Conserv Genet
DOI 10.1007/s10592-008-9578-1

RESEARCH ARTICLE

Non-invasive genetic study of the endangered Cantabrian brown bear (*Ursus arctos*)

Trinidad Pérez · Fernando Vázquez · Javier Naves · Alberto Fernández · Ana Corao · Jesús Albornoz · Ana Domínguez

Population minimale / survie : voir surligné jaune et traduction pages 8 et 9

Received: 18 October 2007 / Accepted: 22 March 2008

Abstract The Brown Bear (*Ursus arctos*) population present in the Cantabrian Mountains has suffered a dramatic decline in recent centuries and is now threatened with extinction. This situation has led to the development and implementation of a species recovery plan. To accomplish this plan, we need to improve our knowledge about the ecology, demography and genetics of this population. This paper presents the genetic analysis of the Cantabrian brown bear population using non-invasive samples (faeces and hairs) collected between 2004 and 2006. It was necessary to optimize a set of 18 microsatellite loci and a sex marker (several new multiplex reactions were developed) to obtain a suitable probability of identity among genotypes to work with this small, deeply structured population. Genotyping of 48 individuals was carried out using a two-step PCR protocol to increase the quality of the multilocus genotypes. Validation of genotypes was performed using a multi-tube approach combined with different software programmes to measure their error rate and reliability. Diversity in the Cantabrian population was low ($H_e = 0.51$) and the population was markedly subdivided into two subpopulations (western and eastern) without current gene flow between them. The level of divergence between the two subpopulations ($F_{st} = 0.41$) and the extremely low diversity in the eastern group ($H_e = 0.25$) indicate that this has had an extremely low effective population size and had been isolated from the main group during the last century. Connectivity between the two subpopulations will be of prime importance for the long-term survival of this species in the Cantabrian Mountains.

Keywords *Ursus arctos* · Microsatellites · Non-invasive samples · Individual identification · Population structure

T. Pérez (✉) · F. Vázquez · A. Corao · J. Albornoz · A. Domínguez
Departamento de Biología Funcional (Genética), Facultad de Medicina 7°Planta, Universidad de Oviedo, C/Julián Clavería s/n, 33071 Oviedo, Spain
e-mail: mtpmendez@yahoo.es

J. Naves · A. Fernández
Departamento de Biología de Organismos y Sistemas (Ecología), Universidad de Oviedo, C/Catedrático Rodrigo Uría s/n, 33071 Oviedo, Spain

J. Naves · A. Fernández
Department of Applied Biology, Estación Biológica de Doñana, Consejo Superior de Investigaciones Científicas, Avda. de María Luisa s/n, Pabellón del Perú, 41013 Sevilla, Spain

Introduction

Brown bears were once found throughout Europe and even inhabited the British Isles until the tenth century. However, since the mid-1800's, populations in Europe have been severely reduced due to habitat destruction and overexploitation by humans (Servheen 1990). Four very small, isolated populations, all of which are endangered, can still be found in southern and western Europe (the Pyrenees, Southern Alps, Cantabrian Mountains and Apennine Mountains), representing the remnants of a once widespread brown bear population (Zedrosser et al. 2001). The Cantabrian population is found in two areas of the Spanish Cantabrian Mountains separated by 30–50 km of mountainous terrain. Interchange between both subpopulations is unlikely (Naves and Nores 1997), due to unsuitable habitat and a high speed railway and motorway that bisect the area (Fig. 1). The population estimate for the western subpopulation is around 50–60 individuals (Wiegand et al. 1998).

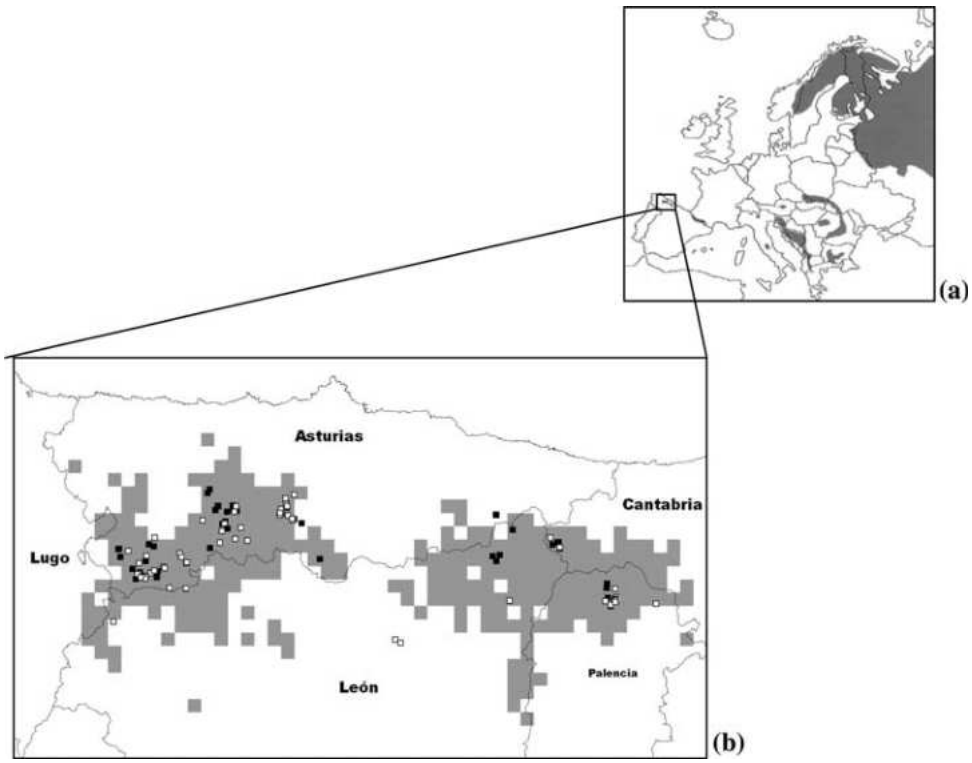

Fig. 1 Distribution map of the brown bear in Europe: (**a**) present distribution of the brown bear in Europe, and (**b**) distribution of the brown bear in the Cantabrian Mountains (based on Naves et al. 2003) and location of the 133 samples analysed in this study. White squares represent samples amplified for more than 14 markers, black squares represent samples with less than 14 loci amplified

The eastern Cantabrian Mountains subpopulation is estimated to contain around 20 bears (Naves et al. 1999). Both Cantabrian bear populations face similar conservation problems.

Mitochondrial DNA studies have shown that the Cantabrian population belongs to the Iberian refugia clade of the western lineage of European brown bears. Only the populations of the Cantabrian and Pyrenean Mountains and small populations from the south of Sweden and Norway belong to this clade (Taberlet and Bouvet 1994). It is important to preserve this population if we wish to maintain the most ancient lineage of the European brown bear.

In order to design an effective conservation plan, it is necessary to understand the structure of threatened populations, particularly those which, like this one, exist in degraded or fragmented habitats (Lande and Barrowclough 1987; Simberloff 1988; Hanski and Gilpin 1997; Taylor and Dizon 1999; Kraaijeveld-Smith et al. 2005). It is well known that small populations are more vulnerable to genetic factors, demographic and environmental stochasticity, genetic drift and inbreeding and have an increased probability of extinction (Soulé 1987). When small populations become fragmented and migration between subpopulations decreases or is eliminated, consequent increases in inbreeding and loss of genetic diversity can have serious negative effects on the long-term viability of population fragments and, by extension, of the population as a whole (Keller et al. 1994; Lacy 1997; Bjilsma et al. 2000; Sherwin and Moritz 2000; Coulon et al. 2004).

To avoid the extinction of this species in the Cantabrian Mountains, the different regional governments from the areas where it is present have developed plans for the conservation of brown bears since 1989. These plans have been coordinated at a national level. In order to be effective, however, an action plan should be based on reliable biological data, such as trustworthy estimates of population size, population genetic status and connectivity with other populations (Bellemain et al. 2007).

To assess the genetic status of the Cantabrian population and provide guidelines for the conservation and management of this population, we used non-invasive genetic techniques which have been successfully used in other studies (Taberlet et al. 1997; Kohn et al. 1999; Woods et al. 1999; Waits et al. 2000; Frantz et al. 2003; Smith et al. 2006; Bellemain et al. 2007). However, this non-invasive approach has a major drawback, namely the problems associated with low DNA quantity and quality (Taberlet and Luikart 1999). Under these circumstances, the probability of critical genotyping errors (primarily allelic dropout and false alleles) is high (Taberlet and Luikart 1999; Taberlet et al. 1996, 1999). The solutions proposed in these studies for addressing and reducing the severity of genotyping errors and increasing the reliability of genotypes include protocols for replicating amplification and programmes for determining the magnitude of error in a dataset and for calculating the reliability of genotype data (Taberlet et al. 1996; Miller et al. 2002; Bonin et al. 2004; McKelvey and Schwartz 2004).

The goal of the present study was to ascertain the current genetic status of the Cantabrian brown bear population using non-invasive genotyping techniques. The results of this work will help us to provide guidelines for the conservation and management of this population. To achieve this goal, we optimized a set of suitable loci microsatellite markers for carrying out population and individual identification studies in this population.

Methods

Sampling

This study presents the genetic data for the Cantabrian brown bear population using non-invasive samples (faeces and hairs) collected in the field between 2004 and 2006 using two sampling methods: opportunistic and systematic. Most of the samples ($n = 106$) were collected in an opportunistic manner from daily routine field work of rangers and field biologists. Some samples ($n = 27$) were collected through systematic surveys carried out seasonally (summer: mid August–mid September; autumn: mid October–mid November) over a grid (2.5 km^2) in the western subpopulation in an area of roughly 750 km^2 covering around 630 km of on-foot surveys each season. Although we were mainly looking for scats, some hair samples were also collected. Since we wished to cover the maximum area of distribution possible, we had to include many opportunistic samples because the systematic surveys only cover a partial area of the species distribution. Figure 1b shows where the samples were collected.

Faeces samples were dry stored with silica after ethanol soaking upon collection following Nsubuga et al. (2004) and Roeder et al. (2004). Hairs were placed in individual envelopes, no further manipulation being necessary until extraction. A total of 133 samples were analysed: 88 from the western subpopulation and 45 from the eastern subpopulation. Together with these non-invasive samples, 13 tissue/blood samples belonging to either dead or captured individuals were used (all from the western subpopulation, eight of which were found dead between 1989 and 2002). The numbers from each type of sample used for the study are listed in Table 1.

Table 1 Number of samples analysed and percent of the total (between parentheses)

	Faeces	Hair	Skin/blood	Total
No. DNA	25 (27.2%)	0 (0%)	1 (7.7%)	26 (17.8%)
Discarded genotypes[a]	27 (29.3%)	6 (14.6%)	3 (23.1%)	36 (24.7%)
Accepted genotypes[b]	40 (43.5%)	35 (85.4%)	9 (69.2%)	84 (57.5%)
Total number samples	92	41	13	146

[a] Include all the incomplete genotypes (with less than 14 loci amplified)

[b] Genotypes with more than 14 loci amplified (all the hair, skin/blood samples and 24 of the faeces samples in this category were genotyped for the 18 loci)

DNA extraction and typing

DNA extraction

DNA was extracted from faeces samples using the 'Qiamp DNA stool kit' (Qiagen, Hilden, Germany) specially developed for this type of material and following the manufacturer's instructions. The only change introduced was the incubation of the samples in ASL buffer ON at 25°C. For hair DNA extraction, 5–10 hair roots were used; these were incubated ON with PCR buffer and proteinase K (Allen et al. 1998; Vigilant 1999). Finally, tissue/blood samples were extracted using the 'DNeasy Tissue kit' (Qiagen, Hilden, Germany) following the manufacturer's instructions. All hair and faeces extractions were carried out in a separate room used exclusively for processing samples of this type.

Genotyping

For the genotyping, we used a total of 18 microsatellite markers (G1A, G1D, G10B, G10C, G10J, G10L, G10O, G10P, G10X (Paetkau and Strobeck 1994; Paetkau et al. 1995) and MU05, MU09, MU10, MU23, MU50, MU51, MU59, MU61, MU64 (Taberlet et al. 1997)), as well as the sex marker SRY (Bellemain and Taberlet 2004), all of which were selected from markers previously used in genetic tracking of Pyrenean brown bears (Taberlet et al. 1997).

DNA amplifications were performed in a two-step PCR following Taberlet et al. (1997) for faeces samples and using only one PCR for hair and tissue/blood samples. To avoid the problem of running out of DNA template (because of the low amount of DNA available) before completing the genotyping for all the microsatellite markers, the amplification was carried out following the multiplex preamplification method (Piggot et al. 2004; Bellemain and Taberlet 2004). This method was optimized for six microsatellite loci and for the sex marker (MU10, MU23, MU50, MU51, MU59, G10L and SRY) by Bellemain and Taberlet (2004). We designed three new multiplex PCRs to amplify all the remaining loci (MU64 + G1A + G10C + G10P, MU61 + G10J + G10O + G10X, MU05 + MU09 + G1D + G10B). A total number of four first-step PCRs, where we amplified

between 4 and 6 loci simultaneously, and nine second-step PCRs, where the number of loci amplified simultaneously varied between 1 and 3, were sufficient to amplify the 18 loci microsatellites and the sex marker (Table 2).

For faeces samples, the first-step PCRs or preamplifications were prepared in a 25 µl volume containing 5 µl template DNA, 0.01 µM of each primer and 12.5 µl of "Qiagen Multiplex PCR Kit" (Qiagen, Hilden, Germany). The second-step PCRs or amplifications were prepared in a 13 µl volume containing 3 µl preamplified product, 0.1 mM of each dNTP, 0.5 µM of each primer, 2 mM $MgCl_2$, 0.5 U *Taq* DNA Polymerase from Qiagen (Qiagen, Hilden, Germany) and 1× *Taq* Qiagen buffer (Qiagen, Hilden, Germany). Amplifications were performed on a GeneAmp PCR 9600 (Applied Biosystems) under the following conditions: for the first-step PCR 15 min at 95°C, 40 cycles composed of 30 s denaturing at 94°C, 90 s annealing at 60°C, 1 min extension at 72°C, and as a final extension step, 30 min at 60°C. For the second-step PCR, 3 min at 94°C, 35 cycles composed of 30 s denaturing at 94°C, 30 s annealing at 60°C, 1 min extension at 72°C, and as a final extension step, 7 min at 72°C.

For hair and tissue/blood samples, nine one-step PCRs were carried out using the primers of the second-step PCRs in a 20 µl volume containing 5 µl (2 µl for tissue/blood samples) template DNA, 0.5 µM of each primer and 10 µl of "Qiagen Multiplex PCR Kit" (Qiagen, Hilden, Germany). Amplifications were performed on a GeneAmp PCR 9600 (Applied Biosystems) under the following conditions: 15 min at 95°C, 40 cycles composed of 30 s denaturing at 94°C, 90 s annealing at 60°C, 1 min extension at 72°C, and as a final extension step, 30 min at 60°C.

PCR products were checked in a 2% agarose gel and the product diluted between 0 and 100 times depending on the intensity of the signal. One microlitre of this product was added to a 12 µl mix of formamide and ROX 400HD (12:0.2) and then loaded on an automatic sequencer ABI310 (Applied Biosystems). Microsatellite patterns were examined both visually and using GENESCAN ANALYSIS 3.1 and GENOTYPER 2.5 software (Applied Biosystems).

Reliability of genotyping results

To test the reliability of our genotyping results and to reduce tracking error in our dataset, we followed different recommendations already proposed in different studies addressing non-invasive genotyping. The suggestions by Bonin et al. (2004) for limiting potential errors in the genotyping process were followed. All the genetic typing was performed using a combination of the multi-tube approach and software packages that assign a reliability value to each multilocus genotype. Three positive PCRs (for both types of non-invasive samples, hair and faeces) were first analysed, a consensus genotype was assigned

Table 2 Observed number of alleles (*A*) with the number of single alleles of each subpopulation between parentheses (Ae), observed heterozygosity (H_o), Nei's estimated heterozygosity (H_e) and deviations from Hardy–Weinberg equilibrium (HWE) by locus for each subpopulation with sample size in parentheses

Locus (i,j)	Eastern subpopulation ($n = 8$)			Western subpopulation ($n = 39$)		
	A (Ae)	H_e	H_o	*A* (Ae)	H_e	H_o
MU10 (1,1)	2 (0)	0.50	0.50	4 (2)	0.42	0.28*
G10L (1,1)	1 (0)	0.00	0.00	6 (5)	0.74	0.64*
MU50 (1,2)	1 (0)	0.00	0.00	5 (4)	0.67	0.69
MU23 (1,2)	3 (1)	0.63	0.88	3 (1)	0.65	0.72
MU59 (1,3)	1 (0)	0.00	0.00	4 (3)	0.56	0.59
MU51 (1,3)	2 (0)	0.49	0.63	4 (2)	0.37	0.38
G10C (2,1)	1 (0)	0.00	0.00	3 (2)	0.50	0.41
MU64 (2,1)	1 (1)	0.00	0.00	3 (3)	0.46	0.33*
G1A (2,2)	2 (1)	0.30	0.38	3 (2)	0.51	0.64
G10P (2,2)	2 (1)	0.22	0.25	2 (1)	0.45	0.49
G10J (3,1)	2 (0)	0.38	0.50	3 (1)	0.65	0.69
G10X (3,2)	2 (1)	0.49	0.13*	4 (3)	0.29	0.26
MU61 (3,2)	1 (0)	0.00	0.00	2 (1)	0.44	0.49
G10O (3,2)	1 (0)	0.00	0.00	1 (0)	0.00	0.00
MU05 (4,1)	3 (1)	0.63	0.75	4 (2)	0.58	0.56
G1D (4,1)	2 (1)	0.30	0.38	1 (0)	0.00	0.00
MU09 (4,2)	2 (1)	0.50	0.75	4 (3)	0.66	0.51*
G10B (4,2)	1 (0)	0.00	0.00	3 (2)	0.17	0.13
Mean	1.67	0.25	0.28	3.33	0.45	0.44

Next to the locus name, between parentheses, the number of the preamplification (i) and amplification (j) PCR

*P-value significant $P < 0.05$

using the GIMLET v.1.3.2 software (Valière 2002) and its reliability was tested using the RELIOTYPE software (Miller et al. 2002). If we found a multilocus genotype with reliability lower than 95%, more repetitions were carried out until achieving said level of reliability. All the samples that could not be reliably typed for at least 14 out of the 18 loci after the entire process was completed were discarded. To further identify any genotyping errors and the relative magnitude of a problem within our multilocus scores, we performed the tests proposed by McKelvey and Schwartz (2004): Examining Bimodality (EB) and Difference in Capture History (DCH). Finally, we determined genotypic mismatches between all scores.

Probability of identity

Using the software GIMLET version 1.3.2 (Valière 2002), we computed the probability of identity (P_{ID}), which is the probability that two individuals drawn at random from a given population share identical genotypes at all typed loci (Paetkau and Strobeck 1994). We also computed the probability of identity among siblings ($P_{ID\text{-}Sib}$) (Waits et al. 2001). This value is the upper limit of the possible ranges for the probability of identity in a population and thus provides the most conservative number of loci required to resolve all bears, including relatives. These calculations were carried out for each subpopulation.

Population genetic parameters and structure

We ran population genetic analyses using the software programmes GENEPOP version 3.4 (Raymond and Rousset 1995), GENETIX version 4.02 (Belkhir et al. 1996–2004) and STRUCTURE version 2.1 (Pritchard et al. 2000). Nuclear genetic diversity was measured as the number of alleles per locus (A), the observed heterozygosity (H_o), as well as Nei's unbiased expected heterozygosity (H_e) (Nei 1978). Deviations from Hardy–Weinberg equilibrium were tested using an exact test. Global tests across loci for heterozygote deficiency and heterozygote excess and pairwise tests for linkage disequilibrium were performed using Fisher's method (Sokal and Rohlf 1994) with 10,000 batches and 10,000 iterations per batch. Correspondence among individual genotypes was studied by means of Factorial Correspondence Analysis (FCA), performed with the GENETIX software. Population substructure was detected with the programme STRUCTURE, which uses a Markov Chain Monte Carlo (MCMC) algorithm to cluster individuals into populations on the basis of multilocus genotype data (Pritchard et al. 2000). We used different values of K, from one to five. For each K tested, we ran STRUCTURE 20 times for 100,000 steps, after a burn-in period of 50,000 steps. The correct value of K was estimated following Evanno et al. (2005). The programme also calculates the fractional membership of each individual in each cluster (Q). Quantification of variation among subpopulations was performed with GENEPOP using Wright's F-statistics (Weir 1996).

Results

Reliability of non-invasive genotyping for the identification of the Cantabrian bear

A total number of 133 non-invasive samples were analysed (92 faeces samples and 41 hair samples). We managed to obtain some amplification for all the hair samples and for 67 faeces samples (72.8%), but only those samples successfully amplified for more than 14 markers were included in the analysis (Table 1). Another nine genotypes were obtained from tissue/blood samples. For the genotyping, we initially tested 24 loci previously used in genetic tracking of Pyrenean bears (Taberlet et al. 1997; Bellemain and Taberlet 2004). Six out these 24 loci did not give a scorable product and so were discarded. Finally, we used the remaining 18 loci and the sex marker; all the loci were amplified using multiplex PCRs, three of which were specifically developed for this study (Table 2). The probability of identity values were 3.28E-09 (P_{ID}) and 1.16E-04 ($P_{ID\text{-}Sib}$) for the western subpopulation and 7.45E-05 and 8.64E-03 for the eastern subpopulation. The P_{ID} and $P_{ID\text{-}Sib}$ values for each marker and subpopulation are shown in Table 3.

Once we had completed three repetitions per sample, we found that out of the 67 faeces that gave a product, only 40 (59.7%) worked for at least 14 markers. For hair samples, 35 out of 41 (85.4%) gave a complete profile. The error rates calculated using the GIMLET programme show that both the number of failed PCRs and percentage of total error are higher for faeces samples than for hair samples (see Fig. 2). These results indicate that the DNA recovered from hair samples has better quality than that recovered from faeces samples. Results from the RELIOTYPE programme showed that 76% of the samples reach 95% reliability after three repetitions, 16% needed a fourth repetition to reach this level of reliability and 8% needed more than four repetitions to reach it (one of the samples had to be repeated up to seven times to achieve 95% reliability).

The results from the EB test, which examines the distribution of the genetic differences between samples, and the DCH test, which determines whether the number of new individuals in the sample increased faster than would be expected when additional loci are added (McKelvey and Schwartz 2004), indicated that our dataset had a low rate of genotyping error (unimodal distribution for the EB test and

Table 3 Probability of identity (P_{ID}) and Probability of identity for siblings ($P_{ID\text{-}sib}$) by locus for each subpopulation with sample size in parentheses and accumulated values

Locus (i,j)	Eastern subpopulation ($n = 8$)		Western subpopulation ($n = 39$)	
	P_{ID}	$P_{ID\text{-}sib}$	P_{ID}	$P_{ID\text{-}sib}$
MU10 (1,1)	3.75E−01	5.94E−01	3.73E−01	6.34E−01
G10L (1,1)	1.00E+00	1.00E+00	1.10E−01	4.10E−01
MU50 (1,2)	1.00E+00	1.00E+00	1.60E−01	4.56E−01
MU23 (1,2)	2.12E−01	4.87E−01	1.93E−01	4.71E−01
MU59 (1,3)	1.00E+00	1.00E+00	2.71E−01	5.40E−01
MU51 (1,3)	3.79E−01	5.99E−01	4.32E−01	6.71E−01
G10C (2,1)	1.00E+00	1.00E+00	3.61E−01	5.89E−01
MU64 (2,1)	1.00E+00	1.00E+00	3.54E−01	6.10E−01
G1A (2,2)	5.30E−01	7.30E−01	3.58E−01	5.84E−01
G10P (2,2)	6.34E−01	7.99E−01	4.02E−01	6.24E−01
G10J (3,1)	4.61E−01	6.78E−01	1.96E−01	4.73E−01
G10X (3,2)	3.79E−01	5.99E−01	5.18E−01	7.33E−01
MU61 (3,2)	1.00E+00	1.00E+00	4.09E−01	6.32E−01
G10O (3,2)	1.00E+00	1.00E+00	1.00E+00	1.00E+00
MU05 (4,1)	2.12E−01	4.87E−01	2.38E−01	5.20E−01
G1D (4,1)	5.30E−01	7.30E−01	1.00E+01	1.00E+01
MU09 (4,2)	3.75E−01	5.94E−01	1.82E−01	4.68E−01
G10B (4,2)	1.00E+00	1.00E+00	7.07E−01	8.44E−01
Accumulated	7.45E−05	8.64E−03	3.28E−09	1.16E−04

Next to the locus name, between parentheses, the number of the preamplification (i) and amplification (j) PCR

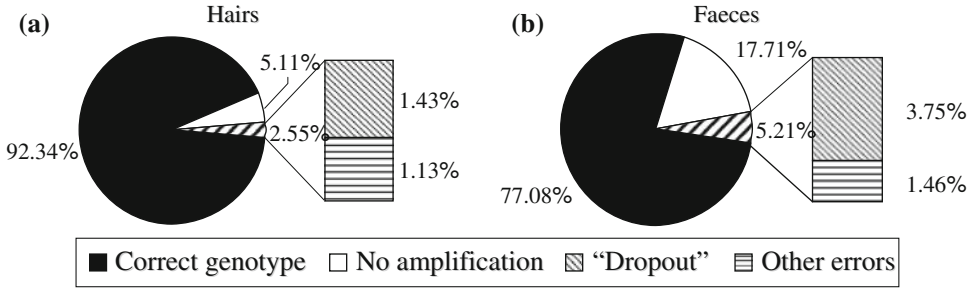

Fig. 2 Percentage of errors (failed PCR, dropout and other types of error) and correct genotypes depending on the non-invasive sample type: (**a**) hair samples, and (**b**) faeces samples

no locus-added new individuals for the DHC test). The presence of a bimodal structure (Fig. 3) in the distribution of the number of loci at which individuals differed (EB test) for the eastern subpopulation is not necessarily a result of error; in this case, it indicates that one of the individuals is highly different from the rest, showing that it could be a migrant. The DHC test could not be carried out on the eastern subpopulation, since all the polymorphic loci tested are included in the genetic tag.

The average number of loci at which individuals differed was 10.62 ± 1.78 for the western subpopulation and 7.69 ± 4.2 for the eastern, although if we remove the individual that appears to be different from the rest, the average is lower (5.68 ± 1.94).

A final number of 31 individuals (16 females, 15 males) out of 45 samples were identified in the western subpopulation and 9 (5 females, 4 males) out of 30 samples in the eastern subpopulation.

Microsatellite diversity and population substructure

The number of alleles for the total population was 67; of these, 45 were uniquely sampled from one of the two subpopulations (8 for the eastern subpopulation and 37 for the western one). The null hypothesis of uniform allelic and genotypic frequencies in the two subpopulations was rejected at 15 out of 17 polymorphic loci. The number of alleles per locus for the western subpopulation ranged between 1 and 6, with an average of 3.33 ± 1.28 and between 1 and 3 for the eastern subpopulation with an average of 1.67 ± 0.67 (Table 2). Two loci were monomorphic for the western subpopulation (G10O and G1D), but up to eight loci were

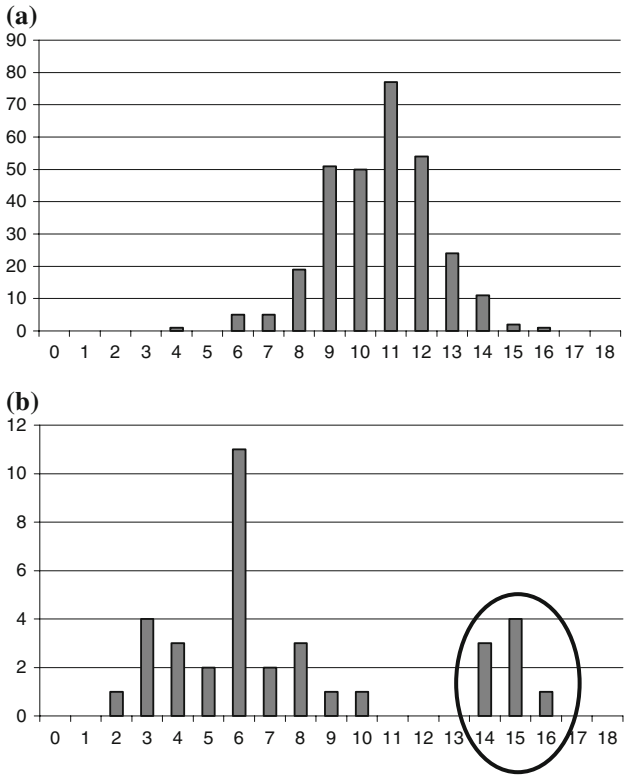

Fig. 3 Number of loci at which one individual differed from all the other individuals from the same population based on 18 loci for both populations: (**a**) western and (**b**) eastern. Only individuals with completed genotypes were used in this calculation (25 for the western subpopulation and 9 for the eastern one). All the comparisons involving one of the individuals found in the eastern subpopulation (in a circle) differed in a higher number of loci from all the rest of the individuals from the same subpopulation

monomorphic for the eastern subpopulation (G10L, Mu50, Mu59, G10C, Mu64, Mu61, G10O, and G10B). The locus G10O was monomorphic for both subpopulations. The average observed and expected heterozygosities for the total population were 0.51 and 0.43, respectively, and were significantly different ($P < 0.001$). The partition of genetic diversity among subpopulations was 42.96%; the values of diversity found for each subpopulation were 0.45 for the western subpopulation and 0.25 for the eastern one (Table 2).

Global tests showed that both subpopulations are in Hardy–Weinberg equilibrium, although four loci (Mu10, G10L, Mu64, Mu09) for the western subpopulation and one locus (G10X) for the eastern subpopulation had a significant deficiency in heterozygotes at the $P < 0.05$ level (Table 2). Statistical tests for linkage disequilibrium were computed for all pairs of loci, though none of these were significant.

The canonical analysis based on factorial correspondence analysis (FCA) divided all the samples in two clearly differentiated groups (Fig. 4). Each group corresponded with the two subpopulations present in the Cantabrian Mountains (western and eastern subpopulations). One individual that was detected in samples collected in the eastern area groups together with the individuals from the western subpopulation. This is the same individual that showed a strange pattern when we checked the distribution of the number of loci at which two individuals differed for the eastern subpopulation (Fig. 3). Furthermore, the programme STRUCTURE detected that Cantabrian brown bears are structured in two subpopulations which correspond with the eastern–western groupings (Fig. 4). The mean Q value (cluster membership) is 0.998 for the western subpopulation and 0.888 for the eastern one. Once more, one individual is detected in the eastern subpopulation that is more similar to the individuals from the western group. The Q value of this individual is 0.010 for the eastern subpopulation and 0.998 for the western one, indicating that this individual (a male) is a migrant. If we remove this individual, the mean Q for the eastern subpopulation is 0.998. For all the aforementioned reasons, this individual was removed from the basic statistical calculations regarding population differentiation.

The degree of genetic differentiation between the two subpopulations was considerable. The proportion of variation in genetic frequencies, θ (F_{st}), accounted for by subdivision was 0.41, while Nei's standard distance between both subpopulations was 0.47.

Discussion

Quality of the genetic data

In order to maximize the success of the non-invasive sample DNA extracts, we used storage and extraction techniques that have previously shown their effectiveness. We obtained 72.8% amplification success for faeces and 100% for hairs; these values are in the range of values described in the literature (see Nsubuga et al. 2004; Chu et al. 2006). However, these success values dropped to 43.48% for faeces and 85.37% for hairs when we considered only those samples that gave us a multilocus profile for at least 14 out the 18 loci used.

Taberlet et al. (1999) suggest that non-invasive studies should include a pilot study to assess the confidence level of the final result by quantifying the genotyping error rate. The results of studies of this type are unique and cannot be transferred to another species or even to another population with different heterozygosity or sample quality. By combining several methodological and statistical methods for tracking and reducing error previously used in different non-invasive studies (Frantz et al. 2003; Bellemain and Taberlet 2004; Smith et al. 2006), we were able to thoroughly assess the level of error in our data set (2.55% for

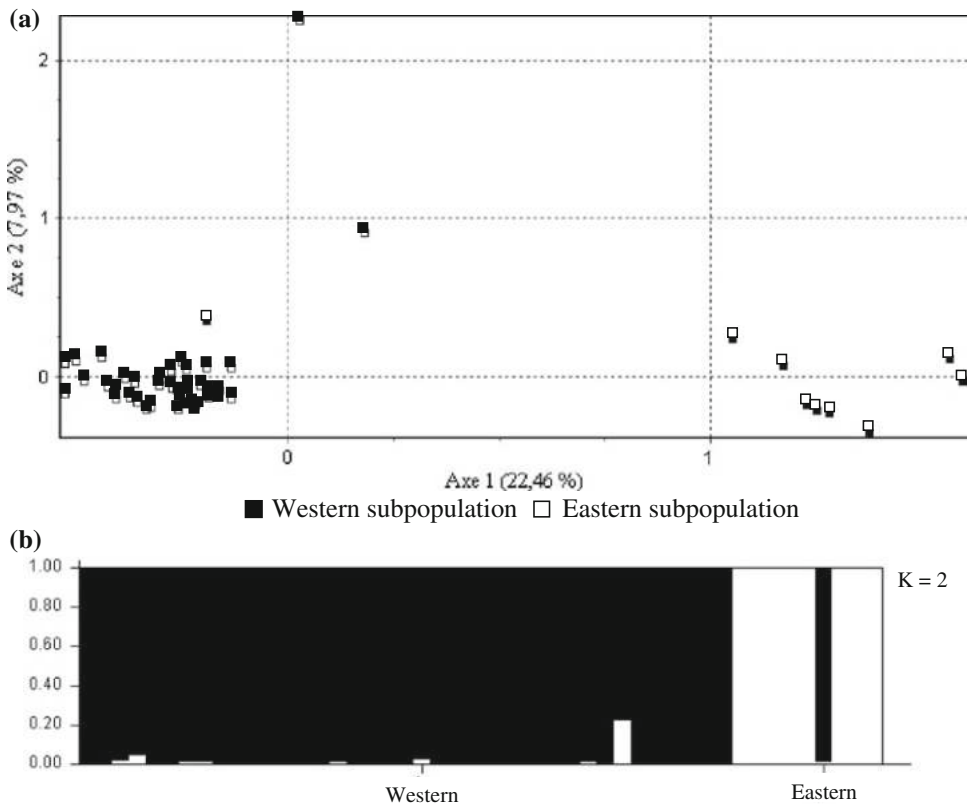

Fig. 4 (**a**) Distribution of individuals according to genotype based on factorial correspondence analysis (FCA), and (**b**) Graphical representation of the STRUCTURE programme. In both cases individuals belonging to the eastern and western populations are indicated in white and black, respectively

hair samples and 5.21% for faeces samples). These error values are similar to the lowest found in the literature for these types of samples (Bayes et al. 2000; Bellemain and Taberlet 2004; Smith et al. 2006). The reliability of the multilocus genotypes after three repeats was quite high (76% of the samples had >95% reliability). A fourth repetition increases this percentage up to 92%, which is in concordance with the result previously presented by Bellemain and Taberlet (2004) in the brown bear.

The low diversity found in the Cantabrian population, mainly in the eastern subpopulation, together with the fact that the most informative loci are not the same in the two subpopulations, makes it necessary to use a high number of markers if we wish to unambiguously identify individuals for future demographic studies.

Genetic status of the bear population in the Cantabrian Mountains

The genetic diversity of the Cantabrian population is lower than in other bear populations considered to have a good conservation status (Paetkau et al. 1998; Waits et al. 1998) and is comparable with those found in Yellowstone ($H_e = 0.55$; Paetkau et al. 1998) and Deosai National Park (Pakistan) ($H_e = 0.55$; Bellemain et al. 2007). Yellowstone and Cantabrian populations have had a very similar history. Both populations have gone from being embedded in a very large continuous population to being an isolated remnant, separated from other brown bears for 300 years in the case of the Cantabrian population (Naves et al. 1999) and 100 years for the Yellowstone population (Paetkau et al. 1998). In both cases, there is no prospect of renewed connections with other populations. The low heterozygosity values can be explained either by a founder effect, which is not the case for the Cantabrian brown bear, or for a sharp decline in population size. The whole population probably began to lose genetic diversity about 300 years ago, when it began to decline in size from a larger population.

The brown bear habitat in the Cantabrian Mountains has decreased considerably from ~9,000 km^2 at the turn of the twentieth century to ~5,000 km^2 at present (Naves and Nores 1997), while the population subdivided into two apparently isolated subpopulations, the western and the eastern. Genetic differentiation ($F_{st} = 0.41$, Nei's standard $D = 0.47$) between the two subpopulations was found to be very high. These values can be compared with the reported microsatellite-based estimates of differentiation among other bear populations. For example, the degree of genetic differentiation is considerably higher than the values reported for pairwise comparisons among subpopulations in Scandinavia, where F_{st} ranged between 0.01 and 0.14 and Nei's standard distance ranged between 0.03 and 0.38. The genetic distance between the two Cantabrian

Traduction abrégée Les populations de Yellowstone et des Monts cantabriques ont une histoire très D•abord parties d•un vaste ensemble étendu en continuité, puis séparées des autres depuis 300 a Asturies, 100 pour Yellowstone, sans possibilité de connexion avec d•autres populations. Fa d•hétérozygotie /ƒ/ Pour les ours cantabriques, perte de diversité génétique depuis environ 300 ans /ƒ/

subpopulations, which are only 30 km apart, are comparable with the values reported for the most distant areas within the continuous distribution of brown bears in North America, which are several thousands of kilometres apart (Paetkau et al. 1998). This result also shows that habitat discontinuities such as roads and farmland play a larger role in genetic substructuring of population than linear distance, which has also been seen in populations of Cross River gorilla (Bergl and Vigilant 2007).

The huge genetic differentiation shown between Cantabrian subpopulations may be related to total isolation between them in conjunction with an extremely low population size in the eastern subpopulation. Diversity in this subpopulation is among the lowest found in the literature and can only be compared with the value reported in the isolated population of the Kodiak islands in Alaska ($H_e = 0.26$; Paetkau et al. 1998). The eastern subpopulation, with a diversity of 0.25, was detached from the main group a few generations ago. Assuming that diversity at the time of the split was comparable with the diversity of the western subpopulation, the relative loss of diversity in the eastern subpopulation ($H_{subpopulation}/H_{initial}$) is 0.56. This leads to a fixation index ($1 - H_{subpopulation}/H_{initial}$) equal to 0.44, quite close to the estimated value of F_{st}. If we assume a continent–island model (given that the relative effect of drift in the western subpopulation is low) with no migration, $F_{st} = 1 - (1 - 1/2N_e)^t$, where N_e is the effective size and t the time in generations. Considering that both subpopulations have been isolated for 10 generations (50–75 years with a generation time for the brown bear of 10–15 years, Allendorf and Servheen 1986; Craighead et al. 1995), the N_e for the eastern subpopulation that would explain the obtained F_{st} would be 4.99 individuals per generation. These figures show that the two subpopulations have probably been totally isolated without any effective migration during the last few generations. Although the data point towards a total lack of genetic flow between the two subpopulations, a migrant male from the western subpopulation into the eastern grouping was identified. This could either be interpreted as the first signal that connectivity between both subpopulations is starting to occur after this long period of isolation, or as the existence of a certain rate of migration, though not effective gene flow. It will be necessary to check in the future whether more migrants are present in the population and whether or not this migration is effective in terms of genetic flux. The main genetic consequence of migration would be a drop in genetic differentiation between the two subpopulations of the Cantabrian brown bear. If we once more assume a continent–island model, the F_{ST} at equilibrium between gene flow due to migration and genetic drift due to the small population size is $\hat{F}_{ST} \approx 1/(4N_em + 1)$, where N_e is the effective size and m is the migration rate per generation. This implies that with one effective migrant per generation ($N_em = 1$), the equilibrium F_{ST} is 0.20, and 0.11 should there be two effective migrants per generation. That would be sufficient to prevent the huge effects of genetic drift in the oriental subpopulation, despite its reduced population size.

Conclusions and recommendations

The results of this study show that the population of Cantabrian brown bears is effectively split into two subpopulations with a very high level of differentiation. Applying this unique criteria, it could be thought that they should be treated as separate management units (MU) (Moritz 1994) However, bearing in mind their contiguous distribution range, the reduced population size of both subpopulations and the real possibility of migration, these two subpopulations need to be managed as a unique unit. The rate of inbreeding per generation in the eastern subpopulation is around 10%, a value far exceeding the maximum tolerable rate of 1% given for domestic animals (Franklin 1980). Therefore, connectivity of both subpopulations should be the highest priority if we wish to maintain the diversity afforded by the eastern nucleus which is in risk of immediate extinction.

The western subpopulation shows a moderate level of diversity in the lowest range of values found in the species, probably due to a sharp decline in population size that began around 300 years ago, when it was isolated from a larger population. A population size of 50–60 individuals was estimated for this nucleus on the basis of the number of females with cubs (Wiegand et al. 1998). It was suggested (Franklin 1980) that the minimum effective size for a population to be viable in the short-term should be 50. Bearing in mind that the ratio of N_e/N found in the brown bear from Yellowstone was 0.27 (Miller and Waits 2003) and using this value in our case, the minimum size of population to be viable in the short-term is $\sim$186 individuals, and even the western subpopulation is far from this number.

If the whole population is far from the minimum number to be viable over the short-term, an even more difficult situation can be depicted in the context of long-term protection of adaptative potential where the minimum effective population size that has been put forward is between 500 and 5,000 (Franklin 1980; Frankham and Franklin 1998; Lynch and Lande 1998). It will be necessary to monitor the whole population to obtain a more accurate estimate of population size and its trend. Habitat loss and human-caused mortality should be avoided in order to facilitate an increase in population size.

Acknowledgements This work was funded by Grant CN-05-030 from the "Consejería de Medio Ambiente y Ordenación del Territorio e Infraestructuras del Principado de Asturias" We wish to thank the "Junta de Castilla y León" for their economic support

Traduction =On a suggéré (Franklin, 1980) que la taille minimale pour qu•une population soit viable à court terme être 50. Vu que le rapport Ne / N chez l'ours brun du Yellowstone est de 0,27 (Miller et Waits 2003) et en utilisant c dans notre cas, la taille minimale de la population pour assurer cette viabilité à court terme est de 186 individus (voir pag précédente justification du parallèle avec Yellowstone) - BBC

and for providing some of the samples; Miguel Rico, Juan Seijas and the "Guardería Rural del Principado de Asturias", who also provided samples; Dr. Lissette Waits and the two anonymous reviewers who helped to improve this manuscript with their valuable comments, and Paul Barnes, who thoroughly reviewed the English.

References

Allen M, Engström AS, Meyers S et al (1998) Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities. J Forensic Sci 43:453–464

Allendorf FW, Servheen C (1986) Genetics and the conservation of grizzly bears. Trends Ecol Evol 1:88–89

Bayes MK, Smith KL, Alberts SC, Brudford MW (2000) Testing the reliability of microsatellite typing from faecal DNA in the savannah baboon. Conserv Genet 1:173–176

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996–2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France)

Bellemain E, Taberlet P (2004) Improved noninvasive genotyping method: application to brown bear (*Ursus arctos*) faeces. Mol Ecol Notes 4:519–522

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

Bergl RA, Vigilant L (2007) Genetic analysis reveals population structure and recent migration within the highly fragmented range of the Cross River gorilla (*Gorilla gorilla diehli*). Mol Ecol 16:501–516

Bjilsma R, Bundgaard J, Boerema A (2000) Does inbreeding affect the extinction risk of small population? Predictions from *Drosophila*. J Evol Biol 13:502–514

Bonin A, Bellemain E, Eidesen PB et al (2004) How to track and assess genotyping errors in population genetics studies. Mol Ecol 13:3261–3273

Chu JH, Lin YS, Wu HY (2006) Applicability of non-invasive sampling in population genetic study of Taiwanese macaques (*Macaca cyclopis*). Taiwania 51:258–265

Coulon A, Cosson JF, Angibault JM et al (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. Mol Ecol 13:2841–2850

Craighead L, Paetkau D, Reynolds HV, Vyse ER, Strobeck C (1995) Microsatellite analysis of paternity and reproduction in Artic grizzly bears. J Hered 86:255–261

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Frankham R, Franklin IR (1998) Response to Lynch and Lande. Anim Conserv 1:73

Franklin IR (1980) Evolutionary change in small populations. In: Soulé ME, Wilcox BA (eds) Conservation biology: an evolutionary-ecological perspective. Sinauer, Sunderland, MA, pp 135–150

Frantz AC, Pope LC, Carpenter PJ et al (2003) Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. Mol Ecol 12:1649–1661

Hanski IA, Gilpin M (1997) Metapopulation biology: ecology and evolution. Academic Press, Toronto

Keller L, Arcese P, Smith J, Hochachka WM, Stearns SC (1994) Selection against inbred song sparrows during a natural population bottleneck. Nature 372:356–357

Kohn MH, York EC, Kamradt DA et al (1999) Estimating population size by genotyping faeces. Proc R Soc Lond B 266:657–663

Kraaijeveld-Smith FLJ, Beebee TJC, Griffiths RA, Moore RD, Schley L (2005) Low gene flow but high genetic diversity in the threatened Mallorcan midwife toad *Alytes muletensis*. Mol Ecol 14:3307–3315

Lacy RC (1997) The importance of genetic variation to the viability of mammalian populations. J Mammal 78:320–335

Lande R, Barrowclough G (1987) Effective population size, genetic variation and their use in population management. In: Soulé ME (ed) Viable populations for conservation. Cambridge University Press, New York, pp 87–123

Lynch M, Lande R (1998) The critical effective size for a genetically secure population. Anim Conserv 1:70–72

McKelvey KS, Schwartz MK (2004) Genetic errors associated with population estimation using non-invasive molecular tagging: problems and new solutions. J Wildl Manage 68:439–448

Miller CR, Waits LP (2003) The history of effective population size and genetic diversity in the Yellowstone grizzly (*Ursus arctos*): implication for conservation. Proc Natl Acad Sci USA 7:4334–4339

Miller CR, Joyce P, Waits LP (2002) Assessing allelic drop-out and genotype reliability using maximum likelihood. Genetics 160:357–366

Moritz C (1994) Defining 'Evolutionary significant units' for conservation. Trends Ecol Evol 9:373–375

Naves J, Nores C (1997) Status of the brown bear in western Cantabria, Spain. In: Servheen C, Herrero S, Peyton B (eds) Bears: status survey, conservation action plan. International Union for the Conservation of Nature, Natural Resources. Gland, Switzerland, pp 104–111

Naves J, Wiegand T, Fernandez A, Stephan T (1999) Riesgo de extinción del oso pardo cantábrico La población occidental. Fundación Oso de Asturias, Oviedo, Spain

Naves J, Wiegand T, Revilla E, Delibes M (2003) Endangered species constrained by natural and human factors: the case of brown bears in northern Spain. Conserv Biol 17:1276–1286

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Nsubuga AM, Robbins MM, Roeder AD et al (2004) Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. Mol Ecol 13:2089–2094

Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. Mol Ecol 3:489–495

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. Mol Ecol 4:347–354

Paetkau D, Waits LP, Clarkson P et al (1998) Variation in genetic diversity across the range of North American brown bears. Conserv Biol 12:418–429

Piggot M, Bellemain E, Taberlet P, Taylor AC (2004) A multiplex pre-amplification method that significantly improves microsatellite amplification and error rates for faecal DNA in limiting conditions. Conserv Genet 5:417–420

Pritchard JK, Stephens P, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Roeder AD, Archer FI, Poinar HN, Morin PA (2004) A novel method for collection and preservation of faeces for genetic studies. Mol Ecol Notes 4:761–764

Servheen C (1990) The status and conservation of the bears of the world. In: International conference of bear research and management monography series, vol. 2, pp. 1–32

Sherwin WB, Moritz C (2000) Managing and monitoring genetic erosion. In: Young AJ, Clarke G (eds) Demography genetics viability of fragmented populations. Cambridge University Press, New York

Simberloff D (1988) The contribution of population and community biology to conservation science. Annu Rev Ecol Syst 19:473–511

Smith DA, Ralls K, Hurt A et al (2006) Assessing reliability of microsatellite genotypes from kit fox faecal samples using genetic and GIS analyses. Mol Ecol 15:387–406

Sokal RR, Rohlf FJ (1994) Biometry: the principles and practice of statistics in biological research. WH Freeman, New York

Soulé ME (ed) (1987) Viable populations for conservation. Cambridge University Press, Cambridge, UK

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 255:195–200

Taberlet P, Luikart G (1999) Non-invasive genetic sampling and individual identification. Biol J Linn Soc 68:41–55

Taberlet P, Griffin S, Goossens B et al (1996) Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Res 26:3189–3194

Taberlet P, Camarra JJ, Griffin S et al (1997) Noninvasive genetic tracking of the endangered Pyrenean brown bear population. Mol Ecol 6:869–876

Taberlet P, Waits LP, Luikart G (1999) Noninvasive genetic sampling: look before you leap. Trends Ecol Evol 14:323–327

Taylor BL, Dizon AE (1999) First policy then science: why a management unit based solely on genetic criteria cannot work. Mol Ecol 8:S11–S16

Valière N (2002) GIMLET, a computer program for analyzing genetic individual identification data. Mol Ecol Notes 2:377–379

Vigilant L (1999) An evaluation of techniques for the extraction and amplification of DNA from naturally shed hairs. Biol Chem 380:1329–1331

Waits LP, Talbot S, Ward RH, Shields GF (1998) Mitochondrial DNA phylogeography of the North American brown bear and implications for conservation. Conserv Biol 12:408–417

Waits LP, Taberlet P, Swenson JE, Sandegren F (2000) Nuclear DNA microsatellite analysis of genetic diversity and gene flow in the Scandinavian brown bear (*Ursus arctos*). Mol Ecol 9:421–431

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. Mol Ecol 10:249–256

Weir BS (1996) Genetic data analysis II. Sinauer Associates Inc., Massachusetts

Wiegand T, Naves J, Stephan T, Fernández A (1998) Assessing the risk of extinction for the brown bear (*Ursus arctos*) in the Cordillera Cantabrica, Spain. Ecol Monogr 68:539–571

Woods JG, Paetkau D, Lewis D et al (1999) Genetic tagging of free-ranging black and brown bears. Wildl Soc B 27:616–627

Zedrosser A, Dahle B, Swenson JE, Gerstl N (2001) Status and management of the brown bear in Europe. Ursus 12:9–20