Eur J Wildl Res (2011) 57:1231–1235
DOI 10.1007/s10344-011-0556-4

TECHNICAL NOTES

Genotyping faeces of red pandas (*Ailurus fulgens*): implications for population estimation

Yu Guo · Yibo Hu · Dunwu Qi · Xiangjiang Zhan · Michael W. Bruford · Fuwen Wei

Received: 19 August 2010 / Revised: 28 April 2011 / Accepted: 13 June 2011 / Published online: 23 June 2011

Abstract The red panda (*Ailurus fulgens*) is an endangered species distributed in the Himalaya and Hengduan Mountains and extremely difficult to monitor because it is elusive, wary and nocturnal. However, recent advances in noninvasive genetics are allowing conservationists to indirectly estimate population size of this animal. Here, we present a pilot study of individual identification of wild red pandas using DNA extracted from faeces. A chain of optimal steps in noninvasive studies were used to maximize genotyping success and minimize error rate across sampling, selection of microsatellite loci, DNA extraction and amplification and data checking. As a result, 18 individual red pandas were identified successfully from 33 faecal samples collected in the field using nine red panda-specific microsatellite loci with a low probability of identity of 1.249×10^{-3} for full siblings. Multiple methods of tracking genotyping error showed that the faecal genetic profiles possessed very few genotyping errors, with an overall error rate of 1.12×10^{-5}. Our findings demonstrate the feasibility and reliability of using faeces as an effective source of DNA for estimating and monitoring wild red panda populations.

Communicated by P. C. Alves

Y. Guo · Y. Hu · D. Qi · X. Zhan · F. Wei (✉)
Key Lab of Animal Ecology and Conservation Biology,
Institute of Zoology, Chinese Academy of Sciences,
1-5 Beichen West Road,
Beijing 100101, China
e-mail: weifw@ioz.ac.cn

M. W. Bruford
Organisms and Environment Division,
Cardiff School of Biosciences, Cardiff University,
Cardiff CF10 3AX, UK

Keywords Faeces · Microsatellite · Noninvasive genetic sampling · Population estimation · Red panda

Introduction

The red panda (*Ailurus fulgens*) is a highly specialized carnivore that feeds primarily on bamboo. Historically, red pandas were distributed throughout Eurasia (Roberts and Gittleman 1984); however, due to widespread habitat loss and fragmentation and poaching, they are now confined to the Himalaya and Hengduan Mountains. Little is known about the exact population size of red pandas because they are elusive, wary and essentially nocturnal. Previous studies have surveyed for feeding sites and faeces along transects and used interviews with villagers and officials to estimate the population size (Wei et al. 1999; Choudhury 2001); however, these methods are not accurate due to lack of direct individual identification. Recent advances in molecular genetics have allowed the use of noninvasive samples such as faeces or hairs for population estimation of endangered or elusive species (see review in Beja-Pereira et al. 2009). Moreover, the list of taxa studied by noninvasive genetics is increasing. To date, this tool has not been applied to red pandas and their conservation.

It has been argued that genetic studies based on noninvasive samples should be preceded by a pilot study to assess the probability of identity, feasibility and reliability of the method before any large-scale study (Taberlet et al. 1999). A pilot study should demonstrate error rates encountered during the genotyping process using noninvasive DNA sources (Gagneux et al. 1997). Here, we collected red panda faeces from the wild and performed a pilot study to establish a protocol for obtaining reliable genetic profiles from faecal samples. Specifically, we

determined whether faecal DNA analysis is an effective approach for identifying individuals and potentially estimating population size in this species.

Materials and methods

Sample collection and preservation

To test for faecal genotyping reliability, matched samples from blood and fresh faeces (N=3) were obtained from three captive red pandas in the Fengtongzhai Nature Reserve. Faeces were immediately collected after deposition. Blood was stored in EDTA vacutainers, and faeces were stored in sterile tubes. All samples were frozen at −20°C.

To assess genotyping reliability of faeces from the wild, 33 faecal samples were collected from the Heizugou, Meigu-Dafengding, Mabian-Dafengding and Shenguozhuang Nature Reserves in the Liangshan Mountains from April–July in 2006. Most faecal samples were less than 2 weeks old based on the freshness of the mucosal outer layer (e.g. Zhan et al. 2006). Whole faeces were collected using disposable plastic gloves in the wild. Given that most intestinal cells cling to the outer layer of faeces, we peeled 5–10 g of faeces from the outer layer and stored them in 99.7% ethanol at room temperature.

Molecular analysis

Blood samples were extracted using the phenol/chloroform method (Sambrook et al. 1989). Faecal DNA was extracted following Zhang et al. (2006). Blank controls were included in extraction and downstream amplification. A 551-bp mitochondrial DNA control region (mtDNA CR) was amplified to corroborate whether extracts are from red pandas, using a pair of primers (the forward by Su et al. 2001 and the reverse, H16781, by Li et al. 2005). Amplification was carried out in 20 μl containing 1–2 μl DNA, 10 μl Premix *Ex Taq* (TaKaRa), 0.4 μM forward and reverse primers and 0.5 μg/μl bovine serum albumin (BSA) (Sigma). The PCR procedure was as follows: initial denaturation at 94°C for 5 min, followed by 35 cycles of PCR (94°C/30 s, 54°C/40 s, 72°C/50 s) and a final step of 72°C for 10 min. The PCR product (5 μl) was electrophoresed on 2% agarose gel. Positive PCR products were purified and subsequently sequenced using an ABI 3100 Sequencer (Applied Biosystems).

Faecal samples successfully amplified by mtDNA were further analysed with a set of microsatellite loci. Nine microsatellite loci with alleles less than 220 bp: Aifu-1, Aifu-2 and Aifu-23 (Liang et al. 2007), RP-11, RP-102 and RP-137 (Liu et al. 2005), RP-108, RP-133 and RP-215 (redesigned), were chosen for this study according to allele length, polymorphism, amplification performance and scoring (Table 1). A modified multi-tube approach (Taberlet et al. 1996; Zhan et al. 2006) was used to obtain reliable consensus genotypes. Firstly, each extract was amplified twice simultaneously, and the locus that produced the same heterozygous genotype was accepted as heterozygous. Otherwise, a third repeat was conducted, and loci in which two alleles occurred at least twice were determined as heterozygous. If not, four additional positive repeats were performed. PCR was performed in 20 μl containing 1–2 μl DNA, 10 μl HotStar *Taq* Master Mix (Qiagen), 0.2 μM forward primer end labelled with a fluorescent dye (HEX, FAM or TAMRA), 0.2 μM reverse primer and 0.5 μg/μl BSA (Sigma). The first step began with 94°C for 15 min, followed by a touchdown PCR (a total of 35–40 cycles of 94°C/30 s, T_{anneal}/40 s, 72°C/50 s) and a final step of 72°C for 15 min. T_{anneal} was decreased by 2°C every second cycle from 60°C to a final temperature (48–50°C), which was used for the following 25 cycles. Positive products were genotyped using an ABI 3730 sequencer (Applied Biosystems). Genotypes were scored using GeneScan v3.7 and Genotyper v3.7 (Applied Biosystems).

Data analysis

To directly test reliability of faecal consensus genotypes derived from the multi-tube method, the faecal consensus genotypes of three captive individuals were compared with respective reference genotypes generated from blood DNA.

For faecal samples collected from the field, unique genotypes were identified following Zhan et al. (2006). Genotypes from different samples were considered to be from the same individual when all alleles at all loci were identical. Based on the set of unique genotypes of wild faecal samples, the probabilities of pairs of individuals and full siblings bearing an identical multilocus genotype, P(ID) and P(sibs), and genetic polymorphism were estimated using GIMLET 1.3.3 (Valière 2002), and the Hardy–Weinberg equilibrium was tested for each locus using GENEPOP 4.0 (Rousset 2008).

Multiple methods were performed to detect whether genotyping errors in the dataset were reduced to a non-significant level. First, the confidence of each observed multilocus genotype was estimated using RELIOTYPE (Miller et al. 2002). To be conservative, only genotypes with ≥99% reliability were considered ‘acceptable’. Second, a mathematical method (Zhan et al. 2010) applicable to the multi-tube genotyping was used to estimate overall genotyping error rate among successfully genotyped samples, considering allele dropout and false allele rates. Finally, for the final individual identification result, MICRO-CHECKER (Van Oosterhout et al. 2004) was used

Table 1 Information of nine microsatellite loci in the order of decreasing polymorphic information content, including their genetic polymorphisms and allele dropout rates based on faecal genotypes obtained from wild samples

Locus	Primer sequences or sources	PIC	*A*	H_E	H_O	P(ID)	P(sibs)	Dropout rate
RP-215	5′-GGGAACATTACTTTTGCT-3′ 5′-TACCACATCTGTACTTTTATTT-3′	0.749	6	0.78	0.56	0.0561	0.3784	0.0714
Aifu-23	Liang et al. (2007)	0.729	7	0.76	0.78	0.0577	0.3918	0.0746
RP-108	5′-GCAAAAGCACTGTTCAGAAT-3′ 5′-ACCTTGTCACTCCCAGCCTCC-3′	0.717	5	0.76	0.78	0.0727	0.3961	0.05
Aifu-1	Liang et al. (2007)	0.689	7	0.72	0.78	0.0739	0.416	0.0921
RP-137	Liu et al. (2005)	0.673	5	0.72	0.78	0.0952	0.4211	0.0908
Aifu-2	Liang et al. (2007)	0.518	4	0.60	0.78[a]	0.2139	0.5102	0.0313
RP-102	Liu et al. (2005)	0.506	4	0.57	0.83	0.2083	0.5275	0.0551
RP-133	5′-GAGAATGGAAGCCTGGAAC-3′ 5′-TGAGTCAGCAGGAAACACCA-3′	0.468	3	0.54	0.78	0.2435	0.5523	0.0952
RP-11	Liu et al. (2005)	0.178	2	0.20	0.11	0.6209	0.8171	0.00

PIC polymorphic information content, *A* number of alleles, H_E expected heterozygosity, H_O observed heterozygosity, *P(ID)* unbiased probability of identity for each locus, *P(sibs)* probability of identity for full siblings, *Dropout rate* rate of allele dropout at heterozygous locus (false allele rate was not reported due to no detection of false alleles)

[a] Significantly deviated from the Hardy–Weinberg equilibrium

to detect the presence of null alleles and genotyping errors such as large allele dropout or stuttering.

Results

Blood and fresh faecal samples of three captive red pandas were successfully extracted and amplified with clear target DNA bands (100% success) using mtDNA CR and nine microsatellite loci. Of 33 faecal samples from the field, 25 samples (76% success) were successfully amplified with mtDNA CR and were verified from red pandas by NCBI-BLAST. Of the 25 faecal samples, 20 samples were successfully amplified by at least seven microsatellite loci. Five samples were excluded in subsequent analyses as the three most informative loci RP-215, Aifu-1 and Aifu-23 were unsuccessfully amplified. No negative controls for extractions or PCRs produced positive results.

Comparison of the faecal genotypes of three captive individuals with respective reference genotypes derived from blood samples (Table 2) showed that there was only one case of false amplification at locus RP-215, which generated a low genotyping error rate per allele of 0.617% relative to the total number of alleles genotyped. The result indicates the reliability of faecal consensus genotypes derived from the multi-tube approach, although of a small sample size.

Pairwise comparisons of 20 genetic profiles from faecal samples of wild red pandas identified 18 unique genotypes (namely individuals) which differed by at least three alleles. Samples LS05 and LS08, with an empty locus each, did not match other consensus genotypes for at least three loci and were therefore considered unique genotypes. The genetic profiles of sample pairs LS21/LS22 and LS26/LS27 completely matched each other. Although both pairs involved one or two empty loci, P(ID) analyses based on genotypes of seven or eight loci suggested they are not likely from the same individual (data not shown). Based on the genotypes of 18 wild individuals, the unbiased P(ID) over the nine loci was 1.115×10^{-8}, and even for full siblings, there was a low P(sibs) of 1.249×10^{-3} (Fig. 1). These results demonstrate that the set of nine microsatellite loci has sufficient power for individual discrimination of wild red pandas.

Ignoring failed reactions, a total of 785 PCRs were performed for the 20 samples, and 175 consensus locus genotypes were obtained, with an average of 4.5 reactions per locus genotype. RELIOTYPE analysis showed that all of the 20 faecal genotypes were deemed with 99% reliability and needed no more PCR replicates. Following Zhan et al. (2010), the mean genotyping error rate per locus was 1.25×10^{-6}, and the overall error rate across nine loci was 1.12×10^{-5}, showing very low error rate among these faecal genetic profiles. MICRO-CHECKER indicated no evidence of genotyping errors such as large allele dropout or stuttering in our final dataset, but detected a signature of null allele at locus RP-215, suggesting this locus should be carefully monitored in future population genetics studies.

Discussion

Our study is the first attempt to use faeces for genetic studies of wild red pandas. The results show that 18

Table 2 Comparison of consensus genotypes generated from faecal samples of three captive red pandas with reference genotypes from blood of the same animals

Individual number	Sample type	Aifu-1	Aifu-23	RP-215	RP-108	RP-137	RP-102	Aifu-2	RP-133	RP-11
A032	Blood	154/158	133/145	142/146	191/201	158/158	199/209	137/141	117/127	132/132
	Faeces 1	154/158	133/145	142/146	191/201	158/158	199/209	137/141	117/127	132/132
	Faeces 2	154/158	133/145	142/146	191/201	158/158	199/209	137/141	117/127	132/132
	Faeces 3	154/158	133/145	142/146	191/201	158/158	199/209	137/141	117/127	132/132
A034	Blood	154/174	133/141	142/146	191/197	154/154	199/199	137/141	117/117	132/132
	Faeces 1	154/174	133/141	142/146	191/197	154/154	199/199	137/141	117/117	132/132
	Faeces 2	154/174	133/141	142/146	191/197	154/154	199/199	137/141	117/117	132/132
	Faeces 3	154/174	133/141	142/146	191/197	154/154	199/199	137/141	117/117	132/132
A038	Blood	154/170	129/141	138/146	197/201	154/154	199/209	129/141	117/127	132/132
	Faeces 1	154/170	129/141	138/146	197/201	154/154	199/209	129/141	117/127	132/132
	Faeces 2	154/170	129/141	138/146	197/201	154/154	199/209	129/141	117/127	132/132
	Faeces 3	154/170	129/141	138/*144*	197/201	154/154	199/209	129/141	117/127	132/132

The italicized allele '144' indicates one allele mismatch with the reference genotype

individuals were successfully identified from 20 faecal samples collected from the field. Importantly, these genetic profiles possess very few genotyping errors, with an overall error rate of 1.12×10^{-5}, demonstrating the reliability of using faeces to perform individual identification of wild red pandas.

For noninvasive genetic studies, optimal combination of a chain of steps from pre-PCR sampling to post-PCR data checking has been implicated in high-quality genetic datasets (Beja-Pereira et al. 2009). The quality of faecal samples is especially critical for amplification success and reliability of target DNA sequences. In order to obtain good faecal DNA, we performed some optimal steps applicable to field research. First, we mainly collected fresh faeces based on the fresh status of the mucosal outer layer because the time interval between deposition and sample collection can influence DNA quality and quantity. In the wild, faecal DNA is easily degraded due to humid environments and/or a long stay before collection. Second, we only used the mucosal outer layer of faeces to maximally enrich the intestinal shed cells of red pandas because the outer layer possesses obvious intestinal mucosal layer due to direct contact with the intestine. Third, we stored faecal samples in ethanol upon collection in the field. Faeces stored in ethanol have been shown to have the highest amplification success of both mtDNA and nuclear DNA as compared to other drying or buffer methods (Murphy et al. 2002). In addition to the optimal steps involved in faecal collection, we used proven methods in extraction and amplification of faecal DNA. For the extraction, we used a proven extraction protocol of faeces DNA developed in our laboratory (Zhang et al. 2006; Zhan et al. 2006). For the microsatellite amplification, we used a popular multi-tube method and high-specificity hot start polymerase to obtain reliable consensus genotypes.

The number and characteristics of microsatellite loci are very important for accurate population estimation. A set of unsuitable loci can result in population underestimation or overestimation due to a shadow effect or an increase of genotyping error rate (Creel et al. 2003). In this study, we performed a test of probability of identity to determine discrimination power of the set of loci genotyped and found

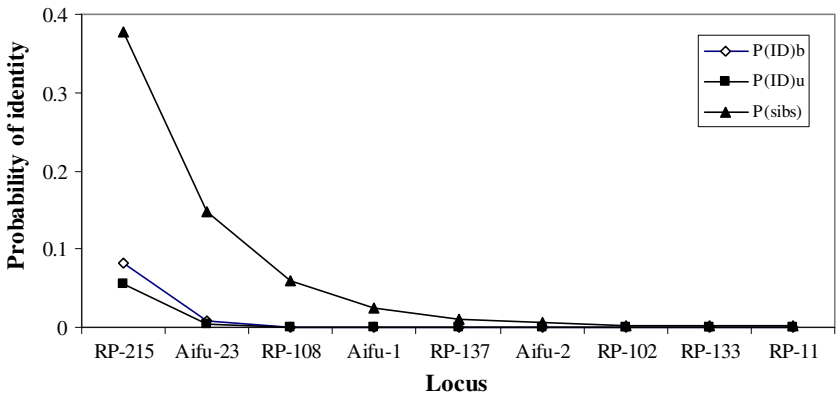

Fig. 1 Probabilities of identity for pairs of individuals and full siblings, P(ID) and P(sibs), in the order of increasing probability of identity of nine microsatellite loci, based on 18 unique genotypes of faecal samples from the wild. *P(ID)b* biased estimation of P(ID), *P(ID)u* unbiased estimation of P(ID) corrected for sample size

that nine loci are sufficient to discriminate 18 individuals, even for full siblings. But when applied to a large sample size, the number and power of loci need further evaluation. Additionally, we selected suitable microsatellite loci according to allele length, polymorphism, amplification performance and scoring to maximize noninvasive genotyping success and minimize error rate.

In summary, this pilot study demonstrates the feasibility and reliability of using red panda faeces to generate reliable genetic profiles for individual identification. This methodology could be applied to studies on population estimation and conservation genetics of red pandas in the field.

Acknowledgments This research was supported by the National Natural Science Foundation of China (30770399), Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-EW-Z-4) and the Royal Society. We thank two anonymous referees for helpful comments on this manuscript.

References

Beja-Pereira A, Oliveira R, Alves PC, Schwartz MK, Luikart G (2009) Advancing ecological understanding through technological transformations in noninvasive genetics. Mol Ecol Resour 9:1279–1301

Choudhury A (2001) An overview of the status and conservation of the red panda in India, with reference to its global status. Oryx 35:250–259

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

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

Li M, Wei FW, Goossens B, Feng ZJ, Tamate HB, Bruford MW, Funk SM (2005) Mitochondrial phylogeography and subspecific variation in the red panda (*Ailurus fulgens*): implications for conservation. Mol Phylogenet Evol 36:78–89

Liang X, Zhang ZH, Zhang L, Zhang WP, Shen FJ, Yang Z, Hou R, He W, Wei K, Yue BS (2007) Isolation and characterization of 16 tetranucleotide microsatellite loci in the red panda (*Ailurus fulgens*). Mol Ecol Notes 7:1012–1014

Liu ZJ, Zhang BW, Wei FW, Li M (2005) Isolation and characterization of microsatellite loci for the red panda, *Ailurus fulgens*. Mol Ecol Notes 5:27–29

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

Murphy MA, Waits LP, Kendall KC, Wasser SK, Higbee JA, Bogden R (2002) An evaluation of long-term preservation methods for brown bear (*Ursus arctos*) faecal DNA samples. Conserv Genet 3:435–440

Roberts MS, Gittleman JL (1984) *Ailurus fulgens*. Mammal Species 222:1–8

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Resour 8:103–106

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual (2nd edition). Cold Spring Harbour Press, Cold Spring Harbour

Su B, Fu YX, Wang YX, Jin L, Chakraborty R (2001) Genetic diversity and population history of the red panda (*Ailurus fulgens*) as inferred from mitochondrial DNA sequence variations. Mol Biol Evol 18:1070–1076

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

Taberlet P, Waits LP, Luikart G (1999) Noninvasive genetic sampling: look before you leap. Trends Ecol Evol 14:323–327

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

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Wei FW, Feng ZJ, Wang ZW, Hu JC (1999) Current distribution, status and conservation of wild red pandas (*Ailurus fulgens*) in China. Biol Conserv 89:285–291

Zhan XJ, Li M, Zhang ZJ, Goossens B, Chen YP, Wang HJ, Bruford MW, Wei FW (2006) Molecular censusing doubles giant panda population estimate in a key nature reserve. Curr Biol 16:451–452

Zhan XJ, Zheng XD, Bruford MW, Wei FW, Tao Y (2010) A new method for quantifying genotyping errors for noninvasive genetic studies. Conserv Genet 11:1567–1571

Zhang BW, Li M, Ma LC, Wei FW (2006) A widely applicable protocol for DNA isolation from faecal samples. Biochem Genet 44:503–512