Species identification refined by molecular scatology in a community of sympatric carnivores in Xinjiang, China

Alice LAGUARDIA¹, Jun WANG¹, Fang-Lei SHI¹,² Kun SHI¹,*, Philip RIORDAN¹,³

¹ The Wildlife Institute, Beijing Forestry University, Beijing 100083, China
² The Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China
³ Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS UK

ABSTRACT

Many ecological studies and conservation management plans employ noninvasive scat sampling based on the assumption that species’ scats can be correctly identified in the field. However, in habitats with sympatric similarly sized carnivores, misidentification of scats is frequent and can lead to bias in research results. To address the scat identification dilemma, molecular scatology techniques have been developed to extract DNA from the donor cells present on the outer lining of the scat samples. A total of 100 samples were collected in the winter of 2009 and 2011 in Taxkorgan region of Xinjiang, China. DNA was extracted successfully from 88% of samples and genetic species identification showed that more than half the scats identified in the field as snow leopard (Panthera uncia) actually belonged to fox (Vulpes vulpes). Correlation between scat characteristics and species were investigated, showing that diameter and dry weight of the scat were significantly different between the species. However it was not possible to define a precise range of values for each species because of extensive overlap between the morphological values. This preliminary study confirms that identification of snow leopard feces in the field is misleading. Research that relies upon scat samples to assess distribution or diet of the snow leopard should therefore employ molecular scatology techniques. These methods are financially accessible and employ relatively simple laboratory procedures that can give an indisputable response to species identification from scats.

Keywords: DNA analysis; Snow leopard; Scats; Noninvasive genetics; Carnivore

INTRODUCTION

Information gathered from scats, such as diet, distribution, abundance and community dynamics, is widely used in many ecological studies and conservation management plans and requires reliable identification of scats (Gibbs, 2000; Long et al, 2008). However the identification of species from field signs alone is not always accurate and misidentification of scats is frequent, in particular in habitats with sympatric, similarly sized carnivore species (Farrell et al, 2000). Traditional methods of identifying scats in the field include size (length and/or diameter), shape (segmented, pointed ends, etc.), color (black, brown, white, etc.), pH or smell, in addition to the co-occurrence of other signs at the site such as pugmarks, scratches or hair (Danner & Dodd, 1982; Green & Flinders, 1981; Jackson & Hunter, 1996). However, these field methods have proven unreliable for the following reasons: body size can vary greatly within species which affects scat dimensions (e.g. between juvenile and adult, and male and female), and other species may investigate the scat sites of the target species, confusing identification by leaving their signs.

To address the scat identification dilemma, molecular scatology techniques have been developed to extract DNA from the donor cells present on the outer lining of the scat samples (Foran et al, 1997; Höss et al, 1992). Target DNA, that enables species identification, is then amplified using polymerase chain reaction (PCR) based methods. These techniques have been employed for scat recognition in numerous studies and for a variety of species, for example from identification of European brown bears (Ursus arctos) (Kohn et al, 1995), to sympatric confamilial groups (Mills et al, 2001), and discerning species in an entire community of carnivores (Fernandes et al, 2008) or even telling apart hybrids (Adams et al, 2003). The use of molecular scatology has also revealed and confirmed that species identification of scats in the field is often inaccurate (Davison et al, 2002; Farrell et al, 2000; Harrington et al, 2009; Prugh & Ritland, 2005; Zuercher et al, 2003).

Received: 19 May 2014; Accepted: 20 November 2014

Foundation items: This study was supported by the Second National Survey of Terrestrial Wildlife in China, State Forestry Administration of China; the Darwin Initiative; and the Robertson Foundation.

*Corresponding author, E-mail: kunshi@bjfu.edu.cn
Snow leopards are very elusive, extremely rare and inhabit harsh environments (Jackson & Ahlborn, 1988; Sunquist & Sunquist, 2002) making research through direct observation almost impossible. Research on this species therefore relies on gathering information in the absence of direct contact, primarily through non-invasive techniques (Ma et al., 2011; Wei et al., 2001; Xu et al., 2005). One of the main survey techniques used is scat collection with the following objectives: to determine prey preference (Bagchi & Mishra, 2006; Liu et al., 2003; Oli, 1993; Shehzad et al., 2012; Wang et al., 2014), to estimate population size (Ale et al., 2007; Fox et al., 1991; Hussain, 2003; Karmacharya et al., 2011; McCarthy et al., 2008), to analyse community structure (Lovari et al., 2013), etc. In most of their ranges, snow leopards co-occur with other similarly sized carnivores, such as wolves, foxes and jackals (Canis aureus), potentially creating uncertainty in the field identification of scat. It is not surprising that evidence of extensive scat misidentification confirmed by molecular techniques has been reported for snow leopards in previous publications. Janečka et al. (2008) first observed a high level of scat misidentification, ranging from 35% in Ladakh, India to 54% in South Gobi, Mongolia. Of 71 samples collected in the study by Karmacharya et al. (2011) in Nepal, 42% were misidentified as snow leopard in the field, but in fact belonged to other carnivores. Anwar et al. (2011) in Pakistan found 52% of scats to be correctly identified as snow leopard while the rest were from other sympatric species. Likewise Shehzad et al. (2012) in Mongolia correctly identified in the field 43% of putative snow leopard scats as confirmed by DNA analysis, while 57% were excluded from the study because they actually belonged to other species.

Despite misidentification being a source of significant bias, scat sampling still remains a fundamental tool in the study and conservation of snow leopards, so it is important to address and reduce the described inaccuracies.

This study aims to: (1) assess the extent of misidentification in the field for snow leopard scat; (2) determine if variables associated with the scat, such as morphological characteristics of the feces, are correlated with species identification; and (3) understand if accurate field identification of snow leopard scats is possible, if so establish more specific field collection protocols that include morphological features statistically relevant for successful species identification of scats.

MATERIALS AND METHODS

Study area and sample collection
Scats were collected from February 22, to March 12, 2009, and March 12 to April 12 in 2011 in Taxkorgan Nature Reserve (TNR; E74°30'-77°00', N36°38'-37°30'), located in the east plateau of the Pamir Mountains, Xinjiang Uygur Autonomous Region, China. TNR has a mean elevation of approximately 4000 m and it is characterized by a cold desert climate, with long very cold winters. The average temperature during the survey months is 1.03 °C, with 3.27 mm average precipitation per month. The carnivore guild includes snow leopards, wolf (Canis lupus), red fox (Vulpes vulpes) and lynx (Lynx lynx). Pallas’ cat (Otocolobus manul) may also be present, but extended surveys using camera traps have never recorded in our study sites.

Data for this study were collected in two sites, Mariang and Mazar; both sites had minimum altitudes above 3 000 m. The study sites straddled the southeastern and eastern boundaries of the TNR respectively, with some transects extending as far as 40 km outside the reserve (Figure 1). In

Figure 1 Location of Taxkorgan Nature Reserve including the two study sites (Mariang and Mazar)
Each sample at time of collection was placed individually in a labeled zip lock plastic bag to avoid contamination. To prevent degradation the samples were air-dried in the field by placing the open bags in a ventilated, cool and dry environment away from direct sunlight. In the laboratory, samples were then transferred to a freezer at −4 °C and finally to −80 °C for long-term storage. The morphological scat data recorded included number of segments (S), mean segment length (ML, cm), total length of scat (TL, cm), mean diameter (MD, cm) and weight after drying (DW, g). Original weight at collection was also noted but not used in these analyses, as it is greatly dependent on the age of the scat (fresh scats are heavier due to a higher water content, while old scats are lighter) and location of the scat (exposure to ice and snow may increase water content of old scats). Weight after drying was preferred as it does not present this constraint and was considered more appropriate.

Scat samples (number of samples in brackets next to the species Latin name in bold) were assigned to reference species based on nucleotide diversity <0.03 and node bootstrap value >90%.

**DNA extraction and species identification**

Laboratory analysis was performed at the Key Laboratory of Animal Ecology and Conservation Biology (Institute of Zoology, Chinese Academy of Sciences, Beijing). The DNA extraction was performed using the QIAamp DNA stool minikit (Qiaigen). A 146 bp region of the mitochondrial cytochrome b gene was amplified by PCR using carnivore specific primers from Farrell et al (2000) (5'-AACTGCAAGCCCTCAGAATG ATATTGTCCCTCA-3' and 5'-TATTCTTTATCGCTATACAT RC ACG-3'). It is not possible to distinguish between wolf and domestic dog using this primer due to their close phylogenetic relationship, so they will be referred to as "wolf/dog" in the rest of this paper. Amplifications were conducted following the protocol developed by Janecka et al (2008), although changes to the volume of reagents (total 50 µL instead of 10 µL) and thermo cycling conditions were made as follows: 5X PrimeSTAR buffer 10 µL, 4 µL of dNTP mixture (2.5 mmol/L), forward primer 1 µL (10 µmol/L), reverse primer 1 µL (10 µmol/L), PrimeSTAR HS DNA polymerase 0.5 µL (2.5 U/µL), DNA extract 4 µL and DNA-free water 29.5 µL. The PCR conditions included an initial denaturing step of 94 °C for 1 min, followed by 40 cycles of 94 °C for 30 sec, 54 °C for 30 sec, 72 °C for 30 sec, and a final extension step of 72 °C for 2 min. PCR products (3 µL amplified DNA plus 1.5 µL of loading dye) were fractionated on a 2% agarose gel, stained with ethidium bromide, and visualized under ultraviolet light. Primers were provided by Invitrogen (Beijing, China) and all reagents by TaKaRa Biotechnology Co. (Dalian, China). PCR products were sent for sequencing to SinoGenoMax Company Ltd. (Beijing, China). Sequences obtained from one strand were examined in 4PEAKS version 1.7.2 (©2006 Mek&Tosj.com) and submitted to a BLAST search (Madden, 2002) to be compared with entries in GenBank sequences in order to identify the species for each sample (Zhang et al, 2000). Sequence matches with an E-value equal to 0.0 and maximum identity value ≥90%, were considered positive species identifications (DeMatteo et al, 2014; Keehner, 2009; Naidu et al, 2011; Rozhnov et al, 2011). In addition, a phylogenetic tree was constructed with reference species (downloaded from GenBank with accession numbers: KJ637144, JCO003577, AB303951, AB194817, EF689046, AY928671, KF990330, EF551002, AF053050, JF357970, JF357968, EF551004, KF661088) to double-check species identification, using the neighbor-joining algorithm based on the Kimura 2-parameter model (Kimura, 1980) with MEGA v6.06 (Tamura et al. 2013). Node support was evaluated using 1 000 bootstrap replicates.

**Variables associated with species identification**

The genetic species identification results were used to calculate field identification error for snow leopard feces (i.e. the percentage of scats that were misidentified as snow leopard in the field). A confusion matrix (or contingency table) was built to summarize the results of the misidentification for all collected samples belonging to the carnivore guild.

A one-way ANOVA test was used to look for correlation between the presumed snow leopard samples (categorical independent variables) and the morphological variables

### Table 1: Scat morphological parameters used in the field as guidance to identify the species.

<table>
<thead>
<tr>
<th>Species</th>
<th>Color</th>
<th>Shape</th>
<th>Number of segments (n)</th>
<th>Length (cm)</th>
<th>Diameter (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snow Leopard</td>
<td>Dark</td>
<td>Long pointed tail</td>
<td>≤3</td>
<td>5-6</td>
<td>&gt;2</td>
</tr>
<tr>
<td>Lynx</td>
<td>Dark</td>
<td>Long pointed tail</td>
<td>≤3</td>
<td>4-5</td>
<td>~2</td>
</tr>
<tr>
<td>Wolf/dog</td>
<td>White or grey</td>
<td>Round end, no tail</td>
<td>≤3</td>
<td>5</td>
<td>&gt;2</td>
</tr>
<tr>
<td>Red Fox</td>
<td>Dark, grey old samples</td>
<td>Round end, short pointed tail</td>
<td>2-5</td>
<td>2-4</td>
<td>&lt;2</td>
</tr>
</tbody>
</table>


(continuous dependent variables) associated with the scat as previously done by Anwar et al. (2011). For those variables that were found to be significant, a post-hoc test of least significant difference (LSD) to explore all possible pair-wise comparisons of means comprising a factor using the equivalent of multiple t-tests.

To analyze the relationship between the variables associated with the presumed snow leopard scat and the actual species, principal component analysis (PCA) was used. This method is a true eigenvector-based multivariate analysis, which reduces the effective dimensionality of a multivariate data set by producing linear combinations of the original variables that summarize the predominant patterns in the data (Peres-Neto et al. 2003).

All analyses were performed using software R 2.11.1 (R Development Core Team, 2013). The total number of samples included in the analysis varied as some samples had missing values (i.e. in 2009 total length, mean segment length and number of segments were not recorded) and could not be used in certain statistical tests.

RESULTS

Identification of species from scat

During the survey 100 scats were collected and were attributed to a species in the field: 51 presumed snow leopard scats, 40 presumed fox scats and 9 could not be identified (Figure 2). No wolf/dog samples or other carnivores were categorized in the field. DNA was successfully extracted, amplified and sequenced for species identification from 88% of scats and all species were categorized with maximum identity value ≥ 90% (Appendix 1, available online). Nucleotide diversity (number of nucleotide base substitutions per site averaged over all sequence pairs within each species) was 0.005 for snow leopard, 0.009 for fox and 0.024 for wolf/dog. Samples were assigned to reference species based on pairwise distance <0.03 and node bootstrap value >90% (Figure 3; Appendix 2, available online). Unidentified scats (12 samples) could either be from species for which the primers were not appropriate

Figure 2 Species identification of scat samples (“Field species ID”) in the field compared to results of genetic identification (“Genetic species ID”)

Figure 3 Condensed neighbor-joining tree with cut-off value of 90% built using Kimura 2-parameter model

(90-carnivore species) or represent samples with low DNA quality not suitable for genetic analysis. The discrepancy between species identification in the field and identification by genetic analysis is shown in Figure 2.

Among the samples identified in the field as snow leopard scats (51 samples), the majority was genetically identified as actually belonging to fox (28 samples, 55%), 4 were wolf/dog (8%), while 16 were snow leopard (31%). The remaining 3 were unidentified samples (6%). Therefore snow leopard scat identification error, calculated based on genetically verified scats for snow leopard scats in our study is 67%.

Genetic analysis showed that scats categorized as fox in the field were correctly identified in all cases, except when DNA analysis was unsuccessful (Table 2, total 40 samples, 32 confirmed fox and 8 were unidentified). A confusion matrix was built (Table 3), showing that while Type 1 errors were made in 32 occasions (scats identified as snow leopard in the field were not snow leopard scats), Type 2 errors didn’t occur (snow leopard scats were not present amongst scats identified in the field as other species).

Table 2 Morphological characteristics (diameter and dry weight) for snow leopard, fox and wolf/dog.

<table>
<thead>
<tr>
<th>Species</th>
<th>Total samples</th>
<th>Diameter (cm)</th>
<th>Dry weight (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snow leopard</td>
<td>16</td>
<td>2.5±0.5</td>
<td>16.2±15.1</td>
</tr>
<tr>
<td>Fox</td>
<td>67</td>
<td>1.8±0.4</td>
<td>6.8±7.0</td>
</tr>
<tr>
<td>Wolf/dog</td>
<td>5</td>
<td>2.4±0.6</td>
<td>14.8±8.3</td>
</tr>
</tbody>
</table>

Table 3 Confusion matrix summarizing the results of 2009 and 2011 field seasons

<table>
<thead>
<tr>
<th>Predicted class</th>
<th>Snow leopard</th>
<th>Fox</th>
<th>Wolf/dog</th>
<th>Unidentified</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snow leopard</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Fox</td>
<td>28</td>
<td>32</td>
<td>0</td>
<td>7</td>
<td>67</td>
</tr>
<tr>
<td>Wolf/dog</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Unidentified</td>
<td>3</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>51</td>
<td>40</td>
<td>0</td>
<td>9</td>
<td>100</td>
</tr>
</tbody>
</table>

Type 1 errors were made (n=32) but no Type 2 errors were made when trying to identify snow leopard scats in the field.
The misidentification rate of snow leopard scats indicates that field identification is problematic, so erroneous conclusions on the overlapping diet of these two species and incorrect prey preference were likely to be made if our study had proceeded without first employing molecular analysis. It was not possible, however, to define precise ranges for each species because of extensive overlap between the scat morphological values. Other variables related to the surrounding environment of the scat, such as presence of other signs, location and substrate type, may also play a role in correct species identification, and should be recorded and verified in future studies (Ma et al, 2005).
Davison et al (2002) have emphasized that scat morphological identification methods need to be more rigorous when used in surveys and suggest using multi-evidence approaches involving a variety of methods to correctly identify species presence. This is also the case for snow leopard surveys as scat misidentification can lead to confusion about the species’ presence and risk of overestimating populations. It is important that research not rely solely on the identification of scats in the field, but scat identification be improved by including other techniques such as molecular analysis, scat detection dogs and camera trapping (Janecka et al, 2011; Long et al, 2007a; Long et al, 2007b). An extensive sampling survey required to collect sufficient snow leopard scat samples can incur high financial and time costs, so an additional cost of 70¥ (about US$12) per sample for molecular analysis is a reasonable price to make these surveys more accurate. These costs will likely decrease in the future and more cost-effective analyses can be used such as real-time PCR or PCR-RFLP (Cossios & Angers, 2006; Harrington et al, 2009; Mukherjee et al, 2010; Rodgers & Janecka, 2012).

This preliminary study confirms that there is a high rate of misidentification of snow leopard scats in the field and that morphological characteristics of scats can’t be used to reliably differentiate between sympatric carnivore species. Therefore any research project that requires species identification (diet studies, sign surveys etc.) is advised to employ noninvasive DNA testing of scats in order to avoid serious bias in results. Genetic methods involve rather straightforward laboratory procedures, are relatively inexpensive and provide indisputable species identification of scats.

It is also important to note that only a small portion of samples collected actually belonged to the target species, and since small sample size is already considered to be a problem when studying elusive animals such as the snow leopard, it is important to remind researchers that they must take this into further account and plan accordingly by increasing the study area, extending duration of surveys or using multiple non-invasive sampling methods.

Finally, molecular species identification is not only a verification tool to be used in the laboratory. The results, in combination with other information on scat morphology, can also provide valuable feedback to field workers to improve collection guidelines and ultimately create effective conservation action plans.

ACKNOWLEDGEMENTS

This project was supported by State Forestry Administration of China (SFA), Snow Leopard Network, the Snow Leopard Trust, Panthera, Beijing Forestry University and the University of Oxford. Many thanks go to Dr Ming Li from the Institute of Zoology, Chinese Academy of Science, for access to a high quality laboratory and Dr Hui-Juan PAN from Beijing Forestry University for technical advice. We appreciate the great support from the Forestry Department of Xinjiang Uygur Autonomous Region, Wildlife Conservation Division of Kashi Region, Taxkorgan Nature Reserve Administration, Livestock Husbandry Bureau of Taxkorgan County, and government of Mariang village. We greatly thank our local guide Kemeli Taklashur for his tremendous contribution in the field.

REFERENCES

Adams JR, Kelly BT, Wait LP. 2003. Using faecal DNA sampling and GIS to monitor hybridization between red wolves (Canis rufus) and coyotes (Canis latrans). Molecular Ecology, 12(8): 2175-2186.


