2005

Evaluating mixed samples as a source of error in non-invasive genetic studies using microsatellites

David A. Roon
University of Idaho, droon@uidaho.edu

Miranda E. Thomas
University of Idaho

Katherine C. Kendall
USGS-NRMSC Science Center, kkendall@usgs.gov

Lisette P. Waits
University of Idaho, lwaits@uidaho.edu

Molecular Ecology (2005) **14**, 195–201 doi: 10.1111/j.1365-294X.2004.02385.x

Evaluating mixed samples as a source of error in non-invasive genetic studies using microsatellites

DAVID A. ROON,* MIRANDA E. THOMAS,† KATHERINE C. KENDALL‡ and LISETTE P. WAITS*
Department of Fish and Wildlife, University of Idaho, Moscow, Idaho 83844–1136, USA, †2290 South Race #629, Denver, CO 80210, ‡USGS-NRMSC Science Center, c/o Glacier National Park, West Glacier, Montana, 59936–0128, USA

Abstract

The use of noninvasive genetic sampling (NGS) for surveying wild populations is increasing rapidly. Currently, only a limited number of studies have evaluated potential biases associated with NGS. This paper evaluates the potential errors associated with analysing mixed samples drawn from multiple animals. Most NGS studies assume that mixed samples will be identified and removed during the genotyping process. We evaluated this assumption by creating 128 mixed samples of extracted DNA from brown bear (*Ursus arctos*) hair samples. These mixed samples were genotyped and screened for errors at six microsatellite loci according to protocols consistent with those used in other NGS studies. Five mixed samples produced acceptable genotypes after the first screening. However, all mixed samples produced multiple alleles at one or more loci, amplified as only one of the source samples, or yielded inconsistent electropherograms by the final stage of the error-checking process. These processes could potentially reduce the number of individuals observed in NGS studies, but errors should be conservative within demographic estimates. Researchers should be aware of the potential for mixed samples and carefully design gel analysis criteria and error checking protocols to detect mixed samples.

Keywords: genotype error, individual identification, mark-recapture, microsatellites, noninvasive genetic sampling, *Ursus arctos*

Received 30 June 2004; revision received 22 September 2004; accepted 22 September 2004

Introduction

Noninvasive genetic sampling (NGS) of wildlife populations is a powerful tool for assessing demography, gene flow, and population structure. Multilocus genotyping, particularly using polymorphic microsatellites, allows for the identification of individuals from highly dilute DNA sources. Materials such as scat, sloughed skin, urine and hair can be collected in many field settings, using a range of methodologies (Taberlet *et al*. 1999). NGS studies, however, are typically carried out with limited observation of the actual study organism. As samples are collected, it is generally assumed that each extracted sample is drawn from a single individual.

Although the assumption that an NGS sample contains the DNA of only one animal is defensible in many cases, there are instances where an NGS sample might originate from more than one individual. Social species may defecate in fixed latrines (Frantz *et al*. 2003; Wilson *et al*. 2003) and territorial animals often urinate in fixed sites (Valiere & Taberlet 2000). Hair traps (Woods *et al*. 1996) and rub tree stations (Kendall *et al*. 1997) are often sampled at broad intervals (Woods *et al*. 1999; Poole *et al*. 2001) allowing for visitation by multiple animals between sampling occasions. Because of the limited amount of DNA in a single hair follicle, NGS researchers frequently pool multiple hairs in a single extraction (Woods *et al*. 1999; Kendall *et al*. 1997; Alpers *et al*. 2003), potentially including DNA from multiple individuals. Similar issues can arise through hair collection at primate nesting sites, which may be used by multiple animals (Gerloff *et al*. 1995).

NGS studies typically select polymorphic microsatellite loci based on probability of identity statistics ($PI_{(ID)}$) or similar measures of allele sharing (Waits *et al*. 2001; Paetkau 2003). When allele sharing is minimized, mixed samples should theoretically display three or more alleles at one or more loci (Alpers *et al*. 2003). Given a reliable program of

Correspondence: David A. Roon, Fax: 208885–9080. E-mail: roon8505@uidaho.edu

genetic screening, there are two circumstances under which mixed samples could add false data points to NGS studies. A mixed sample could contain DNA from individuals with high levels of allele sharing (either through identity by descent, or by chance). If the combined number of alleles per locus was less than three for all amplified loci, then the amplified mixed sample could resemble a legitimate, unique genotype (referred to as Composite Type I, synthesis). The likelihood of such an event should decrease as population-wide allele sharing declines, approaching zero when a sufficient number of polymorphic loci is used.

A more pernicious problem could occur as a consequence of allelic dropout. NGS DNA template is often dilute and / or degraded (Taberlet *et al*. 1999). Allelic dropout (Navidi *et al*. 1992; Taberlet *et al*. 1996; Goossens *et al*. 1998) could generate legitimate-looking genotypes (two or fewer alleles per locus), even in cases where the true number of alleles per locus exceeded two for some amplified loci (Composite Type II, allelic dropout). Thus, the incidence of this type of 'composite' genotype would not necessarily become inconsequential at low levels of allele sharing. Theoretically, dropout should be detected through repeated amplification. However, specific alleles from one of the two source samples might consistently fail to amplify for NGS samples, especially given the competitive nature of the polymerase chain reaction (PCR).

Although the potential for mixed samples has been acknowledged in multiple NGS studies (Table 1), only two studies have experimentally evaluated their potential impact. Wasser *et al*. (1997) 'spiked' a scat from a known sun bear individual with zero, five, and 10 hairs from another known individual, extracted DNA, and demonstrated coamplification of RFLP bands from the two individuals. Alpers *et al*. (2003) used a simulation approach to test the strengths

Table 1 Noninvasive genetic sampling studies and potential bias resulting from mixed samples

Template	# Hairs	Focal species	Deposition observed*	Risk of E.A.†	E.A. observed	Author
Urine	NA	Wolf & dog	No	Yes	Yes	Valiere & Taberlet (2000)
Shed Skin	NA	Whale	No	Yes	No	Valsecchi *et al*. (1998)
Scat‡	NA	Wolf	No	Unclear	No	Valiere *et al*. (2003)
Scat	NA	Langur	Yes	No	No	Launhardt *et al*. (1998)
Scat	NA	Bear	No	Unclear	No	Taberlet *et al*. (1997)
Scat	NA	Gibbon	Yes	No	No	Oka & Takeneka (2001)
Scat	NA	Lynx	No	No	No	Palomares *et al*. (2002)
Scat	NA	Chimpanzee	Yes	No	No	Morin *et al*. (2001)
Scat	NA	Coyote	No	Unclear	No	Kohn *et al*. (1999)
Scat	NA	Bonobo	Yes	No	No	Gerloff *et al*. (1995)
Scat	NA	Rhinoceros	Yes	No	No	Garnier *et al*. (2001)
Scat	NA	Chimpanzee	No	Unclear	No	Constable *et al*. (2001)
Scat	NA	Bonobo	Yes	No	No	Gerloff *et al*. (1995)
Scat	NA	Baboon	Yes	No	No	Bayes *et al*. (2000)
Scat	NA	Wombat	No	Unclear	No	Banks *et al*. (2003)
Scat	NA	Badger	No	No	No	Wilson *et al*. (2003)
Scat	NA	Seal	No	Unclear	No	Reed *et al*. (1997)
Scat	NA	Orangutan	Unclear	Unclear	No	Goossens *et al*. (2000)
Scat	NA	Elephant	No	Unclear	No	Eggert *et al*. (2003)
Scat	NA	Otter	Unclear	Unclear	Unclear	Dallas *et al*. (2003)
Hair	Single hair	Wombat	No	No	No	Sloane *et al*. (2000)
Hair	Single hair	Bear	No	No	No	Taberlet *et al*. (1997)
Hair	Single hair	Chimpanzee	Yes	No	No	Morin *et al*. (1994)
Hair	Single hair	Chimpanzee	Yes	No	No	Morin *et al*. (2001)
Hair	Multiple hairs	Marmot	Yes	No	No	Goossens *et al*. (1998)
Hair	Unclear	Gibbon	No	Unclear	No	Oka *et al*. (2001)
Hair	Multiple hairs	Chimpanzee	Yes	Unclear	No	Constable *et al*. (2001)
Hair	Multiple hairs	Marten	No	Unclear	No	Mowat & Paetkau (2002)
Hair	Multiple hairs	Bear	No	Yes	No	Poole *et al*. (2001)
Hair	Multiple hairs	Bear	No	Yes	Yes	Paetkau (2003)
Hair	Multiple hairs	Bear	No	Yes	Yes	Mowat & Strobeck (2000)
Hair	Multiple hairs	Bear	No	Yes	Yes	Woods *et al*. (1999)

*In studies marked 'deposition observed', samples were collected immediately after a known subject deposited scat or hair in an unambiguous location; †'E.A.' signifies 'excess alleles' (three or four alleles at a single locus); ‡For scat studies, we considered there to be no risk of E.A. when scat deposition was observed, or when preventitive measures (such as the use of builders' chalk) were incorporated into sampling.

and weaknesses of pooling hair samples in wombats, and concluded that pooling increased microsatellite genotyping success. They predicted that mixed samples could easily be detected in populations with high to moderate diversity, but could be a problem for populations with low diversity.

To assess the potential for mixed samples to bias NGS studies, we combined DNA extracted from clumps of bear hair. We amplified six microsatellite loci for these artificial mixed samples, and evaluated how many composite genotypes (Type I & Type II) were generated. We then applied a common error-checking protocol to assess whether these composites would be detected in most NGS studies, and reviewed the literature to determine the prevalence of mixed sample issues in NGS studies.

Materials and methods

Samples for our study were drawn from the Greater Glacier Bear DNA project, a large-scale NGS study of brown bear (*Ursus arctos*) and black bear (*Ursus americanus*) within the Northern Continental Divide Ecosystem (Kendall *et al*. 1997). During this 3-year study, over 3000 brown bear samples were genotyped at six microsatellite loci; G1A, G10B, G10C, G10L, G10M and G10P (Paetkau & Strobeck 1994; Paetkau *et al*. 1995), using the error-checking protocols described below. Global heterozygosity for this paper was 0.71, with allele counts ranging from five (G10L) to 12 (G10B) per locus (D. Roon, K. Kendall, L. Waits, unpublished data). DNA was extracted from five to 10 hair follicles per sample ($\mu = 8$). For the samples used in this paper, mean probability of identity $P_{(ID)}$ was 0.000013, while the mean sibling probability of identity $P_{(ID)sib}$ was 0.0084.

We created each mixed sample by combining 50 μL from two previously genotyped brown bear hair extractions. The exact DNA concentrations for the Glacier Project samples have not been quantified; thus, relative concentrations of DNA varied between mixed samples. We considered standardizing DNA concentrations (Morin *et al*. 2001) but felt that a coarse approach would better reflect the sample conditions within most studies.

We selected three sets of sample pairs with pairwise kinship values (Queller *et al*. 1993) of approximately 0.0 ($N = 44$), 0.3 ($N = 44$), and 0.6 ($N = 40$). These kinship levels were chosen to represent low, moderate, and high levels of allele sharing. A pairwise kinship value of 0.6 might be observed in first-order relatives; values of zero might be observed in unrelated individuals (Queller *et al*. 1993). Individuals with a higher pairwise kinship should display elevated levels of allele sharing; mixed samples drawn from such individuals will have a higher theoretical probability of generating type I and type II composite genotypes.

We obtained six-locus genotypes for each mixed sample. PCR reactions were carried out in 15 μL volumes in an MJ PTC-100 thermal cycler. Microsatellites G1A, G10B, G10C, and G10L were initially amplified as a multiplex; reaction mixes contained 2 μL template DNA, 1.5 μL 10x Goldtaq buffer (Sigma), 1.5 U GoldTaq (Sigma), 150 μM of each dNTP, 0.2 μM of G1A, G10B, and G10C, and 0.3 μM of G10L. After an initial denaturation of 10 min at 95 °C, 39 cycles of 30 s at 95 °C, 30 s at 57.5 °C, and 40 s at 72 °C were performed followed by a final extension of 2 min at 72 °C. Microsatellites G10M and G10P were initially amplified as a multiplex; reaction mixes contained 2 μL template DNA, 1.5 μL Goldtaq buffer (Sigma), 1.0 U GoldTaq (Sigma), 150 μM of each dNTP, 0.4 μM of G10M and 0.3 μM of G10P. After an initial denaturation of 10 min at 95 °C, 39 cycles of 30 s at 95 °C, 30 s at 52 °C, and 40 s at 72 °C were performed followed by a final extension of 2 min at 72 °C. For subsequent screening amplifications, microsatellite sets G1A/G10B and G10C/G10L were amplified as duplexes, while G10M and G10P were amplified solo. Primer concentrations and PCR conditions were the same as described for the multiplexes. Primers were fluorescently labelled with dyes, TET or 6-FAM or HEX (ABI). Negative controls were incorporated into all amplification and electrophoresis runs.

Amplified products were evaluated using an ABI 377 system, according to the protocols described in Woods *et al*. (1999). In cases where a potential allele peak was observed in a leading position (two base pairs longer than another potential allele peak), we required that the leading peak be > 200 florescent units in intensity, and between 0.25 times and 0.90 times the intensity of the following peak.

Each mixed sample was assigned a unique tracking number; gel analysis was performed under a double-blind design by a technician who was not familiar with the source genotypes. If more than four amplifications for any one locus were required to complete an initial six-locus genotype, the sample was considered unreliable, and eliminated from further analysis.

Once six-locus genotypes were obtained, we validated genotypes as follows. All loci were amplified at least twice. If all loci for a mixed sample were amplified twice with no inconsistent results, we accepted the genotype as valid. If we observed an incongruous result at a locus (e.g. initial homozygous amplification, subsequent heterozygous amplification), we performed additional amplifications. For heterozygous loci, we accepted results if two of three amplifications matched. For apparent homozygous loci, we required at least three matching amplifications to overrule a single inconsistent result. If more than four PCRs were required to complete and validate any locus, we considered the sample to be unreliable, and removed it from our data set.

These requirements are not as stringent as some multiple tubes approaches proposed in the literature (Taberlet *et al*. 1996; Miller *et al*. 2002; Frontz *et al*. 2003). As one example, Morin *et al*. (2001) suggested using between two and seven PCRs per homozygous locus, depending on

template concentration (as established through the use of quantitative PCR). In contrast, Sloane *et al*. (2000) suggested that low allelic dropout rates (0.3%) eliminated the need for multiple-tubes; they did suggest that genotypes with single-locus mismatches (across 12 loci) should be treated as matching. Other studies have emphasized the re-amplification of unique genotypes, or of genotypes that differ at single locus; not all of these studies have stipulated criteria for including or rejecting genotypes with inconsistent histories (Woods *et al*. 1999; Mowat & Paetkau 2002). Our methodology serves as an intermediate approach, and a standard for comparison.

Once genotyping was complete, we reviewed the source samples, and crosschecked these data against our genotyping results. In cases where mixed samples yielded an apparent valid genotype, that genotype was checked against the genotypes of both source samples.

Results

For our data set of 128 mixed samples and six microsatellite loci, we generated 768 single-locus amplifications during initial amplification and error checking. Based on the source samples, 315 of these amplification events encompassed mixed sample loci with three or more alleles (143 at kinship = 0.0, 105 at kinship = 0.3, 67 at kinship = 0.6). Mixed sample loci with three or more alleles did not always display the same number of alleles upon amplification (Table 2). For mixed sample loci with three alleles, 35% of recorded amplifications produced only two alleles; 5% produced one. For mixed sample loci with four alleles, 10% of amplifications produced three alleles, 28% produced only two alleles and 5% produced one (Table 2).

When samples with more than three alleles produced alleles less than that number, it was predominantly because one of the two source samples completely failed to amplify (Table 2). Only nine mixed loci amplified as Type I composites, while 25 amplified as Type II composites (Table 2). It should be noted, however, that there were only 14 total loci in our data set where Type I composites could have occurred. Thus Type I error occurred in greater than 60% of the amplifications when the error was possible.

For the 128 'mixed samples' amplified during the study, 123 carried more than three alleles at one or more loci. Five of the 123 'multiple allele' samples produced valid-appearing composite genotypes (Type II, allelic dropout) upon initial amplification. All of these samples were registered as mixed during the subsequent error-checking process.

Ninety-nine of the 123 samples with more than three alleles displayed clear evidence of more than three alleles upon initial amplification, while 11 displayed inconsistent amplification histories, or unacceptable electropherograms. Eight of the 123 'multiple allele' samples yielded results that were an exact match to one of the two source samples (kinship = 0.0 (n = 2), 0.3 (n = 4), and 0.6 (n = 2)), as a result of complete amplification failure of the other source samples.

Five of the 128 'mixed samples' amplified during the study theoretically could have produced Type I (synthesis) composite genotypes. Pairwise kinship values for these samples were zero (n = 1), 0.3 (n = 2), and 0.6 (n = 2). Four of these samples produced unacceptable electropherograms, with spectral intensities less than 200, or deviations from a classic microsatellite 'stepwise' pattern at one or more loci. The fifth sample initially amplified cleanly at six loci, producing a composite genotype (Table 3). During the validation process, Locus G10L amplified twice as a heterozygote for this sample, and twice as a homozygote. Following our reliability criteria, the overall result for this sample was rejected. Of all the amplified mixed samples in this paper, this one sample (0.0–40) came closest to yielding a false 'composite' genotype (Table 3), even under scrutiny.

Table 2 Overall occurrences of different amplification events, grouped according to the number of alleles available to be amplified

			Percentage success for amplifying available alleles				Occurrence of single-sample dropouts and composite genotypes (observed / possible)		
# Alleles*	Total amplifications	Failure	One allele	Two alleles	Three alleles	Four alleles	Single sample dropout	Type I composite	Type II composite
1	197	35%	65%	0	0	0	NA	NA	NA
2	842	25%	20%	55%	0	0	53% (168/310)	64% (9/14)	NA
3	640	20%	5%	35%	40%	0	35% (223/640)	NA	3% (19/640)
4	88	18%	1%	28%	10%	43%	20% (18/88)	NA	7% (6/88)

*'# Alleles' indicates the total number of alleles available to be amplified at a given locus. 'One Allele' through 'Four Alleles' indicates the number of alleles that actually amplified.'single sample dropout' indicates that one of two source samples failed to amplify, while the second of two source samples amplified well. For single-sample dropouts and composite genotypes, results are expressed as the number of observed occurrences over the number of instances in which a given event could have theoretically occurred.

Table 3 Known source sample genotypes and amplification history for mixed sample 0.0–40

	Locus G1A	Locus G10B	Locus G10C	Locus G10L*	Locus G10M	Locus G10P
Source sample A:	185–189	149–155	104–110	153–153	209–213	161–161
Source sample B:	179–189	155–155	104–104	155–155	209–213	155–155
Amplification 1:	179–189	155–155	104–104	153–155	Failed	155–155
Amplification 2:	179–189	Failed	104–104	155–155	209–213	155–155
Amplification 3:	NA	155–155	NA	155–155	209–213	NA
Amplification 4:	NA	NA	NA	153–155	209–213	NA

*Acceptance of initial results, or less stringent error-checking criteria at locus G10L, would have resulted in the acceptance of a Type II (allelic dropout). NA = Not Attempted.

Discussion

The probability of mixed sample occurrence will be a function of study design. We surveyed 32 NGS studies that used hair (n = 13), scat (n = 18), sloughed skin (n = 1) and urine (n = 1) sampling methods (Table 1). Our survey suggests that mixed sample risks will be minimal in most scat studies, higher in sloughed skin studies and studies where multiple hairs are extracted, and potentially highest in urine studies. In 14 of the 19 scat studies, sampling techniques or scat characteristics effectively eliminated mixed samples as a concern (Table 1). Scats were often collected directly following deposition (Gerloff *et al*. 1995; Bayes *et al*. 2000), or, in one case, after a badger latrine had been dusted with builder's chalk (Wilson *et al*. 2003). In five scat studies, sampling techniques were not explained in enough detail for mixed-sample risks to be assessed (Table 1).

In the single sloughed skin study, no multiple alleles were observed. However, the investigators noted that contact between whales could result in the transfer of cellular material (Valsecchi *et al*. 1998). In seven of 13 hair studies, researchers included multiple hairs in extractions. Four of these studies acknowledged the possibility of mixed samples, but argued that genetic screening could identify all such cases. Putative mixed samples were detected in three of these studies (Table 1). For the single urine study, mixed samples were identified as a concern, and detected during microsatellite amplification (Valiere & Taberlet 2000).

The potential impact of mixed samples could be heightened by the presence of Type II (allelic dropout) composite genotypes, as these could occur even at minimal levels of allele sharing. In the absence of error-checking, our study would have generated such composites for 3.9% of the amplified samples (n = 5). However, we observed no composite genotypes in our study after error checking. Our artificial 'mixed samples' encompassed five potential Type I (synthesis) situations; results for these samples were uniformly eliminated because of amplification inconsistencies.

We did see many instances where one source sample out-amplified another within a mixed sample as was documented in Alpers *et al*. (2003). This pattern could potentially result in the creation of Type II (allelic dropout) composite genotypes (Table 3). In our analyses, sample 0.0–40 was the only mixed sample that came close to producing a Type II composite (Table 3). Although the methodology of Morin *et al*. (2001) should have identified 0.0–40 as a mixed sample, it is conceivable that the types of protocols outlined in recent NGS census papers (Poole *et al*. 2001; Mowat & Paetkau 2002) could have initially allowed for the acceptance of a genotype history like 0.0–40. However, an emphasis on careful scrutiny of all single locus mismatch samples (Paetkau 2003) should correct this problem. Our sample size is modest; this probably accounts for the lack of meaningful differences between the number of composite genotypes produced at kinship values of zero, 0.3, and 0.6. Overall, however, our results suggest that events of this type should not occur with enough frequency to impact NGS study results significantly, if rigorous error-checking protocols are utilized.

In eight cases, a mixed sample yielded a genotype from only one of the source samples. Given the competitive nature of PCR, we suggest that one of the two source DNA stocks was more concentrated and overwhelmed the other during the amplification process. In a research project, the amplification of only one of two source genotypes from a 'mixed sample' would mask the presence of one potential data point. A false data point would not, however, be added. The loss of a data point could reduce observed recapture rates (White *et al*. 1982); however, at worst, the resultant error should be conservative, adding a potential negative bias to demographic estimates.

The observed 14% rate of per locus amplification failure (109 of 768 loci) was higher than expected, given that all source samples amplified well during the original study (unpublished data). The presence of DNA from two different animals may have impacted PCR dynamics, or DNA may have degraded in storage (Lindahl 1993). We should note, as well, that many of our PCRs were carried out in duplex or triplex and the use of multiplex PCR may have increased our failure rate.

Although PCR and gel analysis were carried out using a double-blind protocol, the technician involved in gel analysis was aware that samples were 'mixed'. Thus, gels were scrutinized with a heightened level of sensitivity to the presence of anomalous peaks. During the analysis phase of a large-scale NGS study, it would be very easy to ignore extraneous peaks (or just exclude the locus with the questionable result when all other loci amplify cleanly). Errors of this type could be caused by simple analysis fatigue (Paetkau 2003); at the same time, DNA from single animals can sometimes yield patterns that resemble patterns from mixed samples.

We recognize that our experimental conditions do not perfectly mimic the type of events that would occur in a field study, where material from two animals would be extracted as a single unit. However, we argue that a mixture of two extractions should provide a reasonable facsimile. The source DNA extracts that we used had variable concentrations. An experiment that standardized DNA concentrations might yield different results. However, it is unlikely that field collection and extraction would consistently yield template with equal concentrations of DNA from two different animals.

Population-level genetic diversity is a key factor in predicting the potential risk of not detecting mixed samples. We caution that our results are based on a population with relatively high genetic diversity and computer simulations have suggested that mixed samples will be more difficult to detect in small or isolated populations with low diversity (Alpers *et al*. 2003). $P_{(ID)sib}$ is often used as a 'rule of thumb' for defining the number of loci needed to avoid the risk of false genotype matches, the 'shadow effect' defined by Mills *et al*. (2000), and it can also be used as a lower bound on the number of loci needed to detect mixed samples (Alpers *et al*. 2003). However, up to a 67% increase in the number of amplified loci could be required to identify mixed samples in cases where allelic dropout is high and mixed samples are prevalent (Alpers *et al*. 2003).

Overall, our results suggest that mixed samples are not an overwhelming concern for NGS studies. We did see multiple cases where extraneous alleles could have potentially been overlooked, or discounted. Researchers must be meticulous in scrutinizing their data for the presence of multiple alleles. Additionally, we stress that mixed sample detection would be lower in any study where unique genotypes are not re-amplified (Paetkau 2003). However, the overall incidence of mixed samples appears to be low in most field studies (Table 1) and can be reduced through careful study design and collection methods. Carefully selecting a diverse suite of polymorphic loci can minimize overall levels of allele sharing. Researchers should be alert for the presence of mixed samples, but need not view mixed samples as a crippling pitfall in the collection of NGS data.

Acknowledgements

We acknowledge J. Stetz and the many technicians and volunteers who have contributed to the Greater Glacier Bear DNA project. Support for this research was provided by the McIntire-Stennis Foundation, the United States Geological Survey, and a NSF REU grant to the University of Idaho Environmental Science Program (#0097833).

References

Alpers DL, Taylor AC, Sunnucks P, Bellman SA, Sherwin WB (2003) Pooling hair samples to increase DNA yield for PCR. *Conservation Genetics*, **4**, 779–788.

Banks SC, Hoyle SD, Horsup A, Sunnucks P, Taylor AC (2003) Demographic monitoring of an entire species the northern hairy-nosed wombat (*Lasiorhinus krefftii*) by genetic analysis of non-invasively collected material. *Animal Conservation*, **6**, 101–107.

Bayes M, Smith K, Alberts S, Altmann J, Bruford M (2000) Testing the reliability of microsatellite typing from faecal DNA in the savannah baboon. *Conservation Genetics*, **1**, 172–175.

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

Dallas JF, Coxon KE, Sykes T, *et al*. (2003) Similar estimates of population genetic composition and sex ratio derived from carcases and feces of Eurasian otter *Lutra lutra*. *Molecular Ecology*, **12**, 275–282.

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

Frantz AC, Pope LC, Carpenter PJ, Roper TJ, Wilson GJ, Delaney RJ, Burke T (2003) Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. *Molecular Ecology*, **12**, 1649–1661.

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

Gerloff U, Schlotterer C, Rassman K (1995) Amplification of hypervariable simple sequence repeats (microsatellites) from excremental DNA of wild living Bonobos (*Pan paniscus*). *Molecular Ecology*, **4**, 515–518.

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

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

Kendall KC, Waits LP, Schirokauer D (1997) Using DNA to monitor grizzly and black bear populations. *Draft Study Plan*.

Kohn M, York E, Kamradt D, Haught G, Sauvajot R, Wayne R (1999) Estimating population size by genotyping feces. *Proceedings of the Royal Society of London*, **226**, 657–663.

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

Lindahl T (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.

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

Mills LS, Citta JJ, Lair KP, Schwartz MK, Tallmon DA (2000) Estimating animal abundance using noninvasive DNA sampling: promises and pitfalls. *Ecological Applications*, **10**, 283–294.

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.

Morin P, Wallis J, Moore J, Woodruff D (1994) Paternity exclusion in a community of wild chimpanzees using hypervariable simple sequence repeats. *Molecular Ecology*, **3**, 469–478.

Mowat G, Paetkau D (2002) Estimating marten *Martes americanus* population size using hair capture and genetic tagging. *Wildlife Biology*, **8**, 201–208.

Mowat G, Strobeck C (2000) Estimating population size of grizzly bears using hair capture, DNA profiling, and mark-recapture analysis. *Journal of Wildlife Management*, **64**, 183–193.

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*, **7**, 347–359.

Oka T, Takenaka O (2001) Wild gibbon's parentage tested by noninvasive DNA sampling and PCR-amplified polymorphic microsatellites. *Primates*, **42**, 67–73.

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

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

Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology*, **3**, 489–495.

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

Poole K, Mowat G, Fear D (2001) DNA-based population estimate for grizzly bears *Ursus arctos* in northeastern British Columbia, Canada. *Wildlife Biology*, **7**, 105–115.

Queller D, Srassman J, Hughes C (1993) Microsatellites and kinship. *Trends in Ecology and Evolution*, **8**, 285–288.

Reed J, Tollit D, Thonpson P, Amos W (1997) Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal feces. *Molecular Ecology*, **6**, 225–234.

Sloane MA, Sunnucks P, Alpers D, Beheregaray LB, Taylor AC (2000) Highly reliable genetic identification of individual northern hairy-nosed wombats from single remotely collected hairs: a feasible censusing method. *Molecular Ecology*, **9**, 1233–1240.

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

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

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

Valiere N, Fumagalli L, Gielly L, *et al.* (2003) Long-distance wolf recolonization of France and Switzerland inferred from non-invasive genetic sampling over a period of 10 years, *Animal Conservation*, **6**, 83–92.

Valiere N, Taberlet P (2000) Urine collected in the field as a source of DNA for species and individual identification. *Molecular Ecology*, **9**, 2150–2154.

Valsecchi E, Glockner-Ferrari D, Ferrari M, Amos W (1998) Molecular analysis of the efficiency of sloughed skin sampling in whale population genetics. *Molecular Ecology*, **7**, 1419–1422.

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

Wasser SK, Houston CS, Koehler GM, Cadd GG, Fain SR (1997) Techniques for application of fecal DNA methods to field studies of Ursids. *Molecular Ecology*, **6**, 1091–1097.

White GC, Anderson DR, Burnham KP, Otis DL (1982) *Capture-Recapture and Removal Methods for Sampling Closed Populations*. Los Alamos National Laboratory, Los Alamos, New Mexico. La-8787-NERP.

Wilson GJ, Frantz AC, Pope LC, *et al.* (2003) Estimation of badger abundance using fecal DNA typing, *Journal of Applied Ecology*, **40**, 658–666.

Woods J, McLellan B, Paetkau D, Strobeck C, Proctor M (1996) DNA fingerprinting applied to mark-recapture bear studies. *International Bear News*, **5**, 9–10.

Woods J, Paetkau D, Lewis D, McLellen B, Proctor M, Strobeck C (1999) Genetic tagging of free-ranging black and brown bears. *Wildlife Society Bulletin*, **27**, 616–627.

David Roon completed his PhD at the University of Idaho in June of 2004, working on the conservation genetics of brown and black bears in the Greater Glacier Area of Montana. Katherine C. Kendall is an ecologist with the USGS at the Glacier Field Station. Miranda Thomas is a recent graduate of the University of Denver. Lisette P. Waits is an associate professor at the University of Idaho, and cofounder of the Laboratory of Ecological and Conservation Genetics.
