Conserv Genet (2007) 8:69–78
DOI 10.1007/s10592-006-9149-2

ORIGINAL PAPER

Importance of a pilot study for non-invasive genetic sampling: genotyping errors and population size estimation in red deer

Nathaniel Valière · Christophe Bonenfant · Carole Toïgo · Gordon Luikart · Jean-Michel Gaillard · François Klein

Received: 5 September 2005 / Accepted: 12 January 2006 / Published online: 3 June 2006

Abstract Population size information is critical for managing endangered or harvested populations. Population size can now be estimated from non-invasive genetic sampling. However, pitfalls remain such as genotyping errors (allele dropout and false alleles at microsatellite loci). To evaluate the feasibility of non-invasive sampling (e.g., for population size estimation), a pilot study is required. Here, we present a pilot study consisting of (i) a genetic step to test loci amplification and to estimate allele frequencies and genotyping error rates when using faecal DNA, and (ii) a simulation step to quantify and minimise the effects of errors on estimates of population size. The pilot study was conducted on a population of red deer in a fenced natural area of 5440 ha, in France. Twelve microsatellite loci were tested for amplification and genotyping errors. The genotyping error rates for microsatellite loci were 0–0.83 (mean=0.2) for allele dropout rates and 0–0.14 (mean=0.02) for false allele rates, comparable to rates encountered in other non-invasive studies. Simulation results suggest we must conduct 6 PCR amplifications per sample (per locus) to achieve approximately 97% correct genotypes. The 3% error rate appears to have little influence on the accuracy and precision of population size estimation. This paper illustrates the importance of conducting a pilot study (including genotyping and simulations) when using non-invasive sampling to study threatened or managed populations.

Keywords Non-invasive sampling · Microsatellites · Genotyping error · Pilot study · *Cervus elaphus* · Population size

N. Valière (✉) · C. Bonenfant · J.-M. Gaillard
UMR CNRS 5558 Biométrie et Biologie, Evolutive, Université Claude Bernard Lyon I, Villeurbanne, France
e-mail: valiere@biomserv.univ-lyon1.fr

C. Toïgo · F. Klein · C. Bonenfant
Office National de la Chasse et de la Faune Sauvage, Centre National d'Etudes et de Recherches Appliquées, Cervidés-Sanglier, Erstein, France

G. Luikart
UMR CNRS 5553 Laboratoire d'Ecologie Alpine, Université Joseph Fourrier, Grenoble, France

Introduction

The effective management of animal populations requires knowledge of population sizes (Cederlund et al. 1998). This is particularly relevant in the case of threatened species for which monitoring population status and maintaining population viability is the priority goal, or for game species for which population harvest and control is of interest (Cederlund et al. 1998; Wemmer 1997). An abundant literature dealing with the estimation of population size is available to biologists and managers (see Schwartz and Seber 1999; Seber 1982 for reviews). One can distinguish between (i) indices of relative abundance such as faecal pellet counts (see for example Marques et al. 2001), for which a low accuracy can be problematic, (ii) abundance (animal counts), and (iii) absolute population size estimators (line-transect or mark-recapture methods). The latter methods provide the most accurate and robust results (Schwartz and Seber 1999; Williams et al. 2002), however animal tagging for individual identification is required. Depending both on species characteristics (behaviour, size and a priori abundance) and managers investment capabilities, marking many individuals can become very problematic or unrealistic.

New approaches that use genetic identification of individuals from non-invasive samples (see for example reviews in Kohn and Wayne 1997) are now available to estimate both population size and structure (e.g. sex ratio, genetic diversity) without the need for tagging. Population size estimates from faeces have been conducted successfully for coyote (*Canis latrans*, Kohn et al. 1999), wolf (*Canis lupus*, Creel et al. 2003), African forest elephant (*Loxodonta cyclotis*, Eggert et al. 2003) and the European badger (*Meles meles*, Wilson et al. 2003). However, several pitfalls of these new approaches can lead to spurious results. Both false allele occurrence (additional allele amplified or scored) and allele dropout (failure of allele to amplify or score) are common errors during the genotyping process.

Simulations have already shown that genotyping inaccuracies can bias population size estimates as high as 200% (Waits and Leberg 2000). Such errors must be accounted for, and some authors have proposed the use of multiple multilocus genotyping from which one consensus genotype is derived, supposedly free of errors (Navidi et al. 1992; Taberlet et al. 1996). As this approach is costly in time and money, a pilot study (Taberlet and Luikart 1999; Valière 2002b) prior to the overall analysis is useful to determine (i) genotyping error rates, (ii) likely effects of errors on population parameter estimators, and (iii) the effort required to sufficiently minimise genotyping errors. The pilot study should provide the conditions needed for an accurate estimate of population size. In this context, the pilot study is a preliminary test to determine if a study is feasible. This includes tests of amplification from samples, assessment of error rates and marker variability, genotyping design (optimal simplification of the design) and data analysis.

Here we present a pilot study conducted on an enclosed red deer (*Cervus elaphus*) population in France. Funding for this pilot study was limited to 5000 euros and was composed of two steps. We first tested and optimised amplification for 12 microsatellite loci. In a second step we performed simulations to determine the strategy (number of samples, number of PCRs per sample) required to provide reliable estimates of population size.

Material and methods

Sample collection

Tissue samples

All samples were collected in a closed area (Domaine National de Chambord) in France. We collected tissue samples from 40 animals (8 fawns, 32 adults) of both sexes (15 females, 25 males) during autumn–winter 2002–2003. Out of these 40 individuals, 27 had been net-trapped and 13 were harvested during the hunting season. Tissue samples (<0.5 cm^2) were collected from ear biopsies for both living and dead animals and were stored in 96% ethanol. Thereafter samples were stored in 35-mL plastic or glass tubes at room temperature and sheltered from light for up to 6 months until DNA extraction.

Faeces samples

Samples used for the pilot study were randomly chosen among two sets of faeces collected in the field in February and April 2003 along line transects distributed systematically over the entire area. Whenever possible, at least 3 droppings were collected in a 35-mL plastic tube using wooden sticks (for example) to avoid cross-contamination between samples and potential human contamination. At the end of the sampling day, 96% ethanol was added to the collection tubes (ethanol entirely covered the samples). Tubes were sheltered from light and stored at room temperature until DNA extraction.

DNA extraction

Tissue samples

DNA was extracted using a Chelex based method (Walsh et al. 1991). Small pieces of each tissue sample (handled with new disposable blade) were added to 600 µL of 10–40% Chelex solution. Seven µl of a 20 mg/ml solution Proteinase K were added and the mix was incubated two hours at 56°C and 15 min at 94°C. Only the supernatant was used for PCR amplification. Control extraction blanks (extraction procedure without sample) were included.

Faeces samples

DNA extractions were conducted in a UV-sterilised room where no red deer tissue samples were previously treated. Each sample was handled with a new disposable blade. At least two control extraction blanks (extraction procedure without sample) were included for each extraction session.

Extractions were performed as follows: Mucus on the dropping surface containing intestinal cells was washed in a lysis washing buffer (0.1 M Tris–HCl, 0.1 M EDTA, 0.01 M NaCl, *N*-lauroyl sarcosine 1%, pH 7.5–8). One

dropping from each individual was washed in 600 µl of washing buffer for 15 min in a 7 ml sterilised tube. The supernatant (approximately 300 µl) was used for DNA extraction using the DNeasy Blood Kit (QIAGEN GmbH, Germany) following the manufacturer's procedures.

DNA amplification

Microsatellite genotyping

Extractions and amplifications were performed in separate rooms. Amplification mixtures were prepared in a UV-sterilised cabinet. Pipetting was performed using filtered tips and one control PCR blank was included for each set of amplifications.

We originally selected the set of loci from Bonnet et al. (2002) because: (i) it includes 11 microsatellite loci previously tested on several cervid species, (ii) it requires only three multiplexed PCR reactions and one sequencer run and (iii) the loci are on different chromosomes (except for two pairs). We included one additional locus (TGLA126), which can be added in the third multiplexed PCR reaction (Bonnet 2001). In the table published in Bonnet et al. (2002) the primers sequences were not correctly matched to loci. The corrected information concerning the 12 loci are given in Table 1 in agreement with authors of Bonnet et al. (2002).

The PCR conditions given in Bonnet et al. (2002) for the three multiplexed PCR reactions were retested and re-optimised for $MgCl_2$ concentration, relative primer concentration, and annealing temperature to achieve intense bands on agarose gels with ethidium bromide staining for all loci in each multiplex.

Optimised PCR amplifications were performed in a total volume of 12 µl with 3 µl of DNA extract, 1× PCR buffer (PE Applied Biosystems), 0.2 mM dNTP (each), 2.5 mM MgCl, 0.2 mg/ml BSA and 2 U of AmpliTaq Gold™ polymerase (PE Applied Biosystems). Primers concentrations are provided in Table 1. PCR amplifications were conducted in a PTC100 thermal controller (MJ Research, Inc.) under the following conditions: an initial denaturation at 95°C for 10 min, followed by 40 cycles (for tissue samples) or 55 cycles (for faeces samples) of 40 s at 95°C, 40 s at 59/58/56°C (respectively for Multiplex 1, 2 and 3) and 60 s at 72°C. A final elongation step was conducted at 72°C for 10 min. PCR products for all loci and each sample were mixed at equal final concentrations, and the mix for each sample was run in a single capillary of a MEGABACE 1000 (Pharmacia) sequencer. Fragment analysis was

Table 1 Characteristics of the 12 microsatellite loci multiplexed in three fluorescent sets (Dye) used in this study

Locus	Dye	Seq. 5′–3′	BOVMAP name	Size range	Conc. (µM)	Chr.
TGLA57	6-FAM	GCTTTTTAATCCTCAGCTTGCTG GCTTCCAAAACTTTACAATATGTAT	D1S8	80–100*	0.8	1
INRA121	6-FAM	GGAAACCCATTGGAGGATTTG CTTCACTATTCCCCACAAAGC	D18S6	113–167	0.8	18
IDVGA55	6-FAM	GTGACTGTATTTGTGAACACCTA TCTAAAACGGAGGCAGAGATG	D18S16	195–209	0.2	18q24dist
BMC1009	6-FAM	GCACCAGCAGAGAGGACATT ACCGGCTATTGTCCATCTTG	D5S15	282–306	0.2	5q23
VH110	HEX	CTCTAGAGGATCACAGAGAGTCGG GCAGAAACATTTTTTTCCTTCAATATAGTTTCCC	OARVH110	119–132	0.4	29
BM757	HEX	TGGAAACAATGTAAACCTGGG TTGAGCCACCAAGGAACC	D9S13	163–207	0.8	9
BL42	HEX	CAAGGTCAAGTCCAAATGCC GCATTTTTGTGTTAATTTCATGC	D13S11	253–269	0.8	13q22
BM848	HEX	TGGTTGGAAGGAAAACTTGG CCCTCTGCTCCTCAAGACAC	D15S9	355	0.8	15
TGLA53	NED	GCTTTCAGAAATAGTTTGCATTCA ATCTTCACATGATATTACAGCAGA	D16S3	151–193	0.8	16
BM203	NED	GGGTGTGACATTTTGTTCCC CTGCTCGCCACTAGTCCTTC	D27S1	214–236	0.8	27q23–q24
CSSM43	NED	AAAACTCTGGGAACTTGAAAACTA GTTACAAATTTAAGAGACAGAGTT	D27S4	210–300*	0.4	27
TGLA126	NED	CTAATTTAGAATGAGAGAGGCTTCT TTGGTCTCTATTCTCTGAATATTCC	D20S1	100–130*	0.8	20

Primers labelled with the same dye were amplified in a single PCR reaction. Primer sequences (Seq. 5′–3′) have been corrected from BOVMAP database compared to Bonnet 2001. Ranges for fragments length (Size range) are those found in the present study except for loci marked by * for which range is that indicated in Bonnet (2001) and Bonnet et al. (2002). Primer concentration (Conc. µM) have been modified compared to Table 3 in Bonnet et al. 2002. Reference number of chromosome (Chr.) for each locus is given from BOVMAP database

performed with the software packages associated with the MEGABACE 1000 instrument (Genetic Profiler, Pharmacia). Only alleles that were scored unambiguously were selected (unclear microsatellite profile were discarded).

Based on previous studies (Taberlet et al. 1996) and on our personal experience, we assumed that four genotypings were sufficient to estimate error rates for samples of faeces and eight genotypings were sufficient for tissue samples. Thus eight of the 40 tissue samples were randomly chosen and were amplified four times per locus to estimate error rates for tissue samples. As errors were nearly absent for tissue samples (see Results), the remaining 32 tissue samples were amplified only once per locus. Twelve faeces samples were randomly chosen, six among the first sampling session and six among the second sampling session. The faeces samples were genotyped eight times per locus to estimate error rates.

Data analysis

Consensus genotypes were constructed from replicate genotypings through the multiple PCR approach. This procedure had to correct the genotyping error by replicating experiments for each sample and locus (see Navidi et al. 1992; Taberlet et al. 1996). Here we followed the threshold rules: alleles had to appear at least twice out of four PCR repetitions for tissue samples and alleles had to appear at least three times out of eight PCR repetitions for faeces samples. Error rates were therefore quantified by comparison of each replicate genotype with the consensus genotype. The construction of the consensus genotype and the quantification of error rates were performed using the ''Error rates estimation'' module in GIMLET v1.3.4 (Valière 2002a).

The nomenclature we used was the following:

- A positive PCR amplification (PCR+) was a PCR amplification for which a product of the expected length was obtained after amplification and for which a microsatellite profile could be easily scored.
- A consensus genotype (CG) was based on the threshold rule (see above) for replicate genotypings.
- An ambiguous genotype (AG) was a case in which multiple alleles were selected for a (diploid) genotype or when no allele could be selected based on the threshold rules.
- An allelic dropout (ADO) was a case in which one of the alleles was not amplified or not scored for a heterozygous individual.
- A false allele (FA) was a case in which one allele differed from the consensus allele that was amplified or scored for a homozygous individual (false alleles in heterozygotes were easily detected and this error was not considered).

Other types of errors (not explained by ADO or FA) can occur but were not taken into account because of their very low frequencies. The calculation of allele frequencies for the 40 tissue samples was performed using GIMLET v1.3.4 (Valière 2002a) after constructing CGs for replicated tissue samples.

As a test of the power of multilocus genotyping to discriminate between individuals, we calculated the probability of identity (P_{ID}). P_{ID} calculations were performed with GIMLET v1.3.4 with both the unbiased equation for small sample size and the equation for siblings. As advocated in Waits et al. (2001), the observed probability in the population ranged from unbiased P_{ID} to P_{ID} for siblings. Allele diversity, heterozygosities and Hardy–Weinberg tests were performed on the same data using GIMLET v1.3.4 and GENEPOP v3.3 (Raymond and Rousset 1995).

Simulations and population size estimation

To optimise the genotyping procedure (number of PCR amplifications per locus and sample), we used the computer program GEMINI v1.4.0 (Valière et al. 2002). This software uses error rates, and allele frequency estimates to simulate a virtual population, sampling, genotyping and subsequent analysis. Based on simulations of different conditions, we have chosen the strategy that maximises the number of case where the genotyping error (ADO and FA) have been corrected through the multiple PCR procedure. Parameters were set according to the conditions of the study. For our study, the population size was set to 1000. This figure was based on field observations from the Chambord wildlife managers. The sampling effort was set to 500 samples collected during two sampling occasions to match the actual sampling effort in the field. All individuals had the same probability of being sampled. Number of loci and error rates were based on the genotyping part of the study.

To account for differences between estimation methods, we tested two approaches: the rarefaction curve method (RCM, see Kohn et al. 1999) and the Capture–Mark–Recapture (CMR) models (Schwartz and Seber 1999). In the first case, population size estimates were performed using files constructed by GEMINI v1.4.0 and thereafter analysed in the R program (Ihaka and Gentleman 1996). As the RCM is highly sensitive to the order in which the faeces were analysed (Kohn et al. 1999), we randomised 20 times the order of analysed samples and then estimated 20 times the population size. We calculated the mean value from these 20 estimates to get on value for each RCM. For

CMR, the files saved by GEMINI were analysed with the CAPTURE software (Otis et al. 1978) using the null model (M_0).

Results

DNA amplification and optimisation for microsatellite loci

Among the 12 loci tested, only nine (BMC1009, IDVGA55, INRA121, BM203, TGLA53, BL42, BM757, BM848 and VH110) were successfully genotyped and analysed in our study. Locus TGLA57 could not be scored because of a problem during the electrophoresis on the MEGABACE 1000 (Amersham) sequencer. As capillary sequencing is generally sensitive to the amount of PCR products, we suspect that large amount of primer dimers in capillaries might have influenced fragment migration for this locus, although a reduction in primer concentration did not decrease the problem (results not shown). Locus TGLA126 also produced a pattern of several peaks that was difficult to interpret. Locus CSSM43 could not be successfully amplified (as in Bonnet et al. 2002). Finally, it was difficult to amplify the VH110 locus from faeces (after 55 cycles). Only tissue samples (40 cycles) could be properly analysed with this locus.

Genotyping errors for tissue

Genotyping errors from tissue were estimated from eight samples. No ambiguous genotypes were detected. The proportion of PCR+s among loci varied between 88 and 97% (mean 95%) among loci and from 67 to 89% among samples. Only two ADO were detected, both at locus TGLA53 (0.7%).

Allele frequencies and genetic variability

Determination of allele frequencies for eight loci (BMC1009, IDVGA55, INRA121, BM203, TGLA53, BL42, BM757 and VH110) was based on tissue samples from 40 different individuals. All loci (except BM848) were polymorphic in the population. Average number of alleles per polymorphic locus was 8, ranging from 5 (BL42) to 12 (TGLA53). Mean observed heterozygosity over all polymorphic loci was 0.52 and mean expected heterozygosity was 0.72 (Table 2). Loci INRA121, TGLA53, BL42, BM757 and VH110 showed a significant heterozygote deficit (Hardy–Weinberg Exact test, $P < 0.05$). Four pairs of loci (always involving locus BM203, in association with INRA121, IDVGA55, BM757, and VH110) were significantly associated (genotypic disequilibrium test, $P<0.05$).

Probability of identity was calculated on the basis of allele frequencies estimated from the 40 tissue samples. The overall P_{ID} using all loci was 4.27×10^{-9} and the overall $P_{ID\text{-}sibs}$ was 9.45×10^{-4}. If loci VH110, TGLA53 and INRA121 are excluded, which are the three loci that were difficult to type for faeces samples, the overall P_{ID} became 5.57×10^{-6} and the overall $P_{ID\text{-}sibs}$ was 1.17×10^{-2}. Thus $P_{ID\text{-}sibs}$ was again close to the 0.01 threshold necessary to prevent the shadow effect, i.e. the presence of two or more individuals with the same multilocus genotype (Mills et al. 2000).

Genotyping error for faeces samples (microsatellite loci)

Genotyping error rates for faeces were estimated from 12 samples. In 13 of 96 pairs (12 loci×8 individuals), we were not able to resolve consensus genotypes from repeated assays (for example when a majority of PCR reactions did not produce a usable amplicon). Error types other than allele dropout (ADO) or false allele (FA), such as a

Table 2 Number of alleles (Number alleles), expected (Hexp) and observed (Hobs) heterozygosities, rates of positive PCR (PCR+), droupout (ADO) and false allele (FA) for each locus used in this study

Locus	Multiplex	Number alleles	Hexp	Hobs	PCR+	ADO	FA	ADO max	FA max
INRA121*	1	8	0.61	0.22	0.72	0.03	0	0.13	0
IDVGA55	1	7	0.73	0.80	0.80	0.12	0.07	0.25	0.14
BMC1009	1	10	0.81	0.80	0.86	0.27	0	0.83	0
VH110*	2	5	0.61	0.05	–	–	–	–	–
BM757*	2	9	0.70	0.57	0.93	0.05	0.03	0.13	0.13
BL42*	2	5	0.73	0.55	0.88	0.23	0	0.75	0
BM848	2	1	0.00	0.00	0.74	–	0.01	–	–
TGLA53*	3	12	0.85	0.57	0.65	0.32	0.02	0.5	0.13
BM203	3	8	0.68	0.65	0.86	0.30	0	0.5	0
Mean		8	0.72	0.52	0.81	0.19	0.02		

Columns ADO max and FA max indicate maximum values of ADO and FA rates for loci used in the simulations. Means for the number of alleles are computed over polymorph loci. Stars indicate loci for which the heterozygotes deficit is significant

heterozygous genotype misidentified as a different heterozygous genotype were also detected but they appeared at very low rates (<0.001) and we considered them as negligible. The average proportion of positive PCRs was 81% (Table 2) and varied among loci from 65 to 93% and among samples from 41 to 100%.

The median ADO rate was 0.125 (Mean = 0.200) for non-ambiguous heterozygous consensus genotypes (N = 52) ranging from 0 to 0.833. The ADO rates were not homogenous in each of the multiplexed reactions (see Table 2). Nine samples out of the 12 tested showed ADO rates higher than 0.3 for four loci, and four out of these nine samples showed ADO rates higher than 0.5 for three loci. False allele rates averaged 0.02 (Table 2) among all PCR for homozygote samples with non ambiguous consensus genotype (N = 29) and ranged from 0 to 0.143.

Simulation and population size estimation

We performed simulations to define the minimum number of PCR amplifications required per locus and sample to reach the lowest incorrect identification rate after simulation in GEMINI. Simulations were performed using the maximum values shown in Table 2 for each locus and all samples had the same error rates. As error rates can differ among loci in a multiplexed reaction (see Table 2), we set the same number of PCRs for all loci to correct errors.

Figure 1 shows that the proportion of correct genotypes (multilocus CG match the true multilocus genotype) increased with the number of PCR amplifications per sample. With 6 PCR amplifications per locus, the proportion of correct multilocus genotyping was about 97.1% (±0.83) and this proportion was 99.2% (±0.37) for 8 amplifications.

Results for simulations of population size estimation are shown in Table 3. CMR (Model M_0) provided the least biased estimates of population size when no genotyping errors were introduced. RCM with the Eggert et al. (2003) or Chessel (Valière 2002a) equations gave more biased estimates than M_0, but relative biases were lower than 3%. When genotyping errors were introduced and when these errors were corrected using the simple threshold rule among six or eight PCR amplifications, it seems that RCM with the Eggert et al. (2003) or Chessel equations provided estimates with less bias (relative bias from true population size: 8–10%) and less variance than the CMR methods (M_0 model) or the RCM with the Kohn et al. (1999) equation. The last method (RCM using Kohn et al. 1999 equation) gave estimates with upward biases around 90% under all of the tested conditions. Relative bias decreased by about 5% when 8 (instead of 6) PCR amplifications were conducted per sample and locus.

Discussion

Of the 52 samples analysed, eight loci for all tissue samples (N = 40) and for even faeces samples (out of 12) could be amplified. It was possible to amplify at least six or seven loci for three additional faecal samples. For three loci, having short allele lengths (around 100 bp), we could not amplify and score unambiguous microsatellite profiles. As capillary sequencers are sensitive to DNA quantities and because PCR product purification does not completely remove primer dimers, we suspect that the difficulty amplifying these loci was due to primer dimers present in the mix loaded in the capillaries.

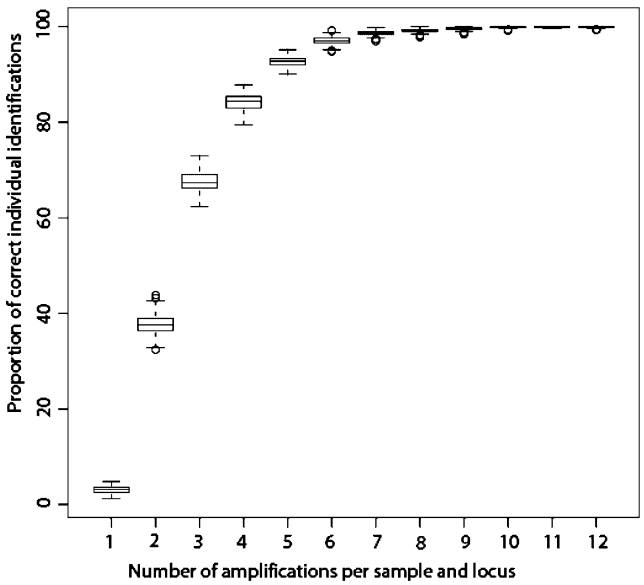

Fig. 1 Plot of percentiles (over 200 iterations of simulations) of the proportion of correct genotyping (consensus genotype was identical to the true genotype of the sample among all loci) versus the number of PCR amplifications per locus and sample. Error rates used for conducting simulations are the higher values for each locus appearing in Table 2

Table 3 Results of population size estimation from simulated population of 1000 individuals

Methods	Without error	From CG after 6 PCR	From CG after 8 PCR
CMR (M)	1007.2 ± 109.2	1401.17 ± 246.0	1374.13 ± 231.9
RCM Kohn	1858.3 ± 173.8	1997.4 ± 199.1	1877.7 ± 170.3
RCM Chessel	1013.7 ± 83.8	1079.6 ± 92.2	1027.81 ± 81.0
RCM Eggert	1025.7 ± 87.9	1096.1 ± 100.6	1035.6 ± 86.1

Population size was estimated using CMR and RCM using the three available equations (see text) from genotyped samples. Estimates from either genotypes without error and consensus genotypes after 6 or 8 PCR amplifications are shown. Standard deviation indicated are those among the 200 simulations

The $P_{\text{ID-Sibs}}$ using the seven polymorphic loci for which we can have reliable data (BMC1009, IDVGA55, INRA121, BM203, TGLA53, BL42, BM757) is close to the 0.01 threshold necessary to prevent the shadow effect (see Mills et al. 2000). This discrimination power with only seven loci is sufficient to ensure the correct discrimination of the individuals (without considering genotyping error). Development of optimal conditions for the remaining loci (e.g. TGLA57, TGLA126 and VH110) should obviously enhance this. The adequacy of this discrimination power was supported by the absence of identical genotypes among the 52 samples used in this pilot study (even when accounting for missing alleles) and by the match probabilities (Woods et al. 1999; match probability calculated for sibs with GIMLET v1.3.4; data not shown) for each tissue sample, which were lower than 0.01 (except for three samples; N=40). Additionally, sex information about samples can increase the power of discrimination of individual by reducing the probability of identity. In our case, it was also possible to correctly determine gender of faecal samples using protocol in Huber et al. 2002 (results not shown).

Results from 40 tissue samples from different individuals showed a heterozygote deficit for five of the eight polymorphic loci. These deficiencies were unlikely due to dropout errors because PCR replications for four samples did not show ADO except 3 cases for locus TGLA53. Estimates of null allele frequencies for these loci, computed with CERVUS program (Marshall et al. 1998), were between 0.09 and 0.66. However, it is presumably unlikely to find out null alleles for all 5 of 7 polymorphic loci and we can safely advance that the probability of finding null alleles for all these loci is close to zero. Finally, without arguments for a potential selection against heterozygotes, the deviation from HWE is likely mostly caused by substructure in the population. More precisely, the genetic isolation probably lead to an increase of inbreeding and hence to the loss of heterozygosity without loss of a lot of alleles (loci have a relatively high level of allele diversity). However, genetic drift could also lead to the apparent heterozygotes deficit as illustrated by the case of locus BM848, which is monomorphic among the 40 individuals. Further analyses are needed to confirm these findings and to estimate genetic variability more precisely.

Error rates lie in the range of rates found in other studies based on non-invasive sampling (see Table 4). However, as mentioned in Creel et al. (2003) in the context of the population size estimation, it should be more appropriate to consider the highest error rate among samples to be sure that all errors in the data set will be corrected. In our study, several samples showed a dropout rate higher than or equal to 0.5 for some loci. We suspect that these high error rates were either due to the quality of some samples (low quantity and/or quality of DNA) and to the high number of PCR cycles for the faecal DNA amplification. These conditions can considerably inflate the dropout rate because the higher the number of cycles, the larger the difference between concentrations of alleles at the end of the PCR if one allele is ''dropped out''. In such case, ADO is more likely to occur than if fewer cycles were ran. However, for the few tests in which we performed only 40 cycles (versus 50) we did not see significantly better results.

Simulations indicated that at least eight amplifications per sample and loci are necessary to obtain an acceptable proportion of correct identification (i.e. >0.99). However, conducting six PCR amplifications per sample will provide a reliable individual identification with a proportion of 97% of correct identifications. Whereas the common procedure (Taberlet et al. 1996) used in studies using non-invasive sampling consists of conducting three amplifications and then four additional amplifications (depending on results of the first three PCR), our approach consists of applying the same number of PCR amplifications to all samples and loci. This genotyping design has three main advantages:

(i) This procedure allows a simple standardised error correction procedure and thus prevents the multiple manipulations of the samples and the subsequent risk of contamination or mishandling during each step. However, our approach does not exclude reanalysing some problematic samples (see for example Paetkau 2003).
(ii) As loci were pooled in amplification multiplexes and did not show homogenous error rates, it is possible to

Table 4 Error rates found in several studies based on noninvasive sampling (dropout ADO and false allele FA)

References	Species	Type	N loci	ADO	FA
Bayes et al. (2000)	*Papio* sp.	Faeces	8	8% (P)	1.2% (P)
Constable et al. (2001)	*Pan troglodytes*	Faeces	16	30% (S)	
Morin et al. (2001)	*Pan troglodytes*	Faeces	9	24% (P)	
Launhardt et al. (1998)	*Presbytis entellus*	Faeces	5	5% (S)	0% (S)
Lathuillière et al. (2001)	*Macaca sylvanus*	Faeces	3	0–6% (P)	13–20% (P)
Goossens et al. (2000)	*Pongo pygmaeus*	Faeces	1	4.2% (S)	3% (S)
Ernest et al. (2000)	*Puma concolor*	Faeces	12	8% (P)	
Kohn et al. (1999)	*Canis latrans*	Faeces	3	2.6% (P)	2.6% (P)
Lucchini et al. 2002	*Canis lupus*	Faeces	6	18% (P)	≈0% (P)
			1	0–33% (P)	
Frantz et al. (2003)	*Meles meles*	Faeces	7	27% (P)	8%[c]
Murphy et al. (2002)	*Ursus arctos*	Faeces	1	≈7% (S)[a]	≈6% (S)[a]
				≈18% (S)[b]	
Taberlet et al. (1996)	*Ursus arctos*	Faeces	1	53% (P)	<5% (P)
Bonin et al. (2004)	*Ursus arctos*	Faeces	6	1% (A)	0.13% (A)
Garnier et al. (2001)	*Diceros bicornis*	Faeces	10	30% (P)	≈0% (P)
Parsons (2001)	*Tursiops truncatus*	Faeces	3	0% (P)	0.97% (P)
Piggot et al. (2004)	*Petrogale penicilata*	Faeces	6	0.92% (P)	3.72% (P)
				0% (P)*	0.02% (P)*
	Dasyurus viverrinus	Faeces	6	0.41% (P)	2.56% (P)
				0.21% (P)*	0% (P)*
	Dasyurus maculatus	Faeces	6	2.77% (P)	7.59% (P)
				0% (P)*	0% (P)*
Gagneux et al. (1997)	*Pan troglodytes*	Hairs	11	31.3% (P)	
Goossens et al. (1998)	*Marmota marmota*	1 Hair	1	14.3% (P)	4% (P)
		3 Hairs		4.9% (P)	1.4% (P)
		10 Hairs		0.4% (P)	0% (P)
Mowat and Paetkau (2002)	*Martes americana*	Hairs	6	1.4% (S)	1.5% (S)
Sloane et al. (2000)	*Lasiorhinus krefftii*	Hairs	12	0% (P)	0.3% (P)
Valière and Taberlet (2000)	*Canis lupus*	Urine	3	4.2% (P)	0% (P)
			2 spl.	74.6% (P)	0% (P)
Segelbacher (2002)	*Tetrao urogallus*	Plumes	10	0.77% (P)	0% (P)

P=among PCR amplifications, S=among samples, A=among analysed alleles

*After multiplex PCR (see Piggot et al. 2004)

[a]Approximation over several DNA preservation methods

[b]Multiple alleles error (see Murphy et al. 2002)

[c]Wrong genotypes (see Frantz et al. 2003)

refine the procedure using different numbers of PCRs for each of the multiplexed reactions.

(iii) The proportion of correct genotyping was higher than 97% in our approach for a total of 21,000 PCRs (seven loci, 500 samples and 8 PCR attempts) against about 50% in Taberlet et al. (1996) for a total of about 17,000 PCRs (results from simulations using GEMINI, data not shown). The latter procedure requires many samples to achieve a satisfactory proportion of correct genotypes and limits automation compared to our approach where we applied the same number of replications for all samples and loci.

Simulations indicated that our residual errors (about 1–3% of individual misidentification) led to small relative biases of population size (less than 10%) using either the capture–recapture or rarefaction methods. The maximum standard deviation was about 10% of the true population size for the CMR method and was 27.3% of the true population size for the RCM method (this latter standard deviation was calculated among the 20 trials—see Materials and methods—and thus only reflects the variability due to sampling order but not the deviation due to the method). This means that the relative bias led by residual genotyping errors for the two methods was lower than the standard deviation led by the estimation itself. Conducting six PCR amplifications per sample and per locus will correct up to 95% of the genotyping errors, leading to a percent of correct multilocus genotyping of 97%. We showed that the residual errors should not greatly affect the population size estimation. A more accurate procedure would be to conduct eight PCRs per sample and locus, but this will decrease only slightly (5%) the relative bias of population size estimation compared to six PCR protocol and requires up to 7000 additional PCRs. The six PCR procedure is then more optimal for our needs.

In conclusion, we suggest that a pilot study is important in any population study using non-invasive samples (Taberlet et al. 1996). This is also true for all other applications. The power required to correctly discriminate between individuals and the ability to correct genotyping errors are both of prime importance but will depend on the application of the study (see Mills et al. 2000 and Taberlet and Luikart 1999). In the case of population size estimation, the probability of identity (i.e. a measure of the power to identify individuals) should be often lower than 0.01, as is the case in our study. Also, as the bias due to genotyping errors could be very high for population size estimates (see Waits and Leberg 2000), multiple PCR amplifications per sample and locus should be conducted. The number of PCR repetitions depends on the error rates and thus on the estimation of the error rates via the pilot study. It is of prime importance to correctly assess the optimal number of PCR amplifications per sample and locus, because on the one hand one has to correct the genotyping error, but on the other hand conducting fewer PCR amplifications per sample allows the typing of more loci either to increase the power to discriminate between individuals or to add information for subsequent studies of genetic diversity or genetic structure. This can decrease the overall number of PCR reactions used.

All these points could be assessed with simulations, as in our study. Here, we have shown that, with a limited number of faecal and tissue samples, we could:

- test the usefulness of primers for loci identified in cattle (Bovidae) with a cervid,
- estimate allele frequencies to assess the power of individual discrimination of the optimised loci,
- estimate genotyping error rates to assess the minimum number of PCR amplifications needed per sample and locus (via a simulation approach) to achieve sufficiently low multilocus rates.
- assess the bias in population size estimation method using real data and simulations.

Acknowledgements The authors would thanks staff of Office National des Forêts (ONF) and Office National de la Chasse (ONCFS), Aurélie Cohas and Ondine Lucas for their help in the field. We also greatly thanks Francis Forget (ONF), Christian Gambier (ONF) and Dominique Odier (ONCFS) for their cooperation. We thanks E. Petit, Joe Niegel and two anonymous reviewers for their comments on previous version of this manuscript. Funding for this project was provided by the ONCFS (contract CNRS-ONCFS no 2002/08).

References

Bayes M, Smith K, Alberts S, Altmann J, Bruford M (2000) Testing the reliability of microsatellites typing from faecal DNA in the savannah baboon. Cons Genet 1:173–176

Bonin A, Bellemain E, Bronken Eidsen P, Pompanon F, Brachmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. Mol Ecol 13:3261–3273

Bonnet A (2001) Etude chromosomique et moléculaire de la sous-famille des Cervinae: application des résultats à la gestion des troupeaux en captivité d'espèces sauvages. PhD Thesis, University of Tours, Tours, France

Bonnet A, Thévenon S, Maudet F, Maillard JC (2002) Efficiency of semi-automated fluorescent multiplex PCRs with 11 microsatellite markers for genetic studies of deer populations. Anim Genet 33:343–350

Cederlund G, Bergqvist J, Kjellander P, Gill R, Gaillard J-M, Boiseaubert B, Ballon P, Duncan P (1998) Managing roe deer and their impact on the environment: maximising the net benefits to society. In: Andersen R, Duncan P, Linnell JDC (eds) The European roe deer: the biology of success. Scandinavian University Press, Oslo, pp 337–372

Constable J, Ashley M, Goodall J, Pusey A (2001) Noninvasive paternity assignment in Gombe chimpanzees. Mol Ecol 10:1279–1300

Creel S, Spong G, Sands JL, Rotella J, Zeigle J, Joe L, Murphy KM, Smith D (2003) Population size estimation in Yellowstone wolves with error-prone noninvasive microsatellite genotypes. Mol Ecol 12:2003–2009

Eggert LS, Eggert JA, Woodruff DS (2003) Estimating population sizes for elusive animals: the forest elephants of Kakum National Park, Ghana. Mol Ecol 12:1389–1402

Ernest H, Penedo M, May B, Syvanen M, Boyce W (2000) Molecular tracking of mountain lions in the Yosemite Valley in California: genetic analysis using microsatellites and faecal DNA. Mol Ecol 9:433–441

Frantz A, Pope L, Carpenter P, Roper T, Wilson G, Delahay R, Burke T (2003) Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. Mol Ecol 12:1649–1661

Gagneux P, Boesch C, Woodruff S (1997) Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. Mol Ecol 6:861–868

Garnier J, Bruford M, Goossens B (2001) Mating system and reproductive skew in the black rhinoceros. Mol Ecol 10:2031–2041

Goossens B, Chikhi L, Utami, S. de Ruiter J, Bruford M (2000) A multi-samples, multi-extracts approach for microsatellite analysis of faecal samples in an arboreal ape. Cons Genet 1:157–162

Huber S, Bruns U, Arnold W (2002) Sex determination of red deer using polymerase chain reaction of DNA from feces. Wildl Soc Bull 30:208–212

Ihaka R, Gentleman R (1996) R: a language for data analysis and graphics. J Comput Graph Stat 5:299–314

Kohn MH, Wayne RK (1997) Facts from feces revisited. Trends Ecol Evol 12:223–227

Kohn MH, York E, Kamradt DA, Haught G, Sauvajot R, Wayne RK (1999) Estimating population size by genotyping faeces. Proc R Soc Lond Ser B 266:657–663

Lathuillière M, Ménard N, Gautier-Hion A, Crouau-Roy B (2001) Testing the reliability of noninvasive genetic sampling by comparing analyses of blood and fecal samples in Barbary Macaques (*Macaca sylvanus*). Am J Primatol 55:151–158

Launhardt K, Epplen C, Epplen J, Winkler P (1998) Amplification of microsatellites adapted from human systems in faecal DNA of wild Hanuman langurs (*Presbytis entellus*). Electrophoresis 19:1356–1361

Marques FFC, Buckland ST, Goffin D, Dixon CE, Borchers DL, Mayle BA, Peace AJ (2001) Estimating deer abundance from line transect surveys of dung: sika deer in southern Scotland. Ecology 38:349–363

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 7:639–655

Mills LS, Citta JJ, Lair K, Schwartz M, Tallmon D (2000) Estimating animal abundance using non-invasive DNA sampling: promise and pittfalls. Ecol Appl 10:283–294

Morin P, Chambers K, Boesch C, Vigilant L (2001) Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). Mol Ecol 10:1835–1844

Mowat G, Paetkau D (2002) Estimating marten population size using hair capture and genetic tagging in Southeastern British Columbia. J Wildlife Manage 8(3)

Navidi W, Arnheim N, Waterman MS (1992) A multitubes approach for accurate genotyping of very small DNA sample using PCR: statistical considerations. Am J Hum Genet 50:347–359

Otis DL, Burnham KP, White GC, Anderson DR (1978) Statistical inference from capture data on closed animal populations. Wildl Monogr 62

Paetkau D (2003) An empirical exploration of data quality in DNA-based population inventories. Mol Ecol 12:1375–1387

Parsons K (2001) Reliable microsatellite genotyping of dolphin DNA from faeces. Mol Ecol Notes 1:341–344

Piggot MP, 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. Cons Genet 5:417–420

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

Schwartz CJ, Seber AF (1999) A review of estimating animal abundance. Stat Sci 14:427–456

Seber AF (1982) The estimation of animal abundance and related parameters. C. Griffin & Co. Ltd., London

Segelbacher G (2002) Noninvasive genetic analysis in birds: testing reliability of feather samples. Mol Ecol Notes 2(3):367–369

Sloane M, Sunnucks P, Alpers D, Beheregaray L, Taylor A (2000) Highly reliable genetic identification of individual northern hairy-nosed wombats from single remotely collected hairs: a feasible censuring method. Mol Ecol 9:1233–1240

Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J (1996) Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Res 24:3189–3194

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

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

Valière N (2002b) Amélioration et optimisation des méthodes non-invasives et des marqueurs microsatellites en Biologie des Populations et de la Conservation. PhD Thesis, Claude Bernard Lyon I, Lyon, France.

Valière 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 J, Leberg P (2000) Biases associated with population estimation using molecular tagging. Anim Conserv 3:191–199

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

Walsh PS, Metzger DA, Higushi R (1991) Chelex100 as a medium for simple extraction of DNA for PCR-based typing forensic material. BioTechniques 10:506–513

Wemmer C (1997) The challenge of conserving large mammals, with an emphasis on deer. In: McShea MJ, Underwood HB, Rappole JH (eds) The science of overabundance: deer ecology and population managment. Smithsonian Institution Press, Washington, D.C., pp 55–66

Williams BK, Nichols JD, Conroy M (2002) Analysis and management of animal populations: modeling, estimation, and decision making. Academic Press, San Diego

Wilson GJ, Frantz AC, Pope LC, Roper TJ, Burke TA, Cheeseman CL, Delahay RJ (2003) Estimation of badger abundance using faecal DNA typing. J Appl Ecol 40:658–666

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