Universität Lüneburg

Fakultät III Umwelt & Technik

Umweltwissenschaften

&

Technische Universität München

Fachgebiet für Wildbiologie und Wildtiermanagement

Diplomarbeit

Establishment of molecular genetic methods for the analysis of non-invasive samples, using alpine ibex (*Capra ibex*) as an example

vorgelegt von

Kristina Salzer

August 2006

Prüfer: PD Dr. Ralph Kühn
Fachgebiet für Wildbiologie und Wildtiermanagement
Technische Universität München

Prof. Dr. Thorsten Aßmann
Institut für Ökologie und Umweltchemie
Universität Lüneburg

Betreuer: Dipl.Biol. Roland Hausknecht, M.Sc. & PD Dr. Ralph Kühn
Fachgebiet für Wildbiologie und Wildtiermanagement

Contents

List of Figures

List of Tables

List of Abbreviations

A	absorbance
acc. to	according to
AD	Allelic Dropout
bp	base pair
BSA	Bovine Serum Albumin
°C	degrees Celsius
cyt b	cytochrome b
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphate
EDTA	ethylene diamine tetraacetic acid
FA	False Allele
Fig.	figure
Int	intercept
KV	Kilovolt
log	logarithm
$MgCl_2$	Magnesium chloride
mg	milligram
mtDNA	mitochondrial DNA
NaCl	Sodium chloride
NTC	non-template control
ng	nanogram
nm	nanometre
PCR	Polymerase Chain Reaction
pg	picogram
qPCR	quantitative PCR
RNAse	ribonuclease
RFLP	Restriction Fragment Length Polymorphism
s	second
SD	standard deviation
µl	micro-litre

Abstract

Non-invasive samples are of increasing importance for the molecular genetic analysis of free-ranging animal populations. However, the low quality and quantity of DNA contained in such extracts poses a risk for Polymerase Chain Reaction (PCR) amplification methods and associated genotyping results. Genotyping errors, such as allelic dropout and false alleles at microsatellite loci, can seriously distort biological inferences. In this study, the reliability of multi-locus microsatellite genotyping, using multiple faecal samples from alpine ibex (*Capra ibex*) was assessed. Five loci were examined in repeated genotyping experiments. The amount of nuclear DNA contained in 34 faecal extracts was determined using quantitative PCR (qPCR) and the SYBR Green fluorescence detection system, revealing comparatively high concentrations of nuclear DNA in ibex pellets. Genotyping errors were identified through direct comparison to reference blood samples, and analyses exhibited a strong dependence of erroneous microsatellite genotypes on starting template amount. The most commonly encountered genotyping errors were due to the amplification of false alleles. Variability in reliability was obvious between individual loci, resulting in a relatively high threshold value of minimum DNA amount necessary for reliable genotyping. Including a method for the discrimination of ibex pellets from samples obtained from related species, this study presents an assay for pre-selection of faecal extracts by amplifiable nuclear DNA content. Thus, molecular genetic analyses using ibex faeces are facilitated, with minimised genotyping errors and less time, effort and material expenditure, and without the risks of injuries related to capturing and handling of animals.

1 Introduction

The genetic variability of alpine ibex (*Capra ibex*) populations is currently among the lowest reported from microsatellites in mammal species (Maudet *et al.* 2002). After a severe decline in numbers throughout the 16th to 18th centuries, protective measures led to the recovery of the last ibex population (N<200 animals), remaining in the Gran Paradiso Mountains of Italy. Many re-introduction experiments throughout the 20th century resulted in the re-establishment of this species across large parts of the European Alps (acc. to Maudet *et al.* (2002): currently around 40,000 individuals). However, all animals are offspring of the rescued Gran Paradiso population, which has led to severe genetic bottlenecks among alpine ibex (Maudet *et al.* 2002). Thus, consideration of genetic aspects is a necessity for future conservation management of this caprine species. In this context, genetic identification of individuals is of prime importance, requiring the analysis of highly variable DNA sequences such as microsatellites.

However, large-scale monitoring of alpine ibex, based on conventional samples like blood or tissue is logistically difficult, as the species' habitat is usually not easily accessible. Recent developments in molecular genetics, though, seem to offer great potential to overcome these sampling constraints. Non-invasive samples, like faeces or shed hairs are increasingly recognised as a valuable source of DNA (reviewed in Waits & Paetkau 2005), allowing exploration of free-ranging animal populations without having to catch, disturb or even observe them (Taberlet & Luikart 1999). Various studies based on faecal-derived DNA have been conducted examining aspects such as population size (Kohn *et al.* 1999; Frantz *et al.* 2003; Bellemain *et al.* 2005), tracking of individuals (Ernest *et al.* 2000), species distribution (Palomares *et al.* 2002; Smith *et al.* 2006), gene flow and genetic diversity (Epps *et al.* 2005), paternity assignment and mating system (Constable *et al.* 2001; Garnier *et al.* 2001) as well as population structure (Piggott *et al.* 2006).

Despite its huge potential, molecular genetic analysis of non-invasive samples does pose several challenges and numerous factors can lead to erroneous results using this approach (for a review, see Pompanon *et al.* 2005). Genotyping errors, principally ´allelic dropout` [AD] (Navidi *et al.* 1992; Walsh *et al.* 1992; Taberlet *et al.* 1996, 1999) and ´false alleles` [FA]

(Taberlet *et al*. 1996, 1999; Goossens *et al.* 1998; Bradley & Vigilant 2002), can lead to mis-interpretation of genotypes and result in distortion of biological inferences. For example, Creel *et al.* (2003) indicated substantial overestimation of population size, and Hoffman & Amos (2005) demonstrated a strong influence of genotyping errors on paternal exclusion. Therefore, potential genotyping errors need to be examined and an unbiased quantification of error rates is necessary, in order to validate the results of microsatellite genotyping using non-invasive samples (Taberlet & Luikart 1999; Bonin *et al.* 2004; Broquet & Petit 2004; Hoffmann & Amos 2005; Pompanon *et al.* 2005).

With increasing awareness of the limitations of non-invasive approaches, different methods have been proposed to achieve more accurate data. Numerous studies have attempted to optimise methods for the collection and preservation of samples (Maudet *et al.* 2004b; Nsubuga *et al.* 2004; Roeder *et al.* 2004) as well as DNA extraction (Wehausen *et al.* 2004; Flagstad *et al.* 1999) and amplification strategies (Piggott *et al.* 2004; Bellemain & Taberlet 2004). Quality control procedures have been developed, including the application of the multiple tube approach (Navidi *et al.* 1992; Taberlet *et al*. 1996) to minimise stochastic errors and to examine reproducibility of results through independent genotyping replication. Morin *et al.* (2001) suggest a pre-screening of samples by amplifiable nuclear DNA amount, using quantitative PCR. In addition to these approaches, different methods have been suggested to avoid mis-interpretation of genotyping results, including the use of simulations (Taberlet *et al.* 1996; Valiere *et al.* 2002; Roon *et al.* 2005), the maximum-likelihood approach (Miller *et al.* 2002) and the pair-wise mismatching method (Paetkau 2003).

DNA extracted from faeces varies substantially in quality and quantity among species (Taberlet & Luikart 1999). Therefore, species specific methods need to be established to examine specific biases and deduce guidelines for deriving reliable genotyping results for the species of interest (Taberlet & Luikart 1999; Parsons 2001; Waits & Paetkau 2005; Pompanon *et al.* 2005). Studies have shown that DNA extracted from herbivore faecal samples can provide DNA suitable for molecular genetic analyses (Garnier *et al.* 2001; Huber *et al.* 2003; Epps *et al.* 2005). However, no study to date has been initiated using microsatellite genotyping from non-invasive samples of a free-ranging goat-like species (*Capra spp*.).

The overall objective of this study is to establish a feasible method allowing reliable individual identification of ibex, using faecal samples to amplify a panel of five microsatellite loci. Genotyping errors are identified through the comparison of faecal-based genotypes with those obtained from reference blood samples (similar to Ernest *et al.* 2000; Lathuilliere *et al*. 2001; Parsons 2001; Huber *et al.* 2003). This allows the calculation of error rates that are examined in relation to the amount of amplifiable nuclear DNA.

Three primary questions will be addressed:

(i) How much amplifiable nuclear DNA is present in a faecal sample of *Capra ibex*?
(ii) Are there different success rates between loci, depending on varied template concentrations?
(iii) What are minimum template amounts for reliable genotyping based on five microsatellite loci?

In addition, a final secondary question will be addressed: is it possible to differentiate defecator species, in order to exclude those pellets from analysis which originate not from ibex, but from related species?

2 Materials and Methods

2.1 Sampling and sample preservation

Blood and faecal samples were collected from seven alpine ibex (*Capra ibex*), of which five were captive animals from the Hellabrunn Zoo, Munich, Germany, and two which were caught in the Swiss National Park. Faecal samples, consisting of five pellets for six and four pellets for one ibex, were removed manually from the colon of the animals, in order to avoid cross-contamination between individuals. Consequently, each faecal sample could be linked to the appropriate reference blood sample. Blood samples were collected in EDTA vacutainers, while faecal pellets were sampled into plastic bags. All faecal samples were preserved dry and stored at -20°C (between four and fifty months) prior to DNA extraction.

2.2 Laboratory procedures

2.2.1 DNA extraction from blood samples

For DNA extraction from blood samples, three methods were tested: a conventional phenol-chloroform extraction method according to Hogan *et al.* (1986), the NucleoSpin® Blood Kit and finally, the NucleoSpin® Blood QuickPure Kit (Macherey-Nagel, Germany). As the latter produced the best results, all blood samples reported in this study were processed using this kit. Extractions were carried out for not more than three blood samples at once, together with one negative control containing no blood. Purification was performed according to the manufacturer's instructions, with the following modifications: besides adding 20µl RNAse to the first lysis step, columns containing precipitated DNA were dried for 10 minutes at room temperature to volatilise the ethanol. DNA was then recovered in a 50µl pre-warmed elution buffer. The extraction step was repeated for each sample to achieve enough solution for subsequent experiments. The DNA content was measured using a spectrophotometer (A_{260}nm). For genotyping, DNA solutions were adjusted to 10ng/µl, aliquoted and stored at -20°C before preparation for further analysis.

2.2.2 DNA extraction from faecal samples

Faecal samples were processed with QIAamp® DNA Stool Mini Kit (Qiagen GmbH, Germany). Single pellets were treated as individual samples and were entirely used in the extraction. All pellets available for an individual were used (four or five pellets for each animal respectively). No more than five samples and one non-template control (NTC) without faeces were handled in one extraction to allow for fast processing of DNA purification.

The manufacturer's instructions for DNA extraction were modified as follows: one whole pellet was processed, weighing from 80mg up to 570mg (for details, see Appendix I). Each frozen pellet was washed in 1.6ml of ASL buffer and vortexed for 1 minute to release shed epithelial cells of the intestinal lining. 1.4ml of the supernatant was equally portioned between two tubes and mixed with 350µl 5M NaCl. Half an InhibitEx tablet was added to each tube and processed according to the manufacturer's instructions. Both DNA isolations were combined in one column, washed and dried for ten minutes at room temperature. DNA was recovered in 100µl elution buffer, aliquoted and stored at -20°C before preparation for further analysis.

2.2.3 Species differentiation

A combination of the polymerase chain reaction (PCR) with restriction fragment length polymorphism analysis (RFLP) was used for species determination. A set of universal primers amplifying a fragment of the conserved region of the vertebrate mitochondrial cytochrome b (cyt b) gene (Kocher *et al.* 1989) was used, with forward primer (L 14841): 5′-AAA AAG CTT CCA TCC AAC ATC TCA GCA GGA TGA AA and reverse primer (H 15149): 5′-AAA CTG CAG CCC CTC AGA ATG ATA TTT GTC CTC A-3`. Amplification of the 375bp cyt b fragment was performed in a total reaction volume of 25µl, containing 2mM $MgCl_2$ (Solis BioDyne, Inc., Estland), 0.2mM dNTPs, 1x PCR buffer (Invitrogen GmbH, Germany: 20mM Tris-HCl (pH 8.4), 50mM KCl), 0.2µM of each primer, 0.04µg bovine serum albumin (BSA, Fermentas GmbH, Germany), 0.04 U Taq Polymerase (FIREPol®, Solis BioDyne, Inc., Estland) and 2µl DNA extracted from blood or faeces, respectively. DNA samples from alpine ibex (*Capra ibex*), chamois (*Rupicapra rupicapra*), goat (*Capra hircus*) and red deer (*Cervus elaphus*) were analysed. PCR amplification conditions on a UNO II thermal cycler (Biometra) were as follows: initial denaturation at 94°C for 3 min,

followed by 35 cycles of 30s at 94°C, 30s at 55°C, 30s at 72°C, and a final elongation step at 72°C for 3 minutes. Two NTC were included in every experiment.

Sequence information on the cyt b gene of the four species tested was obtained from the NCBI database (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db). Sites for a restriction enzyme cleaving the 375bp fragment were identified using the programme NEB Cutter Version 2.0 (Vincze *et al.* 2003) and are displayed in Fig. 1, along with alignment data of the expected cyt b regions. The amplified cyt b fragment was digested using restriction endonuclease *Alu*I (target sequence: AG↓CT), because this enzyme was found to be discriminatory for ibex (93/45/237 bp) versus goat (-), chamois (-) and red deer (-) DNA. Endonuclease digestion was carried out in 25µl digestion volumes, containing 18µl of PCR product, 1x restriction buffer (*TANGO,* Fermentas, Germany) and 0.8 U of restriction enzyme. Incubation was conducted according to the manufacturer's recommendations (Fermentas, Germany): a 3 hour incubation at 37°C was followed by an enzyme inactivating step at 65°C for 20 min.

```
                    *        20         *        40         *        60         *
Ca_AF03473 : ttccaaccccaccaaatatctcatcatgatgaaactttggatccctcctaggcacttgcctaattttacagatcc :  75
Ch_X56289  : tcccaaccccatcaaacatctcatcatgatgaaactttggatccctcctaggaatttgcctaatcttacaaatcc :  75
Ce_AJ00002 : tcccagccccatcaaatatttcatcctgatgaaatttcggctcattactaggagtctgtctaatcctacaaatcc :  75
Ru_AF03472 : tcccagctccatcaaacatctcatcatgatgaaactttggctccctcttaggcatctgcttaattttacagatcc :  75

                  80         *       100         *       120         *       140         *
Ca_AF03473 : tgacaggcctattcctagctatacactacacatctgacacaataacagcattctcctctgtagctcacatttgtc : 150
Ch_X56289  : taacaggcctattcctagcaatacactatacatccgacacaataacagcattttcctctgtaactcacatttgtc : 150
Ce_AJ00002 : tcacaggcctattcctagcgatacactatacatctgatacaataacagcattctcctctgtcacccatatctgtc : 150
Ru_AF03472 : taacaggcctattcctagcaatacactacacatccgatacagcaacagcattctcctctgtaacccacatttgcc : 150

                 160         *       180         *       200         *       220
Ca_AF03473 : gagatgtaaattacggctgaattatccgatacatacacgcaaacggagcatcaatattctttatctgcctattta : 225
Ch_X56289  : gagatgtaaattatggctgaattatccgatacatacacgcaaacggagcatcaatattctttatctgcctattca : 225
Ce_AJ00002 : gagatgtcaattatggctgaattattcgatatatacacgcaaacggggcatcaatatttttcatctgtctattca : 225
Ru_AF03472 : gagatgtaaactacggctgaatcatccgatacatacatgcaaatggagcatcaatatttttcatctgcctattca : 225

                   *       240         *       260         *       280         *       300
Ca_AF03473 : tgcatgtcggacgaggtctatattatggatcatatacctttctagaaacatgaaatattggagtaatcctcctat : 300
Ch_X56289  : tacatatcggacgaggtctatattatggatcatatacctttctagaaacatgaaacattggagtaatcctcctgc : 300
Ce_AJ00002 : tacatgtagggcgaggcctgtactacggatcatatacttttctagagacgtgaaacatcggagtagttcttctat : 300
Ru_AF03472 : tacatgtaggacgaggcctatattacggatcatacacttttctagaaacatgaaacatcggagtaatcctcctac : 300

                    *       320         *       340         *       360         *
Ca_AF03473 : tcgcaacaatagccacagcattcataggctacgtcctaccatgaggacaaatatcattctgaggggcaacagtca : 375
Ch_X56289  : tcgcaacaatggctacagcattcataggctatgttttaccatgaggacaaatatcattttgaggggcaacagtca : 375
Ce_AJ00002 : ttacagttatagccacagcattcgtaggatatgtcctaccatgaggacaaatatcattctgaggagcaacagtca : 375
Ru_AF03472 : tcacaacaatagccacagcgtttatgggctacgtcctaccatgaggacagatatcattctggggagcaacagtta : 375
```

Fig. 1 Sequences of a 375bp cyt b fragment of alpine ibex (Ca*: Capra ibex,*), aligned with sequence data from three species of the suborder Ruminantia (Ch: *Capra hircus*; Ce: *Cervus elaphus*; Ru: *Rupicapra rupicapra*). The corresponding accession numbers are given according to the NCBI database. Positions of restriction endonuclease *Alu*I that cleaves ibex DNA are highlighted in black.

2.2.4 Microsatellite amplification

In a preliminary screening, a panel of 27 primer pairs was tested in *Capra ibex* (for details, see Appendix II). Five primer pairs, showing the best results in amplification of faecal-derived DNA, were selected for genotyping: SR-CRSP 1, SR-CRSP 5, ETH 225, TGLA 122 and ETH 10. All markers used in this study have already been tested with blood-derived DNA from alpine ibex (Maudet *et al.* 2002, 2004a), and the latter two have been found to be some of the most polymorphic microsatellites in alpine populations (Maudet *et al.* 2002). For each microsatellite, the forward primer was labelled with fluorescent 6-FAM to allow detection on the ABI Prism 377 automated sequencer. Each of the five markers was amplified three times in independent PCR reactions.

Table 1 Characteristics of five primer pairs used for amplifications of microsatellite loci in *Capra ibex*. The forward (F) primer was labelled with fluorescent 6-FAM. All loci are dinucleotide.

Locus	Primer sequence	Product size range (bp)	Annealing temp. °C	Primer reference	Transferred on *Capra ibex*
ETH 10	F: GTTCAGGACTGGCCCTGCTAACA R: CCTCCAGCCCACTTTCTCTTCTC	211-213	60	a	Maudet *et al.* 2002
ETH 225	F: GATCACCTTGCCACTATTTCC T R: ACATGACAGCCAGCTGCTACT	148	60	a	Maudet *et al.* 2004a
SR-CRSP 1	F: TGCAAGAAGTTT TTCCAGAGC R: ACCCTGGTTTCACAAAAGG	134-136	60	b	Maudet *et al.* 2004a
SR-CRSP 5	F: GGACTCTACCAACTGAGCTACAAG R: TGAAATGAAGCTAAAGCAATGC	171	60	b	Maudet *et al.* 2004a
TGLA 122	F: CCCTCCTCCAGGTAAATCAGC R: ATCACATGGCAAATAAGTACATAC	137-155	60	c	Maudet *et al.* 2002

(a) Kappes *et al.* 1997; (b) Arevalo *et al*. 1994; (c) Bishop *et al.* 1994.

Individual PCR amplifications were performed in a total reaction volume of 15µl containing either 2µl (20ng DNA per reaction volume) of blood-derived or 2.4µl of faecal-derived DNA, 0.4µM of each primer, 0.2mM dNTPs, 1x-PCR buffer (Invitrogen GmbH, Germany: 20mM Tris-HCl (pH 8.4), 50mM KCl), 3mM $MgCl_2$ (Solis BioDyne, Inc., Estland), 0.04µg bovine serum albumin (BSA, Fermentas GmbH, Germany) and 0.04 U of Taq-Polymerase (FIREPol®, Solis BioDyne, Inc., Estland). PCR amplification conditions on a UNO II thermal cycler (Biometra, Germany) were as follows: initial denaturation at 94°C for 3 min,

followed by 40 cycles of 30s at 94°C, 30s at the primer specific annealing temperature (Table 1), 30s at 72°C, and a final elongation step at 72°C for 3 minutes.

Two NTC were included in every polymerase chain reaction to monitor contamination of reagents. Additionally, all sets of faecal-derived DNA amplifications included one positive control containing DNA extracted from blood. Agarose gel electrophoresis and ethidium bromide staining were used to investigate whether the reactions were successful.

2.2.5 Genotyping of blood- and faeces-derived DNA extracts

Microsatellite genotyping was conducted on 6% polyacrylamide gels (0.2mm) using an ABI Prism 377 automated sequencer (Perkin Elmer), with analysis using GENESCAN® 3.1.2 and GENOTYPER® 2.5 (Applied Biosystems, Foster City, CA, USA).
DNA extracted from blood was genotyped once for every microsatellite, whereas each of the multiple faeces-derived extracts of all seven animals was examined three times per locus. PCR products were diluted depending on the band patterns visualised on the agarose gel image. Alleles were sized relative to an internal size standard (GS 350-500 with ROX label; Applied Biosystems, Foster City, CA, USA). 2.4µl of diluted PCR product was mixed with 1.7µl of the internal standard mix (1.5µl formamide loading buffer and 0.3µl ROX GS 350-500), heated at 94°C for three minutes and flash cooled on ice. Prior to loading PCR products, gels were pre-run at 3 KV for one hour to reduce electrophoresis artefacts (Fernando *et al.* 2001). In order to check for shifting in measured allele lengths due to gel irregularities, one positive control from a blood-derived DNA sample with known size for the locus examined was separated on the gel in every genotyping experiment.

Allele sizing was conducted employing the default settings in GENESCAN® 3.1.2, including a peak amplitude threshold of 50. Results were corrected manually only if more than two alleles had been marked. In that case, the two peaks with the highest signal intensities were chosen, of which the first displayed the higher signal. Basepairs that differed ± 0.99 were concluded to be the same allele. For details of allele scoring, see Table A3 in Appendix III. In the rare case that the internal size standard was not usable in the analysis, the PCR product was genotyped again.

2.2.6 Quantitative PCR

The quantitative PCR (qPCR) targeted a 133-bp fragment of the protein kinase c iota gene from *Capra ibex* (Ropiquet & Hassanin 2005, corresponding GenBank accession number: AY846798). The marker was designed using PRIMER3 software (Whitehead Institute for Biomedical Research, Cambridge, MA, USA; http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi), with the following oligonucleotide sequences: forward (RupruSTD1-1) 5`-TCA AGG GAG AGT ACT TGA GC-3` and reverse (RupruSTD1-2) 5´-AAC TGG GCT AAA CAT CAA AA-3`. To evaluate the success of amplification, ibex PCR products were sequenced and found to be complementary to the excerpt between the 273 and 405bp part of the target gene.

Quantitative PCRs were conducted on a LightCycler (Roche Diagnostics GmbH, Mannheim, Germany), using the following conditions: an initial denaturation step at 95°C for ten minutes was followed by 45 amplification-cycles (15s at 95°C, 10s at 60°C, 20s at 72°C), with a single acquisition of fluorescence at the end of each cycle. PCR products were melted in a final temperature slope (0.1°C/s) from 65°C (15s) up to 99°C with continuous measurement of the fluorescence. The final step was at 40°C for 30s. Measurement of fluorescence levels was conducted using the SYBR Green detection system.

Amplifications were performed in a total reaction volume of 10µl containing 1µl of DNA, 0.6µM of each primer, 0.04µg BSA (Fermentas) and 1x LC Fast Start DNA Master^{PLUS} (Roche Diagnostics GmbH, Germany; containing $MgCl_2$, dNTPs, HotStartTaq Polymerase and SYBR Green). One NTC was included in every experiment.

Eleven dilutions were created using one blood-derived DNA sample from ibex. The DNA content was first quantified by absorbance (A_{260}nm) in a spectrophotometer. Using these standard samples three independent qPCR experiments were conducted. Observed C_t values were averaged and plotted against the log of the DNA concentrations in order to create a standard curve. DNA concentrations in the standards were: 10ng/µl, 7.5ng/µl, 5ng/µl, 2.5ng/µl, 1ng/µl, 750pg/µl, 500pg/µl, 250pg/µl, 100pg/µl, 50pg/µl and 10pg/µl.

Quantities of faecal-derived DNA were examined using the standard curves and calculations in Microsoft Excel spreadsheets. According to Morin *et al.* (2001), DNA concentrations were calculated from the C_t value of the sample, and slope and *Y*-intercept (*Y* int) of the trend line from the standard curve, plotted as the log of DNA concentration versus averaged C_t values:

$$\text{DNA concentration (pg/µl)} = 10^{\left(\frac{Ct - Y\,\text{int}}{slope}\right)}.$$

To check for irregularities occurring during amplification process, four positive controls of known DNA concentrations (derived from the three qPCRs conducted with the standard dilutions) were included in every run. Examination of qPCR specificity was performed using LightCycler®3 software and melting curve analysis (Ririe *et al.* 1997). Additionally, agarose gel electrophoresis was performed to monitor for primer-oligomers.

2.2.7 Contamination control

In order to avoid cross-contamination, extraction, amplification and genotyping processes were performed in physically separated rooms. In every step, disposable gloves and sterile filter tips were used, and the PCR bench was decontaminated by UV irradiation before each PCR was conducted. At least one negative control was used for extraction (containing no blood or faeces) and amplification (containing no DNA) to check for contamination, either by PCR products or by genomic DNA. Each negative control of every DNA extraction was checked for contamination in PCR amplification (Taberlet & Luikart 1999).

2.3 Data Analyses

2.3.1 Quantifying error rates

In order to examine the extent of amplification failure and different genotyping errors, the proportion of positive PCRs (i.e. PCRs with evident amplification) was calculated, see eqn. (1) and (8), respectively. Within this, the proportion of allelic dropout (i.e. one of two alleles missing in heterozygous loci), see eqn. (2) and (9), and the proportion of false alleles (i.e. allele size >1 base pair (bp) different from correct allele size), see eqn. (4), (6) and (10), were calculated. These values were examined for (a), each locus and (b), each individual sample, over all loci among three genotyping replications. Additionally, the mean error rate per locus, see eqn. (7), along with ratios of AD and FA over all loci were calculated, see eqn. (3) and

(5), respectively. Examinations of error rates were conducted according to Broquet & Petit (2004) and Pompanon *et al.* (2005).

(a) Quantifying error rates per locus

To examine and depict differences in reliability between loci, success rates and genotyping error rates were calculated for five individual microsatellite loci. Per single locus, a maximum of three genotyping replications of 34 faecal extracts was included in the analyses (34 x 3 = 102 PCR experiments).

I. Proportion of positive PCR per single locus

The proportion of positive PCR **per locus *j*** (*Aj*) can be defined as

$$Aj = \frac{Bj}{Nj}, \qquad (1)$$

the ratio of the number of positive PCR at locus *j* (*Bj;* number of PCR with evident amplification) on the total number of PCR at locus *j* (*Nj*).

(after Broquet & Petit 2004)

II. Frequency of genotyping errors due to allelic dropout (AD) per single locus

The proportion of erroneous genotypes due to mis-amplification of one allele (AD) **per locus *j*** (*Pj*) can be defined as

$$Pj = \frac{Dj}{A_{hetj}}, \qquad (2)$$

the ratio of the number of observed AD at locus *j* (*Dj*; number of amplifications involving the loss of one allele) on the number of positive amplifications of individuals determined as heterozygous according to their reference blood genotype at locus *j* (A_{hetj}).

(acc. to Broquet & Petit 2004)

The ratio of observed AD **over all loci L** (*pL*) is then

$$pL = \frac{\sum_{j=1}^{L} Dj}{\sum_{j=1}^{L} A_{hetj}}, \qquad (3)$$

the ratio of the total number of observed AD over L loci (*Dj*) on the total number of successful amplifications of heterozygous genotypes among L loci (A_{hetj}).
(acc. to Broquet & Petit 2004)

III. Frequency of genotyping errors due to false alleles (FA) per single locus

The proportion of erroneous genotypes due to the occurrence of false alleles (FA) **per locus *j*** (*fj*) can be defined as

$$fj = \frac{Fj}{Aj}, \quad (4)$$

the ratio between the number of amplifications leading to the creation of one or two false alleles at locus *j* (*Fj*) on the number of positive PCRs for locus *j* (*Aj*) (of either homozygous or heterozygous samples).
(acc. to Broquet & Petit 2004)

The overall FA rate **over all loci L** (*fL*) is then:

$$fL = \frac{\sum_{j=1}^{L} Fj}{\sum_{j=1}^{L} Aj}, \quad (5)$$

the ratio of the total number of FA among L loci (*Fj*) on the total number of successful amplifications among L loci (*Aj*).
(acc. to Broquet & Petit 2004)

IV. Proportion of false alleles per replicated alleles and single locus

The proportion of false alleles (FA) per locus *j* was additionally calculated **per replicated allele** and locus *j* (fj_a). For diploid individuals, two alleles are typed in each successful amplification of locus *j* (Pompanon *et al.* 2005), of which the first can be the same or different from the second. Of these alleles, one or both can differ >1 bp from the correct genotype (= FA).

Hence, the proportion of FA per replicated alleles and locus *j* (fj_a) can be defined as

$$fj_a = \frac{NFj}{2Aj}, \quad (6)$$

the ratio between *NFj*, i.e. the number of false alleles counted among three independent genotyping experiments over all samples (heterozygous or homozygous genotypes) and locus *j* on *2Aj*, i.e. the number of replicated alleles in successful amplifications.
(acc. to Pompanon *et al.* 2005)

V. Proportion of repeatedly occurring false alleles per single locus

Within fj_a the proportion of repeatedly occurring false alleles (i.e. alleles that were scored at least twice among all extracts and three genotyping results) was calculated **per locus**.

VI. Mean error rate per locus

The mean error rate **per locus** (*mj*) can be defined as

$$mj = \frac{Mj}{Aj}, \qquad (7)$$

the ratio between *Mj*, i.e. the number of single-locus genotypes at locus *j*, including at least one allelic mismatch on the total number of positive PCRs at locus *j* (*Aj*).
(acc. to Pompanon *et al.* 2005)

(b) Quantifying error rates per individual sample over all loci

To examine and depict the dependence between template DNA amount and the occurrence of genotyping errors, the proportion of positive PCRs, allelic dropout and false alleles was additionally calculated per individual sample over three replications of five loci (3 x 5 = 15 PCR experiments).

I. Proportion of positive PCR per individual sample *s* over all loci *l* (Al_s)

The proportion of positive PCR experiments **per individual sample *s* over all loci *l*** (Al_s) can be defined as

$$Al_s = \frac{Bl_s}{Nl_s}, \qquad (8)$$

the ratio of the number of positive PCR over all loci *l* for individual sample *s* (Bl_s) on the total number of reactions over all loci *l* for *s* (Nl_s).

II. Frequency of genotyping errors due to allelic dropout (AD) per individual sample over all loci

The proportion of erroneous genotypes due to mis-amplification of one allele (AD) **per individual sample *s* over all loci *l*** (Pl_s) can be defined as

$$Pl_s = \frac{Dl_s}{Al_{hets}}, \quad (9)$$

the ratio of the number of observed AD over all loci *l* for sample *s* (Dl_e; number of amplifications involving the loss of one allele over all loci *l* for sample *s*) on the number of positive amplifications of microsatellite loci determined as heterozygous for sample *s* according to reference blood genotype (Al_{hets}).

III. Frequency of genotyping errors due to false alleles (FA) per individual sample

The proportion of erroneous genotypes due to occurrence of false alleles (FA) **per individual sample *s* over all loci *l*** (fl_s) can be defined as

$$fls = \frac{Fl_s}{Bl_s}, \quad (10)$$

the ratio of the number of false alleles among all loci *l* for sample *s* (Fl_s) on the number of positive PCRs among all loci and sample *s* (Bl_s).

2.3.2 Classification of samples

In order to compare the success rate and accuracy of microsatellite amplification with the DNA template amount, three genotyping results of each extract for individual loci were classified into four categories: (0) three failed amplifications (i.e. no evident amplification) or three incorrect genotypes; (1) one correct genotype and two failed amplifications or one correct genotype and two incorrect genotypes; (2) two correct genotypes and one failed amplification or two correct genotypes and one incorrect genotype; (3) three correct genotypes. Categorized results were examined using STATISTICA® (StatSoft, Europe). Lower and upper quartiles as well as medians were calculated for each error category and locus. Error categories are plotted against log template concentration in pg/µl.

3 Results

3.1 Species differentiation

RFLP analysis was conducted to determine defecator species, using PCR products obtained from blood-derived DNA of goat (*Capra hircus)*, red deer (*Cervus elaphus)*, chamois (*Rupicapra rupicapra*) and alpine ibex (*Capra ibex*). To examine the applicability of the approach to faecal samples, additional experiments were performed using three faecal extracts from chamois and ibex. PCR amplified fragments (375bp) of the mitochondrial cytochrome b gene were digested using restriction endonuclease *Alu*I (Fig. 2). Exclusively ibex amplicons were cleaved, resulting in two specific band patterns (237/93 bp). The third fragment, with an expected size of 45bp, could not be visualised due to its short length. Because of partial digestion of the 375bp fragment, a band pattern with the original fragment size is still present after restriction in ibex amplicons.

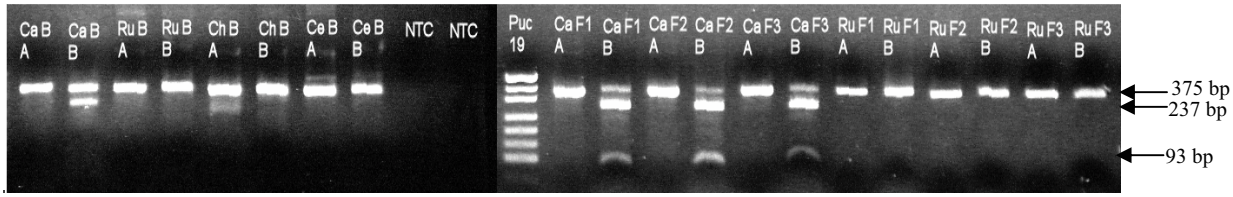


Fig. 2 RFLP analysis of PCR amplified fragments of cytochrome b from different species of the suborder Ruminantia. (A) PCR amplification resulted in a 375bp fragment. (B) Fragments digested using *Alu*I restriction enzyme. From left to right: blood-derived amplicon from alpine ibex (Ca B); chamois (Ru B); goat (Ch B); red deer (Ce B); NTC: non-template control; Puc 19: size marker; faecal-derived amplicon from alpine ibex (Ca F1-3); chamois (Ru F1-3).

3.2 DNA quantification

In order to quantify the concentration of nuclear DNA extracted from faecal pellets, a set of standard samples was measured in three independent qPCR experiments using the SYBR Green detection system. All C_t values obtained in these three experiments were used to calculate average C_t values. Plotted against log DNA concentration, these values produced a standard curve with a trend line correlation coefficient of 0.9949 (Fig. 3).

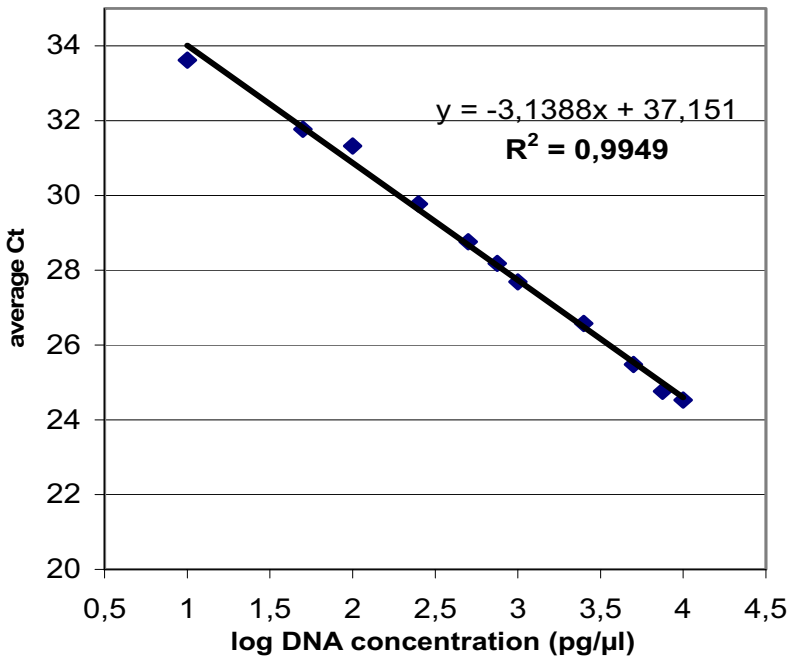


Fig. 3 Plot of the average C_t values for each DNA concentration from three standard curves, calculated from three independent qPCR experiments using primer RupruSTD1.

Standard deviations of individual C_t values for each DNA concentration of the three qPCR experiments are displayed in Table 2. Additionally, average values for DNA concentrations are shown, calculated from the trend line equation from the average standard curve.

The amount of amplifiable nuclear DNA present in 34 faecal samples was measured via qPCR amplification. Results of template quantification calculated with measured C_t values and the averaged C_t values of the three standard curves (displayed in Fig. 3) are shown in Appendix I. Median template amount was 593pg/µl (average template amount: 908pg/µl), ranging from approximately 3pg/µl up to 7.6ng/µl. High variance in DNA concentration

between individual animal extracts could be observed (33-7500pg/µl difference in DNA amount between extracts of the same individual).

Table 2 Measured C_t values and standard deviations from three independent qPCR experiments using blood-derived DNA, summarized in Fig. 3. Calculated DNA concentrations are based on the trend line equation in Fig. 3.

DNA amount		**C_t values**		**Calculated DNA amount**
pg/µl	Log (pg/µl)	Average	SD	Average (pg/µl)
10000	4.000	24.53	0.20	10490
7500	3.875	24.76	0.05	8840
5000	3.699	25.48	0.06	5225
2500	3.398	26.58	0.19	2337
1000	3.000	27.69	0.04	1035
750	2.875	28.18	0.16	719
500	2.699	28.76	0.13	471
250	2.398	29.78	0.20	223
100	2.000	31.32	0.76	72
50	1.699	31.77	0.56	52
10	1.000	33.62	1.66	13

Distribution of DNA concentration in faecal extracts (N=34) is shown in Fig. 4. More than 60% of all extracts exceeded 50pg/µl (N=22), of which 50% contained more than 1ng/µl (N=11). Three samples were measured as containing <0.3pg/µl (8.8%), of which two pellets had disintegrated completely during the first step of extraction.

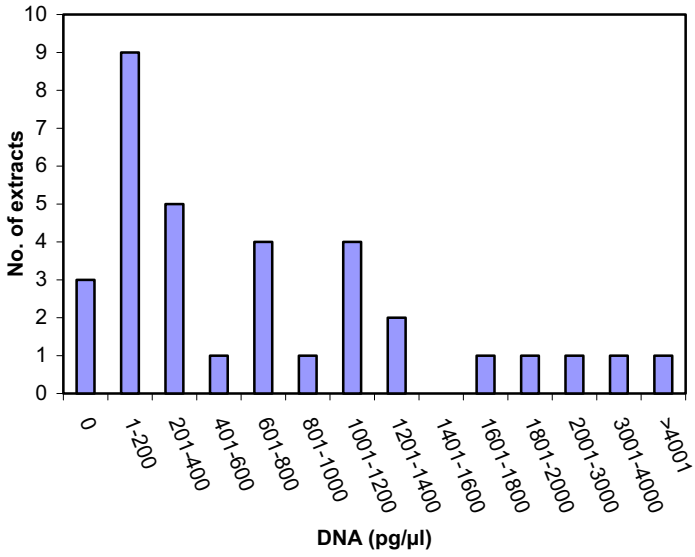


Fig. 4 Distribution of DNA concentration in 34 extracts from ibex faeces. All extracts contained a total volume of 100µl.

For each qPCR experiment, the melting curves of all extracts were examined. Denaturation of double-stranded amplicons, leading to an immediate decrease of fluorescence levels, occurred at 76°C for both blood and faecal extracts (Fig. 5). However, one faecal extract exhibited a melting peak at a lower temperature (74.5°C, highlighted with a red arrow in Fig. 5b). Lines displaying no peak depict NTC that produced no amplicon (Fig. 5a and b) or faecal samples that were measured as zero, respectively (Fig. 5b).

a) blood-derived DNA b) faecal-derived DNA

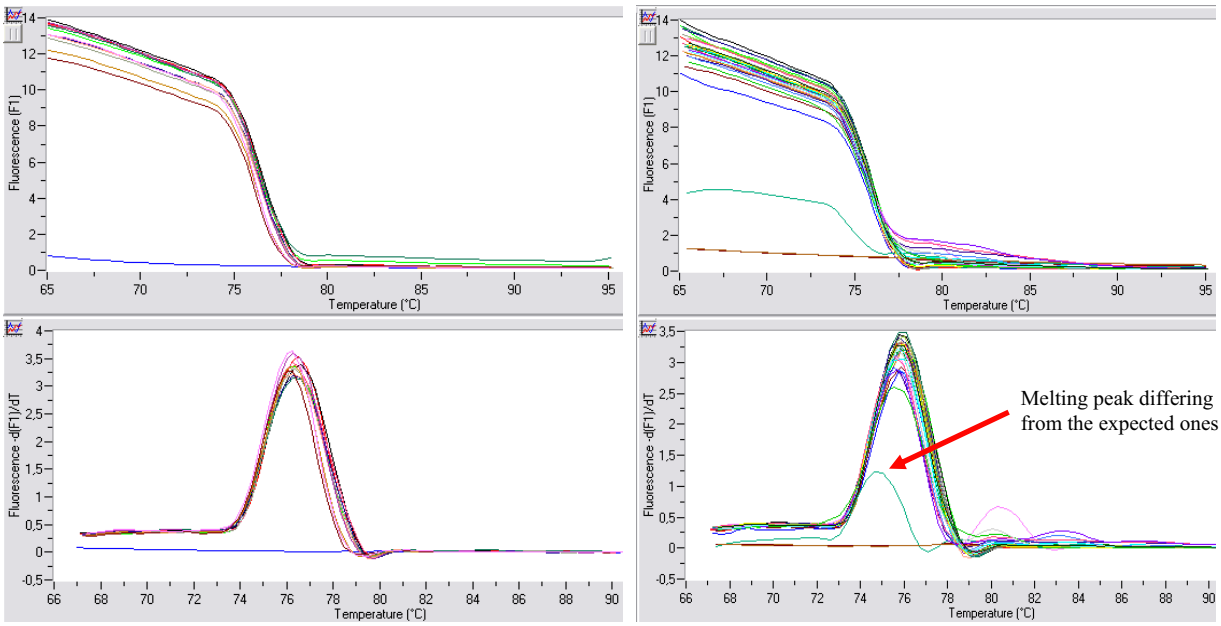


Fig. 5 Melting curves of a qPCR experiment conducted with a) the standard samples derived from ibex blood and b) multiple ibex faecal extracts along with four positive controls with blood-derived DNA

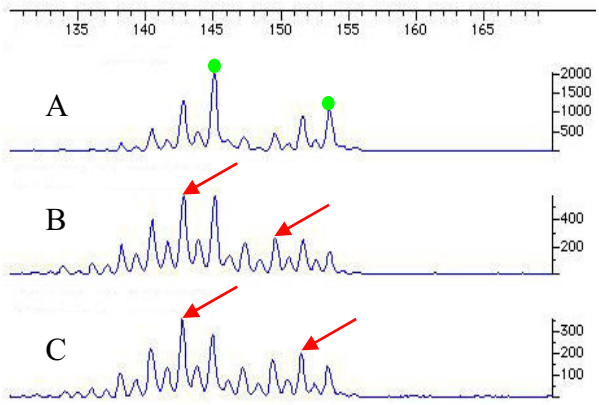
3.3 Data Analyses

3.3.1 Quantifying error rates

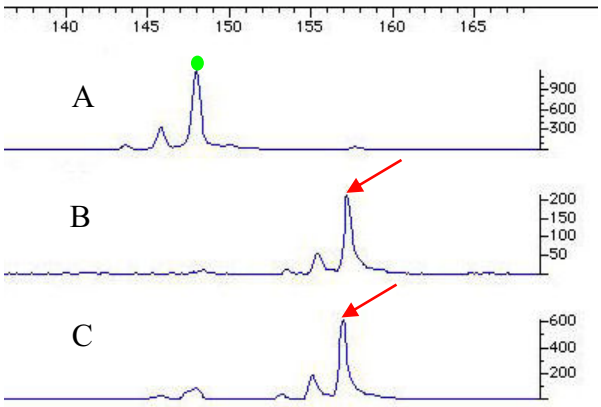
The success rate and accuracy of microsatellite genotyping was examined for five loci. All faeces-based genotypes used in this analysis were derived from multiple DNA extracts for each individual: processing of four to five pellets per animal gave a total of 34 extractions, and three amplifications from each extract over five loci produced a total of 510 (34 x 3 x 5) PCRs. Through direct comparison to genotypes determined from reference blood samples, genotyping errors were identified. The extent of amplification success and resulting error rates was calculated per locus (102 PCR experiments) and per individual sample (15 PCR

experiments), respectively (for details, see Materials and Methods section). Fig. 6 displays examples of electropherogram patterns observed during genotyping experiments, showing two types of false alleles and allelic dropout as genotyping errors.

a) Misscoring



b) False alleles



c) Allelic dropout

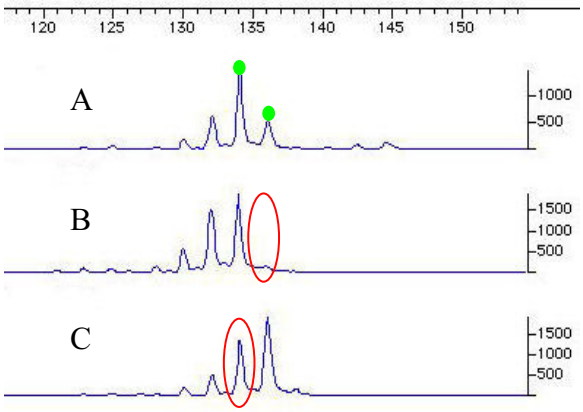


Fig. 6 Electropherogram patterns of microsatellite loci a) TGLA 122, b) ETH 225 and c) SR-CRSP 1. Lanes marked 'A' depict true alleles, derived from extracts containing high amounts of nuclear DNA, highlighted with green dots. Lanes 'B' and 'C' display erroneous genotypes derived from low quality faecal extracts. False alleles are displayed for TGLA 122 and ETH 225, whereas allelic dropout is shown using locus SR-CRSP 1 as an example. The horizontal scale represents the number of bp, while the vertical scale expresses the number of fluorescence units.

(a) Quantifying error rates per locus

The extent of successful amplification, allelic dropout and false alleles along with mean error rate per individual locus is shown in Table 3. Note that the proportion of successful amplifications is expressed relative to the number of PCRs for the replicates over each locus, whereas allelic dropout is depicted relative to the number of successful reactions for heterozygous genotypes. The rate of genotyping errors due to one or two false alleles is calculated on the number of successful reactions for both heterozygous and homozygous genotypes. Additionally, total numbers of false alleles were examined relative to the number of alleles in successful amplifications. Within those, the proportion of false alleles that occurred repeatedly (i.e. false alleles that were scored at least twice among all extracts and three genotyping replications per locus) was calculated. Mean error rate per locus (calculated for both homozygous and heterozygous loci) includes occurrence of AD and FA.

Table 3 Microsatellite loci examined, along with calculated extents of successful amplification, allelic dropout, false alleles, proportions of reproducible false alleles and mean error rate per locus.

Locus	Rate of successful amplification	Rate of genotyping error due to AD	Rate of genotyping error due to FA	Proportion of FA per replicated alleles	Proportion of repeatedly occurring FA	mean error rate per locus
ETH 10	**95.1%** (97/102)	**0%** (0/73)	**28.9%** (28/97)	**27.8%** (54/194)	**88.9%** (48/54)	**28.9%** (28/97)
ETH 225	**98%** (100/102)	**0%** (0/0)	**11%** (11/100)	**7.5%** (15/200)	**93.3%** (14/15)	**11%** (11/100)
SR-CRSP 1	**93.1%** (95/102)	**7.4%** (7/95)	**4.2%** (4/95)	**3.7%** (7/190)	**85.7%** (6/7)	**11.6 %** (11/95)
SR-CRSP 5	**96.1%** (98/102)	**0%** (0/0)	**3.1%** (3/98)	**2.6%** (5/196)	**60%** (3/5)	**3.1%** (3/98)
TGLA 122	**91.2%** (93/102)	**4.4%** (3/68)	**11.8%** (11/93)	**8.1%** (15/186)	**86.7%** (13/15)	**15.1%** (14/93)
TOTAL	**94.7%** (483/510)	**4.2%** (10/236)	**11.8%** (57/483)	**9.9%** (96/966)	**87.5%** (84/96)	**13.9%** (67/483)

The rate of amplification success varied between loci, ranging from 98% (ETH 225) to less than 92% (TGLA 122). In total, 94.7% of 510 individual PCRs showed evident amplification (N=483).

All animals showed heterozygous genotypes on locus SR-CRSP 1. Loci ETH 10 and TGLA 122 were heterozygous in five ibex. The overall extent of allelic dropout at these

heterozygous loci averaged 4.2%, as calculated from 236 positive reactions. Proportions ranged from 0% (ETH 10) up to 7.4% (SR-CRSP 1). For markers ETH 225 and SR-CRSP 5 the proportion of allelic dropout cannot be indicated, as each animal was found to be homozygous on those microsatellite loci.

Genotyping errors that were caused by the occurrence of false alleles were scored at all microsatellite loci, but to substantially different extents. Locus SR-CRSP 5 displayed the lowest rate (3.1%), whereas FA occurred in almost 30% of genotypes for ETH 10. Among all loci, 57 of 483 single-locus genotypes exhibited at least one allelic mismatch, resulting in an overall FA genotyping error rate of 11.8%.
Calculations based on the allelic level show variation in total numbers of false alleles between individual loci, with locus SR-CRSP 5 showing the lowest (3 out of 190 replicated alleles) and ETH 10 displaying the highest (54 out of 194 replicated alleles) numbers of false alleles observed among successful amplifications. Among 483 positive PCRs over all loci that produced 966 alleles in total, 96 alleles were scored as different from true alleles (9.9%). On average, 87.5% of these 96 false alleles occurred at least twice among all extracts and three genotyping replications per individual locus.

Mean error rates per locus ranged from 3% (SR-CRSP 5) up to almost 29% (ETH 10), resulting in a multi-locus error rate of 13.9% for these five microsatellite loci and the set of 34 samples. It should be noted that the calculation of mean error rate per locus can slightly underestimate the true error rate, as allelic dropout can only be detected at heterozygous loci. An unbiased overall genotyping error rate can therefore exclusively be given for SR-CRSP 1, where A_{hetj} corresponds to A_j (for details, see Materials and Methods section).

(b) Quantifying error rates per individual sample over all loci

The frequency of successful amplification, allelic dropout and false alleles was examined for each individual sample (N=34) among three replications of five loci. Results are displayed in Fig. 7, plotted against starting amounts of DNA per µl reaction volume. Three faecal extracts, which led to >300pg starting template per µl reaction mix, are not shown. For these samples, all amplifications were successful and no genotyping errors were observed.

A substantial number of successful PCRs were obtained from extracts containing very low amounts of template (Fig. 8): 58.3% of reactions starting with a DNA concentration of less than 10pg/µl showed evident amplification in all fifteen experiments. However, false alleles were exhibited in at least one (6.7%) and up to six (40%) of these positive reactions per sample. In total, all successful amplifications that started with less than 10pg DNA per µl exhibited 33.5% erroneous genotypes due to false alleles. Moreover, the extent of allelic dropout increased rapidly with a decrease in DNA concentration: 50% of the samples containing less than 10pg template per µl exhibited allelic dropout at least once among heterozygous loci. Hence, a frequency of 12.2% false homozygotes could be observed in successful amplifications of heterozygous loci with initial DNA concentrations lower than 10pg/µl.

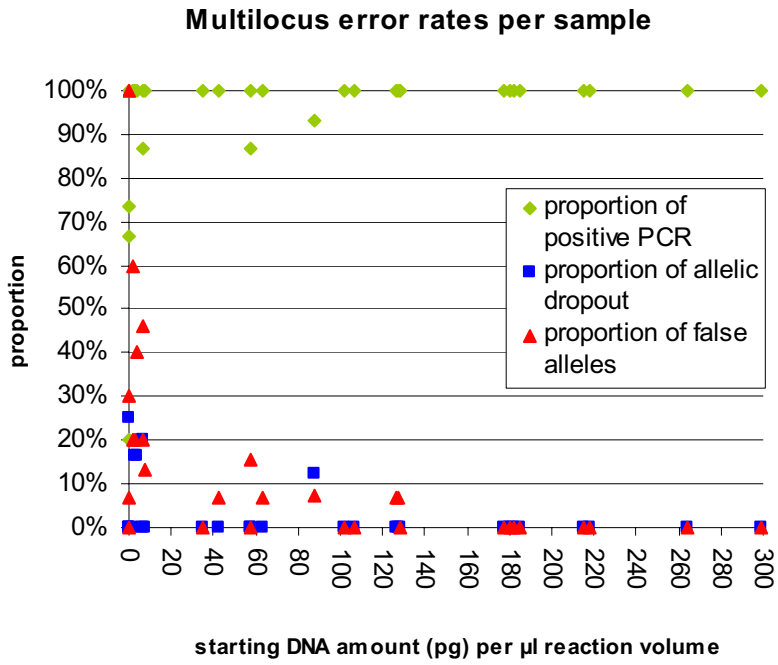


Fig. 7 Dependence of amplification success (green diamonds) and error rates (i.e. proportion of allelic dropout and of false alleles) over all loci on the starting amount of template. Each point represents one extract that was amplified and genotyped independently three times for five loci. For each extract, between six and nine amplifications of heterozygous loci were examined for allelic dropout (blue squares). The proportion of false alleles was investigated among 15 amplifications for each extract (red triangles).

To summarize the results, almost half (45.7%) of the genotypes determined from these low concentrated DNA templates did not correspond to the correct genotypes obtained from blood samples, due to either false alleles or the mis-amplification of one allele at heterozygous loci.

Fig. 7 indicates that with increasing template concentration (>10 pg/µl), the frequency of false alleles remained below 16%. No false alleles occurred if the starting DNA amount exceeded 130pg/µl. Among higher template concentrations (>10 pg/µl), allelic dropout was observed once. It should be noted that amplification failure could not be excluded for template concentrations lower than 88pg/µl.

3.3.2 Classification of samples

Genotyping results for each of the five loci were examined and classified into four categories (for details, see Materials and Methods section). Fig. 8 shows category lists for each locus plotted against template concentration. Quartiles and median values were calculated for each error category per individual locus.

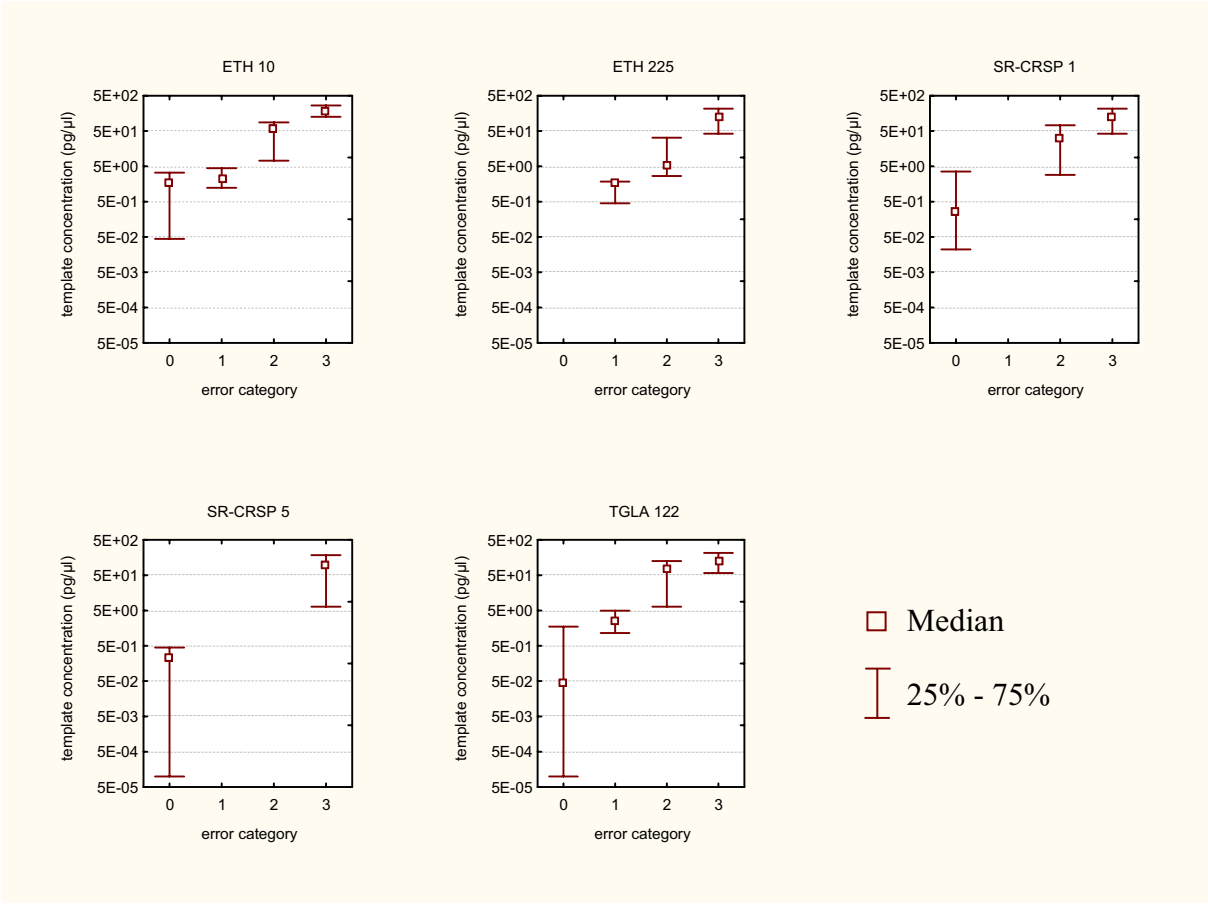


Fig. 8 Distribution of error categories calculated for each locus. Categories are plotted against template concentration in pg/µl PCR. Single samples that were classified into one error category are not displayed in this graph. Medians calculated for each error category are shown as squares.

For all microsatellite loci, the increase of template concentration led to higher reliability of genotyping results (i.e. to a classification of extracts into higher categories). However, there were differences between individual loci in the median and quartile values of each error category. Microsatellite locus ETH 10 shows that correct genotypes were derived in all replications (i.e. category 3) with a median starting DNA amount of 128pg/µl. Other loci exhibited lower medians for error category 3, with SR-CRSP 5 displaying the lowest among individual loci (75pg/µl starting template amount). Moreover, microsatellite locus ETH 10 exhibited the highest values for the lower quartiles of error category 3. For this marker, 75% of extracts classified into this category exceeded 62pg/µl, whereas for locus SR-CRSP 1, the lower quartile suggests that 75% of extracts contained more than 3pg/µl. This value was the lowest among all loci for category 3. However, loci SR-CRSP 5, ETH 225 and TGLA 122 also displayed similar results for lower quartiles.

4 Discussion

4.1 Species differentiation

Results obtained by Kuehn *et al.* (2000) revealed successful differentiation of closely related species through PCR and RFLP analyses, targeting cyt b fragments derived from non-invasive samples. Confirming these findings, endonuclease cleavage of mitochondrial DNA (mtDNA) amplicons facilitated discrimination of ibex samples from those obtained from red deer, chamois and goat, related species that:

(i) potentially occur in or near alpine ibex habitat;
(ii) that produce similar pellet material, and;
(iii) whose DNA is very likely to be amplifiable on microsatellite loci used in this study.

Therefore, the application of this method allows exclusion of ambiguous pellet material prior to further analysis. Furthermore, RFLP analysis facilitates non-invasive sampling of alpine ibex material that is independent of observing individuals. In contrast to nuclear DNA, mitochondrial DNA (mtDNA) is present in high copy numbers per cell, resulting in higher template amounts for PCR amplification and, hence, leading to an increased restriction and differentiation success, even in non-invasive samples. Thus, pellet material can be used as a valuable source of DNA for discrimination of target species, as demonstrated here with ibex and chamois faecal-derived DNA.

4.2 DNA quantification

DNA quantification using qPCR and the SYBR Green detection system showed that most faecal samples from alpine ibex contained relatively high amounts of nuclear DNA, as compared to other studies that have applied qPCR to faecal samples of the chimpanzee (*Pan troglodytes verus*: Morin *et al.* 2001) and mountain gorilla (*Gorilla beringei beringei:* Nsubuga *et al.* 2004), respectively. Variability could be due to differences in fluorescence detection systems: a potential source of error affecting the SYBR Green detection system includes a lack of specificity. Due to the fact that this fluorescence detection system measures

double-stranded PCR products (Ririe *et al.* 1997), potential contaminating and co-amplified DNA (e.g. plant and bacterial DNA in faecal samples) along with the presence of primer-oligomers, can lead to overestimation of DNA amounts. In this study, the use of an optimal primer concentration in qPCR reduced the formation of primer-dimers to a minimum, which was verified through agarose gel electrophoresis and the distinct melting curves observed. Furthermore, melting curve analysis revealed no differences in either denaturation temperatures (76°C) or shape of the melting peaks between blood and faecal-derived qPCR products. Thus, exact concordance in length, GC content and sequence can be assumed (Ririe *et al.* 1997). As blood-derived amplicons were sequenced and found to be homologous to the 133bp fragment of the target gene, it can also be inferred that exclusively ibex DNA was amplified in faecal samples. In contrast, a single faecal-derived amplicon exhibited a melting peak that was lower and differed -1.5°C from the expected melting temperature. Considering that the pellet had disintegrated completely during extraction, inhibitory compounds and/or contaminating DNA from inner pellet material were likely to have been present and probably exceeded the ibex template amounts in the extract, resulting in non-specific amplification. Due to the differing denaturation temperature, this qPCR product was presumably not congruent with the target products (Ririe *et al.* 1997). Hence, the amount of DNA calculated from the C_t value probably does not represent the amount of target DNA contained in the extract. This finding is consistent with that reported by Ririe *et al.* (1997), who indicated ´limited sensitivity by non-specific amplification at low initial template number`.

To summarise, it can be inferred that differences in DNA amounts between this study and those presented by Morin *et al.* (2001) and Nsubuga *et al.* (2004) do not result from non-specific quantification in the present study. More probably, they are caused by species-specific variability (Taberlet & Luikart 1999; Fernando *et al.* 2003), due to differences in diet, digestive systems and living conditions (Fernando *et al.* 2003, Nsubuga *et al.* 2004). This enhances the necessity for preliminary studies that investigate species-specific attributes (Taberlet & Luikart 1999).

Furthermore, different sampling, storage, DNA extraction and amplification methods were applied in the studies. Note that in this study, ibex pellets which exhibited high amounts of DNA were freshly collected from the colon of the animals. Thus, degradation of ibex DNA was reduced to a minimum compared to sampling in the natural environment, where faeces may have been in varying stages of decomposition (Taberlet *et al.* 1999). Lower amounts of

DNA were calculated for samples that had either disintegrated completely during DNA extraction, or that had been stored for more than four years prior to purification. These findings are broadly consistent with results from previous studies: Nsubuga *et al.* (2004) demonstrated continuous template degradation with temporally extended preservation, and Flagstad *et al.* (1999) and Wehausen *et al.* (2004) indicated a negative correlation between the amount of inner pellet material included in extraction and amplification success. The highly erroneous genotyping results of these samples further support the assumption that inhibitory compounds and/or degraded DNA were contained in the faecal extracts.

QPCR experiments exhibited high variability in the amount of DNA between individuals of the same species (similar to Morin *et al.* 2001; Fernando *et al.* 2003; Nsubuga *et al.* 2004) and between extracts of the same individual. This could be due to variations in the accuracy and stochastic processes between single experiments (Lockey *et al.* 1998). However, the results also displayed low variance in C_t values for standard samples between independent qPCR experiments, as well as high correlation coefficients for the three standard curves (R^2: between 0.9663 and 0.9925). Thus, variation between single experiments is negligible as a potential source of error. More probably, variability in faecal-derived DNA amounts between individuals resulted from differing physiological states of the animals sampled (Fernando *et al.* 2003). High variance in template concentrations between extracts of the same individual were probably due to the physical condition of faecal material, with a positive connection between the firmness of pellets and amount of DNA obtained, supporting findings described by Flagstad *et al.* (1999) and Wehausen *et al.* (2004). This hypothesis needs to be taken into account when conducting future studies using ibex faeces. In order to increase the likelihood of obtaining extracts that yield high amounts of template, several pellets should be sampled per individual, especially if samples are moist and less compact.

Taken together, these findings suggest that quantitative PCR using SYBR Green represents a powerful tool to pre-screen ibex faecal extracts for DNA content. QPCR results are congruent with observed error rates for faecal samples, indicating that amplification using RupruSTD1 is sufficiently specific to allow exclusion of low-quality extracts. Eluates containing degraded DNA, inhibitory compounds and/or a high amount of non-target DNA can be identified via mis-amplification or through investigation of melting curves. In future studies, faecal-derived

qPCR products could moreover be sequenced and aligned to the target sequence to allow extra confirmation of data quality (Bradley & Vigilant 2002).

4.3 Assessment of microsatellite genotyping reliability

The aim of this study was to assess the reliability of microsatellite genotyping using faecal samples as a source of ibex DNA. Among multiple extracts and using a panel of five dinucleotide markers, the extent of mis-amplification and genotyping errors along with their dependence on template amount was examined.

4.3.1 Positive PCR

The observed amplification success rate among multiple faecal DNA extracts was similar to those obtained by Huber *et al.* 2003, and Flagstad *et al.* (1999), who achieved positive PCR amplifications in >95% of all genotyping experiments. By contrast, Taberlet *et al.* (1996), Gerloff *et al.* 1995, Frantzen *et al.* 1998, Constable *et al.* 2001 and Morin *et al.* (2001) reported amplification success in lower percentages (up to 64%), whereas Parsons (2001) and Fernando *et al.* (2003) revealed positive PCRs in almost 100% of experiments.

The prevailing reasons for the non-detection of alleles are absolute low template concentrations and the presence of PCR inhibitory compounds (Fernando *et al.* 2003). Therefore, differences might be explained by: (i), variation in DNA quantity and quality of faecal samples between species (Taberlet & Luikart 1999), along with (ii), different sensitivities of the primers used in these studies (He *et al.* 1994). Supporting these assumptions, the present study revealed mis-amplification principally in PCR experiments that started with low amounts of target DNA, as indicated in other studies (Morin *et al.* 2001; Fernando *et al.* 2003). The comparatively high amplification success rate in this study could further be explained by the solitary use of dinucleotide markers. In general, these repeats are shorter and hence, easier to amplify, compared to tri- or tetranucleotide markers (Taberlet & Luikart 1999). Results furthermore confirmed that the amount of available target DNA varied widely between faecal extracts (similar to Morin *et al.* 2001), leading to mis-amplification of samples that were below the threshold value for successful amplification. Additionally, these threshold values obviously differed between loci, which was exhibited by varying degrees of amplification failure.

4.3.2 Allelic dropout

With decreasing amounts of target DNA, several factors are of increasing importance and thus considered to be the main causes of the occurrence of AD. These are: (i) stochastic pipetting errors, (ii) varying copy numbers of each allele, (iii) differential amplification of one of two alleles or (iv) mismatches between primer and target sequences (Taberlet & Luikart 1999; Lathuilliere 2001).

Confirming these assumptions, 90% of AD in this study was observed in amplifications that started with particularly low template concentrations (<10pg/µl). The observed rate of false homozygous genotypes was relatively low (average: 4.2%) as compared to other studies on faeces (e.g. Buchan et al. 2005: 19-21%; Morin et al. 2001: 24%). Huber et al. (2003) and Flagstad et al. (1999), however, reported a similar AD rate (<3%). Furthermore, Fernando et al. (2003) observed dropout in only 0.4% of all experiments, whereas Parsons (2001) reported a total absence of false homozygotes. These variations are again likely due to different properties of primer pairs and to the differing quantity or quality of DNA in faecal extracts (Taberlet & Luikart 1999; Pompanon et al. 2005) used in these studies. For instance, Morin *et al.* (2001) obtained an average DNA amount of 192 pg/µl with a related AD rate of 24%. By contrast, ibex faeces obtained 908pg/µl DNA on average and AD in 4.2% of heterozygous genotypes. Fernando *et al.* (2003) reported false homozygous genotypes in 0.4% of cases, related to an average DNA amount of 4100 pg/µl. Furthermore, differences in allele scoring methods could also impact the rate of AD. Most studies do not report peak minimum values for peak intensities that limited scoring of alleles. Thus, if amplitude thresholds for analysis of electropherograms differ, the extent of AD as well as rates of positive PCR could be influenced.

It is suggested that allelic dropout is less likely to be detected at homozygous loci, and therefore AD error rates are not directly comparable between loci or populations that vary in heterozygosity (Pompanon *et al.* 2005). Furthermore, AD rates in this study were observed among a small number of successful amplifications of heterozygous loci (N=236). However, the trend for an inverse relationship between the concentration of amplifiable DNA in the extract and dropout was obvious and consistent with the findings of other studies (Taberlet *et al.* 1996; Goossens *et al.* 1998; Morin *et al.* 2001; Fernando *et al.* 2003).

4.3.3 False alleles

As expected according to previous studies, the incidence of false alleles increased with a decrease of template amount (Goossens *et al.* 1998). The extent of erroneous single-locus genotypes resulting from false alleles (11.8%) was higher than in most other faecal DNA studies (Morin *et al.* 2001, Parsons 2001: <1%; Taberlet *et al.* 1999: <5%; Fernando *et al.* 2003: 1%), but similar to the extent of FA observed by Huber et al. (2003) in analysis of faeces from red deer (*Cervus elaphus*) and hares (*Lepus europaeus*). Huber *et al.* (2003) presumed that the high percentages (approximately 15%) were facilitated by sporadic contamination, similar to assumptions made by Navidi *et al.* (1992) who estimated that up to 7% of erroneous results might be caused by contamination. In the present study, only 3.1% (N=3) of false alleles were known to exist in the study population and hence, could be derived from cross-contamination with DNA from other individuals. As they were observed only once per extract among three replications, contamination would probably have occurred during PCR setup. The similarities between error rates calculated for ibex and red deer, lends support to the idea that specific characteristics of ruminant faeces might impact the reliability of microsatellite genotyping.

It needs to be mentioned that alleles which fell outside the normal range were scored repeatedly among several extracts and replications (average: 87.5%). This contradicts findings by other researchers who have reported the spurious and non-reproducible occurrence of false alleles (e.g. Morin *et al.* 2001). It should be noted that all loci used in this study were dinucleotides, which are assumed to be prone to error from false alleles, generated by polymerase slippages and resulting stutter bands (Taberlet *et al.* 1996; Goossens *et al.* 1998; Taberlet & Luikart 1999; Morin *et al.* 2001). According to Fernando *et al.* (2003), slippage is uni-directional, leading to products smaller in size (bp) than true alleles. Such scoring difficulties were principally observed for loci ETH 10 and TGLA 122 that revealed stutter bands even in blood-derived PCR products. Here, large peak intensity differences in amplicons derived from low template amounts, resulted in mis-scoring of smaller products as true alleles (see Fig. 6a on page 20). This confirms the assumptions made by Fernando *et al.* (2003) and supports results shown by Goossens *et al.* (1998). Loci that were easier to score (ETH 225, SR-CRSP 1 and SR-CRSP 5) revealed substantially lower rates of FA with such a

pattern. This is congruent with the fact that Morin *et al.* (2001), who used tri- and tetranucleotide repeats, reported a negligible extent of false alleles.

Nevertheless, FAs with a different structure were still present among all loci. Here, patterns were similar to those observed by Bradley & Vigilant (2002): alleles that fell outside the potential stutter and normal size ranges were scored repeatedly among several extracts and replications. Furthermore, they were present as peaks of normal shape and intensity compared to those of true alleles (Fig. 6b on page 20). Two extracts in particular that had disintegrated during extraction and could not be measured in qPCR, exhibited a very high percentage of false alleles (78.6%) with such a pattern. This suggests that compounds co-purified from the inner pellet material might have been amplified and led to erroneous results, lending support to the findings described by Bradley & Vigilant (2002). The fact that the extent of erroneous genotypes increased markedly with a decrease in template concentration seems to further support this idea: as the ratio between target and contaminating DNA changes towards the latter, co-amplification of contaminant DNA is likely to increase, resulting in higher rates of FA. This could further be enhanced by mismatches between primer and target sequences: Bradley & Vigilant (2002) assumed that primers transferred from a different species might not be wholly complementary to the template of the new species and hence, lead to reduced amplification specificity.

Concerning different calculation methods of FA rates, it is to be noted that examination ´per replicated alleles` revealed slightly underestimated FA error rates, even if ratios between loci remained similar. This confirms statements made by Broquet & Petit (2004). For this study, the extent of erroneous single-locus genotypes is of prime importance, rather than the total number of false alleles per positive reaction. Therefore, preference should be given to the calculation of error rates per locus (similar to Broquet & Petit 2004, Pompanon *et al.* 2005).

To summarise, multi-locus genotyping of ibex faeces revealed ubiquitous and comparatively high rates of FA among extracts containing low DNA amounts. These erroneous genotypes were presumably due to: (i) Taq slippages in dinucleotide loci, resulting in mis-scoring of alleles, and (ii) contaminant or inhibitory compounds that were co-purified from faeces, resulting in unspecific amplification. In order to further investigate these potential sources of error: (i) tri- or tetranucleotide repeats could be analysed according to the present approach,

along with (ii) sequencing and sequence alignment of PCR products that revealed erroneous genotypes (Bradley & Vigilant 2002).

4.4.5 Applicability to future studies and perspectives

Classification analysis facilitated assessment of primer efficacy. Interlocus variability was clearly exhibited by the classification of genotyping results into error categories, revealing a strong dependence of genotyping reliability on initial template amount for all loci. Results showed that a minimum DNA amount of 128pg template per µl reaction (i.e. 790pg/µl eluate, 38.2% of extracts) was required to reliably determine genotypes for the panel of five loci presented in this study. Considering that all three amplifications exceeding this DNA concentration produced correct and acceptable results, a maximum of two replications would be sufficient. In order to expand the range of usable extracts, solely loci ETH 225, SR-CRSP 1 and -5 and TGLA 122 could be used for genotyping ibex, because the minimum DNA amounts are lower for reliable genotyping of these markers (100pg/µl reaction, i.e. 630pg/µl eluate, 47.1% of extracts). However, ETH 10 was found to be one of the most polymorphic loci among alpine ibex (Maudet *et al.* 2002) and hence, probably needs to be included to allow better resolution between individuals. For extracts containing DNA concentrations below these threshold values (classified into category 2), a minimum of three replications will be essential to obtain reliable genotypes. Here, the guidelines described for the multiple tube approach (Navidi *et al.* 1992; Taberlet *et al.* 1996) are recommended to be followed. Even low genotyping error rates can seriously influence biological inferences (Waits & Leberg 2003), especially for questions concerning individual identification (Creel *et al.* 2003) or paternity assignment (Hoffman & Amos 2005). As these issues will be of prime importance considering ibex conservation management, extracts with particularly low DNA concentrations (classified into category 1 and 0) should be excluded from such analyses. A new extract should be prepared and quantified in its place.

The relatively high mean error rates per locus (nearly 30% for ETH 10, along with an average value of 13.9% over all five loci) furthermore support the conclusion that a pre-screen of samples for the amount of nuclear DNA is inevitable. If molecular genetic analyses based solely on non-invasive samples, without any reference samples, were conducted on a panel of such highly error-prone loci, correct biological inferences would be almost impossible.

According to Taberlet *et al.* (1996), the overall genotyping error rate is recommended to be lower than 1%. This recommendation can only be met with the exclusive use of samples that exceed the minimum DNA amount necessary for reliable genotyping.
It should be noted that samples in this study were removed from the colon of subject animals and were effectively freshly collected, so do not truly represent ´natural conditions`. Thus, lower template amounts contained in faecal extracts can be expected if samples are collected several hours or days after defecation and exposure to weather conditions (Wehausen *et al.* 2004), due to higher degradation. In order to achieve amplifications that contain the required amount of DNA for genotyping the panel of five loci, the concentration of template per reaction could be adjusted. Nonetheless, this might simultaneously increase the amount of contaminant DNA or inhibitory compounds in the reaction (Taberlet *et al.* 1999) and thus, does not necessarily lead to improved results. Employment of NaCl in DNA extraction along with the use of BSA in microsatellite amplification in this study, however, facilitated an increase of template amount in PCR without decreasing amplification success (data not shown), due to binding of inhibitors (Pääbo 1999). It should still be considered though, that this adjustment poses the risk of exhausting sample material. Hence, the possible number of genotyping experiments would be reduced, which would presumably need to be high for resolution of individual animals within bottlenecked ibex populations.

In order to increase the feasibility of non-invasive approaches, this study could be extended to screen further microsatellite loci. As the results demonstrated, there is variability in the minimum amount of DNA necessary for reliable genotyping of microsatellites. Consequently, the range of loci that could be used for amplifying faecal extracts with lower amounts of DNA is likely to be increased with the number of markers examined. Preferentially, tri- or tetranucleotide microsatellite loci should be selected, in order to reduce potential stutter bands and hence, false alleles that can complicate the analysis of dinucleotide loci. Due to small sample size and relatively low numbers of microsatellite loci assessed, examination of significant correlation between allele size and extent of genotyping errors was limited. Nevertheless, preference should be given to loci yielding shorter products, as these are considered to be less error-prone (Buchan *et al.* 2005; Hoffman & Amos 2005). A trend confirming these statements could be observed in the present results.

It is worth noting that the minimum DNA amount necessary for reliable microsatellite genotyping of two replications (128pg template per µl reaction) differs substantially from that presented by Morin *et al.* (2001). In their study, they indicated a threshold value of 13.4 pg template per µl reaction. This variation, presumably caused by different characteristics of the microsatellite markers and quality of extracts, indicates that results cannot necessarily be transferred to other species. This supports the recommendations of Taberlet & Luikart (1999) and Buchan *et al.* (2005), who strongly advocate using pilot studies to explore the locus- and species-specific properties of genotypes.

5 Conclusions

Despite the relatively small number of samples, this study clearly demonstrates the potential of faecal samples in the non-invasive genetic exploration of ibex populations. Multi-locus microsatellite genotyping of ibex samples revealed the ubiquitous presence of genotyping errors among the panel of five loci used, with varying degrees of AD and FA between loci and false alleles representing the most commonly encountered errors. However, genotyping error rates exhibited a strong dependence on amplifiable template amount. Hence, if a threshold value of DNA amount is exceeded in extracts, reliable genotyping is facilitated. Through quantification of nuclear DNA and examination of related genotyping errors, the minimum DNA concentration required for reliable genotyping could be determined for the panel of five microsatellite loci. Thus, for future studies, classification of extracts by nuclear DNA content will allow exclusion of extracts that are not suitable for analysis, due to their low template concentrations and associated poor amplification success and reliability. Consequently, fruitless attempts to genotype low quality extracts can be avoided.

In conclusion, this study showed that if a certain threshold value for template concentration can be met in the extract, ibex faeces are a source of DNA providing reliable amplification of microsatellite markers. Above threshold values, PCR inhibition and primer characteristics are not problematic, and two PCR replicates per sample are sufficient to eliminate genotyping errors.

6 Acknowledgments

This work was conducted at the Technische Universität München, funded by the working group ´Molecular Ecology and Conservation Genetics`. I would like to thank the Hellabrunn Zoo in Munich and Flurin Filli from the Swiss National Park for providing samples from alpine ibex and chamois. I would also like to thank Ralph Kühn and Roland Hausknecht for supervision. Many thanks to my friends Eike Müller, Stephanie Pietsch and Helmut Bayerl for fruitful discussions and all their support during the last few months. Furthermore, I would like to thank Rob Gregory, Dirk Rothenstein and Vanessa Salzer for great help and very helpful comments on the manuscript and Stephan Masepohl and my parents for their patience and their motivating words.

7 References

Arevalo E, Holder DA, Derr JN*, et al.* (1994) Caprine microsatellite dinucleotide repeat polymorphisms at the SR-CRSP-1, SR-CRSP-2, SR-CRSP-3, SR-CRSP-4 and SR-CRSP-5 loci. *Animal Genetics* **25**, 202.

Barendse W, Vaiman D, Kemp SJ*, et al.* (1997) A medium-density genetic linkage map of the bovine genome. *Mammalian Genome* **8**, 21-28.

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

Bellemain E, Swenson JE, Tallmon D, Brunberg S, Taberlet P (2005) Estimating population size of elusive animals with DNA from hunter-collected feces: four methods for Brown Bears. *Conservation Biology* **19**, 150-161.

Bhebe E, Kogi J, Holder D*, et al.* (1994) Caprine microsatellite dinucleotide repeat polymorphisms at the SR-CRSP-6, SR-CRSP-7, SR-CRSP-8, SR-CRSP-9, SR-CRSP-10 loci. *Animal Genetics* **25**, 203.

Bishop MD, LKappes SM, Keele JW*, et al.* (1994) A Genetic Linkage Map for Cattle. *Genetics* **136**, 619-639.

Bonin A, Bellemain E, Bronken Eidesen P*, et al.* (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* **13**, 3261-3273.

Bradley BJ, Vigilant L (2002) False alleles derived from microbial DNA pose a potential source of error in microsatellite genotyping of DNA from faeces. *Molecular Ecology Notes* **2**, 602-605.

Broquet T, Petit E (2004) Quantifying genotyping errors in noninvasive population genetics. *Molecular Ecology* **13**, 3601-3608.

Buchan JC, Archie EA, Van Horn RC, Moss CJ, Alberts SC (2005) Locus effects and sources of error in noninvasive genotyping. *Molecular Ecology Notes* **5**, 680-683.

Buchanan FC, Crawford AM (1993) Ovine microsatellites at the OarFCB11, OarFCB128, OarFCB193, OarFCB266 and OarFCB304 loci. *Animal Genetics* **24**, 145.

Buchanan FC, Galloway SM, Crawford AM (1994) Ovine microsatellites at the OarFCB5, OarFCB19, OarFCB20, OarFCB48, OarFCB129 and OarFCB226 loci. *Animal Genetics* **25**, 60.

Constable JL, Ashley MV, Goodall J, Pusey AE (2001) Noninvasive paternity assignment in Gombe chimpanzees. *Molecular Ecology* **10**, 1279-1300.

Creel S, Spong G, Sands JL*, et al.* (2003) Population size estimation in Yellowstone wolves with error-prone noninvasive microsatellite genotypes. *Molecular Ecology* **12**, 2003-2009.

Epps CW, Palsboll PJ, Wehausen JD*, et al.* (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology Letters* **8**, 1029-1038.

Ernest HB, Penedo MC, May BP, Syvanen M, Boyce WM (2000) Molecular tracking of mountain lions in the Yosemite valley region in California: genetic analysis using microsatellites and faecal DNA. *Molecular Ecology* **9**, 433-441.

Farrell LE, Roman J, Sunquist ME (2000) Dietary separation of sympatric carnivores identified by molecular analysis of scats. *Molecular Ecology* **9**, 1583-1590.

Fernando P, Evans BJ, Morales JC, Melnick DJ (2001) Electrophoresis artefacts - a previously unrecognized cause of error in microsatellite analysis. *Molecular Ecology Notes* **1**, 325-328.

Fernando P, Vidya TN, Rajapakse C, Dangolla A, Melnick DJ (2003) Reliable noninvasive genotyping: fantasy or reality? *Journal of Heredity* **94**, 115-123.

Flagstad O, Roed K, Stacy JE, Jakobsen KS (1999) Reliable noninvasive genotyping based on excremental PCR of nuclear DNA purified with a magnetic bead protocol. *Molecular Ecology* **8**, 879-883.

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

Frantzen MA, Silk JB, Ferguson JW, Wayne RK, Kohn MH (1998) Empirical evaluation of preservation methods for faecal DNA. *Molecular Ecology* **7**, 1423-1428.

Garnier JN, Bruford MW, Goossens B (2001) Mating system and reproductive skew in the black rhinoceros. *Molecular Ecology* **10**, 2031-2041.

Gerloff U, Schlötterer C, Rassmann K*, et al.* (1995) Amplification of hypervariable simple sequence repeats (microsatellites) from excremental DNA of wild living bonobos (*Pan paniscus*). *Molecular Ecology* **4**, 515-518.

Goossens B, Waits LP, Taberlet P (1998) Plucked hair samples as a source of DNA: reliability of dinucleotide microsatellite genotyping. *Molecular Ecology* **7**, 1237-1241.

He Q, Marjamaki M, Soini H, Mertsola J, Viljanen MK (1994) Primers are decisive for sensitivity of PCR. *Biotechniques* **17**, 82-87.

Hoffman J, Amos, W. (2005) Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology* **14**, 599-612.

Hogan BLM, Constantini F, Lacy E (1986) Manipulating the mouse embryo, Section D. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Huber S, Bruns U, Arnold W (2003) Genotyping herbivore feces facilitating their further analyses. *Wildlife Society Bulletin* **31**, 692-697.

Kappes SM, Keele JW, Stone RT*, et al.* (1997) A second-generation linkage map of the bovine genome. *Genome Research* **7**, 235-249.

Kocher TD, Thomas WK, Meyer A*, et al.* (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 6196-6200.

Kogi J, Yeh CC, Bhebhe E*, et al.* (1995) Caprine microsatellite dinucleotide repeat polymorphisms at the SR-CRSP-11, SR-CRSP-12, SR-CRSP-13, SR-CRSP-14 and SR-CRSP-15 loci. *Animal Genetics* **26**, 449.

Kohn MH, York EC, Kamradt DA*, et al.* (1999) Estimating population size by genotyping faeces. *Proceedings of the Royal Society of London, Series B: Biological Science* **266**, 657-663.

Kuehn R, Schwab G, Schroeder W, Rottmann O (2000) Differentiation of *Castor fiber* and *Castor canadensis* by non-invasive molecular methods. *Zoo Biology* **19**, 511-515.

Lathuilliere M, Menard 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*). *American Journal of Primatology* **55**, 151-158.

Lockey C, Otto E, Long Z (1998) Real-time fluorescence detection of a single DNA molecule. *Biotechniques* **24**, 744-746.

Maudet C, Miller C, Bassano B*, et al.* (2002) Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine ibex [*Capra ibex (ibex)*]. *Molecular Ecology* **11**, 421-436.

Maudet C, Beja-Pereira A, Zeyl E*, et al.* (2004a) A standard set of polymorphic microsatellites for threatened mountain ungulates (Caprini, Artiodactyla). *Molecular Ecology Notes* **4**, 49-55.

Maudet C, Luikart G, Dubray D, von Hardenberg A, Taberlet P (2004b) Low genotyping error rates in wild ungulate faeces sampled in winter. *Molecular Ecology Notes* **4**, 772-774.

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

Morin PA, Chambers KE, 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*). *Molecular Ecology* **10**, 1835-1844.

Navidi W, Arnheim N, Waterman MS (1992) A multiple-tubes approach for accurate genotyping of very small DNA samples by using PCR: statistical considerations. *American Journal of Human Genetics* **50**, 347-359.

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

Pääbo S, Gifford JA, Wilson AC (1988) Mitochondrial DNA sequences from a 7000-year old brain. *Nucleic Acids Research* **16**, 9775-9787.

Paetkau D (2003) An empirical exploration of data quality in DNA-based population inventories. *Molecular Ecology* **12**, 1375-1387.

Palomares F, Godoy JA, Piriz A, O'Brien SJ, Johnson WE (2002) Faecal genetic analysis to determine the presence and distribution of elusive carnivores: design and feasibility for the Iberian lynx. *Molecular Ecology* **11**, 2171–2182.

Parsons KM (2001) Reliable microsatellite genotyping of dolphin DNA from faeces. *Molecular Ecology Notes* **1**, 341-344.

Piggott 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. *Conservation Genetics* **5**, 417-420.

Piggott MP, Banks SC, Stone N, Banffy C, Taylor AC (2006) Estimating population size of endangered brush-tailed rock-wallaby (*Petrogale penicillata*) colonies using faecal DNA. *Molecular Ecology* **15**, 81-91.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics* **6**, 847-859.

Ririe KM, Rasmussen RP, Wittwer CT (1997) Product Differentiation by Analysis of DNA Melting Curves during the Polymerase Chain Reaction. *Analytical Biochemistry* **245**, 154-160.

Roeder AD, Archer FI, Poinar HN, Morin PA (2004) A novel method for collection and preservation of faeces for genetic studies. *Molecular Ecology Notes* **4**, 761-764.

Roon DA, Waits LP, Kendall KC (2005) A simulation test of the effectiveness of several methods for error-checking non-invasive genetic data. *Animal Conservation* **8**, 203-215.

Ropiquet A, Hassanin A (2005) Molecular evidence for the polyphyly of the genus Hemitragus (Mammalia, Bovidae). *Molecular Phylogenetics and Evolution* **36**, 154-168.

Smith DA, Ralls K, Hurt A*, et al.* (2006) Assessing reliability of microsatellite genotypes from kit fox faecal samples using genetic and GIS analyses. *Molecular Ecology* **15**, 387-406.

Taberlet P, Griffin S, Goossens B*, et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research* **24**, 3189-3194.

Taberlet P, Luikart G (1999) Non-invasive genetic sampling and individual identification. *Biological Journal of the Linnean Society* **68**, 41-55.

Taberlet P, Waits LP, Luikart G (1999) Noninvasive genetic sampling: look before you leap. *Trends in Ecology & Evolution* **14**, 323-327.

Thieven U, Solinas-Toldo S, Friedl R*, et al.* (1997) Polymorphic CA-microsatellites for the integration of the bovine genetic and physical map. *Mammalian Genome* **8**, 52-55.

Vaiman D, Mercier D, Moazami-Goudarzi K*, et al.* (1994) A set of 99 cattle microsatellites: characterization, synteny mapping, and polymorphism. *Mammalian Genome* **5**, 288-297.

Valiere N, Berthier P, Mouchiroud D, Pontier D (2002) GEMINI: software for testing the effects of genotyping errors and multitubes approach for individual identification. *Molecular Ecology Notes* **2**, 83-86.

Vincze T, Posfai J, Roberts RJ (2003) NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Research* **31**, 3688-3691.

Waits JL, Leberg PL (2000) Biases associated with population estimation using molecular tagging. *Animal Conservation* **3**, 191-199.

Waits LP, Paetkau D (2005) Noninvasive genetic sampling tools for wildlife biologists: a review of applications and recommendations for accurate data collection. *Journal of Wildlife Management* **69**, 1419-1433.

Wehausen JD, Ramey RR, Epps CW (2004) Experiments in DNA extraction and PCR amplification from bighorn sheep feces: the importance of DNA extraction method. *Journal of Heredity* **95**, 503-509.

Yeh CC, Kogi JK, Holder MT*, et al.* (1997) Caprine microsatellite dinucleotide repeat polymorphisms at the SR-CRSP21, SR-CRSP22, SR-CRSP23, SR-CRSP24, SR-CRSP25, SR-CRSP26 and SR-CRSP27 loci. *Animal Genetics* **28**, 380-381.

APPENDIX I

Table A1 Weight of 34 faecal pellets (mg) used in DNA extraction, observed C_t values, DNA concentrations measured in these extracts (in pg/µl) and initial template concentration contained in 1µl PCR volume (in pg). All extracts contained a total volume of 100µl. Pellets that had disintegrated completely during the first step in extraction are marked in the third column ([d]).

Animal	Extract	pellet (mg)	C_t value	Calculated DNA concentration (pg/µl)	Template concentration (pg/µl) per reaction
1	A	360	29.12	362	58
	B	330	28.05	793	127
	C	320	29.55	264	42
	D	180	28.06	787	126
	E	280	29.12	362	58
2	A	100	28.03	805	129
	B	240	29.80	220	35
	C	310	27.31	1364	218
	D	180	29.01	392	63
	E	360	28.35	636	102
3	A	230	25.96	3673	588
	B	160	24.97	7594	1215
	C	140 [d]	> 40	0.00	0.000
	D	330 [d]	35.74	2.81	0.450
	E	280 [d]	> 40	0.00	0.000
4	A	200	27.56	1136	182
	B	160	27.33	1345	215
	C	150	28.29	665	106
	D	210	26.88	1870	299
	E	80	27.59	1111	178
5	A	150	28.55	549	88
	B	240	27.57	1128	180
	C	310	27.05	1651	264
	D	290	27.54	1153	184
	E	290	26.33	2800	448
6	A	120	33.01	21	3
	B	380	32.12	40	6
	C	390	33.7	13	2
	D	460	33.45	15	2
	E	570	31.95	45	7
7	A	410	38.9	0.28	0.04
	B	300	32.07	42	7
	C	210	33.81	12	2
	D	290	33.89	11	2

APPENDIX II

Table A2 Panel of microsatellite loci tested in amplifications with blood- and faecal-derived DNA from *Capra ibex.* The second and third column display whether the cross-species amplification was successful (+) or not (-).

Locus	Primer reference	Transferable on blood-derived DNA from *Capra ibex*	Transferable on faecal-derived DNA from *Capra ibex*
SR-CRSP 8	Bhebe *et al.* 1994	+	+
SR-CRSP 24	Yeh *et al.*1997	+	+
BM 1225	Bishop *et al.* 1994	+	+
INRA 36	Vaiman *et al.* 1994	+	+
OarFCB 304	Buchanan & Crawford 1993	+	+
CSSM 36	Kappes *et al.* 1997	+	-
CSSM 6	Kappes *et al.* 1997	+	-
CSSM 16	Kappes *et al.* 1997	+	-
SR-CRSP 7	Bhebe *et al.* 1994	+	-
SR-CRSP 12	Kogi *et al.* 1995	+	-
SR-CRSP 14	Kogi *et al.* 1995	+	-
ETH 152	Kappes *et al.* 1997	+	-
INRA 35	Vaiman *et al.* 1994	+	-
Oar FCB 20	Buchanan *et al.* 1994	+	-
BM 1818	Kappes *et al.* 1997	+	-
MM 12	Barendse *et al.* 1997	+	-
IDVGA 30	Kappes *et al.* 1997	+	-
INRA 23	Vaiman *et al.* 1994	-	-
CSSM 66	Bishop *et al.* 1994	-	-
HAUT 24	Thieven *et al.* 1997	-	-
ILSTS 6	Kappes *et al.* 1997	-	-
ILSTS 30	Kappes *et al.* 1997	-	-

APPENDIX III

Table A3 Macro written in GENOTYPER® 2.5 to analyse electropherogram patterns and to score alleles of microsatellite loci amplicons.

```
Remove labels from peaks whose height is less than 32% of the
highest peak in a category range;
then remove labels from peaks preceded by a higher, labelled
peak within 0.00 to 1.60 bp;
then remove labels from peaks followed by a higher, labelled
peak within 0.00 to 3.00 bp.
```