Conserv Genet
DOI 10.1007/s10592-010-0115-7

RESEARCH ARTICLE

Addressing challenges in non invasive capture-recapture based estimates of small populations: a pilot study on the Apennine brown bear

V. Gervasi · P. Ciucci · F. Davoli · J. Boulanger · L. Boitani · E. Randi

Received: 19 October 2009 / Accepted: 21 July 2010

Abstract It is often difficult to determine optimal sampling design for non-invasive genetic sampling, especially when dealing with rare or elusive species depleted of genetic diversity. To address this problem, we ran a hair-snag pilot study on the remnant Apennine brown bear population. We used occupancy models to estimate the performance of an improved field protocol, a meta-analysis approach to indirectly model capture probability, and simulations to evaluate the effect of genotyping errors on the accuracy of capture-recapture population estimates. In spring 2007 we collected 70 bear hair samples in 15 5 × 5 km cells, using 5 10-day trapping sessions. Bear detectability was higher in 2007 than in a previous attempt on the same population in 2004, reflecting improved field protocols and sampling design. However, individual capture probability was 0.136 (95% CI = 0.120–0.152), still below the minimum requirements of capture-mark-recapture closed population models. We genotyped hair samples ($n = 63$) at 9 microsatellite loci, obtaining 94% Polymerase Chain Reaction success, and 13 bear genotypes. Estimated P_{IDsib} was 0.00594, and per-genotype error rate was 0.13, corresponding to a 99% probability of correct individual identification. Simulation studies showed that the effect of non-corrected or filtered genetic errors on the accuracy of population estimates was negligible only when individual capture probability was >0.2. Our results underline how the interaction among field protocols, sampling strategies and genotyping errors may affect the accuracy of DNA-based estimates of small and genetically depleted populations, and warned us about the feasibility of a survey using only traditional hair-snag sampling. In this and similar cases, indications from pilot studies can provide cost-effective means to evaluate the efficiency of designed sampling and modelling procedures.

Keywords Apennine brown bear · Capture-mark-recapture · DNA sampling · Genotyping error · Occupancy models · Pilot study

V. Gervasi (✉) · P. Ciucci · L. Boitani
Dipartimento di Biologia Animale e dell'Uomo, Università "La Sapienza" di Roma, Viale dell'Università 32, 00185 Rome, Italy
e-mail: vincenzo.gervasi@uniroma1.it

F. Davoli · E. Randi
Istituto Nazionale per la Fauna Selvatica, Via Ca' Fornacetta 9, 40064 Ozzano Emilia, Bologna, Italy

J. Boulanger
Integrated Ecological Research, 924 Innes, Nelson, BC V1L 5T2, Canada

Introduction

The estimation of population size with non invasive genetic sampling and capture-mark-recapture (CMR) models is a complex multi-step process, that involves an appropriate sampling design (Boulanger et al. 2002; Lindberg and Rexstad 2002), effective field techniques, a reliable protocol of genetic individual identification (Paetkau 2003; Roon et al. 2005; Waits and Paetkau 2005), and adequate CMR modelling for population estimation. During each of these phases, errors and inefficiencies can occur, exerting a strong influence on the accuracy of the estimates. In fact, genotyping errors can cause up to 200% bias in population estimation (Roon et al. 2005; Waits and Leberg 2000), whereas sparse sample sizes, resulting from inadequate sampling intensity, can even prevent a reliable application of CMR models (Boulanger et al. 2002, 2004). Therefore,

Published online: 04 August 2010

an adequate planning of any attempt to estimate population size through non invasive genetics should start from the identification of study-specific requirements and assumptions of CMR models, and adopt a study design that fulfils them (Boulanger et al. 2002; Lindberg and Rexstad 2002). Moreover, high performance genotyping protocols must be defined a priori, to minimize the probability of individual misidentification during the genotyping process. Nevertheless, the main challenge is that often the expected sampling and genotyping efficiencies are unknown, which makes the a-priori optimization process problematic. To assess the expected performance of sampling and genotyping procedures for population estimation studies, simulations, meta-analyses, and pilot studies are generally considered as the most appropriate tools (Boulanger et al. 2002, 2004; Pompanon et al. 2005; Valiere et al. 2007). However, few studies have attempted to jointly address all the main issues related to sampling, genotyping, and modelling capture-recapture data, especially when dealing with rare or elusive species.

The Apennine brown bear lives in a small and isolated population (Ciucci and Boitani 2008; Gervasi et al. 2008), and is highly depleted of genetic variability (Lorenzini et al. 2004; Randi et al. 1994). In 2004, we produced a preliminary, hair-snag based estimate of 43 bears (95% CI = 35–67) for this population, but low capture probability and limited power to model heterogeneity affected precision and reliability of the population estimate (Gervasi et al. 2008). In addition, the low genetic variability of this population (Lorenzini et al. 2004; Randi et al. 1994) potentially reduces our power in distinguishing individual genotypes as observed in similar studies on brown bears (Paetkau 2004; Waits et al. 2001), thus requiring an increased number of microsatellite loci.

To address the above problems, we carried out a hair-snag pilot study in spring-summer 2007. Our aim was to address not only the specific problems at the sampling, genotyping, and modelling levels, but also those potentially arising from the interaction of these 3 co-dependent stages of a CMR-based population estimation.

Therefore, our work had the 3 following objectives:

1) to evaluate the expected efficiency of an improved field protocol and sampling design, with respect to the 2004 sampling protocol (Gervasi et al. 2008);
2) to indirectly estimate the expected individual capture probability provided by this design for our bear population;
3) to evaluate how the interaction between sampling performance and PCR error rates could affect our population estimates.

We believe the indications of our study extend beyond the particular species and population we studied, and are particularly relevant for small populations, whose detectability and genetic variability are expected to be low.

Materials and methods

Study area

The remnant distribution of the Apennine brown bear is mostly included inside the borders of the Abruzzo, Lazio and Molise National Park (PNALM, 41°50′N; 13°54′E, Italy) and its outer buffering area, which encompass 1,294 km^2 (Ciucci and Boitani 2008; Falcucci et al. 2009). The study area hosts a largely intact mountain forest ecosystem, with elevations ranging from 1,000 to 2,200 m a.s.l. Deciduous forests, mainly dominated by beech (*Fagus sylvatica*) and oak (*Quercus* spp.), cover mid-elevations, whereas almost 30% of the study area is above timberline, with bare rocks and high elevation pastures. Valley bottoms are characterized by a mixture of small human settlements, an extensive network of asphalt and gravel roads, fragmented woodland, and pastures.

Sampling design

We referenced our previous 2004 sampling when designing the 2007 pilot study, to allow comparison between the sampling efficiencies of the 2 studies (Gervasi et al. 2008). We therefore defined 3 sampling strata according to the 2004 sampling success over the entire 44 5 × 5 km cells sampling grid (no. of bear samples/cell): 21 cells (no bear samples in 2004) were included in stratum 1; 13 cells (1–3 bear samples in 2004) were included in stratum 2; 10 cells (>3 bear samples in 2004), were included in stratum 3. We then randomly subsampled 15 cells from the entire grid, proportionally allocating them to the number of cells within each stratum. In accordance with other hair-snag studies (Kendall et al. 2008; Mowat and Strobek 2000; Woods et al. 1999), in 2007 we collected hair samples in spring and summer. This was different from the 2004 survey, which was conducted during fall (Gervasi et al. 2008).

Field methods

In 2007 we more effectively selected hair-trap locations by favouring sites close to topographical constrictions and to edges between forested and open habitats, where bear movements are more likely to be intercepted. In addition, we used a recently developed GIS-based bear occurrence model (Falcucci et al. 2009) to guide the selection of trap locations, in habitat-types with the highest probability of occurrence. For the 2004 survey we more opportunistically

selected trap-sites, based on their accessibility and proximity to roads (Gervasi et al. 2008; Potena et al. 2004).

We built hair traps following standard procedures for brown bear hair snag sampling (Woods et al. 1999; Mowat and Strobek 2000; Boulanger et al. 2004). We sampled each cell for 5 consecutive 10-day sessions, moving hair traps within cells at the end of each session. We baited traps using 2–3 l of beef blood, rotted for 3 months in closed barrels, mixed with an equal proportion of rancid fish oil, which we poured on the central mound. Silica desiccant (10 g) was used to preserve each sample. Sampling methods for 2004 differed slightly in that bait only consisted of fish oil and hair samples were collected in 25 ml of 90% ethanol.

Genetic analyses

We extracted DNA in a dedicated room within a sterile UV hood, using a guanidine thiocyanate/silica protocol (Gerloff et al. 1995), and used the Amelogenin AMG gene (Ennis and Gallagher 1994) for sex identification. We obtained individual genotypes through replicated PCRs at nine autosomal microsatellite loci: MU50, MU15, G10C, MU51, MU05, MU11, MU59, G10B, G10L (Paetkau et al. 1995; Taberlet et al. 1997). We followed a multiple-tube procedure: (i) a screening step, designed to identify and remove low quality DNA samples, in which all samples were amplified 4 times at 3 loci (MU15, MU50 and AMG); (ii) all samples, which produced a detectable amplification of at least one "true" brown bear allele at the typed locus in >50% of PCRs, were further processed at the remaining 7 loci (G10C, MU51, MU05, MU11, MU59, G10B, G10L), thus obtaining 4 PCR replicates per locus; (iii) a reliability score R for each multilocus genotype was computed with RELIOTYPE (Miller et al. 2002); (iv) reliable genotypes ($R > 0.95$) were used to identify the consensus genotypes, using GIMLET V.133 (Valière 2002) with the threshold method: an allele was retained in consensus genotype if it was scored at least 2 times. In a post-process control of the genetic database, we computed the pairwise expected and observed mismatch (MM) values, using the software MM-DIST (Kalinowski et al. 2006), and identified and purged the laboratory record of all genotype pairs differing at only 1 (1-MM) or 2 alleles (2-MM). Each locus was amplified separately in a volume of 8 µl containing 4 µl of template solution (*c*. 20 DNA nanograms), 1.5 mM $MgCl_2$, 0.8 µl of 0.2% BSA, 0.225 µl of each 10 µM primer solution, 0.36 µl of 10 mM dNTPs, 0.2 units of *Taq* polymerase (Eppendorf), and the following thermal profiles: (1) 94°C for 2 min; 55 cycles at 94°C for 15 s, 52.5°C for 30 s and 72°C for 30 s; 72°C for 10 min (for microsatellites); (2) 94°C for 2 min; 55 cycles at 94°C for 15 s, 57.5°C for 30 s and 72°C for 30 s; 72°C for 10 min (for AMG). All PCRs were done under a sterile airflow hood cleaned with UV light. Negative controls were used during each step from DNA extractions to PCR. Positive controls (samples with known genotypes) were added to each PCR session. The same genotyping protocol and error-checking protocols were used in the 2 studies. Fragments were analysed using an ABI 3130 XL automated sequencer and GENEMAPPER V.3.0 (Applied Biosystems). In addition to the procedures described above, and as recommended by Paetkau (2003) and Roon et al. (2005), we replicated the genotyping process for all 1-MM and 2-MM pairs, and for all genotypes identified though a single hair sample. In addition, we checked if genetic sampling locations of 6 adult bears, fitted with a GPS-collar during the sampling period, were consistent with their observed home ranges and movements, and reanalysed all inconsistent samples. Moreover, we blindly compared the genotypes obtained from blood samples of all bears livetrapped during the study period with hair samples collected on the same bears during immobilization.

We estimated per genotype allelic dropout (ADO) rates as the ratio between the number of single-locus genotypes including at least one allelic mismatch, and the number of replicated single-locus genotypes (Pompanon et al. 2005), whereas false alleles (FA) were estimated as follows: (1) alleles whose molecular weight was outside the range of the reference invasive genotypes; (2) alleles detected only once in at least 7 replicated PCRs (Taberlet et al. 1996).

Estimation of detection probability

The first objective of our pilot study was to test for whether the 2007 design improved sampling efficiency over that used in 2004. For this purpose, we used cells (hereafter sites) as sample units, and applied single season occupancy models (McKenzie et al. 2002) using the Program PRESENCE (Hines 2006) to test for a difference in the average level of detection probability between 2004 and 2007. This approach does not yield an estimate of the average probability of each bear in the study area to be genetically sampled (i.e., capture probability), but rather of the probability of detecting bear presence in each site, given that at least one bear was present in the site (i.e. detection probability). As we could not directly estimate capture probability based on the small 2007 pilot study hair sample (see below), we used the estimated detection probability as an alternative measure of the relative efficiency of the 2007 vs. the 2004 sampling effort. As our sampling effort did not last longer than 90 days in either 2004 or 2007, constant occupancy by bears during each sampling period seemed a reasonable assumption (cf Boulanger et al. 2008b), where occupancy probability is defined as the probability that a given site is occupied

by at least one individual of the study species during the study period (McKenzie et al. 2002). In addition to comparing the 2004 vs. the 2007 sampling success (i.e., detection probability), we tested the hypothesis of no change in the occupancy probability between the 2 years, by comparing models that included and excluded a year effect on the occupancy probability. We therefore included the year of sampling as a group covariate (2004 vs. 2007), to test for its effect both on the occupancy and detection probabilities, whereas the distance of each site from the borders of the sampling grid was used as an individual covariate in modelling the occupancy probability of bears. Moreover, to explore the potential effect of a set of spatial variables on the detection probability of bears, we also measured for each trap the distances from forest edge, closest human settlement, closest gravel or asphalt road, and the altitude.

We evaluated model support using the sample size corrected Akaike Information Criterion (AICc) of model fit, and considered the model with the lowest value of the AICc to be the most parsimonious (Burnham and Anderson 2002). To account for the degree of uncertainty in model selection, we averaged the estimates of parameters from different models, using the Akaike weights (Burnham and Anderson 2002) as an index of their relative support.

Estimation of capture probability

We used detection probability as a measure of relative efficiency of the 2007 vs. the 2004 sampling strategy. However, it is individual capture probability and its heterogeneity that directly affect the performance of closed CMR models (Boulanger et al. 2004; White et al. 1982). Due to sub-sampling of the study area in 2007, the consequent violation of geographic closure, and the small sample obtained, a direct CMR estimation of capture probability was not feasible (Boulanger and McLellan 2001; White and Shenk 2001). Therefore, we used an indirect approach to estimate the individual capture probability expected under our field conditions and sampling efficiency. Using individual capture probabilities reported in 14 hair-snag studies both in Europe and in North America (Table 1), we estimated, through a linear regression model, the relationship between capture probability in each of these bear populations and a corresponding measure of catch per unit effort (CPUE, Romain-Bondi et al. 2004), that is the number of bear captures per 1,000 trap nights. Because CPUE is influenced not only by the average capture probability, but also by the density of bears on the sampling grid, we standardized the CPUE index according to the estimated bear density in each study area. The resulting regression function was then used to estimate the expected average level of individual capture probability in our bear population for the 2007 sampling design.

Simulations of PCR error rates

To test the joint effect of sampling performances and undiscovered genotyping errors on the accuracy of CMR based population estimates, we performed a set of simulations using GEMINI 1.3.0 (Valiére et al. 2002). We sampled for 5 sessions a virtual population of 50 bears, which was

Table 1 Summary statistics of 14 hair-snag projects in Europe and North America

Project	No. traps	Trap nights	Bear captures	CPUE[a]	Density (bears/100 km^2)	Stand. CPUE[b]	$\hat{P}$
British Columbia UCR 1996 (Woods et al. 1999)	256	2,653	73	27.516	2.6	10.436	0.16
British Columbia (Mowat and Strobek 2000)	381	4,245	134	31.56	2.6	11.88	0.1
Alberta (Mowat and Strobek 2000)	321	4,494	46	10.23	1.4	6.95	0.13
British Columbia (Poole et al. 2001)	515	6,180	142	22.97	1.7	13.23	0.19
South Selkirks	408	4,080	45	11.02	1.9	5.80	0.09
North Selkirks	362	3,620	91	25.13	4.8	5.23	0.08
Prophet Plateau	334	4,008	42	10.479	1.0	10.47	0.13
Prophet mountains	192	2,304	100	43.40	3.1	14.00	0.21
Yellowhead	198	2,772	51	18.39	1.7	10.27	0.16
Parnsip plateau	188	2,632	24	9.11	1.0	5.49	0.12
Parnsip mountains	402	5,628	299	53.12	4.3	12.35	0.22
Great Glacier 1998 (Kendall et al. 2008)	626	8,764	222	25.33	3.2	7.68	0.12
Great Glacier 2000 (Kendall et al. 2008)	633	8,862	234	26.40	3.3	7.92	0.12
Abruzzo 2004 (Gervasi et al. 2008)	219	2,628	12	4.56	2.9	1.55	0.03

We used these studies to derive a regression function between CPUE and the average level of the individual capture probability

[a] CPUE is the number of bear captures obtained per 1,000 trap nights

[b] The standardized CPUE is the ratio between CPUE and the estimated bear density of each study area

deemed as a reasonable approximation of the actual population size, based on a previous survey (Gervasi et al. 2008). To simulate different levels of capture probability, we set an increasing number of collected samples (i.e., 25, 50, 75 and 100). Individual variation in capture probability was first defined by an intercept-only model, corresponding to a uniform distribution. Then, to introduce the effect of heterogeneity in capture probabilities, we simulated 2 mixtures inside the population, one with high and one with low capture probability, with a CV of capture probability = 0.3 (Carothers 1973). Allele frequencies for the 9 loci were derived from the genotype analysis of tissue and blood samples of 11 live-captured or dead bears from our population (Randi, pers. comm.). The overall PCR per-genotype error rate (ER; Broquet and Petit 2004) was set to 5.5%, as estimated by the analysis of 821 hair samples collected during the period 2004–2005 (Randi et al. 2006). Nevertheless, because the degree of genetic misidentification in previously analyzed samples can produce an underestimation of the actual frequency of occurrence of ADO and FA (Pompanon et al. 2005), we conservatively ran the same set of simulations with doubled (ER = 11%) and tripled (ER = 16.5%) error rates. Being applied in the context of a multiple-tube approach, the error rates we used for the simulations are clearly meant to be applied before the application of consensus genotype rules (Pompanon et al. 2005). That is, they do not necessarily reflect the true final percentage of expected individual misidentifications, thereby representing a conservative scenario. We simulated the genotyping lab process with 4 PCR repetitions per sample, and a consensus threshold of 2, and measured the resulting proportion of wrong genetic identifications, which corresponded to the actual post-process, residual error rate. As a last step, we used likelihood-based closed population mark-recapture models (Pledger 2000), implemented in the Program MARK (White and Burnham 1999), to produce population estimates from the simulated encounter histories. This type of model separates the sample of genotyped individuals into a finite number of mixtures, based on their differential individual capture probability, thus allowing to model its variance, and to produce unbiased estimates of population size when individual heterogeneity is present in the data (Pledger 2000). We then compared these error-prone estimates, in terms of accuracy and confidence interval coverage, with error-free reference estimates generated using the same population parameters. We performed 1,000 iterations for each set of simulated parameters. Because we extracted samples from the virtual population with replacement, allowing multiple samples of the same genotype to be extracted during one session, the number of collected samples was not proportional to the individual capture probability. Therefore, we averaged the 1,000 capture probability estimates, resulting from MARK closed capture analysis of each set of simulated parameters, thus obtaining a more reliable indication of capture probability.

Results

We started hair sampling on 5 May and completed it by 28 July 2007, totally activating 70 hair traps. Eighteen (25%) hair traps proved positive, and provided 91 hair samples (Table 2), 21 of which were macroscopically attributed to non target species. Genetic analyses were therefore attempted on all remaining 70 bear samples.

Overall costs for the pilot study were less than € 15,000, and projection of both field and laboratory costs, had we performed the survey on the whole study area (Table 3), would have amounted to € 23,173 and € 22,000, respectively, for a total of € 45,173.

Genetic analyses

After the screening step, we selected 63 out of the 70 samples that showed >50% of positive PCRs (PCR+) at the 3 screening loci. The selected hair samples were genotyped at the additional 7 loci, providing an overall PCR success rate of 0.94 (SE = 0.01). Error rates ranged from

Table 2 Summary statistics of the hair-snag pilot study, conducted in spring-summer 2007 on the Apennine brown bear population in the Abruzzo, Lazio and Molise National Park (Central Italy)

Session	Period[a]	No. traps	No. traps with bear samples	No. collected samples				Bear samples/trap
				Bear	Cervid	Canid	Horse	
1	5 May–14 May	15	4	13	1	5	0	0.90
2	15 May–24 May	15	3	16	0	0	5	1.06
3	25 May–09 June	14	3	7	0	0	0	0.35
4	11 June–27 June	15	4	7	0	3	6	0.46
5	01 July–17 July	11	4	27	0	1	0	2.45
Total		70	18	70	1	9	11	1.00

[a] Date intervals refer to trap activation period, where the last day of each session corresponds to the date of activation of the last trap

Table 3 Comparison of sampling statistics between the 2007 vs. the 2004 hair-snagging effort on the Apennine brown bear population in the National Park of Abruzzo Lazio and Molise (Central Italy)

	2004	2007
Sampling period	4 Oct.–25 Nov.	5 May–28 July
No. cells sampled	44	15
No. traps	219	70
No. positive traps (%)	13 (6%)	18 (25%)
No. bear samples	24	70
Bear samples/trap (mean ± SD)	0.11 ± 0.63	1.00 ± 2.39
Bear genotypes (FF:MM)	11 (10:1)	13 (8:5)
Detection probabilities[a] (±95% CI)	0.099 (0.057–0.173)	0.366 (0.325–0.496)
Field costs (€)	23,000	7,900
Lab costs (at €100/sample)	2,400	7,000
Total cost/sample (€)	1,058	212

[a] Model averaged estimates from occupancy analysis (Table 4)

FA = 0.01 (SE = 0.002) to ADO = 0.12 (SE = 0.004), for an overall pre-checking per-genotype error rate (ER) of 0.13 (SE = 0.006). We obtained reliability scores $R > 0.95$ in 92% (58/63) of the samples, which provided 13 different genotypes (8 females and 5 males) (Table 3). Eight bears were sampled in one session, 4 in two sessions, and 1 in 3 sessions, for a total of 19 bear captures.

No 1-MM or 2-MM pairs were detected among the revealed bear genotypes, and all hair samples attributed to radiomarked bears ($N = 27$) fell inside the individual, GPS-based 95% Minimum Convex Polygon home ranges. In addition, invasively detected genotypes (blood samples) of 3 live-trapped male bears matched those blindly derived from 6 hair samples (2 for each bear) collected during handling. The observed and expected heterozygosity of the 2007 genetic sample was Ho = 0.46 and He = 0.44; the sample was in Hardy–Weinberg equilibrium, and the probabilities of identity were $P_{ID} = 0.000035$, and $P_{IDsib} = 0.00594$, respectively (Fig. 1).

Estimation of detection probability

The occupancy analysis indicated a marginal year effect on the variation of occupancy probability (Table 4), thus showing that no relevant changes in bear occupancy had occurred in the study area between 2004 and 2007. Detection probabilities were mainly influenced by the year of sampling (2007 vs. 2004), as this term was included as a main effect in all most supported models (Table 4). Model averaged estimates of detection probabilities were 0.099 (95% CI = 0.057–0.173) and 0.366 (95% CI = 0.325–0.496) in 2004 and 2007, respectively (Table 3). Under constant occupancy conditions, this confirms the higher efficiency of the improved sampling protocol in 2007.

Moreover, we detected an effect of hair trap distance from the forest edge on detection probability, with higher sampling probability for traps placed in ecotones (Fig. 2). Location of hair traps with reference to altitude, distance

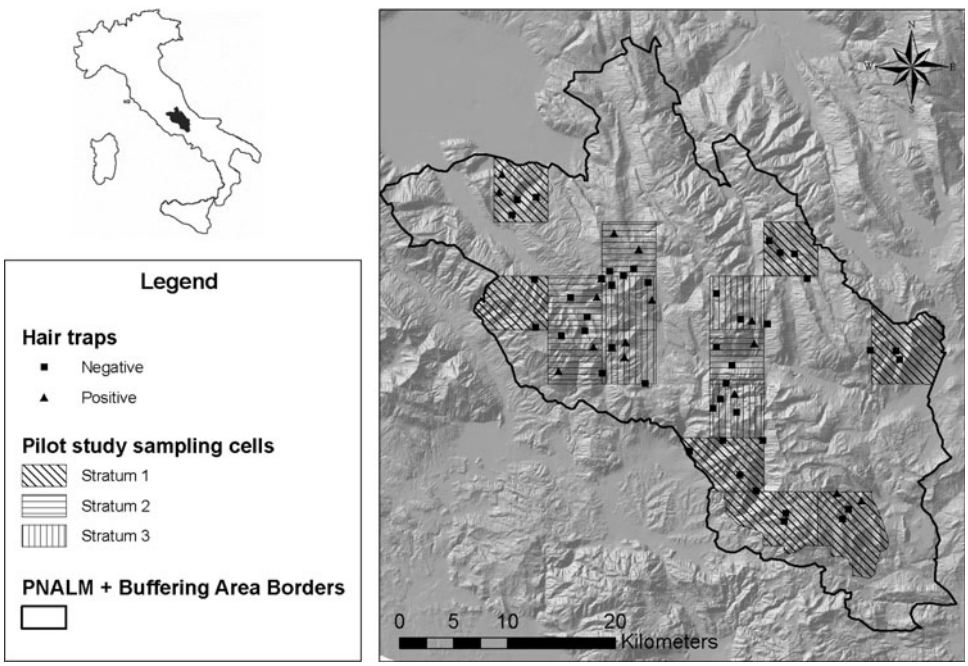

Fig. 1 Distribution of the 15 grid cells (5 × 5 km) sampled for a hair-snag pilot study to assess feasibility of a DNA-based capture-mark-recapture estimation of the Apennine brown bear population size (National Park of Abruzzo Lazio and Molise, May–July 2007). Cells were sampled according to 3 strata based on the sampling success of a hair-snag survey carried out in 2004 (Gervasi et al. 2008)

Table 4 AIC model selection for occupancy estimation models, applied to the genetic sampling of the Apennine brown bear population in 2004 and 2007

Model No.	Detection probability	Occupancy	AIC_c	No. parameters	ΔAIC_c	w_i
1	Year + forest edge	Grid edge	143.43	5	0.00	0.868
2	Year + forest edge	Constant	148.61	4	5.18	0.065
3	Year + forest edge	Year	150.05	5	6.62	0.031
4	Year	Grid edge	151.99	4	8.56	0.012
5	Year + road	Grid edge	153.01	5	9.58	0.007
6	Year + antro	Grid edge	153.43	5	10.00	0.005
7	Year + altitude	Grid edge	153.94	5	10.51	0.004
8	Forest edge	Grid edge	154.11	4	10.68	0.004
9	Year	Constant	157.79	3	14.36	0.000
10	Year	Year	159.21	4	15.78	0.000
11	Forest edge	Constant	159.34	3	15.91	0.000
12	Road	Constant	163.69	3	20.26	0.000
13	Antro	Constant	168.97	3	25.54	0.000
14	Constant	Constant	169.96	2	26.53	0.000
15	Altitude	Constant	170.43	3	27.00	0.000

Covariates included: year (year of sampling); the linear distance between the geometric centre of each cell and the border of the sampling grid (grid edge); linear distance of each hair trap from closest edge between a forested and open area (forest edge); linear distance of each hair trap from the closest gravel or paved road (road); linear distance between each trap and the closest human settlement (antro); altitude (m a.s.l.) of each trap

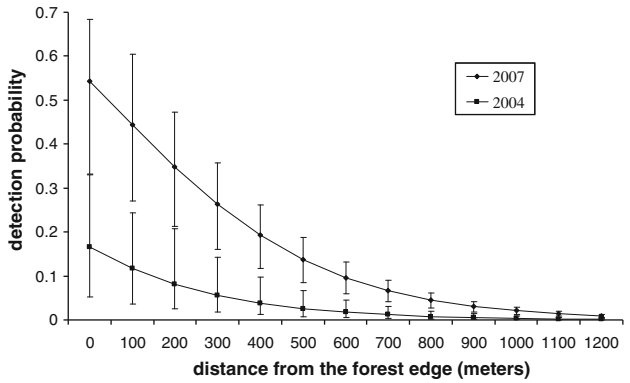

Fig. 2 Detection probability as a function of hair trap distance from forest edge for both the 2004 survey and the 2007 pilot study. The dependent variable is the probability that bear presence is detected inside a sampling cell, given that the cell is actually occupied throughout the sampling period by at least one bear. Estimates obtained from the most supported occupancy model (model 1 in Table 4). *Error bars* are 95% confidence intervals of the estimates

from roads, and distance from human settlements had a marginal effect on variation in the detection probability (Table 4).

Estimation of capture probability

For the CPUE analysis ($N = 14$ hair-snag studies), the intercept term in the regression function between CPUE and capture probability was not significantly different from 0 ($t = 0.674$, $p = 0.511$). The linear regression function through the origin ($y = 0.0147x$) explained most ($R^2 = 0.886$) of the variation in the average level of the individual capture probability. Using this model, and the CPUE index from our 2007 sampling effort, we estimated a corresponding expected, average capture probability of our hair-snag sampling design of 0.136 (95% CI = 0.120–0.152).

Simulations of PCR error rates

With 4 PCR amplifications, identification success rates were above 98% for all the 3 simulated error rates. The average number of new false individuals generated per 100 genotyped samples ranged from 0.12 (ER = 5.5%) to 1.86 (ER = 16.5%), which corresponds to the post-process final probability of individual misidentification.

As from simulations of CMR model performances (Fig. 3), population estimates based on the observed PCR error rate (ER = 5.5%) showed a 13.9% positive bias for low values of the average capture probability ($\widehat{P} = 0.10$). However, simulations with 0 error rate displayed a bias of 10.4%, showing that only 3.5% of bias at this level of capture probability was actually resulting from the effect of genetic errors, whereas it was mainly a consequence of the low performance of the Pledger model at low levels of capture probability (Pledger 2005). Nevertheless, higher genetic error rates caused a substantial decrease of accuracy in population estimates, with genetic-based bias ranging from 5.0% (ER = 11.0%) to 8.8% (ER = 16.5%).

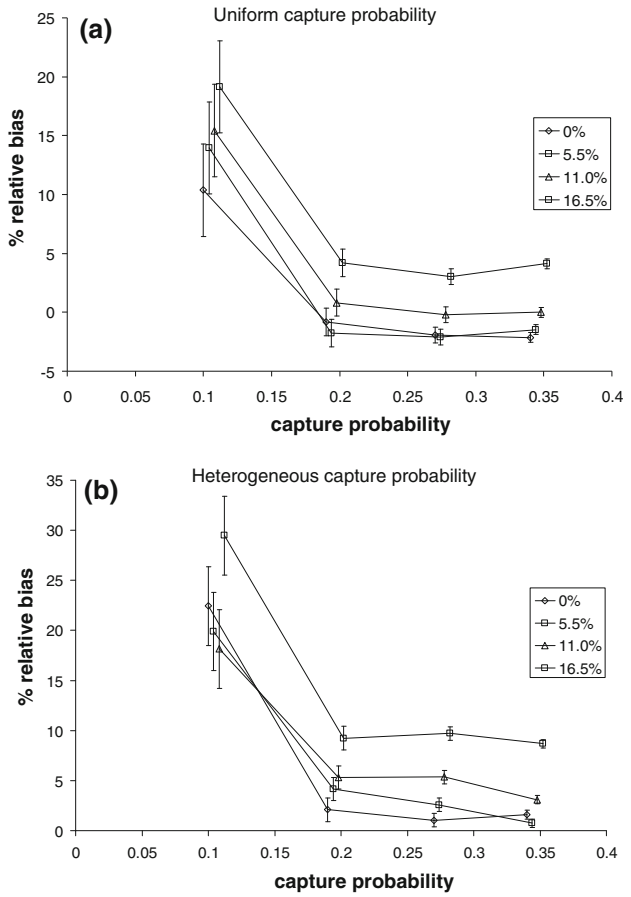

Fig. 3 Simulations of the effect of genotyping error rates on population size estimates based on closed CMR models (program MARK). With reference to a theoretical population of 50 bears, we simulated an increasing number of collected samples (25, 50, 75, and 100) at average rates of genotyping errors from 0 to 16.5%. Individual encounter histories were generated, using program GEMINI, by an intercept model (Fig. 3a), or by a 2 mixture model (Fig. 3b). *Error bars* are 95% confidence intervals of the estimates. We removed all population estimates with bias >100%, to prevent a few outliers affecting the average estimate of population size accuracy

Increasing error rates did not greatly affect CI coverage, which was between 91.2 and 95.8% for all simulations. When capture probability increased to $\widehat{P} = 0.19$, the same 5.5% PCR error rate resulted in only 1.7% biased population estimates (Fig. 3), thus revealing that a better performing sampling protocol can dramatically improve the accuracy of population estimates by both mitigating the effects of genotyping errors, and ensuring adequate performance of the Pledger model. This was also confirmed using the most extreme level of simulated error rates (ER = 16.5%), providing estimates of only 3.0% positive bias at a high average capture probability ($\widehat{P} = 0.27$).

Under the heterogeneity model, the effect of genetic errors on the accuracy of population estimates was higher, but again was less pronounced at increasing values of the individual capture probabilities (Fig. 3).

Discussion

Demographic studies based on non invasive genetic sampling often fail due to inappropriate study designs, resulting in poor sampling performance (Boulanger et al. 2002), or due to inadequate genotyping protocol, which reduces the power of correctly discriminating individual animals (Mills et al. 2000; Waits and Leberg 2000; Waits and Paetkau 2005). Accordingly, past applications of hair-snag sampling to the Apennine brown bear population (Potena et al. 2004) suffered from poor sampling performance, which prevented this technique to be satisfactorily used as a stand-alone tool for population estimation (cf. Gervasi et al. 2008). Our pilot study aimed to address the problems above at all relevant stages, from field protocols, to sampling strategies, to genotyping protocols, to modelling procedures. By using single-season occupancy models (McKenzie et al. 2002) we were first able to verify that the improvements adopted in 2007 sampling design and field techniques had indeed a positive effect on our ability to non invasively sample bears from this small population, and to increase individual capture probability in future hair-snag applications. In this perspective, our results demonstrate the added value of occupancy models not only in evaluating species distribution patterns (McKenzie et al. 2002), but also as a powerful tool to design grid-based population sampling for a variety of taxonomic groups, especially when detection probability needs to be maximized for analytical purposes. In fact, occupancy models are originally meant to estimate species occupancy rates (e.g., Kroll et al. 2007; Mortelliti and Boitani 2007), and detection probability is therefore nothing more than a nuisance parameter. However, they can also be applied to test hypotheses concerning the main factors influencing detection probability, as recently done for bears (Boulanger et al. 2008b) and other species as well (*Pecari tajatu*: Longoria and Weckerly 2007; butterflies: Pellet 2008). Although we were able to estimate the probability to detect bear presence at each sampling cell, this did not provide any quantitative information about capture probability of each individual bear in the population, which is the key parameter influencing the performance of CMR closed capture models (Huggins 1991; Pledger 2000). For this reason, we used a meta-analysis approach, based on an ad hoc developed regression function, to indirectly estimate the individual capture probability expected under a systematic hair snag sampling strategy. We acknowledge that this approach is not a substitute for the direct estimation of capture probability using CMR models and that, as such, it suffers from theoretical limitations. First, our capture probability regression model does not provide any indication concerning extent and causes of heterogeneity of capture probability in our bear population, a factor exerting

a strong influence on the performance of CMR models (Boulanger et al. 2004; White et al. 1982). Second, we used the CMR-based density estimates to standardize the CPUE index, as independently derived estimates of bear density for each study were not available to develop the regression model. This likely resulted in a partial correlation between bear density (predictor variable) and capture probability (the response variable), with a likely overestimation of model support. Thirdly, sample sizes were insufficient to consider the effect of other design parameters such as cell size, closure violation, and seasonality on capture probability estimates. We nevertheless believe this approach is robust enough for our general aim and the needed resolution of our analysis. It could be considered an operational alternative to a direct CMR estimation of capture probability, in cases where the objective of the study is not to directly estimate population size (cf. Romain-Bondi et al. 2004), or when a pilot study, as in our case, is used as a tool to assess the feasibility of population-wide CMR modelling studies. Although the individual capture probability we indirectly estimated through the regression function ($\widehat{P} = 0.136$) is much higher than the one estimated for the Apennine bear population in the 2004 hair snag survey ($\widehat{P} = 0.03$; Gervasi et al. 2008), it is still below the minimum requirement, needed to produce reliable estimates for small populations ($\widehat{P} > 0.2$; Boulanger et al. 2002; White et al. 1982). Therefore, even though our 2007 hair snag pilot study showed an improved efficiency over previous efforts, the enhanced sampling success in this bear population will not necessarily yield levels of capture probabilities for reliable CMR applications, using hair snag data alone. It should be noted that other long-term brown bear genetic sampling projects showed a progressive increase of capture probability over the years (Boulanger et al. 2002, 2006), suggesting that enhanced expertise of field personnel through experience is a key factor contributing to any further increase of trapping success. Nevertheless, our findings caution us about the feasibility of the systematic hair-snagging survey as a stand-alone technique to estimate the size of this small bear population. In our case, the integration of multiple data sources into a global CMR framework (Boulanger et al. 2008a; Kendall et al. 2008) represents a promising opportunity. For example, along with systematic hair-snagging, rub-tree (Kendall et al. 2008, 2009), scats (Solberg et al. 2006), phototraps (Karanth et al. 2006), and other non invasive sampling techniques could be all used not only simultaneously, but also according to different and complementary sampling strategies (Gervasi et al. 2008).

Besides the implementation of an efficient sampling technique, the application of a reliable lab protocol for genetic identification of sampled individuals is a fundamental requirement in the multi-step process of a CMR based population estimation (Paetkau 2003; Waits and Paetkau 2005). In our case, the quality of DNA samples collected from hair-snag traps was usually good enough to allow for reliable laboratory performances, as only 10% of the samples were discarded and 92% of the remaining ones successfully led to the identification of individual genotypes. Dealing with a population depleted of genetic diversity, our main concern was the power to correctly distinguish all the individuals in the population (Paetkau 2004; Waits et al. 2001). However, the P_{IDsib} estimated on the set of 9 polymorphic loci was well below the threshold of 0.01, theoretically required to prevent the risk of shadow effect (Mills et al. 2000). In addition, the maximum number of distinguishable genotypes was $1/P_{IDsib} = 1/0.00594$, or 168 bears, which is much more than the expected population size (Gervasi et al. 2008). Therefore, considering both the reduced genetic variability of the Apennine brown bear population and its expected size, the set of 9 loci that we used plausibly provided enough power to distinguish individuals. It should be noted that we used a higher number of polymorphic loci than other similar applications on brown bears, usually using from 6 (Bellemain et al. 2005; Mowat and Strobek 2000; Paetkau 2003) to 7 (Kendall et al. 2008, 2009) loci, and therefore we increased the risk of occurrence of genotyping errors (Paetkau 2003). Nevertheless, our P_{IDsib} values indicate that this was an inevitable choice to reliably identify individual bears in our population. A higher and unacceptable risk of shadow effect would in fact have been introduced by using a lower number of loci (i.e., $P_{IDsib} = 0.0344$ and $P_{IDsib} = 0.0209$ with 6 and 7 loci, respectively).

We experienced similar PCR success rates during 2004 (83%) and 2007 (84%) genetic sampling projects, but the initial ADO and FA rates we detected during 2007 were higher than in previous non invasive genetic applications on the same bear population. Differences in genotyping error rates can be related to several factors, from the amount of DNA available (Taberlet et al. 1996), to the storage system adopted (Murphy et al. 2002; Roon et al. 2003). Both season of sampling and storage system differed between non invasive surveys of 2004 and 2007, probably causing a different quality of DNA templates. However, there is no evidence of a seasonal effect on DNA quality of bear hair samples, nor has any storage method been proven to be most efficient in preserving DNA for individual identification (Roon et al. 2003; Waits and Paetkau 2005). Therefore, we suspect that the differences in ADO and FA rates we observed are due more to the amount of DNA available than to other processing factors. In fact, the majority of hair samples, used in previous genetic studies on this population, provided on average a higher number of bear hairs per sample, because they were opportunistically and more frequently collected at

buckthorn (*Rhamnus alpina*) aggregations (Gervasi et al. 2008). In any case, given the observed ADO and FA rates, our simulations confirmed that the number of PCR repetitions was still sufficient to correctly identify more than 98% of samples, thus producing an estimate of only one false new individual, after 147 correctly genotyped samples.

The simulated effect of residual genetic misidentifications on the accuracy of CMR based population estimates was rather low, especially when average individual capture probability was >0.2. However, the same PCR error rate caused a relevant bias in population estimates, when associated with poor sampling performance (i.e., $0.1 \leq \widehat{P} \leq 0.2$), showing that there is a direct and strong relationship between the efficiency of sampling protocols, the reliability of genotyping procedures, and the performance of CMR models for population estimation. Likewise, other authors (Roon et al. 2005; Waits and Leberg 2000) concluded that when a higher number of samples was collected in each session, residual genetic errors have less relevant negative effects on the accuracy of population estimates. However, in contrast to our findings, the same authors reported a negative bias in population estimates for sparse data. The likelihood-based estimators we used in our simulations are believed to provide better performance in modeling heterogeneity when sample size is small (Norris and Pollock 1996; Pledger 2000), and they have a broad applicability in current CMR-based population estimation projects. In contrast, previous simulations used a former M_0 estimator implemented in program CAPTURE (Waits and Leberg 2000), or a jackknife estimator (Roon et al. 2005), the latter having an inherently low bias when data are sparse (Chao 1987, 1989). However, our simulations showed that the Pledger model can also produce biased results when capture-recapture data are small and sparse, and in these conditions it might under- or overestimate population size, depending on the distribution of capture probabilities in the population (Pledger 2005). Accordingly, at low levels of capture probability, high bias in population estimates was only partially due to the effect of residual genetic misidentification, as a relevant role was played by instability of the Pledger model. In any case, our simulations further cautioned us on the effective feasibility of a large-scale systematic hair-snag survey on our bear population, without further enhancements of sampling performance, not only because capture probability might be below the theoretical requirements of CMR models, but also because it might negatively interact with residual genetic error rates, thus further inflating bias of the population estimate. One potential method to offset the issue of genotyping error would be to use the models of Lukacs and Burnham (2005). In this case genotyping error rate would be entered into the model directly from lab estimates rather than estimated from the mark-recapture data.

Indications and implications of this study are not limited to the special case of the Apennine brown bear population. We believe that when DNA-based CMR population estimates are needed, a pilot study coupled with other analytical tools can provide a cost-effective way to a-priori evaluate the feasibility of the intended sampling design for population estimation. By using this approach, we saved about 67% of the costs which would have been otherwise involved in a population-wide systematic hair-snag sampling of the Apennine bear population of doubtful reliability and, accordingly, we realized that alternative methods should be contemplated.

Acknowledgements We are grateful to the Scientific Service of the Abruzzo, Lazio and Molise National Park, and in particular to C. Sulli and R. Latini, for their contribution in planning and organizing field activities during 2007 sampling season. We thank the many park wardens of the PNALM for sharing with us the field work. Other technicians and students contributed to the field work, among which R. Bucci, G. Di Domenico, L. Grottoli, G. Manno, B. Molnar, D. Pagliaroli, F. Quattrociocchi. We are particularly indebted to G. Potena, L. Sammarone, and M. Posillico of the Forest Service (Ufficio Territoriale per la Biodiversità di Castel di Sangro) for their logistic support. Thanks to Stefano Focardi, and to two anonymous referees, who provided insightful comments and ideas on a previous draft of this paper.

References

Bellemain E, Swenson JE, Tallmon D, Brunberg S, Taberlet P (2005) Estimating population size of elusive animals with DNA from hunter-collected faeces: four methods for brown bears. Conserv Biol 19:150–161

Boulanger J, McLellan BN (2001) Closure violation in DNA-based mark-recapture estimation of grizzly bear populations. Can J Zool 79:642–651

Boulanger J, White GC, McLellan BN et al (2002) A meta-analysis of grizzly bear DNA mark-recapture projects in British Columbia, Canada. Ursus 13:137–152

Boulanger J, McLellan BN, Woods JG, Proctor MF, Strobek C (2004) Sampling design and bias in DNA-based capture-mark-recapture population and density estimates of grizzly bears. J Wildl Manag 68:457–469

Boulanger J, Proctor M, Himmer S, Stenhouse G, Paetkau D, Cranston J (2006) An empirical test of DNA mark-recapture sampling strategies for grizzly bears. Ursus 17:149–158

Boulanger J, Kendall KC, Stetz JB et al (2008a) Use of multiple data sources to improve DNA-based mark-recapture population estimates of grizzly bears. Ecol Appl 18:577–589

Boulanger J, White GC, Proctor M et al (2008b) Use of occupancy models to estimate the influence of previous live captures on DNA-based detection probabilities of grizzly bears. J Wildl Manag 72:589–595

Broquet T, Petit E (2004) Quantifying genotyping errors in noninvasive population genetics. Mol Ecol 13:3601–3608

Burnham KP, Anderson DR (2002) Model selection and multimodel inference—a practical information—theoretic approach, 2nd edn. Springer-Verlag, New York

Carothers AD (1973) The effects of unequal catchability on Jolly-Seber estimates. Biometrics 29:79–100
Chao A (1987) Estimating population size for capture-recapture data with unequal catchability. Biometrics 43:783–791
Chao A (1989) Estimating population size for sparse data in capture-recapture experiments. Biometrics 45:427–438
Ciucci P, Boitani L (2008) The Apennine brown bear: a critical review of its status and conservation problems. Ursus 19:130–145
Ennis S, Gallagher TF (1994) A PCR-based sex determination assay in cattle based on the bovine amelogenine locus. Anim Genet 25:425–427
Falcucci A, Ciucci P, Maiorano L et al (2009) Assessing habitat quality for conservation using an integrated occurrence-mortality model. J Appl Ecol 46:600–609
Gerloff U, Schlotterer C, Rassmann K et al (1995) Amplification of hypervariable simple sequence repeats (microsatellites) from excremental of wild living Bonobos (Pan paniscus). Mol Ecol 4:515–518
Gervasi V, Ciucci P, Boulanger J et al (2008) A preliminary estimate of the Apennine Brown Bear population size based on hair-snag sampling and multiple data-source mark-recapture Huggins models. Ursus 19:103–121
Hines JE (2006) PRESENCE2—software to estimate patch occupancy and related parameters. USGS-PWRC. http://www.mbr-pwrc.usgs.gov/software/presence.html
Huggins RM (1991) Some practical aspects of a conditional likelihood approach to capture experiments. Biometrics 47: 725–732
Kalinowski ST, Sawaya MA, Taper ML (2006) Individual identification and distribution of genotypic differences between individuals. J Wildl Manag 70:1148–1150
Karanth KU, Nichols JD, Kumar NS, Hines JE (2006) Assessing tiger population dynamics using photographic capture-recapture sampling. Ecology 87:2925–2937
Kendall KC, Stetz JB, Roon DA et al (2008) Grizzly bear density in Glacier National Park, Montana. J Wildl Manag 72:1693–1705
Kendall KC, Stetz JB, Boulanger JB et al (2009) Demography and genetic structure of a recovering grizzly bear population. J Wildl Manag 73:3–17
Kroll AJ, Duke SD, Runde DE, Arnett EB, Austin KA (2007) Modeling habitat occupancy of orange-crowned warblers in managed forests of Oregon and Washington, USA. J Wildl Manag 71:1089–1097
Lindberg M, Rexstad E (2002) Capture-recapture sampling designs. In Encyclopedia of environments, vol 1. Wiley, Chichester, pp 251–262
Longoria MP, Weckerly FW (2007) Estimating detection probabilities from sign of collared peccary. J Wildl Manag 71:652–655
Lorenzini R, Posillico M, Lovari S, Petrella A (2004) Noninvasive genotyping of the endangered Apennine brown bear: a case study not to let one's hair down. Anim Conserv 7:199–209
Lukacs PM, Burnham KP (2005) Estimating population size from DNA-based closed capture-recapture data incorporating genotyping error. J Wildl Manag 69:396–403
McKenzie D, Nichols JD, Lachman GB et al (2002) Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255
Miller CR, Joyce P, Waits LP (2002) Assessing allelic dropout and genotype reliability using maximum likelihood. Genetics 160:357–366
Mills LS, Citta JJ, Lair KP, Schwarz MK, Tallmon T (2000) Estimating animal abundance using non-invasive sampling. Promises and pitfalls. Ecol Appl 10:283–294
Mortelliti A, Boitani L (2007) Estimating species' absence, colonization and local extinction in patchy landscapes: an application of occupancy models with rodents. J Zool 273:244–248
Mowat G, Strobek C (2000) Estimating population size of grizzly bears using hair capture, DNA profiling, and mark-recapture analysis. J Wildl Manag 64:183–193
Murphy MA, Waits LP, Kendall KC et al (2002) An evaluation of long-term preservation methods for brown bear (*Ursus arctos*) faecal DNA samples. Conserv Genet 3:435–440
Norris JL, Pollock KH (1996) Nonparametric MLE under two closed capture-recapture models with heterogeneity. Biometrics 52:639–649
Paetkau D (2003) An empirical exploration of data quality in DNA-based population inventories. Mol Ecol 12:1375–1387
Paetkau D (2004) The optimal number of markers in genetic capture-mark-recapture studies. J Wildl Manag 68:449–452
Paetkau D, Calvert W, Stirling I, Strobek C (1995) Microsatellite analysis of population structure in Canadian polar bears. Mol Ecol 4:347–354
Pellet J (2008) Seasonal variation in detectability of butterflies surveyed with Pollard walks. J Insect Conserv 12:155–162
Pledger S (2000) Unified maximum likelihood estimates for closed models using mixtures. Biometrics 56:434–442
Pledger S (2005) The performance of mixture models in heterogeneous closed population capture-recapture. Biometrics 61: 868–876
Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. Nat Rev 6:847–859
Poole KG, Mowat G, Fear DA (2001) DNA-based population estimate for grizzly bears Ursus arctos in Northeastern British Columbia, Canada. Wildl Biol 7:105–115
Potena G, Sammarone L, Randi E et al (2004) Relazione finale sul censimento della popolazione, sullo status genetico e demografia—dinamica di popolazione. Prodotto identificabile del Progetto LIFENAT99/IT/006244, Ministero delle Politiche Agricole e Forestali and European Commission, Bruxelles, Belgium (in Italian)
Randi E, Gentile L, Boscagli G, Huber D, Roth HU (1994) Mitochondrial DNA sequence divergence among some west European brown bear (*Ursus arctos*) populations: lessons for conservation. Heredity 73:480–489
Randi E, Pierpaoli L, Potena G, Sammarone L, Petrella A, Posillico M (2006) Relazione finale sul conteggio della popolazione, sullo status genetico e demografia—dinamica di popolazione. Prodotto identificabile del Progetto LIFENAT99/IT/006244. Ministero delle Politiche Agricole e Forestali—European Commission, Bruxelles, Belgium (in Italian)
Romain-Bondi KA, Wielgus RB, Waits LP et al (2004) Density and population size estimates for North Cascade grizzly bears using DNA hair-sampling techniques. Biol Conserv 117:417–428
Roon DA, Waits LP, Kendall KC (2003) A quantitative evaluation of two methods for preserving hair samples. Mol Ecol Notes 3:163–166
Roon DA, Waits LP, Kendall KC (2005) A simulation test of the effectiveness of several methods for error-checking non-invasive genetic data. Anim Conserv 8:203–215
Solberg KH, Bellemain E, Drageset O, Taberlet P, Swenson JE (2006) An evaluation of field and non-invasive genetic methods to estimate brown bear (*Ursus arctos*) population size. Biol Conserv 128:158–168
Taberlet P, Griffin SB, Goossens B et al (1996) Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Res 24:3189–3194
Taberlet P, Camarra JJ, Griffin S et al (1997) Non invasive genetic tracking of the endangered Pyrenean brown bear population. Mol Ecol 6:869–876
Valière N (2002) GIMLET: a computer program for analysing genetic individual identification data. Mol Ecol Notes 2:377–379

Valiere N, Bonenfant C, Toigo C et al (2007) Importance of a pilot study for non-invasive genetic sampling: genotyping errors and population size estimation in red deer. Conserv Genet 8:69–78

Valiére N, Berthier P, Mouchiroud D, Pontier D (2002) GEMINI: software for testing the effects of genotyping errors and multitubes approach for individual identification. Mol Ecol Notes 2:83–86

Waits LP, Leberg PL (2000) Biases associated with population estimation using molecular tagging. Anim Conserv 3:191–199

Waits LP, Paetkau D (2005) Non invasive sampling tools for wildlife biologists: a review of applications and recommendations for accurate data collection. J Wildl Manag 69:1419–1433

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

White GC, Burnham KP (1999) Program MARK: survival estimation from population of marked animals. Bird Study 46(supplement):120–138

White GC, Shenk TM (2001) Population estimation with radio marked animals. In: Millspaugh JJ, Marzluff JM (eds) Design and analysis of radio telemetry studies. Academic Press, San Diego, pp 329–350

White GC, Anderson DR, Burnham KP, Otis DL (1982) Capture-recapture and removal methods for sampling closed populations. Los Alamos National Laboratory, Los Alamos, N. Mex

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