Fishing for mammals: landscape-level monitoring of terrestrial and semi-aquatic communities using eDNA from lotic ecosystems

Naiara Guimarães Sales[1]*, Maisie B. McKenzie[1]*, Joseph Drake[2]*, Lynsey R. Harper[3], Samuel S. Browett[1], Ilaria Coscia[1], Owen S. Wangensteen[4], Charles Baillie[1], Emma Bryce[5], Deborah A. Dawson[6], Erinma Ochu[1], Bernd Hänfling[3], Lori Lawson Handley[3], Stefano Mariani[1,7], Xavier Lambin[5], Christopher Sutherland[2] and Allan D. McDevitt[1#]

[1]Environment and Ecosystem Research Centre, School of Science, Engineering and Environment, University of Salford, Salford, UK

[2]Department of Environmental Conservation, University of Massachusetts-Amherst, Amherst, USA

[3]Department of Biological and Marine Sciences, University of Hull, Kingston upon Hull, UK

[4]Norwegian College of Fishery Science, University of Tromsø, Tromsø, Norway

[5]School of Biological Sciences, University of Aberdeen, Aberdeen, UK

[6]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK

[7]School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool, UK

*These authors contributed equally to this work

#Corresponding author: a.mcdevitt@salford.ac.uk

Environment and Ecosystem Research Centre, School of Science, Engineering and Environment, University of Salford, Salford, UK

Abstract

1. Environmental DNA (eDNA) metabarcoding has revolutionised biomonitoring in both marine and freshwater ecosystems. However, for semi-aquatic and terrestrial animals, the application of this technique remains relatively untested.

2. We first assess the efficiency of eDNA metabarcoding in detecting semi-aquatic and terrestrial mammals in natural lotic ecosystems in the UK by comparing sequence data recovered from water and sediment samples to the mammalian communities expected from historical data. Secondly, using occupancy modelling we compared the detection efficiency of eDNA metabarcoding to multiple conventional non-invasive survey methods (latrine surveys and camera trapping).

3. eDNA metabarcoding detected a large proportion of the expected mammalian community within each area. Common species in the areas were detected at the majority of sites. Several key species of conservation concern in the UK were detected by eDNA sampling in areas where authenticated records do not currently exist, but potential false positives were also identified.

4. Water-based eDNA metabarcoding provided comparable results to conventional survey methods in per unit of survey effort for three species (water vole, field vole, and red deer) using occupancy models. The comparison between survey 'effort' to reach a detection probability of ≥0.95 revealed that 3-6 water replicates would be equivalent to 3-5 latrine surveys and 5-30 weeks of single camera deployment, depending on the species.

5. *Synthesis and Applications.* eDNA metabarcoding can be used to generate an initial 'distribution map' of mammalian diversity at the landscape level. If conducted during times of peak abundance, carefully chosen sampling points along multiple river

courses provide a reliable snapshot of the species that are present in a catchment area. In order to fully capture solitary, rare and invasive species, we would currently recommend the use of eDNA metabarcoding alongside other non-invasive surveying methods (i.e. camera traps) to maximize monitoring efforts.

Keywords: biomonitoring, camera trapping, eDNA metabarcoding, latrine surveys, mammals, occupancy modelling, rivers

Introduction

Environmental DNA (eDNA) metabarcoding (the simultaneous identification of multiple taxa using DNA extracted from an environmental sample, e.g. water, soil, based on short amplicon sequences) has revolutionised the way we approach biodiversity monitoring in both marine and freshwater ecosystems (Valentini et al., 2016; Deiner et al. 2017). Successful applications include tracking biological invasions, detecting rare and endangered species and describing entire communities (Holman et al., 2019). Most eDNA metabarcoding applications on vertebrates to date have focused on monitoring fishes and amphibians (Hänfling et al., 2016; Valentini et al., 2016). What has become apparent from studies in lentic systems (ponds and lakes) is that semi-aquatic and terrestrial mammals can also be detected (Hänfling et al., 2016; Harper et al., 2019). As a result, there has been an increasing focus on the use of both vertebrate (Harper et al., 2019) and mammal-specific primer sets (Ushio et al., 2017; Leempoel et al., 2019; Sales et al., 2019) for detecting mammalian communities using eDNA metabarcoding.

Mammals include some of the most imperiled taxa, with over one fifth of species considered to be threatened or declining (Visconti et al., 2011). Monitoring of mammalian biodiversity is therefore essential. Given that any optimal survey approach is likely to be species-specific, very few species can be detected at all times when they are present. This imperfect detection (even greater for elusive and rare species) can lead to biased estimates of occurrence and hinder species conservation (Mackenzie et al., 2002). For mammals, repeated surveys using several monitoring methods are usually applied. These include indirect observations such as latrines, faeces, hair, or tracks, or direct observations such as live-trapping or camera trapping surveys over short time intervals such that closure/invariance can be assumed and detectability

estimated (Nichols et al., 2008). Each of these methods has associated efficiency, cost and required expertise trade-offs, which become more challenging as the spatial and temporal scales increase.

eDNA sampling yields species-specific presence/absence data that are likely to be most valuable for inferring species distributions using well established analytical tools such as occupancy models (MacKenzie et al., 2002). These models resolve concerns around imperfect detection of difficult to observe species. When coupled with location-specific detection histories, these can be used to infer true occurrence states, factors that influence occupancy rates, colonization-extinction probabilities, and estimates of detection probability (MacKenzie et al., 2017). The use of eDNA sampling to generate species-specific detection data has unsurprisingly increased in recent years, and in many cases has outperformed or at least matched conventional survey methods (Lugg et al., 2018; Tingley et al., 2019). Although comparisons between eDNA analysis and conventional surveys for multi-species detection are numerous (see Table S1 in Lugg et al., 2018), studies focusing on detection probability estimates for multiple species identified by metabarcoding are rare (Abrams et al., 2019; Valentini et al., 2016).

The aim of this study was to assess the efficiency of eDNA metabarcoding for detecting semi-aquatic and terrestrial mammals in natural lotic systems in the UK. We conducted eDNA sampling in rivers and streams in two areas (Assynt, Scotland and Peak District National Park, England). Together these locations have the majority of UK semi-aquatic and terrestrial mammalian species present (Table S1). Our objectives were two-fold: first, we sought to establish whether eDNA metabarcoding is a viable technique for monitoring semi-aquatic and terrestrial mammals by comparing it to the mammalian communities expected from historical data, a group for which

106 eDNA sampling has rarely been evaluated in a natural setting. Secondly, we evaluate
107 the detection efficiency of water- and sediment-based eDNA sampling in one of these
108 areas (Assynt) for multiple species compared to multiple conventional non-invasive
109 survey methods (latrine surveys and camera trapping).

Material and Methods

Latrine surveys

Assynt, a heather-dominated upland landscape in the far northwest of the Scottish Highlands, UK (Fig. 1A), is the location of an ongoing 20-year metapopulation study of water voles (*Arvicola amphibius*) led by the University of Aberdeen (Fig. S1). Here, we mainly focus only on data collected in 2017. The metapopulation is characterized by 116 discrete linear riparian habitat patches (ranging from 90 m to nearly 2.5 km) distributed sparsely (4% of waterway network) throughout the 140 km^2 study area (Sutherland et al., 2014). Water voles use prominently placed latrines for territory marking (Fig. S2A). Using latrine surveys, a reliable method of detection (Sutherland et al., 2014), water vole occupancy status was determined by the detection of latrines that are used for territory marking (Sutherland et al., 2013). During the breeding season (July and August), latrine surveys were conducted twice at each site. In addition to water vole latrines, field vole (*Microtus agrestis*) pellets are also easily identifiable, and so field vole detections were also recorded along waterways as a formal part of the latrine survey protocol. Live-trapping was then carried out at patches deemed to be occupied by water voles according to latrine surveys to determine their abundances (this was used to determine which sites were sampled for eDNA; Fig. 1A).

Camera Trap Data

Camera traps were deployed at the beginning of July and thus overlapped temporally with the latrine survey in Assynt. Data were collected from cameras deployed at seven of these patches. Within each of these patches, cameras were deployed at the midpoint of the areas where active signs (latrines, grass clipping, burrows) were

detected, and if no signs were detected, at the midpoint of historical water vole activity (J. Drake, C. Sutherland and X. Lambin, *pers. comm*.). These will also capture images of any species present in the area that come within close proximity of the camera (Fig. S3A-F).

Cameras were deployed approximately 1 m above ground on iron 'u-posts' to avoid flooding, prevent knock-down by wind/wildlife, and optimize both depth of field and image clarity. Cameras (Bushnell HD Trophy Cam, Overland Park, KA) were set at normal detection sensitivity (to reduce false-triggers from grass/shadows), low night time LED intensity (to prevent image white out in near depth of field), three shot burst (to increase chance of capturing small, fast moving bodies), and 15 min intervals between bursts (to increase temporal independence of captures and decrease memory burden). The area each camera photographed was approximately 1-2 m^2. Animals were identified on images and information was stored as metadata tags using the R (R Core Team, 2018) package camtrapR following the procedures described in Niedballa et al. (2018). Independence between detections was based on 60-minute intervals between species-specific detections.

eDNA sampling

A total of 18 potential water vole patches were selected for eDNA sampling in Assynt from 25-27th October 2017. The time lag between the latrine/live-trapping and eDNA surveys was because of two main reasons: (i) legitimate concerns around cross-site DNA contamination during latrine/live-trapping where researchers moved on a daily basis between sites as well as regularly handled and processed live animals (for decontamination procedures see the Supplementary Material) and (ii) the selection of eDNA sampling sites was based on the latrine surveys and abundance data provided

by live-trapping so could only occur after this was completed by August 6th. Water and sediment samples were collected from patches where water voles were determined to be absent (five sites; A1-A5); with 1-2 individuals present (three sites; A9, A16 and 18); 3-5 individuals (five sites; A6, A8, A11, A14 and A17); and 7-11 individuals (five sites; A7, A10, A12, A13 and A15; Fig. 1A). Each of these streams/rivers differed in their characteristics (in terms of width, depth and flow) and a representation of the sites is depicted in Fig. S4A-D. Three water (two litres each) and three sediment (~25mL) replicates were taken at each patch (further details of sample collection are provided in Appendix S1).

In addition to Assynt, eDNA sampling was also conducted on a smaller scale in the Peak District National Park, England (Fig. S5) to incorporate additional mammals that are not known to be present in Assynt (Table S1). Here, the occurrence of water vole was identified by the presence of latrines in two sites (P1 and P2) at the time of eDNA sampling (Fig. S2A), whilst no latrines were identified at one site (P3). At site P1, an otter (*Lutra lutra*) spraint was identified at the time of eDNA sampling (Fig. S2B). These three sites were sampled in March 2018 using the same methodology as in Assynt but were taken in close proximity (<50cm) to water vole latrines where present (Fig. S2A).

eDNA Laboratory Methods

DNA was extracted from the sediment samples using the DNeasy PowerMax Soil kit and from the water samples using the DNeasy PowerWater Kit (both QIAGEN Ltd.) following the manufacturer's instructions in a dedicated eDNA laboratory in the University of Salford. In order to avoid the risk of contamination during this step, DNA extraction was conducted in increasing order of expected abundance of water voles in

the eDNA samples (all field blanks were extracted first, followed by the sites with supposedly zero water vole abundance, up to the highest densities last). Along with field blanks (Assynt = 8, Peak District = 2), six lab extraction blanks were included (one at the end of each daily block of extractions). A decontamination stage using a Phileas 25 Airborne Disinfection Unit (Devea SAS) was undertaken before processing samples from different locations. Additional information regarding decontamination measures and negative controls can be found in the Supplementary Material.

A complete description of PCR conditions, library preparation and bioinformatic analyses are provided in Appendix S1. Briefly, eDNA was amplified using the MiMammal 12S primer set (MiMammal-U-F, 5′- GGGTTGGTAAATTTCGTGCCAGC-3′; MiMammal-U-R, 5′- CATAGTGGGGTATCTAATCCCAGTTTG-3′; Ushio et al., 2017) targeting a ˜170bp amplicon from a variable region of the 12S rRNA mitochondrial gene. A total of 147 samples, including field collection blanks (10) and laboratory negative controls (12, including six DNA extractions blanks and six PCR negative controls), were sequenced in two multiplexed Illumina MiSeq runs. To minimize bias in individual reactions, PCRs were replicated three times for each sample and subsequently pooled. Illumina libraries were built using a NextFlex PCR-free library preparation kit according to the manufacturer's protocols (Bioo Scientific) and pooled in equimolar concentrations along with 1% PhiX (v3, Illumina). The libraries were run at a final molarity of 9pM on an Illumina MiSeq platform using the 2 x 150bp v2 chemistry.

Bioinformatic analysis were conducted using OBITools metabarcoding package (Boyer et al., 2016) and the taxonomic assignment was conducted using ecotag against a custom reference database (see Appendix 1). To exclude MOTUs/reads putatively belonging to sequencing errors or contamination, the final dataset included

only MOTUs that could be identified to species level (>98%), and MOTUs containing less than 10 reads and with a similarity to a sequence in the reference database lower than 98% were discarded (Cilleros et al., 2019). The maximum number of reads detected in the controls for each MOTU in each sequencing run were removed from all samples (Table S7). For water voles, field voles and red deer (the most abundant wild mammals in terms of sequence reads in our dataset), this equated to a sequence frequency threshold of ≤0.17%, within the bounds of previous studies on removing sequences to account for contamination and tag jumping (Cilleros et al., 2019; Hänfling et al., 2016; Schnell, Bohmann, & Gilbert, 2015).

Occupancy/Detection Analysis in Assynt

The data collection from the different survey types described above (water-based eDNA, sediment-based eDNA, latrine and camera traps) produced the following site-specific detection/non-detection data:

(a) Latrine: two latrine surveys at 116 patches.
(b) w-eDNA: three water-based eDNA samples at 18 of the 116 patches surveyed.
(c) s-eDNA: three sediment-based eDNA samples at 18 of the 116 patches surveyed.
(d) Camera: six one-week occasions of camera trapping data at seven of the 18 patches surveyed by both Latrine and eDNA (w-eDNA + s-eDNA) surveys.

We chose to focus on three species that were detected by at least three of the four methods: water voles, field voles and red deer (*Cervus elaphus*). Water voles and field voles were recorded using all four survey methods and had detection histories for 14 surveying events ((Latrine × 2) + (w-eDNA × 3) + (s-eDNA × 3) + (Camera × 6)).

Red deer were not recorded during latrine surveys and had detection histories for 12 surveying events ((w-eDNA × 3) + (s-eDNA × 3) + (Camera × 6)). To demonstrate the relative efficacy of the four surveying methods, we restricted the analyses to the 18 sites where both latrine surveys were conducted and eDNA samples were taken, seven of which had associated camera trapping data. Although each surveying method differs in terms of effort and effective area surveyed, each are viable surveying methods that are readily applied in practice. A unit of survey effort here is defined as one latrine survey, one w-eDNA replicate, one s-eDNA replicate or one week of camera trapping. So, while the specific units of effort are not directly comparable, the relative detection efficacy per surveying method-specific unit of effort is of interest and will provide important context for designing future monitoring studies and understanding the relative merits of each surveying method. Analyzing the data using occupancy models allowing for method-specific detectability enables such a comparison in per unit effort efficacy between eDNA metabarcoding and multiple conventional survey methods.

A single season occupancy model (MacKenzie et al., 2002) was applied to the ensemble data where detection histories were constructed using each of the surveying events as sampling occasions (MacKenzie et al., 2017). The core assumption here is that the underlying occupancy state (i.e. occupied or empty) is constant over the sampling period, and therefore, every sampling occasion is a potentially imperfect observation of the true occupancy status. Because occasions represent method-specific surveying events, we used “surveying method” as an occasion-specific covariate on detection (Latrine, w-eDNA, s-eDNA and Camera). Our primary objective was to quantify and compare method-specific detectability, so we did not consider any

other competing models. For comparing the methods, we compute accumulation curves as (MacKenzie & Royle, 2005):

$$p^*_{smk} = 1 - (1 - \hat{p}_{sm})^k$$

Where p^*_{smk} is the cumulative probability of detecting species *s*, when species *s* is present, using method *m* after *k* surveying events based on the estimated surveying method-specific detection probability for each species ($\hat{p}_{sm}$). We vary k from 1 to a large number and find the value of *k* that results $p^*_{smk} \geq 0.95$. We conducted the same analysis separately for water voles, field voles, and red deer. Analysis was conducted in R (R Core Team, 2018) using the package unmarked (Fiske & Chandler, 2011).

Results

Mammal Detection via eDNA metabarcoding

The two sequencing runs generated 23,276,596 raw sequence reads and a total of 15,463,404 sequences remained following trimming, merging, and length filtering. After bioinformatic analysis, the final 'filtered' dataset contained 23 mammals (Tables S2 and S3). For mammals, ~12 million reads were retained after applying all quality filtering steps (see Appendix 1). Reads from humans, cattle (*Bos taurus*), pig (*Sus scrofa*), horse (*Equus ferus*), sheep (*Ovis aries*) and dog (*Canis lupus familiaris*), were not considered further as the focus of this study was on wild mammals (Table S4). *Felis* was included because of the potential of it being wildcat (*Felis silvestris*) or domestic cat (*F. catus*)/wildcat hybrids. A final dataset comprising ~5.9 million reads was used for the downstream analyses (Table S4).

In Assynt, the wild species identified were the red deer (18/18 sites); water vole (15/18); field vole (13/18); wood mouse (*Apodemus sylvaticus* - 9/18); pygmy shrew (*Sorex minutus* - 4/18); wild/domestic cat (*Felis* spp. - 4/18); mountain hare (*Lepus timidus* - 4/18); rabbit (*Oryctolagus cuniculus* - 3/18); water shrew (*Neomys fodiens* - 3/18); common shrew (*Sorex araneus* - 2/18); edible dormouse (*Glis glis* - 2/18); grey squirrel (*Sciurus carolinensis* - 1/18); pine marten (*Martes martes* - 1/18); brown rat (*Rattus norvegicus* - 1/18); red fox (*Vulpes vulpes* - 1/18) and badger (*Meles meles* - 1/18; Fig. 1B). All of these species are distributed around/within Assynt (Table S1), with the exception of the edible dormouse and the grey squirrel. These are unequivocally absent from the region. The edible dormouse is only present in southern England and the grey squirrel is not distributed that far north in Scotland (Mathews et al., 2018).

Of the wild mammals in the Peak District, the water vole, field vole, wood mouse and otter were found in two sites (P1 and P2). The red deer, pygmy shrew, common shrew, water shrew, red squirrel (*Sciurus vulgaris*), grey squirrel, pine marten and badger were each found at a single site (Fig. S5). Only rabbit was found in site P3. All species identified are currently distributed within the Park (Table S1), except the red squirrel and pine marten. The pine marten, which is critically endangered in England, has only two reliable records that have been confirmed in the Park since 2000 and the red squirrel has not been present for over 18 years (Alston et al. 2012).

Overall, water samples yielded better results than sediment samples regarding species detection and read count for both areas sampled (Figs 1B and S5). In Assynt, only the wild/domestic cat was exclusively detected in sediment samples (four sites), whereas water samples recovered eDNA for ten additional species not found in the sediment samples. The red deer, water vole, field vole, mountain hare and pygmy shrew were also found in sediment samples in Assynt (Fig. 1B), and water vole and wood mouse in the Peak District sediment samples (Fig. S5).

Occupancy Analysis

Of the 18 sites where both latrine and eDNA surveys were conducted, water voles were detected at 13, and field voles were detected at 11. A total of seven wild mammals were recorded at the seven sites with a camera trap from July 10th to October 25th, 2017 (Fig. S3 and Table S5). There were several incidences where a shrew could not be identified to species level using camera traps. For camera traps, water voles were recorded at all sites, red deer at five out of seven, field voles and weasels at three sites, water shrews and otters at two, and a red fox at a single site.

For the 18 sites in Assynt, estimated site occupancy (with 95% confidence intervals) from the combined surveying methods was 0.91 (0.63 – 0.98) for water voles and 0.88 (0.57 – 0.98) for field voles. Red deer were observed at every patch by at least one of the methods, and therefore occupancy was 1 (Table 1). For all three species, per sample detection probability was higher for eDNA taken from water than for eDNA taken from sediment (Table 1, Fig. 2). The surveying method specific efficacy pattern was similar for water voles and field voles (Table 1, Fig. 2): latrine surveys had the highest probability of detecting the species (0.77 and 0.52 respectively), followed by eDNA from water (0.57 and 0.40 respectively), then camera trapping (0.50 and 0.20 respectively), and finally eDNA from sediment (0.27 and 0.02 respectively). Detection probability was higher for water voles than field voles using all four methods (Table 1, Fig. 2). No effort was made to record red deer presence during latrine surveys. Like the water voles and field voles, red deer detection has higher using eDNA from water (0.67, CI: 0.53 – 0.78) compared to eDNA from sediment (0.10, CI: 0.04 – 0.21). Unlike the voles, which were more detectable by cameras than sediment eDNA, red deer detection on cameras was similar to sediment eDNA (0.10, CI: 0.04 – 0.24).

The patterns described above detail surveying event-specific detectability. We also computed the cumulative detection probability for each method and each species ($\hat{p}_{sm}$). The cumulative detection curves over 15 surveying events are shown in Fig. 2. The number of surveying events, *k*, required to achieve $p^*_{psm} \geq 0.95$ for water voles was 3 surveys, 4 samples, 10 samples, and 5 weeks, for latrines, water eDNA, sediment eDNA, and cameras respectively. The number of surveying events, *k*, required to achieve $p^*_{psm} \geq 0.95$ for field voles was 5 surveys, 6 samples, 141 samples, and 14 weeks, for latrines, water eDNA, sediment eDNA, and cameras respectively.

343 The number of surveying events, *k*, required to achieve $p^*_{psm} \geq 0.95$ for red deer was
344 3 samples, 30 samples, and 29 weeks, for water eDNA, sediment eDNA, and cameras
345 respectively (see also Fig. 2).

Discussion

Despite the increasing potential of eDNA metabarcoding as a biomonitoring tool (Deiner et al., 2017), its application has largely been focused on strictly aquatic or semi-aquatic animals, thus restricting management and conservation efforts of the wider ecosystem (Williams et al., 2018). Here, we demonstrate the ability of eDNA metabarcoding to provide a valuable 'terrestrial dividend' for mammals from freshwater lotic ecosystems, with a large proportion of the expected species from the wider landscape being detected in each of the two study locations. In particular, we have demonstrated that water-based eDNA sampling offers a promising and complementary tool to conventional survey methods for the detection of whole mammalian communities.

Detection of mammalian communities using eDNA metabarcoding

Of the species known to be common in both Assynt and the Peak District, eDNA metabarcoding readily detected the water vole, field vole and red deer at the majority of sites surveyed (Figs. 1B and S5). Pygmy, common and water shrews, wood mice and mountain hares were also detected by eDNA metabarcoding at multiple sites in Assynt (Fig. 1B). A higher eDNA detection rate is expected for aquatic and semi-aquatic mammals compared to terrestrial mammals in aquatic environments due to the spatial and temporal stochasticity of opportunities for terrestrial mammals to be in contact with the water (Ushio et al., 2017). The semi-aquatic water vole was generally detected by eDNA metabarcoding where we expected to find it and at relatively high read numbers (Figs. 1B, S1 and S5). This is in line with previous studies in lentic systems (Harper et al., 2019). However, the red deer was the only terrestrial species

370 detected by eDNA sampling at all sites in Assynt, and the terrestrial field vole at over
371 70% of surveyed sites.

372 In addition to lifestyle (semi-aquatic or terrestrial), the number of individuals of
373 each species (i.e. group-living) may be important for eDNA detection (Williams et al.,
374 2018). As a counter example to this, otters and weasels were notably absent in the
375 eDNA samples in Assynt despite being captured by camera traps (Fig. S3 and Table
376 S5). Otters were present in the water eDNA samples at two sites in the Peak District,
377 albeit at a lower number of reads in comparison to most of the other species detected
378 (Fig. S5; Table S2). This mirrors previous studies where eDNA analysis has performed
379 relatively poorly for otter detection in captivity and the wild (Harper et al., 2019;
380 Thomsen et al., 2012). Carnivores were generally detected on fewer occasions (e.g.
381 red foxes, badgers and pine martens; Figs. 1B and S5) or not at all (e.g. stoats and
382 American mink in addition to those discussed above) in comparison to smaller
383 mammals and red deer, and a similar pattern has been shown with North American
384 carnivores in a recent study using eDNA from soil samples (Leempoel et al., 2019).
385 For some of these species, species ecology/behavior such as a relatively large home
386 range and more solitary nature (e.g. red foxes) may go some way towards explaining
387 a lack of, or few, eDNA records. Furthermore, as demonstrated by Ushio et al. (2017)
388 poor efficiency for amplifying some mammal species might be associated to
389 suboptimal experimental conditions (e.g. inadequate primer design, primer bias, DNA
390 concentration, species masking and/or annealing temperatures).

391 Regarding the sampling medium for eDNA, we demonstrated that water is a
392 more effective method for detection of mammal eDNA than sediment (Table 1; Figs.
393 1B and S5). For one of our focal species, the water vole, 75% of sites which were
394 deemed unoccupied by latrine surveys and those with ≤2 individuals (8 sites) in

Assynt, returned a non-detection for sediment eDNA as opposed to 37.5% of sites for water (Figs. 1A, 1B and S1). Distinct temporal inferences are provided by eDNA recovered from water and sediment samples. DNA bound to sediments can remain detectable for a longer period (i.e. up to hundreds of years) and provide historical data, whereas, eDNA retrieved from water samples provide more contemporary data due to a faster degradation in the water column (Turner et al., 2015). It is worth investigating further if sediment eDNA could indicate the presence of a more 'established' population, where a certain threshold of individuals and long-term occupation (i.e. historical) is required for detection in sediment (Fig. S1; Turner et al., 2015; Leempoel et al., 2019).

Importantly, sparse or single eDNA records should be carefully verified. The edible dormouse and grey squirrel sequences identified within the Assynt samples (Fig. 1B) and red squirrel within the Peak District (Fig. S5) highlights the caveats associated with this technique. If management decisions had relied on eDNA evidence alone, false positives for these species could lead to unnecessary resources being allocated for management/eradication programmes as the edible dormouse and grey squirrel are classified as invasive species within Great Britain. These potentially arose due to sample carryover from a previous sequencing run on the same instrument (a known issue with Illumina sequencing platforms; Nelson et al., 2014) which included those species for the reference database construction. Controlling for false positives is certainly a huge challenge in eDNA metabarcoding and the need to standardize and optimize thresholds for doing so is an ongoing debate (Ficetola et al., 2015; Harper et al., 2019).

Even with these concerns around false positives highlighted, two records are potentially noteworthy in a conservation context for UK mammals because of the

relatively high read number associated with these records (Tables S2 and S3). The first of these is the *Felis* records in sediment samples in multiple sites in Assynt (Fig. 1B). Even with 'pure' *F. silvestris* as reference sequences, it was not possible to distinguish between the wild and domesticated species for this 12S fragment (data not shown). Despite ongoing conservation efforts, there may now be no 'pure' Scottish wildcats left in the wild in the UK but isolated populations (perhaps of hybrid origin) may exist in this region (Sainsbury et al., 2019). Given that these eDNA detections were all from sediment samples, it is possible that they may be historical rather than contemporary (see above). The other significant eDNA record was the pine marten in the Peak District. The pine marten (*Martes martes*) is known to occur in the Scottish Highlands but had disappeared from most of the UK and recently has been recovering from historical persecution, including a potential expansion of its range. Still, authentic records from northern England are scarce or lacking altogether (Alston et al., 2012; Sainsbury et al., 2019). However, a record of a recent roadkill exists from just outside the Park's boundary (BBC News, 2018). The high number of reads recovered for the Peak District sample (4293 reads versus 25 in the Assynt sample) adds credence to this positive eDNA detection but further investigations are warranted into the potential presence of this species in the area.

Comparisons between surveying methods

Comparisons of species detection by traditional survey approaches and eDNA analysis are now numerous in the literature, and mainly focus on what is and what is not detected within and across different methods (Hänfling et al., 2016; Leempoel et al., 2019). Yet, there has been growing incorporation of occupancy modelling to estimate the probability of detecting the focal species, in comparison to one other

survey method, either for a single species (Lugg et al., 2018; Tingley et al., 2019) or multiple species (Valentini, et al., 2016; Abrams et al., 2019). Simultaneous multi-method comparisons for multiple species have been lacking and this study directly addresses this for the first time.

The probability of detecting the water vole and field vole was higher for the latrine surveys than eDNA sampling (both water and sediment) and camera traps (Table 1; Fig. 2). However, when considering confidence intervals, there was considerable overlap between latrine, water-based eDNA metabarcoding and camera traps for both species, with only sediment-based eDNA metabarcoding yielding a low probability of detection (Table 1). Detection probabilities for water-based eDNA metabarcoding and camera traps were similar for water voles, with camera traps less likely to detect the field vole than water-based eDNA. For the red deer (for which no latrine survey was undertaken), water-based eDNA metabarcoding had a much higher probability of detection than either sediment-based eDNA metabarcoding or camera traps (which performed similarly; Table 1). Despite the increasing adoption of camera traps in providing non-invasive detections for mammals (Hofmeester et al., 2019), camera traps were outperformed by water-based eDNA metabarcoding for the three focal species in this component of the study. Here, camera traps were deployed so as to sample the habitat of the water vole (see Fig. S3), which may explain lower detection for other terrestrial species in comparison to eDNA metabarcoding (see above). Studies focusing on a single species often report that eDNA analysis outperforms the conventional survey method in terms of detection probabilities (e.g. Lugg et al., 2018). For metabarcoding, there is clearly a need to carefully consider the potential for cross contamination between samples and how false positives (and negatives) could impact detection probabilities using occupancy modelling with eDNA data (Brost et al., 2018;

Lahoz-Monfort et al., 2016). Among the recommendations made by Lahoz-Monfort et al. (2016) to account for these uncertainties, one was the simultaneous collection of data from more conventional surveying methods. Here, we have demonstrated general congruence between surveying methods for the water vole (Table S5; Fig. S1) and using certain species to apply a multiple detection methods model would be appropriate in further studies (Lahoz-Monfort et al., 2016). Alternatively, using repeated sampling and known negative controls in occupancy models that fully incorporate false-positive errors could be applied in the absence of other surveying data (Brost et al., 2018). Overall, multi-species metabarcoding studies may trade-off a slightly lower (but comparable) detection probability than other survey methods for individual species (Fig. 2) in favor of a better overall "snapshot" of occupancy of the whole mammalian community (Ushio et al., 2017).

The comparison between survey 'effort' for the four methods to reach a probability of detection of ≥0.95 is highly informative and provides a blueprint for future studies on mammal monitoring. Focusing on the water vole for example, three latrine surveys would be required. A total of four water-based and 10 sediment-based eDNA replicates or five weeks of camera trapping would be required to achieve the same result (Fig. 2). This increases for the field vole in the same habitat, with five latrine surveys and six water-based eDNA replicates. Sediment-based eDNA metabarcoding would be impractical for this species and camera trapping would take 14 weeks. What is important here is the spatial component and the amount of effort involved in the field. Taking 4-6 water-based eDNA replicates from around one location within a patch could provide the same probability of detecting these small mammals with three latrine surveys. In many river catchments, there may be 100s to 1000s of kilometers to survey that would represent suitable habitat, and only a fraction of that may be occupied by

any given species. This is particularly relevant in the context of recovery of water vole populations post-translocation or in situations where remnant populations are bouncing back after invasive American mink (*Neovison vison*) control has been instigated. On a local scale, finding signs of water voles through latrine surveys is not necessarily difficult, but monitoring the amount of potential habitat (especially lowland) for a species which has undergone such a massive decline nationally is a huge undertaking (Morgan et al., 2019).

The use of eDNA metabarcoding from freshwater systems to generate an initial, coarse and rapid 'distribution map' for vertebrate biodiversity (and at a relatively low cost) could transform biomonitoring at the landscape level. For group-living (i.e. deer) and small mammal species, carefully chosen sampling points (with at least five water-based replicates) along multiple river courses could provide a reliable indication of what species are present in the catchment area if conducted during times of peak abundance (i.e. Summer and Autumn). Then, on the basis of this, practitioners could choose to further investigate specific areas for confirmation of solitary, rare or invasive species (e.g. carnivores) with increased effort in terms of both the number of sampling sites and replicates taken. At present, we would recommend the use of eDNA metabarcoding alongside other non-invasive surveying methods (e.g. camera traps) when monitoring invasive species or species of conservation concern to maximize monitoring efforts (Abrams et al., 2019; Sales et al., 2019).

It is clear that eDNA metabarcoding is a promising tool for monitoring semi-aquatic and terrestrial mammals in both lotic (this study) and lentic systems (Harper et al., 2019; Ushio et al., 2017). We detected a large proportion of the expected mammalian community (Table S1). Water-based eDNA metabarcoding is comparable or out-performs other non-invasive survey methods for several species (Fig. 2).

However, there remain challenges for the application of this technique over larger spatial and temporal scales. Technical issues of metabarcoding in laboratory and bioinformatic contexts have been dealt with elsewhere (Harper et al., 2019) but understanding the distribution of eDNA transport in the landscape and its entry into natural lotic systems is at an early stage (and incorporating such variables in occupancy modelling approaches). This clearly requires more detailed and systematic eDNA sampling than undertaken here, particularly in an interconnected river/stream network with organisms moving between aquatic and terrestrial environments. Leempoel et al. (2019) recently demonstrated the feasibility for detecting terrestrial mammal eDNA in soil samples but this study has shown that sampling a few key areas in freshwater ecosystems (e.g. larger rivers and lakes) within a catchment area could potentially provide data on a large proportion (if not all) of the mammalian species within it, even when some species are present at low densities (Deiner et al., 2017). In this regard, future studies might also investigate the value of citizen science, where trained volunteers can contribute to data collection at key sites, thus scaling up the reach of research whilst raising public awareness and the significance of mammalian conservation concerns (Parsons et al., 2018).

Data accessibility

Data will be made available in Dryad upon acceptance.

Authors contributions

ADM, XL, CS, OSW, IC, SM, NGS, SSB, EO, BH and LLH conceived the study. Monitoring and live-trapping of water voles was part of XL, CS, EB and JD's ongoing work in Assynt. JD analysed the camera trap data. DAD advised on primer set/data validation and provided information and data on mammals in the Peak District. ADM, NGS, SSB and MBM carried out the eDNA sampling. MBM, NGS, SSB, CB and ADM performed the laboratory work. NGS, OSW, LRH, MBM, CB and ADM carried out the bioinformatic analyses. NGS, ADM, IC and MBM analysed the eDNA data. CS and JD conducted the occupancy modelling. ADM, NGS, CS, JD, MBM and LRH wrote the paper, with all authors contributing to editing and discussions.

Acknowledgements

The eDNA component of this project was funded by the British Ecological Society (grant no. SR17/1214) awarded to ADM. JD was supported by University of Massachusetts Organismal and Evolutionary Biology Research Grant and Spring 2018 Graduate School Fieldwork Grant. We thank Kristy Deiner for enlightening conversations about these results. We are grateful to Jerry Herman and Andrew Kitchener for the tissue samples from National Museums Scotland. Christine Gregory, Douglas Ross and Sarah Proctor provided water vole and otter information for sampling in the Peak District and Sara Peixoto provided sequence assemblies. We thank the various landowners for permission to sample on their property. The authors

562 declare that no conflict of interest exists. We thank Brittany Mosher and the
563 anonymous reviewers for significantly improving the manuscript.

References

Abrams, J. F., Hoerig, L., Brozovic, R., Axtner, J., Crampton-Platt, A., Mohamed, A., … Wilting, A. (2019). Shifting up a gear with iDNA: from mammal detection events to standardized surveys. Journal of Applied Ecology, 56, 1637-1648.

Alston, D., Mallon, D.P. & Whiteley, D. (2012). *The Mammals of Derbyshire*. Derbyshire Mammal Group and Sorby Natural History Society. England.

BBC News (2018). Mystery of pine marten found dead '100 miles from home'. Available at: https://www.bbc.co.uk/news/uk-england-derbyshire-45331872. Accessed April 25th, 2019.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac., E. (2016). obitools: A unix inspired software package for DNA metabarcoding. Molecular Ecology Resources, 16, 176–182.

Brost, B. M, Mosher, B. A., & Davenport, K. A. (2018). A model-based solution of observational errors in the laboratory. Molecular Ecology Resources, 18, 580-589.

Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., … Brosse, S. (2019). Unlocking biodiversity and conservation studies in high diversity environments using environmental DNA (eDNA): a test with Guianese freshwater fishes. Molecular Ecology Resources, 19, 27-46.

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, 5872–5895.

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., …

586 Taberlet, P. (2015). Replication levels, false presences and the estimation of the
587 presence/absence from eDNA metabarcoding data. Molecular Ecology Resources
588 15, 543–556.

589 Fiske, I., & Chandler, R. (2011). unmarked: An R Package for Fitting Hierarchical
590 Models of Wildlife Occurrence and Abundance. Journal of Statistical Software, 43(10),
591 1-23.

592 Hänfling, B., Handley, L. L., Read, D. S., Hahn, C., Li, J., Nichols, P., … Winfield, I. J.
593 (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term
594 data from established survey methods. Molecular Ecology, 25(13), 3101–3119.

595 Harper, L. R., Handley, L. L., Carpenter, A. I., Murray-Dickson, G., Di Muri, C.,
596 Macgregor, C. J., … Hänfling, B. (2019). Environmental DNA (eDNA) metabarcoding
597 of pond water as a tool to survey conservation and management priority mammals.
598 Biological Conservation, 238, 108225.

599 Hofmeester, T. R., Cromsigt, J. P. G. M., Odden, J., Andrén, H., Kindberg, J., & Linnell,
600 J. D. C. (2019). Framing pictures: A conceptual framework to identify and correct for
601 biases in detection probability of camera traps enabling multi-species comparison.
602 Ecology and Evolution, 9, 2320-2336.

603 Holman, L. E., de Bruyn, M., Creer, S., Carvalho, G., Robidart, J., & Rius, M. (2019).
604 The detection of introduced and resident marine species using environmental DNA
605 metabarcoding of seawater and sediment. Scientific Reports, 9, 11559.

606 Lahoz-Monfort, J. L, Guillera-Arroita, G., & Tingley, R. (2016). Statistical approaches
607 to account for false-positive errors in environmental DNA samples. Molecular Ecology
608 Resources, 16, 673-685.

609 Leempoel, K., Herbert, T., & Hadly, E. A. (2019). A comparison of eDNA to camera
610 trapping for assessment of terrestrial mammalian diversity. bioRxiv, doi:
611 10.1101/634022.

612 Lugg, W. H., Griffiths, J., van Rooyen, A. R., Weeks, A. R., & Tingley, R. (2018).
613 Optimal survey designs for environmental DNA sampling. Methods in Ecology and
614 Evolution, 9, 1049–1059.

615 MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J.A., & Langtimm,
616 C. A. (2002). Estimating site occupancy rates when detection probabilities are less
617 than one. Ecology, 83: 2248-2255.

618 MacKenzie, D. I., & Royle, J. A. (2005). Designing occupancy studies: general advice
619 and allocating survey effort. Journal of Applied Ecology, 42(6) 1105-1114.

620 MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L., & Hines, J. E.
621 (2017). *Occupancy estimation and modeling: inferring patterns and dynamics of*
622 *species occurrence*. Elsevier.

623 Mathews, F., Kubasiewicz, L. M., Gurnell, J., Harrower, C. A., McDonald, R. A., &
624 Shore, R. F. (2018). *A Review of the Population and Conservation Status of British*
625 *Mammals: Technical Summary*. Natural England. Peterborough.

626 Morgan, W., Cornulier, T., & Lambin, W. (2019). Colonisation dynamics during range
627 expansion is poorly predicted by dispersal in the core range. Ecography, 42, 1142-
628 1151
629
630 Nelson, M. C., Morrison, H. G., Benjamino, J., Grim, S. L., & Graf, J. (2014). Analysis,
631 optimization and verification of Illumina-generated 16s rRNA gene amplicon surveys.
632 PLoS ONE, 9, e0094249
633
634 Nichols, J. D., Bailey, L. L., O'Connell Jr, A. F., Talancy, N. W., Grant, E. H. C., Gilbert,
635 A. T., ... Hines, J. E. (2008). Multi-scale occupancy estimation and modelling using
636 multiple detection methods. Journal of Applied Ecology, 45(5), 1321-1329.
637
638 Niedballa, J., Courtiol, A., & Sollmann, R. (2018). camtrapR: Camera Trap Data
639 Management and Preparation of Occupancy and Spatial Capture-Recapture
640 Analyses. R package version 1.0.

641 Parsons, A. W., Goforth, C., Costello, R., & Kays, R. (2018). The value of citizen
642 science for ecological monitoring of mammals. PeerJ, 6, e4536.

643 R Core Team. (2018). R: A language and environment for statistical computing. R
644 Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

645 Sainsbury, K. A., Shore, R. F., Schofield, H., Croose, E., Campbell, R. D., & McDonald,
646 R. A. (2019). Recent history, current status, conservation and management of native
647 mammalian carnivore species in Great Britain. Mammal Review, 49, 171-188.

648 Sales, N. G., da Cruz Kaizer, M., Coscia, I., Perkins, J. E., Highlands, A., Boubli, J. P.,

649 … McDevitt, A. D. (2019). Assessing the potential of environmental DNA
650 metabarcoding for monitoring Neotropical mammals: a case study in the Amazon and
651 Atlantic Forest, Brazil. Mammal Review, In Press.

652 Schnell, I.B., Bohmann, K., & Gilbert, M.T.P. (2015). Tag jumps illuminated - reducing
653 sequence-to-sample misidentifications in metabarcoding studies. Molecular Ecology
654 Resources, 15(6), 1289–1303.

655 Sutherland, C., Elston, D.A., & Lambin, X. (2013). Accounting for false positive
656 detection error induced by transient individuals. Wildlife Research, 40(6), 490-498.

657 Sutherland, C. S., Elston, D. A., & Lambin, X. (2014). A demographic, spatially explicit
658 patch occupancy model of metapopulation dynamics and persistence. Ecology, 95:
659 3149-3160.

660 Tingley, R., Greenlees, M., Oertel, S., van Rooyen, A. R., & Weeks, A. R. (2019).
661 Environmental DNA sampling as a surveillance tool for cane toad *Rhinella marina*
662 introductions on offshore islands. Biological Invasions, 21, 1-6.

663 Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T.,
664 … Willerslev, E. (2012). Monitoring endangered freshwater biodiversity using
665 environmental DNA. Molecular Ecology, 21, 2565-2573.

666 Turner, C. R., Uy, K. L., & Everhart, R. C. (2015). Fish environmental DNA is more
667 concentrated in aquatic sediments than surface water. Biological Conservation, 183,
668 93–102.

669 Ushio, M., H. Fukuda, T. Inoue, K. Makoto, O. Kishida, K. Sato, … Miya, M. (2017).
670 Environmental DNA enables detection of terrestrial mammals from forest pond water.
671 Molecular Ecology Resources, 17(6), e63-e75.

672 Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., …
673 Dejean, T. (2016). Next-generation monitoring of aquatic biodiversity using
674 environmental DNA metabarcoding. Molecular Ecology, 25(4), 929–942.

675 Visconti, P., Pressey, R. L., Giorgini, D., Maiorano, L., Bakkenes, M., Boitani, L., …
676 Rondinini, C. (2011). Future hotspots of terrestrial mammal loss. Philosophical
677 transactions of the Royal Society of London. Series B, Biological Sciences, 366(1578),
678 2693-702.

679 Williams, K. E., Huyvaert, K. P., Vercauteren, K. C., Davis, A. J., & Piaggio, A. J.
680 (2018). Detection and persistence of environmental DNA from an invasive, terrestrial
681 mammal. Ecology and Evolution, 8(1), 688–695.

682

683 **Tables**

684

685 Table 1. Estimated site occupancies and detection probabilities obtained for water-
686 based eDNA (w-eDNA), sediment-based eDNA (s-eDNA) and conventional survey
687 methods (Latrine and Camera) in Assynt.

		Detection probability			
Species	Occupancy	*Latrine*	*w-eDNA*	*s-eDNA*	*Camera*
Water vole	0.91 (0.63 – 0.98)	0.77 (0.59 – 0.89)	0.57 (0.43 – 0.71)	0.27 (0.16 – 0.41)	0.50 (0.35 – 0 .65)
Field vole	0.89 (0.57 – 0.98)	0.52 (0.34 – 0.69)	0.40 (0.26 – 0.55)	0.02 (0.00 – 0.14)	0.20 (0.10 – 0.37)
Red deer	1.00 (1.00 – 1.00)	--	0.67 (0.53 – 0.78)	0.10 (0.04 – 0.21)	0.10 (0.09 – 0.24)

688

689

690

FIGURES

Figure 1. Environmental DNA (eDNA) sampling sites in Assynt, Scotland (A). Categorical values for water vole abundance at each site based on live-trapping data. In (B), a bubble graph representing presence-absence and categorical values of the number of reads retained (after bioinformatic filtering) for eDNA (water in blue and sediment in orange) from each wild mammal identified in each site in Assynt (A1-A18).

Figure 2. The detection probabilities of each survey method (sediment-based eDNA, water-based eDNA, latrine and camera) for each of three focal species (from top to bottom on the left); water vole; field vole and red deer. On the right, the accumulation curves for each species for the number of sampling events for each survey method to provide a ≥0.95 probability of detection.

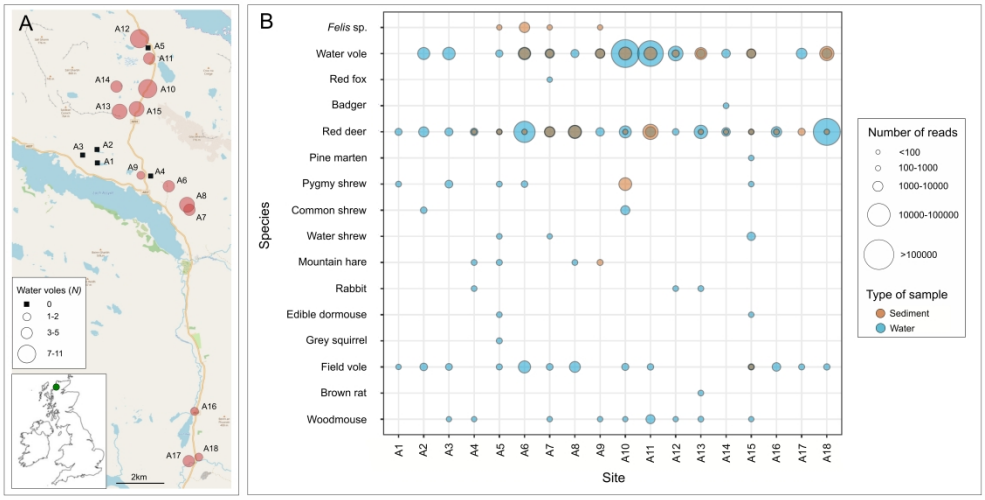

Figure 1. Environmental DNA (eDNA) sampling sites in Assynt, Scotland (A). Categorical values for water vole abundance at each site based on live-trapping data. In (B), a bubble graph representing presence-absence and categorical values of the number of reads retained (after bioinformatic filtering) for eDNA (water in blue and sediment in orange) from each wild mammal identified in each site in Assynt (A1-A18).

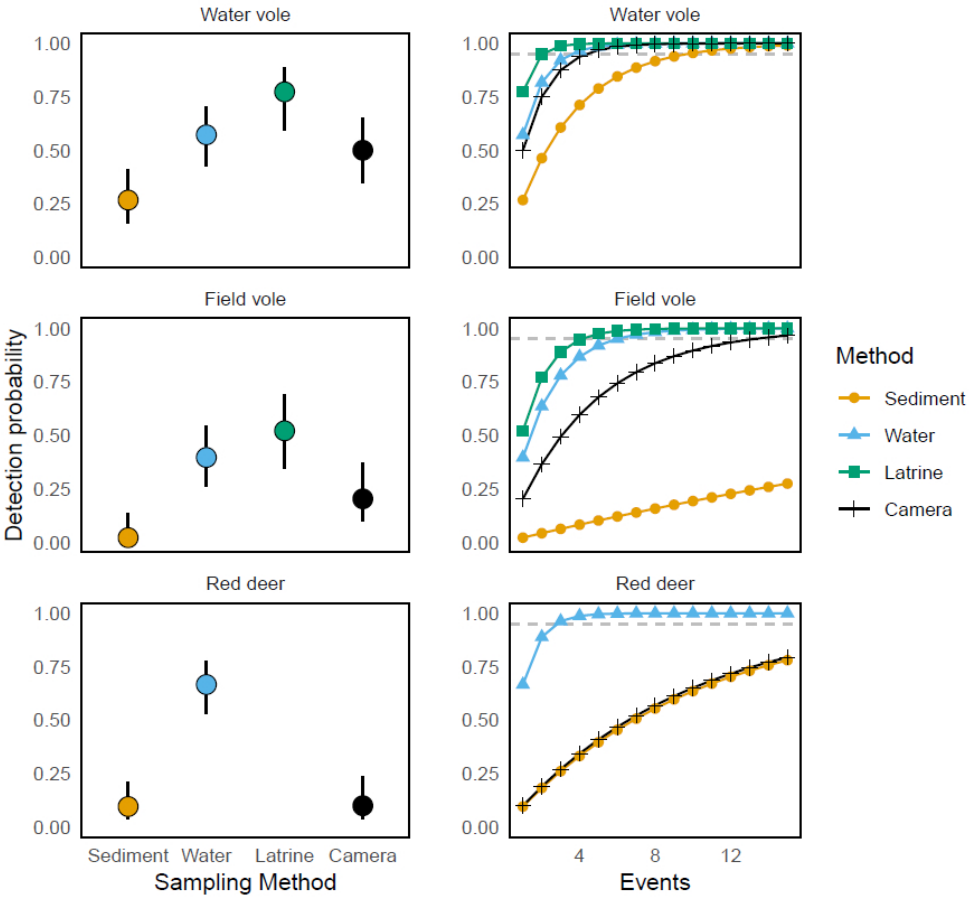

Figure 2. The detection probabilities of each survey method (sediment-based eDNA, water-based eDNA, latrine and camera) for each of three focal species (from top to bottom on the left); water vole; field vole and red deer. On the right, the accumulation curves for each species for the number of sampling events for each survey method to provide a ≥0.95 probability of detection.

SUPPLEMENTARY MATERIAL

Fishing for mammals: landscape-level monitoring of terrestrial and semi-aquatic communities using eDNA from lotic ecosystems

Naiara Guimarães Sales[1]*, Maisie B. McKenzie[1]*, Joseph Drake[2]*, Lynsey R. Harper[3], Samuel S. Browett[1], Ilaria Coscia[1], Owen S. Wangensteen[4], Charles Baillie[1], Emma Bryce[5], Deborah A. Dawson[6], Erinma Ouchu[1], Bernd Hänfling[3], Lori Lawson Handley[3], Stefano Mariani[1,7], Xavier Lambin[5], Christopher Sutherland[2] and Allan D. McDevitt[1#]

[1]Environment and Ecosystem Research Centre, School of Science, Engineering and Environment and Life Sciences, University of Salford, Salford, UK

[2]Department of Environmental Conservation, University of Massachusetts-Amherst, Amherst, USA

[3]Department of Biological and Marine Sciences, University of Hull, Kingston upon Hull, UK

[4]Norwegian College Fishery Science, University of Tromsø, Tromsø, Norway

[5]School of Biological Sciences, University of Aberdeen, Aberdeen, UK

[6]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK

[7]School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool, UK

Appendix 1

eDNA sample collection

Three water sample replicates (two litres each) and three sediment sample replicates (50 ml falcon tube, approximately half-filled) were taken at each site in Assynt, always within a reachable distance from the river's edge and at a depth where sediment samples could be taken (Fig. S4A). Water samples were filtered on site using a Sterivex 0.45 µm filter unit (Merck Millipore) and filters were stored in silica beads in the field (1-3 days; Majaneva et al., 2018) then frozen until DNA extraction. Sediment samples were stored in 100% ethanol. Appropriate decontamination precautions were taken including the use of disposable gloves and decontamination of all equipment and surfaces by using 50% bleach solution). Samples from the Peak District were filtered within 5 hours in the University of Salford laboratory facilities due to its close proximity to the sampling locations. A single filter was used for each replicate in Assynt and the Peak District, and the volume filtered varied between each, ranging from 150 ml to 2 L (see Tables S2 and S3). Negative field controls were taken in both Assynt (N= 8) and the Peak District (N= 2) and were obtained by collecting, preserving and processing distilled water in exactly the same way as the field samples. The amount of sediment collected also varied, with 4 to 10g used in the extractions. A Pearson's correlation was performed to determine if the amount of water/sediment influenced the amount of retained reads for mammals after bioinformatic filtering.

Reference database

Given that this project proposed to use mammal-specific primers (MiMammal-U, Ushio et al., 2017) to target the same region of 12S as the MiFish primers (Miya et al., 2015),

an *in silico* evaluation was first performed using ecoPCR (Ficetola et al., 2010) of the MiMammal-U primer set against a custom, phylogenetically curated reference database for mammals distributed in the UK and Ireland. This database was one of several databases constructed for UK vertebrates and used in an eDNA metabarcoding study of pond biodiversity (see Harper et al. 2019 for details). The mammal database was updated in July 2018 for the purposes of the present study. Parameters were set to allow a fragment size of 50-250 bp and different number of mismatches (0, 1, 2, 3) between each primer and each sequence in the reference database. Reference sequence data was available for 103 mammal species (91.96%) in the UK. The nine species that were not represented were either cetaceans or bats. Of those species with reference sequence data (N = 103), 44 (42.72%), 65 (63.11%), 72 (69.90%), and 82 (79.61%) mammals were amplified when 0, 1, 2, and 3 primer-sequence mismatches were allowed respectively. Species that did not amplify under any scenario due to the lack of an appropriate reference sequence for the specific 12S region being targeted for MiMammal (and of relevance to this study) were the European water vole (*Arvicola amphibius*), greater white-toothed shrew (*Crocidura russula*), Millet's shrew (*Sorex coronatus*), Eurasian pygmy shrew (*Sorex minutus*), field vole (*Microtus agrestis*), common vole (*Microtus arvalis*), grey squirrel (*Sciurus carolinensis*), and European polecat (*Mustela furo*).

Because certain focal mammalian species were missing from online reference databases, a new reference database of 32 UK terrestrial mammals targeting this fragment of the 12S gene was created from ethanol-preserved tissues samples obtained from National Museums Scotland (Table S6). DNA was extracted using the ISOLATE II kit according to the manufacturer's protocol. These DNA samples were then included in a large vertebrate barcoding project using the MiFish (Miya et al.,

2015) primers (O. Wangensteen et al., *unpublished data*). Although these primers were originally designed to amplify fishes, they are known to amplify mammals also and target the exact same region as the MiMammal primers (Ushio et al., 2017). This was conducted to save on sequencing costs and the prior knowledge that these primers would generate reference sequences for the majority of UK mammals (Hänfling et al., 2016). Of these mammals, only *Sorex araneus* and *Neomys fodiens* failed to generate reference sequences. PCRs were then carried out on a subset of the tissue-extracted DNA (see Table S6) and Sanger-sequenced (Macrogen Inc.) using the MiMammal-U primers (Ushio et al., 2017) to confirm the results obtained with the MiFish primers.

eDNA Laboratory Methods

Field and Laboratory controls

In order to avoid the risk of contamination, clean and consistent field and laboratory protocols are paramount. Besides the decontamination measures taken, three types of negative controls (field, extraction and PCR) were included. Field blanks comprised of distilled water which was preserved and processed using exactly the same protocols and equipment as the field samples. These were processed first to ascertain if contaminations arose in the field (either during the water/sediment sampling or during the filtering process). DNA extraction blanks, represented by empty tubes included in the extraction step, were undertaken at the end of each batch of extractions to ascertain the potential for contaminations arising from reagents and the laboratory environment. Finally, no-template amplification controls (NTC) were included during the amplification step (PCR) of the actual samples through the inclusion of several reactions lacking DNA to account for putative contamination during this procedure.

The chronology of DNA extraction followed an increasing order of expected abundance in the eDNA samples (all field blanks extracted first, followed by the sites with supposedly zero water vole abundance, up to the highest densities last). Field blanks were processed at the beginning of the DNA extraction to try to tease apart the potential contamination between field and lab contaminations. The implementation of this chronology was due to the fact that it is the first time a study focusing on using eDNA with terrestrial and semi-aquatic mammals has been undertaken like this in multiple sites that were sampled in the same session, with the researchers moving around in the habitat (terrestrial) of the target group of organisms.

eDNA amplification and sequencing

A set of 96 primers pairs with seven-base sample-specific MIDs and a variable number (2-4) of fully degenerate positions (leading Ns) to increase variability in amplicon sequences were used. PCR amplification was conducted using a single-step protocol and to minimize bias in individual reactions, PCRs were replicated three times for each sample and subsequently pooled. The PCR reaction consisted of a total volume of 20 µl including 10 µl Amplitaq; 0.16 µl of BSA; 1.0 µl of each of the two primers (5 µM); 5.84 µl of ultra-pure water, and 2 µl of DNA template. The PCR profile included an initial denaturing step of 95°C for 10 min, 40 cycles of 95°C for 30s, 60°C for 45s, and 72°C for 30s and a final extension step of 72°C for 5 min. Amplification were checked through electrophoresis in a 1.5% agarose gel stained with GelRed (Cambridge Bioscience). PCR products were pooled in two different sets and a left-sided size selection was performed using 1.1x Agencourt AMPure XP (Beckman Coulter). Illumina libraries were built from each set, using a NextFlex PCR-free library preparation kit according to the manufacturer's protocols (Bioo Scientific). Libraries

were then quantified by qPCR using a NEBNext qPCR quantification kit (New England Biolabs) and pooled in equimolar concentrations along with 1% PhiX (v3, Illumina). The libraries were run at a final molarity of 9pM on an Illumina MiSeq platform using the 2 x 150bp v2 chemistry.

Bioinformatic analysis

OBITools metabarcoding package (Boyer et al., 2016) was used for the bioinformatic analysis. Quality of the reads was assessed using FastQC, paired-end reads were aligned using illuminapairedend and the ngsfilter command was used for dataset demultiplexing. Short fragments originated from library preparation artefacts (primer-dimer, non-specific amplifications) and reads containing ambiguous bases were removed applying a length filter selecting fragments of 140-190bp using obrigrep. Clustering of strictly identical sequences was performed using obiuniq and a chimera removal step was applied in vsearch (Rognes et al., 2016) through the uchime-denovo algorithm (Edgar et al., 2011). The taxonomic assignment was conducted using ecotag.

A stringent approach was applied to our analyses to avoid false positives and exclude MOTUs/reads putatively belonging to sequencing errors or contamination. The final dataset included only MOTUs that could be identified to species level (>0.98), and MOTUs containing less than 10 reads and with a similarity to a sequence in the reference database lower than 98% were discarded (Cilleros et al., 2019). Singleton reads within individual replicates were also discarded. The maximum number of reads detected in the controls for each MOTU in each sequencing run were removed from all samples (Table S7). For water voles, field voles and red deer (the most abundant wild mammals in terms of sequence reads in our dataset), this equated to a sequence

149 frequency threshold of ≤0.17%, within the bounds of previous studies on removing
150 sequences to account for contamination and tag jumping (Cilleros et al., 2018; Schnell,
151 Bohmann, & Gilbert, 2015). The number of retained reads per replicate was not
152 significantly correlated with the volume of water filtered (Pearson's correlation: r =
153 0.213; p = 0.094) or the amount of sediment collected (Pearson's correlation: r = 0.076;
154 p = 0.556).

TABLES

Table S1. Species (and the Order to which they belong) that are expected to be found within Assynt (based on Matthews et al. 2018) and the Peak District (Alston et al. 2012) and whether or not they were detected by eDNA. A * indicates species where presence is uncertain from Matthews et al. (2018).

Common name	Scientific name	Order	eDNA
Assynt			
Red deer	*Cervus elaphus*	*Artiodactyla*	**Yes**
Sika deer	*Cervus nippon*	*Artiodactyla*	No
Roe deer	*Capreolus capreolus*	*Artiodactyla*	No
Water vole	*Arvicola amphibius*	*Rodentia*	**Yes**
Field vole	*Microtus agrestis*	*Rodentia*	**Yes**
Wood mouse	*Apodemus sylvaticus*	*Rodentia*	**Yes**
Bank vole*	*Myodes glareolus*	*Rodentia*	No
Brown rat	*Rattus norvegicus*	*Rodentia*	**Yes**
Pygmy shrew	*Sorex minutus*	*Eulipotyphla*	**Yes**
Water shrew	*Neomys fodiens*	*Eulipotyphla*	**Yes**
Common shrew	*Sorex araneus*	*Eulipotyphla*	**Yes**
Hedgehog*	*Erinaceus europaeus*	*Eulipotyphla*	No
European mole	*Talpa europaea*	*Eulipotyphla*	No
Mountain hare	*Lepus timidus*	*Lagomorpha*	**Yes**
European rabbit	*Oryctolagus cuniculus*	*Lagomorpha*	**Yes**
Stoat	*Mustela erminea*	*Carnivora*	No
Weasel	*Mustela nivalis*	*Carnivora*	No
Badger	*Meles meles*	*Carnivora*	**Yes**
Otter	*Lutra lutra*	*Carnivora*	No
Red fox	*Vulpes vulpes*	*Carnivora*	**Yes**
Pine marten	*Martes martes*	*Carnivora*	**Yes**
Wildcat*	*Felis silvestris*	*Carnivora*	?
Peak District			
Red deer	*Cervus elaphus*	*Artiodactyla*	**Yes**
Roe deer	*Capreolus capreolus*	*Artiodactyla*	No
Fallow deer	*Dama dama*	*Artiodactyla*	No
Water vole	*Arvicola amphibius*	*Rodentia*	**Yes**
Field vole	*Microtus agrestis*	*Rodentia*	**Yes**
Wood mouse	*Apodemus sylvaticus*	*Rodentia*	**Yes**
Bank vole	*Myodes glareolus*	*Rodentia*	No
Brown rat	*Rattus norvegicus*	*Rodentia*	No
House mouse	*Mus musculus*	*Rodentia*	No
Grey squirrel	*Sciurus carolinensis*	*Rodentia*	**Yes**
Harvest mouse*	*Micromys minutus*	*Rodentia*	No
Pygmy shrew	*Sorex minutus*	*Eulipotyphla*	**Yes**
Water shrew	*Neomys fodiens*	*Eulipotyphla*	**Yes**
Common shrew	*Sorex araneus*	*Eulipotyphla*	**Yes**
Hedgehog	*Erinaceus europaeus*	*Eulipotyphla*	No
European mole	*Talpa europaea*	*Eulipotyphla*	No

Mountain hare	*Lepus timidus*	*Lagomorpha*	No
Brown hare	*Lepus europaeus*	*Lagomorpha*	No
European rabbit	*Oryctolagus cuniculus*	*Lagomorpha*	**Yes**
Stoat	*Mustela erminea*	*Carnivora*	No
Weasel	*Mustela nivalis*	*Carnivora*	No
Badger	*Meles meles*	*Carnivora*	**Yes**
Otter	*Lutra lutra*	*Carnivora*	**Yes**
Red fox	*Vulpes vulpes*	*Carnivora*	No
American mink	*Neovison vison*	*Carnivora*	No
Pine marten	*Martes martes*	*Carnivora*	**Yes**
Polecat	*Mustela putorius*	*Carnivora*	No

162

163

164 **Table S2.** Species identified (with at least 98% identity to the reference database) and
165 their associated number of reads after bioinformatic filtering in each site (Assynt A1-
166 A18 and Peak District P1-P3) and in each of three replicates (_1 to _3) for water-based
167 eDNA. The volume of water filtered is indicated for each replicate.
168 *Additional file: TableS2_Reads_Water.xlsx*
169
170 **Table S3.** Species identified (with at least 98% identity to the reference database) and
171 their associated number of reads after bioinformatic filtering in each site (Assynt A1-
172 A18 and Peak District P1-P3) and in each of three replicates (_1 to _3) for sediment-
173 based eDNA. The weight of sediment used for the DNA extraction is indicated for each
174 replicate.
175 *Additional file: TableS3_Reads_Sediment.xlsx*
176

177 **Table S4.** Number of reads obtained after all filtering steps applied to remove non-
178 target MOTUs.

WATER	Total
Total Reads	13,336,064
After removing reads from the blanks	10,709,199
After removing non-mammal reads	10,262,851
After removing human reads	8,508,564
After removing domestic animals (*Sus, Bos, Equus, Ovis, Canis*)	5,544,208
MOTUs with minimum identity of 0.98	5,414,427

SEDIMENT	Total
Total Reads	3,309,866
After removing reads from the blanks	1,684,433
After removing non-mammal reads	1,543,826
After removing human reads	649,499
After removing domestic animals (*Sus, Bos, Equus, Ovis, Canis*)	500,473
MOTUs with minimum identity of 0.98	465,997

179
180
181
182

183 **Table S5**. Mammalian species recorded at seven camera traps in Assynt. Boxes

184 shaded in grey represent sites where each species was recorded.

185

Common name	Scientific name	Site						
		A5	A10	A11	A12	A13	A14	A15
Water vole	*Arvicola amphibius*	■	■	■	■	■	■	■
Red deer	*Cervus elaphus*		■	■		■	■	■
Field vole	*Microtus agrestis*		■		■		■	
Water shrew	*Neomys fodiens*		■		■			
Weasel	*Mustela nivalis*		■	■			■	
Otter	*Lutra lutra*		■	■				
Red fox	*Vulpes vulpes*			■				
Unidentified Shrew	-		■		■			

186

187

188 **Table S6.** List of tissue samples from mammals used for generating a local reference
189 database using MiFish primers (Miya et al. 2015). All species were tested for
190 amplification using MiMammal-U primers (Ushio et al. 2017) and those highlighted in
191 bold were Sanger-sequenced.

192

Common name	Scientific name	ID
Wood mouse	*Apodemus sylvaticus*	Z.2009.101.1025
Wood mouse	*Apodemus sylvaticus*	Z.2009.101.1149M
House mouse	*Mus domesticus*	Z.2009.101.593M
House mouse	*Mus domesticus*	Z.2009.101.426
Field Vole	***Microtus agrestis***	**Z.2009.101.1045**
Field Vole	***Microtus agrestis***	**Z.2009.101.1994M**
Bank Vole	*Myodes glareolus*	Z.2009.101.97M
Bank Vole	*Myodes glareolus*	Z.2009.101.696M
Weasel	*Mustela nivalis*	Z.2009.101.664
Weasel	*Mustela nivalis*	Z.2009.101.363
Yellow-necked mouse	*Apodemus flavicollis*	Z.2009.101.983M
Yellow-necked mouse	*Apodemus flavicollis*	Z.2009.101.984M
Water shrew	***Neomys fodiens***	**Z.2009.101.141M**
Water shrew	***Neomys fodiens***	**Z.2009.101.1915M**
Pygmy shrew	***Sorex minutus***	**Z.2009.101.1162M**
Pygmy shrew	***Sorex minutus***	**Z.2009.101.458M**
Common shrew	***Sorex araneus***	**Z.2009.101.611M**
Common shrew	***Sorex araneus***	**Z.2009.101.126M**
Common Vole	***Microtus arvalis***	**Z.2009.101.991**
Common Vole	***Microtus arvalis***	**Z.2009.101.917**
Brown Rat	*Rattus norvegicus*	Z.2009.101.931
Brown Rat	*Rattus norvegicus*	Z.2009.101.1026
Grey Squirrel	*Sciurus carolinensis*	23/24
Grey Squirrel	*Sciurus carolinensis*	23/10
Water Vole	***Arvicola amphibius***	**23/15**
Water Vole	***Arvicola amphibius***	**23/17**
Edible dormouse	*Glis glis*	23/16
Edible dormouse	*Glis glis*	23/35
Brown hare	*Lepus europaeus*	23/22
Mountain hare	*Lepus timidus*	23/20
Mountain hare	*Lepus timidus*	23/1
Hedgehog	*Erinaceus europaeus*	23/19
Mole	*Talpa europaea*	23/13
Mole	*Talpa europaea*	23/14
Red fox	*Vulpes vulpes*	23/25
Badger	*Meles meles*	23/12
Badger	*Meles meles*	23/34
Otter	***Lutra lutra***	**23/7**
Otter	***Lutra lutra***	**23/33**
Polecat	*Mustela putorius*	23/5
Polecat	*Mustela putorius*	23/6
Red deer	*Cervus elaphus*	23/31
Red deer	*Cervus elaphus*	23/32
Sheep	*Ovis aries*	23/9
Horse	*Equus caballus*	24/31
Red Squirrel	*Sciurus vulgaris*	1/24
Red Squirrel	*Sciurus vulgaris*	1/31

Pine marten	*Martes martes*	1/1
Pine marten	*Martes martes*	1/13
Coypu	*Myocastor coypus*	62/12
Coypu	*Myocastor coypus*	22/13
Brown hare	*Lepus europaeus*	22/7
Stoat	*Mustela erminea*	22/31
Stoat	*Mustela erminea*	22/33
Red fox	*Vulpes vulpes*	21/28
Hedgehog	*Erinaceus europaeus*	72/32
Sika	*Cervus nippon*	57/31
Horse	*Equus caballus*	57/24
Beaver	*Castor fiber*	63/25
Sheep	*Ovis aries*	58/31
American mink	***Neovison vison***	AMX01
American mink	***Neovison vison***	AMX02
Wildcat	*Felis silvestris*	Z.2015.118.1
Wildcat	*Felis silvestris*	Z.2015.118.2

193

194 **Table S7**. Maximum number of reads subtracted to control for contamination and/or
195 tag switching for each wild species in each eDNA sampling type (water or sediment)
196 and the type of blank in which the reds were identified (Field, Extraction and PCR).
197 Species indicated by * were not identified as eDNA positive records.

198

Common name	Scientific name	Blank	Reads
Red deer	*Cervus elaphus*	Field	164
Water vole	*Arvicola amphibius*	Extraction	7479
Field vole	*Microtus agrestis*	Field	324
Wood mouse	*Apodemus sylvaticus*	None	0
Brown rat	*Rattus norvegicus*	None	0
Pygmy shrew	*Sorex minutus*	Field	1
Water shrew	*Neomys fodiens*	Extraction	1
Common shrew	*Sorex araneus*	Field	2
Mountain hare	*Lepus timidus*	Field	76
European rabbit	*Oryctolagus cuniculus*	Field	38
Stoat*	*Mustela erminea*	Field	68
Badger	*Meles meles*	None	0
Otter	*Lutra lutra*	Extraction	1
Red fox	*Vulpes vulpes*	None	0
Pine marten	*Martes martes*	None	0
Cat	*Felis* spp.	None	0
American mink*	*Neovison vison*	Extraction	343
Red squirrel	*Sciurus vulgaris*	Extraction	1
Grey squirrel	*Sciurus carolinensis*	None	0
Edible dormouse	*Glis glis*	None	0
Human 1	*Homo sapiens*	Field	547
Human 2	*Homo sapiens*	Field	110107
Human 3	*Homo sapiens*	Field	1
Cattle	*Bos* spp.	Extraction	1630
Sheep	*Ovis* spp.	Field	122
Pig	*Sus scrofa domesticus*	Field	99
Dog	*Canis lupus familiaris*	Field	135
Horse	*Equus przewalskii*	None	0

FIGURES

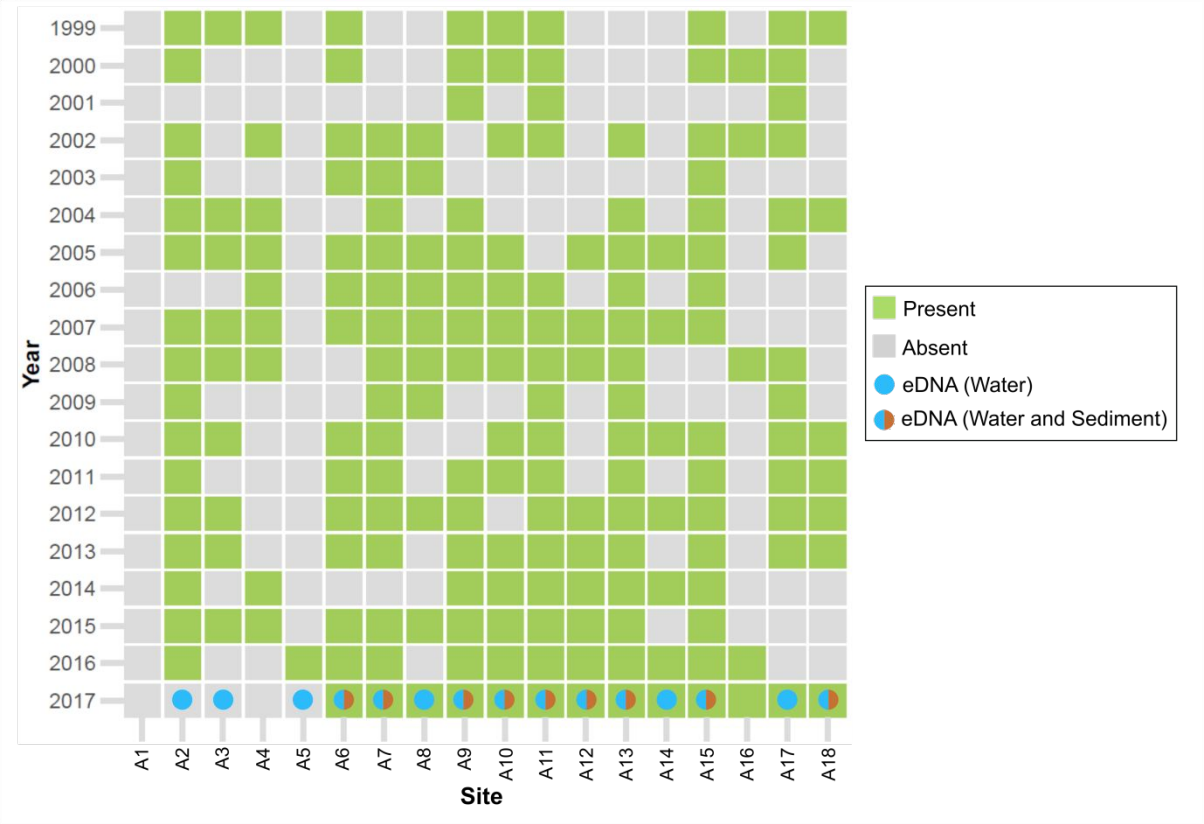

Figure S1. Presence and absence of water voles (*Arvicola amphibius*) from 1999-2017 using latrine surveys (X. Lambin, *unpublished data*) from sites A1-A18. Positive detections using environmental DNA (eDNA; water; and water and sediment) indicated in 2017.

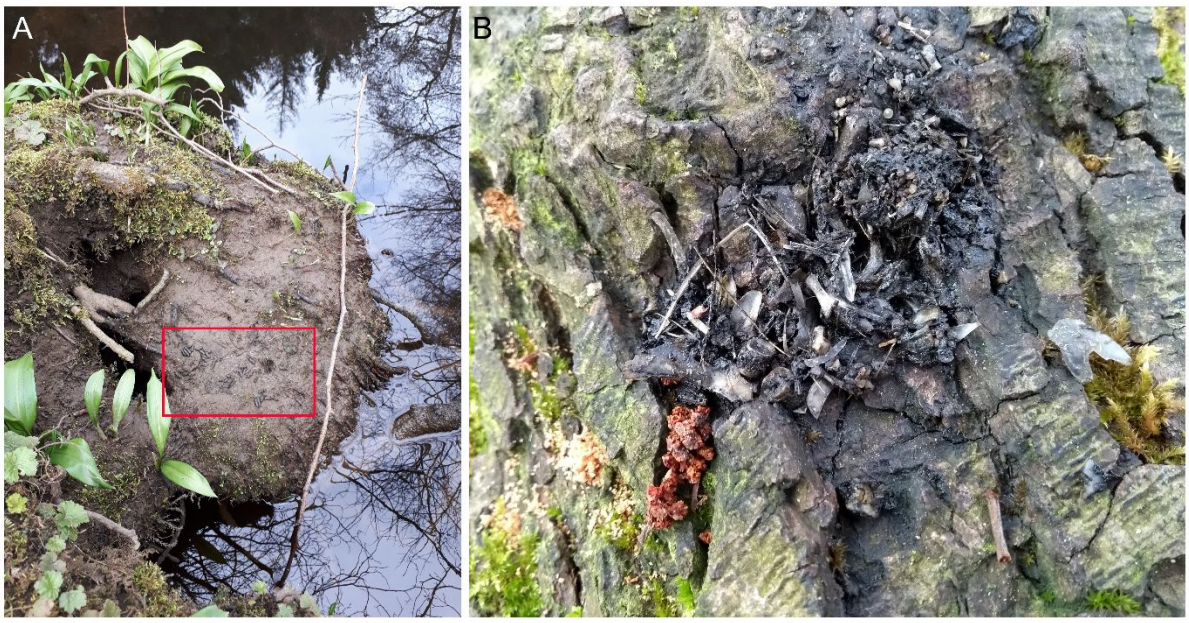

Figure S2. Example of a water vole latrine with faecal pellets, highlighted in the red rectangle in (A), and an otter spraint in (B). Both are from site P1 in the Peak District.

Figure S3. Examples of camera trap photographs for six species. Photographs have been manually adjusted to increase visibility of the species. Red boxes are used to highlight where the smaller mammals are positioned within the photograph. A: weasel (*Mustela nivalis*); B: red deer (*Cervus elaphus*); C: water vole (*Arvicola amphibius*); D: field vole (*Microtus agrestis*); E: Eurasian otter (*Lutra lutra*) and F: water shrew (*Neomys fodiens*).

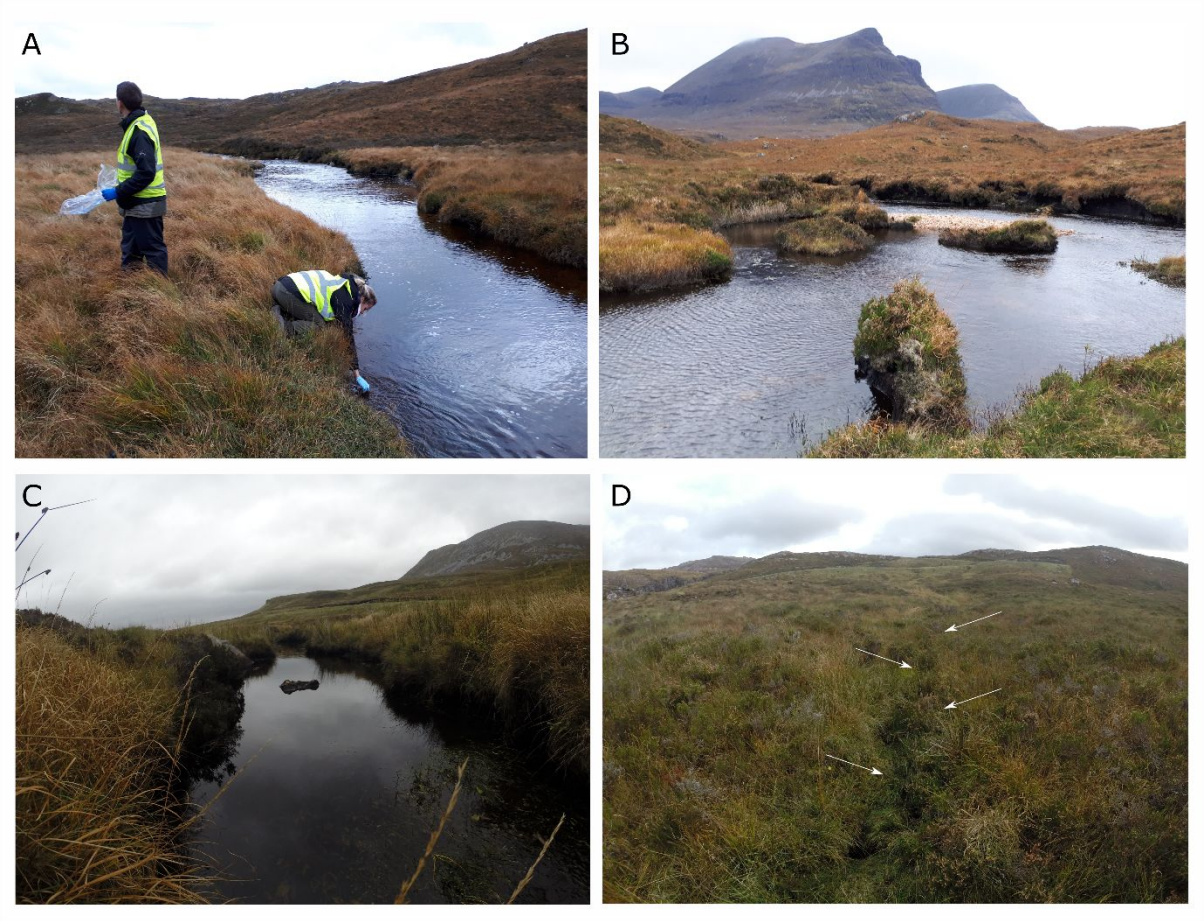

Figure S4. Examples of four sampling areas for environmental DNA (eDNA): A = A8; B = A12; C = A16 and D = A11. Sites A8, A11 and A12 returned positive eDNA records for the water vole, site A16 was negative. Sampling at site A11 was conducted in a narrow stream that is not visible here but is indicated by the white arrows (D). Sampling methodology for eDNA is indicated in (A), where sampling was conducted along the edge of the river/stream for both water and sediment samples.

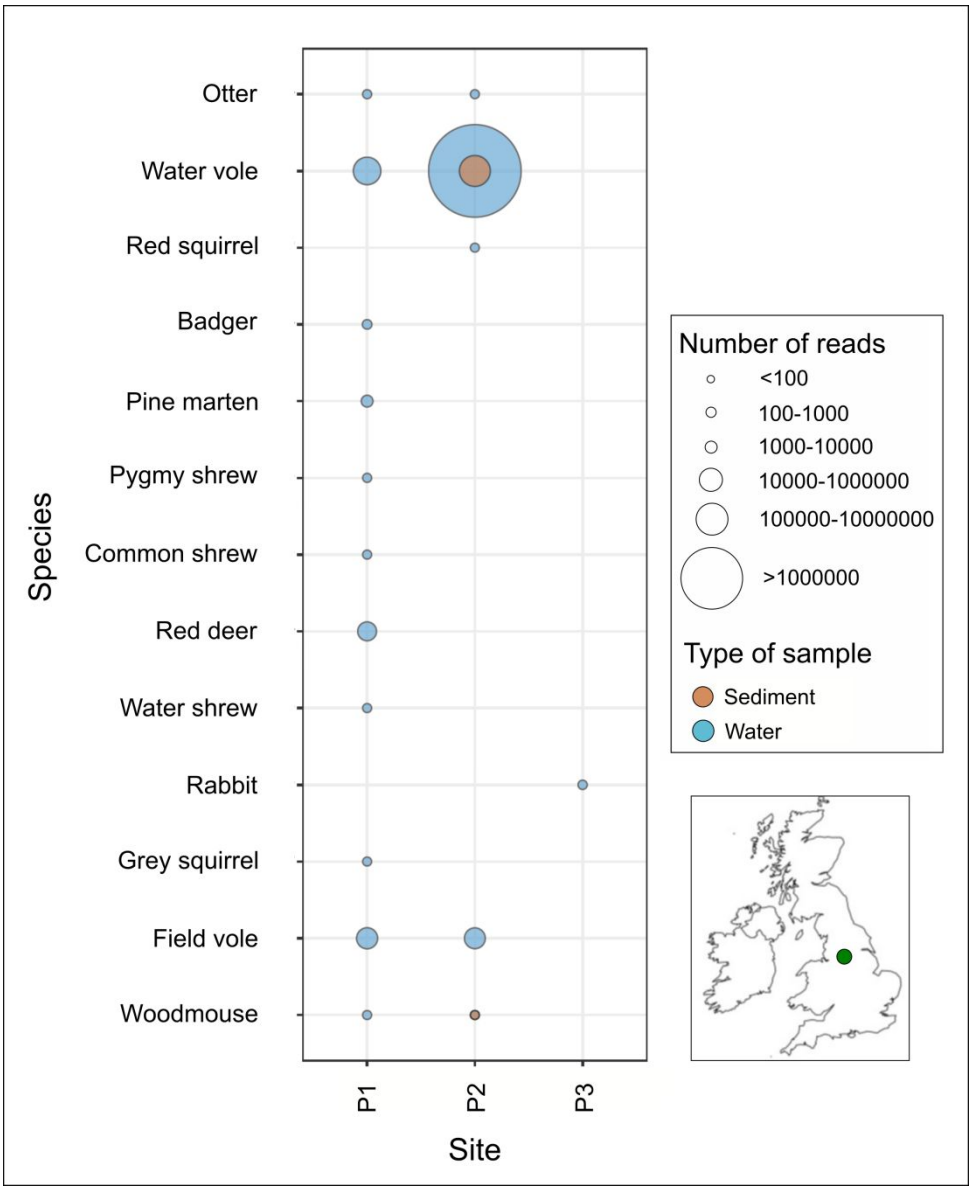

Figure S5: A bubble graph representing presence-absence and categorical values of the number of reads retained (after bioinformatic filtering) for eDNA (water in blue and sediment in orange) from each wild mammal identified in each site (P1-P3) in the Peak District National Park. The location of the Peak District is indicated in the inset map but the actual sampling sites can not be disclosed due to conservation and persecution concerns around certain protected species.

References:

Alston, D., Mallon, D.P., & Whiteley, D. (2012). *The Mammals of Derbyshire*. Derbyshire Mammal Group and Sorby Natural History Society. England.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac., E. (2016). obitools: A unix inspired software package for DNA metabarcoding. Molecular Ecology Resources, 16, 176–182.

Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., … Brosse, S. (2019). Unlocking biodiversity and conservation studies in high diversity environments using environmental DNA (eDNA): a test with Guianese freshwater fishes. Molecular Ecology Resources, 19, 27-46.

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27, 2194–2200.

Ficetola, G. F., Miaud, C., Pompanon, P., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. Biology Letters, 4, 423–425.

Hänfling, B., Handley, L. 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.

Majaneva, M., Diserud, O. H., Eagle, S. H. C., Boström, E., Hajibabaei, M., & Ekrem, T. (2018). Environmental DNA filtration techniques affect recovered biodiversity. Scientific Reports, 8, 4682.

Mathews, F., Kubasiewicz, L. M., Gurnell, J., Harrower, C. A., McDonald, R. A., & Shore, R. F. (2018). A Review of the Population and Conservation Status of British Mammals: Technical Summary. Natural England. Peterborough.

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., … Iwasaki, W. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes. Royal Society Open Science, 2, 150088

Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahe, F. (2016). VSEARCH: a versatile open source tool for metagenomics. PeerJ, 4, e2584.

Schnell, I. B., Bohmann, K., & Gilbert, M. T. P. (2015). Tag jumps illuminated - reducing sequence to sample misidentifications in metabarcoding studies. Molecular Ecology Resources, 15, 1289–1303.

Ushio, M., H. Fukuda, T. Inoue, K. Makoto, O. Kishida, K. Sato, … Miya, M. (2017). Environmental DNA enables detection of terrestrial mammals from forest pond water. Molecular Ecology Resources, 17, e63-e75.