MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 1951–1965 doi: 10.1111/j.1365-294X.2011.05424.x

Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan

WASIM SHEHZAD,* TIAYYBA RIAZ,* MUHAMMAD A. NAWAZ,*† CHRISTIAN MIQUEL,* CAROLE POILLOT,* SAFDAR A. SHAH,‡ FRANÇOIS POMPANON,* ERIC COISSAC* and PIERRE TABERLET*

**Laboratoire d'Ecologie Alpine, CNRS-UMR 5553, Université Joseph Fourier, BP 53, F-38041 Grenoble Cedex 9, France, †Snow Leopard Trust (Pakistan Program), 17-Service Road North, I-8/3, Islamabad, Pakistan, ‡Wildlife Department, Khyber Pakhtunkhwa, Pakistan*

Abstract

Diet analysis is a prerequisite to fully understand the biology of a species and the functioning of ecosystems. For carnivores, traditional diet analyses mostly rely upon the morphological identification of undigested remains in the faeces. Here, we developed a methodology for carnivore diet analyses based on the next-generation sequencing. We applied this approach to the analysis of the vertebrate component of leopard cat diet in two ecologically distinct regions in northern Pakistan. Despite being a relatively common species with a wide distribution in Asia, little is known about this elusive predator. We analysed a total of 38 leopard cat faeces. After a classical DNA extraction, the DNA extracts were amplified using primers for vertebrates targeting about 100 bp of the mitochondrial 12S rRNA gene, with and without a blocking oligonucleotide specific to the predator sequence. The amplification products were then sequenced on a next-generation sequencer. We identified a total of 18 prey taxa, including eight mammals, eight birds, one amphibian and one fish. In general, our results confirmed that the leopard cat has a very eclectic diet and feeds mainly on rodents and particularly on the Muridae family. The DNA-based approach we propose here represents a valuable complement to current conventional methods. It can be applied to other carnivore species with only a slight adjustment relating to the design of the blocking oligonucleotide. It is robust and simple to implement and allows the possibility of very large-scale analyses.

Keywords: blocking oligonucleotide, DNA metabarcoding, mitochondrial DNA, ribosomal DNA, species identification

Received 8 June 2011; revision received 16 November 2011; accepted 24 November 2011

Introduction

The nature of trophic interactions is a fundamental question in ecology and has commanded the attention of biologists for decades. Dietary behavioural studies provide key data for understanding animal ecology, evolution and conservation (Symondson 2002; Krahn *et al.* 2007). Wild felids are among the keystone predators and have significant effects on ecosystem functioning, despite their relatively low biomass (Mills *et al.* 1993; Power *et al.* 1996). The modal mass concept (Macdonald *et al.* 2010) proposes that each felid species focuses on large-as-possible prey to maximize their intake relative to their energy expenditure for each catch, provided that such prey can be safely killed.

Owing to their elusive behaviour, scientific knowledge of South Asian wild cats is limited (Nowell & Jackson 1996). The leopard cat (*Prionailurus bengalensis*) is a small felid (weight 1.7–7.1 kg; Sunquist & Sunquist 2009), with a wide range in Asia (8.66×10^6 km^2; Nowell & Jackson 1996). Beginning in Pakistan and

Correspondence: Pierre Taberlet, Fax: +33(0)4 76 51 42 79; E-mail: pierre.taberlet@ujf-grenoble.fr

parts of Afghanistan in the west, the leopard cat occurs throughout Southeast Asia, including the islands of Sumatra, Borneo, and Taiwan. It extends into China, Korea, Japan and the Far East of Russia. (Macdonald *et al.* 2010). The leopard cat's flexible habitat selection and prey choices favour its distribution throughout the range (Watanabe 2009; Mukherjee *et al.* 2010). It is found in very diverse environments, from semideserts to tropical forests, woodlands to pine forests and scrubland to agriculture land (Sunquist & Sunquist 2002). It prefers to live in habitats near sources of water and can be found in the close proximity to human population (Scott *et al.* 2004).

The population status of the leopard cat is not uniform throughout its range. The cat is relatively secure in China (Lau *et al.* 2010) and in India (Nowell & Jackson 1996), endangered in Korea (Rho 2009) and most endangered in Japan (Mitani *et al.* 2009). In Pakistan, this species is categorized by the IUCN as ''data deficient'' as no information exists about the extent of its occurrence, nor its occupancy, population and habitat (Sheikh & Molur 2004). Major threats to the species include hunting, habitat loss and fragmentation because of the human population expansion in addition to competition for prey with other sympatric carnivores (Izawa & Doi 1991). Commercial exploitation for the fur trade is a significant threat throughout its range (Sheikh & Molur 2004); in China, the annual pelt harvest was estimated at to be 400 000 animals in mid-1980s (Nowell & Jackson 1996).

Despite being a relatively common species with a wide distribution, comparatively little information is available about the diet of the leopard cat in general, and no information at all specific to Pakistan, where this predator is rare. Faeces analysis by hair mounting and bone examination is used extensively and can provide information about the diet (e.g. Oli *et al.* 1994; Gaines 2001; Bagchi & Mishra 2006; Lovari *et al.* 2009). Muridae (mainly *Rattus* spp. and *Mus* spp.) seem to represent the main prey items throughout the leopard cat distribution range, supplemented by a wide variety of other prey including small mammals such as shrews and ground squirrels, birds, reptiles, frogs and fish (Tatara & Doi 1994; Grassman *et al.* 2005; Austin *et al.* 2007; Rajaratnam *et al.* 2007; Watanabe 2009; Fernandez & de Guia 2011).

Molecular analysis of faeces (Höss *et al.* 1992; Kohn & Wayne 1997) provides an alternative noninvasive approach to study animal diet, but prey DNA in faeces is often highly degraded, preventing the amplification of long fragments (Zaidi *et al.* 1999; Jarman *et al.* 2002). Until 2009, most of the molecular-based studies to analyse diet were carried out using traditional sequencing approaches (e.g. Deagle *et al.* 2005a, 2007; Bradley *et al.* 2007). These methods require cloning PCR products and subsequent Sanger sequencing of these clones by capillary electrophoresis. However, this approach is both time-consuming and expensive (Pegard *et al.* 2009).

Next-generation sequencing is revolutionizing diet analysis based on faeces (Valentini *et al.* 2009b), because sequence data from very large numbers of individual DNA molecules in a complex mixture can be studied without the need for cloning. Valentini *et al.* (2009a) have presented a universal approach for the diet analysis of herbivores. The methodology consists of extracting DNA from faeces to amplify it using the universal primers *g* and *h*, which amplify the short P6 loop of the chloroplast *trn*L (UAA) intron (Taberlet *et al.* 2007), and in sequencing the PCR products using a next-generation sequencer.

While such an approach has been successfully implemented for herbivores, the analysis of carnivore diet presents a real challenge when using primers for mammals or vertebrates, as predator DNA can be simultaneously amplified with prey DNA (Deagle *et al.* 2005b; Jarman *et al.* 2006). Furthermore, prey fragments might be rare in the DNA extract from faeces, and consequently be prone to being missed during the early stages of PCR, resulting in a PCR product almost exclusively containing the dominant sequences of predators (Jarman *et al.* 2004, 2006; Green & Minz 2005). Various methods have been proposed to avoid amplifying predator DNA. Species-specific or group-specific primers have been specially designed to avoid priming on predator DNA and to specifically amplify the target prey species (Vestheim *et al.* 2005; Deagle *et al.* 2006; King *et al.* 2010). This is not a convenient strategy if the prey are taxonomically diverse, which makes the design of suitable primers difficult (Vestheim & Jarman 2008). Another strategy involves cutting predator sequences with restriction enzymes before and/or during and/or after PCR amplification (Blankenship & Yayanos 2005; Green & Minz 2005; Dunshea 2009). However, these approaches can only be implemented with *a priori* knowledge of the potential prey.

The ideal system for studying carnivore diet using DNA in faeces lies in combining, in the same PCR, primers for vertebrates and a blocking oligonucleotide with a 3-carbon spacer (C3-spacer) on the 3' end that specifically reduces the amplification of the predator DNA. Such a blocking oligonucleotide must be specifically designed to target predator DNA and thus bind preferentially with predator sequences, limiting their amplification. This concept has been effectively used in the field of clinical chemistry (Kageyama *et al.* 2008; Wang *et al.* 2008; Li *et al.* 2009) and in environmental microbiology (Liles *et al.* 2003). However, the application of blocking oligonucleotide in trophic studies is relatively recent. Vestheim & Jarman (2008) first used a

blocking oligonucleotide to assess the diet of Antarctic krill. More recently, Deagle *et al*. (2009, 2010) investigated the diet of Australian fur seals (*Arctocephalus pusillus*) and penguins (*Eudyptula minor*) by combining a blocking oligonucleotide approach with 454 GS-FLX pyrosequencing technologies.

The main aim of this study was to analyse the leopard cat diet in two distinct environments in Pakistan by developing a method that would give the vertebrate diet profile of a carnivore without any *a priori* information about the prey species. This method is based on the use of recently designed primers for vertebrates (Riaz *et al*. 2011) together with a blocking oligonucleotide specific to the leopard cat and employing a high-throughput next-generation sequencer. However, such an approach cannot detect the cases of infanticide and possible cannibalism that have been documented in Felidae (e.g. Natoli 1990).

Materials and methods

General strategy for diet analysis of the leopard cat

Figure 1 outlines the general strategy we followed for the diet analysis of the leopard cat. After the faeces collection and DNA extraction, the samples were confirmed to be those of leopard cat by using leopard cat–specific primers. Selected samples were amplified in two series of experiments, one with primers for vertebrates and the other with the same primers plus a blocking oligonucleotide specific to the leopard cat. These PCR products were subsequently sequenced using the Illumina sequencing platform GA IIx. The amplified sequences of prey taxa were identified by comparison with reference databases (GenBank/EMBL/DDBJ), taking into account prey availability according to their geographic distributions.

Sample collection and preservation

Putative felid faeces were collected in two areas: Ayubia National Park (ANP) and Chitral Gol National Park (CGNP). Both national parks are located in the Khyber Pakhtunkhwa province and represent two extremities of the leopard cat range in Pakistan (Fig. 2). These national parks have disparate environments. The ANP is comprised of moist temperate forests, subalpine meadows and subtropical pine forests. Mean temperatures range between 4.2 °C in January to 26 °C in July. The altitudinal variation ranges from 1050 to 3027 m, and the mean annual rainfall is between 1065 and 1424 mm. It has ~200 species of birds, 31 species of

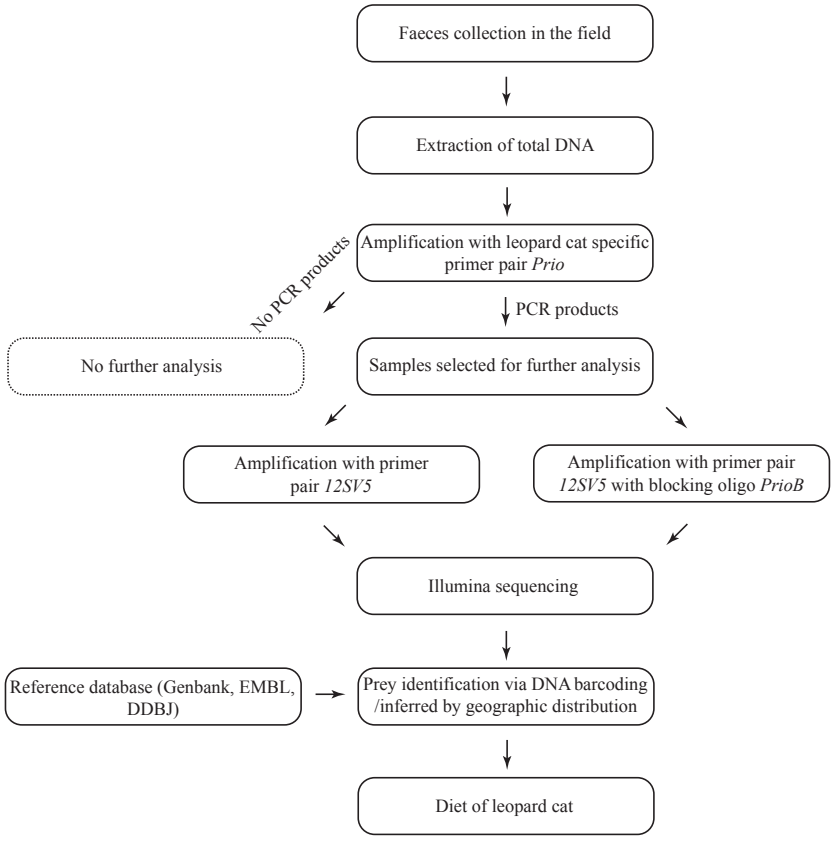

Fig. 1 Flowchart diagram showing the various steps involved in the diet analysis of the leopard cat. The samples in the dotted box were discarded from further experimentation.

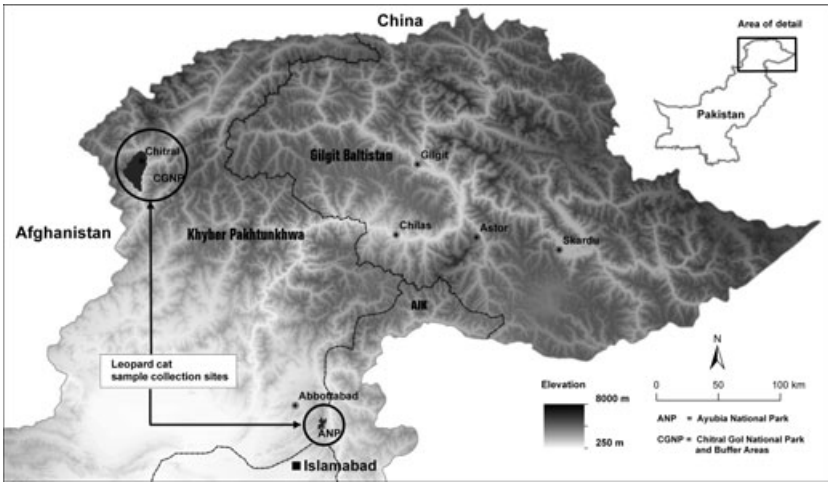

Fig. 2 Sampling locations of leopard cat faeces in northern Pakistan.

mammals, 16 species of reptiles and three species of amphibians (Farooque 2007).

The CGNP generally falls into a subtropical zone with vegetation classified as dry temperate forests. Forests of the park are growing under the extremes of climatic and edaphic factors, and tree canopy is rarely closed. Mean temperature of the valley ranges between 1 °C in January to 24 °C in July, and average annual rainfall varies between 450 and 600 mm. The park supports 29 mammals, 127 birds and nine reptiles (GoN-WFP & IUCN 1996; Mirza 2003).

We collected 114 faecal samples from ANP and 67 from the CGNP. The samples were preserved first in 90% ethanol and then shifted into silica gel for transportation to LECA (Laboratoire d'Ecologie Alpine), Université Joseph Fourier, Grenoble, France, for diet analysis.

DNA extraction

All extractions were performed in a room dedicated to degrade DNA extractions. Total DNA was extracted from about 15 mg of faeces using the DNeasy Blood and Tissue Kit (QIAgen GmbH). Each 15 mg faecal sample was incubated for at least 3 h at 55 °C with a lysis buffer (Tris–HCl 0.1 M, EDTA 0.1 M, NaCl 0.01 M and *N*-lauroyl sarcosine 1% with pH 7.5–8), before following the manufacturer's instructions. The DNA extracts were recovered in a total volume of 250 μL. Blank extractions without samples were systematically performed to monitor possible contaminations.

Selection/designing of primer pairs for the leopard cat diet study

Identification of faecal samples as leopard cat. We used the leopard cat–specific primer pair *PrioF/PrioR*, amplifying a 54-bp fragment (without primers) of the mitochondrial 12S gene (Table 1). The specificity of this primer pair was validated both by empirical experiments (Ficetola *et al.* 2010) and by the program *ecoPCR* (Bellemain *et al.* 2010; Ficetola *et al.* 2010), with parameters to prevent mismatches on the two last nucleotides of each primer, and designed to tolerate a maximum of three mismatches on the remaining part of the primers. The goal of such an experimental validation was to distinguish leopard cat faeces from those from the two other felid species potentially occurring in the study areas, i.e. the common leopard (*Panthera pardus*) in ANP and the snow leopard (*Panthera uncia*) in CGNP. The primary identification of samples was carried out on the basis of the presence of a PCR product of the suit-

Table 1 Sequences of the primer pairs used in the study. The length of amplified fragments (excluding primers) with *Prio* & *12SV5* was 54 and ~100 bp, respectively

Name	Primer sequence (5–3′)	References
PrioF	CCTAAACTTAGATAGTTAATTTT	Ficetola *et al.* (2010)
PrioR	GGATGTAAAGCACCGCC	Ficetola *et al.* (2010)
12SV5F	TAGAACAGGCTCCTCTAG	Riaz *et al.* (2011)
12SV5R	TTAGATACCCCACTATGC	Riaz *et al.* (2011)
PrioB	CTATGCTTAGCCCTAAACTTAGATAGTTAATTTTAACAAAACTATC-C3	This study

able length as revealed by electrophoresis on a 2% agarose gel. The samples successfully amplified using *PrioF/PrioR* were selected for further analyses.

The PCRs were carried out in a total volume of 20 μL with 8 mM Tris–HCl (pH 8.3), 40 mM KCl, 2 mM $MgCl_2$, 0.2 mM of each dNTP, 0.2 μM of each primer, BSA (5 μg), 0.5 U of AmpliTaq Gold® DNA polymerase (Applied Biosystems) using 2 μL of DNA extract as a template. The PCR conditions were set as an initial 10-min denaturation step at 95 °C to activate the polymerase, followed by 45 cycles of denaturation at 95 °C for 30 s and annealing at 53 °C for 30 s, without elongation steps as the amplified fragment was very short.

Blocking oligonucleotide specific to leopard cat sequences. The *PrioB* (Table 1) blocking oligonucleotide specific to leopard cat sequences was designed as suggested by Vestheim & Jarman (2008). This blocking oligonucleotide was used to limit the amplification of leopard cat sequences when using the primers targeting all vertebrates. Table 2 presents a sequence alignment of *PrioB* with the main groups of vertebrates. This blocking oligonucleotide might also slightly block the amplification of other felid species, but will not prevent the amplification of other vertebrate groups.

Primer pair for vertebrates. We used the primer pair *12SV5F/12SV5R* designed by the *ecoPrimers* program (Riaz *et al.* 2011). *ecoPrimers* scans whole genomes to find new barcode markers and their associated primers, by optimizing two quality indices measuring the taxonomical coverage and the discrimination power to select the most efficient markers, according to specific experimental constraints such as marker length or targeted taxa. This primer pair for vertebrates represents the best choice found by *ecoPrimers* among short barcodes, as derived from the available vertebrate whole mitochondrial genomes currently available. It amplifies a ~100-bp fragment of the V5 loop of the mitochondrial 12S gene, with the ability to amplify short DNA fragments such as those recovered from faeces, and has a high taxonomic resolution, despite its short size. Using the *ecoPCR* program (Bellemain *et al.* 2010; Ficetola *et al.* 2010), and based on the release 103 of the EMBL database, this fragment unambiguously identifies 77% of the species and 89% of the genera as recorded by this EMBL release (Riaz *et al.* 2011).

DNA amplification for diet analysis

All DNA amplifications were carried out in a final volume of 25 μL, using 2 μL of DNA extract as template. The amplification mixture contained 1 U of AmpliTaq Gold® DNA Polymerase (Applied Biosystems), 10 mM Tris–HCl, 50 mM KCl, 2 mM of $MgCl_2$, 0.2 mM of each dNTP, 0.1 μM of each primer (*12SV5F/12SV5R*) and 5 μg of bovine serum albumin (BSA; Roche Diagnostic). The PCR mixture was denatured at 95 °C for 10 min, followed by 45 cycles of 30 s at 95 °C and 30 s at 60 °C; as the target sequences are ~100 bp long, the elongation step was removed to reduce the +A artefact (Brownstein *et al.* 1996; Magnuson *et al.* 1996) that might decrease the efficiency of the first step of the sequencing process (blunt-end ligation). Using the aforementioned conditions, the DNA extracts were amplified twice, first with *12SV5F/12SV5R* (0.1 μM each) and second with *12SV5F/12SV5R/PrioB* (0.1 μM for *12SV5F* and *12SV5R*, 2 μM for *PrioB*). These primer concentrations have been chosen after a series of test experiments, with various concentrations of *PrioB* (data not shown).

Table 2 Sequence alignment showing the specificity of the *PrioB* blocking oligonucleotide. The first six nucleotides of the *PrioB* blocking oligonucleotide overlap with the *12SV5R* amplification primer. This sequence alignment contains two other Felidae species (*Felis catus* and *Panthera tigris*), another carnivore species from the Ursidae family (*Ursus arctos*), two rodents (*Rattus rattus* and *Microtus kikuchii*), one insectivore (*Crocidura russula*), one bird (*Gallus gallus*), one amphibian (*Rana nigromaculata*) and one fish (*Cyprinus carpio*)

Accession number	Species name	Sequences (5′–3′)
PrioB blocking oligonucleotide		CTATGCTTAGCCCTAAACTTAGATAGTTAATTTTAACAAAACTATC
HM185183	*Prionailurus bengalensis*	.....
NC_001700	*F. catus*	.....CCC.A.....
JF357967	*P. tigris*	.....C.....CCCA.....
NC_003427	*U. arctos*	.....T.....A..A...A..T...AA.CA...TTAT..
NC_012374	*R. rattus*	.....C.TA...A...CA.C..CA....TAT.T
NC_003041	*M. kikuchii*	.....C.TAG..A..AAAAC.A...TA.T.G..
NC_006893	*C. russula*	.....A.A.C.A.C..A.AAC.AG.CTG.TCG
NC_007236	*G. gallus*	......C.....TC.....CC.CCCA.C.CAC.TGTATC.
NC_002805	*R. nigromaculata*	T.....C.....GT....AATC.ACTCAC.CCAACCA.CGC.AGGG
NC_001606	*C. carpio*	.....C....G.....C...C.TCC.GC.AC..TT.G.TGTC.

The primers for vertebrates, *12SV5F* and *12SV5R*, were modified by the addition of specific tags on the 5′ end to allow the assignment of sequence reads for the relevant sample (Valentini *et al.* 2009a). All of the PCR products were tagged identically on both ends. These tags were composed of CC on the 5′ end followed by seven variable nucleotides that were specific to each sample. The seven variable nucleotides were designed using the *oligoTag* program (http://www.prabi.grenoble.fr/trac/OBITools) to have at least three differences among the tags, to contain no homopolymers longer than two and to avoid a C on the 5′ end so as to allow the detection of a possible deletion within the tag. All of the PCR products from the different samples were first purified using the MinElute PCR purification kit (QIAGEN GmbH), titrated using capillary electrophoresis (QIAxel; QIAgen GmbH) and finally mixed together in equimolar concentration before sequencing.

DNA sequencing

The sequencing was carried out on the Illumina Genome Analyzer IIx (Illumina Inc.), using the Paired-End Cluster Generation Kit V4 and the Sequencing Kit V4 (Illumina Inc.), following the manufacturer's instructions. A total of 108 nucleotides were sequenced on each extremity of the DNA fragments.

Sequence analysis and taxon assignation

The sequence reads were analysed separately with and without the blocking oligonucleotide, using the OBITools (http://www.prabi.grenoble.fr/trac/OBITools). First, the direct and reverse reads corresponding to a single molecule were aligned and merged using the *solexaPairEnd* program, taking into account data quality during the alignment and the consensus computation. Primers and tags were then identified using the *ngsfilter* program. Only sequences with a perfect match on tags and a maximum of two errors on primers were recorded for the subsequent analysis. The amplified regions, excluding primers and tags, were kept for further analysis. Strictly, identical sequences were clustered together using the *obiuniq* program, keeping the information about their distribution among samples. Sequences shorter than 60 bp, or containing ambiguous nucleotides, or with occurrence lower or equal to 100 were excluded using the *obigrep* program. Taxon assignation was achieved using the *ecoTag* program (Pegard *et al.* 2009). *EcoTag* relies on a dynamic programming global alignment algorithm (Needleman & Wunsch 1970) to find highly similar sequences in the reference database. This database was built by extracting the relevant part of the mitochondrial 12S gene from EMBL nucleotide library using the *ecoPCR* program (Bellemain *et al.* 2010; Ficetola *et al.* 2010). A unique taxon was assigned to each unique sequence. This unique taxon corresponds to the last common ancestor node in the NCBI taxonomic tree of all the taxids of the sequences of the reference database that matched against the query sequence. Automatically assigned taxonomic identification was then manually curated to further eliminate those sequences that were the likely result of PCR artefacts (including chimeras, primer dimers or nuclear pseudogenes) or from obvious contaminations. Usually, chimeras can be easily identified by their low identity (<0.9) over the entire query sequence length with any known sequence and by their low frequency when compared with the main prey items. Finally, the prey items were tentatively identified by correlating sequence data with the potential leopard cat vertebrate prey known to be present in the two regions where the faeces were collected, with the constraint that such potential prey must be phylogenetically close to the prey identified in the public database by the *ecoTag* program. The significance of diet differences between ANP and CGNP was assessed by Pearson's chi-squared tests with simulated *P*-values based on 10^6 replicates, using the frequency of occurrence of prey in faeces. Results of such a test have to be analysed carefully because categories used in the contingency table are prey and several prey are detected in each faeces (Wright 2010). This potentially induced a bias if we consider that two prey in the same faeces cannot be considered as independently sampled. If it really exists, the dependency between prey count leads us to overestimate the true number of degrees of freedom. This is a main problem if the test is not rejecting the null hypothesis, but in case of the rejection of this null hypothesis, this places us on the conservative side of the decision.

Rarefaction analysis of prey in faeces originating from ANP and CGNP

We used species rarefaction curve to estimate the total number of prey species likely to be eaten by the leopard cat in the two study areas. The species accumulation, based on the faecal samples, was computed using the analytical formulas of Colwell *et al.* (2004) in EstimateS (Version 8.2, R. K. Colwell, http://purl.oclc.org/estimates).

Results

Of 181 putative felid faeces collected in the field, 38 samples were confirmed to be that of leopard cat with species-specific primers (22 from ANP of 114, and 16

from CGNP of 67). The next-generation sequencing generated about 0.6 and 0.5 million sequences for the samples without and with the blocking oligonucleotide (Table 3), respectively. After applying different filtering programmes, we finally obtained 232 and 141 sequences from the run without and with blocking oligonucleotides, respectively. Sequences within a sample having either a low frequency (e.g. <0.01 when compared with the most frequent sequence) or being very similar to a highly represented sequence were considered to be amplification/sequencing errors and were discarded. All faeces identified as leopard cat with the species-specific primers were confirmed by sequencing. The leopard cat sequence (accession numbers FR873685 and FR873686) was found with a frequency superior to 0.5 in all samples when using only the *12SV5* primer pair (Fig. 3). As in similar experiments (e.g. Deagle *et al.* 2009), we found some human contaminations corresponding to 0.2% and 5.4% of the sequences without and with the blocking oligonucleotide, respectively. A few PCR artefacts with very short sequences were also observed when using the blocking oligonucleotide, but not without blocking.

Effect of blocking oligonucleotide on predator/prey amplification

When amplifications were carried out only with *12SV5* primers, sequences of the leopard cat represented 91.6% of the total count, eight samples (sample 1–8; Fig. 3) exclusively yielded the leopard cat sequence, and 11 different prey taxa were observed in the diet. The blocking oligonucleotide *PrioB* drastically reduced the amplification of the leopard cat sequences, down to 2.2% of the total sequence count, with no leopard cat sequences observed in 31 samples. Under blocking nucleotide conditions, we recorded the amplification of seven additional prey items not previously detected when the same samples were amplified using the *12SV5* primers. The amplification failed in three samples when using the blocking oligonucleotide. The comparison of amplifications without and with blocking oligonucleotide is shown in Fig. 3.

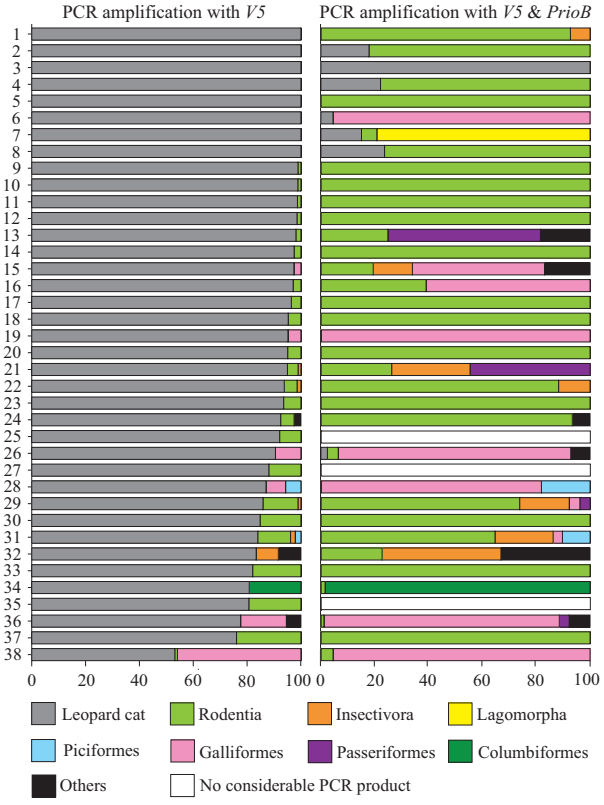

Fig. 3 Comparison of the amplifications of leopard cat and its prey sequences with *12SV5* primers for vertebrates without and with blocking oligonucleotide. The prey items are shown up to the order rank; fish and amphibians are grouped together in the ''others'' category. Each horizontal bar corresponds to the analysis of a single faeces using the *12SV5* primers, either without blocking oligonucleotide (on the left) or with blocking oligonucleotide (on the right). On each bar, the different colours represent the sequence count (%) of predator and prey items present in the sample. Samples 25, 27 and 35 did not show any considerable PCR products with blocking oligonucleotide amplification.

Table 3 Overview of the sequence counts at different stages of the analysis

Primer pair used	*12SV5F/12SV5R*	*12SV5F/12SV5R/PrioB*
Number of properly assembled sequences*	592 648	498 595
Number of unique sequences	44 441	73 414
Number of unique sequences, longer than 60 bp	44 066	46 765
Number of unique sequences, longer than 60 bp, with occurrence in the whole data set higher or equal to 100 (corresponding percentage of properly assembled sequences*)	232 (56.91%)	141 (44.84%)

*Direct and reverse sequence reads corresponding to a single DNA molecule were aligned and merged, producing what we called a ''properly assembled sequence''.

Diet composition of leopard cat

A total of 18 different prey taxa were identified in the diet of the leopard cat, seven of which were identified without ambiguity up to species level (Table 4). A maximum of seven prey items were observed within the same faeces sample, while 15 samples had only a single prey. We were not able to recover any prey DNA from only a single faeces: the experiments without and with blocking oligonucleotide with that sample produced only leopard cat sequences.

The diet composition of the leopard cat from ANP was eclectic; we observed 15 different prey taxa in 22 faeces samples. The house rat predominated the diet (in 68% of the faeces), followed by Asiatic white-toothed shrew (32%) and Murree hill frog (27%). We observed seven prey items (Himalayan wood mouse, Kashmir flying squirrel, Murree vole, Asiatic white-toothed shrew, chicken, kalij pheasant and jungle crow) within a single faeces, whereas six faeces indicated only a single prey. Overall, Rodentia dominates the diet at ANP with a presence in 91% of the faeces (Fig. 4a). Table 5 gives an overview of the leopard cat diet in Pakistan compared with previous studies.

Eight prey taxa were identified in 16 faeces from CGNP. The house rat predominated the diet (in 44% of the faeces), followed by Kashmir flying squirrel (31%) and Himalayan wood mouse (19%). Rodentia with five different prey species also dominated the diet at CGNP with a presence in 81% of the faeces (Fig. 4b).

While the leopard cat diet in both ANP and CGNP is composed mainly of rodents, the differences between these two areas were significant, both when considering all prey species independently (*P*-value: 0.01; χ^2 test with simulated *P*-value based on 10^6 replicates) and when grouping prey according to their taxonomy (Rodentia, Insectivora, Lagomorpha, Aves, Batracia and Teleostei; *P*-value: 0.03; χ^2 test with simulated *P*-value based on 10^6 replicates). As discussed in the study by Wright (2010), using Pearson chi-squared test for such data can lead to misinterpretation because of the overestimation of the degrees of freedom. By overestimating the degrees of freedom, it is more difficult to reject the null hypothesis. Consequently, rejecting the null hypothesis, as we did, places us on the conservative side of the decision.

Results of the rarefaction analysis are presented in Fig. 5. The number of prey species expected in the pooled faecal samples, based on the rarefaction curve, was 15 (95% CI: 13.91–16.09) and 8 (95% CI: 4.14–11.86) for the ANP and CGNP, respectively. In the case of ANP, 13 of 15 species with a cumulative frequency of 93% in the diet were detected in the first 11 samples. In CGNP, all of the documented prey species were detected in first 13 samples and the rest of the samples reflected their repeats.

Discussion

The leopard cat diet

All documented studies, including the present study, suggest that the order Rodentia is the primary prey base for the leopard cat (presence in 81.2–96.0% of the faeces in six studies, Table 5). Within *Rodentia*, the Muridae family dominates, with a presence in 50.0–86.4% of the faeces in Pakistan and up to 96% in other localities. The arboreal behaviour of the leopard cat (Nowell & Jackson 1996) broadens its trophic niche by enabling it to hunt tree-nesting birds and even flying squirrels in Pakistan. Birds and herpetofauna (reptiles and amphibians) are apparently the other main food groups after mammals. Birds have been reported in all studies, although the highest frequency was observed in Pakistan (presence in 18.7–45.5% of the faeces). In contrast to previous studies, where conventional methods did not allow species identification for birds, we are reporting eight distinct taxa. This specificity is an evident advantage of DNA-based diet methods recently developed. We also report fish in the diet, which have only once been reported previously (Inoue 1972). Our method did not allow the detection of invertebrates or plants, although these have been reported in other studies.

The results of the rarefaction analysis show the efficiency of the molecular method for detecting prey; this is advantageous for studying rare species that inhabit difficult terrains and that do not allow for collecting a large number of samples. Our sample size is smaller than what is generally recommended for classical diet studies; previously, 80 samples have been suggested for common leopards (Mukherjee *et al.* 1994). However, considering the greater detection efficiency of the new method, supported by the rarefaction estimates, our sample size seems to be adequate for estimating the vertebrate diet diversity of the leopard cat in the two studied regions.

The higher diversity of prey detected in samples from ANP as compared to those from CGNP probably reflects the higher productivity and diversity of temperate forests in the former park. The Kashmir flying squirrel prefers to nest on dead trees and is found in both national parks. Its frequency as a prey item was significantly higher in CGNP, the open forests of which probably make flying squirrel more susceptible to predation.

Surprisingly, the leopard cat seems to predate on prey with larger adult body size in Pakistan than in southern parts of its range (Table 5). Larger prey was

Table 4 List of prey taxa found in leopard cat diet (ANP: Ayubia National Park; CGNP: Chitral Gol National Park)

MOTU number	Accession number	Number of sequence reads	Number of occurrence		Most similar sequence(s) in public databases using BLAST				Putative taxon identification taking into account the locations where the faeces samples were collected	
			ANP 22 faeces	CGNP 16 faeces	Species name(s)	Accession number(s)	Query coverage (%)	Maximum identity (%)	Scientific name	Common name
1	FR873673	66680	15	7	*R. tanezumi/rattus*	EU273712/EU273707	100	100	*R. rattus*	House rat
2	FR873674	23746	4	0	*Microtus lusitanicus/ pyrenaicus/ duodecimcostatus/savii*	AJ972919/AJ972916/ AJ972915/AJ972914	100	95	*Hyperacrius wynnei* (?)	Murree vole (?)
3	FR873675	10848	4	0	*Phasianus colchicus/ versicolor*	FJ752430/AB164626	100	99	*Lophura leucomelanos* (?)	Kalij pheasant (?)
4	FR873676	10077	2	5	*Eoglaucomys fimbriatus*	AY227562	100	100	*E. fimbriatus*	Kashmir flying squirrel
5	FR873677	9902	5	3	*Apodemus uralensis*	AJ311128	100	100	*Apodemus rusiges*	Himalayan wood mouse
6	FR873678	9827	2	0	*Pucrasia macrolopha*	FJ752429	100	100	*P. macrolopha*	Koklass pheasant
7	FR873679	9361	7	0	*Crocidura gueldenstaedti*	AF434825	97	100	*Crocidura pullata* (?)	Asiatic white-toothed shrew (?)
8	FR873680	8700	0	1	*Columba livia*	GQ240309	100	99	*C. livia* (?)	Rock pigeon (?)
9	FR873681	8469	1	2	*Alectoris chukar*	FJ752426	100	100	*A. chukar*	Chukar partridge
10	FR873682	3626	6	0	*Nanorana parkeri*	AY322333	100	97	*Paa vicina* (?)	Murree hill frog (?)
11	FR873683	3329	0	1	*Lepus* spp.	AY292707	100	94	*Lepus capensis* (?)	Cape hare (?)
12	FR873684	2762	2	0	*Gallus gallus*	GU261719	100	100	*G. gallus*	Chicken
13	FR873687	2049	2	0	Timaliidae	AF376932	100	100	Timaliidae	Babblers
14	FR873688	1770	2	0	*Pica pica; Corvus macrorhynchos/corone/ frugilegus/albus*	HQ915867; AB042345/ AF386463/Y18522/ U38352	100	100	*C. macrorhynchos*	Jungle crow
15	FR873689	1034	2	0	*Picus viridis*	EF027325	100	97	*Dendrocopos* sp. (?)	Woodpecker (?)
16	FR873690	542	0	1	*Dryomys nitedula*	D89005	100	94	*D. nitedula* (?)	Forest dormouse (?)
17	FR873691	434	1	0	*Cephalosilurus apurensis; Liobagrus obesus*	EU179838; DQ321752	100	93	Siluriformes (?)	Cat fish (?)
18	FR873692	105	3	1	*Mus musculus castaneus*	EF108342	100	100	*Mus musculus*	House mouse

The question mark (?) after a taxon name means that the proposed identification must be confirmed by comparison with a local DNA database.

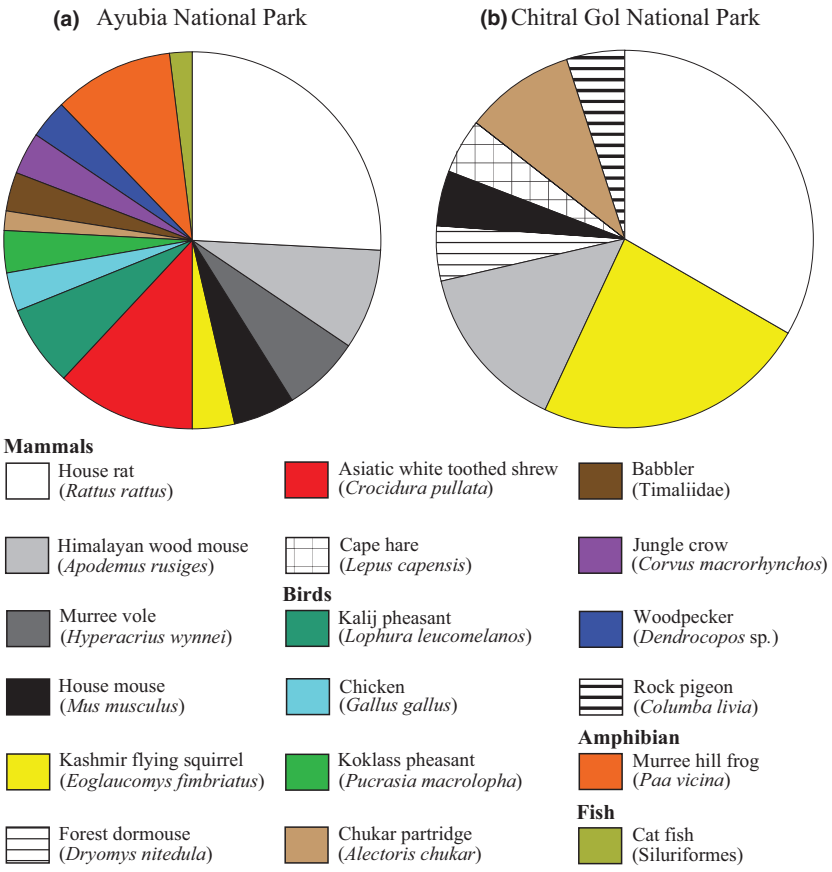

Fig. 4 Composition and comparison of the various prey items consumed and their relative frequency in the diet of the leopard cat at (a) Ayubia National Park and (b) Chitral Gol National Park.

Table 5 Comparison of leopard cat diet across its range in Asia. Except the present study, all other references estimated the diet using traditional morphology-based methods

	Occurrence in faeces, %							
Locality	Rodentia	Other mammals	Birds	Reptiles and amphibians	Fish	Invertebrates	Plant matter	References
ANP, Pakistan	90.9	31.8	45.5	27.3	4.5	Not recorded	Not recorded	Present study
CGNP, Pakistan	81.2	6.2	18.7	0.0	0.0	Not recorded	Not recorded	Present study
Negros-Panay Faunal Region, Philippines	96.0	8.0	8.0	—	—	—	12.0	Fernandez & de Guia (2011)
Khao Yai National Park, Thailand	93.8	24.5	8.2	8.2	—	36.7	—	Austin *et al.* (2007)
Sabah, Malaysian Borneo	93.1	4.2	5.6	19.4	—	11.1	11.1	Rajaratnam *et al.* (2007)
North-central Thailand	89.0	17.0	4.0	—	—	21.0	—	Grassman *et al.* (2005)
Tsushima islands, Japan	91.3	0.3	36.5	22.3	—	24.3	78.8	Tatara & Doi (1994)

usually the house rat (140–280 g), but even bigger prey were occasionally reported. Grassman *et al.* (2005) found remains of Java mouse deer (*Tragulus javanicus*; 1.18–1.28 kg from Weathers & Snyder (1977) and Endo *et al.* 2002) in leopard cat faeces, and Austin *et al.* (2007) once recorded a large ungulate (*Cervus unicolor*; 70.5–112 kg from Idris *et al.* 2000). In Pakistan, many large prey were found in the diet, including the Kashmir flying squirrel (560–734 g; Hayssen 2008), the cape hare (2.10–2.30 kg; Lu 2000), the chukar partridge (450–800 g; del Hoyo *et al.* 1994), the kalij pheasant (564–1150 g; del Hoyo *et al.* 1994), the koklass pheasant

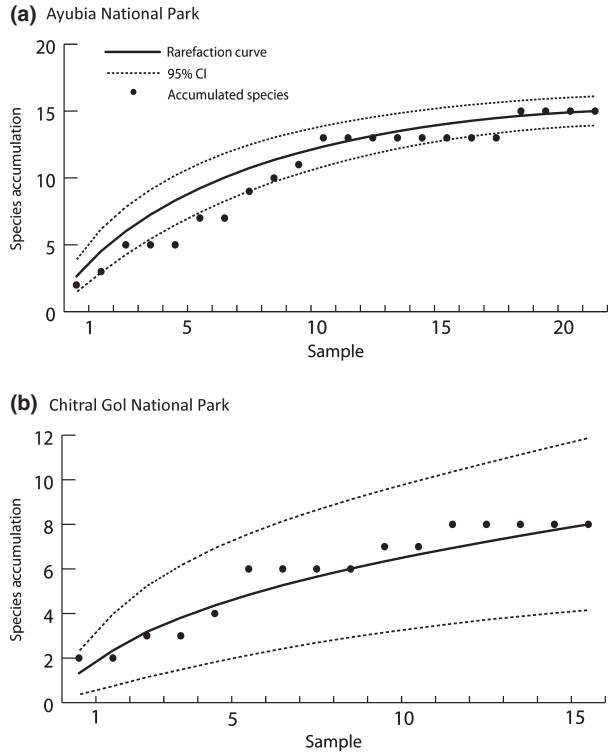

Fig. 5 Species accumulation curves based on the prey species identified in the faecal samples of leopard cat collected in ANP (a) and CGNP (b).

(930–1415 g; del Hoyo *et al.* 1994) and the jungle crow (570–580 g; Matsubara 2003). Two nonexclusive explanations can be proposed to explain such a diet shift towards larger species. First, only juveniles of the larger species may have been captured. It is important to note that remains of juveniles might be difficult to identify in faeces using traditional approaches. DNA-based methods allow straightforward taxon identification, but obviously not the age of prey. Second, the body size of the leopard cat in Pakistan might be larger than in southern areas of its distribution range, possibly explaining their ability to catch larger prey. This last hypothesis tends to be supported by the fact that the leopard cat is known to show considerable variation in size across its geographic distribution, with larger animals in China and Russia (Sunquist & Sunquist 2009), but cannot be confirmed because of the scarcity of data in Pakistan.

We conclude that the results of the present study are in general agreement with previous diet studies of the leopard cat indicating a very eclectic diet. However, the present study highlighted a possible broadening of the diet to include larger prey and provided more precise information by resolving major diet groups to a lower taxonomical level, which was not previously possible using conventional methods.

Conservation implications

The current extent of occurrence of the leopard cat in Pakistan is not resolved (Sheikh & Molur 2004). It historic range started from Chitral and extended to the eastern border of Pakistan, including areas of Swat, Hazara and Ayubia National Park (Nowell & Jackson 1996; Roberts 2005). In the north, it occupied parts of Gilgit Baltistan probably up to an elevation of 3000 m (Habibi 1977). The present study documents its current occurrence in two extremities of its historic range. A leopard cat was photographed in CGNP (SLT 2008), and authors have collected evidence of its presence in Machiara National Park, Azad Jammu and Kashmir, and western parts of the Gilgit Baltistan. This evidence suggests that the historic range of the cat in Pakistan is probably intact, although its population status needs to be determined.

Among the 18 taxa eaten by the cat in Pakistan, four (*Apodemus rusiges*, *Dryomys nitedula*, *Eoglaucomys fimbriatus* and *Lepus capensis*) are categorized as vulnerable (Sheikh & Molur 2004). Because the leopard cat is highly adaptable and appears to be widespread in Pakistan, it may be a potential threat to these species, which have a cumulative frequency of 44.7% of occurrence in faeces. A population assessment of the leopard cats is needed to evaluate the magnitude of this possible threat and to tailor an appropriate management strategy for both prey and predator.

A DNA-based approach for studying carnivore diet

Diet analysis combining next-generation sequencing and vertebrate primers with blocking oligonucleotides has tremendous potential for large-scale studies on carnivore diet. This approach is very robust and presents the complete diet profile of the vertebrate prey consumed. It is highly accurate and discriminates between closely related species in most of the cases. Moreover, a priori knowledge of prey items consumed is not essential, as it is when designing more specific DNA-based approaches. However, such analyses can yield a substantial amount of artefactual sequences including chimeras, nuclear pseudogenes and primer dimers, especially when using the blocking oligonucleotide. As our primers target highly conserved DNA regions in vertebrates, it seems unlikely that a nuclear pseudogene will better match with the *12SV5* primers than the true mitochondrial copies. Furthermore, as mitochondrial copies are much more frequent than nuclear copies, the number of occurrences of any pseudogene sequence should be much lower than the corresponding mitochondrial sequence. With regard to these possible artefacts, we recommend keeping stringent PCR conditions

as described in the Materials and Methods section and treating as significant only sequences showing a strong correspondence with a known sequence (at least >0.9) together with a relatively high frequency.

An ongoing debate on DNA-based diet studies concerns the quantification of different prey items consumed and their relative presence in sequence counts. This issue has been highlighted in several recent DNA-based dietary studies (e.g. Deagle *et al.* 2009, 2010; Soininen *et al.* 2009; Valentini *et al.* 2009a). The sequence count cannot be interpreted as quantitative for a few reasons. Biased amplification of some species has been observed when PCR was carried out of a known mixture (Polz & Cavanaugh 1998). Strong biases will occur in dietary studies when primers mismatch with certain prey sequences, resulting in the amplification inclined towards the perfect matches. The two highly conserved regions targeted by the primers *12SV5F* and *12SV5R* make the approach less susceptible to PCR biases. Deagle *et al.* (2010) suggested that differences in the density of mitochondrial DNA in tissues can also bias the sequence count. In the present study, we avoided quantitative interpretations from the results of our sequence counts and recorded only the presence/absence of the different prey in the different faeces.

The blocking oligonucleotide approach has considerable potential for its use in trophic analyses. The design of a blocking oligonucleotide specific to the leopard cat requires knowing the leopard cat sequence for the target DNA region. In this study, the blocking oligonucleotide technique not only inhibited the amplification of the leopard cat DNA, but also uncovered seven more prey taxa in the diet that had not been amplified previously without the blocking oligonucleotide. We used a high concentration of *PrioB* (2 μM) compared with *12SV5F* and *12SV5R* primers (0.1 μM each). For each faeces sample, we systematically ran amplifications without and with blocking oligonucleotide, as amplification with such a relatively high *PrioB* concentration might fail.

One limitation of the approach with the *12SV5F* and *12SV5R* primers proposed here is that it only identifies vertebrate prey. Many carnivores have a more diverse diet, including invertebrates and plants. For example, the Eurasian badger (*Meles meles*) exploits a wide range of food items, especially earthworms, insects and grubs. It also eats small mammals, amphibians, reptiles and birds as well as roots and fruits (Revilla & Palomares 2002). For instance, to study the badger's diet, we suggest complementing the primers for vertebrates with several additional systems, such as primers targeting plant taxa (e.g. Taberlet *et al.* 2007; Valentini *et al.* 2009a) or earthworms (Bienert *et al.* 2012).

One more limitation of this approach for identifying vertebrates is that cases of cannibalism cannot be detected. In such a situation, the predator DNA cannot be distinguished from the prey DNA that belongs to the same species. This limitation was not acknowledged in previous DNA-based diet analyses for vertebrate predators, despite the cases of cannibalism have been documented, for example, in Otariidae (e.g. Wilkinson *et al.* 2000). However, if cannibalism is important from a behavioural point of view, it represents a marginal phenomenon when studying the diet.

Another potential difficulty concerns species identification. In some cases, we had to combine the best match using public databases together with expert knowledge about the available prey in the location where the faeces were collected. For example, in our study, the best match (99%) for MOTU number 3 in public databases corresponded to two species of the genus *Phasianus* (*P. colchicus* and *P. versicolor*). These two species are not recorded in ANP, and thus, we identify this MOTU as the closest relative (Huang *et al.* 2009; Shen *et al.* 2010) occurring in ANP, the kalij pheasant (*Lophura leucomelanos*). If the identification of the kalij pheasant seems reliable, some other putative identifications are more problematic, particularly those having a relatively low identities with known sequences in public databases (i.e. *Hyperacrius wynnei*, *Paa vicina*, *L. capensis*, *Dendrocopos* sp., and *D. nitedula*). To remove such uncertainties, we recommend constructing a local reference database when possible.

The results of the present study correspond to summer diet and may not reflect the complete diet profile of the leopard cat in Pakistan. In future, it would be interesting to collect samples throughout the year, with the attendant possibility of revealing more prey taxa than what we have observed in this study.

Conclusion

Noninvasive sampling is the only way to study the diet of elusive animals like the leopard cat. In Pakistan, we obtained results confirming the eclectic characteristics of this predator, together with an extension of the diet towards larger prey. The DNA-based approach has a better resolution than conventional approach-based identification of prey from hair and bone remains. While DNA-based methods cannot assess prey ages, conventional approaches might reveal the potential ages of the prey when necessary, possibly determining whether juveniles or adults of larger prey were consumed. As a consequence, DNA-based diet analysis can provide a valuable complement to conventional methods.

The DNA-based approach we propose here is particularly robust and simple to implement and allows the possibility of very large-scale analyses. It can be applied

to other carnivore species with only a slight adjustment concerning the design of the blocking oligonucleotide.

Acknowledgements

The authors are thankful to A. Valentini, D. Rioux and L. Gielly for their valuable suggestions and comments throughout the course of this study. For English correction in the manuscript, we express our gratitude to M. MacFarlane and M. Reynolds. W. Shehzad and T. Riaz were supported by PhD scholarship from the HEC (Higher Education Commission), Government of Pakistan. The Himalayan Wildlife Foundation and Snow Leopard Trust provided financial support for sample collections in Ayubia and Chitral Gol National Parks. The staff of the Khyber Pakhtunkhwa Wildlife Department and Snow Leopard Trust assisted in sample collection.

References

Austin SC, Tewes ME, Grassman Jr LI, Silvy NJ (2007) Ecology and conservation of the leopard can *Prionailurus bengalensis* and clouded leopard *Neofelis nebulosa* in Khao Yai National Park, Thailand. *Acta Zoologica Sinica*, **53**, 1–14.

Bagchi S, Mishra C (2006) Living with large carnivores: predation on livestock by the snow leopard (*Uncia uncia*). *Journal of Zoology*, **268**, 217–224.

Bellemain E, Carlsen T, Brochmann C *et al.* (2010) ITS as DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiology*, **10**, 189.

Bienert R, de Danieli S, Miquel C *et al.* (2012) Tracking earthworm communities from soil DNA. *Molecular Ecology*, **21**, doi: 10.1111/j.1365-294X.2011.05407.x.

Blankenship LE, Yayanos AA (2005) Universal primers and PCR of gut contents to study marine invertebrate diets. *Molecular Ecology*, **14**, 891–899.

Bradley BJ, Stiller M, Doran-Sheehy DM *et al.* (2007) Plant DNA sequences from feces: potential means for assessing diets of wild primates. *American Journal of Primatology*, **69**, 699–705.

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by *Taq* polymerase: primer modification that facilitate genotyping. *BioTechniques*, **20**, 1004–1010.

Colwell RK, Mao CX, Chang J (2004) Interpolating, extrapolating, and comparing incidence-based species accumulation curves. *Ecology*, **85**, 2717–2727.

Deagle BE, Jarman SN, Pemberton D, Gales NJ (2005a) Genetic screening for prey in the gut contents from a giant squid (*Architeuthis* sp.). *Journal of Heredity*, **96**, 417–423.

Deagle BE, Tollit DJ, Jarman SN *et al.* (2005b) Molecular scatology as a tool to study diet: analysis of prey DNA in scats from captive Steller sea lions. *Molecular Ecology*, **14**, 1831–1842.

Deagle BE, Eveson JP, Jarman SN (2006) Quantification of damage in DNA recovered from highly degraded samples – a case study on DNA in faeces. *Frontier in Zoology*, **3**, 11.

Deagle BE, Gales NJ, Evans K *et al.* (2007) Studying seabird diet through genetic analysis of faeces: a case study on macaroni penguins (*Eudyptes chrysolophus*). *PLoS ONE*, **2**, e831.

Deagle BE, Kirkwood R, Jarman SN (2009) Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular Ecology*, **18**, 2022–2038.

Deagle BE, Chiaradia A, McInnes J, Jarman SN (2010) Pyrosequencing faecal DNA to determine diet of little penguins: is what goes in what comes out? *Conservation Genetics*, **11**, 2039–2048.

Dunshea G (2009) DNA-based diet analysis for any predator. *PLoS ONE*, **4**, e5252.

Endo H, Kimura J, Sasaki M *et al.* (2002) Functional morphology of the mastication muscles in the lesser and greater mouse deer. *Journal of Veterinary Medical Science*, **64**, 901–905.

Farooque M (2007) *Management Plan of Ayubia National Park*. Khyber Pakhtunkhwa Wildlife Department, Peshawar.

Fernandez DAP, de Guia APO (2011) Feeding habits of Visayan leopard cats (*Prionailurus bengalensis rabori*) in sugarcane fields of Negros Occidental, Philippines. *Asia Life Sciences*, **20**, 143–154.

Ficetola GF, Coissac E, Zundel S *et al.* (2010) An *in silico* approach for the evaluation of DNA barcodes. *BMC Genomics*, **11**, 434.

Gaines WL (2001) Large carnivore surveys in the North Karakorum Mountains, Pakistan. *Natural Areas Journal*, **21**, 168–171.

GoNWFP & IUCN (1996) *Sarhad Provincial Conservation Strategy*. Sarhad Programme Office, IUCN–The World Conservation Union, Peshawar.

Grassman LI, Tewes ME, Silvy NJ, Kreetiyutanont K (2005) Spatial organization and diet of the leopard cat (*Prionailurus bengalensis*) in north-central Thailand. *Journal of Zoology*, **266**, 45–54.

Green SJ, Minz D (2005) Suicide polymerase endonuclease restriction, a novel technique for enhancing PCR amplification of minor DNA templates. *Applied and Environmental Microbiology*, **71**, 4721–4727.

Habibi K (1977) *The Mammals of Afghanistan: Their Distribution and Status*. UNDP, FAO and Ministry of Agriculture, Kabul.

Hayssen V (2008) Patterns of body and tail length and body mass in Sciuridae. *Journal of Mammalogy*, **89**, 852–873.

Höss M, Kohn M, Pääbo S, Knauer F, Schröder W (1992) Excrement analysis by PCR. *Nature*, **359**, 199.

del Hoyo J, Elliot A, Sargatal J (1994) *Handbook of the Birds of the World*, Vol. 2. New World Vultures to Guinea Fowl Lynx Edicions, Barcelona.

Huang ZH, Liu NF, Xiao YA *et al.* (2009) Phylogenetic relationships of four endemic genera of the Phasianidae in China based on mitochondrial DNA control-region genes. *Molecular Phylogenetics and Evolution*, **53**, 378–383.

Idris I, Moin S, Sulah S, Jiwan D (2000) Some physical characteristics of Sambar deer (*Cervus unicolor*). *Pertanika Journal of Tropical Agricultural Science*, **33**, 55–59.

Inoue T (1972) The food habit of Tsushima leopard cat, *Felis bengalensis* spp., analysed from their scats. *Journal of the Mammalogical Society of Japan*, **5**, 155–169.

Izawa M, Doi T (1991) Status of conservation and management of two species of felidae in Japan (in Japanese). *Mammalian Science*, **31**, 15–22.

Jarman SN, Gales NJ, Tierney M, Gill PC, Elliott NG (2002) A DNA-based method for identification of krill species and its

application to analysing the diet of marine vertebrate predators. *Molecular Ecology*, **11**, 2679–2690.

Jarman SN, Deagle BE, Gales NJ (2004) Group-specific polymerase chain reaction for DNA-based analysis of species diversity and identity in dietary samples. *Molecular Ecology*, **13**, 1313–1322.

Jarman SN, Redd KS, Gales NJ (2006) Group-specific primers for amplifying DNA sequences that identify Amphipoda, Cephalopoda, Echinodermata, Gastropoda, Isopoda, Ostracoda and Thoracica. *Molecular Ecology Notes*, **6**, 268–271.

Kageyama T, Sato Y, Nishizawa S, Teramae N (2008) Competitive binding of small ligands to nucleobases in AP site-containing DNA duplexes. *Nucleic Acids Symposium Series*, **52**, 119–120.

King RA, Vaughan IP, Bell JR, Bohan DA, Symondson WOC (2010) Prey choice by carabid beetles feeding on an earthworm community analysed using species- and lineage-specific PCR primers. *Molecular Ecology*, **19**, 1721–1732.

Kohn MH, Wayne RK (1997) Facts from feces revisited. *Trends in Ecology & Evolution*, **12**, 223–227.

Krahn MM, Herman DP, Matkin CO *et al.* (2007) Use of chemical tracers in assessing the diet and foraging regions of eastern North Pacific killer whales. *Marine Environmental Research*, **63**, 91–114.

Lau MWN, Fellowes JR, Chan BPL (2010) Carnivores (Mammalia: Carnivora) in South China: a status review with notes on the commercial trade. *Mammal Review*, **40**, 247–292.

Li MJ, Sato Y, Nishizawa S *et al.* (2009) 2-aminopurine-modified abasic-site-containing duplex DNA for highly selective detection of theophylline. *Journal of the American Chemical Society*, **131**, 2448–2449.

Liles MR, Manske BF, Bintrim SB, Handelsman J, Goodman RM (2003) A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Applied and Environmental Microbiology*, **69**, 2684–2691.

Lovari S, Boesi R, Minder I *et al.* (2009) Restoring a keystone predator may endanger a prey species in a human-altered ecosystem: the return of the snow leopard to Sagarmatha National Park. *Animal Conservation*, **12**, 559–570.

Lu X (2000) Body weights of the Cape hare *Lepus capensis* in the northern China. *Acta Theriologica*, **45**, 271–280.

Macdonald DW, Loveridge AJ, Nowell K (2010) Dramatis personae: an introduction to the wild felids. In: *Biology and Conservation of Wild Felids* (eds Macdonald DW and Loveridge AJ), pp. 3–58. Oxford University Press, New York.

Magnuson VL, Ally DS, Nylund SJ *et al.* (1996) Substrate nucleotide-determinated non-templated addition of adenine by Taq DNA polymerase: implications for PCR-based genotyping and cloning. *BioTechniques*, **21**, 700–709.

Matsubara H (2003) Comparative study of territoriality and habitat use in syntopic Jungle Crow (*Corvus macrorhynchos*) and Carrion Crow (*C. corone*). *Ornithological Science*, **2**, 103–111.

Mills LS, Soule ME, Doak DF (1993) The keystone-species concept in ecology and conservation. *BioScience*, **43**, 219–224.

Mirza ZB (2003) *Biological Baseline Study of Chitral Gol National Park*. Protected Areas Management Project, Islamabad.

Mitani N, Mihara S, Ishii N, Koike H (2009) Clues to the cause of the Tsushima leopard cat (*Prionailurus bengalensis euptilura*) decline from isotopic measurements in three species of Carnivora. *Ecological Research*, **24**, 897–908.

Mukherjee S, Goyal SP, Chellam R (1994) Standardization of scat analysis techniques for leopard (*Panthera pardus*) in Gir National Park, western India. *Mammalia*, **58**, 139–143.

Mukherjee S, Krishnan A, Tamma K *et al.* (2010) Ecology driving genetic variation: a comparative phylogeography of jungle cat (*Felis chaus*) and leopard cat (*Prionailurus bengalensis*) in India. *PLoS ONE*, **5**, 16.

Natoli E (1990) Mating strategies in cats – a comparison of the role and importance of infanticide in domestic cats, *Felis catus* L and lions, *Panthera leo* L. *Animal Behaviour*, **40**, 183–186.

Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, **48**, 443–453.

Nowell K, Jackson P (1996) *Wild Cats, Status Survey and Conservation Action Plan*. IUCN, Gland, Switzerland, 382p.

Oli MK, Taylor IR, Rogers ME (1994) Snow leopard (*Panthera Uncia*) predation of livestock – an assessment of local perceptions in the Annapurna Conservation area, Nepal. *Biological Conservation*, **68**, 63–68.

Pegard A, Miquel C, Valentini A *et al.* (2009) Universal DNA based methods for assessing the diet of grazing livestock and wildlife from faeces. *Journal of Agricultural and Food Chemistry*, **57**, 5700–5706.

Polz MF, Cavanaugh CM (1998) Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology*, **64**, 3724–3730.

Power ME, Tilman D, Estes JA *et al.* (1996) Challenges in the quest for keystones. *BioScience*, **46**, 609–620.

Rajaratnam R, Sunquist M, Rajaratnam L, Ambu L (2007) Diet and habitat selection of the leopard cat (*Prionailurus bengalensis borneoensis*) in an agricultural landscape in Sabah, Malaysian Borneo. *Journal of Tropical Ecology*, **23**, 209–217.

Revilla E, Palomares F (2002) Does local feeding specialization exist in Eurasian badgers? *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, **80**, 83–93.

Rho P (2009) Use of GIS to develop a multivariate habitat model for the leopard cat (*Prionailurus bengalensis*) in mountainous region of Korea. *Journal of Ecology & Field Biology*, **32**, 229–236.

Riaz T, Shehzad W, Viari A *et al.* (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, **39**, e145.

Roberts TJ (2005) *Field Guide to the Large and Medium Sized Mammals of Pakistan*. Oxford University Press, Oxford, UK.

Scott DM, Gemita E, Maddox TM (2004) Small cats in human modified landscapes in Sumatra. *Cat News*, **40**, 23–25.

Sheikh KM, Molur S (2004) Status and red list of Pakistan's mammals. Based on the Conservation Assessment and Management Plan Workshop. IUCN Pakistan.

Shen YY, Liang L, Sun YB *et al.* (2010) A mitogenomic perspective on the ancient, rapid radiation in the Galliformes with an emphasis on the Phasianidae. *BMC Evolutionary Biology*, **10**, 132.

Snow Leopard Trust (2008) Population monitoring of large carnivores in Chitral Gol National Park. Progress update for 2007–2008, submitted to NWFP Wildlife Department, Peshawar, Pakistan.

Soininen EM, Valentini A, Coissac E *et al.* (2009) Analysing diet of small herbivores: the efficiency of DNA barcoding

coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures. *Frontiers in Zoology*, **6**, 9.

Sunquist ME, Sunquist FC (2002) *Wild Cats of the World*. University of Chicago Press, Chicago, Illinois, 416pp.

Sunquist ME, Sunquist FC (2009) Family Felidae (cats). In: *Handbook of the Mammals of the World. Vol. 1. Carnivores* (eds Wilson DE and Mittermeier RA), pp. 54–168. Lynx Edicions, Barcelona.

Symondson WOC (2002) Molecular identification of prey in predator diets. *Molecular Ecology*, **11**, 627–641.

Taberlet P, Coissac E, Pompanon F *et al*. (2007) Power and limitations of the chloroplast *trn*L(UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, **35**, e14.

Tatara M, Doi T (1994) Comparative analyses on food-habits of Japanese marten, Siberian weasel and leopard cat in The Tsushima islands, Japan. *Ecological Research*, **9**, 99–107.

Valentini A, Miquel C, Nawaz MA *et al.* (2009a) New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the *trn*L approach. *Molecular Ecology Resources*, **9**, 51–60.

Valentini A, Pompanon F, Taberlet P (2009b) DNA barcoding for ecologists. *Trends in Ecology and Evolution*, **24**, 110–117.

Vestheim H, Jarman SN (2008) Blocking primers to enhance PCR amplification of rare sequences in mixed samples – a case study on prey DNA in Antarctic krill stomachs. *Frontiers in Zoology*, **5**, 11.

Vestheim H, Edvardsen B, Kaartvedt S (2005) Assessing feeding of a carnivorous copepod using species-specific PCR. *Marine Biology*, **147**, 381–385.

Wang YF, Ng MTT, Zhou TY *et al*. (2008) C3-Spacer-containing circular oligonucleotides as inhibitors of human topoisomerase I. *Bioorganic & Medicinal Chemistry Letters*, **18**, 3597–3602.

Watanabe S (2009) Factors affecting the distribution of the leopard cat *Prionailurus bengalensis* on East Asian islands. *Mammal Study*, **34**, 201–207.

Weathers WW, Snyder GK (1977) Hemodynamics of the lesser mouse deer, *Tragulus javanicus*. *Journal of Applied Physiology*, **42**, 679–681.

Wilkinson IS, Childerhouse SJ, Duignan PJ, Gulland FMD (2000) Infanticide and cannibalism in the New Zealand sea lion, *Phocarctos hookeri*. *Marine Mammal Science*, **16**, 494–500.

Wright BE (2010) Use of chi-square tests to analyze scat-derived diet composition data. *Marine Mammal Science*, **26**, 395–401.

Zaidi RH, Jaal Z, Hawkes NJ, Hemingway J, Symondson WOC (1999) Can multiple-copy sequences of prey DNA be detected amongst the gut contents of invertebrate predators? *Molecular Ecology*, **8**, 2081–2087.

T.R., E.C. and P.T. are co-inventors of a pending French patent on the primer pair named *12SV5F* and *12SV5R* and on the use of the amplified fragment for identifying vertebrate species from environmental samples. This patent only restricts commercial applications and has no impact on the use of this method by academic researchers.

Data accessibility

DNA sequences of the V5 loop of the mitochondrial 12S gene: GenBank accessions FR873673–FR873692.

Fasta file and filtered data deposited in the Dryad repository: doi: 10.5061/dryad.443t4m1q.