Received: 11 March 2019 | Revised: 15 July 2019 | Accepted: 22 July 2019

DOI: 10.1111/1755-0998.13069

RESOURCE ARTICLE

Terrestrial mammal surveillance using hybridization capture of environmental DNA from African waterholes

Peter Andreas Seeber[1] | **Gayle K. McEwen**[1] | **Ulrike Löber**[1,2,3,4] | **Daniel W. Förster**[5] | **Marion Linda East**[6] | **Jörg Melzheimer**[7] | **Alex D. Greenwood**[1,8]

[1]Department of Wildlife Diseases, Leibniz-Institute for Zoo and Wildlife Research, Berlin, Germany

[2]Experimental and Clinical Research Center, Charité-Universitätsmedizin Berlin and Max Delbrück Center for Molecular Medicine, Berlin, Germany

[3]Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany

[4]Berlin Center for Genomics in Biodiversity Research (BeGenDiv), Berlin, Germany

[5]Department of Evolutionary Genetics, Leibniz-Institute for Zoo and Wildlife Research, Berlin, Germany

[6]Department of Ecological Dynamics, Leibniz-Institute for Zoo and Wildlife Research, Berlin, Germany

[7]Department of Evolutionary Ecology, Leibniz-Institute for Zoo and Wildlife Research, Berlin, Germany

[8]Department of Veterinary Medicine, Freie Universität Berlin, Berlin, Germany

Correspondence
Peter Andreas Seeber, Limnologal Institute, University of Constance, Mainaustrasse 252, Konstanz 78467, Germany.
Email: seeber.pa@gmail.com

Alex D. Greenwood, Department of Wildlife Diseases, Leibniz-Institute for Zoo and Wildlife Research, Alfred-Kowalke-Strasse 17, Berlin 10315, Germany.
Email: greenwood@izw-berlin.de

Funding information
Leibniz Gemeinschaft, Grant/Award Number: SAW-2015-IZW-1 440

Abstract

Determining species distributions can be extremely challenging but is crucial to ecological and conservation research. Environmental DNA (eDNA) approaches have shown particular promise in aquatic systems for several vertebrate and invertebrate species. For terrestrial animals, however, eDNA-based surveys are considerably more difficult due to the lack of or difficulty in obtaining appropriate sampling substrate. In water-limited ecosystem where terrestrial mammals are often forced to congregate at waterholes, water and sediment from shared water sources may be a suitable substrate for noninvasive eDNA approaches. We characterized mitochondrial DNA sequences from a broad range of terrestrial mammal species in two different African ecosystems (in Namibia and Tanzania) using eDNA isolated from native water, sediment and water filtered through glass fibre filters. A hybridization capture enrichment with RNA probes targeting the mitochondrial genomes of 38 mammal species representing the genera/families expected at the respective ecosystems was employed, and 16 species were identified, with a maximum mitogenome coverage of 99.8%. Conventional genus-specific PCRs were tested on environmental samples for two genera producing fewer positive results than hybridization capture enrichment. An experiment with mock samples using DNA from non-African mammals showed that baits covering 30% of nontarget mitogenomes produced 91% mitogenome coverage after capture. In the mock samples, over-representation of DNA of one species still allowed for the detection of DNA of other species that was at a 100-fold lower concentration. Hybridization capture enrichment of eDNA is therefore an effective method for monitoring terrestrial mammal species from shared water sources.

KEYWORDS
biodiversity, eDNA, mitogenome, species monitoring, terrestrial

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2019 The Authors. *Molecular Ecology Resources* published by John Wiley & Sons Ltd

1 | INTRODUCTION

Information on species distribution and abundance is important in field ecology and species conservation in situ. Survey methods using environmental DNA (eDNA) have recently demonstrated that eDNA-based surveys can, in some cases, outperform traditional survey approaches to determine species presence (Civade et al., 2016; Deiner, Fronhofer, Mächler, Walser, & Altermatt, 2016; Hänfling et al., 2016).

Most studies using eDNA for species detection have focused on aquatic organisms. DNA is shed into the environment and subsequently becomes homogenously distributed (Deiner, Bik, et al., 2017; Deiner, Renshaw, et al., 2017; Thomsen & Willerslev, 2015). In contrast, eDNA surveys of terrestrial species have been applied less comprehensively and typically have relied on faecal samples or gut content of blood-feeding invertebrates (Deiner, Bik, et al., 2017). Environmental bulk samples have been used for eDNA approaches targeting arthropods (Shokralla et al., 2016); however, for terrestrial organisms, the lack of a substrate that is comparable to water in terms of ease of sampling and representativeness of eDNA content may generally hamper the efficiency of eDNA-based surveys in terrestrial habitats. Many mammal species living in seasonally water-limited ecosystems depend (to a varying extent) on drinking from common water sources. DNA shed from an organism diffuses rapidly in water; therefore, in principle eDNA may be detected by sampling at any remote point of a stagnant water body (Rees, Maddison, Middleditch, Patmore, & Gough, 2014). In practice, however, DNA readily binds to suspended particles that diffuse less easily and eventually will sediment (Turner et al., 2014). Hence, both water and sediment are potential sources of eDNA in water-limited ecosystems.

The amount of mitochondrial DNA (mtDNA) by far exceeds that of nuclear DNA in shed cells (Thomsen & Willerslev, 2015). Numerous regions within the mitochondrial genome (mitogenome) can be used for mammalian taxonomic classification and phylogenetics. Most eDNA studies have employed metabarcoding using PCR amplification of one or few loci (e.g. cytochrome b, COI, 12S rRNA or the mitochondrial D-loop) for species detection and discrimination (Aylagas, Borja, Irigoien, & Rodríguez-Ezpeleta, 2016; Dougherty et al., 2016; Hänfling et al., 2016; Hunter et al., 2015; Mächler, Deiner, Steinmann, & Altermatt, 2014; Olds et al., 2016; Shokralla et al., 2016; Ushio, Murata, et al., 2017; Valentini et al., 2016). However, a metabarcoding approach may be error-prone in terms of primer bias and higher likelihood of false-negative results when the target DNA is highly degraded or diluted (Deiner, Bik, et al., 2017). Hybridization capture enrichment (hybridization capture, henceforth) effectively avoids problems such as primer bias and can substantially increase the yield of target DNA retrieved from environmental samples (Dowle, Pochon, Banks, Shearer, & Wood, 2016; Wilcox et al., 2018). Hybridization capture efficiently enriches target DNA and is based on hybridization of sample DNA with synthetic complementary DNA or RNA oligonucleotide baits (Gasc, Peyretaillade, & Peyret, 2016). These baits are designed from the targeted reference sequence, and multiple target

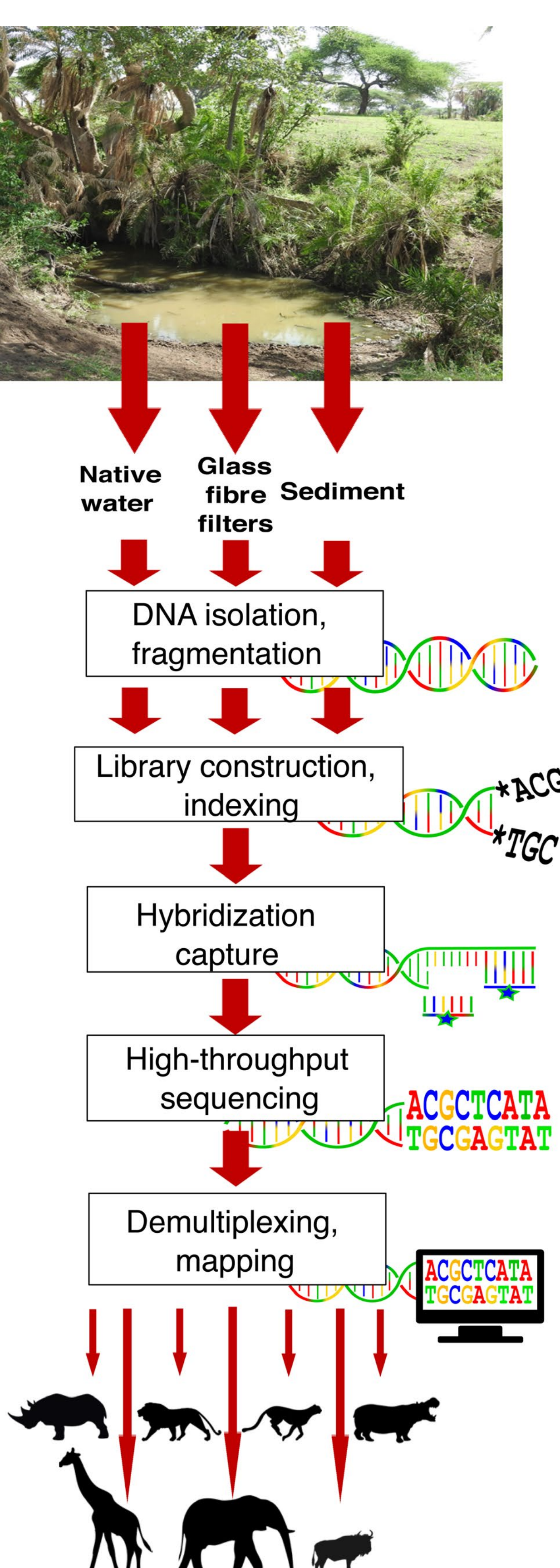

FIGURE 1 Workflow from collection of environmental samples to assignment of species sequences

genomes can be targeted in a single capture reaction. Furthermore, hybridization capture can tolerate high divergence between baits and target. Successful capture has been shown with baits which are 20% divergent from their target sequences (Hawkins et al., 2016); however, capture efficiency may decrease at around 5%–10% divergence (Paijmans, Fickel, Courtiol, Hofreiter, & Förster, 2016).

In the current study, a hybridization capture approach targeting the mitogenomes of 38 mammalian species was implemented to recover eDNA from water and sediment from waterholes in two African countries. Samples were collected in the rainy and in the dry season to assess whether animal aggregations at shrinking water bodies would be reflected in the abundance of eDNA in water and sediment samples. We compared the hybridization capture results with those of genus-specific conventional PCRs. Furthermore, we tested hybridization capture concentration-dependent efficiency using experimentally diluted target DNA samples of non-African mammals to determine the methodological limitations. Our results demonstrate that hybridization capture enrichment can be used effectively for monitoring the presence of terrestrial mammal species from shared water sources.

2 | MATERIALS AND METHODS

2.1 | Sampling

Water and sediment were sampled from natural and pumped water holes in two regions in Namibia (Etosha National Park [NP] in northern Namibia and ||Karas region in southern Namibia), and from water bodies in the Serengeti NP, Tanzania. In Namibia, samples were collected in November 2015 and April 2016, both during a dry period. In Tanzania, samples were collected in February, June/July and early October 2016, which corresponded to the rainy season and early and late dry season, respectively. At each sampling site ($N = 14$), water was collected from the surface in sterile 50 ml tubes, and sediment samples were collected by filling a sterile 50 ml tube with sediment from the submerged top layer at 20–50 cm from the perimeter (Figure 1). Additionally, at five sites in the Serengeti NP native water was filtered using glass fibre filters (Whatman, GE Healthcare) to concentrate the total amount of DNA. A disposable 50 ml syringe was attached to a filter holder (Swinnex, Merck Millipore) containing a glass microfibre filter (nominal pore size approximately 0.7 µm, diameter 25 mm). Filtration was carried out manually and repeated for each filter until a total volume of 300 ml was passed through the filter. Each glass fibre filter was then removed from the filter holder using sterile forceps, folded and stored in a 2 ml cryotube. All water and sediment samples and water filters were stored on ice packs during the respective field trip and frozen at −20°C upon return to the field station (within <6 hr). The GPS coordinates of each sampling point were made available online (NCBI BioProject accession number PRJNA515605).

2.2 | Mock samples

Mock samples with known DNA concentrations of different mammal species were generated to determine the sensitivity of the hybridization capture method to background DNA and dilution. Three non-African species, koala (*Phascolarctos cinereus*), musk ox (*Ovibos moschatus*) and polar bear (*Ursus maritimus*), were selected in order to identify any potential cross-contamination of environmental samples with DNA from mock samples. DNA was isolated from tissue samples using a commercially available extraction kit (NucleoSpin® tissue kit, Macherey-Nagel) following the manufacturer's instructions and pooled diluted DNA extracts in (a) equal concentrations (20 ng/µl each; 'mock sample #1') and (b) different concentrations (koala 100 ng/µl, musk ox 10 ng/µl and polar bear 1 ng/µl; 'mock sample #2').

2.3 | eDNA isolation and genomic library preparation

DNA from sediment samples was isolated using a commercial kit (NucleoSpin® Soil kit, Macherey-Nagel) following the manufacturer's instructions. As the major fraction of DNA in turbid water was likely bound to suspended solids (Turner et al., 2014), we isolated eDNA from water samples by centrifuging 25 ml aliquots of each sample at 4,000 G for 45 min, discarding the supernatant and processing the pellet the same way as sediment samples. Environmental DNA from glass fibre filters was isolated by manually shredding each filter using sterile forceps and then subjecting it to a DNA isolation using the extraction protocol for sediment. With each batch of DNA isolations ($N = 6$), we included two negative controls (using 20 µl of DNA-free water) that were subsequently also included in the library preparation process.

Genomic library preparation from African and mock samples was based on the protocol of Meyer and Kircher (2010) for building Illumina sequencing libraries for multiplexed target capture and sequencing, with modifications as follows. Genomic DNA was physically sheared to an average size of 400 nucleotides using a Covaris M220 ultrasonicator (Covaris). Fragmentation efficiency and fragment size distribution were visualized by running each sample on an Agilent 2,200 TapeStation (Agilent) using a D1000 chip. For all downstream processing steps, we used polypropylene low DNA-binding reaction tubes (Sarstedt) to minimize DNA retention by the plastic surface. Genomic libraries were built from sheared DNA using a NEBNext dual-index kit (New England Biolabs, NEB) following the manufacturer's instructions. Samples were dual-indexed with unique index combinations using five PCR cycles. The final library concentration after indexing was quantified on the Agilent 2,200 TapeStation. The negative controls had no measurable DNA after the DNA extractions and did not produce amplification products after library preparation.

2.4 | Oligonucleotide bait design, hybridization capture and sequencing

To design RNA oligonucleotide baits for hybridization capture, we compiled the mitochondrial genome sequences of 38 selected mammal species (representing 26 families in 15 orders; Table 1) that are

TABLE 1 Reference mitogenomes used for bait design and mapping. Representative species expected at specific sites (target species) of which no mitogenome was available were replaced by closely related species (replacement species)

order	family	Target species	Replacement species	NCBI accession no.
Afrosoricida	Chrysochloridae	*Eremitalpa granti*		AM904729.1
		Chrysochloris sp.	*Chrysochloris asiatica*	NC_004920.1
Proboscidea	Elephantidae	*Loxodonta africana*		DQ316069.1
Hyracoidea	Procaviidae	*Procavia* sp.		AB096865.1
Primates	Galagidae	*Otolemur crassicaudatus*		KJ434961.1
	Cercopithecidae	*Papio Anubis*		NC_020006.2
Rodentia	Pedetidae	*Pedetes surdaster*	*Pedetes capensis*	HE983623.1
	Muridae	various	*Acomys cahirinus*	NC_020758.1
Lagomorpha	Leporidae	*Lepus capensis*		GU937113.1
Pholidota	Manidae	*Smutsia temminckii*		KP125951.1
Tubulidentata	Orycteropodidae	*Orycteropus afer*		Y18475.1
Eulipotyphla	Erinaceidae	*Atelerix sp.*	*Erinaceus europaeus*	NC_002080.2
Macroscelididae	Macroscelididae	*Elephantulus* sp.		NC_004921.1
Chiroptera	Pteropodidae	*Pteropus vampyrus*		NC_026542.1
	Microbats	various	*Myotis myotis*	KT901455.1
Soricomorpha	Soricidae	*Crocidura attenuata*		KP120863.2
Perissodactyla	Equidae	*Equus quagga*		JX312733.2
		Equus zebra		JX312724.1
	Rhinocerotidae	*Diceros bicornis*		FJ905814.1
Artiodactyla	Hippopotamidae	*Hippopotamus amphibius*		NC_000889.1
	Suidae	*Phacochoerus africanus*		DQ409327.1
	Giraffidae	*Giraffa camelopardalis*		AP003424.1
	Bovidae	*Syncerus caffer*		EF536353.1
		Tragelaphus scriptus		JN632707.1
		Nanger granti		JN632666.1
		Kobus ellipsiprymnus		JN632651.1
		Connochates taurinus		JN632628.1
		Aepyceros melampus		JN632592.1
Carnivora	Viverridae	*Genetta* sp.	*Genetta servalina*	NC_024568.1
		Civetticis civetta		NC_033378.1
	Herpestidae	*various*	*Herpestes javanicus*	KY117548.1
	Hyaenidae	*Crocuta crocuta*		NC_020670.1
		Hyaena hyaena		NC_020669.1
	Felidae	*Panthera pardus*		EF551002.1
		Acinonyx jubatus		AY463959.1
	Canidae	*Canis mesomelas*	*Canis aureus*	KT448273.1
		Lycaon pictus		CM007554.1
	Mustelidae	*Mellivora capensis*	*Lutra lutra*	FJ236015.1

known to be present in the regions studied (National Museum of History, 2005). Species that were not represented in the NCBI database were replaced by closely related taxa (genus or family level; *N* = 7; 'replacement species', Table 1). Species were selected in order to cover every family which was represented at the field sites with at least one species, apart from the Chiroptera and Rodentia which contain multiple families (two species were chosen from each). The Tenrecidae were represented in the bait set by a member of their sister family Chrysochloridae as no mitogenome sequence within the Tenrecidae family was available. The compiled mitochondrial sequences were submitted to Arbor Biosciences for custom design of 80 bp baits. In total, 19,496 unique baits were produced;

the bait set was made available in an online repository (https://doi.org/10.17632/zsd3w5vttk.1; Seeber et al., 2019). Due to the expected degree of target DNA degradation in environmental samples (Pilliod, Goldberg, Arkle, & Waits, 2014), baits were tiled threefold to allow for overlap and ensure sufficient coverage.

For each hybridization capture, we pooled two genomic libraries of approximately equal DNA concentration. Mock sample captures were performed in individual reactions (i.e.. one genomic library per reaction). The hybridization capture was performed according to the MYbaits protocol with the following modifications: a total amount of 150 ng of baits per capture reaction was used, and library pools were incubated for hybridization with the baits at 60°C for 42 hr. Following capture, the beads were removed and the libraries were amplified in three replicates with Herculase II Fusion polymerase enzyme (Agilent) using P5/P7 bridge primers (Illumina) and 15 amplification cycles. The final library concentration was measured on the Agilent 2,200 TapeStation using a D1000 chip. PCR products were then purified using the MinElute PCR Purification Kit (Qiagen). For sequencing, the enriched capture products were pooled such that they were equimolar. The enriched library pool was sequenced on an Illumina MiSeq platform in a single MiSeq flow cell using v2 chemistry (2 × 250 bp paired-ends).

2.5 | Bioinformatic analyses

The raw sequence data were demultiplexed with bcl2fastq v. 2.20 (Illumina, Inc.), and adapter sequences were removed using CUTADAPT v. 1.15 (Martin, 2011) with an error rate of 0.15. The reads were then quality-trimmed with TRIMMOMATIC v. 0.32 (Bolger, Lohse, & Usadel, 2014) using a sliding window approach of 10 bp and a minimum PHRED score of 20. Forward and reverse paired reads (with read lengths between 100 and 500 bp) were merged using FLASH (Magoc & Salzberg, 2011) with a minimum overlap of 10 bp. To eliminate duplicate sequences, the data set was de-replicated using the VSEARCH (Rognes, Flouri, Nichols, Quince, & Mahé, 2016) filter algorithm fastx_uniques with default settings. The resulting sequences were mapped against the reference mitogenomes using BWA (Li & Durbin, 2009) and SAMtools v. 1.9 (Li et al., 2009); additionally, the sequences were blasted against the NCBI's nucleotide database using default parameters. Sample pairs (water and sediment) or triplets (water, sediment and glass fibre filters) were considered positive for a species when at least five sequences were obtained with a total coverage of at least 300 bp of the respective mitogenome. This coverage can be considered sufficient to obtain a reliable consensus. The entire workflow is shown in Figure 1.

2.6 | PCR-based approach

As a comparison with the hybridization capture results, a duplex PCR was performed using MyTaq™ DNA Polymerase (Bioline) and specific primers to amplify fragments of the mitochondrial D-loop of *L. africana* (376 bp; Eggert, Rasner, & Woodruff, 2002) and *Equus* sp. (590 bp; Seeber, Soilemetzidou, East, Walzer, & Greenwood, 2017), respectively. In each PCR, we used bovine serum albumin at a concentration of 0.6 µg/µl to improve yield. The cycling conditions were 95°C for 2 min, followed by 40 cycles of 95°C for 20 s, 60°C for 30 s and 72°C for 45 s, and a final step of 72°C for 2 min.

3 | RESULTS

A total of 13,309,849 reads were generated from environmental and mock samples. The 14 water and sediment samples produced 4,336,426 (mean per sample 289,095 ± 289,694) and 3,809,261 (mean per sample 253,951 ± 247,905) reads, respectively. The five glass fibre filters produced 2,746,613 reads (mean 549,323 ± 566,836), and the two mock samples produced 765,145 reads (mean 382,573 ± 53,299). A total of 1,652,404 reads could not be assigned due to missing indices, and two sample pairs (water and sediment) produced no reads.

Out of the 38 target species from which the bait set was designed, we identified 16 species in our samples (Table 2); overall, 13 species were identified from water samples, 10 species from sediment and nine species from water filters. Glass fibre filters generally produced higher numbers of sequences per sample and species (mean 74 ± 104) than water (mean 47 ± 130) or sediment samples (mean 30 ± 54; Table 2). For the two best-covered species, *Loxodonta africana* and *Hippopotamus amphibius*, a maximum mitogenome coverage of 87.9% and 99.8%, respectively, was observed. Fewer numbers of species-specific positives were obtained from five sample pairs (water and sediment) collected during the wet season in Tanzania (January–June; seven positives) than from the three sample pairs collected during the dry season (July/October; 21 positives). BLAST searches against the NCBI nucleotide database showed that the remaining sequences (about 10.9 million reads) most closely matched nontarget organisms such as other abundant eukaryotes (about 8 million reads) including algae, land plants and nontarget metazoa, but also matched species of other kingdoms such as bacteria (e.g. about 2.7 million reads), viruses and archaea.

The mitogenome sequences used in the mock experiments were not closely related to the species on which the bait set was based. The coverage of the mitogenomes of the three mock species by the used baits was predicted by mapping the entire bait set against the mitogenome of each species using the default settings of BWA; thereby, coverage values of 89%, 60% and 30% were predicted for *O. moschatus*, *U. maritimus* and *P. cinereus*, respectively (Figure 2). However, in the actual hybridization capture, coverage of 97%, 99% and 91% of the respective mitogenome was obtained. The total number of mapped sequences per species differed between the mock samples—in mock sample #1 with equal amounts of input DNA, a substantially higher number of sequences mapping to *O. moschatus*, were obtained than for the other two species (*U. maritimus* and *P. cinereus*; Table 3). In mock sample #2, the high concentration of *P. cinereus* DNA produced a somewhat higher number of sequences than in mock sample #1, and a 16-fold smaller amount of *U. maritiums* DNA produced a lower number of sequences than in mock sample #1.

TABLE 2 Number of sequences per water (W), sediment (S) or glass fibre filter (gff) sample in each geographical region mapping to the selected target species mitogenomes. Only samples that in total (water plus sediment plus glass fibre filter) generated ≥ 5 mapped sequences with a coverage of at least 300 bp assigned to the respective target mitogenomes are shown

		Namibia				*Tanzania*							
		Etosha National Park		*‖Karas region*		*Serengeti National Park*							
	month	*Nov*	*Nov*	*Apr*	*Apr*	*Jan*	*Feb*	*Feb*	*Jun*	*Jun*	*Jul*	*Oct*	*Oct*
	sample	**1**	**2**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Species	type												
Loxodonta africana	W	11	30								6	7	5
	S	78	77									7	
	gff											25	10
Myotis myotis	W												
	S												
	gff											290	12
Orycteropus afer	W												21
	S												
	gff											13	41
Equus quagga	W		10								5		12
	S		15								11		8
	gff									18		8	6
Equus zebra	W		5										
	S		9		12								
	gff												
Hippopotamus amphibius	W								630		100	9	
	S								252		6	38	
	gff												
	W		15										
	S		54										
Syncerus caffer	W										11		
	S										11	5	6
	gff												
Oryx gazella	W												
	S				8								
	gff												
Connochates taurinus	W					5							
	S												
	gff												
Crocuta crocuta	W					117							
	S				8								
	gff											5	
Panthera pardus	W												
	S												
	gff											240	9
Panthera leo	W							8				9	
	S		6		9							5	
	gff											150	
Acinonyx jubatus	W												5
	S												
	gff											276	
Canis aureus	W						18						
	S						9						
	gff												
Lycaon pictus	W											5	
	S												
	gff											7	

Note: Blank cells indicate < 5 sequences; grey shading indicates that no respective sample was collected. Underlined months indicate the respective dry season.

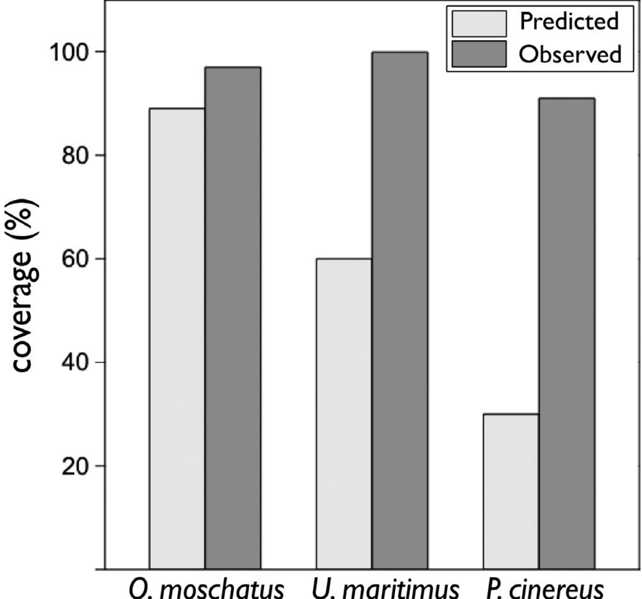

FIGURE 2 Coverage of baits (‘*predicted*’) and of sequencing sequences produced from the hybridization capture (‘*observed*’) relative to the mitogenomes of the respective species

The duplex PCR amplified the *L. africana* D-loop fragment in 3 out of 33 subsamples (water, sediment and filters), whereas 11 produced species-specific sequences using hybridization capture (Table 4). The equid D-loop fragment was amplified from 4 out of 33 subsamples, whereas 10 subsamples yielded *E. quagga* or *E. zebra* sequences after hybridization capture (Table 4). The remaining subsamples produced only unspecific products or failed to amplify.

4 | DISCUSSION

Metabarcoding relies on the occurrence of relatively intact DNA for PCR amplification, and the use of degenerate primers to cover multiple species may bias the results (Taberlet, Bonin, Zinger, & Coissac, 2018). Metabarcoding of terrestrial animals from eDNA has been performed using environmental sample material of animals in captivity (Harper et al., 2019; Rodgers & Mock, 2015), and to a lesser extent, also in the wild (Egeter et al., 2018; Ushio, Fukuda, et al., 2017). Under natural conditions, however, this approach is typically hampered by the presumably low amount of target DNA in any given environmental substrate compared to the high amount of background DNA. Also, long-term occupied zoo enclosures can be expected to contain a substantially higher amount of DNA of the inhabitants, compared to the respective natural environments. In water-limited ecosystems, however, water sources are an important resource for many species and are thus potential sources of eDNA.

Glass fibre filters, in general, produced more sequences per mammal species using hybridization capture than the corresponding water or sediment samples, which may partially be attributed to the larger volume of water that was filtered (300 ml) compared to the volume used for DNA isolation from native water (25 ml). However, this effect was not entirely consistent in the present study. For example, we found sequences of *Hippopotamus amphibious* in three water samples and their corresponding sediment samples, but not in the respective glass fibre filters. Due to the observed high variability between sample types and sampling sites, we would recommend collection of different sample types and multiple samples per site to increase detection rates. In previous studies on eDNA isolated from water samples, large volumes of water (e.g. 5 L; Wilcox et al., 2018) were used for filtration, and a study using a comparable volume of water (250 ml water for filtration) obtained variable results (Grey et al., 2018). It would thus seem advisable, based on our results and those of Grey et al. (2018) that a large volume of water be utilized. In the present study, filtration was performed manually using considerably turbid water, which lead to rapid clogging of the filters. In order to filter larger amounts of water (i.e.>300 ml), multiple filters per sampling site should be used, particularly when turbid water is to be sampled. Compared with eDNA isolation from native water, the use of glass fibre filtration is likely to be more convenient in a field setting than the collection, transport and storage of large volume water samples.

Those species which were abundant and may be expected to shed larger amounts of DNA due to their behaviour at water holes (such as full-body submergence in elephants and hippos) yielded the most on-target reads and produced the highest coverage of the mitogenome. In contrast, relatively few sequences were obtained from species that are rare and may be expected to shed minute amounts of DNA into the water, if any (e.g. few sequences of cheetahs and no sequences of rodents were obtained). Therefore, the success of eDNA capture from water may be limited for certain taxa, but it is a very promising approach for those species which can be expected to shed a sufficient amount of DNA into the water.

A higher number of sequences were produced from samples collected during the mid- and late dry season (June/July and October, respectively), whereas relatively few positive results were obtained from samples collected during the rainy season in Tanzania (February). This pattern is likely due to the dilution effect by frequent precipitation but may also reflect the increasing scarcity of

TABLE 3 Mock samples: amount of input DNA for library building and sequences of each species after hybridization capture

	Mock sample #1		Mock sample #2	
Species	DNA input (ng)	Sequences	DNA input (ng)	Sequences
Phascolarctos cinereus	250	33,162	1,500	52,239
Ovibos moschatus	250	114,390	150	123,100
Ursus maritimus	250	90,742	15	26,604

TABLE 4 PCR of specific mitochondrial D-loop fragments of *L. africana* and *Equus* sp. from water (W), sediment (S) and glass fibre filter (gff) samples. Shown are the numbers of specific sequences obtained from the hybridization capture, as in Table 2

target species	sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Loxodonta africana	W	*11*	*30*										6	7	5
	S	*78*	*77*				*6*							7	
	gff													25	10
Equus sp.	W		*15*										*5*		*12*
	S		*24*				*12*						*11*		*8*
	gff											*18*		*8*	*6*

Note: Cells with outside borders indicate positive PCR results. Blank cells indicate < 5 sequences per sample and less than 300 bp coverage; grey shading indicates that no respective sample was collected.

available water bodies over the dry season and higher numbers of congregating animals. Despite the drought at the time of sampling in Namibia, the respective water samples produced a relatively low number of mammal sequences. Most of the sampled water holes in Namibia were pumped and thus constantly refilled, which may be why mammalian DNA was not as concentrated as in natural water bodies that typically shrink over the course of a dry season.

The results of the current study likely reflect relatively recent shedding of DNA into the respective water source, as environmental DNA is typically dilute and once shed into the environment, DNA is exposed to harsh conditions such as microbial metabolism and extensive UV radiation, leading to fragmentation, hydrolytic damage (i.e. cytosine deamination) and oxidation (Hawkins et al., 2016); thus, degradation of eDNA can be a limiting factor in such analyses (Pilliod et al., 2014). However, a recent study demonstrated that UV radiation does not negatively affect eDNA-based detection rates from water samples (Mächler, Osathanunkul, & Altermatt, 2018). Under optimal conditions, the stability of DNA in the environment facilitates the tracing of mammals in environmental samples for a considerable amount of time which may produce a bias in the accuracy of results (Dejean et al., 2011); however, in natural water bodies in the tropics, eDNA degrades within a comparably short amount of time, that is within a few days (Eichmiller, Best, & Sorensen, 2016)..

The mock sample experiment demonstrated that successful hybridization capture is possible despite considerable divergence between baits and target species. The baits were designed exclusively from eutherian mammals, and although only 30% of the mitogenome of the marsupial *P. cinereus* was predicted to be covered by the baits set, the hybridization capture produced sequences covering about 91% of the mitogenome. It has been previously shown that the amount of captured DNA reflects initial target DNA abundance (Wilcox et al., 2018). In the present study, we found that mock samples of known DNA concentrations produced results which were not entirely in line with the DNA input concentration, which is likely a result of divergence between baits and target species. Using different concentrations of input DNA, the sequences of *O. moschatus* outnumbered those of the other two species, even though, in mock experiment #2, the amount of input DNA of this species was ten times lower than that of *P. cinereus*. No ursid species were used for bait design, and the predicted coverage based on homology for *U. maritimus* by the baits was 60%. However, the captured sequences produced a coverage of 99% and the number of sequences decreased by only 30% between the two mock samples despite a 20-fold reduction in DNA input. Our results confirm successful hybridization capture for considerably divergent species and demonstrate that despite over-representation of DNA of one species, the DNA of other distantly related species at a 100-fold lower concentration can still be detected. In the 33 environmental sample libraries, a total of 1, 1 and 2 sequences mapped to the species *O. moschatus*, *P. cinereus* and *U. maritimus*, used in the mock samples, respectively; therefore, cross-contamination and/or misindexing was minimal in this study. Given these findings, we argue that the chosen species identification threshold of 300 bp coverage and a minimum of five sequences ensures that contamination or misindexing did not influence our results.

Compared to hybridization capture, conventional PCR approaches to amplify D-loop fragments of two abundant genera were less successful (three and four positives by PCR compared to 11 and 10 positives by hybridization capture). This effect is likely due to DNA degradation and unspecific primer binding which may have prevented amplification of the respective 376–590 bp fragments in several eDNA samples. Although for PCR-based eDNA studies typically short fragments of about 100 bp are targeted to account for DNA degradation, longer fragments have been successfully amplified (e.g. 658 bp; Deiner et al., 2016).

Hybridization capture of eDNA from water, sediment or water filters can target multiple species and is thus a promising method for biodiversity monitoring by generating data on the presence of certain types of mammalian species. In order to produce more comprehensive survey results, multiple samples per site should be collected to account for patchy distribution of eDNA in water and sediment. As the cost of analysis per sample may be a limiting factor, hybridization

capture allows for the pooling of multiple genomic libraries per capture reaction which reduces the per sample costs, among other cost-saving measures (Förster et al., 2018). Using an amount of 100–150 ng baits per reaction (Hawkins et al., 2016; present study) and pooling up to four libraries per capture reaction can limit the cost for baits to about 15 US dollars per library. Although this additional cost can be omitted when using a metabarcoding approach, hybridization capture can be used to recover large portions of the entire mitogenome of the target organism, which in the case of fragmented DNA may be superior to PCR-based methods which depend on the integrity of the respective target fragment. The selection of target sequences of target species and subsequent bait design are key issues for increasing the efficiency and specificity of the hybridization capture. However, hybridization capture clearly tolerates high levels of mismatch between baits and targets as observed in the mock DNA experiments performed in the current study.

For broad surveys of multiple species or taxa, conserved sites that yield little information may be omitted as bait templates in order to increase the coverage and potential sequencing depth of variable sites and thereby improve taxonomic resolution. However, captured DNA fragments can substantially exceed the length of baits and, furthermore, anchored hybrid enrichment can produce sequences that are not complementary to the RNA bait, but that hybridize to the unbound portion of a captured target DNA molecule (Lemmon, Emme, & Lemmon, 2012; Tsangaras et al., 2014). Therefore, even baits derived from conserved sites may increase the overall coverage of informative sites.

Hybridization capture of eDNA from water sources is a promising tool to assess the presence of terrestrial mammals in water-limited ecosystems. Divergent baits may broaden the efficiency of the process. However, divergence between baits and target sequences may result in a high amount of nontarget reads and may decrease the overall efficiency of the approach. It is therefore likely that baits which are highly specific to the target taxa will improve the overall capture efficiency for eDNA of common and rare species.

ACKNOWLEDGEMENTS

This study was supported by a grant from the Leibniz Gemeinschaft (SAW-2015-IZW-1 440) and the Leibniz Institute for Zoo and Wildlife Research. We thank the Namibian Ministry of Environment and Tourism, the Tanzania Commission of Science and Technology, Tanzania National Parks and the Tanzania Wildlife Research Institute for their support. We extend our gratitude to Ruben Portas, Dirk Bockmühl, Sonja Metzger, Heribert Hofer and Annie Francis for assistance at the field sites and to Niccoló Alfano, Jan Axtner, Anisha Dayaram and John Galindo Puentes for assistance in the laboratory.

AUTHOR CONTRIBUTIONS

Environmental samples were collected by PAS and JM. Laboratory procedures were performed by PAS. Sequence data were analysed by PAS, GKM and UL. ADG initiated the study. All authors contributed to study design and writing the manuscript.

ETHICAL APPROVAL

Permission to conduct research in Namibia was granted by the Ministry of Environment and Tourism (permit No. 2094/2016), and permission to conduct research in Tanzania was granted by the Tanzania Commission for Science and Technology (permit No. 2015-168-NA-90-130) and the Tanzanian Wildlife Research Institute. All protocols adhered to the laws and guidelines of Namibia, Tanzania and Germany, respectively.

DATA AVAILABILITY STATEMENT

Sequences of the bait set and alignments were deposited in the Mendeley Data repository (https://doi.org/10.17632/zsd3w5vttk.1). Raw sequence data were made available as an NCBI BioProject under the accession number PRJNA515605.

ORCID

Peter Andreas Seeber https://orcid.org/0000-0003-4401-8248
Alex D. Greenwood https://orcid.org/0000-0002-8249-1565

REFERENCES

Aylagas, E., Borja, Á., Irigoien, X., & Rodríguez-Ezpeleta, N. (2016). Benchmarking DNA Metabarcoding for biodiversity-based monitoring and assessment. *Frontiers in Marine Science*, *3*(June), 96. https://doi.org/10.3389/fmars.2016.00096

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Civade, R., Dejean, T., Valentini, A., Roset, N., Raymond, J.-C., Bonin, A., ... Pont, D. (2016). Spatial representativeness of environmental DNA Metabarcoding signal for fish biodiversity assessment in a natural freshwater system. *PLoS ONE*, *11*(6), e0157366. https://doi.org/10.1371/journal.pone.0157366

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., ... Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. https://doi.org/10.1111/mec.14350

Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J. C., & Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nature Communications*, *7*, 12544. https://doi.org/10.1038/ncomms12544

Deiner, K., Renshaw, M. A., Li, Y., Olds, B. P., Lodge, D. M., & Pfrender, M. E. (2017). Long-range PCR allows sequencing of mitochondrial genomes from environmental DNA. *Methods in Ecology and Evolution*, *8*(12), 1888–1898. https://doi.org/10.1111/2041-210X.12836

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., & Miaud, C. (2011). Persistence of environmental DNA in freshwater ecosystems. *PLoS ONE*, *6*(8), e23398. https://doi.org/10.1371/journal.pone.0023398

Dougherty, M. M., Larson, E. R., Renshaw, M. A., Gantz, C. A., Egan, S. P., Erickson, D. M., & Lodge, D. M. (2016). Environmental DNA

(eDNA) detects the invasive rusty crayfish Orconectes rusticus at low abundances. *Journal of Applied Ecology*, *53*(3), 722–732. https://doi.org/10.1111/1365-2664.12621

Dowle, E. J., Pochon, X., C. Banks, J., Shearer, K., & Wood, S. A. (2016). Targeted gene enrichment and high-throughput sequencing for environmental biomonitoring: A case study using freshwater macroinvertebrates. *Molecular Ecology Resources*, *16*(5), 1240–1254. https://doi.org/10.1111/1755-0998.12488

Egeter, B., Peixoto, S., Brito, J. C., Jarman, S., Puppo, P., & Velo-Antón, G. (2018). Challenges for assessing vertebrate diversity in turbid Saharan water-bodies using environmental DNA. *Genome*, *61*(11), 807–814. https://doi.org/10.1139/gen-2018-0071

Eggert, L. S., Rasner, C. A., & Woodruff, D. S. (2002). The evolution and phylogeography of the African elephant inferred from mitochondrial DNA sequence and nuclear microsatellite markers. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1504), 1993–2006. https://doi.org/10.1098/rspb.2002.2070

Eichmiller, J. J., Best, S. E., & Sorensen, P. W. (2016). Effects of temperature and trophic state on degradation of environmental DNA in lake water. *Environmental Science and Technology*, *50*(4), 1859–1867. https://doi.org/10.1021/acs.est.5b05672

Förster, D. W., Bull, J. K., Lenz, D., Autenrieth, M., Paijmans, J. L. A., Kraus, R. H. S., … Fickel, J. (2018). Targeted resequencing of coding DNA sequences for SNP discovery in nonmodel species. *Molecular Ecology Resources*, *18*(6), 1356–1373. https://doi.org/10.1111/1755-0998.12924

Gasc, C., Peyretaillade, E., & Peyret, P. (2016). Sequence capture by hybridization to explore modern and ancient genomic diversity in model and nonmodel organisms. *Nucleic Acids Research*, *44*(10), 4504–4518. https://doi.org/10.1093/nar/gkw309

Grey, E. K., Bernatchez, L., Cassey, P., Deiner, K., Deveney, M., Howland, K. L., … Lodge, D. M. (2018). Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. *Scientific Reports*, *8*(1), 8843. https://doi.org/10.1038/s41598-018-27048-2

Hänfling, B., Lawson Handley, L., Read, D. S., Hahn, C., Li, J., Nichols, P., … Winfield, I. J. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, *25*(13), 3101–3119, https://doi.org/10.1111/mec.13660

Harper, L. R., Handley, L. L., Carpenter, A. I., Murray-Dickson, G., Muri, C. D., Macgregor, C. J., … Hänfling, B. (2019). Environmental DNA (eDNA) metabarcoding of pond water as a tool to survey conservation and management priority mammals. *BioRxiv*, *546218*, https://doi.org/10.1101/546218

Hawkins, M. T. R., Hofman, C. A., Callicrate, T., McDonough, M. M., Tsuchiya, M. T. N., Gutiérrez, E. E., … Maldonado, J. E. (2016). In-solution hybridization for mammalian mitogenome enrichment: Pros, cons and challenges associated with multiplexing degraded DNA. *Molecular Ecology Resources*, *16*(5), 1173–1188. https://doi.org/10.1111/1755-0998.12448

Hunter, M. E., Oyler-McCance, S. J., Dorazio, R. M., Fike, J. A., Smith, B. J., Hunter, C. T., … Hart, K. M. (2015). Environmental DNA (eDNA) sampling improves occurrence and detection estimates of invasive Burmese Pythons. *PLoS ONE*, *10*(4), e0121655. https://doi.org/10.1371/journal.pone.0121655

Lemmon, A. R., Emme, S. A., & Lemmon, E. M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, *61*(5), 727–744. https://doi.org/10.1093/sysbio/sys049

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., & Homer, N. … 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Mächler, E., Deiner, K., Steinmann, P., & Altermatt, F. (2014). Utility of environmental DNA for monitoring rare and indicator macroinvertebrate species. *Freshwater Science*, *33*(4), 1174–1183. https://doi.org/10.1086/678128

Mächler, E., Osathanunkul, M., & Altermatt, F. (2018). Shedding light on eDNA: Neither natural levels of UV radiation nor the presence of a filter feeder affect eDNA-based detection of aquatic organisms. *PLoS ONE*, *13*(4), 1–15. https://doi.org/10.1371/journal.pone.0195529

Magoc, T., & Salzberg, S. L. (2011). FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, *27*(21), 2957–2963. https://doi.org/10.1093/bioinformatics/btr507

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, *2010*(6), pdb.prot5448–pdb.prot5448. https://doi.org/10.1101/pdb.prot5448

Natural Museum of History, S. I.(2005). *Mammal species of the world* (E. Wilson & D. Reeder, Eds.). Retrieved from https://www.departments.bucknell.edu/biology/resources/msw3/

Olds, B. P., Jerde, C. L., Renshaw, M. A., Li, Y., Evans, N. T., Turner, C. R., … Lamberti, G. A. (2016). Estimating species richness using environmental DNA. *Ecology and Evolution*, *6*(12), 4214–4226. https://doi.org/10.1002/ece3.2186

Paijmans, J. L. A., Fickel, J., Courtiol, A., Hofreiter, M., & Förster, D. W. (2016). Impact of enrichment conditions on cross-species capture of fresh and degraded DNA. *Molecular Ecology Resources*, *16*(1), 42–55. https://doi.org/10.1111/1755-0998.12420

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., & Waits, L. P. (2014). Factors influencing detection of eDNA from a stream-dwelling amphibian. *Molecular Ecology Resources*, *14*(1), 109–116. https://doi.org/10.1111/1755-0998.12159

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R. M., & Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology*, *51*(5), 1450–1459. https://doi.org/10.1111/1365-2664.12306

Rodgers, T. W., & Mock, K. E. (2015). Drinking water as a source of environmental DNA for the detection of terrestrial wildlife species. *Conservation Genetics Resources*, *7*(3), 693–696. https://doi.org/10.1007/s12686-015-0478-7

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *4*, e2584. https://doi.org/10.7717/peerj.2584

Seeber, P. A., Dayaram, A., Sicks, F., Osterrieder, N., Franz, M., & Greenwood, A. D. (2019). Noninvasive Detection of Equid Herpesviruses in Fecal Samples. *Appl Environ Microbiol*, *85*(3), e02234–18. https://doi.org/10.1128/AEM.02234-18

Seeber, P. A., Soilemetzidou, S. E., East, M. L., Walzer, C., & Greenwood, A. D. (2017). Equine behavioral enrichment toys as tools for non-invasive recovery of viral and host DNA. *Zoo Biology*, *36*(5), 341–344. https://doi.org/10.1002/zoo.21380

Shokralla, S., Gibson, J. F., King, I., Baird, D. J., Janzen, D. H., Hallwachs, W., & Hajibabaei, M. (2016). Environmental DNA barcode sequence capture: Targeted, PCR-free Sequence capture for biodiversity analysis from bulk environmental samples. *BioRxiv*, 1–28, https://doi.org/10.1101/087437

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA: For Biodiversity Research and Monitoring*. Oxford Scholarship Online, ISBN-13: 9780198767220. https://doi.org/10.1093/oso/9780198767220.001.0001

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA - An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, *183*, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019

Tsangaras, K., Wales, N., Sicheritz-Pontén, T., Rasmussen, S., Michaux, J., Ishida, Y., … Greenwood, A. D. (2014). Hybridization capture using short PCR products enriches small genomes by capturing flanking sequences (CapFlank). *PLoS ONE*, *9*(10), e109101. https://doi.org/10.1371/journal.pone.0109101

Turner, C. R., Barnes, M. A., Xu, C. C. Y., Jones, S. E., Jerde, C. L., & Lodge, D. M. (2014). Particle size distribution and optimal capture of aqueous macrobial eDNA. *Methods in Ecology and Evolution*, *5*(7), 676–684. https://doi.org/10.1111/2041-210X.12206

Ushio, M., Fukuda, H., Inoue, T., Makoto, K., Kishida, O., Sato, K., … Miya, M. (2017). Environmental DNA enables detection of terrestrial mammals from forest pond water. *Molecular Ecology Resources*, *17*(6), e63–e75. https://doi.org/10.1111/1755-0998.12690

Ushio, M., Murata, K., Sado, T., Nishiumi, I., Takeshita, M., Iwasaki, W., & Miya, M. (2017). Demonstration of the potential of environmental DNA as a tool for the detection of avian species. *Scientific Reports*, *8*(1), 199208. https://doi.org/10.1038/s41598-018-22817-5

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., … Dejean, T. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, *25*(4), 929–942. https://doi.org/10.1111/mec.13428

Wilcox, T. M., Zarn, K. E., Piggott, M. P., Young, M. K., McKelvey, K. S., & Schwartz, M. K. (2018). Capture enrichment of aquatic environmental DNA: A first proof of concept. *Molecular Ecology Resources*, *June*, 1–10. https://doi.org/10.1111/1755-0998.12928

How to cite this article: Seeber PA, McEwen GK, Löber U, et al. Terrestrial mammal surveillance using hybridization capture of environmental DNA from African waterholes. *Mol Ecol Resour.* 2019;19:1486–1496. https://doi.org/10.1111/1755-0998.13069