Environmental DNA (eDNA) metabarcoding of pond water as a tool to survey conservation and management priority mammals

Lynsey R. Harper[a,b*], Lori Lawson Handley[a], Angus I. Carpenter[c], Muhammad Ghazali[d], Cristina Di Muri[a], Callum J. Macgregor[e], Thomas W. Logan[a], Alan Law[f], Thomas Breithaupt[a], Daniel S. Read[g], Allan D. McDevitt[h], and Bernd Hänfling[a]

[a] Department of Biological and Marine Sciences, University of Hull, Hull, HU6 7RX, UK

[b] Illinois Natural History Survey, Prairie Research Institute, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

[c] Wildwood Trust, Canterbury Rd, Herne Common, Herne Bay, CT6 7LQ, UK

[d] The Royal Zoological Society of Scotland, Edinburgh Zoo, 134 Corstorphine Road, Edinburgh, EH12 6TS, UK

[e] Department of Biology, University of York, Wentworth Way, York, YO10 5DD, UK

[f] Biological and Environmental Sciences, University of Stirling, Stirling, FK9 4LA, UK

[g] Centre for Ecology & Hydrology (CEH), Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, UK

[h] Ecosystems and Environment Research Centre, School of Science, Engineering and Environment, University of Salford, Salford, M5 4WT, UK

***Corresponding author:** lynsey.harper2@gmail.com

Lynsey Harper, Illinois Natural History Survey, Prairie Research Institute, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

©2019, Elsevier. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/

24 ## Abstract

25

26 Environmental DNA (eDNA) metabarcoding can identify terrestrial taxa utilising aquatic habitats
27 alongside aquatic communities, but terrestrial species' eDNA dynamics are understudied. We
28 evaluated eDNA metabarcoding for monitoring semi-aquatic and terrestrial mammals,
29 specifically nine species of conservation or management concern, and examined
30 spatiotemporal variation in mammal eDNA signals. We hypothesised eDNA signals would be
31 stronger for semi-aquatic than terrestrial mammals, and at sites where individuals exhibited
32 behaviours. In captivity, we sampled waterbodies at points where behaviours were observed
33 ('directed' sampling) and at equidistant intervals along the shoreline ('stratified' sampling). We
34 surveyed natural ponds (N = 6) where focal species were present using stratified water
35 sampling, camera traps, and field signs. eDNA samples were metabarcoded using vertebrate-
36 specific primers. All focal species were detected in captivity. eDNA signal strength did not differ
37 between directed and stratified samples across or within species, between semi-aquatic or
38 terrestrial species, or according to behaviours. eDNA was evenly distributed in artificial
39 waterbodies, but unevenly distributed in natural ponds. Survey methods deployed at natural
40 ponds shared three species detections. Metabarcoding missed badger and red fox recorded by
41 cameras and field signs, but detected small mammals these tools overlooked, e.g. water vole.
42 Terrestrial mammal eDNA signals were weaker and detected less frequently than semi-aquatic
43 mammal eDNA signals. eDNA metabarcoding could enhance mammal monitoring through
44 large-scale, multi-species distribution assessment for priority and difficult to survey species, and
45 provide early indication of range expansions or contractions. However, eDNA surveys need high

spatiotemporal resolution and metabarcoding biases require further investigation before routine implementation.

Key-words: camera traps, field signs, lentic, monitoring, semi-aquatic mammals, terrestrial mammals

52 ## 1. Introduction

53

54 Mammals are a highly threatened taxon, with 25% of species at risk of extinction globally due
55 to harvesting, habitat degradation/loss, non-native species or perception as pests (Visconti et
56 al., 2011). Most species lack long-term, systematic monitoring, with survey efforts biased
57 towards rare species (Massimino, Harris, & Gillings, 2018). Data deficiency prevents robust
58 estimation of mammalian range expansions/declines and population trends (Bland, Collen,
59 Orme, & Bielby, 2015). Therefore, effective and evidence-based strategies for mammal
60 conservation and management are urgently needed (Mathews et al., 2018).

61 Many mammals are nocturnal and elusive thus monitoring requires non-invasive,
62 observational methods such as camera traps and field signs, e.g. footprints, scat (Caravaggi et
63 al., 2018; Harris & Yalden, 2004; Kinoshita et al., 2019; Sadlier, Webbon, Baker, & Harris, 2004).
64 Camera trapping is cost-efficient, standardised, reproducible, and produces data suited to site
65 occupancy modelling, but only surveys a fraction of large, heterogeneous landscapes. Trap
66 placement can substantially influence species detection probabilities, and traps often miss
67 small species (Burton et al., 2015; Caravaggi et al., 2018; Ishige et al., 2017; Leempoel, Hebert,
68 & Hadly, 2019). Field sign surveys are inexpensive, but resource-intensive for broad geographic
69 coverage (Kinoshita et al., 2019; Sadlier et al., 2004). Species can have similar footprints and
70 scat, increasing the potential for misidentification (Franklin et al., 2019; Harris & Yalden, 2004).
71 Mammal survey methods can be species-specific, thus multiple methods are necessary for
72 large-scale, multi-species monitoring schemes (Massimino et al., 2018; Sales et al., 2019).

73 Environmental DNA (eDNA) analysis is a recognised tool for rapid, non-invasive, cost-

74 efficient biodiversity assessment across aquatic and terrestrial ecosystems (Deiner et al., 2017).
75 Organisms transfer genetic material to their environment via secretions, excretions, gametes,
76 blood, or decomposition, which can be isolated from environmental samples (Thomsen &
77 Willerslev, 2015). Studies using eDNA analysis to target specific semi-aquatic and terrestrial
78 mammals have employed PCR or quantitative PCR (qPCR) (e.g. Franklin et al., 2019; Lugg,
79 Griffiths, van Rooyen, Weeks, & Tingley, 2017; Rodgers & Mock, 2015; Thomsen et al., 2012;
80 Williams, Huyvaert, Vercauteren, Davis, & Piaggio, 2018). eDNA metabarcoding can screen
81 entire communities using PCR combined with high-throughput sequencing (Deiner et al., 2017;
82 Thomsen & Willerslev, 2015), but mammalian assessments are uncommon (Klymus, Richter,
83 Thompson, & Hinck, 2017; Kinoshita et al., 2019; Leempoel et al., 2019; Sales et al., 2019; Ushio
84 et al., 2017). Tropical mammal assemblages have been obtained by metabarcoding invertebrate
85 blood meals (e.g. Tessler et al., 2018) and salt licks (Ishige et al., 2017), but samples from the
86 physical environment have tremendous potential to reveal mammal biodiversity over broad
87 spatiotemporal scales (Sales et al., 2019; Ushio et al., 2017).

88 In aquatic ecosystems, eDNA metabarcoding has predominantly been applied to
89 characterise fish (e.g. Evans et al., 2017; Hänfling et al., 2016; Lawson Handley et al., 2018;
90 Valentini et al., 2016) and amphibian (e.g. Bálint et al., 2018; Valentini et al., 2016)
91 communities. However, mammals also leave eDNA signatures in water that metabarcoding can
92 detect (Harper et al., 2019; Klymus et al., 2017; Sales et al., 2019; Ushio et al., 2017). Ponds in
93 particular provide drinking, foraging, dispersive, and reproductive opportunities for semi-
94 aquatic and terrestrial mammals (Klymus et al., 2017). Samples from these waterbodies could
95 uncover biodiversity present in the wider environment (Deiner et al., 2017; Harper et al., 2019).

96 Drinking is a major source of eDNA deposition due to the release of saliva, but mammals may
97 also swim, wallow, urinate or defecate in water (Rodgers & Mock, 2015; Ushio et al., 2017;
98 Williams et al., 2018). Furthermore, arboreal mammals may use ponds less than semi-aquatic
99 and ground-dwelling species, non-territorial mammals may visit ponds less than territorial
100 species, and group-living species may deposit more eDNA than solitary species (Williams et al.,
101 2018). Despite evidence for eDNA deposition by semi-aquatic and terrestrial mammals in
102 freshwater ecosystems, little is known about the influence of mammal behaviour on the
103 distribution and strength of the eDNA signal left behind (defined here as proportional read
104 counts).

105 In this study, we conducted two experiments under artificial and natural conditions to
106 evaluate eDNA metabarcoding of pond water as a tool for monitoring semi-aquatic, ground-
107 dwelling, and arboreal mammals of conservation or management concern. The first
108 experiment, carried out on nine focal species housed at two wildlife parks, examined the role of
109 sampling strategy, mammal lifestyle, and mammal behaviour on eDNA detection and signal
110 strength under artificial conditions. Mammal eDNA detection is expected from enclosure water
111 that is frequently used by individuals for drinking, swimming and bathing. We hypothesised
112 that: (1) eDNA would be unevenly distributed, thus directed sampling would yield stronger
113 eDNA signals (i.e. higher proportional read counts) for mammals than stratified sampling; (2)
114 semi-aquatic mammals would have stronger eDNA signals than ground-dwelling or arboreal
115 mammals; and (3) mammal behaviours involving water contact would generate stronger eDNA
116 signals. The second experiment validated eDNA metabarcoding against camera trapping and
117 field sign searches for mammal identification at natural ponds, and investigated spatiotemporal

118 variation in mammal eDNA signals. Mammal eDNA detection is unpredictable at natural
119 waterbodies that can be extensive, subject to environmental fluctuations, and used rarely or
120 not at all by individuals. We hypothesised that: (1) eDNA metabarcoding would detect more
121 mammals than camera trapping or field signs; (2) semi-aquatic mammals would be readily
122 detected and their eDNA evenly distributed in ponds in comparison to terrestrial mammals; and
123 (3) temporal sampling would reveal that terrestrial mammal eDNA is detectable for short
124 periods in comparison to fully aquatic vertebrates.

125

126

127 2. Materials and methods

128

129 2.1 Study species

130

131 We studied nine mammal species that are the focus of European conservation or management
132 (Mathews et al., 2018): European water vole (*Arvicola amphibius*), European otter (*Lutra lutra*),
133 Eurasian beaver (*Castor fiber*), European hedgehog (*Erinaceus europaeus*), European badger
134 (*Meles meles*), red deer (*Cervus elaphus*), Eurasian lynx (*Lynx lynx*), red squirrel (*Sciurus*
135 *vulgaris*), and European pine marten (*Martes martes*). Water vole, otter, red squirrel, pine
136 marten and hedgehog are UK Biodiversity Action Plan species (Joint Nature Conservation
137 Committee, 2018). Water vole, otter, and beaver are semi-aquatic, red squirrel and pine
138 marten are arboreal, and the other species are ground-dwelling. Badger and red deer live in
139 groups whereas the other species are predominantly solitary.

140

141 ## 2.2 Experiment 1: eDNA detection and signal strength in artificial systems

142

143 Behavioural observation and eDNA sampling were conducted between 18th – 21st September
144 2017 at Wildwood Trust (WT), Kent, England, and 10th – 11th October 2017 at Royal Zoological
145 Society of Scotland (RZSS) Highland Wildlife Park (HWP), Kingussie, Scotland. Sixteen categories
146 of behaviour were defined based on potential contact with waterbodies and species lifestyle,
147 and the frequency and duration of behaviours recorded (Table 1, Appendix A: Table A1). The
148 number of individuals in each enclosure was recorded alongside waterbody size (Table 2).
149 Beaver, lynx, red deer, and red squirrel were present at both wildlife parks, whereas other
150 captive species were only present at WT. Each species was observed for one hour on two
151 separate occasions except nocturnal mammals (badger and beaver), which were observed
152 overnight using camera traps (Bushnell Trophy Cam Standard, Bushnell Corporation, KS, USA).
153 One camera trap per enclosure was positioned perpendicular to the ground (1 m height, 2 m
154 from shoreline) to capture water and shoreline. Cameras took 30 s videos (1920 x 1080) when
155 triggered (30 s interval between triggers) at high sensitivity. Behavioural observation was not
156 undertaken for WT water voles as animals were under quarantine or HWP red squirrels as
157 individuals were wild. Photos of waterbodies in animal enclosures are provided in Appendix B.

158 Water samples were collected from enclosures within 3 hrs of the second behavioural
159 observation period. Up to six directed or stratified samples were collected, but sample number
160 varied by species according to waterbody size and observed behaviours (Tables A1, A2).
161 Enclosure drinking containers were also sampled and classed as ‘other’ samples. Bathing and

162 drinking bowls were sampled where enclosures contained no artificial waterbodies (WT water
163 vole, red squirrel, and hedgehog). The HWP beaver enclosure was empty for 24 hrs before
164 sampling. Water was sampled from a RZSS Edinburgh Zoo (EZ) enclosure containing beavers
165 and classed as 'other'. A sample was collected from a water bath in the HWP woods to capture
166 wild red squirrels and classed as 'other'.

167 Directed samples (2 L surface water taken approximately where behaviours were
168 observed) were collected before stratified samples (2 L surface water [8 x 250 ml pooled
169 subsamples] taken at equidistant points [access permitting] around the waterbody perimeter)
170 to minimise disturbance to the water column and cross-contamination risk. Samples were
171 collected using sterile Gosselin™ HDPE plastic bottles (Fisher Scientific UK Ltd, UK) and
172 disposable gloves. A field blank (1 L molecular grade water [MGW]) was taken into each species
173 enclosure, opened, and closed before artificial water sources were sampled. Samples (n = 80)
174 collected from WT and HWP were transported alongside field blanks (n = 13) in sterile
175 coolboxes with ice packs to the University of Kent (UoK) and EZ respectively, where ice was
176 added to coolboxes.

177 Samples and blanks were vacuum-filtered within 6 hrs of collection in a UoK wet
178 laboratory and within 24 hrs of collection in an EZ staff room. Surfaces and equipment were
179 sterilised before, during, and after set-up in temporary work areas. Surfaces and vacuum
180 pumps were wiped with 10% v/v chlorine-based commercial bleach (Elliott Hygiene Ltd, UK)
181 solution. Non-electrical equipment was immersed in 10% bleach solution for 10 minutes,
182 followed by 5% v/v MicroSol detergent (Anachem, UK), and rinsed with purified water. Up to
183 500 ml of each 2 L sample was vacuum-filtered through sterile 0.45 µm mixed cellulose ester

184 membrane filters with pads (47 mm diameter; Whatman, GE Healthcare, UK) using Nalgene™
185 filtration units. One hour was allowed for each sample to filter and a second filter used if
186 clogging occurred. A filtration blank (1 L MGW) was processed during each filtration round (*n* =
187 12), and equipment sterilised after each filtration round. After 500 ml had filtered or one hour
188 had passed, filters were removed from pads using sterile tweezers, placed in sterile 47 mm
189 petri dishes (Fisher Scientific UK Ltd, UK), sealed with parafilm (Sigma-Aldrich Company Ltd,
190 UK), and stored at -20 °C. The total water volume filtered per sample was recorded for
191 downstream analysis (Table A2; Fig. A1).
192

193 2.3 Experiment 2: eDNA detection and signal strength in natural systems

194

195 At three sites where focal species were present based on cumulative survey data, we selected
196 two ponds (range 293-5056 m^2, average 1471 m^2) within 4 km of each other. The Bamff Estate
197 (BE), Alyth, Scotland, was selected for beaver, otter, badger, red deer, and red squirrel, but roe
198 deer (*Capreolus capreolus*) and red fox (*Vulpes vulpes*) were also present. Otter, water vole, and
199 badger were present at Tophill Low Nature Reserve (TLNR), Driffield, East Yorkshire, alongside
200 American mink (*Neovison vison*), stoat (*Mustela erminea*), weasel (*Mustela nivalis*), rabbit
201 (*Oryctolagus cuniculus*), brown hare (*Lepus europaeus*), red fox, roe deer, and grey squirrel
202 (*Sciurus carolinensis*). We selected Thorne Moors (TM), Doncaster, South Yorkshire, for red deer
203 and badger, but stoat, weasel, red fox, roe deer, and Reeve's muntjac (*Muntiacus reevesi*) were
204 also present. Camera traps (Bushnell Trophy Cam Standard/Aggressor, Bushnell Corporation,
205 KS, USA) were deployed at TM (one per pond) and BE (three per pond) one week prior to eDNA

206 sampling and collected once sampling was completed. At TLNR, camera traps (two to three per
207 pond) were deployed one day before a 5-day period of eDNA sampling and collected one week
208 after sampling was completed. Camera traps were positioned perpendicular to the ground (1 m
209 height, 0.3-1 m from shoreline) to capture water and shoreline. Cameras took three
210 photographs (5 megapixel) when triggered (3 s interval between triggers) at high sensitivity.

211 Ten stratified samples were collected from the shoreline of each pond (TM: 17th April
212 2018; BE: 20th April 2018; TLNR: 23rd – 27th April 2018) and a field blank (1 L MGW) included as
213 in Experiment 1. TLNR ponds were sampled every 24 hrs over 5 days to investigate
214 spatiotemporal variation in mammal eDNA signals. TM and TLNR samples were transported on
215 ice in sterile coolboxes to the University of Hull (UoH) eDNA facility, and stored at 4 °C. BE
216 samples were transported in sterile coolboxes with ice packs to BE accommodation. Surfaces
217 and equipment were sterilised before, during, and after set-up as in Experiment 1. Samples (n =
218 140) and field blanks (n = 14) were vacuum-filtered within 4 hrs of collection as in Experiment 1
219 with minor modifications to maximise detection probability as follows. The full 2 L of each
220 sample was vacuum-filtered where possible, two filters were used for each sample, and
221 duplicate filters were stored in one petri dish at -20 °C. A filtration blank (1 L MGW) was
222 processed during each filtration round (n = 21). The total water volume filtered per sample was
223 recorded (Table A3).

224

225 ## 2.4 DNA extraction

226

227 DNA was extracted within 2 weeks of filtration at the UoH eDNA facility using the Mu-DNA
228 water protocol (Sellers, Di Muri, Gómez, & Hänfling, 2018). The full protocol is available at:
229 https://doi.org/10.17504/protocols.io.qn9dvh6. Duplicate filters from samples in Experiment 1
230 were lysed independently and the lysate from each loaded onto one spin column. As more
231 samples were collected in Experiment 2, duplicate filters were co-extracted by placing both in a
232 single tube for bead milling. An extraction blank, consisting only of extraction buffers, was
233 included for each round of DNA extraction (*n* = 17). Eluted DNA (100 µl) was stored at -20 °C
234 until PCR amplification.

235

236 ## 2.5 eDNA metabarcoding

237

238 Our eDNA metabarcoding workflow is fully described in Appendix A. Briefly, we performed
239 nested metabarcoding using a two-step PCR protocol, where Multiplex Identification (MID) tags
240 were included in the first and second PCR for sample identification (Kitson et al., 2019). The first
241 PCR amplified eDNA in triplicate with published 12S ribosomal RNA (rRNA) primers 12S-V5-F (5'-
242 ACTGGGATTAGATACCCC-3') and 12S-V5-R (5'-TAGAACAGGCTCCTCTAG-3') (Riaz et al., 2011).
243 Harper et al. (2018) validated these primers *in silico* for UK vertebrates, and found 91/112
244 mammal species listed on the Natural History Museum Checklist of Mammalia v1 (subspecies
245 excluded) could be distinguished. Nine indistinguishable species lacked reference sequences,
246 whereas 12 had reference sequences but did not amplify. PCR positive controls (two per PCR

247 plate; *n* = 16) were exotic cichlid (*Maylandia zebra*) DNA (0.05 ng/µl), and PCR negative controls

248 (two per PCR plate; *n* = 16) were MGW (Fisher Scientific UK Ltd, UK). PCR products were pooled

249 to create sub-libraries (Fig. A2) and purified with Mag-BIND® RxnPure Plus magnetic beads

250 (Omega Bio-tek Inc, GA, USA), following the double size selection protocol established by

251 Bronner et al. (2009). Ratios of 0.9x and 0.15x magnetic beads to 100 µL of each sub-library

252 were used. Eluted DNA (30 µL) was stored at -20 °C until the second PCR could be performed.

253 The second PCR bound pre-adapters, MID tags, and Illumina adapters to the sub-libraries. PCR

254 products were purified with Mag-BIND® RxnPure Plus magnetic beads (Omega Bio-tek Inc, GA,

255 USA), following the double size selection protocol established by Bronner et al. (2009). Ratios of

256 0.7x and 0.15x magnetic beads to 50 µL of each sub-library were used. Eluted DNA (30 µL) was

257 stored at 4 °C until quantification and normalisation. The library was purified again, quantified

258 by qPCR using the NEBNext® Library Quant Kit for Illumina® (New England Biolabs® Inc., MA,

259 USA), and fragment size (330 bp) and removal of secondary product verified using an Agilent

260 2200 TapeStation and High Sensitivity D1000 ScreenTape (Agilent Technologies, CA, USA). The

261 library (220 eDNA samples, 27 field blanks, 33 filtration blanks, 17 extraction blanks, 16 PCR

262 negative controls, and 16 PCR positive controls) was sequenced on an Illumina MiSeq® using a

263 MiSeq Reagent Kit v3 (600-cycle) (Illumina, Inc, CA, USA). Raw sequence reads were

264 demultiplexed using a custom Python script. metaBEAT v0.97.11 (https://github.com/HullUni-

265 bioinformatics/metaBEAT) was used for quality trimming, merging, chimera removal, clustering,

266 and taxonomic assignment of sequences against our UK vertebrate reference database (Harper

267 et al., 2018) which contains sequences for 103 UK mammals. Taxonomic assignment used a

268 lowest common ancestor approach based on the top 10% BLAST matches for any query that

269 matched a reference sequence across more than 80% of its length at minimum identity of 98%.

270

271 **2.6 Data analysis**

272

273 Analyses were performed in R v.3.4.3 (R Core Team, 2017). The total unrefined read counts (i.e.
274 raw taxonomically assigned reads) per sample were calculated and retained for downstream
275 analyses. Assignments were corrected: family and genera containing a single UK species were
276 reassigned to that species, species were reassigned to domestic subspecies, and
277 misassignments were corrected, e.g. *Lynx pardinus* and *Lynx lynx*. Manual reassignment
278 duplicated some metaBEAT assignments thus the read count data for these assignments were
279 merged. Taxon-specific sequence thresholds (i.e. maximum sequence frequency of each taxon
280 in PCR positive controls) were used to mitigate cross-contamination and false positives (Table
281 A4, Fig. A3), and remnant contaminants and higher taxonomic assignments removed excluding
282 the following genera. *Anas* (Dabbling ducks) was retained because potential for hybridisation
283 reduced confidence in species-level assignments, and *Emberiza* (Buntings) and *Larus* (White-
284 headed gulls) were retained because reference sequences were missing for several common
285 species. Dataset refinement is fully described in Appendix A. Taxonomic assignments remaining
286 in the refined dataset were predominantly of species resolution and considered true positives.
287 We split the refined dataset by Experiment 1 (artificial waterbodies) and Experiment 2 (natural
288 ponds). Proportional read counts for each species were calculated from the total unrefined
289 read counts per sample. Our proportional read count data were not normally distributed
290 (Shapiro–Wilk normality test: $W = 0.915$, $P < 0.001$), thus we used a Mann-Whitney U test to

291 compare the median proportional read count of stratified and directed samples across species.

292 We employed binomial Generalized Linear Mixed-effects Models (GLMMs) with the logit
293 link function using the package glmmTMB (development version; Brooks et al., 2017) for the
294 following tests. First, we compared the eDNA signals from stratified and directed samples for
295 each mammal species using a hierarchical model including sample type nested within species
296 (fixed) and wildlife park (random) as effects. We tested the influence of species lifestyle on
297 mammal eDNA signals using a model with species lifestyle (fixed) and species nested within
298 wildlife park (random) as effects. Using directed samples, we tested the influence of behaviour
299 on mammal eDNA signals using two hierarchical models, including species nested within
300 wildlife park (random) and specific (e.g. swimming, drinking) or generic (i.e. water contact
301 versus no water contact) behaviour(s) respectively (fixed) as effects. We assessed model fit
302 using diagnostic plots and performed validation checks to ensure model assumptions were met
303 and overdispersion was absent (Zuur, Ieno, Walker, Saveliev, & Smith, 2009).

304 For Experiment 2, we qualitatively compared mammal presence-absence records
305 generated by eDNA metabarcoding, camera trapping, and field signs. TLNR ponds were
306 sampled every 24 hrs for 5 days, thus proportional read counts were averaged across days for
307 comparison to BE and TM ponds (sampled once each). We qualitatively compared the
308 distribution and persistence of eDNA signals between semi-aquatic and terrestrial mammals
309 using tile plots and heat maps of the unaveraged proportional read counts for identified species
310 at TLNR over the 5-day period. All figures were produced using the package ggplot2 v3.0.0
311 (Wickham, 2016).

312

313

314 # 3. Results

315

316 ## 3.1 eDNA metabarcoding

317

318 The sequencing run generated 47,713,656 raw sequence reads, of which 37,590,828 remained
319 following trimming, merging, and length filter application. After removal of chimeras and
320 redundancy via clustering, the library contained 21,127,061 sequences (average read count of
321 64,215 per sample including controls), of which 16,787,750 (79.46%) were assigned a
322 taxonomic rank. Contamination (Fig. A4) was observed in the field blanks (badger, beaver, lynx,
323 pine marten, red squirrel, and water vole) as well as in the filtration and extraction blanks
324 (human [*Homo sapiens*] and cichlid). PCR negative controls were contaminated to different
325 extents with human, cichlid, beaver, and pine marten as well as non-focal species. After
326 threshold application, contaminants remaining in eDNA samples included Gentoo penguin
327 (*Pygoscelis papua*), reindeer (*Rangifer tarandus*), cichlid, and human. The refined dataset
328 contained 59 vertebrate species, including six amphibians, 10 fish, 19 birds, and 24 mammals
329 (Table A5).

330

331 ## 3.2 Experiment 1: eDNA detection and signal strength in artificial systems

332

333 All nine focal species were detected in captivity, of which seven were detected in all water

334 samples taken from their respective enclosures. HWP red deer were not detected in 2 of 5
335 stratified samples, and WT hedgehog was not detected in 1 of 2 drinking bowl samples (Fig. 1).
336 'Other' samples (neither directed nor stratified) were excluded from further comparisons, thus
337 hedgehog, red squirrel, and water vole were omitted in downstream analyses. Across species,
338 stratified samples (0.406) had a higher median proportional read count than directed samples
339 (0.373), but this difference was not significant (Mann-Whitney *U* test: U = 1181.5, P = 0.829).
340 Proportional read counts for directed and stratified samples did not significantly differ (χ^2_6 =
341 0.364, P = 0.999) within species either (Fig. 2a; GLMM: θ = 0.168, χ^2_{53} = 8.915, P = 1.000,
342 pseudo-R^2 = 39.21%). Otter proportional read counts were lower than other species, but not
343 significantly so. Similarly, species lifestyle (semi-aquatic, ground-dwelling, arboreal) did not
344 influence (χ^2_2 = 0.655, P = 0.721) proportional read counts (Fig. 2b; GLMM: θ = 0.213, χ^2_{61} =
345 13.002, P = 1.000, pseudo-R^2 = 11.85%). Proportional read counts did not differ (χ^2_{11} = 1.369, P
346 = 0.999) according to specific behaviours exhibited by species (Fig. 3a; GLMM: θ = 0.355, χ^2_{31} =
347 11.013, P = 0.999, pseudo-R^2 = 9.17%). Likewise, generic behaviour (i.e. water contact versus no
348 water contact) did not influence (χ^2_{11} = 0.002, P = 0.964) proportional read counts (Fig. 3b;
349 GLMM: θ = 0.217, χ^2_{41} = 8.897, P = 1.000, pseudo-R^2 = 8.50%).
350

351 **3.3 Experiment 2: eDNA detection and signal strength in natural systems**

352

353 At natural ponds, eDNA metabarcoding, camera trapping, and field signs all detected beaver,
354 red deer, and roe deer. Camera traps (Fig. 4) and field signs recorded red fox and badger when
355 eDNA metabarcoding did not (Fig. 5). However, eDNA metabarcoding revealed small mammals

356 missed by cameras and field signs, including water vole, water shrew (*Neomys fodiens*), bank
357 vole (*Myodes glareolus*), common shrew (*Sorex araneus*), brown rat (*Rattus norvegicus*), rabbit,
358 grey squirrel, and common pipistrelle (*Pipistrellus pipistrellus*). We observed mice or vole
359 footprints at BE Pond 1, but could not ascertain species. Fig. 5 summarises mammals recorded
360 by different methods at each site with reference to cumulative survey data. Notably, only
361 beaver was found at the same ponds by all methods. Although methods shared species at site
362 level, species were not always detected at the same pond. Detection rates for species captured
363 by at least one survey method are summarised in Table A6.

364 Sampling of natural ponds revealed spatial patterns in eDNA detection and signal
365 strength. eDNA from non-domestic terrestrial mammals (i.e. mammals excluding dog [*Canis*
366 *lupus familiaris*], pig [*Sus scrofa domesticus*], sheep [*Ovis aries*] and cow [*Bos taurus*]) was
367 unevenly dispersed compared with semi-aquatic mammals (Fig. A5). Semi-aquatic beaver and
368 water vole were detected in at least 90% and 60% respectively of water samples (n = 10)
369 collected from single ponds, albeit water shrew was only detected in 10% of samples. Non-
370 domestic terrestrial mammals were routinely detected in <20% of water samples collected from
371 a pond and left relatively weak eDNA signals. Overall, beaver was the most consistently
372 detected mammal with the highest proportional read counts. However, the strongest and most
373 evenly distributed signals belonged to amphibians, particularly common frog (*Rana temporaria*)
374 and great crested newt (*Triturus cristatus*) (Fig. A5).

375 TLNR samples collected over a 5-day period (D01-05) revealed that mammal detection
376 heavily depends on the spatial and temporal resolution of eDNA metabarcoding surveys (Fig.
377 A6). Mammal eDNA signals in pond water were ephemeral, often disappearing within 24-48 hrs

378 of initial detection, as opposed to amphibians that were detected for multiple days and whose
379 eDNA signal increased in strength. The majority of semi-aquatic or terrestrial mammals were
380 only detected in a single sample on each day.
381
382
383 # 4. Discussion
384
385 We have demonstrated the potential of eDNA metabarcoding for monitoring conservation and
386 management priority mammals, but species detection rates are variable. Our experiments have
387 validated this molecular approach and provided new insights that will inform the development
388 and application of mammal eDNA metabarcoding. Sampling strategy, mammal lifestyle, and
389 mammal behaviour did not influence eDNA detection and signal strength in captivity, but all
390 played vital roles in natural ponds. Although semi-aquatic and terrestrial mammals were
391 detected from pond water, their eDNA signals were temporary and weak in comparison to
392 aquatic amphibians and fishes. Nonetheless, this suggests that eDNA is representative of
393 contemporary and local mammal diversity.
394
395 ## 4.1 Influence of sampling strategy and mammal behaviour on eDNA detection
396
397 In Experiment 1, all nine focal species were detected in captivity, and seven were detected in all
398 water samples taken from their respective enclosures. This demonstrates that our method can
399 successfully detect a variety of mammals from pond and drinking water. Surprisingly, we found

400 that neither sampling strategy nor mammal lifestyle nor mammal behaviour influenced eDNA
401 detectability and signal strength in captivity. This included behaviours associated with eDNA
402 deposition, e.g. swimming, drinking, urination, and defecation (Rodgers & Mock, 2015; Ushio et
403 al., 2017; Williams et al., 2018). Enclosures were permanently occupied and artificial
404 waterbodies likely saturated with eDNA, which possibly masked behavioural signals. Modest
405 replication may have limited experimental power, preventing patterns being detected
406 statistically. Nonetheless, our results show that mammal contact with water enables eDNA
407 deposition and detection.

408 Unsurprisingly, given the nature of wild mammal interactions with natural systems
409 versus those in captivity, Experiment 2 results highlight the challenges of mammal eDNA
410 detection. We recorded 17 mammals using three monitoring tools, comparable to the 17
411 mammals expected from cumulative survey data despite discordance. Field signs and camera
412 trapping detected red fox and badger where eDNA metabarcoding did not, but eDNA
413 metabarcoding identified water vole and other small mammals missed on camera or with
414 ambiguous field signs, i.e. mice, voles, shrews. Importantly, camera trap deployment period,
415 height, and positioning may have influenced small mammal detection by this method
416 (Caravaggi et al., 2018). Ishige et al. (2017) achieved comparable mammal detection at salt licks
417 with eDNA metabarcoding and camera trapping, but species presence was inconsistent
418 between salt licks surveyed. Using multi-species occupancy modelling for three mammal
419 species, Sales et al. (2019) observed water-based eDNA metabarcoding provided comparable
420 detection probabilities to conventional survey methods and actually outperformed camera
421 trapping. Similarly, Leempoel et al. (2019) found soil-based eDNA metabarcoding identified the

422 same mammals as camera trapping as well as small mammals rarely seen on camera, albeit the
423 methods differed between sites. Our own results echo all three studies, where despite some
424 inconsistencies, eDNA metabarcoding enhanced species inventories and identified smaller,
425 cryptic taxa.

426 Notably, no survey method captured semi-aquatic otter despite presence at study sites
427 and successful detection in eDNA metabarcoding studies of UK ponds (Harper et al., 2019),
428 lakes (Hänfling et al., 2017), and rivers/streams (Sales et al., 2019). Captive otter also had a
429 weaker eDNA signal than other semi-aquatic mammals studied here. Lower eDNA detection
430 rates for otter, badger, and red fox may stem from species' ecologies (Sales et al., 2019). These
431 mammals are wide-ranging (Gaughran et al., 2018; Thomsen et al., 2012) and may not readily
432 release DNA in water. Otters often spraint on grass or rock substrata outside water and use
433 latrines associated with caves and dens (Ruiz-Olmo & Gosálbez, 1997). As terrestrial mammals,
434 red fox and badger must drink from or enter ponds for eDNA deposition to occur (Rodgers &
435 Mock, 2015; Ushio et al., 2017; Williams et al., 2018). Otter, badger, and red fox detection may
436 require greater spatiotemporal resolution of eDNA sampling. This is reinforced by other eDNA
437 metabarcoding studies where mammal detection was highly variable across sites surveyed
438 (Ishige et al., 2017; Klymus et al., 2017; Leempoel et al., 2019; Sales et al., 2019; Ushio et al.,
439 2017). False negatives may instead be symptomatic of metabarcoding bias, but this is unlikely in
440 our study (section 4.2).

441 eDNA from other semi-aquatic mammals was evenly distributed, being found in most or
442 all samples collected on fine spatial scales within natural ponds, whereas terrestrial mammal
443 eDNA was highly localised and detected in few (<20%) samples. Mammal eDNA signals varied

444 temporally, being detectable for two consecutive days maximum. Depending on the species,
445 mammal eDNA may be spatially and temporally clumped in lentic ecosystems due to the nature
446 and frequency of water contact. Unless non-domestic mammals exhibit behaviours involving
447 prolonged water contact (e.g. swimming, wallowing), they may only be detected at drinking
448 sites (Klymus et al., 2017; Ushio et al., 2017; Williams et al., 2018). Conversely, domestic
449 mammals may have elevated detection rates in ponds due to high occurrence of these
450 waterbodies in agricultural landscapes as well as eDNA transport by rainfall and run-off (Staley
451 et al., 2018). eDNA detection and persistence are further influenced by group size, where eDNA
452 from multiple individuals endures for longer periods in water than eDNA from single individuals
453 (Williams et al., 2018). Detailed investigations incorporating biotic (e.g. population size, body
454 mass, behaviour) and abiotic (e.g. temperature, pH, rainfall) factors are needed to understand
455 the longevity of mammal eDNA signals in aquatic ecosystems (Rodgers & Mock, 2015; Sales et
456 al., 2019; Williams et al., 2018).

457 Our two experiments have shown that sampling strategy influences mammal eDNA
458 detection. Mammal eDNA was evenly distributed in closed, artificial waterbodies, but locally
459 distributed in open, natural ponds. Captive mammal enclosures contained one species
460 (excluding HWP red deer) and a drinking container(s) and/or small waterbody (range 0.01-162
461 m^2, mean 27.4 m^2). Some enclosures housed more individuals of a species than others, thereby
462 increasing eDNA deposition and detection probability (Williams et al., 2018). Wild mammals
463 have an array of freshwater habitats at their disposal and can hold vast territories. Therefore,
464 rates of pond visitation and eDNA deposition are more irregular (Klymus et al., 2017; Ushio et
465 al., 2017), possibly leading to between-sample variation (Williams et al., 2018).

466

467 ## 4.2 Accounting for false positives and false negatives in metabarcoding

468

469 eDNA metabarcoding has potential for inclusion in mammal monitoring schemes (section 4.3),
470 but like existing monitoring tools, may produce false negatives or false positives. Our process
471 controls identified low-level contamination at all stages of metabarcoding, but primarily during
472 sampling or PCR (Appendix A). We applied taxon-specific sequence thresholds to our data to
473 mitigate false positives as in Harper et al. (2019). Remnant contaminants were cichlid
474 (laboratory), Gentoo penguin (environment), reindeer (environment), and human
475 (environment/laboratory). Gentoo penguin is housed at EZ and was identified from EZ beaver
476 enclosure water. The WT red squirrel and reindeer enclosures are in close proximity. DNA
477 transport by wildlife (e.g. waterfowl [Hänfling et al., 2016]) and park staff/visitors may explain
478 this environmental contamination. Human DNA was present across process controls
479 corresponding to artificial and natural waterbodies. Human DNA may be amplified and
480 sequenced instead of focal species, potentially resulting in false negative detections for rare
481 and/or less abundant species. Human DNA blocking primers can prevent this bias, but may
482 impair PCR amplification efficiency (Klymus et al., 2017; Ushio et al., 2017; Valentini et al.,
483 2016). Sequence thresholds are one method of accounting for contamination in metabarcoding
484 datasets, but this is a topic that warrants deeper investigation aimed at researching and
485 refining standardised methods for false positive identification and mitigation, e.g. the R
486 package microDecon (McKnight et al., 2019).

487 In our study, eDNA metabarcoding produced false negatives for otter, badger, and red

488 fox at natural ponds. We selected a 12S metabarcode designed to amplify vertebrate DNA (Riaz
489 et al., 2011). One of four fox reference sequences (NCBI Accession: KF387633.1) possessed one
490 mismatch to the forward primer, and one of three otter reference sequences (NCBI Accession:
491 EF672696.1) possessed one mismatch to the reverse primer. These mismatches did not occur
492 within the first or last four bases of either primer sequence, and there were no primer
493 mismatches with the badger reference sequences (Harper et al., 2018). Therefore, amplification
494 bias was not responsible for these false negatives. DNA from aquatic and more abundant
495 species may have overwhelmed otter, badger, and red fox DNA during amplification and
496 sequencing, i.e. species-masking (Kelly, Port, Yamahara, & Crowder, 2014; Klymus et al., 2017).
497 Species-masking may also arise from use of proportional read counts as an index of eDNA signal
498 strength. High proportional read counts for a species may translate to a weak eDNA signal if the
499 total mammalian eDNA concentration is highly variable between samples or lower than the
500 total eDNA concentration for other taxonomic groups in a sample. Metabarcoding primers
501 targeting mammals (Ushio et al., 2017) or multi-marker (e.g. 12S, 16S, COI) investigations
502 (Evans et al., 2017; Hänfling et al., 2016; Kelly et al., 2014; Klymus et al., 2017) may improve
503 mammal detection in systems with competition from non-target aquatic species and where
504 total mammalian eDNA concentration varies between samples. Similarly, more biological and
505 technical replication may improve species detection probabilities (Evans et al., 2017; Lawson
506 Handley et al., 2019; Sales et al., 2019; Valentini et al., 2016). Importantly, otter also had lower
507 qPCR detection than amphibians and fish (Thomsen et al., 2012). A metabarcoding and qPCR
508 comparison (e.g. Harper et al., 2018; Lacoursière-Roussel, Dubois, Normandeau, & Bernatchez,
509 2016) would confirm whether poor amplification efficiency for otter arises from technical bias

510 or species ecology, and whether eDNA metabarcoding can reliably monitor otter alongside the
511 wider mammalian community.
512
513
4.3 Scope of eDNA metabarcoding for mammal monitoring
514
515 Mammal population assessments are hindered by lack of data and systematic monitoring for
516 many species (Mathews et al., 2018). Distribution and occupancy data are poor for most
517 species, with ongoing survey effort biased toward rare species. Surveys heavily rely on citizen
518 science and casual records (Massimino et al., 2018). Tools that provide standardised, systematic
519 monitoring of mammal populations are needed (Mathews et al., 2018). Despite issues inherent
520 to metabarcoding for biodiversity monitoring (Deiner et al., 2017), this tool has enormous
521 potential to enhance mammal monitoring, conservation, and management. eDNA
522 metabarcoding generates distribution data for multiple species, whether rare, invasive, or
523 abundant, and could track conflicting species simultaneously, e.g. water vole, American mink,
524 and otter (Bonesi & Macdonald, 2004) or red squirrel, grey squirrel, and pine marten (Sheehy,
525 Sutherland, O'Reilly, & Lambin, 2018).
526 eDNA metabarcoding can rapidly survey multitudes of aquatic sites at landscape-scale
527 where camera traps might be resource-intensive, cost-inefficient, and susceptible to
528 theft/damage (Ushio et al., 2017). Field signs require volunteer time and skill (Sadlier et al.,
529 2004) to be employed at comparable spatial scales to eDNA metabarcoding which could
530 provide accurate data for species misidentified from field signs, e.g. mice and voles, otter and
531 mink (Franklin et al., 2019; Harris & Yalden, 2004). However, camera traps and field signs both

532 recorded species that eDNA metabarcoding missed. Therefore, eDNA metabarcoding is
533 complementary and should be incorporated into, not replace, existing monitoring schemes
534 (Leempoel et al., 2019; Sales et al., 2019). This tool could be most effective in mammal
535 monitoring if deployed at the edges of known species distributions, in areas where species
536 presence is unknown, and in areas with isolated species records (Mathews et al., 2018).
537

538 **4.4 Recommendations for mammal survey using eDNA metabarcoding**

539

540 Water-based eDNA metabarcoding shows great promise for mammal monitoring encompassing
541 conservation and management priority species (Sales et al., 2019). However, there are factors
542 to be considered when designing and conducting mammal eDNA surveys that may not be
543 problematic for surveys of fishes or amphibians. Mammal eDNA detection probabilities from
544 natural ponds will likely be high when areas with dense populations are studied, but rigorous
545 sampling strategies will be required to track mammals in areas sparsely populated by
546 individuals. Multiple ponds must be sampled repeatedly, and samples taken at multiple
547 locations within ponds without pooling to enable site occupancy inferences. Importantly, we
548 sampled natural ponds in spring but sampling in other seasons may produce different results,
549 reflective of species' ecologies (Lawson Handley et al., 2019). To account for differential
550 mammal visitation rates and maximise eDNA detection probabilities, we recommend that
551 researchers and practitioners using eDNA metabarcoding for mammal monitoring channel their
552 efforts into extensive sampling of numerous waterbodies in a given area over prolonged
553 timescales. Water-based eDNA appears to be indicative of contemporary mammal presence,

with most mammal eDNA signals lost within 1-2 days. Therefore, eDNA metabarcoding could provide valuable mammalian community "snapshots" that may not be obtained with other survey methods (Ushio et al., 2017). Different sample types (e.g. water, soil, snow, salt licks, feeding traces, faeces, hair, and blood meals) may also offer new insights to mammal biodiversity (Franklin et al., 2019; Ishige et al., 2017; Kinoshita et al., 2019; Leempoel et al., 2019; Sales et al., 2019; Tessler et al., 2018; Ushio et al., 2017).

Data accessibility

Raw sequence reads have been archived on the NCBI Sequence Read Archive (Study: SRP164740; BioProject: PRJNA495011; BioSamples: SAMN10195928 - SAMN10196255; SRA accessions: SRR7986451 - SRR7986778). Jupyter notebooks, R scripts and corresponding data are archived online (https://doi.org/10.5281/zenodo.2561415).

Acknowledgements

The authors declare that they have no conflicts of interest in presenting this work for publication. This work was funded by the University of Hull, and D.S.R was supported by the Natural Environment Research Council award number NE/R016429/1 as part of the UK-SCAPE

575 programme delivering National Capability. We would like to thank Gill Murray-Dickson
576 (National Museums Scotland) and Helen Senn (RZSS Edinburgh Zoo) for feedback on the study
577 design, and Richard Griffiths (DICE, University of Kent) for support with filtration of water
578 samples from Wildwood Trust. We are grateful to staff at Wildwood Trust and RZSS Highland
579 Wildlife Park for assisting with eDNA sampling from animal enclosures. We thank Richard
580 Hampshire and volunteers at Tophill Low Nature Reserve for assisting with camera trap
581 deployment. Tim Kohler (Natural England) and Paul Ramsay kindly gave permission to sample at
582 Thorne Moors and the Bamff Estate respectively.
583
584
585
Author contributions
586
587 L.R.H, B.H, and L.L.H conceived and designed the study. A.I.C and M.G coordinated sampling at
588 Wildwood Trust and RZSS Highland Wildlife Park respectively. L.R.H, C.D.M, C.J.M, and T.L
589 collected and filtered water samples. A.L, T.L, and T.B helped select natural ponds to be
590 surveyed using eDNA, camera trapping, and field signs, and provided camera traps for the
591 study. L.R.H, A.L, and T.L deployed camera traps, which were then collected and footage
592 analysed by L.R.H. L.R.H processed samples in the laboratory with advice from C.D.M and A.M.
593 D.S.R sequenced the final library. L.R.H completed bioinformatic processing of samples, and
594 subsequent data analysis. L.R.H wrote the manuscript, which all authors contributed critically to
595 drafts of and gave final approval for publication.
596

597 ## References

598

599 Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2018) Scrutinizing key steps for
600 reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, *9*,
601 134–147. https://doi.org/10.1111/2041-210X.12849
602 Bálint, M., Nowak, C., Márton, O., Pauls, S. U., Wittwer, C., Aramayo, J. L., Schulze, A.,
603 Chambert, T., Cocchiararo, B., & Jansen, M. (2018). Accuracy, limitations and cost efficiency
604 of eDNA-based community survey in tropical frogs. *Molecular Ecology Resources*, *18*, 1415–
605 1426. https://doi.org/10.1111/1755-0998.12934
606 Bland, L. M., Collen, B., Orme, C. D. L., & Bielby, J. (2015). Predicting the conservation status of
607 data-deficient species: Predicting Extinction Risk. *Conservation Biology*, *29*, 250–259.
608 https://doi.org/10.1111/cobi12372
609 Bonesi, L., & Macdonald, D. W. (2004). Differential habitat use promotes sustainable
610 coexistence between the specialist otter and the generalist mink. *Oikos*, *106*, 509–
611 519. https://doi.org/10.1111/j.0030-1299.2004.13034.x
612 Bronner, I. F., Quail, M. A., Turner, D. J. & Swerdlow, H. (2009) Improved Protocols for Illumina
613 Sequencing. *Current Protocols in Human Genetics*, *18*, 18.2.1–18.2.42.
614 https://doi.org/10.1002/0471142905.hg1802s79
615 Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug,
616 H. J., Machler, M., & Bolker, B. M. (2017). glmmTMB Balances Speed and Flexibility Among
617 Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, *9*, 378–400.
618 https://doi.org/10.32614/RJ-2017-066
619 Burton, A. C., Neilson, E., Moreira, D., Ladle, A., Steenweg, R., Fisher, J. T., Bayne, E., & Boutin,
620 S. (2015). REVIEW: Wildlife camera trapping: a review and recommendations for linking
621 surveys to ecological processes. *Journal of Applied Ecology*, *52*, 675–
622 685. https://doi.org/10.1111/1365-2664.12432
623 Caravaggi, A., Gatta, M., Vallely, M. -C., Hogg, K., Freeman, M., Fadaei, E., Dick, J. T. A.,
624 Montgomery, W. I., Reid, N., & Tosh, D. G. (2018). Seasonal and predator-prey effects on
625 circadian activity of free-ranging mammals revealed by camera traps. *PeerJ*, *6*, e5827.
626 https://doi.org/10.7717/peerj.5827
627 Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S.,
628 Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E., & Bernatchez, L. (2017). Environmental
629 DNA metabarcoding: transforming how we survey animal and plant communities. *Molecular*
630 *Ecology*, *26*, 5872–5895. https://doi.org/10.1111/mec.14350
631 Evans, N. T., Li, Y., Renshaw, M. A., Olds, B. P., Deiner, K., Turner, C. R., Jerde, C. L., Lodge, D. M.,
632 Lamberti, G. A., & Pfrender, M. E (2017). Fish community assessment with eDNA
633 metabarcoding: effects of sampling design and bioinformatic filtering. *Canadian Journal of*
634 *Fisheries and Aquatic Sciences*, *74*, 1362–1374. https://doi.org/10.1139/cjfas-2016-0306

635 Franklin, T. W., McKelvey, K. S., Golding, J. D., Mason, D. H., Dysthe, J. C., Pilgrim, K. L., Squires,
636 J. R., Aubry, K. B., Long, R. A., Greaves, S. E., Raley, C. M., Jackson, S., MacKay, P., Lisbon, J.,
637 Sauder, J. D., Pruss, M. T., Heffington, D., & Schwartz, M. K (2019). Using environmental DNA
638 methods to improve winter surveys for rare carnivores: DNA from snow and improved
639 noninvasive techniques. *Biological Conservation*, *229*, 50–58.
640 https://doi.org/10.1016/j.biocon.2018.11.006
641 Gaughran, A., Kelly, D. J., MacWhite, T., Mullen, E., Maher, P., Good, M., & Marples, N. M.
642 (2018). Super-ranging. A new ranging strategy in European badgers. *PLoS ONE*, *13*,
643 e0191818. https://doi.org/10.1371/journal.pone.0191818
644 Hänfling, B., Lawson Handley, L., Read, D. S., Hahn, C., Li, J., Nichols, P., Blackman, R. C., Oliver,
645 A., & Winfield, I. J. (2016). Environmental DNA metabarcoding of lake fish communities
646 reflects long-term data from established survey methods. *Molecular Ecology*, *25*, 3101–3119.
647 https://doi.org/10.1111/mec.13660
648 Hänfling, B., Lawson Handley, L., Lunt, D., Shum, P., Winfield, I., & Read, D. S. (2017). *A Review
649 of Recent Advances in Genetic Methods to Identify Improvements in CAMERAS Partners
650 Monitoring Activities*. University of Hull, Hull, UK.
651 Harper, L. R., Lawson Handley, L., Hahn, C., Boonham, N., Rees, H. C., Gough, K. C., Lewis, E.,
652 Adams, I. P., Brotherton, P., Phillips, S., & Hänfling, B. (2018). Needle in a haystack? A
653 comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested
654 newt (*Triturus cristatus*). *Ecology and Evolution*, *8*, 6330–6441.
655 https://doi.org/10.1002/ece3.4013
656 Harper, L. R., Lawson Handley, L., Hahn, C., Boonham, N., Rees, H. C., Lewis, E., Adams, I. P.,
657 Brotherton, P., Phillips, S., & Hänfling, B. (2019) Testing ecological hypotheses at the
658 pondscape with environmental DNA metabarcoding: a case study on a threatened
659 amphibian. *bioRxiv*, 278309. https://doi.org/10.1101/278309
660 Harris, S., & Yalden, D. W. (2004). An integrated monitoring programme for terrestrial mammals
661 in Britain. *Mammal Review*, *34*, 157–167. https://doi.org/10.1046/j.0305-1838.2003.00030.x
662 Ishige, T., Miya, M., Ushio, M., Sado, T., Ushioda, M., Maebashi, K., Yonechi, R., Lagan, P., &
663 Matsubayashi, H. (2017). Tropical-forest mammals as detected by environmental DNA at
664 natural saltlicks in Borneo. *Biological Conservation*, *210*, 281–285.
665 https://doi.org/10.1016/j.biocon.2017.04.023
666 Joint Nature Conservation Committee (2018) UK BAP priority terrestrial mammal species.
667 Accessed on 1st December 2018. http://jncc.defra.gov.uk/page-5170
668 Kelly, R. P., Port, J. A., Yamahara, K. M., & Crowder, L. B. (2014). Using environmental DNA to
669 census marine fishes in a large mesocosm. *PLoS ONE*, *9*, e86175.
670 https://doi.org/10.1371/journal.pone.0086175

671 Kinoshita, G., Yonezawa, S., Murakami, S., & Isagi, Y. (2019) Environmental DNA Collected from
672 Snow Tracks is Useful for Identification of Mammalian Species. *Zoological Science*, *36*, 198–
673 207. http://dx.doi.org/10.2108/zs180172
674 Kitson, J. J. N., Hahn, C., Sands, R. J., Straw, N. A., Evans, D. M., & Lunt, D. H. (2019). Detecting
675 host-parasitoid interactions in an invasive Lepidopteran using nested tagging DNA-
676 metabarcoding. *Molecular Ecology*, *28*, 471–483. https://doi.org/10.1111/mec.14518
677 Klymus, K. E., Richter, C. A., Thompson, N., & Hinck, J. E. (2017). Metabarcoding of
678 Environmental DNA Samples to Explore the Use of Uranium Mine Containment Ponds as a
679 Water Source for Wildlife. *Diversity*, *9*, 54. https://doi.org/10.3390/d9040054
680 Lacoursière-Roussel, A., Dubois, Y., Normandeau, E., & Bernatchez, L. (2016). Improving
681 herpetological surveys in eastern North America using the environmental DNA method.
682 *Genome*, *59*, 991–1007. https://doi.org/10.1139/gen-2015-0218
683 Lawson Handley, L., Read, D.S., Winfield, I.J., Kimbell, H., Johnson, H., Li, J., Hahn, C., Blackman,
684 R., Wilcox, R., Donnelly, R., Szitenberg, A., & Hänfling, B. (2019) Temporal and spatial
685 variation in distribution of fish environmental DNA in England's largest lake. *Environmental*
686 *DNA*, **12**, e0176343. https://doi.org/10.1002/edn3.5
687 Leempoel, K., Hebert, T., & Hadly, E. A. (2019) A comparison of eDNA to camera trapping for
688 assessment of terrestrial mammal diversity. *bioRxiv*, 634022.
689 https://doi.org/10.1101/634022
690 Lugg, W. H., Griffiths, J., van Rooyen, A. R., Weeks, A. R., & Tingley, R. (2017). Optimal survey
691 designs for environmental DNA sampling. *Methods in Ecology and Evolution*, *9*, 1049–1059.
692 https://doi.org/10.1111/2041-210X.12951
693 Massimino, D., Harris, S. J., & Gillings, S. (2018). Evaluating spatiotemporal trends in terrestrial
694 mammal abundance using data collected during bird surveys. *Biological Conservation*, *226*,
695 153–167. https://doi.org/10.1016/j.biocon.2018.07.026
696 Mathews, F., Kubasiewicz, L. M., Gurnell, J., Harrower, C. A., McDonald, R. A., & Shore, R. F.
697 (2018) *A Review of the Population and Conservation Status of British Mammals: Technical*
698 *Summary*. Natural England, Peterborough.
699 McKnight, D. T., Huerlimann, R., Bower, D. S., Schwarzkopf, L., Alford, R. A., & Zenger, K. R.
700 (2019) microDecon: A highly accurate read-subtraction tool for the post-sequencing removal
701 of contamination in metabarcoding studies. *Environmental DNA*, *1*, 14–25.
702 https://doi.org/10.1002/edn3.11
703 Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers:
704 inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids*
705 *Research*, *39*, e145. https://doi.org/10.1093/nar/gkr732
706 Rodgers, T. W., & Mock, K. E. (2015). Drinking water as a source of environmental DNA for the
707 detection of terrestrial wildlife species. *Conservation Genetics Resources*, *7*, 693–696.
708 https://doi.org/10.1007/s12686-015-0478-7

709 Ruiz-Olmo, J., & Gosálbez, J. (1997). Observations on the sprainting behaviour of the otter *Lutra*
710 *lutra* in the NE Spain. *Acta Theriologica*, *42*, 259–270.
711 Sadlier, L. M. J., Webbon, C. C., Baker, P. J., & Harris, S. (2004). Methods of monitoring red foxes
712 Vulpes vulpes and badgers Meles meles: are field signs the answer? *Mammal Review*, *34*,
713 75–98. https://doi.org/10.1046/j.0305-1838.2003.00029.x
714 Sales, N. G., McKenzie, M. B., Drake, J., Harper, L. R., Browett, S. S., Coscia, I., Wangensteen, O.
715 S., Baillie, C., Bryce, E., Dawson, D. A., Ochu, E., Hänfling, B., Lawson Handley, L., Mariani, S.,
716 Lambin, X., Sutherland, C., & McDevitt, A. D. (2019) Fishing for mammals: landscape-level
717 monitoring of terrestrial and semi-aquatic communities using eDNA from lotic ecosystems.
718 *bioRxiv*, 629758. https://doi.org/10.1101/629758
719 Sellers, G. S., Di Muri, C., Gómez, A., & Hänfling, B. (2018). Mu-DNA: a modular universal DNA
720 extraction method adaptable for a wide range of sample types. *Metabarcoding and*
721 *Metagenomics*, *2*, e24556. https://doi.org/10.3897/mbmg.2.24556
722 Sheehy, E., Sutherland, C., O'Reilly, C., & Lambin, X. (2018) The enemy of my enemy is my
723 friend: native pine marten recovery reverses the decline of the red squirrel by suppressing
724 grey squirrel populations. *Proceedings of the Royal Society of London B*, *285*, 20172603.
725 https://doi.org/10.1098/rspb.2017.2603
726 Staley, Z. R., Chuong, J. D., Hill, S. J., Grabuski, J., Shokralla, S., Hajibabaei, M., & Edge, T. A.
727 (2018). Fecal source tracking and eDNA profiling in an urban creek following an extreme rain
728 event. *Scientific Reports*, *8*, 14390. https://doi.org/10.1038/s41598-018-32680-z
729 Tessler, M., Weiskopf, S. R., Berniker, L., Hersch, R., Mccarthy, K. P., Yu, D. W., & Siddall, M. E.
730 (2018). Bloodlines: mammals, leeches, and conservation in southern Asia. *Systematics and*
731 *Biodiversity*, *16*, 488–496. https://doi.org/10.1080/14772000.2018.1433729
732 Thomsen, P. F., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., Orlando, L., &
733 Willerslev, E. (2012). Monitoring endangered freshwater biodiversity using environmental
734 DNA. *Molecular Ecology*, *21*, 2565–2573. https://doi.org/10.1111/j.1365-294X.2011.05418.x
735 Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA–an emerging tool in conservation
736 for monitoring past and present biodiversity. *Biological Conservation*, *183*, 4–18.
737 https://doi.org/10.1016/j.biocon.2014.11.019
738 Ushio, M., Fukuda, H., Inoue, T., Makoto, K., Kishida, O., Sato, K., Murata, K., Nikaido, M., Sado,
739 T., Sato, Y., Takeshita, M., Iwasaki, W., Yamanaka, H., Kondoh, M., & Miya, M. (2017).
740 Environmental DNA enables detection of terrestrial mammals from forest pond water.
741 *Molecular Ecology Resources*, *17*(6), e63–e75. https://doi.org/10.1111/1755-0998.12690

742 Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E.,
743 Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G. H.,
744 Geniez, P., Pont, D., Argillier, C., Baudoin, J. -M., Peroux, T., Crivelli, A. J., Olivier, A.,
745 Acqueberge, M., Le Brun, M., Møller, P. R., Willerslev, E., & Dejean, T. (2016). Next-
746 generation monitoring of aquatic biodiversity using environmental DNA metabarcoding.
747 *Molecular Ecology*, *25*, 929–942. https://doi.org/10.1111/mec.13428
748 Visconti, P., Pressey, R. L., Giorgini, D., Maiorano, L., Bakkenes, M., Boitani, L., Alkemade, R.,
749 Falcucci, A., Chiozza, F., & Rondinini, C. (2011). Future hotspots of terrestrial mammal loss.
750 *Philosophical Transactions of the Royal Society B*, *366*, 2693–2702.
751 https://doi.org/10.1098/rstb.2011.0105
752 Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. Springer, New York, USA.
753 http://ggplot2.org
754 Williams, K. E., Huyvaert, K. P., Vercauteren, K. C., Davis, A. J., & Piaggio, A. J. (2018). Detection
755 and persistence of environmental DNA from an invasive, terrestrial mammal. *Ecology and*
756 *Evolution*, *8*, 688–695. https://doi.org/10.1002/ece3.3698
757 Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009) *Mixed effects models*
758 *and extensions in ecology with R*. Springer, New York, USA.
759

760 **Table 1.** Ethogram used to catalogue mammal behaviours that occur in or near artificial
761 waterbodies in captive enclosures. Importantly, this ethogram was designed to catalogue
762 mammal behaviours potentially leading to eDNA deposition. Therefore, it may not be
763 comparable to ethograms typically used in study of captive animals.

764

Behaviour	**Definition**
Swimming	Mammal completely submerged in and moving through waterbody using limbs
Bodypart in water	Mammal partially submerged in waterbody, e.g. foot or tail in water
Drinking	Water taken into mouth and swallowed by mammal
Feeding	Food taken into mouth and swallowed by mammal in or near waterbody, e.g. otter and fish
Scratching	Bodypart or external object in enclosure used by mammal to relieve itch near waterbody
Urinating/scent-marking	Liquid excretion passed by mammal in or near waterbody
Pooing	Solid excretion passed by mammal in or near waterbody
Sniffing	Air visibly drawn through nose of mammal to detect a smell around waterbody, possibly involving contact with water
Standing	Mammal motionless in or near waterbody
Walking	Mammal moving around waterbody at a regular pace by lifting and setting down each foot in turn, never having both feet off the ground at once
Running	Mammal moving around waterbody at a speed faster than a walk, never having both or all the feet on the ground at the same time
Vocalising	Mammal producing sound while in or near waterbody
Grooming	Mammal cleaning fur or skin with its tongue while in or near waterbody
Resting	Mammal lying down or sitting in or near waterbody
Other	Behaviour exhibited in or near waterbody that does not conform to other categories, e.g. chasing tail
Not visible	Mammal moved to part of enclosure not visible to the observer

765

766 **Table 2.** Summary of focal species studied at wildlife parks and their lifestyle. The number of

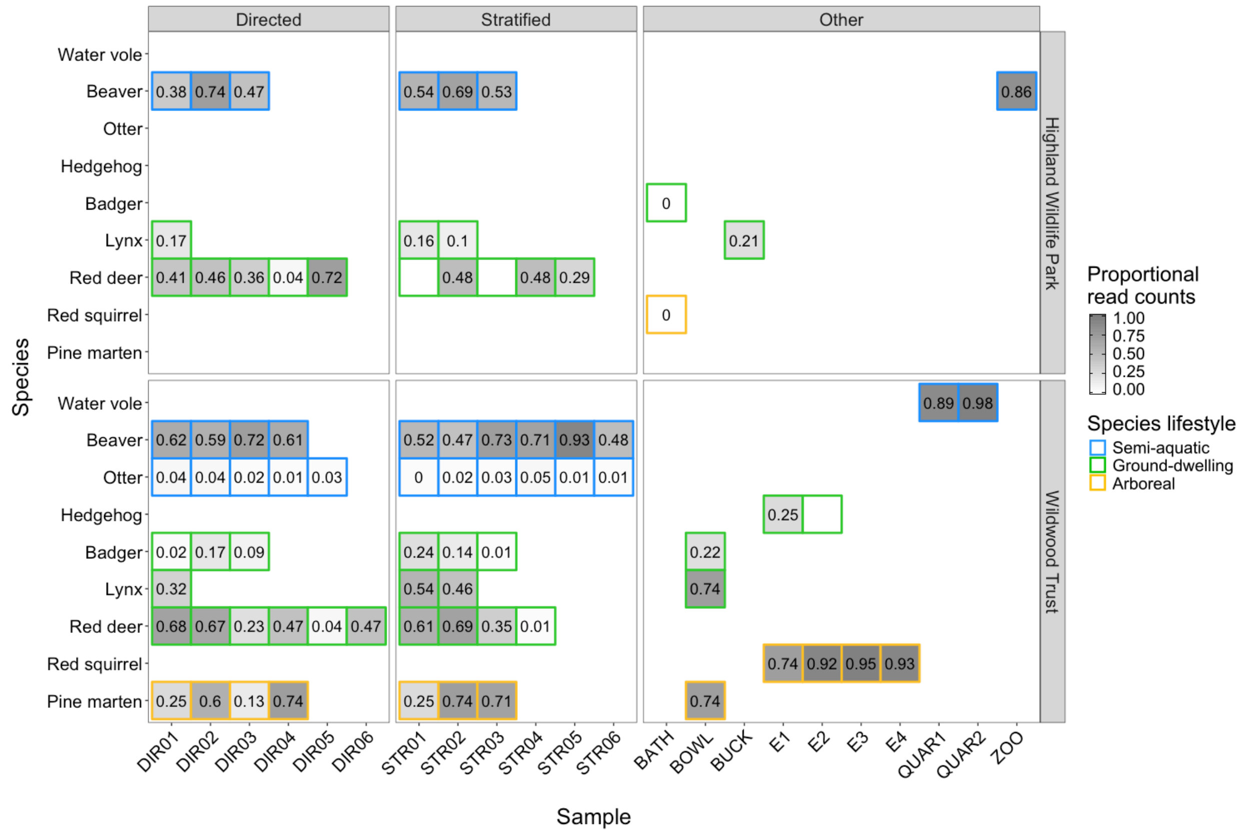
767 individuals present and waterbody size in enclosures is provided.

768

Site	Species	Lifestyle	Enclosure	Number of individuals	Waterbody size (m^2)
Wildwood Trust	European otter (*Lutra lutra*)	Semi-aquatic	1	2	162
	European water vole (*Arvicola amphibius*)	Semi-aquatic	1	4	0.09
			2	1	0.09
	European beaver (*Castor fiber*)	Semi-aquatic	1	2	100
			2	1	100
	European hedgehog (*Erinaceus europaeus*)	Ground-dwelling	1	1	0.04
			2	2	0.04
	European badger (*Meles meles*)	Ground-dwelling	1	4	1.73
	Red deer (*Cervus elaphus*)	Ground-dwelling	1	8	100
	Eurasian lynx (*Lynx lynx*)	Ground-dwelling	1	2	2
	Red squirrel (*Sciurus vulgaris*)	Arboreal	1	2	0.01
			2	3	0.01
			3	3	0.01
			4	2	0.01
	European pine marten (*Martes martes*)	Arboreal	1	1	2
			2	1	0.375
Highland Wildlife Park	Red squirrel (*Sciurus vulgaris*)	Arboreal	NA	NA	0.25
	Eurasian lynx (*Lynx lynx*)	Ground-dwelling	1	8	2
	European beaver (*Castor fiber*)	Semi-aquatic	1	2	50
	Red deer (*Cervus elaphus*)	Ground-dwelling	1	30	NA

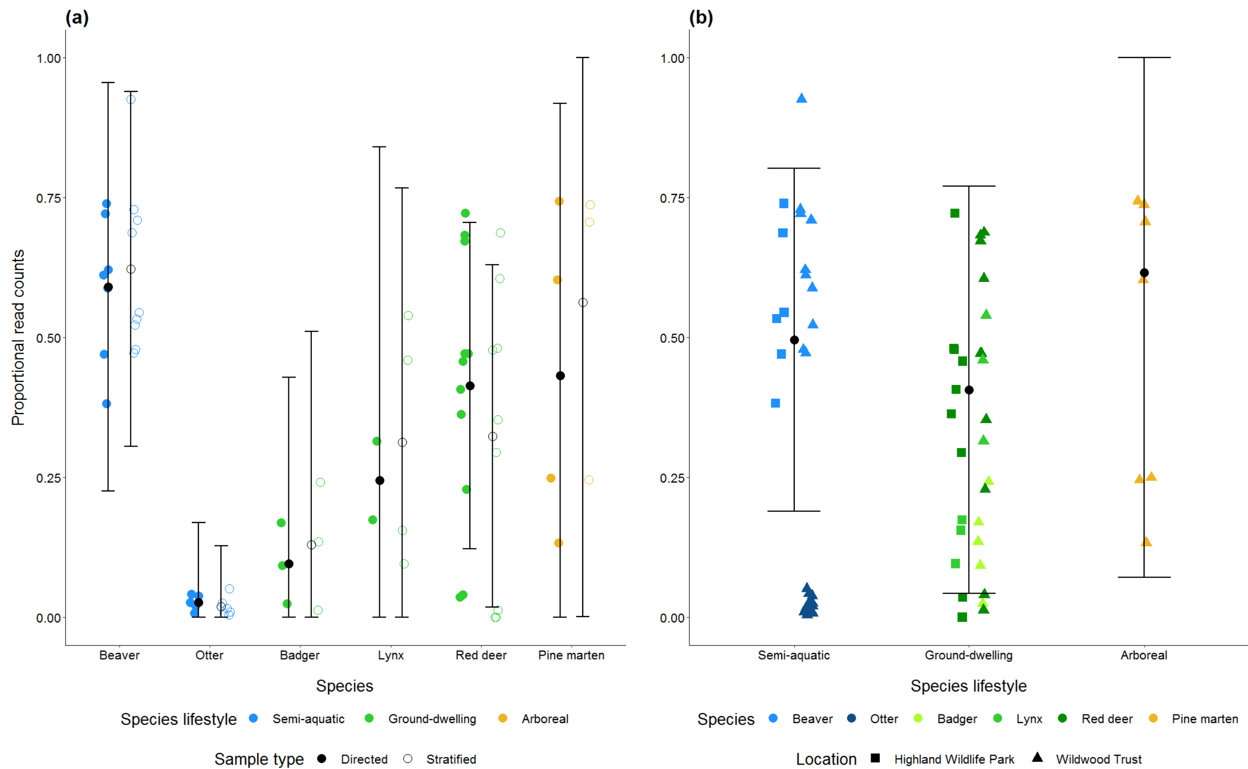
769

770



771

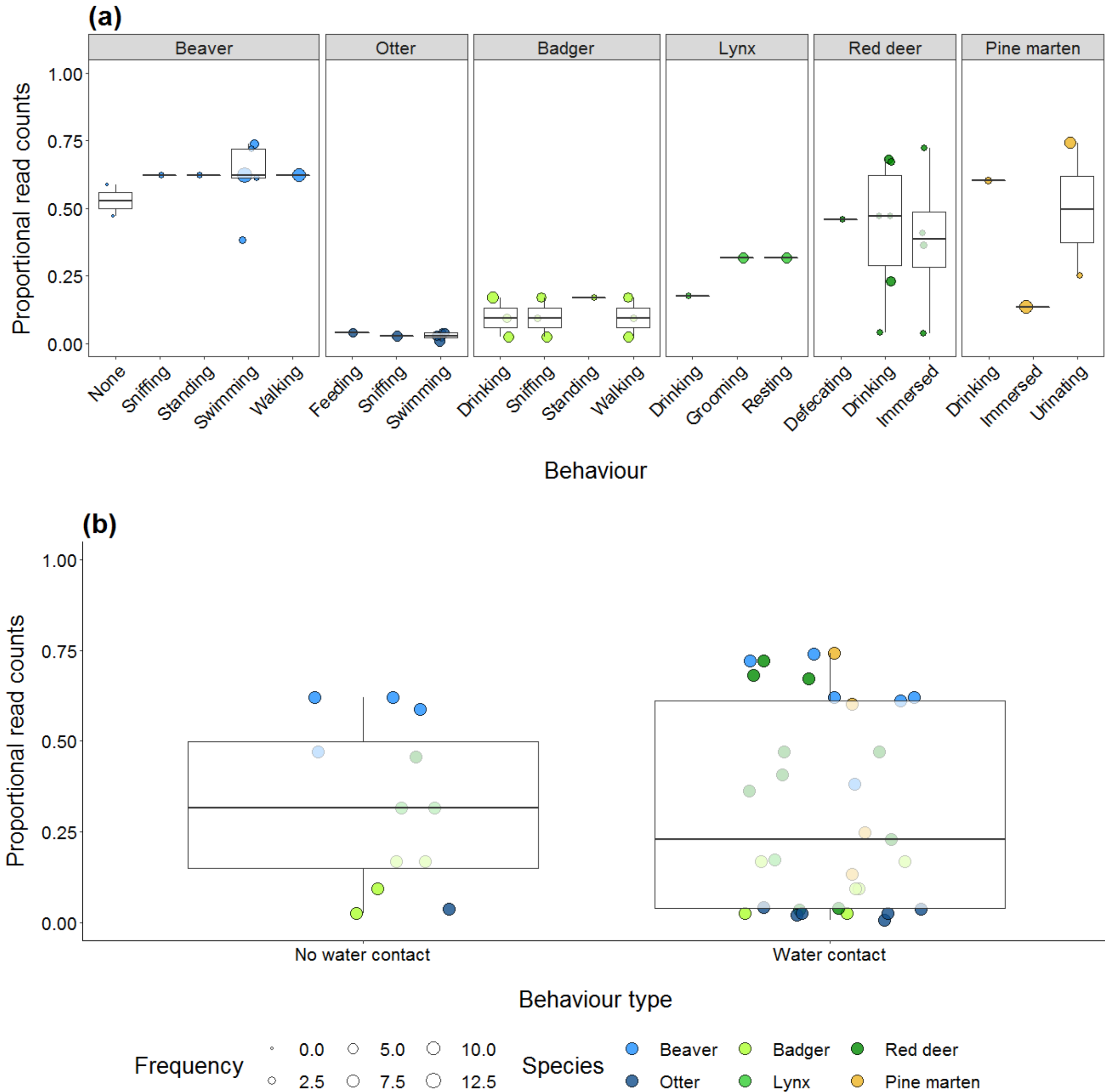
772 **Figure 1.** Heatmap showing proportional read counts for eDNA samples (n = 81) from
773 Experiment 1. The heatmap is faceted by sample type (directed, stratified or other) and wildlife
774 park (Highland Wildlife Park or Wildwood Trust). Each cell represents an individual sample
775 taken from an enclosure containing the focal species in that row. Directed (DIR01-DIR06) and
776 stratified (STR01-STR06) samples were collected for each species from artificial waterbodies.
777 Samples were also collected from drinking containers (E1, E2, E3, E4, BOWL, BUCK), water vole
778 (QUAR1, QUAR2) and RZSS Edinburgh Zoo beaver (ZOO) enclosures, and a water bath (BATH) in
779 RZSS Highland Wildlife Park woods. The maximum proportional read count for each cell (i.e.
780 sample) is 1, if all reads from a particular sample belonged to the focal species. Cells containing
781 0 represent samples with proportional read counts less than 0.01 whereas empty cells are
782 samples with proportional read counts of exactly 0.



783

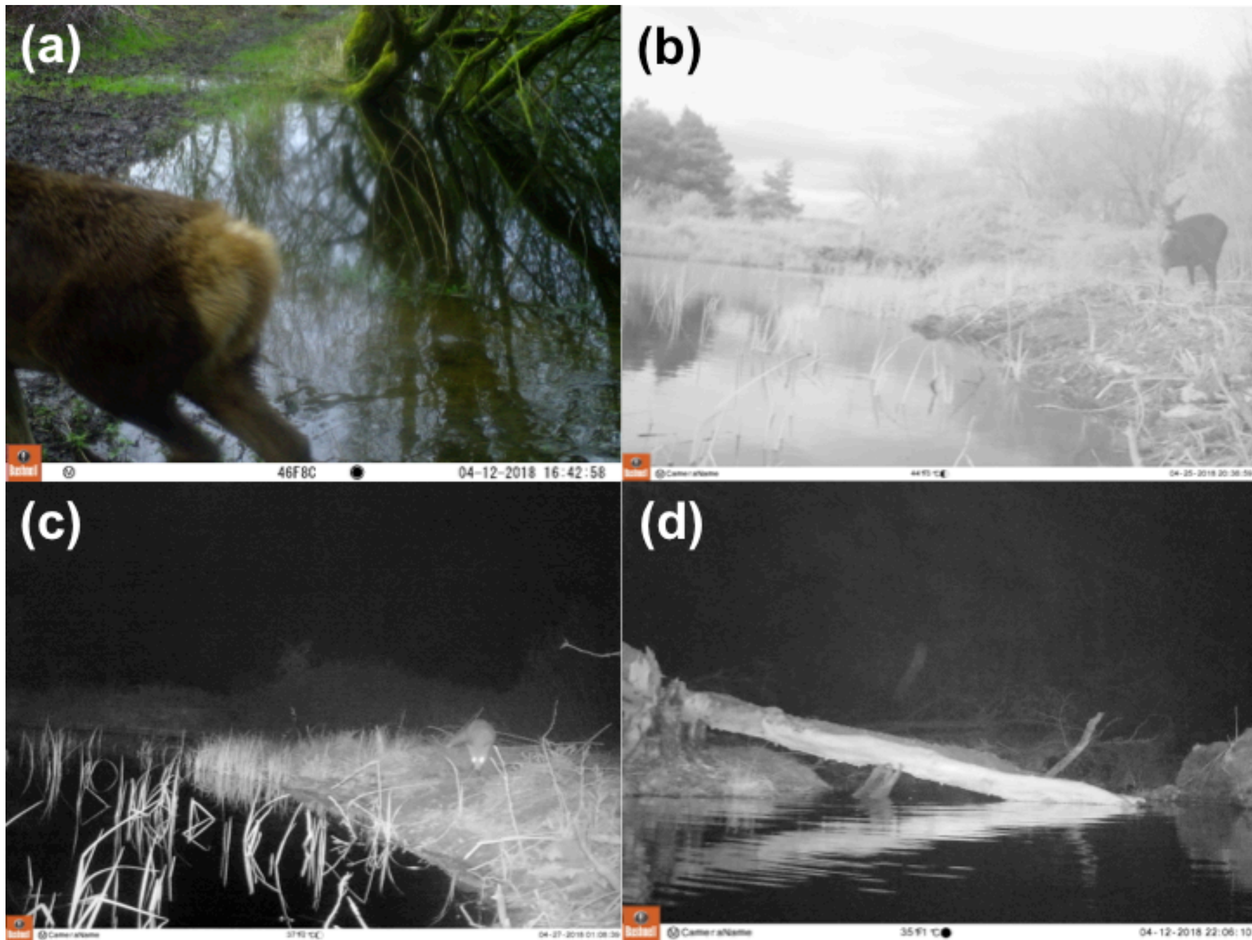
784 **Figure 2.** Relationships predicted by the binomial GLMMs between proportional read counts
785 and sample type nested within species **(a)** or species lifestyle **(b)** for Experiment 1. The
786 observed data (coloured points) are displayed against the predicted relationships (black points
787 with error bars) for each species **(a)** or species lifestyle **(b)**. Points are shaped by sample type **(a)**
788 or wildlife park **(b)**, and coloured by species lifestyle. Error bars represent the standard error
789 around the predicted means.

790



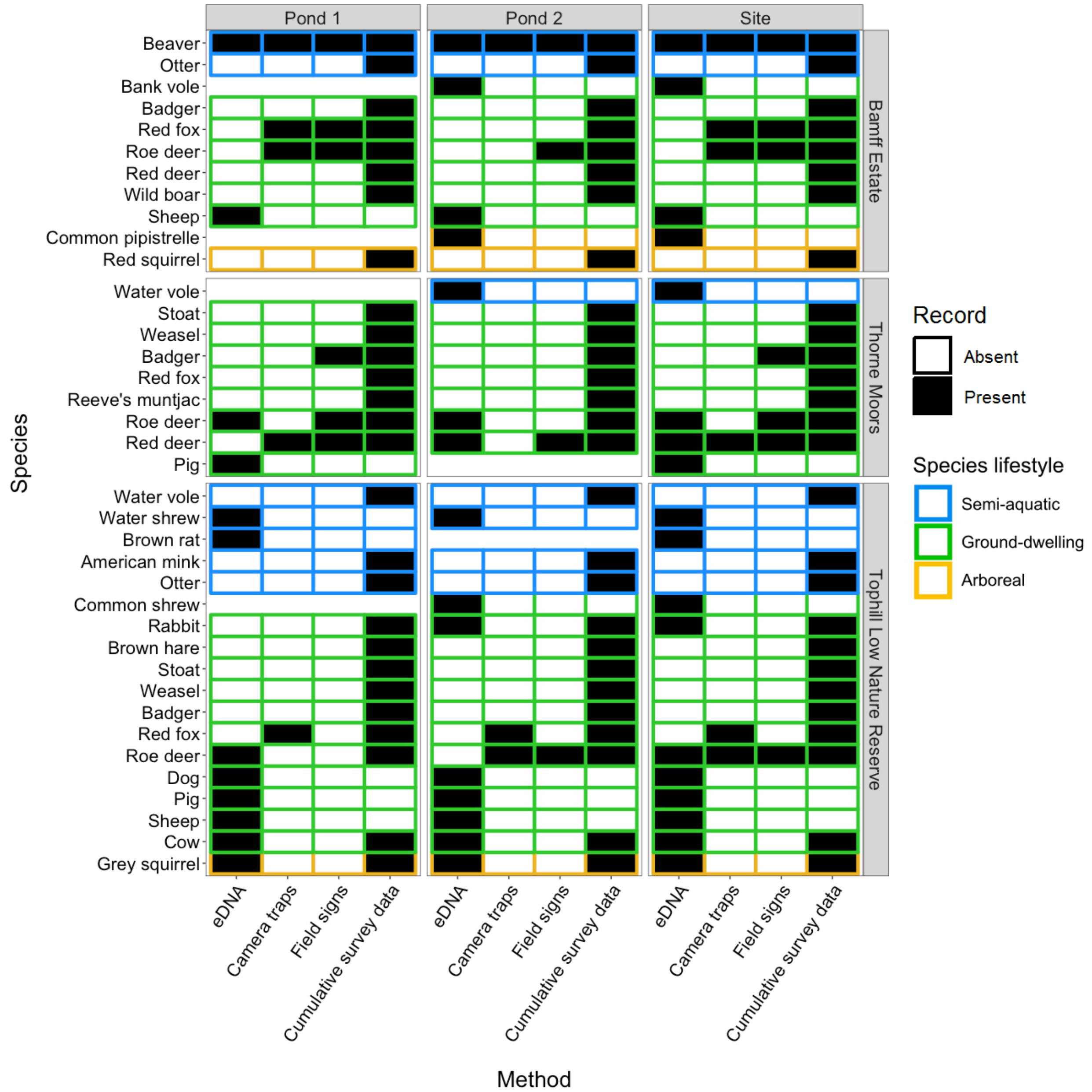
791

792 **Figure 3.** Boxplots showing the mean proportional read counts for specific **(a)** and generic **(b)**
793 behaviour(s) exhibited by focal species in Experiment 1. Boxes show 25th, 50th, and 75th
794 percentiles, and whiskers show 5th and 95th percentiles. Points are coloured by species
795 lifestyle, and each point in **(a)** represents a directed sample sized by frequency of behaviour.
796 The behaviour 'none' for beaver represents occurrences of beaver in water but out of view of
797 camera traps.

798 

799 **Figure 4.** Exemplar camera trap photos taken at natural ponds where focal species were
800 present in Experiment 2. Red deer was recorded at Thorne Moors **(a)**, roe deer **(b)** and red fox
801 **(c)** were recorded at Tophill Low Nature Reserve, and beaver was recorded at the Bamff Estate
802 **(d)**.

803



804

805 **Figure 5.** Tile plot showing species presence-absence at individual pond and site-level as

806 indicated by field signs, camera trapping, and eDNA metabarcoding in Experiment 2. Surveys

807 were performed at sites where focal species presence was confirmed by cumulative survey

808 data.

Appendix A

Contents

1. Materials and methods 3
  1.1 eDNA metabarcoding workflow 3
  1.2 Data analysis 5
    1.2.1 Dataset refinement 5
    1.2.2 Effect of water volume and filtration method 7
2. Results and discussion 8
  2.1 Effect of water volume and filtration method 8
  2.2 Technical pitfalls of eDNA metabarcoding for mammal monitoring 8
3. Tables 9
  Table A1 9
  Table A2 13
  Table A3 15
  Table A4 21
  Table A5 22
  Table A6 25
4. Figures 27
  Figure A1 27
  Figure A2 28
  Figure A3 29
  Figure A4 30
  Figure A5 31
  Figure A6 32
References 33

1. Materials and methods

1.1 eDNA metabarcoding workflow

A two-step PCR protocol was performed on eDNA samples at the University of Hull. Dedicated rooms were available for pre-PCR and post-PCR processes. Pre-PCR processes were performed in a dedicated eDNA laboratory, with separate rooms for filtration, DNA extraction, and PCR preparation of sensitive environmental samples. PCR reactions were set up in a ultraviolet (UV) and bleach sterilised laminar flow hood. To minimise cross-contamination risk between samples, eight-strip PCR tubes with individually attached lids were used instead of 96-well plates (Port et al., 2016) and PCR reactions were sealed with mineral oil (Sigma-Aldrich Company Ltd, Dorset, UK) droplets (Harper et al., 2018). PCR positive (n = 2) and negative controls (n = 2) were included on each PCR run to screen for sources of potential contamination. The DNA (0.05 ng/µL) used for the PCR positive control (n = 16) was *Maylandia zebra* as this is an exotic cichlid not found in UK freshwater habitats. The negative controls (n = 16) substituted molecular grade sterile water (Fisher Scientific UK Ltd, Loughborough, UK) for template DNA.

During the first PCR, the target region was amplified using published 12S rRNA primers 12S-V5-F (5'-ACTGGGATTAGATACCCC-3') and 12S-V5-R (5'-TAGAACAGGCTCCTCTAG-3') (Riaz et al., 2011) that were validated *in silico* for all UK vertebrates by Harper *et al*. (2018). Primers were modified to include MID tags, heterogeneity spacers, sequencing primers, and pre-adapters. During the first PCR, three replicates were performed for each sample to combat amplification bias. PCR reactions were performed in 25 µL volumes, consisting of 3 µL of template DNA, 1.5 µL of each 10 µM primer (Integrated DNA Technologies, Belgium), 12.5 µL of Q5® High-Fidelity 2x Master Mix (New England Biolabs® Inc., MA, USA) and 6.5 µL molecular grade water (Fisher Scientific UK Ltd, UK). PCR was performed on an Applied Biosystems® Veriti Thermal Cycler (Life Technologies, CA, USA) with the following thermocycling profile: 98 °C for 5 mins, 35 cycles of 98 °C for 10 s, 58 °C for 20 s and 72 °C for 30 s, followed by a final elongation step at 72 °C for 7 mins. PCR products were stored at 4 °C until replicates for each sample were pooled, and 2 µL of pooled PCR product was added to 0.5 µL of 5x DNA Loading Buffer Blue (Bioline®, London, UK). PCR product was visualised on 2% agarose gels (1.6 g Bioline® Agarose in 80 mL 1x sodium borate) stained with ethidium bromide, and gels were imaged using Image Lab Software (Bio-Rad Laboratories Ltd, Watford, UK). A PCR product was deemed positive where there was an amplification band on the gel that was of the expected size (200-300 bp). PCR products were stored at -20 °C until they were pooled according to PCR plate to create sub-libraries for purification with Mag-BIND® RxnPure Plus magnetic beads (Omega Bio-tek Inc, GA, USA), following the double size selection protocol established by Bronner et al. (2009). Ratios of

0.9x and 0.15x magnetic beads to 100 µL of each sub-library were used. Eluted DNA (30 µL) was stored at -20 °C until the second PCR could be performed.

The second PCR bound pre-adapters, MID tags, and Illumina adapters to the purified sub-libraries. Two replicates were performed for each sub-library in 50 µL volumes, consisting of 6 µL of template DNA, 3 µL of each 10 µM primer (Integrated DNA Technologies, Belgium), 25 µL of Q5® High-Fidelity 2x Master Mix (New England Biolabs® Inc., MA, USA) and 13 µL molecular grade water (Fisher Scientific UK Ltd, UK). PCR was performed on an Applied Biosystems® Veriti Thermal Cycler (Life Technologies, CA, USA) with the following thermocycling profile: 95 °C for 3 mins, 8 cycles of 98 °C for 20 s and 72 °C for 1 min, followed by a final elongation step at 72 °C for 5 mins. PCR products were stored at 4 °C until duplicates for each sub-library were pooled, and 2 µL of pooled product was added to 0.5 µL of 5x DNA Loading Buffer Blue (Bioline®, London, UK). PCR products were visualised on 2% agarose gels (1.6 g Bioline® Agarose in 80 mL 1x Sodium borate) stained with ethidium bromide, and gels were imaged using Image Lab Software (Bio-Rad Laboratories Ltd, Watford, UK). Again, PCR products were deemed positive where there was an amplification band on the gel that was of the expected size (300-400 bp). Sub-libraries were stored at 4 °C until purification with Mag-BIND® RxnPure Plus magnetic beads (Omega Bio-tek Inc, GA, USA), following the double size selection protocol established by Bronner et al. (2009). Ratios of 0.7x and 0.15x magnetic beads to 50 µL of each sub-library were used. Eluted DNA (30 µL) was stored at 4 °C until normalisation and final purification.

Sub-libraries were quantified on a Qubit™ 3.0 fluorometer using a Qubit™ dsDNA HS Assay Kit (Invitrogen, UK) and pooled proportional to sample size and concentration. The pooled library was purified using the same ratios, volumes, and protocol as second PCR purification. Based on Qubit™ concentration, the library was diluted to 6 nM for quantification by real-time quantitative PCR (qPCR) using the NEBNext® Library Quant Kit for Illumina® (New England Biolabs® Inc., MA, USA). The library was also checked using an Agilent 2200 TapeStation and High Sensitivity D1000 ScreenTape (Agilent Technologies, CA, USA) to verify secondary product had been removed successfully and a fragment of the expected size (330 bp) remained. The library was frozen and transported in a sterile portable freezer (-20 °C) to Centre for Ecology & Hydrology (CEH), Wallingford, for sequencing. The library was sequenced at 11.5 pM with 10% PhiX Control on an Illumina MiSeq® using 2 x 300 bp V3 chemistry (Illumina Inc., CA, USA).

Raw sequence reads were demultiplexed using a custom Python script then processed using metaBEAT (metaBarcoding and Environmental Analysis Tool) v0.97.11 (https://github.com/HullUni-bioinformatics/metaBEAT). Raw reads were quality trimmed from the read ends (minimum per base phred score Q30) and across sliding windows (window size 5bp; minimum average phred score Q30) using Trimmomatic v0.32 (Bolger, Lohse, & Usadel, 2014). Reads were cropped to a maximum length of 110 bp and reads shorter than 90 bp after

quality trimming were discarded. The first 18 bp of remaining reads were also removed to ensure no locus primer remained. Sequence pairs were merged into single high quality reads using FLASH v1.2.11 (Magoč & Salzberg, 2011), provided there was a minimum overlap of 10 bp and no more than 10% mismatch between pairs. Only forward reads were kept for pairs that could not be merged. A final length filter was applied to ensure sequences were reflected the expected fragment size (90-110 bp). Retained sequences were screened for chimeric sequences against a custom reference database for UK vertebrates (Harper et al., 2018) using the uchime algorithm (Edgar, Haas, Clemente, Quince, & Knight, 2011), as implemented in vsearch v1.1 (Rognes, Flouri, Nichols, Quince, & Mahé, 2016). Redundant sequences were removed by clustering at 100% identity ('--cluster_fast' option) in vsearch v1.1 (Rognes et al., 2016). Clusters were considered sequencing error and omitted from further processing if they were represented by less than three sequences. Non-redundant sets of query sequences were then compared against the UK vertebrate reference database (Harper et al., 2018) using BLAST (Zhang, Schwartz, Wagner, & Miller, 2000). Putative taxonomic identity was assigned using a lowest common ancestor (LCA) approach based on the top 10% BLAST matches for any query that matched a reference sequence across more than 80% of its length at minimum identity of 98%. Unassigned sequences were subjected to a separate BLAST search against the complete NCBI nucleotide (nt) database at 98% identity to determine the source via LCA as described above. The bioinformatic analysis has been deposited in the GitHub repository for reproducibility (permanently archived at: https://doi.org/10.5281/zenodo.2561415).

1.2 Data analysis

1.2.1 Dataset refinement

Assignments from different databases were merged, and spurious assignments (i.e. non-UK species, invertebrates and bacteria) removed from the dataset. The family Cichlidae was reassigned to *Maylandia zebra*. The genera *Bison*, *Bos*, *Buteo*, *Castor*, *Meleagris*, *Pelophylax*, *Sprattus*, *Strix*, and *Triturus* were reassigned to European bison (*Bison bonasus*), cow (*Bos taurus*), common buzzard (*Buteo buteo*), Eurasian beaver (*Castor fiber*), marsh frog (*Pelophylax ridibundus*), turkey (*Meleagris gallopavo*), European sprat (*Sprattus sprattus*), tawny owl (*Strix aluco*), and great crested newt (*Triturus cristatus*) respectively based on local knowledge of sampling sites and UK distribution maps (National Biodiversity Network Atlas, 2019). The species *Sus scrofa* and *Canis lupus* were reassigned to pig (*Sus scrofa domesticus*) and (*Canis lupus familiaris*) given the restricted distribution of wild boar (*S. scrofa*) and absence of grey wolf (*C. lupus*) in the UK.

Misassignments included the cichlids *Haplochromis burtoni*, *Oreochromis niloticus*, and *Pundamilia nyererei* which were reassigned to *M. zebra*, and Iberian lynx (*Lynx pardinus*) which was reassigned to Eurasian lynx (*Lynx lynx*). Other potential misassignments were green-winged teal (*Anas carolinensis*), yellow-browed bunting (*Emberiza chrysophrys*), and Iceland gull (*Larus glaucoides*). These are rare migrants that have been infrequently recorded in the UK (British Trust for Ornithology, 2019) but may have been assigned due to potential for hybridisation within the genus *Anas* and missing reference sequences for several common species within the genera *Emberiza* and *Larus*. These species were reassigned to the genera *Anas*, *Emberiza*, and *Larus*. Reads from corrected assignments were then merged with unaltered assignments.

Of 89 process controls included throughout the metabarcoding workflow, 39 produced no reads. Reads generated for 50 of 89 process controls ranged from 3 to 4930, and strength of each contaminant varied (mean = 62.4%, range = 0.3 - 100.0% of the total reads per process control). Environmental contamination was observed in the field blanks (European badger [*Meles meles*], beaver, lynx, European pine marten [*Martes martes*], red squirrel [*Sciurus vulgaris*], and European water vole [*Arvicola amphibius*]) as well as environmental and/or laboratory contamination in the filtration and extraction blanks (human [*Homo sapiens*] and cichlid). PCR negative controls were also contaminated with human, cichlid, beaver, and pine marten as well as non-focal species (Fig. S3). Consequently, we evaluated different sequence thresholds to minimise the risk of false positives in our dataset. These included the maximum sequence frequency of cichlid DNA in eDNA samples (0.308%), maximum sequence frequency of any DNA except cichlid in PCR positive controls (0.064%), and taxon-specific thresholds (maximum sequence frequency of each taxon in PCR positive controls). The different thresholds were applied to the eDNA samples and the results from each compared (Fig. S4). The taxon-specific thresholds (Table S4) retained the most biological information, thus these were selected for downstream analysis. Consequently, taxa were only classed as present at sites if their sequence frequency exceeded taxon-specific thresholds.

Contaminants remaining in eDNA samples after threshold application included Gentoo penguin (*Pygoscelis papua*) and reindeer (*Rangifer tarandus*) which were likely sourced from the environment, cichlid sourced from the laboratory, and human which may have originated from the environment or the laboratory. Gentoo penguin was only detected in water from the beaver quarantine enclosure at RZSS Edinburgh Zoo, and reindeer was only detected in water sampled from a red squirrel enclosure at Wildwood Trust. Human DNA was detected in the majority of eDNA samples, and cichlid DNA was also present at low frequency in some samples. These contaminants and assignments higher than species level were removed from the dataset, excluding the genera *Anas*, *Emberiza*, and *Larus*. Therefore, all taxonomic assignments in the final dataset were predominantly of species resolution and considered real detections. European bison, which is present in the red deer (*Cervus elaphus*) enclosure at RZSS Highland Wildlife Park, was detected in two samples taken from this enclosure. However, these

detections were excluded from downstream analyses as bison was not one of our focal species. Samples belonging to focal species in Experiment 1 were contaminated with DNA of other focal species to different extents (mean = 6.7%, range = 0.0 - 100.0% of the total refined reads per sample). Therefore, any proportional reads for incorrect focal species in each enclosure were set to 0 for the purposes of downstream analysis.

1.2.2 Effect of water volume and filtration method

We tested the hypothesis that volume of water filtered or number of filters used may affect read counts. A hierarchical binomial Generalized Linear Mixed Model (GLMM) with the logit link function from the development version of the R package glmmTMB (Brooks et al., 2017), including volume and number of filters as fixed effects and species nested within wildlife park as a random effect, was used. Validation checks were performed to ensure all model assumptions were met where possible and absence of overdispersion (Zuur, Ieno, Walker, Saveliev, & Smith, 2009). Model fit was assessed visually and with the Hosmer and Lemeshow Goodness of Fit Test (Hosmer & Lemeshow, 2000) using the R package ResourceSelection v0.3-2 (Lele, Keim, & Solymos, 2016). Predictions for each model were obtained using the *predict* function and upper and lower 95% CIs were calculated from the standard error of the predictions. Figures were produced using the R package ggplot2 v3.3.0 (Wickham, 2016).

2. Results and discussion

2.1 Effect of water volume and filtration method

Neither volume of water filtered (χ^2_1 = 2.141 ± 0.143) or number of filters used (χ^2_1 = 0.108 ± 0.742) had a significant effect on the proportional read counts based on the hierarchical model (θ = 0.221, χ^2_{76} = 16.798, *P* = 1.000, pseudo-R^2 = 20.57%). Proportional read counts somewhat decreased (-0.003 ± 0.002, *Z* = -1.389, *P* = 0.165) as water volume filtered increased (Fig. S1a), and marginally increased (0.232 ± 0.703, *Z* = 0.331, *P* = 0.741) where two filters were used for water filtration (Fig. S1b).

2.2 Technical pitfalls of eDNA metabarcoding for mammal monitoring

Our process controls identified low-level contamination from domestic and wild species at all stages of metabarcoding, but contaminants primarily occurred in field and PCR negative controls. We also identified cichlid sequences in field, filtration, and extraction controls, despite this DNA not being handled before PCR. Contaminants found in process controls may stem from PCR contamination (Kelly, Port, Yamahara, & Crowder, 2014) or sequencing error (Hänfling et al., 2016). Indeed, PCR negative controls corresponding to Tophill Low Nature Reserve samples were contaminated with great crested newt (*Triturus cristatus*), a highly abundant pond species at the reserve. This suggests that concentrated samples can contaminate negative controls during metabarcoding. Negative controls at each stage of metabarcoding can identify contaminant introduction (Klymus, Richter, Thompson, & Hinck, 2017; Ushio et al., 2017), but contaminants in these controls can amplify exponentially with no competition affecting inferences (Harper et al., 2018). Innovative approaches are needed to mitigate metabarcoding contamination, such as false positive estimation using occupancy models (Ficetola et al., 2015) or sequencing thresholds, i.e. the number of sequence reads required for a sample to be species-positive (Evans et al., 2017; Hänfling et al., 2016).

3. Tables

Table A1. Behavioural observation data for species housed at wildlife parks, including date, time, weather conditions, behaviour, and frequency and duration of behaviour.

Location	Date	Start time	End time	Weather	Air temperature (°C)	Species	Enclosure	Behaviour	Frequency	Duration (hrs)
Wildwood Trust	18/09/2017	10:29	11:29	Cloudy	13	Red deer (*Cervus elaphus*)	1	Drinking	9	9.37
								Feeding	3	22.37
								Defecating	1	0.5
								Sniffing	1	0.53
								Standing	6	30
								Walking	12	22.27
Wildwood Trust	19/09/2017	09:50	10:50	Sunny	17	Red deer (*Cervus elaphus*)	1	Drinking	2	4.47
								Feeding	2	40.68
								Standing	1	34.8
								Walking	1	5.75
								Resting	1	6.7
Wildwood Trust	18/09/2017	11:38	12:38	Partial sun	16	Eurasian lynx (*Lynx lynx*)	1	Scratching	1	0.08
								Urinating	2	3.42
								Standing	1	0.38
								Walking	9	18.12
								Walking	2	9.17
								Running	1	0.38
								Vocalising	4	4.58
								Resting	4	29.6
								Resting	1	1.32
								Grooming	4	5.03
								Grooming	1	0.08
								Not visible	1	1.88
								Not visible	2	48.25

Location	Date	Start time	End time	Weather	Air temperature (°C)	Species	Enclosure	Behaviour	Frequency	Duration (hrs)
Wildwood Trust	19/09/2017	09:30	10:30	Cloudy	12	Eurasian lynx (*Lynx lynx*)	1	Drinking	3	1.13
								Urinating	1	0.01
								Not visible	1	10.52
								Resting	1	2.42
								Walking	3	46.8
								Walking	3	27.92
								Resting	1	0.77
								Grooming	1	0.45
								Other	1	0.97
								Not visible	2	30.95
Wildwood Trust	18/09/2017	11:36	12:32	Partial sun	16	European pine marten (*Martes martes*)	1	Immersed	10	0.35
								Drinking	2	0.03
								Urinating	1	0.03
								Other	19	0.67
Wildwood Trust	21/09/2017	09:26	09:56	Sunny	17	European pine marten (*Martes martes*)	1	Drinking	1	0.12
								Sniffing	7	6.83
								Standing	1	0.82
								Walking	10	8.45
								Resting	10	12.18
								Playing	9	5.22
								Not visible	1	0.98
Wildwood Trust	21/09/2017	09:56	10:26	Sunny	17	European pine marten (*Martes martes*)	2	Immersed	3	0.2
								Urinating	9	0.27
								Standing	1	0.23
								Running	4	29.57
								Resting	1	0.2

Location	Date	Start time	End time	Weather	Air temperature (°C)	Species	Enclosure	Behaviour	Frequency	Duration (hrs)
Wildwood Trust	18/09/2017	13:25	13:55	Partial sun	17	Red squirrel (*Sciurus vulgaris*)	1	Feeding	5	3.77
								Walking	1	1.15
								Running	4	3.68
								Resting	2	0.8
								Drinking	2	0.93
								Running	4	23.83
								Resting	2	5.27
Wildwood Trust	18/09/2017	14:00	14:30	Partial sun	17	Red squirrel (*Sciurus vulgaris*)	2	Drinking	1	1.4
								Feeding	4	15.5
								Running	3	13.1
								Drinking	2	0.57
								Feeding	3	1.07
								Running	6	27.37
Wildwood Trust	18/09/2017	13:45	14:45	Sunny	16	European otter (*Lutra lutra*)	1	Swimming	12	14.25
								Standing	1	0.35
								Playing	2	5.28
Wildwood Trust	19/09/2017	09:21	10:21	Cloudy	12	European otter (*Lutra lutra*)	1	Swimming	8	16.4
								Feeding	3	5.7
								Sniffing	5	8.33
								Not visible	4	35.43
Wildwood Trust	19/09/2017	19:54	05:04	Cloudy	9	European badger (*Meles meles*)	1	Drinking	14	0.95
								Sniffing	36	8.43
								Walking	36	8.43
o d	19/09/2017	20:47	06:58	Clear	9	Eurasian beaver	1	Swimming	40	13.3

Location	Date	Start time	End time	Weather	Air temperature (°C)	Species	Enclosure	Behaviour	Frequency	Duration (hrs)
						(*Castor fiber*)		Sniffing	1	0.33
								Standing	1	0.15
								Walking	12	1.52
Highland Wildlife Park	09/10/17	15:09	16:09	Partial sun	13	Eurasian lynx (*Lynx lynx*)	1	Drinking	1	46
								Feeding	2	8.77
								Walking	6	19.82
								Resting	2	13.88
								Grooming	1	2
								Other	2	0.92
Highland Wildlife Park	NA	NA	NA	NA	NA	Eurasian beaver (*Castor fiber*)	1	Swimming	6	1.07
								Standing	2	0.5
								Sniffing	1	0.08
								Feeding	1	0.27
								Other	1	0.27
Highland Wildlife Park	09/10/17	09:55	12:25	Cloudy	12	Red deer (*Cervus elaphus*)	1	Drinking	1	0.02
								Feeding	4	32.32
								Walking	7	16.62
								Resting	1	23
								Other	1	0.08
								Not visible	5	42.78

Table A2. Summary of directed, random, or other samples collected for each species at wildlife parks.

Site	Species	Enclosure	Sample type	Number of samples	Volume filtered (ml)
Wildwood Trust	European otter (*Lutra lutra*)	1	Targeted	5	500
			Passive	6	500
	European water vole (*Arvicola amphibius*)	1	Other	1	250
		2	Other	1	250
	Eurasian beaver (*Castor fiber*)	1	Targeted	4	150
			Passive	5	150-200
		2	Passive	1	150
	European hedgehog (*Erinaceus europaeus*)	1	Other	1	250
		2	Other	1	250
	European badger (*Meles meles*)	1	Targeted	3	500
			Passive	3	500
			Other	1	500
	Red deer (*Cervus elaphus*)	1	Targeted	6	10-75
			Passive	4	25-150
	Eurasian lynx (*Lynx lynx*)	1	Targeted	1	500
			Passive	2	500
			Other	1	500
	Red squirrel (*Sciurus vulgaris*)	1	Other	1	250
		2	Other	1	250
		3	Other	1	250
		4	Other	1	250
	European pine marten	1	Targeted	3	500

Site	Species	Enclosure	Sample type	Number of samples	Volume filtered (ml)
	(*Martes martes*)		Passive	2	500
			Other	1	500
		2	Targeted	1	500
			Passive	1	500
Highland Wildlife Park	Red squirrel (*Sciurus vulgaris*)	NA	Other	1	500
	Eurasian lynx (*Lynx lynx*)	1	Targeted	1	500
			Passive	2	500
			Other	1	500
	Eurasian beaver (*Castor fiber*)	1	Targeted	3	500
			Passive	3	500
			Other	1	100
	Red deer (*Cervus elaphus*)	1	Targeted	5	50-200
			Passive	5	125-500

Table A3. Summary of samples collected from natural ponds at locations where target species were confirmed as present.

Site	Date	Pond	Sample	Volume filtered (L)
Thorne Moors	17/04/2018	1	1	1.5
			2	1
			3	0.65
			4	0.8
			5	0.85
			6	0.8
			7	1
			8	1.5
			9	0.15
			10	0.1
		2	1	0.175
			2	0.4
			3	0.4
			4	0.5
			5	0.75
			6	0.75
			7	0.3
			8	0.3
			9	0.4
			10	0.6
Bamff Estate	20/04/2018	1	1	0.75
			2	0.45

Site	Date	Pond	Sample	Volume filtered (L)
			3	0.55
			4	1
			5	2
			6	0.9
			7	1
			8	0.95
			9	1.05
			10	0.6
		2	1	0.95
			2	0.95
			3	0.85
			4	1
			5	0.95
			6	0.95
			7	1.1
			8	0.95
			9	0.95
			10	0.95
Tophill Low Nature Reserve	23/04/2018	1	1	0.6
			2	0.625
			3	0.625
			4	0.675
			5	1.1
			6	0.85
			7	1

Site	Date	Pond	Sample	Volume filtered (L)
			8	0.625
			9	0.6
			10	0.65
		2	1	1
			2	1
			3	1
			4	1
			5	0.9
			6	0.9
			7	0.95
			8	0.95
			9	0.95
			10	0.85
	24/04/2018	1	1	0.625
			2	0.8
			3	0.7
			4	0.65
			5	0.9
			6	0.8
			7	0.75
			8	0.65
			9	0.65
			10	0.625
		2	1	1
			2	0.9

Site	Date	Pond	Sample	Volume filtered (L)
			3	0.9
			4	1
			5	0.875
			6	0.825
			7	0.85
			8	1.1
			9	1.2
			10	1.1
	25/04/2018	1	1	0.65
			2	0.75
			3	0.825
			4	0.55
			5	1
			6	0.8
			7	0.8
			8	0.65
			9	0.725
			10	0.55
		2	1	0.9
			2	0.9
			3	1
			4	0.925
			5	0.85
			6	0.775
			7	0.875

Site	Date	Pond	Sample	Volume filtered (L)
			8	1.1
			9	1.1
			10	1
	26/04/2018	1	1	0.55
			2	0.775
			3	0.7
			4	0.7
			5	0.75
			6	0.85
			7	0.8
			8	0.65
			9	0.65
			10	0.725
		2	1	0.9
			2	0.9
			3	1
			4	0.9
			5	0.9
			6	0.775
			7	0.65
			8	0.75
			9	0.85
			10	0.7
	27/04/2018	1	1	0.7
			2	0.65

Site	Date	Pond	Sample	Volume filtered (L)
			3	0.75
			4	0.65
			5	0.85
			6	0.8
			7	0.85
			8	0.7
			9	0.7
			10	0.75
		2	1	0.85
			2	0.8
			3	0.95
			4	0.75
			5	0.8
			6	0.85
			7	0.8
			8	0.85
			9	1.1
			10	0.9

Table A4. List of taxa detected in PCR positive controls by eDNA metabarcoding and corresponding taxon-specific false positive sequence threshold applied.

Taxonomic assignment	Common name	Threshold
Anas spp.	Dabbling ducks	0.00067132
Anatidae	Ducks, geese, swans	0.000100995
Arvicola amphibius	European water vole	0.000342575
Aves	Birds	0.000054
Castor fiber	European beaver	0.003023912
Columba	Pigeons	0.0000877
Corvidae	Corvids	0.000081
Gasterosteidae	Sticklebacks	0.001862034
Homo sapiens	Human	0.000873784
Lynx lynx	Eurasian lynx	0.0000585
Martes martes	European pine marten	0.000906857
Mus musculus	Mouse	0.000107263
Passeriformes	Songbirds	0.0000202
Pelophylax ridibundus	Marsh frog	0.0000743
Phasianidae	Gamebirds	0.000107263
Phoxinus phoxinus	Common minnow	0.000092
Pungitius pungitius	Ninespine stickleback	0.026399055
Rana temporaria	Common frog	0.064393287
Sus scrofa domesticus	Pig	0.000148423
Triturus cristatus	Great crested newt	0.001758274
unassigned	NA	0.009074043

Table A5. Summary of species detected using eDNA metabarcoding across all samples collected in this study.

Common name	Binomial name	Number of samples (*N* = 220)
Red-legged partridge	*Alectoris rufa*	2
Dabbling ducks	*Anas* spp.	38
European eel	*Anguilla anguilla*	6
Grey heron	*Ardea cinerea*	14
European water vole	*Arvicola amphibius*	12
European bison	*Bison bonasus*	2
Cow	*Bos taurus*	44
Common toad	*Bufo bufo*	22
Common buzzard	*Buteo buteo*	3
Dog	*Canis lupus familiaris*	4
Roe deer	*Capreolus capreolus*	4
European beaver	*Castor fiber*	50
Red deer	*Cervus elaphus*	36
Atlantic herring	*Clupea harengus*	24
Rock dove	*Columba livia*	6
Stock dove	*Columba oenas*	6
Common quail	*Coturnix coturnix*	8
Grass carp	*Ctenopharyngodon idella*	1
Buntings	*Emberiza* spp.	1
Horse	*Equus caballus*	28
European hedgehog	*Erinaceus europaeus*	1
European robin	*Erithacus rubecula*	8
Common moorhen	*Gallinula chloropus*	10

Common name	**Binomial name**	**Number of samples (*N* = 220)**
Eurasian jay	*Garrulus glandarius*	1
Three-spined stickleback	*Gasterosteus aculeatus*	18
White-headed gulls	*Larus* spp.	3
Palmate newt	*Lissotriton helveticus*	4
Smooth newt	*Lissotriton vulgaris*	80
European otter	*Lutra lutra*	16
Eurasian lynx	*Lynx lynx*	22
European pine marten	*Martes martes*	16
Turkey	*Meleagris gallopavo*	3
European badger	*Meles meles*	25
Mouse	*Mus musculus*	11
Bank vole	*Myodes glareolus*	2
Eurasian water shrew	*Neomys fodiens*	7
Red-crested pochard	*Netta rufina*	18
European rabbit	*Oryctolagus cuniculus*	37
European smelt	*Osmerus eperlanus*	2
Sheep	*Ovis aries*	9
Great tit	*Parus major*	10
Marsh frog	*Pelophylax ridibundus*	11
Common pheasant	*Phasianus colchicus*	19
Common minnow	*Phoxinus phoxinus*	10
Eurasian magpie	*Pica pica*	7
Common pipistrelle	*Pipistrellus pipistrellus*	1
Ninespine stickleback	*Pungitius pungitius*	7
Common frog	*Rana temporaria*	20

Common name	**Binomial name**	**Number of samples (*N* = 220)**
Brown rat	*Rattus norvegicus*	10
Brown trout	*Salmo trutta*	12
Grey squirrel	*Sciurus carolinensis*	9
Red squirrel	*Sciurus vulgaris*	13
Common shrew	*Sorex araneus*	2
European sprat	*Sprattus sprattus*	5
Tawny owl	*Strix aluco*	1
Common starling	*Sturnus vulgaris*	2
Pig	*Sus scrofa domesticus*	45
European mole	*Talpa europaea*	1
Great crested newt	*Triturus cristatus*	100
Eurasian wren	*Troglodytes troglodytes*	1
Song thrush	*Turdus philomelos*	10
Northern lapwing	*Vanellus vanellus*	1

Table A6. Summary of detection rates for species which were detected by at least one survey method performed at six ponds across three sites in this study.

Species	Lifestyle	Field signs	Camera trapping	eDNA metabarcoding
Common pipistrelle (*Pipistrellus pipistrellus*)	Arboreal	0/6 (0%)	0/6 (0%)	1/6 (16.67%)
Grey squirrel (*Sciurus carolinensis)*	Arboreal	0/6 (0%)	0/6 (0%)	2/6 (33.33%)
Cow (*Bos taurus*)	Ground-dwelling	0/6 (0%)	0/6 (0%)	3/6 (50%)
Sheep (*Ovis aries*)	Ground-dwelling	0/6 (0%)	0/6 (0%)	4/6 (66.67%)
Pig (*Sus scrofa domesticus*)	Ground-dwelling	0/6 (0%)	0/6 (0%)	3/6 (50%)
Dog (*Canis lupus familiaris*)	Ground-dwelling	0/6 (0%)	0/6 (0%)	2/6 (33.33%)
Roe deer (*Capreolus capreolus*)	Ground-dwelling	4/6 (66.67%)	2/6 (33.33%)	3/6 (50%)
Red deer (*Cervus elaphus*)	Ground-dwelling	2/6 (33.33%)	1/6 (16.67%)	1/6 (16.67%)
Red fox (*Vulpes vulpes*)	Ground-dwelling	1/6 (16.67%)	3/6 (50%)	0/6 (0%)
Badger (*Meles meles*)	Ground-dwelling	1/6 (16.67%)	0/6 (0%)	0/6 (0%)
Bank vole (*Myodes glareolus*)	Ground-dwelling	0/6 (0%)	0/6 (0%)	1/6 (16.67%)
Common shrew (*Sorex araneus*)	Ground-dwelling	0/6 (0%)	0/6 (0%)	1/6 (16.67%)
Rabbit (*Oryctolagus cuniculus*)	Ground-dwelling	0/6 (0%)	0/6 (0%)	1/6 (16.67%)
Water vole (*Arvicola amphibius*)	Semi-aquatic	0/2 (0%)	0/2 (0%)	1/6 (16.67%)

Water shrew (*Neomys fodiens*)	Semi-aquatic	0/6 (0%)	0/6 (0%)	2/6 (33.33%)
Brown rat (*Rattus norvegicus*)	Semi-aquatic	0/6 (0%)	0/6 (0%)	1/6 (16.67%)
Beaver (*Castor fiber*)	Semi-aquatic	2/6 (33.33%)	2/6 (33.33%)	2/6 (33.33%)

4. Figures

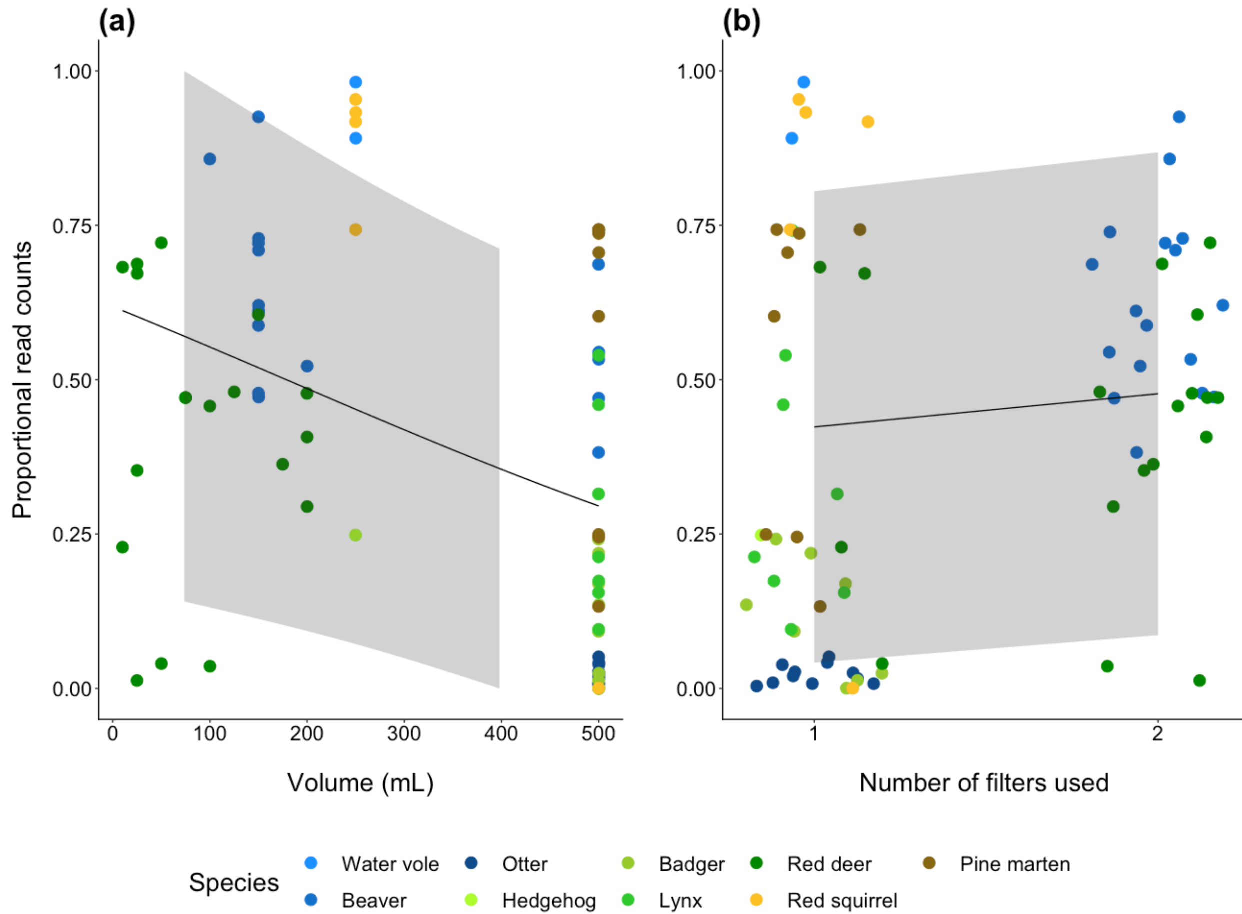


Figure A1. Relationship between the fixed effects (volume and number of filters) and response variable (proportional read counts) as predicted by the binomial GLMM. The 95% CIs, as calculated using the predicted proportional read counts and standard error for these predictions, are given for each relationship. The observed data (points) are displayed against the predicted relationships (lines). Proportional read count marginally decreased as volume of water filtered increased **(a)**, but increased as number of filters used increased **(b)**.

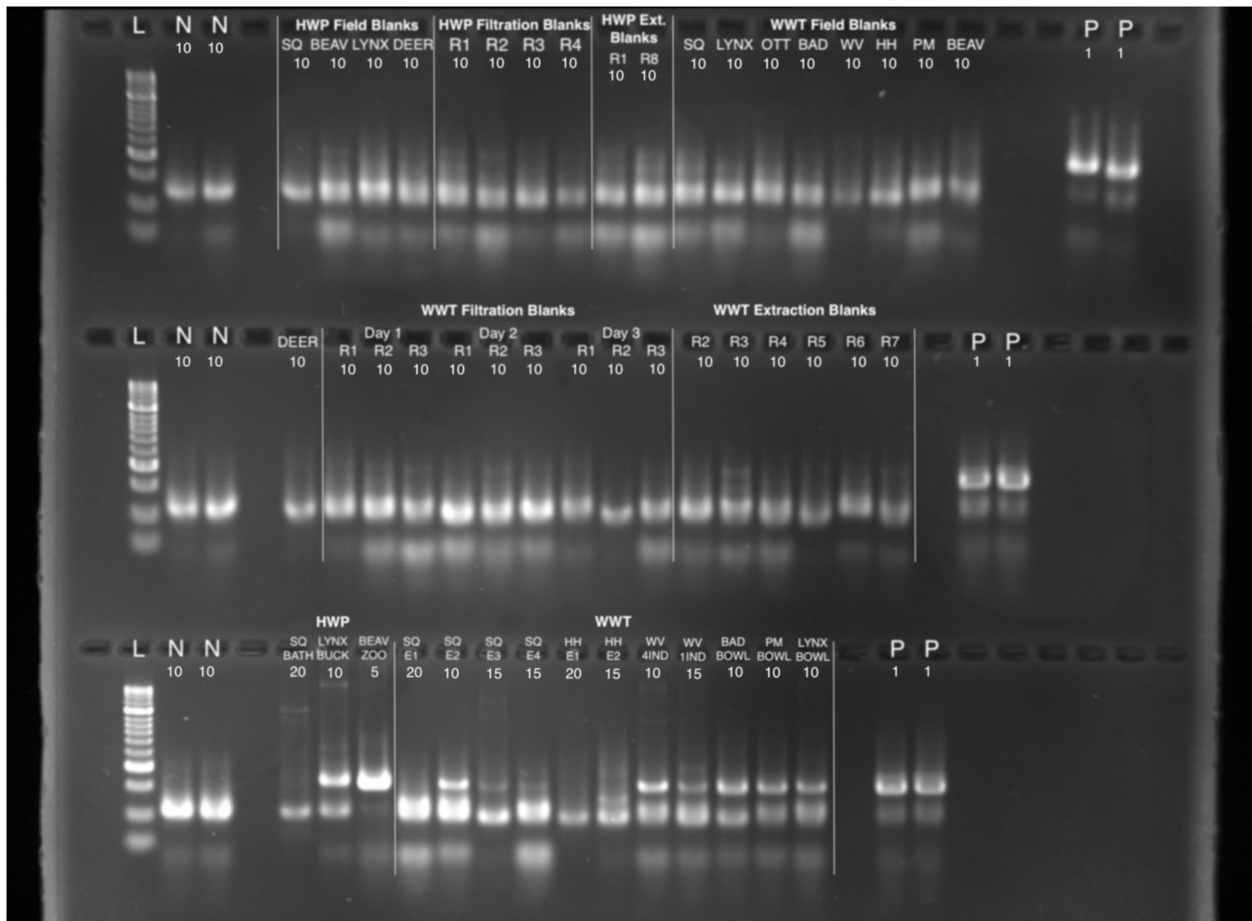


Figure A2. Example gel image of pooled first PCR products which were run on 2% agarose gels with Hyperladder™ 50bp (Bioline®, London, UK) molecular weight marker (L). PCR products were assigned an amplification score based on band strength (0 = no band, 1 = faint band, 2 = bright band, 3 = very bright band). These scores were used to determine how much product should be pooled to create each sub-library (0 = 20 µL, 1 = 15 µL, 2 = 10 µL, 3 = 5 µL). All blanks and PCR negative controls were pooled in consistent volumes (10 µL). Only 1 µL of each PCR positive control was pooled.

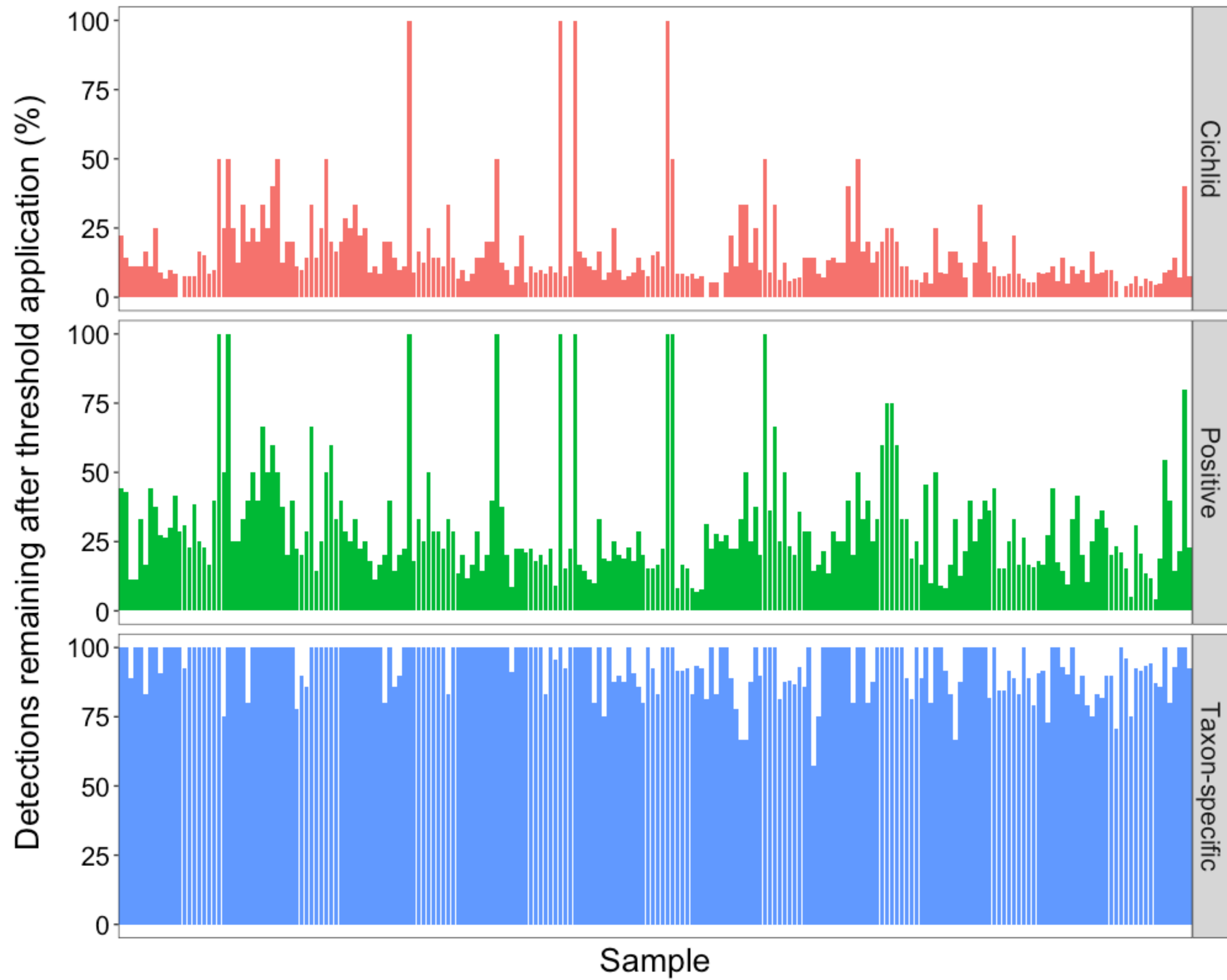


Figure A3. Barplot showing the impact of different false positive sequence thresholds on the proportion of taxa detected in each sample. The taxon-specific thresholds retained the most biological information, thus these were applied to the eDNA metabarcoding data for downstream analyses.

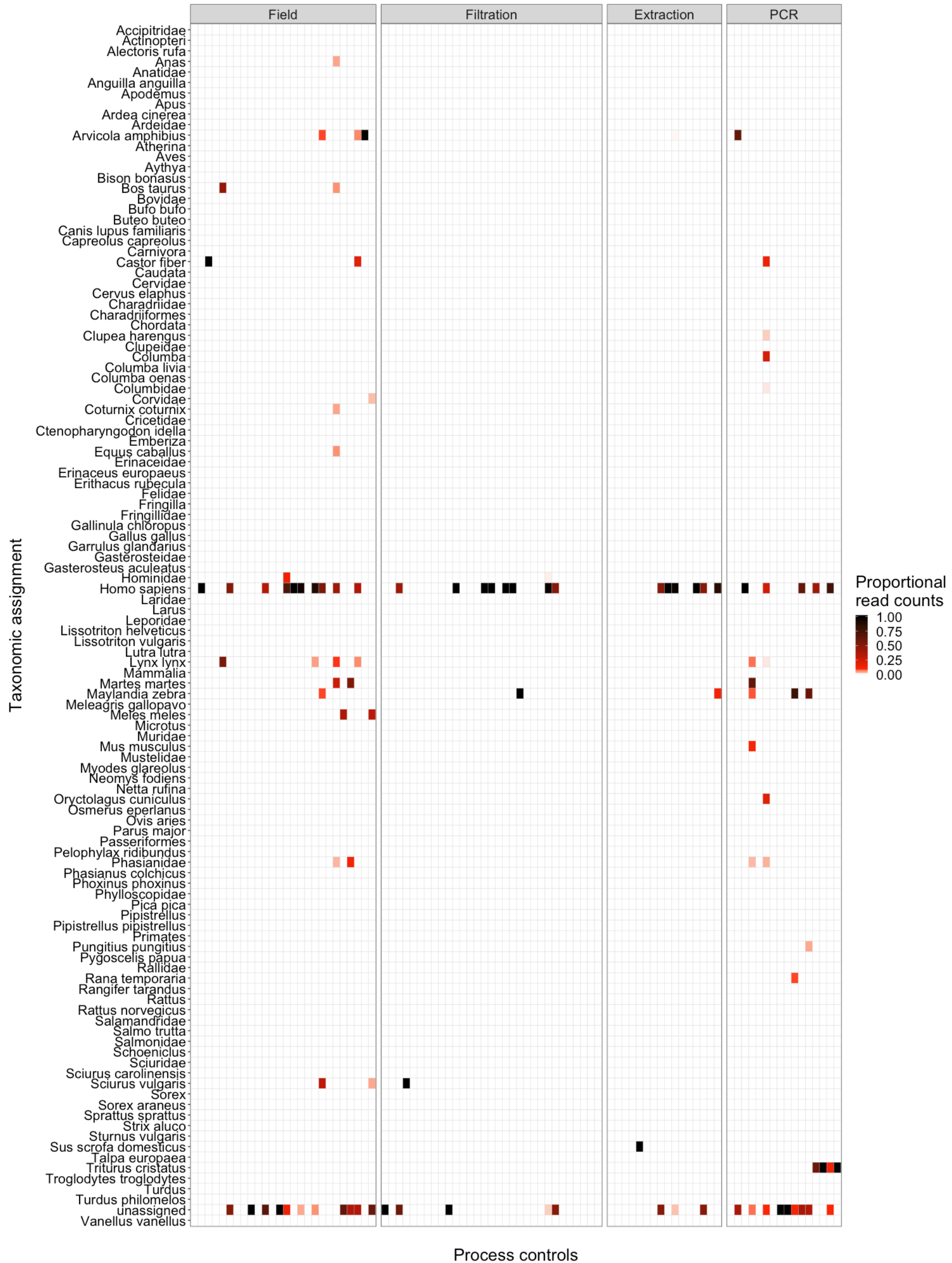


Figure A4. Heatmap showing the frequency of contamination in negative process controls (field blanks, filtration blanks, extraction blanks, and PCR negative controls). Assignments that were not detected in a given process control are coloured white.

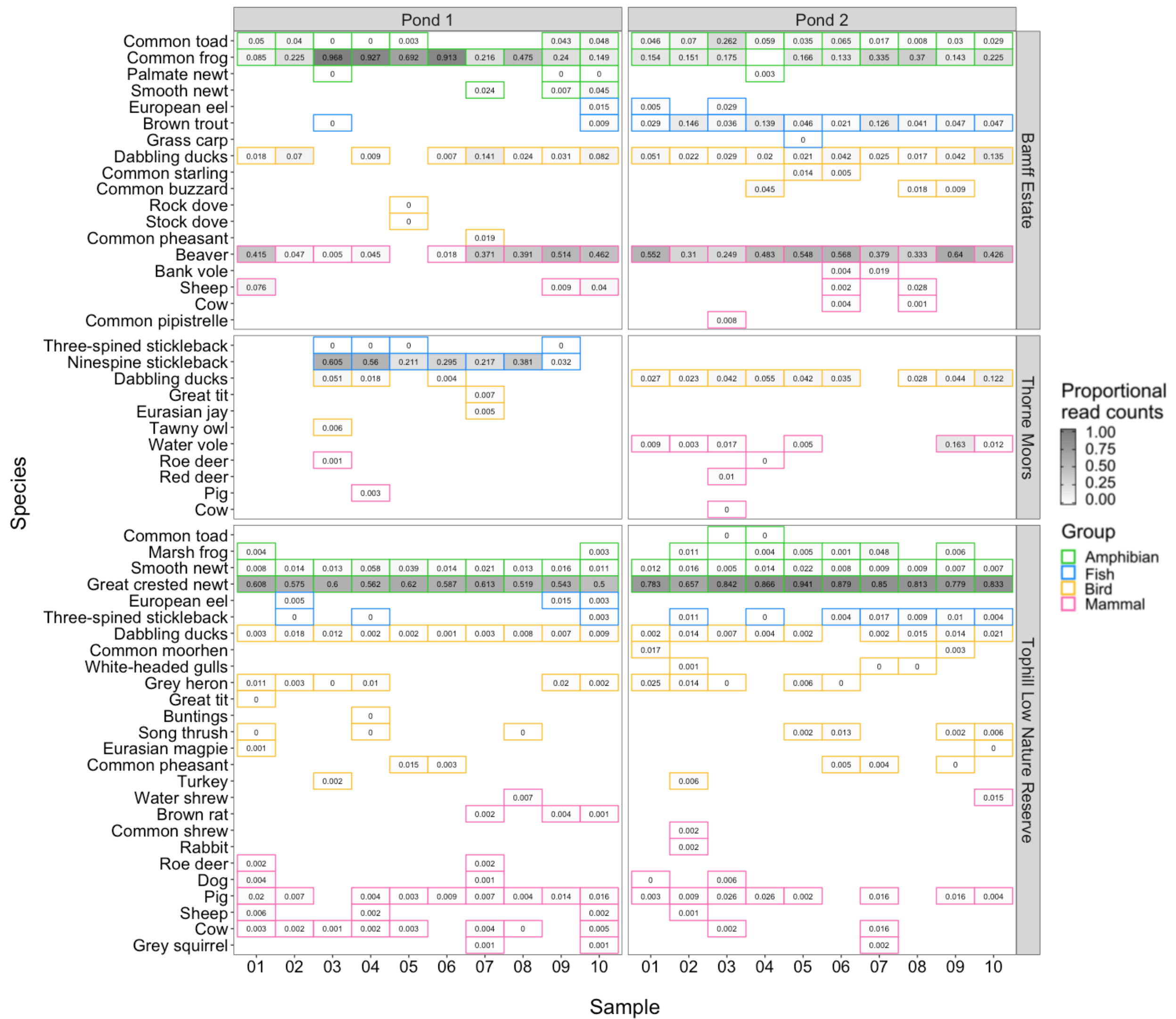


Figure A5. Heatmap showing proportional read counts for samples collected from natural ponds at sites where focal species were present. Each square represents a sample that had reads assigned to a vertebrate species. Species with low proportional read counts (i.e. more than 3 decimal places) are labeled 0.

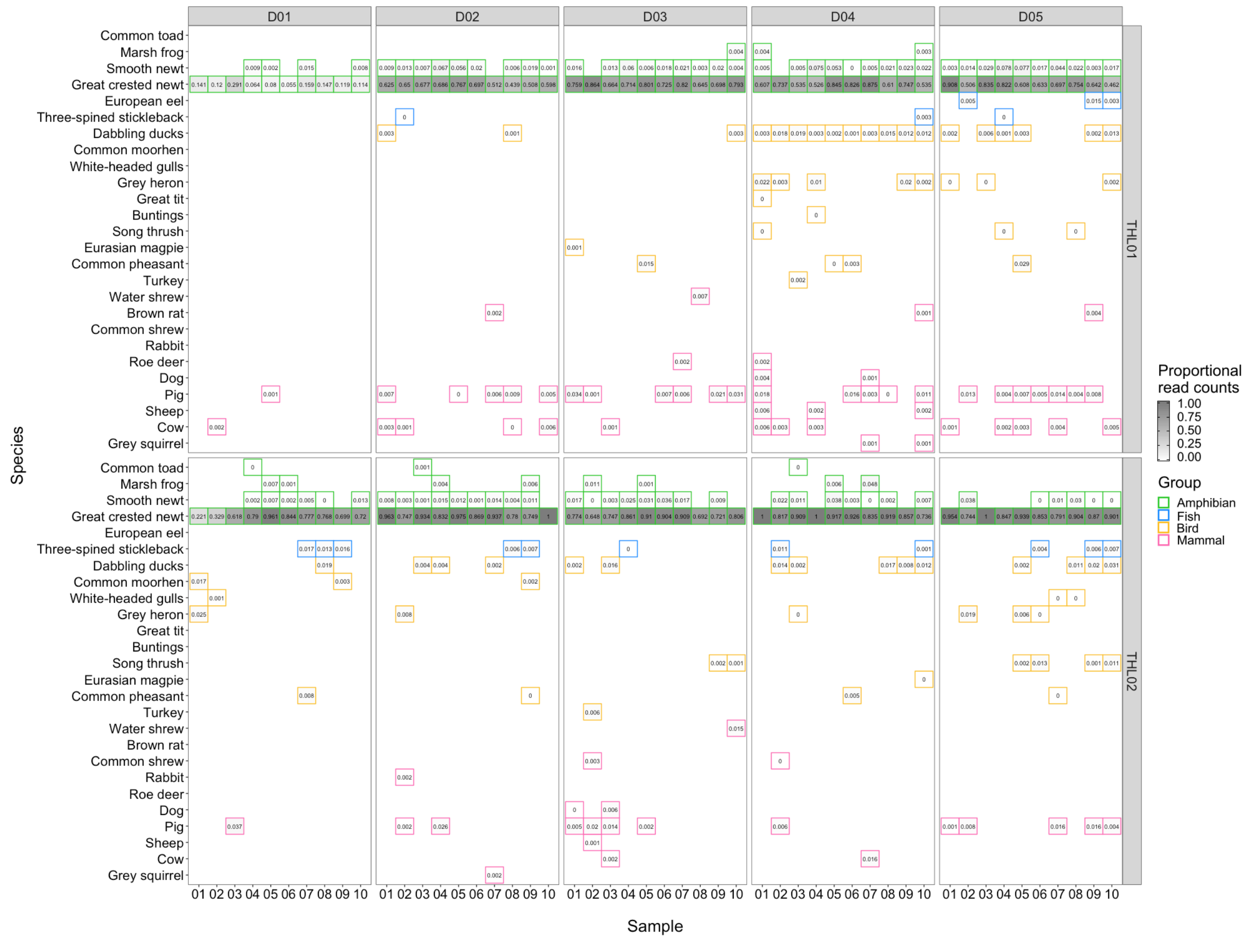


Figure A6. Heatmap showing species detected from samples collected at ponds (THL01 and THL02) within Tophill Low Nature Reserve every 24 hrs over a 5-day period (D01 - D05). Each square represents a sample that had reads assigned to a vertebrate species. Species with low proportional read counts (i.e. more than 3 decimal places) are labeled 0.

References

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

Bronner, I. F., Quail, M. A., Turner, D. J., & Swerdlow, H. (2009). Improved Protocols for Illumina Sequencing. *Current Protocols in Human Genetics*, *18*, 18.2.1–18.2.42. https://doi.org/10.1002/0471142905.hg1802s80

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Machler, M. & Bolker, B. M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, *9*, 378–400. https://doi.org/10.32614/RJ-2017-066

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. https://doi.org/10.1093/bioinformatics/btr381

Evans, N. T., Li, Y., Renshaw, M. A., Olds, B. P., Deiner, K., Turner, C. R., Jerde, C. L., Lodge, D. M., Lamberti, G. A. & Pfrender, M. E (2017). Fish community assessment with eDNA metabarcoding: effects of sampling design and bioinformatic filtering. *Canadian Journal of Fisheries and Aquatic Sciences*, *74*, 1362–1374. https://doi.org/10.1139/cjfas-2016-0306

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., Gielly, L., Lopes, C. M., Boyer, F., Pompanon, F., Rayé, G. & Taberlet, P. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*, *15*, 543–556. https://doi.org/10.1111/1755-0998.12338

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

Harper, L. R., Lawson Handley, L., Hahn, C., Boonham, N., Rees, H. C., Gough, K. C., Lewis, E., Adams, I. P., Brotherton, P., Phillips, S. & Hänfling, B. (2018). Needle in a haystack? A comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (*Triturus cristatus*). *Ecology and Evolution*, *8*, 6330–6441. https://doi.org/10.1002/ece3.4013

Hosmer, D. W., & Lemeshow, S. (2000). Multiple Logistic Regression. In *Applied Logistic Regression* (pp. 31–46). John Wiley & Sons, Inc.

Kelly, R. P., Port, J. A., Yamahara, K. M., & Crowder, L. B. (2014). Using environmental DNA to census marine fishes in a large mesocosm. *PLoS ONE*, *9*, e86175. https://doi.org/10.1371/journal.pone.0086175

Klymus, K. E., Richter, C. A., Thompson, N., & Hinck, J. E. (2017). Metabarcoding of Environmental DNA Samples to Explore the Use of Uranium Mine Containment Ponds as a Water Source for Wildlife. *Diversity*, *9*, 54. https://doi.org/10.3390/d9040054

Lele, S. R., Keim, J. L., & Solymos, P. (2016). ResourceSelection: Resource Selection (Probability) Functions for Use-Availability Data. https://CRAN.R-project.org/package=ResourceSelection

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

Port, J. A., O'Donnell, J. L., Romero-Maraccini, O. C., Leary, P. R., Litvin, S. Y., Nickols, K. J., Yamahara, K. M. & Kelly, R. P. (2016). Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Molecular Ecology*, *25*, 527–541. https://doi.org/10.1111/mec.13481

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, *39*, e145. https://doi.org/10.1093/nar/gkr732

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

Ushio, M., Fukuda, H., Inoue, T., Makoto, K., Kishida, O., Sato, K., Murata, K., Nikaido, M., Sado, T., Sato, Y., Takeshita, M., Iwasaki, W., Yamanaka, H., Kondoh, M. & 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

Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. Springer, New York, USA. http://ggplot2.org

Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology*, *7*, 203–214. https://doi.org/10.1089/10665270050081478

Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A. & Smith, G. M. (2009) *Mixed effects models and extensions in ecology with R*. Springer, New York, USA.

Appendix B

Wildwood Trust, Canterbury, Kent

Otter (*Lutra lutra*)

Beaver (*Castor fiber*)

Water vole (*Arvicola amphibius*)

Hedgehog (*Erinaceus europaeus*)

Badger (*Meles meles*)

Lynx (*Lynx lynx*)

Red deer (*Cervus elaphus*)

Red squirrel (*Sciurus vulgaris*)

Pine marten (*Martes martes*)

Highland Wildlife Park, Kingussie, Scotland

Water bath in woods

Lynx (*Lynx lynx*)

Red deer (*Cervus elaphus*)