Generating and testing ecological hypotheses at the pondscape with environmental DNA metabarcoding: a case study on a threatened amphibian

Lynsey R. Harper[1,2*], Lori Lawson Handley[1], Christoph Hahn[1,3], Neil Boonham[4,5], Helen C. Rees[6], Erin Lewis[4], Ian P. Adams[4], Peter Brotherton[7], Susanna Phillips[7] and Bernd Hänfling[1]

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

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

[3] Institute of Zoology, University of Graz, Graz, Styria, Austria

[4] Fera Science Ltd (Fera), Sand Hutton, York, YO14 1LZ, UK

[5] Newcastle University, Newcastle upon Tyne, NE1 7RU, UK

[6] ADAS, School of Veterinary Medicine and Science, The University of Nottingham, Sutton Bonington Campus, Leicestershire, LE12 5RD, UK

[7] Natural England, Peterborough, PE1 1NG, UK

***Corresponding author: Lynsey R. Harper**

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

E-mail: lynsey.harper2@gmail.com

Running title: Hypothesis testing with eDNA metabarcoding

Abstract

Environmental DNA (eDNA) metabarcoding is revolutionising biodiversity monitoring, but has unrealised potential for ecological hypothesis generation and testing. Here, we validate this potential in a large-scale analysis of vertebrate community data generated by eDNA metabarcoding of 532 UK ponds. We test biotic associations between the threatened great crested newt (*Triturus cristatus*) and other vertebrates as well as abiotic factors influencing *T. cristatus* detection at the pondscape. Furthermore, we test the status of *T. cristatus* as an umbrella species for pond conservation by assessing whether vertebrate species richness is greater in ponds with *T. cristatus* and higher *T. cristatus* Habitat Suitability Index (HSI) scores. *T. cristatus* detection was positively correlated with amphibian and waterfowl species richness. Specifically, *T. cristatus* was positively associated with smooth newt (*Lissotriton vulgaris*), common coot (*Fulica atra*), and common moorhen (*Gallinula chloropus*), but negatively associated with common toad (*Bufo bufo*). *T. cristatus* detection did not significantly decrease as fish species richness increased, but negative associations with common carp (*Cyprinus carpio*), three-spined stickleback (*Gasterosteus aculeatus*) and ninespine stickleback (*Pungitius pungitius*) were identified. *T. cristatus* detection was negatively correlated with mammal species richness, and *T. cristatus* was negatively associated with grey squirrel (*Sciurus carolinensis*). *T. cristatus* detection was negatively correlated with larger pond area, presence of inflow, and higher percentage of shading, but positively correlated with HSI score, supporting its application to *T. cristatus* survey. Vertebrate species richness was significantly higher in *T. cristatus* ponds and broadly increased as *T. cristatus* HSI scores increased. We reaffirm reported associations (e.g. *T. cristatus* preference for smaller ponds) but also provide novel insights, including a negative effect of pond inflow on *T. cristatus*. Our findings demonstrate the prospects of eDNA metabarcoding for ecological hypothesis generation and testing at landscape scale, and dramatic enhancement of freshwater conservation, management, monitoring and research.

Keywords: biodiversity assessment, environmental DNA (eDNA), hypothesis testing, metabarcoding, ponds, species associations, *Triturus cristatus*

52 ## 1. Introduction

53

54 Environmental DNA (eDNA) analysis offers ecologists exceptional power to detect organisms
55 within and across ecosystems. DNA released by organisms into their environment via
56 secretions, excretions, gametes, blood, or decomposition, can be sampled and analysed
57 using different approaches to reveal the distribution of single or multiple species (Rees et
58 al., 2014; Lawson Handley, 2015). eDNA analysis combined with high-throughput
59 sequencing (i.e. eDNA metabarcoding) can yield efficient, comprehensive assessments of
60 entire communities (Deiner et al., 2017), providing a step change in biodiversity monitoring
61 (Hering et al., 2018). eDNA metabarcoding has untapped potential to generate and test
62 ecological hypotheses by enabling biodiversity monitoring at landscape scale with minimal
63 impact to communities under investigation. This potential has already been demonstrated
64 with targeted eDNA analysis by Wilcox et al. (2018), where climate-mediated responses of
65 bull trout (*Salvelinus confluentus*) to biotic and abiotic factors were revealed using
66 quantitative PCR (qPCR) on crowd-sourced eDNA samples. Although eDNA metabarcoding
67 assessments of alpha and beta diversity along environmental gradients are increasing (e.g.
68 Hänfling et al., 2016; Olds et al., 2016; Kelly et al., 2016; Evans et al., 2017; Li et al., 2018a;
69 Nakagawa et al., 2018), this tool is less commonly used for ecological hypothesis testing,
70 such as the impact of environmental stressors (Li et al., 2018b; Macher et al., 2018).

71 Aquatic ecosystems are highly suited to eDNA studies as eDNA exists in multiple
72 states with rapid modes of transport and degradation, increasing detectability of
73 contemporary biodiversity (Rees et al., 2014; Barnes & Turner, 2015). Lentic systems
74 provide further opportunities for eDNA research, being discrete water bodies with variable
75 physicochemical properties that do not experience flow dynamics (Harper et al., 2019).

76 Ponds in particular have enormous biodiversity and experimental virtue that has not been
77 maximised in previous eDNA metabarcoding assessments of this habitat (Valentini et al.,
78 2016; Evans et al., 2017; Klymus et al., 2017; Ushio et al., 2017; Bálint et al., 2018). These
79 small and abundant water bodies span broad ecological gradients (De Meester et al., 2005)
80 and comprise pondscapes - a network of ponds and their surrounding terrestrial habitat (Hill
81 et al., 2018). Pondscapes contribute substantially to aquatic and non-aquatic biodiversity
82 across spatial scales, with ponds supporting many rare and protected species in fragmented
83 landscapes (De Meester et al., 2005; Biggs et al., 2016; Hill et al., 2018). Consequently,
84 ponds are model systems for experimental validation and examination of biogeographical
85 patterns (De Meester et al., 2005). Habitat complexity and tools required for different taxa
86 with associated bias (Evans et al., 2017) and cost (Valentini et al., 2016) once hindered
87 exhaustive sampling of pond biodiversity (Hill et al., 2018), but eDNA metabarcoding may
88 overcome these barriers (Harper et al., 2019).

89 In the UK, the threatened great crested newt (*Triturus cristatus*) is an umbrella
90 species for pond conservation. The extensive literature on *T. cristatus* ecology provides an
91 excellent opportunity to validate ecological patterns revealed by eDNA metabarcoding. Both
92 biotic (e.g. breeding substrate, prey, and predators) and abiotic (e.g. pond area, depth, and
93 temperature) factors are known to influence *T. cristatus* breeding success (Langton, Beckett
94 & Foster, 2001). The *T. cristatus* Habitat Suitability Index (HSI [Oldham et al., 2000; ARG-UK,
95 2010; O'Brien et al., 2017]) accounts for these factors using 10 suitability indices that are
96 scored and combined to calculate a decimal score between 0 and 1 (where 1 = excellent
97 habitat). Larvae are susceptible to fish and waterfowl predation (Edgar & Bird, 2006; Rannap
98 & Briggs, 2006; Skei et al., 2006; Hartel, Nemes & Oellerer, 2010), and adults reportedly
99 avoid ponds containing three-spined stickleback (*Gasterosteus aculeatus*) (McLee & Scaife,

100 1992), ninespine stickleback (*Pungitius pungitius*), crucian carp (*Carassius carassius*), and
101 common carp (*Carassius carpio*) (Rannap, Lõhmus & Briggs, 2009a, b). Conversely, *T.*
102 *cristatus* and smooth newt (*Lissotriton vulgaris*) prefer similar habitat and often co-occur
103 (Rannap & Briggs, 2006; Skei et al., 2006; Rannap et al., 2009a; Denoël et al., 2013; Cayuela
104 et al., 2018). *T. cristatus* individuals thrive in ponds with good water quality as indicated by
105 diverse macroinvertebrate communities (Oldham et al., 2000; Rannap et al., 2009a). Pond
106 networks encourage *T. cristatus* occupancy (Joly et al., 2001; Rannap et al., 2009a; Hartel et
107 al., 2010; Denoël et al., 2013), but larger pond area discourages presence (Joly et al., 2001).
108 Ponds with heavy shading (Vuorio, Heikkinen & Tikkanen, 2013) or dense macrophyte cover
109 (Rannap & Briggs, 2006; Skei et al., 2006; Hartel et al., 2010) are unlikely to support viable
110 populations. *T. cristatus* individuals also depend on terrestrial habitat, preferring open,
111 semi-rural pondscapes (Denoël et al., 2013) containing pasture, extensively grazed and
112 rough grassland, scrub, and coniferous and deciduous woodland (Oldham et al., 2000;
113 Rannap & Briggs, 2006; Rannap et al., 2009a; Gustafson, Malmgren & Mikusiński, 2011;
114 Vuorio et al., 2013).

115 We assessed vertebrate communities at the pondscape using a dataset produced by
116 eDNA metabarcoding for over 500 ponds with comprehensive environmental metadata. We
117 validated eDNA metabarcoding as a tool for ecological hypothesis generation and testing,
118 and compared its outputs to previous results produced by established methods. Specifically,
119 we identified and tested biotic (community presence-absence data) and abiotic
120 (environmental metadata on ponds and surrounding terrestrial habitat) factors influencing
121 *T. cristatus* detection at the pondscape - an impractical task by conventional means.
122 Furthermore, we tested the applicability of the HSI to predict eDNA-based *T. cristatus*
123 detection. Finally, we assessed the umbrella species status of *T. cristatus* by investigating

124 whether *T. cristatus* detection and the *T. cristatus* HSI score can predict vertebrate species
125 richness of ponds.
126
127
128 ## 2. Materials and methods
129
130 **2.1 eDNA samples**
131
132 Samples were previously collected from 508 ponds as part of Natural England's Great
133 Crested Newt Evidence Enhancement Programme and from 24 privately surveyed ponds
134 using established methodology (Biggs et al., 2015), detailed in Supporting Information:
135 Appendix 1. Briefly, 20 x 30 mL water samples were collected from each pond and pooled.
136 Six 15 mL subsamples were taken from the pooled sample and each added to 33.5 mL
137 absolute ethanol and 1.5 mL sodium acetate 3 M (pH 5.2). Subsamples were pooled during
138 DNA extraction to produce one eDNA sample per pond. Targeted qPCR detected *T. cristatus*
139 in 265 (49.81%) ponds (Harper et al., 2018).
140 Environmental metadata were collected for 504 of 532 ponds (Fig. S1) by
141 environmental consultants contracted for Natural England's Great Crested Newt Evidence
142 Enhancement Programme. Metadata included: maximum depth of ponds; pond
143 circumference; pond width; pond length; pond area; pond density (i.e. number of ponds per
144 km^2); terrestrial overhang; shading; macrophyte cover; HSI score (Oldham et al., 2000); HSI
145 band (categorical classification of HSI score [ARG-UK, 2010]); pond permanence; water
146 quality; pond substrate; presence of inflow or outflow; presence of pollution; presence of
147 other amphibians, fish and waterfowl; woodland; rough grass; scrub/hedge; ruderals; other

148 terrestrial habitat (i.e. good quality terrestrial habitat that did not conform to
149 aforementioned habitat types); and overall terrestrial habitat quality (see Table S1).
150
151 **2.2. eDNA metabarcoding**
152
153 We repurposed the taxonomically assigned sequence reads from Harper et al. (2018) that
154 were produced using eDNA metabarcoding of pond water to compare qPCR and eDNA
155 metabarcoding for *T. cristatus* detection. Here, we provide a summary of their eDNA
156 metabarcoding workflow (see Supporting Information: Appendix 1 for details).
157 A custom, phylogenetically curated reference database of mitochondrial 12S
158 ribosomal RNA (rRNA) sequences for UK fish species was previously constructed for eDNA
159 metabarcoding of lake fish communities (Hänfling et al., 2016). Harper et al. (2018)
160 constructed additional reference databases for UK amphibians, reptiles, birds, and
161 mammals. Reference sequences available for species varied across vertebrate groups:
162 amphibians 100.00% (*N* = 21), reptiles 90.00% (*N* = 20), mammals 83.93% (*N* = 112), and
163 birds 55.88% (*N* = 621). Table S2 lists species without database representation, i.e. no
164 records for any species in a genus. Sanger sequences were obtained from tissue of *T.*
165 *cristatus*, *L. vulgaris*, Alpine newt (*Ichthyosaura alpestris*), common toad (*Bufo bufo*), and
166 common frog (*Rana temporaria*) to supplement the amphibian database. The complete
167 reference databases compiled in GenBank format were deposited in a GitHub repository
168 and permanently archived (https://doi.org/10.5281/zenodo.1188710) by Harper et al.
169 (2018). Reference databases were combined for *in silico* validation of published 12S rRNA
170 primers 12S-V5-F (5'-ACTGGGATTAGATACCCC-3') and 12S-V5-R (5'-
171 TAGAACAGGCTCCTCTAG-3') (Riaz et al., 2011) using ecoPCR software (Ficetola et al., 2010).

172 Set parameters allowed a 50-250 bp fragment and three mismatches between each primer
173 and reference sequence. Primers were validated *in vitro* for UK fish by Hänfling et al. (2016)
174 and for six UK amphibian species by Harper et al. (2018) (Fig. S2).

175 eDNA samples were amplified in triplicate with the aforementioned primers, where
176 PCR positive controls (six per PCR plate; *n* = 114) were cichlid (*Rhamphochromis esox*) DNA
177 (0.284 ng/µL) and PCR negative controls (six per PCR plate; *n* = 114) were sterile molecular
178 grade water (Fisher Scientific UK Ltd, UK). PCR products were individually purified using
179 E.Z.N.A® Cycle Pure V-Spin Clean-Up Kits (Omega Bio-tek, GA, USA) following the
180 manufacturer's protocol. Purified product was pooled according to sample for the second
181 PCR which bound Multiplex Identification tags to the purified products. PCR products were
182 individually purified using magnetic bead clean-up and quantified with a Quant-IT™
183 PicoGreen™ dsDNA Assay (Invitrogen, UK). Samples were normalised, pooled, and libraries
184 quantified using a Qubit™ dsDNA HS Assay (Invitrogen, UK). Libraries were sequenced on an
185 Illumina® MiSeq using 2 x 300 bp V3 chemistry (Illumina, Inc, CA, USA) and raw sequence
186 reads processed using metaBEAT (metaBarcoding and Environmental Analysis Tool) v0.97.7
187 (https://github.com/HullUni-bioinformatics/metaBEAT).

188 After quality filtering (phred score Q30) and trimming (length filter of 90-110 bp)
189 using Trimmomatic v0.32 (Bolger, Lohse, & Usadel, 2014), merging (10 bp overlap minimum
190 and 10% mismatch maximum) using FLASH v1.2.11 (Magoč & Salzberg, 2011), chimera
191 detection using the uchime algorithm (Edgar et al., 2011) in vsearch v1.1 (Rognes et al.,
192 2016), and clustering (97% identity with 3 sequences minimum per cluster) with vsearch
193 v1.1 (Rognes et al., 2016), non-redundant query sequences were compared against the
194 reference database using BLAST (Zhang et al., 2000). Putative taxonomic identity was
195 assigned using a lowest common ancestor (LCA) approach based on the top 10% BLAST

196 matches for any query matching with at least 98% identity to a reference sequence across
197 more than 80% of its length. Unassigned sequences were subjected to a separate BLAST
198 against the complete NCBI nucleotide (nt) database at 98% identity to determine the source
199 via LCA as described above. The bioinformatic analysis was archived
200 (https://doi.org/10.5281/zenodo.1188710) by Harper et al. (2018) for reproducibility.
201
202 **2.3 Data analysis**
203
204 *2.3.1 Dataset refinement*
205
206 Analyses were performed in R v.3.4.3 (R Core Team, 2017). Data and R scripts have been
207 deposited in a dedicated GitHub repository for this study, which has been permanently
208 archived (https://doi.org/10.5281/zenodo.3365703). Assignments from different databases
209 were merged, and spurious assignments (i.e. non-UK species, invertebrates and bacteria)
210 removed from the dataset. The family Cichlidae was reassigned to *Rhamphochromis esox*.
211 The green-winged teal (*Anas carolinenisis*) was reassigned to *Anas* (dabbling ducks) because
212 this species is a rare migrant and available reference sequences were identical across our
213 metabarcode to those for mallard (*Anas platyrhynchos*) and Eurasian teal (*Anas crecca*),
214 which are widely distributed across the UK. Scottish wildcat (*Felis silvestris*) does not occur
215 at the sampling localities (Kent, Lincolnshire and Cheshire) and was therefore reassigned to
216 domestic cat (*Felis catus*). Wild boar (*Sus scrofa*) and grey wolf (*Canis lupus*) were reassigned
217 to domestic pig (*Sus scrofa domesticus*) and domestic dog (*Canis lupus familiaris*) given the
218 restricted distribution of *S. scrofa* and absence of *C. lupus* in the wild in the UK. The genus
219 *Strix* was reassigned to tawny owl (*Strix aluco*) as it is the only UK representative of this

genus. Where family and genera assignments containing a single UK representative had reads assigned to species, reads from all assignment levels were merged and manually assigned to that species.

Of the 114 PCR negative controls included, 50 produced no reads (Fig. S3). Reads generated for 64 of 114 PCR negative controls ranged from 0 to 49,227, and strength of each contaminant varied (mean = 0.021%, range = 0 - 100.0% of the total reads per PCR negative control). To minimise risk of false positives, we evaluated different sequence thresholds. These included the maximum sequence frequency of *R. esox* DNA in eDNA samples (100%), maximum sequence frequency of any DNA except *R. esox* in PCR positive controls (83.96%), 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 S3) 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 their threshold. After applying the taxon-specific thresholds, the PCR positive control (*R. esox*), human (*Homo sapiens*), and domestic species (Table S4) were removed from the dataset. Higher taxonomic assignments excluding the genus *Anas* were then removed, thus taxonomic assignments in the final dataset were predominantly of species resolution. The read count data were converted to a binary species matrix for downstream analyses using the package cooccur v1.3 (Griffith, Veech & Marsh, 2016) and Generalized Linear Mixed-effects Models (GLMMs), which were fit by maximum likelihood with the glmer function using the package lme4 v1.1-21 (Bates et al., 2015). County was included as a random effect in each GLMM to account for spatial dependencies between ponds (see Supporting Information: Appendix 1).

244

245 *2.3.2 Hypothesis generation: biotic factors potentially influencing T. cristatus detection*

246

247 We first mined the eDNA metabarcoding data and investigated the influence of vertebrate
248 group richness on *T. cristatus* detection. We used a binomial GLMM with the logit link
249 function that included species richness of other amphibians, fish, waterfowl, terrestrial
250 birds, and mammals as fixed effects (model 1, N = 532). We then performed a preliminary
251 analysis using the package cooccur v1.3 (Griffith, Veech & Marsh, 2016) to identify species
252 associations between *T. cristatus* and other vertebrates (N = 532). Hypotheses generated by
253 the eDNA metabarcoding data relating to factors influencing *T. cristatus* detection are
254 summarised in Table 1.

255

256 *2.3.3 Hypothesis testing: biotic and abiotic factors known to influence T. cristatus detection*

257

258 Identified associations from the cooccur analysis in conjunction with the existing *T. cristatus*
259 literature informed candidate biotic variables to be modelled against *T. cristatus* detection
260 (model 2, n = 504). The existing *T. cristatus* literature informed candidate abiotic variables to
261 be modelled against *T. cristatus* detection (model 3, n = 504). Hypotheses tested using the
262 eDNA metabarcoding data and associated environmental metadata to examine biotic and
263 abiotic factors influencing *T. cristatus* detection are summarised in Tables 2 and S5.
264 Selection of a suitable set of explanatory variables and modelling framework is fully
265 described in Supporting Information: Appendix 1. Briefly, candidate biotic and abiotic
266 explanatory variables were assessed for: collinearity using a Spearman's rank pairwise
267 correlation matrix and examination of variance inflation factors (Zuur et al., 2009), relative

importance using a classification tree within the package rpart v4.1-13 (Therneau, Atkinson & Ripley, 2014), and non-linearity using Generalized Additive Models (GAMs) (Zuur et al., 2009). We constructed separate binomial GLMMs with the logit link function for biotic (model 2) and abiotic (model 3) explanatory variables that passed these assessments.

2.3.4 Hypothesis testing: T. cristatus HSI and umbrella status

We modelled HSI score (fixed effect) separately to prevent HSI score masking variation caused by the individual biotic and abiotic variables it encompasses (model 4). Using a Poisson GLMM, we tested the umbrella species status of *T. cristatus* by modelling vertebrate species richness against *T. cristatus* detection and the *T. cristatus* HSI score (model 5).

2.3.5 Model selection and validation

For each GLMM, we employed an information-theoretic approach using Akaike Information Criterion (AIC) to determine the most parsimonious approximating model to make predictions (Akaike, 1973). Biotic and abiotic models considered respectively (Table S5) were nested thus the best models were chosen using stepwise backward deletion of terms based on Likelihood Ratio Tests (LRTs). Models 1, 4 and 5 were compared to null GLMMs. The final models resulting from stepwise selection were as follows:

Model 1: *T. cristatus* ~ (1|County) + amphibian richness + fish richness + waterfowl richness + terrestrial bird richness + mammal richness

Model 2: *T. cristatus* ~ (1|County) + *L. vulgaris* + *B. bufo* + *C. carpio* + *G. aculeatus* + *Gallinula chloropus* + Waterfowl

Model 3: *T. cristatus* ~ (1|County) + inflow + pond area + shading

Model 4: *T. cristatus* ~ (1|County) + HSI score

Model 5: Vertebrate species richness ~ (1|County) + *T. cristatus* + HSI score

Final models were tested for overdispersion using a custom function testing overdispersion of the Pearson residuals. Model fit was assessed using the Hosmer and Lemeshow Goodness of Fit Test within the package ResourceSelection v0.3-2 (Lele, Keim & Solymos, 2016), quantile-quantile plots, and partial residual plots (Zuur et al., 2009). Conditional R^2 was calculated using the r.squaredGLMM function in the package MuMIn v1.42.1 (Bartoń, 2018). The Benjamini-Hochberg correction was calculated for explanatory variable p-values from the cooccur and GLMM analyses using the p.adjust function in the package stats v3.4.3, and this alpha applied to the results to account for potential Type I error. The original p-values and the adjusted Benjamini-Hochberg p-values are reported for comparison in Tables 1 and 2. Model predictions were obtained using the predictSE function in the package AICcmodavg v2.1-1 (Mazerolle, 2016) and upper and lower 95% CIs were calculated from the standard error of the predictions. Results were plotted using the package ggplot2 v3.1.0 (Wickham, 2016).

313 ## 3. Results

314

315 ### 3.1 eDNA metabarcoding

316

317 Harper et al. (2018) processed a total of 532 eDNA samples and 228 PCR controls across two
318 sequencing runs. The runs generated raw sequence read counts of 36,236,862 and
319 32,900,914 respectively. After quality filtering, trimming, and merging of paired-end reads,
320 26,294,906 and 26,451,564 sequences remained. Following removal of chimeras and read
321 dereplication via clustering, the libraries contained 14,141,237 and 14,081,939 sequences
322 (average read counts of 36,826 and 36,671 per sample respectively), of which 13,126,148
323 and 13,113,143 sequences were taxonomically assigned. The final dataset used here
324 (assignments corrected, thresholds applied, and assignments removed) contained 53
325 vertebrate species (Table S6), including six amphibians, 14 fish, 17 birds, and 16 mammals
326 (Fig. 1).

327

328 ### 3.2 Pondscape biodiversity

329

330 All native amphibians were detected as well as the non-native marsh frog (*Pelophylax*
331 *ridibundus*). *T. cristatus* (*n* = 148), *L. vulgaris* (*n* = 151) and *R. temporaria* (*n* = 122) were
332 widespread, but *B. bufo* (*n* = 42), palmate newt (*Lissotriton helveticus* [*n* = 5]) and *P.*
333 *ridibundus* were uncommon (*n* = 1). The threatened European eel (*Anguilla anguilla* [*n* =
334 15]), European bullhead (*Cottus gobio* [*n* = 14]), and *C. carassius* (*n* = 2) were detected
335 alongside native fishes, such as pike (*Esox Lucius* [n = 17]) and roach (*Rutilus rutilus* [*n* = 71]),
336 but also introduced species, including *C. carpio* (*n* = 40), ruffe (*Gymnocephalus cernua* [*n* =

337 1]), and rainbow trout (*Oncorhynchus mykiss* [*n* = 3]). Some identified waterfowl were
338 ubiquitous, such as common moorhen (*Gallinula chloropus* [*n* = 211]), whereas others were
339 less common, e.g. grey heron (*Ardea cinerea* [*n* = 1]) and Eurasian oystercatcher
340 (*Haematopus ostralegus* [*n* = 1]). Terrestrial fauna were often detected in fewer than five
341 ponds (Figs. 1c, d). Buzzard (*Buteo buteo* [*n* = 4]), Eurasian jay (*Garrulus glandarius* [*n* = 7]),
342 dunnock (*Prunella modularis* [*n* = 4]), and starling (*Sturnus vulgaris* [*n* = 4]) were the most
343 frequently detected terrestrial birds. Introduced mammals (Mathews et al., 2018), such as
344 grey squirrel (*Sciurus carolinensis* [*n* = 57]) and Reeve's muntjac (*Muntiacus reevesi* [*n* = 3]),
345 outweighed native mammals. Nonetheless, we detected several mammals with Biodiversity
346 Actions Plans and/or of conservation concern (Mathews et al., 2018), including otter (*Lutra*
347 *lutra* [*n* = 1]), water vole (*Arvicola amphibius* [*n* = 16]), European polecat (*Mustela putorius*
348 [*n* = 1]), brown hare (*Lepus europaeus* [*n* = 1]) and water shrew (*Neomys fodiens* [*n* = 8]).
349 Notably, the invasive American mink (*Neovison vison*) was not detected despite widespread
350 UK distribution (Mathews et al., 2018). All species and their detection frequencies are listed
351 in Table S6.
352
353 **3.3 Hypothesis generation: biotic factors potentially influencing *T. cristatus* detection**
354
355 *T. cristatus* detection was positively associated with amphibian and waterfowl species
356 richness, yet negatively associated with mammal species richness (Table 1). *T. cristatus*
357 detection was less probable as fish and terrestrial bird species richness increased, but these
358 trends were not significant (Table 1, Fig. 2; GLMM: overdispersion $\theta = 0.995$, $\chi^2_{525} = 522.276$,
359 $P = 0.525$; fit $\chi^2_8 = 1.568$, $P = 0.992$, $R^2 = 12.74\%$). With no correction for Type I error, *T.*
360 *cristatus* had significant positive associations with three species (Table 2, Fig. S5), including

361 *L. vulgaris*, common coot (*Fulica atra*) and *G. chloropus*, and significant negative
362 associations with six species (Table 2, Fig. S5), including *B. bufo*, *C. carpio*, *G. aculeatus*, *P.*
363 *pungitius*, common pheasant (*Phasianus colchicus*) and *S. carolinensis*. However, only *T.*
364 *cristatus* associations with *L. vulgaris* and *G. chloropus* were retained after applying the
365 Benjamini-Hochberg correction (Table 2).

366 *T. cristatus* detection was more likely in ponds with more amphibian species (Table
367 1, Fig. 2a). *T. cristatus* was detected in 51.66% of ponds (*n* = 151) identified as containing *L.*
368 *vulgaris*, but in only 11.91% of ponds (*n* = 42) where *B. bufo* was detected (Fig. 1a).
369 Probability of *T. cristatus* detection was lower in ponds with more fish species (Table 1), and
370 *T. cristatus* was not detected in ponds with more than four fish species (Fig. 2b). *T. cristatus*
371 was only detected in 15.00% (*n* = 40), 14.55% (*n* = 55) and 6.67% (*n* = 15) of ponds identified
372 as containing *C. carpio*, *G. aculeatus* and *P. pungitius* respectively (Fig. 1b). In contrast, *T.*
373 *cristatus* detection was more likely in ponds where more waterfowl species were detected
374 (Table 1, Fig. 2c). *T. cristatus* was detected in 41.67% (*n* = 48) and 36.02% (*n* = 211) of ponds
375 identified as containing *F. atra* and *G. chloropus* respectively (Fig. 1c). *T. cristatus* detection
376 was negatively correlated with higher terrestrial bird species richness, but not significantly
377 so (Table 1, Fig. 2d). However, *T. cristatus* was only detected in 12.00% (*n* = 25) ponds where
378 *P. colchicus* was detected (Fig. 1c). *T. cristatus* detection was less likely in ponds where more
379 mammal species were detected (Table 1, Fig. 2e). Specifically, *T. cristatus* was only detected
380 in 15.79% (*n* = 57) of ponds where *S. carolinensis* was detected (Fig. 1d).

381

382 **3.4 Hypothesis testing: biotic and abiotic factors known to influence *T. cristatus* detection**

383

384 Biotic variables put forward for model selection were eDNA-based detection/non-detection

385 of *L. vulgaris*, *B. bufo*, *C. carpio*, *G. aculeatus*, *P. pungitius*, *G. chloropus*, and *F. atra* as well
386 as presence/absence of other amphibians, fish, and waterfowl as recorded during HSI
387 assessment by environmental consultants (Table S5). Only detection/non-detection of *L.*
388 *vulgaris*, *B. bufo*, *C. carpio*, *G. aculeatus*, and *G. chloropus* as well as waterfowl presence
389 were retained by model selection as explanatory variables for the biotic GLMM of *T.*
390 *cristatus* detection (Table 2, Figs. 3a-f; GLMM: overdispersion $\theta = 1.013$, $\chi^2_{495} = 495.000$, $P =$
391 0.411; fit $\chi^2_8 = 4.437$, $P = 0.8157$, $R^2 = 27.77\%$). Abiotic variables put forward for model
392 selection were max. depth, pond density, presence of inflow, pond area, pond substrate,
393 presence of outflow, percentage of macrophyte cover, water quality, pond permanence,
394 percentage of shading, ruderals, scrub/hedge, and woodland (Table S5). Only presence of
395 inflow, pond area, and percentage of shading were retained by model selection as
396 explanatory variables for the abiotic GLMM explaining *T. cristatus* detection (Table 2, Figs.
397 3g-i; GLMM: overdispersion $\theta = 1.000$, $\chi^2_{499} = 499.126$, $P = 0.490$; fit $\chi^2_8 = 4.898$, $P = 0.768$, R^2
398 $= 10.00\%$). Results are summarised and relationships compared to those previously reported
399 for *T. cristatus* in Table 2. Application of the Benjamini-Hochberg correction for Type I error
400 did not alter variable significance.
401
402 **3.5 Hypothesis testing: *T. cristatus* HSI and umbrella status**
403
404 HSI score positively correlated with *T. cristatus* detection probability (GLMM: overdispersion
405 $\theta = 0.995$, $\chi^2_{501} = 498.372$, $P = 0.525$; fit $\chi^2_8 = 5.395$, $P = 0.715$, $R^2 = 6.66\%$), where *T. cristatus*
406 detection was more likely in ponds with a higher HSI score (Fig. 3j; estimate ± standard error
407 $= 3.094 \pm 0.803$, $\chi^2_1 = 16.020$, $P < 0.001$). Vertebrate species richness was positively
408 associated with *T. cristatus* detection (GLMM: overdispersion $\theta = 1.112$, $\chi^2_{500} = 555.800$, $P =$

409 0.042; fit χ^2_8 = -11.91, P = 1.000, R^2 = 9.14%), with more species detected in ponds identified
410 as containing *T. cristatus* (Fig. 4a; estimate ± standard error = 0.408 ± 0.059, χ^2_1 = 46.265, P
411 < 0.001, adjusted P [Benjamini-Hochberg] < 0.001). However, vertebrate species richness
412 did not significantly increase with the *T. cristatus* HSI score (Fig. 4b; estimate ± standard
413 error = 0.228 ± 0.206, χ^2_1 = 1.227, P = 0.268, adjusted P [Benjamini-Hochberg] = 0.268).
414
415
416 **4. Discussion**
417
418 We have validated eDNA metabarcoding for ecological hypothesis generation and testing
419 using the community data produced by this tool in combination with environmental
420 metadata for ponds. We identified and tested biotic and abiotic factors influencing *T.*
421 *cristatus* detection, whether the HSI can be applied to *T. cristatus* eDNA survey, and
422 whether *T. cristatus* is truly an umbrella species for pond conservation. *T. cristatus*
423 detection was more likely in ponds where *L. vulgaris* and *G. chloropus* were detected, and
424 ponds where *B. bufo*, *C. carpio*, and *G. aculeatus* were not detected. *T. cristatus* detection
425 was also more likely in ponds where waterfowl presence was recorded during HSI
426 assessment. Ponds identified as containing *T. cristatus* were typically small, absent of inflow,
427 and not excessively shaded. The *T. cristatus* HSI correlated with *T. cristatus* detection, but
428 not vertebrate species richness. Nonetheless, more vertebrates were detected in ponds
429 where *T. cristatus* was detected thus presence of this amphibian may indicate good quality
430 habitat for other vertebrates. Our findings demonstrate the power of eDNA metabarcoding
431 to enhance freshwater monitoring and research by providing biodiversity data *en masse* at
432 low cost.

433

434 **4.1 Pondscape biodiversity**

435

436 eDNA metabarcoding detected six amphibian, 14 fish, 17 bird, and 16 mammal species
437 across 532 UK ponds. This diverse species inventory emphasises the importance of ponds as
438 habitat for aquatic taxa, but also as stepping stones for semi-aquatic and terrestrial taxa (De
439 Meester et al., 2005; Hill et al., 2018) through provision of drinking, foraging, dispersive, and
440 reproductive opportunities (Biggs et al., 2016; Klymus et al., 2017). Some species detections
441 may be the result of eDNA transport from water bodies in the surrounding area (Hänfling et
442 al., 2016) to ponds via inflow. However, this signifies the capacity of ponds to provide
443 natural samples of freshwater and terrestrial biodiversity in the wider catchment (Deiner et
444 al., 2017).

445

446 **4.2 Biotic factors influencing *T. cristatus* detection**

447

448 *T. cristatus* detection was more likely in ponds with higher amphibian species richness -
449 particularly ponds identified as containing *L. vulgaris* but not *B. bufo*. *T. cristatus* and *L.*
450 *vulgaris* have similar habitat requirements and tend to breed in the same ponds (Skei et al.,
451 2006; Rannap et al., 2009a; Denoël et al., 2013; Cayuela et al., 2018), with >60% overlap
452 reported (Rannap & Briggs, 2006). However, *L. vulgaris* can inhabit a broader range of
453 habitat (Rannap & Briggs, 2006; Skei et al., 2006) than *T. cristatus*, which depends on larger,
454 deeper ponds with abundant macrophytes and no fish located in open, semi-rural
455 landscapes (Denoël et al., 2013). *B. bufo* can inhabit fish-containing ponds (Manenti &
456 Pennati, 2016) and *T. cristatus* may predate *B. bufo* eggs and larvae (Langton et al., 2001).

457 This may explain the negative association between *B. bufo* and *T. cristatus* as opposed to
458 the positively associated *T. cristatus* and *L. vulgaris*.

459 *T. cristatus* detection marginally decreased with higher fish species richness, and *T.*
460 *cristatus* was negatively associated with *C. carpio*, *G. aculeatus*, and *P. pungitius*. These
461 fishes are common in and typical of ponds. All *T. cristatus* life stages may be predated by
462 fishes (Langton et al., 2001) and negative effects of fish presence-absence on *T. cristatus*
463 occupancy, distribution, and abundance are repeatedly reported (Joly et al., 2001; Rannap &
464 Briggs, 2006; Skei et al., 2006; Denoël & Ficetola, 2008; Rannap et al., 2009a, b; Hartel et al.,
465 2010; Denoël et al., 2013). *G. aculeatus* predates *T. cristatus* eggs and larvae (McLee &
466 Scaife, 1992; Jarvis, 2010), and has non-consumptive effects on *T. cristatus* embryos (Jarvis,
467 2010). *T. cristatus* larvae were also found to alter their behaviour when exposed to
468 predatory *G. aculeatus* but not non-predatory *C. carassius* (Jarvis, 2012), another fish
469 characteristic of ponds.

470 In our study, we detected *T. cristatus* in 50% of ponds where *C. carassius* was
471 detected, but <20% of ponds where large and/or predatory fishes were detected, e.g. *C.*
472 *carpio*, *G. aculeatus*, *E. lucius*. Although we observed fewer detections from ponds for *C.*
473 *carassius* than *C. carpio*, *G. aculeatus* or *E. lucius*, previous research also indicates large
474 and/or predatory fish are more detrimental to *T. cristatus* occurrence (Skei et al., 2006;
475 Hartel et al., 2010; Chan, 2011). *C. carassius* does not hinder *T. cristatus* oviposition, larval
476 behaviour, or recruitment success (Chan, 2011; Jarvis, 2012), or pond invertebrate and
477 macrophyte diversity (Stefanoudis et al., 2017). In contrast, *C. carpio* foraging reduces
478 invertebrate density and macrophyte cover (Maceda-Veiga, López & Green, 2017), which
479 lowers *T. cristatus* reproductive and foraging success and heightens predator exposure
480 (Rannap & Briggs, 2006; Gustafson et al., 2006; Chan, 2011). *C. carassius* and *C. carpio* are

481 both included among fish species assumed to negatively impact *T. cristatus* and whose
482 presence-absence is assessed for the *T. cristatus* HSI (ARG-UK, 2010). However, given that *C.*
483 *carassius* does not directly predate *T. cristatus* or indirectly alter its behaviour, reproductive
484 success, or habitat, we advocate a systematic re-evaluation of problematic fish species for *T.*
485 *cristatus* conservation.

486 *T. cristatus* detection was positively associated with waterfowl species richness,
487 namely *F. atra* and *G. chloropus* detection. These waterfowl species share macrophytes and
488 macroinvertebrates as resources with amphibians, feeding on both directly (Perrow et al.,
489 1997; Paillisson & Marion, 2001; Wallau et al., 2010). *F. atra* and *G. chloropus* crop
490 emergent macrophytes to search for invertebrate prey (Paillisson & Marion, 2001; Wallau et
491 al., 2010), which may indirectly benefit *T. cristatus* foraging. Although *Fulica* spp. can also
492 pull up submerged vegetation and damage vegetation banks (Lauridsen, Jeppesen &
493 Andersen, 1993), diet is macrophyte-dominated in late summer and autumn (Perrow et al.,
494 1997) and unlikely to impact *T. cristatus* breeding in spring (Langton et al., 2001). The
495 positive association identified here between *T. cristatus* and these waterfowl most likely
496 reflects a shared preference for macrophyte-rich ponds.

497 *T. cristatus* detection was less frequent in ponds with higher mammal species
498 richness. Our preliminary cooccur analysis indicated *T. cristatus* had negative associations
499 with *P. colchicus* and *S. carolinensis,* but not when a correction for Type I error was applied.
500 *T. cristatus* associations with *F. atra*, *B. bufo*, *C. carpio*, *G. aculeatus*, and *P. pungitius* were
501 also non-significant when this correction was applied. These associations require further
502 investigation to determine whether they are real, and if so, whether they are direct and
503 reflect competition, predation or shared resources as discussed above, or indirect and
504 reflect land-use or effects of additional species.

505

506 **4.3 Abiotic factors influencing *T. cristatus* detection**

507

508 *T. cristatus* detection was less likely in large ponds with inflow present and a greater
509 percentage of shading. Effects of pond area may depend on the size range of ponds studied.
510 Although our results indicate *T. cristatus* prefers smaller ponds, pond area does not always
511 influence occupancy (Maletzky, Kyek & Goldschmid, 2007; Denoël & Ficetola, 2008;
512 Gustafson et al., 2011) and was deemed a poor predictor of reproductive success (Vuorio et
513 al., 2013). *T. cristatus* has been found to utilise small and large ponds (Rannap & Briggs,
514 2006; Skei et al., 2006); however, very small ponds (<124 m^2) may be unable to support all
515 life stages, and larger ponds may contain fish and experience eutrophication due to
516 agricultural or polluted run-off (Rannap & Briggs, 2006). Inflow to ponds may exacerbate
517 these problems by facilitating entry of agricultural or polluted run-off and connections to
518 streams and rivers containing large, predatory fish (Freshwater Habitats Trust, 2015). Our
519 results corroborate existing research where viable *T. cristatus* populations were unlikely in
520 ponds that were shaded (Vuorio et al., 2013) or had dense macrophyte cover (Rannap &
521 Briggs, 2006; Skei et al., 2006; Hartel et al., 2010).

522 In our study, most environmental metadata available were qualitative, preventing
523 detailed analyses on pond properties and terrestrial habitat in relation to *T. cristatus*
524 detection. Better understanding of *T. cristatus* detection in relation to species interactions
525 and habitat quality could be achieved with quantitative data on pond properties (e.g. water
526 chemistry), terrestrial habitat (e.g. type, density, distance to ponds), and aquatic and
527 terrestrial habitat usage by different vertebrate species. Furthermore, given the
528 metapopulation dynamics of *T. cristatus*, future research should investigate spatial drivers

529 (e.g. pond density and other indices of connectivity, land cover, climate variables, roads,
530 rivers, elevation) of *T. cristatus* detection using innovative modelling approaches, such as
531 individual-based models (Messager & Olden, 2018). However, acquiring this data to perform
532 these models is a phenomenal task for large numbers of ponds across a vast landscape
533 (Denoël & Ficetola, 2008).
534
535 **4.4 *T. cristatus* HSI and umbrella status**
536
537 We found the HSI can predict eDNA-based *T. cristatus* detection at the UK pondscape. This
538 contradicts conventional studies which deemed the index inappropriate for predicting *T.*
539 *cristatus* occupancy or survival probabilities (Unglaub et al., 2015). We detected more
540 vertebrates in ponds identified as containing *T. cristatus*, which may support its status as an
541 umbrella species for pond biodiversity and conservation (Gustafson et al., 2006). We also
542 observed a non-significant increase in vertebrate species richness with increasing *T.*
543 *cristatus* HSI score. An adapted HSI, designed to predict species richness, could help select
544 areas for management and enhancement of aquatic and terrestrial biodiversity. Until then,
545 presence of *T. cristatus* and its HSI may confer protection to broader biodiversity by
546 identifying optimal habitat for pond creation and restoration to encourage populations of
547 this threatened amphibian. The HSI is not without issue due to qualitative data used for
548 score calculation and subjective estimation of indices (Oldham et al., 2000; O'Brien et al.,
549 2017). For future application of this index in *T. cristatus* eDNA survey, we recommend
550 metabarcoding to quantify some qualitatively assessed indices (e.g. water quality via
551 macroinvertebrate diversity, fish and waterfowl presence) alongside *T. cristatus* detection.
552 Provided rigorous spatial and temporal sampling are undertaken, eDNA metabarcoding can

553 also generate site occupancy data to estimate relative species abundance (Valentini et al.,
554 2016; Hänfling et al., 2016; Lawson Handley et al., 2019; Li et al., 2019a).
555
556 **4.5 Limitations of repurposed eDNA samples for metabarcoding applications**
557
558 This study was based on samples that were repurposed from targeted eDNA surveys for *T.*
559 *cristatus*. eDNA sampling, capture, and extraction was conducted in accordance with Biggs
560 et al. (2015), whose methods were chosen based on the eDNA literature at that time. Six 15
561 ml water samples were taken from a homogenised sample (600 ml). Ethanol precipitation
562 was used for eDNA capture, followed by DNA extraction on the combined lysate from the six
563 subsamples. Consequently, ponds were represented by a single eDNA sample, for which 12
564 qPCR replicates were performed (Biggs et al., 2015). However, eDNA metabarcoding may
565 require an entirely different workflow to enhance species detection (Harper et al., 2018;
566 Harper et al., 2019).
567 Independent biological replicates as opposed to pseudoreplicates from a single
568 water sample are key to account for spatial heterogeneity of eDNA from different species
569 (Hänfling et al., 2016; Bálint et al., 2018; Lawson Handley et al., 2019; Li et al., 2019b).
570 Larger volumes should be filtered instead of ethanol precipitation on small volumes to
571 maximise eDNA capture (Harper et al., 2019). Samples should be extracted individually to
572 maintain biological replication followed by independent technical replicates for PCR and
573 sequencing. These levels of replication are required for hierarchical occupancy modelling to
574 identify false positives and species detection probabilities as well as minimise false
575 negatives (Ficetola et al., 2015; Bálint et al., 2018; Dorazio & Erickson, 2018).
576 We performed three PCR replicates for each sample that were pooled prior to

577 sequencing. We included a large number of PCR controls and applied sequence thresholds
578 to control for any false positives in our dataset arising from potential laboratory
579 contamination. Pooling of technical replicates allowed us to screen vertebrate communities
580 from over 500 ponds, but prevented occupancy modelling to verify species detection.
581 Samples collected for targeted eDNA analysis can vary widely in terms of sampling strategy,
582 capture, extraction and storage protocols. Ultimately, these protocols may limit the data
583 generated by eDNA metabarcoding and the analyses that can be performed. This is an
584 important consideration for any future research where eDNA samples are repurposed to
585 address different ecological questions.
586
587 **4.6 Prospects of eDNA metabarcoding for freshwater conservation, management, and**
588 **research**
589
590 We have demonstrated the effectiveness of eDNA metabarcoding for landscape-scale
591 biodiversity monitoring as well as ecological hypothesis generation and testing. We
592 combined metabarcoding with environmental metadata to identify new and revisit old
593 hypotheses relating to biotic and abiotic factors that influence a threatened amphibian at
594 the UK pondscape. Our findings will guide *T. cristatus* conservation in the face of increasing
595 land-use and habitat fragmentation - a poignant issue as protective legislation for this
596 species in the UK is changing. Whilst conservation of threatened species and their habitat
597 should be a priority, the bigger picture should not be ignored. eDNA metabarcoding could
598 enhance our understanding of freshwater networks, particularly pondscapes, to enable
599 more effective monitoring, protection, and management of aquatic and terrestrial
600 biodiversity. We are only now beginning to realise and explore these opportunities.

601

602

603 ## Acknowledgements

604

605 This work was funded by the University of Hull. We would like to thank Jennifer Hodgetts
606 (Fera Science Ltd) for assisting with sample collection, and Jianlong Li (University of Hull) for
607 sequencing primer design and advice on laboratory protocols. Tissue samples for primer
608 validation and Sanger sequencing were provided by Andrew Buxton and Richard Griffiths
609 (DICE, University of Kent) under licence from Natural England, and Barbara Mabel and
610 Elizabeth Kilbride (University of Glasgow).

611

612

613 ## Data Availability

614

615 The taxonomically assigned sequence reads used in this study were generated by Harper et
616 al. (2018). The raw sequence reads were archived on the NCBI Sequence Read Archive
617 (Bioproject: PRJNA417951; SRA accessions: SRR6285413 - SRR6285678). The bioinformatics
618 analysis was deposited in a GitHub repository and permanently archived
619 (https://doi.org/10.5281/zenodo.1188710). R scripts and corresponding data for this study
620 have been deposited in a separate GitHub repository, which has been permanently archived
621 (https://doi.org/10.5281/zenodo.3365703).

622

623

Author Contributions

B.H, L.R.H, L.L.H and N.B conceived and designed the study. H.C.R and N.B contributed samples for processing. L.R.H performed laboratory work and analysed the data. I.P.A and E.L offered advice on and supervised sequencing. C.H assisted with bioinformatics analysis. P.B and S.P contributed datasets for analysis. L.R.H wrote the manuscript, which all authors revised.

References

Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, **60**, 255–265. https://doi.org/10.2307/2334537

ARG-UK. (2010). *ARG UK Advice Note 5: Great Crested Newt Habitat Suitability Index*. Amphibian and Reptile Groups of the United Kingdom.

Bálint, M., Nowak, C., Márton, O., Pauls, S. U., Wittwer, C., Aramayo, J. L., Schulze, A., Chambert, T., Cocchiararo, B., & Jansen, M. (2018). Accuracy, limitations and cost efficiency of eDNA-based community survey in tropical frogs. *Molecular Ecology Resources*, **18**, 1415–1426. https://doi.org/10.1111/1755-0998.12934

Barnes, M. A., & Turner, C. R. (2015). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, **17**, 1–17. https://doi.org/10.1007/s10592-015-775-4

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, **67**, 1–48. https://doi.org/10.18637/jss.v067.i01

Biggs, J., Ewald, N., Valentini, A., Gaboriaud, C., Dejean, T., Griffiths, R. A., Foster, J., Wilkinson, J. W., Arnell, A., Brotherton, P., Williams, P., & Dunn, F. (2015). Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt (*Triturus cristatus*). *Biological Conservation*, **183**, 19–28. https://doi.org/10.1016/j.biocon.2014.11.029

Biggs, J., von Fumetti, S., & Kelly-Quinn, M. (2016). The importance of small waterbodies for biodiversity and ecosystem services: implications for policy makers. *Hydrobiologia*, **793**, 3–39. https://doi.org/10.1007/s10750-016-3007-0

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

Cayuela, H., Grolet, O., & Joly, P. (2018). Context-dependent dispersal, public information, and heterospecific attraction in newts. *Oecologia*, **188**, 1069–1080. https://doi.org/10.1007/s00442-018-4267-3

Chan, K. (2011). Can great crested newts (*Triturus cristatus*) coexist with fish?. Masters Thesis. University College London (UCL).

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E., & Bernatchez, L. (2017). Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Molecular Ecology*, **26**, 5872–5895. https://doi.org/10.1111/mec.14350

De Meester, L., Declerck, S., Stoks, R., Louette, G., Van De Meutter, F., De Bie, T., Michels, E., & Brendonck, L. (2005). Ponds and pools as model systems in conservation biology, ecology and evolutionary biology. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **15**, 715–725. https://doi.org/10.1002/aqc.748

Denoël, M., & Ficetola, G. F. (2008). Conservation of newt guilds in an agricultural landscape of Belgium: the importance of aquatic and terrestrial habitats. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **18**, 714–728. https://doi.org/10.1002/aqc.853

Denoël, M., Perez, A., Cornet, Y., & Ficetola, G. F. (2013). Similar local and landscape processes affect both a common and a rare newt species. *PloS ONE*, **8**, e62727. https://doi.org/10.1371/journal.pone.0062727

Dorazio, R.M. & Erickson, R.A. (2018) ednaoccupancy: An r package for multiscale occupancy modelling of environmental DNA data. *Molecular Ecology Resources*, **18**, 368–380. https://doi.org/10.1111/1755-0998.12735

Edgar, P., & Bird, D. R. (2006). *Action Plan for the Conservation of the Crested Newt Triturus Cristatus Species Complex in Europe*. Council of the European Union, Strassbourg, Germany.

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., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., Taberlet, P., & Pompanon, F. (2010). An *In silico* approach for the evaluation of DNA barcodes. *BMC Genomics*, **11**, 434. https://doi.org/10.1186/1471-2164-11-434

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

Freshwater Habitats Trust (2015). Pond Habitat Survey Method Booklet. Freshwater Habitats Trust, Oxford.

Griffith, D., Veech, J., & Marsh, C. (2016). cooccur: Probabilistic Species Co-Occurrence Analysis in R. *Journal of Statistical Software*, **69**, 1–17. https://doi.org/10.18637/jss.v069.c02

Gustafson, D. H., Pettersson, C. J., & Malmgren, J. C. (2006). Great crested newts (*Triturus cristatus*) as indicators of aquatic plant diversity. *The Herpetological Journal*, **16**, 347–352.

Gustafson, D. H., Malmgren, J. C., & Mikusiński, G. (2011). Terrestrial Habitat Predicts use of Aquatic Habitat for Breeding Purposes — A Study on the Great Crested Newt (*Triturus cristatus*). *Annales Zoologici Fennici*, **48**, 295–307. https://doi.org/10.5735/086.048.0504

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*, 6330–6341. https://doi.org/10.1002/ece3.4013

Harper, L. R., Buxton, A. S., Rees, H. C., Bruce, K., Brys, R., Halfmaerten, D., Read, D. S., Watson, H. V., Sayer, C. D., Jones, E. P., Priestley, V., Mächler, E., Múrria, C., Garcés-Pastor, S., Medupin, C., Burgess, K., Benson, G., Boonham, N., Griffiths, R. A., Lawson Handley, L., & Hänfling, B. (2019). Prospects and challenges of environmental DNA (eDNA) monitoring in freshwater ponds. *Hydrobiologia*, **826**, 25–41. https://doi.org/10.1007/s10750-018-3750-5

Hartel, T., Nemes, S., & Oellerer, K. (2010). Using connectivity metrics and niche modelling to explore the occurrence of the northern crested newt *Triturus cristatus* (Amphibia, Caudata) in a traditionally managed landscape. *The Environmentalist*, **37**, 195–200. https://doi.org/10.1017/S037689291000055X

Hering, D., Borja, A., Jones, J. I., Pont, D., Boets, P., Bouchez, A., Bruce, K., Drakare, S., Hänfling, B., Kahlert, M., Leese, F., Meissner, K., Mergen, P., Reyjol, Y., Segurado, P., Vogler, A., & Kelly, M. (2018). Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. *Water Research*, **138**, 192–205. https://doi.org/10.1016/j.watres.2018.03.003

Hill, M. J., Hassall, C., Oertli, B., Fahrig, L., Robson, B. J., Biggs, J., Samways, M. J., Usio, N., Takamura, N., Krishnaswamy, J., & Wood, P. J. (2018). New policy directions for global pond conservation. *Conservation Letters*, **e12447**. https://doi.org/10.1111/conl.12447

Jarvis, L. E. (2010). Non-consumptive effects of predatory three-spined sticklebacks (*Gasterosteus aculeatus*) on great crested newt (*Triturus cristatus*) embryos. *The Herpetological Journal*, **20**, 271–275.

Jarvis, L. E. (2012). Microhabitat preferences of the great crested newt *Triturus cristatus* in a woodland area. PhD Thesis. Open University.

Joly, P., Miaud, C., Lehmann, A., & Grolet, O. (2001). Habitat Matrix Effects on Pond Occupancy in Newts. *Conservation Biology*, **15**, 239–248. https://doi.org/10.1111/j.1523-1739.2001.99200.x

Bartoń, K. (2018). MuMIn: Multi-Model Inference. R package version 1.42.1. https://CRAN.R-project.org/package=MuMIn

Kelly, R. P., O'Donnell, J. L., Lowell, N. C., Shelton, A. O., Samhouri, J. F., Hennessey, S. M., Feist, B. E., & Williams, G. D. (2016). Genetic signatures of ecological diversity along an urbanization gradient. *PeerJ*, **4**, e2444. https://doi.org/10.7717/peerj.2444

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**, 1-18. https://doi.org/10.3390/d9040054

Langton, T., Beckett, C., & Foster, J. (2001). Great crested newt conservation handbook. Froglife, Halesworth.

Lauridsen, T. L., Jeppesen, E., & Andersen, F. Ø. (1993). Colonization of submerged macrophytes in shallow fish manipulated Lake Væng: impact of sediment composition and waterfowl grazing. *Aquatic Botany*, **46**, 1–15. https://doi.org/10.1016/0304-3770(93)90061-Z

Lawson Handley, L. (2015). How will the "molecular revolution" contribute to biological recording? *Biological Journal of the Linnean Society,* 115, 750–766. https://doi.org/10.1111/bij.12516

Lawson Handley, L., Read, D.S., Winfield, I.J., Kimbell, H., Johnson, H., Li, J., Hahn, C., Blackman, R., Wilcox, R., Donnelly, R., Szitenberg, A. & Hänfling, B. (2019) Temporal and spatial variation in distribution of fish environmental DNA in England's largest lake. *Environmental DNA*, 12, e0176343. https://doi.org/10.1002/edn3.5

Lele, S. R., Keim, J. L., & Solymos, P. (2016). ResourceSelection: Resource Selection (Probability) Functions for Use-Availability Data.

Li, Y., Evans, N. T., Renshaw, M. A., Jerde, C. L., Olds, B. P., Shogren, A. J., Deiner, K., Lodge, D. M., Lamberti, G. A., & Pfrender, M. E. (2018a). Estimating fish alpha- and beta-diversity along a small stream with environmental DNA metabarcoding. *Metabarcoding and Metagenomics*, 2, e24262. https://doi.org/10.3897/mbmg.2.24262

Li, F., Peng, Y., Fang, W., Altermatt, F., Xie, Y., Yang, J., & Zhang, X. (2018b). Application of environmental DNA metabarcoding for predicting anthropogenic pollution in rivers. *Environmental Science & Technology*, 52, 11708–11719. https://doi.org/10.1021/acs.est.8b03869

Li, J., Hatton-Ellis, T.W., Lawson Handley, L.-J., Kimbell, H.S., Benucci, M., Peirson, G. & Hänfling, B. (2019a) Ground-truthing of a fish-based environmental DNA metabarcoding method for assessing the quality of lakes. *Journal of Applied Ecology*, **56**, 1232–1244. https://doi.org/10.1111/1365-2664.13352

Li, J., Lawson Handley, L.J., Harper, L.R., Brys, R., Watson, H.V., Di Muri, C., Zhang, X. & Hänfling, B. (2019b) Limited dispersion and quick degradation of environmental DNA in fish ponds inferred by metabarcoding. *Environmental DNA*, **5**, 133. https://doi.org/10.1002/edn3.24

Maceda-Veiga, A., López, R., & Green, A. J. (2017). Dramatic impact of alien carp *Cyprinus carpio* on globally threatened diving ducks and other waterbirds in Mediterranean shallow lakes. *Biological Conservation*, **212**, 74–85. https://doi.org/10.1016/j.biocon.2017.06.002

Macher, J. -N., Vivancos, A., Piggott, J. J., Centeno, F. C., Matthaei, C. D., & Leese, F. (2018). Comparison of environmental DNA and bulk-sample metabarcoding using highly degenerate cytochrome c oxidase I primers. *Molecular Ecology Resources*, **18**, 1456–1468. https://doi.org/10.1111/1755-0998.12940

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

Maletzky, A., Kyek, M., & Goldschmid, A. (2007). Monitoring status, habitat features and amphibian species richness of crested newt (*Triturus cristatus* superspecies) ponds at the edge of the species range (Salzburg, Austria). *Annales de Limnologie - International Journal of Limnology*, **43**, 107–115. https://doi.org/10.1051/limn/2007017

Manenti, R., & Pennati, R. (2016). Environmental factors associated with amphibian breeding in streams and springs: effects of habitat and fish occurrence. *Amphibia-Reptilia*, **37**, 237–242. https://doi.org/10.1163/15685381-00003040

Mathews, F., Kubasiewicz, L. M., Gurnell, J., Harrower, C. A., McDonald, R. A., & Shore, R. F. (2018). A Review of the Population and Conservation Status of British Mammals: Technical Summary. A report by the Mammal Society under contract to Natural England, Natural Resources Wales and Scottish Natural Heritage. Natural England, Peterborough.

Mazerolle, M. J. (2016). AICcmodavg: Model Selection and Multimodel Inference Based on (Q)AIC(c).

McLee, A. G., & Scaife, R. W. (1992). The colonisation by great crested newts (*Triturus cristatus*) of a water body following treatment with a piscicide to remove a large population of sticklebacks (*Gasterosteus aculeatus*). *Herpetological Bulletin*, **42**, 6–9.

Messager, M. L., & Olden, J. D. (2018). Individual-based models forecast the spread and inform the management of an emerging riverine invader. *Diversity & Distributions*, **24**, 1816–1829. https://doi.org/10.1111/ddi.12829

Nakagawa, H., Yamamoto, S., Sato, Y., Sado, T., Minamoto, T., & Miya, M. (2018). Comparing local- and regional-scale estimations of the diversity of stream fish using eDNA metabarcoding and conventional observation methods. *Freshwater Biology*, **63**, 569–580. https://doi.org/10.1111/fwb.13094

O'Brien, D., Hall, J., Miró, A. & Wilkinson, J. (2017). Testing the validity of a commonly-used habitat suitability index at the edge of a species' range: great crested newt *Triturus cristatus* in Scotland. *Amphibia-Reptilia*, **38**, 265-273. https://doi.org/10.1163/15685381-00003108

Oldham, R. S., Keeble, J., Swan, M. J. S., & Jeffcote, M. (2000). Evaluating the suitability of habitat for the great crested newt (*Triturus cristatus*). *The Herpetological Journal*, **10**, 143–155.

Olds, B. P., Jerde, C. L., Renshaw, M. A., Li, Y., Evans, N. T., Turner, C. R., Deiner, K., Mahon, A. R., Brueseke, M. A., Shirey, P. D., Pfrender, M. E., Lodge, D. M., & Lamberti, G. A. (2016). Estimating species richness using environmental DNA. *Ecology and Evolution*, **6**, 4214–4226. https://doi.org/10.1002/ece3.2186

Paillisson, J. -M., & Marion, L. (2001). Interaction between coot (*Fulica atra*) and waterlily (*Nymphaea alba*) in a lake: the indirect impact of foraging. *Aquatic Botany*, **71**, 209–216. https://doi.org/10.1016/S0304-3770(01)00182-6

Perrow, M. R., Schutten, H. J., Howes, J. R., Holzer, T., Madgwick, F. J., & Jowitt, A. (1997). Interactions between coot (*Fulica atra*) and submerged macrophytes: the role of birds in the restoration process. *Hydrobiologia*, **342**, 241–255. https://doi.org/10.1023/A:1017007911190

Rannap, R., & Briggs, L. (2006). THE CHARACTERISTICS OF GREAT CRESTED NEWT TRITURUS CRISTATUS'BREEDING PONDS. Protection of the Great Crested Newt. Project Report. The experiences of LIFE-Nature project "Protection of *Triturus cristatus* in the Eastern Baltic Region" LIFE04NAT/EE/000070. Ministry of the Environment of the Republic of Estonia, Tallinn.

Rannap, R., Lõhmus, A., & Briggs, L. (2009a). Niche position, but not niche breadth, differs in two coexisting amphibians having contrasting trends in Europe. *Diversity & Distributions*, **15**, 692–700. https://doi.org/10.1111/j.1472-4642.2009.00575.x

Rannap, R., Lõhmus, A., & Briggs, L. (2009b). Restoring ponds for amphibians: a success story. *Hydrobiologia*, **634**, 87–95. https://doi.org/10.1007/s10750-009-9884-8

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

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

Skei, J. K., Dolmen, D., Rønning, L., & Ringsby, T. H. (2006). Habitat use during the aquatic phase of the newts *Triturus vulgaris* (L.) and *T. cristatus* (Laurenti) in central Norway: proposition for a conservation and monitoring area. *Amphibia-Reptilia*, **27**, 309–324. https://doi.org/10.1163/156853806778189972

Stefanoudis, P. V., Sayer, C. D., Greaves, H. M., Davidson, T. A., Robson, H., Almeida, D., & Smith, E. (2017). Consequences of fish for cladoceran, water beetle and macrophyte communities in a farmland pond landscape: implications for conservation. *Fundamental and Applied Limnology*, **190**, 141–156. https://doi.org/10.1127/fal/2017/1004

Therneau, T., Atkinson, B. & Ripley, B. (2014). rpart: Recursive Partitioning and Regression Trees. R package version 4.1-13.

Unglaub, B., Steinfartz, S., Drechsler, A., & Schmidt, B. R. (2015). Linking habitat suitability to demography in a pond-breeding amphibian. *Frontiers in Zoology*, **12**. https://doi.org/10.1186/s12983-015-0103-3

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**, e63-e65. https://doi.org/10.1111/1755-0998.12690

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G. H., Geniez, P., Pont, D., Argillier, C., Baudoin, J. -M., Peroux, T., Crivelli, A. J., Olivier, A., Acqueberge, M., Le Brun, M., Møller, P. R., Willerslev, E. & Dejean, T. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, **25**, 929–942. https://doi.org/10.1111/mec.13428

Vuorio, V., Heikkinen, R. K., & Tikkanen, O. -P. (2013). Breeding Success of the Threatened Great Crested Newt in Boreal Forest Ponds. *Annales Zoologici Fennici*, **50**, 158–169. https://doi.org/10.5735/086.050.0303

Wallau, G. L., Della-Flora, F., Bueno, A. S., Corso, J., Ortiz, M. F., & Cáceres, N. C. (2010). Behaviour of the Common Moorhen in Rio Grande do Sul, Brazil. *Acta Ethologica*, **13**, 127–135. https://doi.org/10.1007/s10211-010-0082-5

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer New York, USA.

Wilcox, T. M., Young, M. K., McKelvey, K. S., Isaak, D. J., Horan, D. L. & Schwartz, M. K. (2018). Fine-scale environmental DNA sampling reveals climate-mediated interactions between native and invasive trout species. *Ecosphere*, **9**, e02500. https://doi.org/10.1002/ecs2.2500

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.

898 **Table 1.** Summary of hypothesis generation using eDNA metabarcoding data, specifically
899 biotic factors that may influence *T. cristatus* detection probability. Direction of observed
900 effects on eDNA-based *T. cristatus* detection identified by the GLMM assessing species
901 richness in each vertebrate group (*N* = 532) and the preliminary cooccur analysis (*N* = 532)
902 are given. Negative and positive effects are listed as - and + respectively. Test statistic is for
903 LRT used and significant P-values (<0.05) are in bold.

Variable	**Analysis**							
	Cooccur			**GLMM**				
	Effect	***P***	**Adjusted *P* (Benjamini-Hochberg)**	**DF**	**Effect size (SE)**	**χ^2**	***P***	**Adjusted *P* (Benjamini-Hochberg)**
Amphibians								
Species richness				1	0.526 (0.140)	14.195	**<0.001**	**<0.001**
L. vulgaris	+	**<0.001**	**<0.001**	1	1.616 (0.232)			
B. bufo	-	**0.010**	0.489	1	-1.793 (0.643)			
Fish								
Species richness				1	-0.163 (0.124)	1.798	0.180	0.225
G. aculeatus	-	**0.012**	0.489					
C. carpio	-	**0.040**	0.714					
P. pungitius	-	**0.049**	0.714					
C. carassius								
Waterfowl								
Species richness				1	0.491 (0.134)	13.530	**<0.001**	**<0.001**
G. chloropus	+	**0.001**	0.012					
F. atra	+	**0.022**	0.237					
Terrestrial birds								
Species richness				1	-0.044 (0.302)	0.021	0.884	0.884

Variable	Analysis							
	Cooccur			GLMM				
	Effect	*P*	Adjusted *P* (Benjamini-Hochberg)	DF	Effect size (SE)	χ^2	*P*	Adjusted *P* (Benjamini-Hochberg)
P. colchicus	-	**0.050**	0.714					
Terrestrial mammals								
Species richness				1	-0.515 (0.188)	8.501	**0.004**	**0.006**
S. carolinensis	-	**0.020**	0.685					

904 **Table 2.** Summary of hypothesis testing using eDNA metabarcoding data and associated
905 environmental metadata, specifically biotic and abiotic factors known to influence *T.*
906 *cristatus* detection probability. Reported effects on *T. cristatus* detection in the literature
907 and hypothesised effects on eDNA-based *T. cristatus* detection are given for each variable.
908 Any variables not explicitly reported in the literature that were identified by the preliminary
909 cooccur analysis are listed as UNK. Direction of observed effects on eDNA-based *T. cristatus*
910 detection identified by each analysis (biotic GLMM, *n* = 504; abiotic GLMM, *n* = 504) are
911 given. Negative and positive effects are listed as - and + respectively. For categorical
912 variables with more than one level, effect size and standard error (SE) are only given for
913 levels reported in the model summary. Test statistic is for LRT used and significant P-values
914 (<0.05) are in bold. Variables included for model selection but not retained in the final
915 models are listed as NR.
916

Variable	**Effect reported**	**Hypothesised effect**	**GLMM**				
			DF	**Effect size (SE)**	**χ^2**	***P***	**Adjusted *P* (Benjamini-Hochberg)**
Amphibians							
L. vulgaris	+	+	1	1.616 (0.232)	50.024	**<0.001**	**<0.001**
B. bufo	UNK		1	-1.793 (0.643)	11.175	**<0.001**	**0.002**
Fish							
G. aculeatus	-	-	1	-1.048 (0.453)	6.300	**0.012**	**0.014**
C. carpio	-	-	1	-1.457 (0.586)	7.854	**0.005**	**0.008**
P. pungitius	-	-	NR				
C. carassius	-	-					
Waterfowl							
G. chloropus	UNK		1	0.893 (0.228)	15.703	**<0.001**	**<0.001**
F. atra	UNK		NR				
Waterfowl presence	-	-	2		6.350	**0.042**	**0.042**
Minor				0.531 (0.234)			
Major				0.808 (0.495)			
Pond area	-/+	-	1	-0.0003 (0.0002)	4.134	**0.042**	**0.042**

Variable	Effect reported	Hypothesised effect	GLMM				
			DF	Effect size (SE)	χ^2	*P*	Adjusted *P* (Benjamini-Hochberg)
Pond density	+	+			NR		
Pond depth	+	+			NR		
Pond substrate	+	+			NR		
Pond permanence	+	+			NR		
Water quality	+	+			NR		
Inflow Present	-	-	1	-0.820 (0.251)	11.437	**<0.001**	**0.001**
Outflow	UNK				NR		
Macrophyte cover	-/+	-			NR		
Shading	-/+	-	1	-0.010 (0.003)	12.266	**<0.001**	**0.001**
Woodland	+	+			NR		
Scrub/hedge	+	+			NR		
Ruderals	UNK				NR		

917

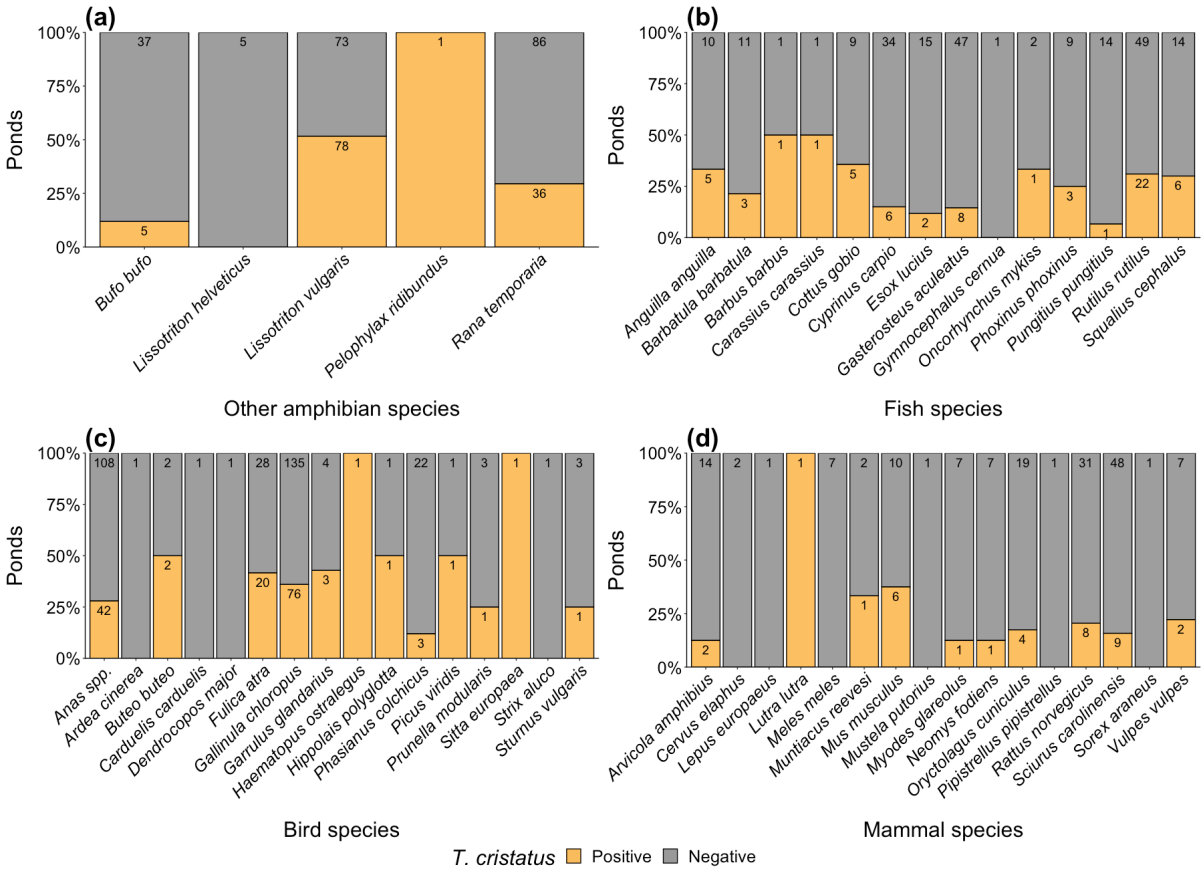

Figure 1. eDNA metabarcoding detection of *T. cristatus* in relation to other vertebrate species (*N* = 532 ponds): **(a)** other amphibians, **(b)** fish, **(c)** birds, and **(d)** mammals. Numbers on each bar are the number of ponds with (orange) and without (grey) *T. cristatus* in which a vertebrate species was detected.

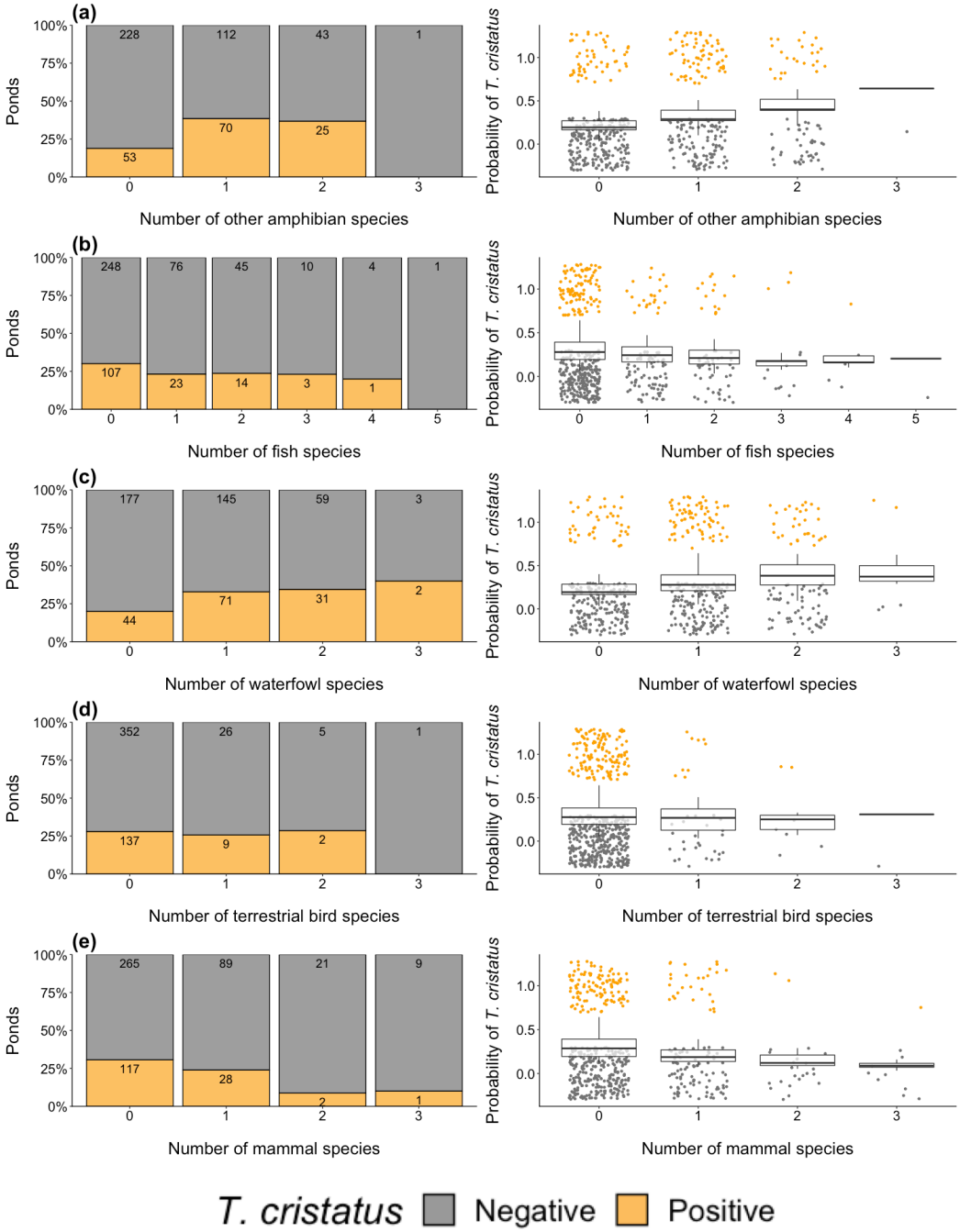

Figure 2. *T. cristatus* detection (orange) and non-detection (grey) in relation to species richness of different vertebrate groups (*N* = 532 ponds): **(a)** other amphibians, **(b)** fish, **(c)** waterfowl, **(d)** terrestrial birds, and **(e)** mammals. Observed proportion of ponds with and without *T. cristatus* (left) is plotted alongside predicted probability of *T. cristatus* detection (right). Numbers on barplots of observed detection are the number of ponds for each category. In plots showing predicted *T. cristatus* detection, the observed data is shown as points (jittered around 0 and 1 to clarify variation in point density) and boxes are the model predictions.

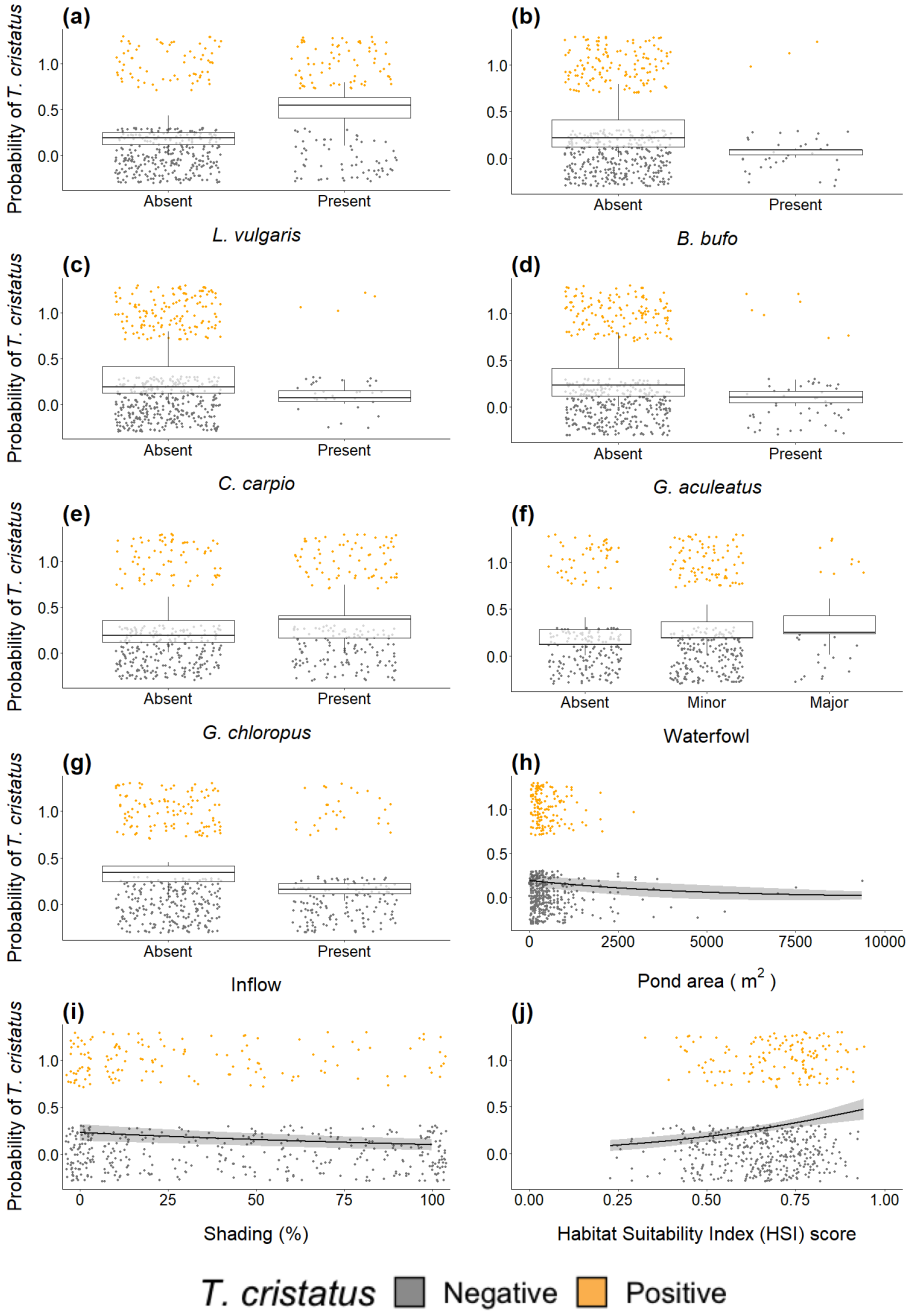

Figure 3. Biotic and abiotic factors influencing *T. cristatus* detection (*n* = 504 ponds): **(a)** *L. vulgaris* detection, **(b)** *B. bufo* detection, **(c)** *C. carpio* detection, **(d)** *G. aculeatus* detection, **(e)** *G. chloropus* detection, **(f)** extent of waterfowl presence, **(g)** presence of inflow, **(h)** pond area, **(i)** percentage of shading, and **(j)** HSI score. The 95% CIs, as calculated using the predicted *T. cristatus* probability values and standard error for these predictions, are given for each relationship. The observed *T. cristatus* detection (orange) and non-detection (grey) data are displayed as points (jittered around 0 and 1 to clarify variation in point density) against the predicted relationships (boxes/lines).

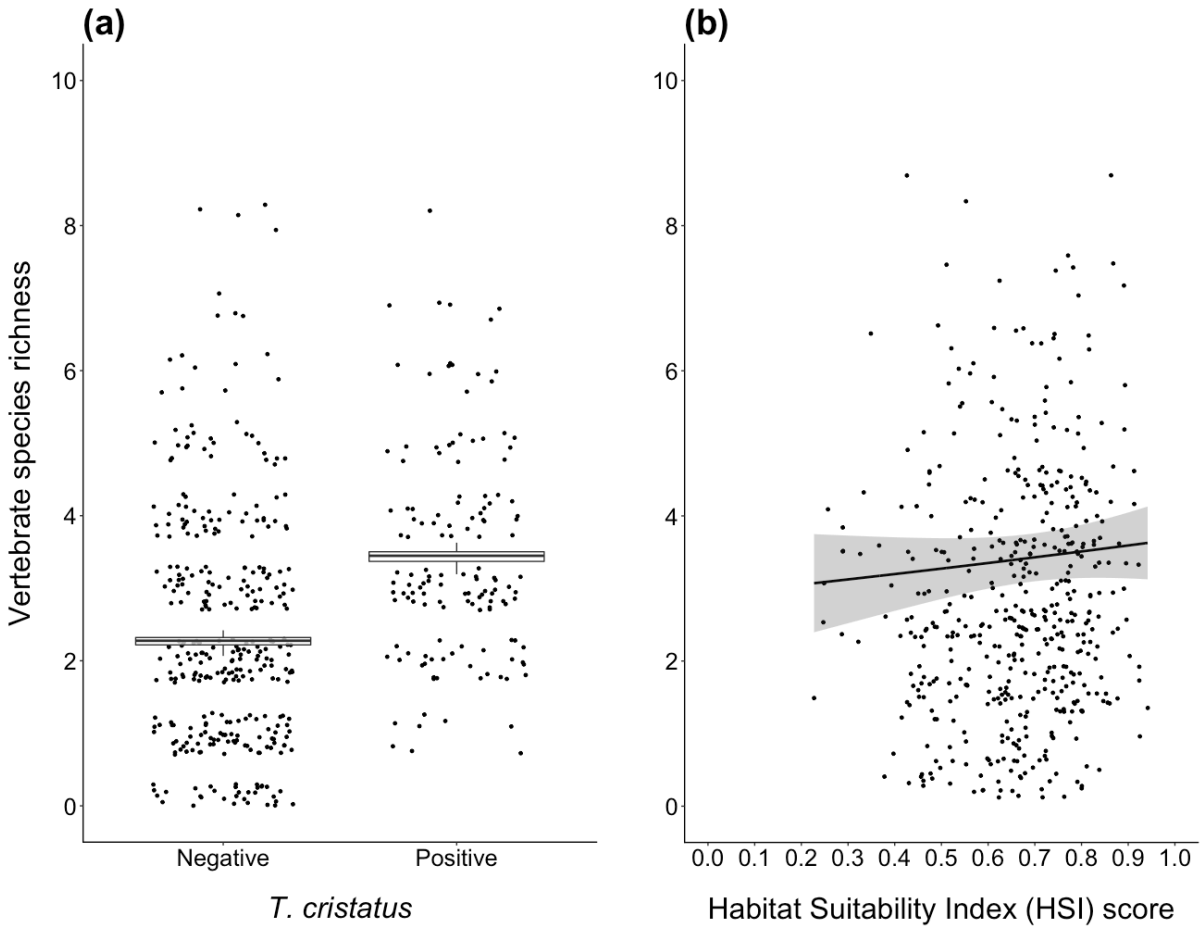

Figure 4. Vertebrate species richness in ponds (*n* = 504) in relation to: **(a)** *T. cristatus* detection and **(b)** the *T. cristatus* HSI score. The 95% CIs, as calculated using the predicted species richness values and standard error for these predictions, are given for each relationship. The observed data are displayed as points (jittered to clarify variation in point density) against the predicted relationships (boxes/lines).