Conserv Genet (2009) 10:2005–2007
DOI 10.1007/s10592-009-9880-6

TECHNICAL NOTE

Sampling error in non-invasive genetic analyses of an endangered social carnivore

Penny A. Spiering · Micaela Szykman Gunther · David E. Wildt · Michael J. Somers · Jesús E. Maldonado

Received: 23 December 2008 / Accepted: 24 February 2009 / Published online: 5 March 2009

Abstract Modern non-invasive genetic technologies are useful in studies of rare and difficult-to-observe species. An examination of endangered African wild dog (*Lycaon pictus*) faecal DNA revealed that 11.4% of samples were assigned incorrectly to an individual. Sampling mistakes in the field are not normally considered in non-invasive genetic assessments, but can be a significant source of error. To ensure meticulous data interpretation, non-invasive genetic studies should track and report sampling inaccuracies.

Keywords *Lycaon pictus* · Non-invasive · Microsatellites · Sampling error · Faecal sampling

Faecal genetic analysis is a powerful tool that eliminates handling of stress-sensitive animals (Kohn and Wayne 1997) while allowing studies of social organization and mating systems (Archie et al. 2008), but only when errors mis-assigning kinship are eliminated. Under rigorous field conditions, investigators may incorrectly identify an individual or scat or erroneously label the sample storage container, resulting in sampling error. As error likelihood increases when studying large, dynamic groups defecating in the same area, non-invasive genetic analyses have been under-exploited for studying kinship within social groups. Furthermore, minimal attention has been dedicated to sampling error (Bonin et al. 2004), potentially resulting in data misinterpretation.

The African wild dogs of KwaZulu-Natal (KZN), South Africa originated from ongoing reintroduction efforts initiated in the 1980s (see Gusset et al. 2008) and currently comprise eight packs distributed across three protected areas. This is one of the most endangered (Woodroffe et al. 1997) and socially complex species in Africa (Estes and Goddard 1967). Wild dogs live as cooperative breeders in large-sized packs consisting of as many as 27 individuals (Creel and Creel 2002). Minimizing pack disruption is essential for these sensitive animals, and faecal sampling provides an excellent opportunity for understanding genetic relatedness.

To investigate errors in faecal genetic sampling techniques, we compared 42 individual tissue/blood DNA genotypes to faecal DNA genotypes to determine the level of sampling error in our field studies. To further reduce inaccuracies, we evaluated multiple samples collected from 22 individuals and screened a large number of microsatellite loci to ensure sufficient power to identify unique individuals in our dataset (Waits et al. 2001).

From January 2003 through January 2008, defecating wild dogs were observed from $\geq$15 m away and samples collected after the animals moved away to minimize

P. A. Spiering (✉) · J. E. Maldonado
Center for Conservation and Evolutionary Genetics, National Zoological Park, Smithsonian Institution, Washington, DC 20008, USA
e-mail: SpieringP@si.edu

P. A. Spiering · M. S. Gunther · D. E. Wildt
Center for Species Survival, Conservation and Research Center, National Zoological Park, Smithsonian Institution, Front Royal, VA 22630, USA

P. A. Spiering · M. J. Somers
Centre for Wildlife Management, University of Pretoria, Pretoria 0002, South Africa

M. S. Gunther
Department of Wildlife, Humboldt State University, Arcata, CA 95521, USA

M. J. Somers
Centre for Invasion Biology, University of Pretoria, Pretoria 0002, South Africa

disturbance. Since reintroduction or birth, each pack of wild dogs has been monitored intensively, and individuals are recognized by unique natural markings (e.g. coat patterns, ear tears) documented in photographs. Samples of questionable identification or samples possibly mixed with another individual's faeces due to over-marking were not collected. Faeces was collected in plastic bags and transported to a −20°C freezer within 4 h of collection. Invasive samples were collected from a subset of the population during translocation and collaring. During anaesthesia, an ear biopsy was taken and placed in DMSO solution, while blood was aspirated from the femoral vein into a vacutainer tube containing EDTA anticoagulant and stored at −20°C. Based on field observations, we predicted that 136 faecal samples were derived from 114 individuals, with 22 sampled twice; 42 animals also contributed an invasive sample (Table 1).

Faecal DNA was extracted with the QIAamp DNA Stool Mini Kit and tissue and blood DNA extracted with the QIAamp Tissue and Blood Kit (QIAGEN) in a separate room from PCR products (Eggert et al. 2005). We screened 19 microsatellite loci from the 2006 International Society for Animal Genetics domestic dog panel (Table 2). For all loci, faecal samples were genotyped at least three times for heterozygotes and five times for homozygotes. Invasive samples were genotyped once for each locus and replicated when allelic scores were questionable or failed to match faecal genotypes. Each 25 µl PCR reaction contained 3.0 µl of template DNA, 1.0 µl of forward and reverse primers, 2.0 µl $MgCl_2$, 0.2 µl *Taq Gold* polymerase, 2.5 µl *Taq* buffer (Applied Biosystems), 2.0 µl BSA and 2.5 µl dNTPs. The PCR profile included 10 min at 95°C, followed by 38 cycles for faecal DNA or 30 cycles for tissue/blood DNA of 1 min at 95°C, 1 min at 60°C and 1 min at 72°C, and an extension of 10 min at 72°C. Samples were run with GS-500 ROX size standard on ABI PRISM 3100 or 3130 automatic sequencers (Applied Biosystems) and alleles determined with Genotyper (Perkin Elmer) or GeneMapper (Applied Biosystems) software.

The mean amplification success rate for non-invasive samples was 63% (Table 2). Errors associated with allelic dropout (37%; Table 2) and false alleles (3%) accounted for most of the genotyping inaccuracies in this study. Consensus faecal genotypes were assessed and genotypes deemed <95% reliable by RelioType software (Miller et al. 2002) were replicated at specific loci. Although genotyping errors were significant in our faecal DNA investigation, as well as those of others (Bonin et al. 2004), it was possible to correct allelic dropout errors to generate consistent genotypes.

Once reliable genotypes were achieved, we checked for further errors that would have occurred in the field at the time of sample collection. Of the 22 individuals with two faecal samples for comparison, we found inconsistent genotypes for three samples (13.6%), but were able to elucidate identities of these individuals with invasive samples. Comparing genotypes from tissue/blood and faeces showed that four of 42 (9.5%) individual donors were sampled in error. All erroneously identified samples matched another individual in the pack. Lastly, CERVUS 2.0 (Marshall et al. 1998) determined that our loci yielded highly significant probability of identity (4.78×10^{-14}) and probability of identity among siblings (2.1×10^{-6}) values. Assessing genotypes of 110 faecal samples labelled as different individuals (excluding the four samples not matching invasive samples) revealed that only 104 were unique individuals, resulting in 5.5% sampling error. In summary, three evaluative methods

Table 1 Summary of sampling error results for 136 faecal samples

	Number of individuals (incorrect samples)	Error (%)
Individuals observed in field	114	
Replicate samples	22 (3)	13.6
Blood/tissue and faecal comparison	42 (4)	9.5
Probability of identity (CERVUS)	110 (6)	5.5
Overall sampling error	114 (13)	11.4

Table 2 Locus names, size ranges (base pairs), amplification success and allelic dropout rates for African wild dog faecal samples

Locus	Size range	Amplification success (%)	Allelic dropout (%)
AHT137	131–147	67	24
AHTh130	117–125	88	9
AHTh171	217–225	42	49
AHTh260	246–254	37	44
AHTk211	89–91	73	27
AHTk253	298–306	65	35
CXX279	116–118	85	22
FH2054	128–140	59	32
FH2328	194–220	68	24
FH2848	232–240	54	41
INRA21	97–101	59	29
INU030	144–150	64	40
INU055	208–216	62	51
LEI004	95–99	57	49
REN54P11	234–246	43	34
REN105L03	235–245	83	41
REN162C04	194–200	72	55
REN169DO1	208–212	68	48
REN247M23	254–256	51	44
Mean	234–246	63	37

detected sampling errors due to incorrectly identifying faeces donors or mislabelling containers 13 times in 114 faecal samples (11.4%; Table 1).

Although genotyping errors from faecal DNA analyses were anticipated, sampling inaccuracies were unexpected. Previously, faecal DNA studies have not considered the confounding influence of observer error and mis-assignment of sample source, although this is well addressed in endocrine (Keay et al. 2006; Stavisky et al. 2001) and demographic studies (Gabriele et al. 2001; Milligan et al. 2003; Kelly 2001). Difficult field conditions common to wildlife research can introduce mistakes to faecal sample recovery. For example, in our study, the investigator must wait for the animal to move away and then leave the vantage point to collect the sample. Also this species lives in a social, tight-knit group, and it is common to encounter a significant number of wild dogs defecating within the same area and in dense bush. A semi-natural captive environment apparently increased the likelihood of misidentification, as four sampling inaccuracies occurred while animals were in large enclosures for reintroduction and defecated in the same area daily.

We demonstrate that sampling error in faecal genetic studies can be reduced and allow the resolution of genotypes of misidentified animals to ensure accurate data. We recommend photographing defecating animals, collecting multiple samples from individuals, comparing invasive and faecal samples, and using sufficient loci for adequate power to detect duplicated individuals.

Acknowledgments This research was supported by Smithsonian Institution, University of Pretoria, Rotterdam Thandiza Fund, Humboldt State University, Association of Zoos and Aquariums, Disney Wildlife Conservation Fund, Knowsley Safari Park, DST-NRF Centre of Excellence for Invasion Biology, Pittsburgh Zoo and Morris Animal Foundation.

References

Archie EA, Maldonado JE, Hollister-Smith JA et al (2008) Fine-scale population genetic structure in a fission–fusion society. Mol Ecol 17:2666–2679. doi:10.1111/j.1365-294X.2008.03797.x

Bonin A, Bellemain E, Eidesen PB et al (2004) How to track and assess genotyping errors in population genetic studies. Mol Ecol 13:3261–3273. doi:10.1111/j.1365-294X.2004.02346.x

Creel S, Creel NM (2002) The African wild dog: behavior, ecology and conservation. Princeton University Press, Princeton

Eggert LS, Maldonado JE, Fleischer RC (2005) Nucleic acid isolation from ecological samples: animal scat and other associated materials. In: Molecular evolution: producing the biochemical data, part B. Methods Enzym 395:73–87

Estes R, Goddard J (1967) Prey selection and hunting behaviour of the African wild dog. J Wildl Manage 31:52–70. doi:10.2307/3798360

Gabriele CM, Straley JM, Mizroch SA et al (2001) Estimating the mortality rate of humpback whale calves in the central North Pacific Ocean. Can J Zool 79:589–600. doi:10.1139/cjz-79-4-589

Gusset M, Ryan SJ, Hofmeyr M et al (2008) Efforts going to the dogs? Evaluating attempts to reintroduce endangered wild dogs in South Africa. J Appl Ecol 45:100–108

Keay JM, Singh J, Gaunt MC, Kaur T (2006) Fecal glucocorticoids and their metabolites as indicators of stress in various mammalian species: a literature review. J Zoo Wildl Med 37:234–244. doi:10.1638/05-050.1

Kelly MJ (2001) Computer-aided photograph matching in studies using individual identification: an example from Serengeti cheetahs. J Mammal 82:440–449. doi:10.1644/1545-1542(2001)082<0440:CAPMIS>2.0.CO;2

Kohn MK, Wayne RK (1997) Facts from feces revisited. Trends Ecol Evol 12:223–227. doi:10.1016/S0169-5347(97)01050-1

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 7:639–655. doi:10.1046/j.1365-294x.1998.00374.x

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

Milligan JL, Davis AK, Altizer SM (2003) Errors associated with using colored leg bands to identify wild birds. J Field Ornithol 74:111–118

Stavisky RC, Whitten PL, Hammett DH, Kaplan JR (2001) Lake pigments facilitate analysis of faecal cortisol and behavior in group-housed macaques. Am J Phys Anthropol 116:51–58. doi:10.1002/ajpa.1101

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. Mol Ecol 10:249–256. doi:10.1046/j.1365-294X.2001.01185.x

Woodroffe R, Ginsberg JR, Macdonald DW (1997) The African wild dog- status survey and conservation action plan. IUCN, Gland