Journal of Heredity 2003:94(2):115–123
DOI: 10.1093/jhered/esg022

Reliable Noninvasive Genotyping: Fantasy or Reality?

P. Fernando, T. N. C. Vidya, C. Rajapakse, A. Dangolla, and D. J. Melnick

From the Center for Environmental Research and Conservation, Columbia University, 1200 Amsterdam Avenue, New York, NY 10027, USA (Fernando and Melnick); Center for Ecological Sciences, Indian Institute of Science, Bangalore, India (Vidya); Department of National Zoological Gardens, Elephant Orphanage, Pinnawela, Sri Lanka (Rajapakse); School of Veterinary Science, University of Peradeniya, Peradeniya, Sri Lanka (Dangolla); and Departments of Anthropology and Biological Sciences, Columbia University, 1200 Amsterdam Avenue, New York, NY 10027, USA (Melnick). We thank L. K. A. Jayasinghe and officials at the Pinnawela elephant orphanage, Sri Lanka, for assistance, and the Department of National Zoological Gardens, Sri Lanka, for permission to collect samples. This research was funded in part by WWF US, WWF International, and USFWS Asian Elephant Conservation Fund grants for genetic analysis of Asian elephants; a USFWS Rhinoceros and Tiger Conservation Fund grant for genetic analysis of the Javan rhino; and additional support from the Center for Environmental Research and Conservation (CERC).

Address correspondence to Prithiviraj Fernando at the address above, or e-mail: pf133@columbia.edu.

Abstract

Noninvasive genotyping has not gained wide application, due to the notion that it is unreliable, and also because remedial measures are time consuming and expensive. Of the wide variety of noninvasive DNA sources, dung is the most universal and most widely used in studies. We have developed collection, extraction, and amplification protocols that are inexpensive and provide a high level of success in amplifying both mitochondrial and nuclear DNA from dung. Here we demonstrate the reliability of genotyping from elephant dung using these protocols by comparing results from dung-extracted DNA to results from blood-extracted DNA. The level of error from dung extractions was only slightly higher than from blood extractions, and conducting two extractions from each sample and a single amplification from each extraction was sufficient to eliminate error. Di-, tri-, and tetranucleotide loci were equally reliable, and low DNA quantity and quality and PCR inhibitors were not a major problem in genotyping from dung. We discuss the possible causes of error in genotyping with particular reference to noninvasive samples and suggest methods of reducing such error.

Recent advances in molecular genetics have led to a proliferation of studies applying genetic analysis to diverse fields such as development, ecology, evolution, behavior, and conservation. One of the factors constraining even broader application of genetic analysis is sample acquisition, especially with regard to rare, endangered, or cryptic fauna and those logistically difficult to sample. Realization that material such as dung, shed hair and feathers, sloughed skin, discarded food wadges, and eggshells were a source of DNA for genetic analysis suggested noninvasive sampling could overcome many sampling constraints (Kohn and Wayne 1997). However, a number of problems with noninvasive samples, such as copurifying contaminants (Litvaitis and Litvaitis 1996), low amounts of DNA (Frantzen et al. 1998) and DNA degradation leading to nonamplification, false alleles, sporadic contamination, and allelic dropout (Gagneux et al. 1997; Kohn and Wayne 1997; Taberlet et al. 1999), were soon identified. Some reports have suggested methods for overcoming these problems, among them the collection of three to six samples per individual (Frantzen et al. 1998), multiple extraction and PCR analysis of samples (Taberlet et al. 1996, 1999), and quantification of target DNA in extracts (Morin et al. 2001). Unfortunately, these approaches substantially increase the time and expense of laboratory analysis (Taberlet et al. 1999), thus partly nullifying the advantages of noninvasive sampling.

To date, dung and to a lesser extent hair have been the most widely employed noninvasive sources of DNA. Dung is of particular interest as all animals defecate regularly, and for many species, finding dung is comparatively simple, and collection, storage, and transport require little technology or expense. While the use of dung as a source of mitochondrial DNA is well established, and a number of comprehensive studies using this technique have been published, its use in nuclear DNA studies has been limited (Taberlet et al. 1999). Few published studies have empirically tested the reliability

of using dung as a source of DNA for nuclear markers. Using protocols optimized for collection, storage, extraction, and amplification of DNA from dung, we have had excellent success in amplifying both mitochondrial (Fernando and Lande 2000; Fernando et al. 2000) and nuclear DNA. Here we report the results of experiments to assess the reliability of amplifying nuclear microsatellite markers from dung and identify the causes of error. We compare amplification of microsatellite alleles from blood and dung of the same individuals at di-, tri-, and tetranucleotide loci to identify any differences in reliability between the two DNA sources and to assess if a particular type of locus is more appropriate for studies using dung.

Materials and Methods

Samples

Paired samples of dung and blood were obtained from 20 captive Asian elephants (*Elephas maximus*) in the "elephant orphanage," Pinnawela, Sri Lanka. Approximately 2 ml of blood or dung were collected into 15-ml tubes; 4 ml of storage buffer (100 mM NaCl, 10 mM Tris, 25 mM EDTA, 0.2 M guanidine thiocyanate, 2% SDS) was added as preservative, and the tubes were stored at ambient temperature. Dung samples were collected from individuals within 6 h after defecation. DNA extraction from dung targets epithelial cells sloughed from the gut lining during gut passage of the dung bolus. As the surface of the dung bolus is the last to have been in contact with the gut lining and the first to dry upon deposition, it would contain the least degraded DNA. Therefore, samples were collected from the surface of dung boli by scraping off the crust.

DNA Extraction, PCR Amplification, and Allele Scoring

Approximately 250 µl volume of sample (blood/dung) was placed in a 1.5-ml microcentrifuge tube, and 1 ml digestion buffer (same as storage buffer minus the guanidine thiocyanate) and 20 µl proteinase K 20 mg/ml solution were added. Samples were digested overnight at 70°C in a shaker at 100 rpm. Digests were spun down at 13000 rpm in a microcentrifuge, and 500 µl supernatant was extracted with 1 ml phenol-chloroform-isoamyl alcohol (PCI 25:24:1) following standard procedure (Sambrook et al. 1989). In recovering the aqueous layer from the PCI, only 400 µl was removed, to minimize pipetting up any denatured protein or PCI solution. The extract was purified using a QIAGEN gel extraction kit (the binding buffer provided has a pH indicator, and hence is useful in adjusting the pH if necessary, as binding of DNA to the silica-impregnated filter is pH dependant) and using the manufacturer's reagents and protocol.

Two extractions were made from each sample, and two PCR amplifications were conducted from each extraction for each locus. Amplification followed protocols published elsewhere (Fernando et al. 2000, 2001b). Samples extracted from blood and dung were amplified for 25 and 45 cycles, respectively. To minimize cross-contamination, extractions, PCR setup, and handling of amplification products were conducted at separate locations with dedicated instruments and reagents, and barrier tips were used for pipetting solutions.

We tested amplification for two each of tetra- (EMX3 and EMX4), tri- (EMX1 and EMX2) (Fernando et al. 2001b), and dinucleotide (Laf MS02 and Laf MS03) (Nyakaana and Arctander 1998) microsatellite loci. PCR products were electrophoresed on ABI 377 machines using TAMRA 500 (ABI) internal size standard. Gels were prerun for 1 h before loading, to overcome electrophoresis artifacts (Fernando et al. 2001a). Results were analyzed using GENESCAN software. Alleles were scored without knowledge of the individual's identity, and genotypes obtained from blood and dung were compared for each individual after all replicates were run.

As scoring alleles from dinucleotide loci may be prone to error from false alleles (stutter bands) generated by polymerase slippage (Schlotterer and Tautz 1992), based on multiple amplification and rerunning of samples, the following observations were made and subsequently used as guidelines in interpreting electropherograms:

Polymerase slippage is unidirectional, resulting in smaller products; hence stutter bands are always smaller in size (bp) than their true alleles;

In an electropherogram of a heterozygote with two alleles different by one or two repeats, the stutter bands of the larger allele are additively superimposed on the smaller allele; hence the peak of the smaller allele is higher/greater in area than the larger allele (e.g., in an individual with allele sizes 233 and 235, stutter bands of the 235 allele are sizes 233 and 231; thus, a peak is observed at 235 from the 235 allele; a peak at 233 composed of the 233 allele plus the first stutter of the 235 allele; a peak at 231 composed of the first stutter of the 233 allele plus the second stutter of the 235 allele, and so on);

In a heterozygote with alleles spaced far enough to eliminate superimposition of stutter bands, the smaller allele (bp) tends to amplify more strongly than the larger allele, but the peak of the larger allele is at least 1/10 the height/area of the smaller allele; and

Where a split peak is observed for an allele (two peaks one bp apart), the cumulative areas and heights of the two peaks represent the true area and height of the allele.

Dilution Experiment

To examine the occurrence of allelic dropout and non-amplification due to low DNA quantity, dung extracts (N = 26) were diluted in multiples of two from 1/5 to 1/1280 and PCR amplified with primers for the dinucleotide locus Laf MS03. Locus Laf MS03 was selected in preference to the other five loci, as it had the highest number of alleles (7) and heterozygotes (17/20) in the sample population; hence it would facilitate detection of allelic dropout. Total DNA in extractions was measured by a fluorometer (TKO 100,

Table 1. Results of PCR amplifications from dung (D) and blood (B) samples

	Locus											
	EMX1		EMX2		EMX3		EMX4		Laf MS02		Laf MS03	
	D	B	D	B	D	B	D	B	D	B	D	B
Correct genotype	79	80	79	80	78	80	77	80	78	78	80	80
False heterozygote	1	0	1	0	2	0	0	0	0	0	0	0
False homozygote	0	0	0	0	0	0	1	0	1	0	0	0
Nondetection of alleles	0	0	0	0	0	0	2	0	0	2	0	0
Multiple alleles	0	0	0	0	0	0	0	0	1	0	0	0

D, dung; B, blood. EMX1 and EMX2 are tri-, EMX3 and EMX4 are tetra-, and LafMS02 and LafMS03 are dinucleotide loci.

Hoefer Scientific Instruments). To examine if target DNA quantity available for PCR was related to total DNA concentration, the first dilution at which an incorrect result was observed was regressed against the DNA concentration measured.

Test of PCR Inhibition

To examine the occurrence of PCR inhibitors and to test if using higher volumes of extract in the PCR resulted in amplification failure from inhibitory copurifying substances, PCR reactions were conducted using 2×, 5×, 10×, 15×, and 30× of dung extracts ($N = 15$). To maintain conditions and reagent concentrations constant, the relevant extract volumes were completely evaporated in a microconcentrator, reconstituted with 2 μl of water, and PCR conducted with primers for locus Laf MS03.

Results

Two extractions each from 20 dung and 20 blood samples gave a total of 80 (40 × 2) extractions, and two PCRs from each extraction over 6 loci produced a total of 960 (80 × 2 × 6) PCRs, half (480) of which were with dung and half with blood-extracted DNA. Thus, for each sample (dung or blood) we had 4 replicate PCRs for each locus (2 extractions × 2 PCRs). The "true" genotype of an individual was taken as the genotype observed in at least 3 out of the 4 PCR reactions from blood-extracted DNA, and a different genotype from the "true genotype" or the nondetection of amplification products was deemed an error in genotyping.

Four errors were observed in our study: (1) False heterozygote: identification of a homozygote as a heterozygote due to the presence of an extra "allele." (2) False homozygote: identification of a heterozygote as a homozygote due to the recognition of only one of the two alleles. (3) Nondetection of alleles: inability to score a genotype due to absence of identifiable amplification product. (4) Multiple alleles: inability to score a genotype due to multiple peaks in an electropherogram.

A low level of error was observed with both dung- and blood-extracted samples. In the 480 PCR amplifications with dung-extracted DNA, the total error was 1.9% ($N = 9$), consisting of 0.8% ($N = 4$) false heterozygotes, 0.4% ($N = 2$) false homozygotes, 0.4% ($N = 2$) nondetection of alleles, and 0.2% ($N = 1$) multiple alleles. The total error observed in the 480 reactions with blood-extracted DNA was 0.4% ($N = 2$), which was due to nondetection of alleles (Table 1). Other than a single instance where both PCRs from the same extract of a dung sample recorded a false heterozygote at the tetranucleotide locus EMX 3, no instances of error at any one locus (from both dung and blood) were traced to a single extract (Table 2). Therefore, conducting two extractions from each sample and one PCR from each extract would suffice to eliminate the error observed in our study.

Dilution Experiment

All 26 dung samples provided correct results (true genotype) before dilution and at a dilution of 1 in 5. The proportion of error in genotyping (wrong genotype or nonamplification) increased with higher dilution, and all samples gave incorrect results at a dilution of 1 in 1280 (Figure 1). Incorrect genotypes were due to allelic dropout rather than amplification artifacts. The concentration of total DNA in the 26 dung extracts ranged from 3 to 113 ng/μl (mean 41, SD 37.3). The regression analysis found no relationship between the total DNA concentration and the first dilution at which an error was observed [$r^2 = .0066$, $P = .6828$ (Figure 2)], thus confirming that the quantity of target DNA available for PCR in dung extracts was not related to the concentration of total DNA (Morin et al. 2001).

Test of PCR Inhibition

The correct genotype was scored in all 15 samples of dung-extracted DNA at concentrations of 2×, 5×, and 10×. At 15× and 30×, no amplification products were detected in one sample, and at 30×, two additional samples had diminished signal, compared to at lower concentrations, but could be scored unambiguously. This suggests that the concentration of PCR inhibitors in our dung-extracted DNA samples was very low.

Discussion

Overall, our results demonstrate a high level of accuracy in genotyping from dung-extracted DNA, especially when com-

Table 2. Details of observed genotyping errors and probable causes

Extract	Locus	True genotype	Observed genotype	Observed error	Probable cause
D2.14	EMX1	134/134	134/151	False heterozygote	LE/SC
D1.13	EMX2	219/219	219/225	False heterozygote	LE/SC
D2.6	EMX3	253/253	237/253	False heterozygote	LE/SC
D2.6	EMX3	253/253	237/253	False heterozygote	LE/SC
D2.7	EMX4	351/387	387	False homozygote	LE/AD
D2.6	EMX4	387/387	–	Nondetection	LE/AF
D1.16	EMX4	387/387	–	Nondetection	LE/AF
D2.10	LAF2	135/137	133/135/137	Multiple alleles	LE/AA/SC
D2.17	LAF2	135/137	135	False homozygote	LE/AD
B1.17	LAF2	135/137	–	Nondetection	LE
B1.18	LAF2	135/137	–	Nondetection	LE

LE, laboratory error; SC, sporadic contamination; AD, allelic dropout; AF, amplification failure; AA, amplification artifacts. Extracts from dung are denoted D- and from blood B-, followed by the extract number (first or second) and the sample number. Five observed types of error may occur in genotyping: (1) False heterozygote: identification of a homozygote as a heterozygote due to the presence of an extra "allele." (2) False homozygote: identification of a heterozygote as a homozygote due to the recognition of only one of the two alleles. (3) Nondetection of alleles: inability to score a genotype due to absence of identifiable amplification product. (4) Multiple alleles: inability to score a genotype due to multiple peaks in electropherogram. (5) Incorrect assignment of alleles, where identified allele or alleles (in a homozygote or heterozygote respectively) differ from the true alleles of the individual.

pared to other studies reporting error from noninvasive genotyping. In addition to the four errors observed in our study (false heterozygote, false homozygote, nondetection of alleles, and multiple alleles), incorrect assignment of alleles, where identified allele or alleles (in a homozygote or heterozygote, respectively) differ from the true alleles of the individual, can also occur in genotyping. These five observed errors could result from a number of causes (Table 3), which are discussed below.

Laboratory Error

Laboratory error results from mistakes, such as switching or mixing of samples, and errors in setting up PCR, loading, or labeling. The observed error in our study from blood-extracted DNA (0.4% nondetection of alleles) is almost certainly due to laboratory error. Because laboratory error can cause any of the five observed errors (Table 3), it also cannot be excluded as a probable cause in many of the other instances of observed error in our study (Table 2). Few studies specifically assess or report levels of laboratory error, making it difficult to compare levels of error between studies (Gagneux et al. 1997). In a study of the reliability of genotyping from shed hair, Gagneux et al. (1997) attributed an error of 5.6% to laboratory error, contamination, and amplification artifacts. In the same study, they found <1% error with plucked hair samples, which suggests that only a small part of error with shed hair was laboratory error.

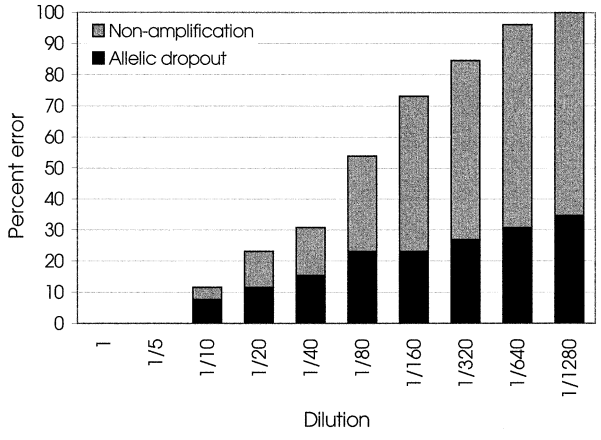

Figure 1. Cumulative error observed from nonamplification and allelic dropout with increasing dilution of dung samples ($N = 26$).

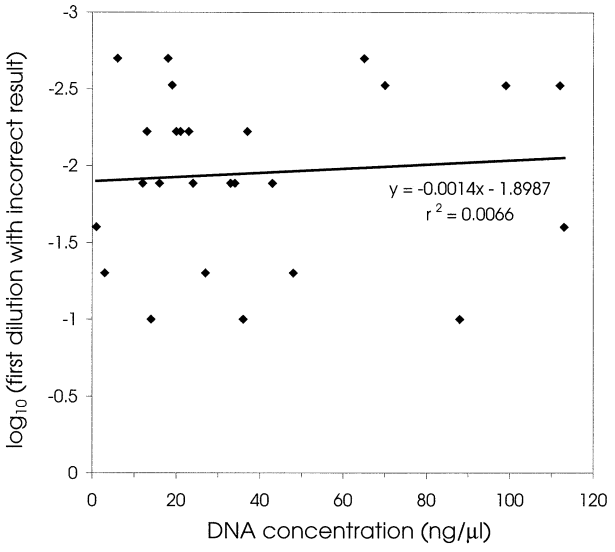

Figure 2. Regression of the first dilution at which an incorrect result was observed against total DNA concentration in dung extracts.

Table 3. Genotyping errors and their causes

Cause	Observed error				
	Nondetection of alleles	False homozygote	False heterozygote	Multiple alleles	Incorrect assignment
Laboratory error	X	X	X	X	X
Allelic dropout		X			
Amplification artifacts			X	X	
Contamination			X	X	
Incorrect size assessment					X
Amplification failure	X				
Electrophoresis artifacts			X	X	X

Morin et al. (2001) reported an error rate of 0.9% from "anomalous and nonreproducible results"—presumably laboratory error. Thus, the level of laboratory error in our study is comparable to the few other published sources. Laboratory error should be independent of the sample source, unless more complicated or longer extraction protocols are used for noninvasive samples, providing greater opportunity for mistakes to occur. In our study, samples from blood and dung were treated similarly. Thus, while laboratory error could have contributed to the observed error with dung-extracted DNA in our study (1.9%), the higher error in comparison to blood-extracted DNA (0.8%) suggests additional causes. The extent of laboratory error is likely to vary with different technicians and may be minimized by training, use of shorter protocols, and following clearly defined procedure in laboratory analyses.

Sporadic Contamination

Sporadic contamination results in false heterozygotes and multiple alleles, from cross-contamination with other samples or PCR products having alleles different from the individual in question. We observed 0.8% ($N = 4$) false heterozygotes and 0.2% ($N = 1$) multiple alleles with dung-extracted DNA. The "additional alleles" observed in each case (Table 2) were present in our population, and, hence, could be from cross-contamination. Two of the false heterozygotes were from extract D2.6, and contamination of the extract was the most likely cause. As the other two instances were from two different extracts, if contamination was the cause, it would have occurred at PCR setup. The instance of multiple alleles could be due to contamination at PCR setup or to amplification artifacts. While few studies assess or report error due to contamination, Navidi et al. (1992) estimated that sporadic contamination could cause up to 7% error in laboratory studies with large sample sizes.

Amplified samples represent millions of very small DNA fragments that are potential templates. Thus, low-level, sporadic contamination of extracts or reactions from indirect aerosol spread of amplified material is possible, especially when large numbers of samples are screened. It is of greater concern with noninvasive samples because of the usually higher number of amplification cycles (45 versus 25 cycles for blood in our study) and lower starting template concentration. High levels of contamination, such as from direct contamination with amplified product, could cause false homozygotes or incorrect alleles by swamping the true signal and could occur with any type of sample, but is better considered under laboratory error as it should not occur if adequate safeguards are taken.

Spatial separation of activities associated with extraction and PCR setup from those associated with handling amplified product; using dedicated instruments, reagents, and other consumables for each activity; and using barrier tips that reduce aerosol contamination of pipettes can decrease cross-contamination from amplified product.

Allelic Dropout

Stochastic sampling of only one allele of a heterozygote, resulting in a "false homozygote," is considered to be the main problem in genotyping noninvasive samples (Morin et al. 2001; Taberlet et al. 1999). The observed rate of false homozygotes for dung-extracted DNA in our study was 0.4%. Previous studies have reported 1.5%–2% (Flagstad et al. 1999) and 24% (Morin et al. 2001) allelic dropout with fecal extracts, 31.3% dropout with shed hair extracts (Gagneux et al. 1997), and 14.29%, 4.9%, and 0.41% with extracts from 1, 3, and 10 plucked hairs, respectively (Goosens et al. 1998). Below a threshold value, the incidence of allelic dropout is inversely related to the concentration of amplifiable DNA in the extract (Taberlet et al. 1999). Morin et al. (2001) demonstrated the use of quantitative PCR to estimate available target DNA in extracts from hair and dung samples. In their study, extracts from fecal and hair samples had an average of 192 and 21 pg/μl target DNA, respectively, with samples <100 pg/μl failing to amplify one-third of the time and exhibiting allelic dropout in almost half of successful amplifications (Morin et al. 2001). In contrast, in our dilution experiment, allelic dropout was first observed at a 1/10 sample dilution, and in some samples dropout was not evident even at a dilution of 1/640 (Figure 1). Although not directly comparable to Morin et al. (2001), our results suggest that a greater amount of target DNA was available for PCR in our extracts, explaining the comparatively low level of allelic dropout observed in our study.

Amplification Failure

We failed to obtain a genotype from 0.4% ($N = 2$) of our PCR reactions from dung-extracted DNA due to nondetection of alleles. Although laboratory error cannot be excluded, the observation that both these instances occurred at the locus with the largest alleles (387 bp) makes amplification failure a probable cause. Flagstad et al. (1999) observed failure rates of 3.5% with sheep and 5% with reindeer, and Morin et al. (2001) reported 21% with chimpanzee fecal samples. Gagneux et al. (1997) reported that "approximately 50% of all shed hair extracts gave amplification products in at least some of the amplifications." Goosens et al. (1998) had only a single PCR failure in 350 PCRs from single-hair extracts and none from extracts from 3 or 10 hairs.

Amplification failure occurs from absolute low amounts of template DNA or the presence of substances that inhibit PCR. Morin et al. (2001) demonstrated that the quantity of target DNA available for amplification in dung extracts varied widely, while that in extracts from hair was uniformly low. Our dilution experiment confirms the wide variation of available target DNA in our extracts but suggests that the minimum available was well above the threshold for amplification failure.

Extracts from dung contain large amounts of polysaccharides, pigments, RNA, and other substances that can inhibit PCR (Flagstad et al. 1999; Morin et al. 2001). It has been suggested that increasing the amount of extract in a PCR reaction will not improve success, as it would also likely increase the concentration of inhibitors (Morin et al. 2001). The results of our test of PCR inhibition supports the presence of inhibitors, but confirms that inhibition of amplification from contaminants was not a major problem. While procedures that diligently remove inhibitors can improve amplification success, there is a tradeoff between cleaning up extracts and DNA quality, as we have found that the greater the processing, the greater the possibility of DNA degradation. We believe the two-stage extraction protocol that we employed strikes an ideal balance between the two, with the phenol/chloroform stage removing contaminants soluble in the organic phase as well as large insoluble fragments, and the silica-based QIAGEN column stage removing contaminants soluble in the aqueous phase. PCR inhibition by any contaminants copurifying with DNA was overcome by the addition of BSA in our sample mix.

Because primer sensitivity can differ by several orders of magnitude (He et al. 1994), it determines the thresholds of DNA and inhibitor concentrations allowing positive amplification. The use of highly sensitive and specific primers and optimization of annealing temperatures, reagent concentrations, and PCR conditions is important in decreasing amplification failure with noninvasive samples.

Amplification Artifacts

Generation of amplification artifacts by polymerase slippage during PCR can cause "false heterozygotes," if a stutter band of a homozygous allele is identified as a true allele, or "multiple alleles," if stutter bands cannot be reliably differentiated from true alleles in a heterozygote. In both instances, the "false alleles" will be the first stutter bands of the true alleles, and hence will be smaller in size by one repeat from the true allele. While we observed 0.8% false heterozygotes and 0.2% multiple alleles in our amplifications from dung samples, amplification artifacts can be excluded as a cause of the false heterozygotes, because the false alleles observed were different by more than one repeat from the true alleles (Table 2). In the case of the single instance of multiple alleles observed, we cannot eliminate amplification artifacts as a cause, since the false allele observed was also an allele found in the population, nor can we exclude laboratory error or sporadic contamination as a cause. Thus, error attributable to amplification artifacts in our study was absent or negligible. Goosens et al. (1998) observed 4%, 1.43%, and 0% amplification artifacts in PCR of extracts from 1, 3, and 10 hairs, respectively. Flagstad et al. (1999) and Morin et al. (2001) do not report amplification artifacts in their studies. Generation of false alleles from polymerase slippage is greatest with di-, less with tri-, and does not occur with tetranucleotide loci (Morin et al. 2001; Schlotterer and Tautz 1992; Taberlet et al. 1999). Therefore, if false alleles were a major cause of error in genotyping, we would expect to see a preponderance of false heterozygotes and multiple alleles with the two dinucleotide loci, which we did not. In a study using all three types of loci, Gagneux et al. (1997) similarly found no major differences in rates of error that could be attributed to amplification artifacts at di- versus tri- or tetranucleotide loci. We believe the use of a consistent set of allele calling guidelines in the current study was instrumental in decreasing or eliminating error from amplification artifacts.

Electrophoresis Artifacts

"False alleles" from electrophoresis artifacts are smaller by 4–6 bp than their true alleles and could give rise to false heterozygotes, multiple alleles, and incorrect assignment. We followed a gel running protocol that eliminated or reduced electrophoresis artifacts (Fernando et al. 2001a), and our results suggest that electrophoresis artifacts were not a problem in our study. However, in contrast to other causes of error, we note that electrophoresis artifacts are more likely to occur with good sources of DNA such as blood, rather than with noninvasive sources such as dung or hair (unpublished data).

Size Assessment Error

Sizing of alleles using GENESCAN, although undoubtedly much more reliable than eyeballing autorads, is still not a perfect science. This is a particular problem with dinucleotide loci when alleles differ by a single repeat, and differences in estimates of the size of an allele can lead to mistakes in allele assignment. We were able to overcome problems due to incorrect allele assignment by running samples multiple times

Table 4. Variation in size estimates for seven alleles at dinucleotide locus Laf MS03

Allele	Assigned size (bp)	Mean	Range	SD	*N*
a	133	133.04	132.65–133.22	0.144	16
b	137	137.3	136.65–138.07	0.237	65
c	139	139.5	138.63–140.51	0.29	73
d	147	146.26	145.07–146.89	0.376	31
e	149	148.61	147.70–148.92	0.266	73
f	151	150.91	150.58–151.00	0.127	17
g	155	154.95	154.64–155.22	0.148	24

and identifying the range of values that could be obtained for a particular allele (Table 4). Knowledge of allele sizes present in the population can help in identifying possible errors in size assessment, which can then be confirmed or refuted by rerunning the samples at several dilutions.

Our success in genotyping noninvasive samples has been achieved through careful optimization of collection, extraction, and amplification protocols. A few important considerations in sample collection and the application of noninvasive genotyping to other species are discussed below.

Age of Dung in Sample Collection

While the freshness of dung in our study likely had a positive effect on the quality and quantity of DNA, the greatest consequences of error in genotyping are in studies requiring individual identification for assessing relatedness and kinship (Taberlet et al. 1999). For such studies, sampling of particular individuals requires observation of defecation, and thus, sample freshness should not be a concern. In studies conducted in our laboratory, we have had success rates of over 95% in amplifying microsatellites with field-collected samples from free-ranging elephants. However, as a rule we try to collect samples that are less than 24 h old and those that are likely to have dried quickly—as in exposed or dry conditions—and not samples that have been moist for long periods or subject to rain, etc. Amplification success from samples that have been in the field for longer periods would be less, depending on environmental variables, such as the ambient temperature, humidity, exposure to sun, etc., that would determine bacterial activity, and hence DNA degradation. In tests carried out to assess the age of dung that will provide amplifiable mitochondrial DNA, samples up to 8 days old [elephant (Fernando et al. 2000)] and 3 months old [lynx (Palomares et al. 2002)] provided positive results. However, genotyping success is determined by the amount of target DNA in samples, which depends on environmental and species-specific variables and the sensitivity of the primers used. Therefore, it is not possible to set universal limits as to what age samples should be collected, other than to suggest that samples less than 24 h old would be preferable.

Sample Preservation

We have found that proper collection and preservation of samples is critical for high success. Methods of preservation such as adding a buffer or 95% ethanol are preferable to freezing or desiccation, both in terms of ease under field conditions and continuation of bacterial activity (hence DNA degradation) in frozen or desiccated samples each time they are thawed or become rehumidified. We have had equal success with using storage buffer and 95% ethanol as preservatives and have successfully amplified from samples stored in either medium for periods of up to 3 years. However, the ratio of preservative to sample is more critical with buffer. Because DNA degradation will continue in the presence of an insufficient volume of preservative, if sample collection is by untrained personnel, ethanol may be preferable.

Application to Other Species

Ongoing studies in our laboratory (e.g., rhinos, elephants, apes, monkeys, and canids) and others that have found a high degree of success in applying noninvasive genotyping from dung to diverse taxa (Flagstad et al. 1999; Frantzen et al. 1998) suggest that our success is not taxa limited. As the target for DNA extraction from dung is epithelial cells sloughed during gut passage of dung, the quantity of DNA obtained may vary with the physiological state of the individual sampled and the diet and digestive system of the species. Therefore, it would be prudent to assess the success of genotyping from dung and determine the optimum quantity of dung that should be used in extractions before embarking on a large-scale study on a new species. In our experience, selection of primers and PCR optimization is perhaps the single most important factor in increasing the success of genotyping with dung. In applying this technique to a new species, we suggest that the extraction protocol presented in this study could be used as a starting point and that a preliminary study be conducted to determine the success of genotyping with dung and to modify the protocol as required. A suggested course of action in adapting this protocol to a new species is presented in Figure 3. We recommend initial screening with mitochondrial DNA, using universal primers to control for problems of primer sensitivity and optimization associated with untested microsatellite primers. While amplification success of microsatellite loci may be less than for mitochondrial fragments, due to the lower copy number, if good mitochondrial amplification is achieved for a majority of samples, a high percentage of them should also amplify for microsatellites.

Conclusion

Concern over the reliability of dung as a source of DNA for genetic analysis stems from the assumption that all noninvasive sources of DNA have similar attributes. For example, the amount of DNA obtainable from noninvasive samples is generally considered to be limiting and to preclude more than one study (Taberlet et al. 1999). However, this does not apply to dung, as the raw material is not limiting, and extractions can be scaled up to obtain large total volumes

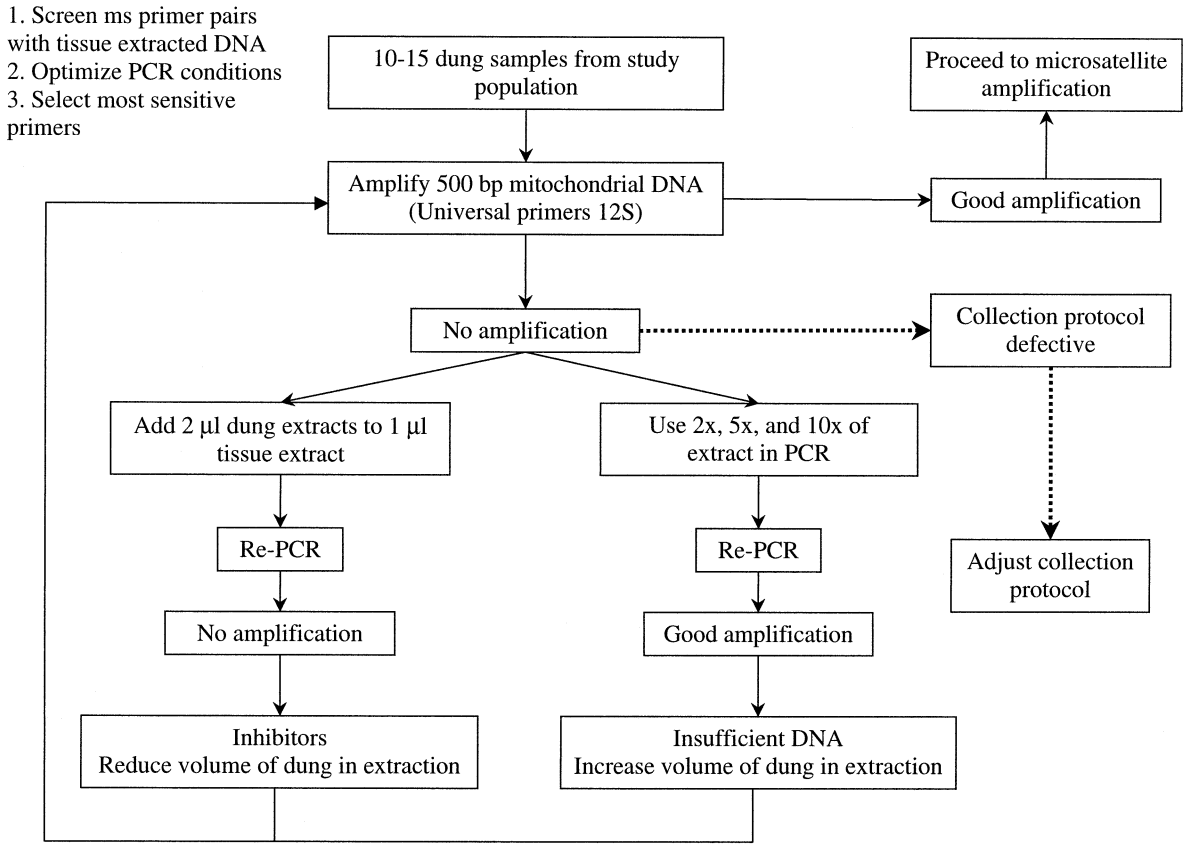

Figure 3. Summary of proposed course of action in starting a study employing genotyping from dung. The dotted lines represent a second round of amplification subsequent to adjustments to extraction protocol.

of DNA. Noninvasive samples encompass a wide range of sources, and it is not surprising that they should differ in their characteristics pertaining to genetic analysis. Of the two widely employed sources of noninvasive samples, hair and dung, many studies attempting to amplify DNA from hair extracts have found it unreliable, possibly because of low DNA yield (Morin et al. 2001), whereas the results from dung have been mixed.

The main causes of error in our study were sporadic contamination/laboratory error, consistent with genotyping from DNA samples of nanogram–microgram (300–300,000 copies of a unique sequence) range (Navidi et al. 1992). When the DNA concentration is very low (picogram range), other causes such as allelic dropout and false alleles become important in addition (Morin et al. 2001; Navidi et al. 1992; Taberlet et al. 1999), and genotyping becomes unreliable. At such low concentrations, methods such as the multiple tubes approach (Navidi et al. 1992; Taberlet et al. 1999) and quantifying target DNA (Morin et al. 2001) are likely to improve results substantially. However, the additional time and money required for such analyses could limit the use of noninvasive samples to situations where obtaining conventional samples is impossible or could entail even greater investments in time and funds.

The level of error we have found in this carefully controlled experiment would have little effect on use of microsatellite markers to investigate questions of population structure, gene flow, or other ecological issues. Similarly, they should have little effect on studies requiring individual identification, such as paternity exclusion, because single locus discrepancies between an offspring and its putative father are usually confirmed by retyping before acceptance. Our study suggests that dung is a source of DNA providing reliable amplification of microsatellite markers, that PCR inhibition from copurifying substances is not a major problem, and that two PCR replicates per sample, ideally from two extracts, are sufficient to eliminate genotyping error. We therefore conclude that dung is a source of DNA that can be reliably genotyped and thus allows geneticists and ecologists to realize the full potential of noninvasive genotyping.

References

Fernando P, Evans B, Morales JC, and Melnick DJ, 2001a. Electrophoresis artifacts—a previously unrecognized cause of error in microsatellite analysis. Mol Ecol Notes 1:325–328.

Fernando P and Lande R, 2000. Molecular genetic and behavioral analyses of social organization in the Asian elephant. Behav Ecol Sociobiol 48:84–91.

Fernando P, Pfrender ME, Encalada S, and Lande R, 2000. Mitochondrial DNA variation, phylogeography, and population structure of the Asian elephant. Heredity 84:362–372.

Fernando P, Vidya TNC, and Melnick DJ, 2001b. Isolation and characterization of tri- and tetranucleotide microsatellite loci in the Asian elephant, *Elephas maximus*. Mol Ecol Notes 1:232–233.

Flagstad O, Roed K, Stacy JE, and Jakobsen KS, 1999. Reliable noninvasive genotyping based on excremental PCR of nuclear DNA purified with a magnetic bead protocol. Mol Ecol 8:879–883.

Frantzen MAJ, Silk JB, Ferguson JWH, Wayne RK, and Kohn MH, 1998. Empirical evaluation of preservation methods for faecal DNA. Mol Ecol 7:1423–1428.

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

Goosens B, Waits LP, and Taberlet P, 1998. Plucked hair samples as a source of DNA: reliability of dinucleotide microsatellite genotyping. Mol Ecol 7:1237–1241.

He Q, Marjamaki M, Soini H, Mertsola J, and Viljanen NK, 1994. Primers are decisive for sensitivity of PCR. Biotechniques 17:82–85.

Kohn M and Wayne RK, 1997. Facts from feces revisited. TREE 12:223–227.

Litvaitis MK and Litvaitis JA, 1996. Using mitochondrial DNA to inventory the distribution of remnant populations of New England cottontails. Wildl Soc Bulletin 24:725–730.

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

Navidi W, Arnheim N, and Waterman MS, 1992. A multiple-tubes approach for accurate Genotyping of very small DNA samples by using PCR: statistical considerations. Am J Hum Genet 50:347–359.

Nyakaana S and Arctander P, 1998. Isolation and characterisation of microsatellite loci in the African elephant (*Loxodonta africana* Blumenbach 1797). Mol Ecol 7:1436–1437.

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

Sambrook E, Fritsch EF, and Maniatis T, 1989. Molecular cloning: a laboratory manual, 4th ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Schlotterer C and Tautz D, 1992. Slippage synthesis of simple sequence DNA. Nucleic Acids Res 20:211–215.

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

Taberlet P, Waits LP, and Luikart G, 1999. Noninvasive genetic sampling: look before you leap. TREE 14:323–327.

Received May 29, 2002
Accepted December 31, 2002

Corresponding Editor: Stephen J. O'Brien