis of the spectacled bear (*Tremarctos ornatus*) in the northern Andean

area. Hereditas 138(2):81–93. doi:10.1034/j.1601-5223.2003.01578.x

Ruiz-Garcia M, Orozco-terWengel P, Castellanos A, Arias L (2005) Microsatellite analysis of the spectacled bear (*Tremarctos ornatus*) across its range distribution. Genes Genet Syst 80(1):57–69. doi:10.1266/ggs.80.57

Shaw C, Wilson P, White B (2003) A reliable molecular method of gender determination for mammals. J Mammal 84(1):123–128. doi:10.1644/1545-1542(2003)084<0123:ARMMOG>2.0.CO;2

Swofford DL (1998) PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, MA

Taberlet P, Bouvet J (1994) Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear *Ursus arctos* in Europe. Proc R Soc Lond B Biol Sci 255(1344):195–200. doi:10.1098/rspb.1994.0028

Taberlet P, Mattock H, Dubois-Paganon C, Bouvet J (1993) Sexing free-ranging brown bears *Ursus arctos* using hairs found in the field. Mol Ecol 2(6):399–403. doi:10.1111/j.1365-294X.1993.tb00033.x

Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N et al (1996) Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Res 24(16):3189–3194. doi:10.1093/nar/24.16.3189

Taberlet P, Camarra JJ, Griffin S, Uhrès E, Hanotte O, Waits LP et al (1997) Noninvasive genetic tracking of the endangered Pyrenean brown bear population. Mol Ecol 6(9):869–876. doi:10.1111/j.1365-294X.1997.tb00141.x

Taberlet P, Waits LP, Luikart G (1999) Noninvasive genetic sampling: look before you leap. Trends Ecol Evol 14(8):323–327. doi:10.1016/S0169-5347(99)01637-7

Teletchea F, Maudet C, Hänni C (2005) Food and forensic molecular identification: update and challenges. Trends Biotechnol 23(7):359–366. doi:10.1016/j.tibtech.2005.05.006

Tucker PK, Lundrigan BL (1993) Rapid evolution of the sex determining locus in Old World mice and rats. Nature 364(6439):715–717. doi:10.1038/364715a0

Waits L, Talbot SL, Ward R, Shields GH (1998) Mitochondrial DNA phylogeography of the North American brown bear and implications for conservation. Conserv Biol 12(2):408–417. doi:10.1046/j.1523-1739.1998.96351.x

Waits LP, Sullivan J, O'Brien SJ, Ward RH (1999) Rapid radiation events in the family Ursidae indicated by likelihood phylogenetic estimation from multiple fragments of mtDNA. Mol Phylogenet Evol 13(1):82–92. doi:10.1006/mpev.1999.0637

Wasser SK, Houston CS, Koehler GM, Cadd GG, Fain SR (1997) Techniques for application of faecal DNA methods to field studies of ursids. Mol Ecol 6(11):1091–1097. doi:10.1046/j.1365-294X.1997.00281.x

Weikard R, Pitra C, Kuhn C (2006) Amelogenin cross-amplification in the family Bovidae and its application for sex determination. Mol Reprod Dev 73(10):1333–1337. doi:10.1002/mrd.20554

Whitfield LS, Lovell-Badge R, Goodfellow PN (1993) Rapid sequence evolution of the mammalian sex-determining gene SRY. Nature 364(6439):713–715. doi:10.1038/364713a0

Wooding S, Ward R (1997) Phylogeography and pleistocene evolution in the North American black bear. Mol Biol Evol 14(11):1096–1105

Woods J, Paetkau D, Lewis D, McLellan BN, Proctor M, Strobeck C (1999) Genetic tagging of free-ranging black and brown bears. Wildl Soc Bull 27(3):616–627

Wozencraft WC (1993) Order Carnivora. In: Wilson DE, Reeder DM (eds) Mammal species of the world—a taxonomic and geographic reference. Smithsonian Institution Press, Washington

Xu X, Lin L, Zhang Z, Shen F, Zhang L, Yue B (2007) A reliable, non-invasive PCR method for giant panda (*Ailuropoda melaneuloca*) sex identification. Conserv Genet (in press)

Yamamoto K, Tsubota T, Komatsu T, Katayama A, Murase T, Kita I et al (2002) Sex identification of Japanese black bear, *Ursus thibetanus japonicus*, by PCR based on amelogenin gene. J Vet Med Sci 64(6):505–508. doi:10.1292/jvms.64.505

Yu L, Li QW, Ryder OA, Zhang YP (2004) Phylogeny of the bears (Ursidae) based on nuclear and mitochondrial genes. Mol Phylogenet Evol 32(2):480–494. doi:10.1016/j.ympev.2004.02.015

Zhan X, Li M, Zhang Z, Goossens B, Chen Y, Wang H et al (2006) Molecular censusing doubles giant panda population estimate in a key nature reserve. Curr Biol 16(12):R451–R452. doi:10.1016/j.cub.2006.05.042