# Testing ecological hypotheses at the pondscape with environmental DNA metabarcoding: a case study on a threatened amphibian

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

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

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

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

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

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

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

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

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

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 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* occupancy 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* occupancy 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* occupancy 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* occupancy was negatively correlated with mammal species richness, and *T. cristatus* was negatively associated with grey squirrel (*Sciurus carolinensis*). *T. cristatus* occupancy was negatively influenced by 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 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*

# 1. Introduction

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

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

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

In the UK, the threatened great crested newt (*Triturus cristatus*) is an umbrella species for pond conservation. The extensive literature on *T. cristatus* ecology provides an excellent opportunity to validate ecological patterns revealed by eDNA metabarcoding. Both biotic (e.g. breeding substrate, prey, and predators) and abiotic (e.g. pond area, depth, and temperature) factors are known to influence *T. cristatus* breeding success (Langton, Beckett & Foster, 2001). The *T. cristatus* Habitat Suitability Index (HSI [Oldham et al., 2000; ARG-UK, 2010]) accounts for these factors using 10 suitability indices that are scored and combined to calculate a decimal score between 0 and 1 (where 1 = excellent habitat). Larvae are susceptible to fish and waterfowl predation (Edgar & Bird, 2006; Rannap & Briggs, 2006; Skei et al., 2006; Hartel, Nemes & Oellerer, 2010), and adults reportedly avoid ponds containing three-spined stickleback (*Gasterosteus aculeatus*) (McLee & Scaife, 1992),

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

We assessed vertebrate communities at the pondscape using a dataset generated by eDNA metabarcoding for over 500 ponds with comprehensive environmental metadata. We validated eDNA metabarcoding as a tool for ecological hypothesis testing, and compared its outputs to previous results generated by established methods. Specifically, we tested biotic (community presence-absence data) and abiotic (environmental metadata on ponds and surrounding terrestrial habitat) determinants of *T. cristatus* at the pondscape - an impractical task by conventional means. Furthermore, we tested the applicability of the HSI to predict eDNA-based *T. cristatus* occupancy. Finally, we assessed the umbrella species

123 status of *T. cristatus* by investigating whether *T. cristatus* presence and the *T. cristatus* HSI
124 score can predict vertebrate species richness of ponds.
125
126
127 ## 2. Materials and methods
128
129 ### 2.1 Samples
130
131 We repurposed the taxonomically assigned sequence reads from Harper et al. (2018) that
132 were produced using eDNA metabarcoding of pond water to compare qPCR and eDNA
133 metabarcoding for *T. cristatus* detection. Samples from 508 ponds included in Natural
134 England's Great Crested Newt Evidence Enhancement Programme were processed using
135 eDNA metabarcoding alongside 24 privately surveyed ponds. Water samples were collected
136 using established methodology (Biggs et al., 2015), detailed in Supporting Information:
137 Appendix 1. Briefly, 20 x 30 mL water samples were collected from each pond and pooled.
138 Six 15 mL subsamples were taken from the pooled sample and each added to 33.5 mL
139 absolute ethanol and 1.5 mL sodium acetate 3 M (pH 5.2). Subsamples were pooled during
140 DNA extraction to produce one eDNA sample per pond. Targeted qPCR detected *T. cristatus*
141 in 265 (49.81%) ponds (Harper et al., 2018).
142 Environmental metadata (Table S1) were collected for 504 of 532 ponds (Fig. S1) by
143 environmental consultants contracted for Natural England's Great Crested Newt Evidence
144 Enhancement Programme. Metadata included: maximum depth of ponds; pond
145 circumference; pond width; pond length; pond area; pond density (i.e. number of ponds per
146 $km^2$); terrestrial overhang; shading; macrophyte cover; HSI score (Oldham et al., 2000); HSI

band (categorical classification of HSI score [ARG-UK, 2010]); pond permanence; water quality; pond substrate; presence of inflow or outflow; presence of pollution; presence of other amphibians, fish and waterfowl; woodland; rough grass; scrub/hedge; ruderals; other terrestrial habitat (i.e. good quality terrestrial habitat that did not conform to aforementioned habitat types); and overall terrestrial habitat quality.

### 2.2 DNA reference database construction

A custom, phylogenetically curated reference database of mitochondrial 12S ribosomal RNA (rRNA) sequences for UK fish species was previously constructed for eDNA metabarcoding of lake fish communities (Hänfling et al., 2016). Harper et al. (2018) constructed additional reference databases for UK amphibians, reptiles, birds, and mammals (Supporting Information: Appendix 1). Reference sequences available for species varied across vertebrate groups: amphibians 100.00% (*N* = 21), reptiles 90.00% (*N* = 20), mammals 83.93% (*N* = 112), and birds 55.88% (*N* = 621). Table S2 lists species without database representation, i.e. no records for any species in a genus. Sanger sequences were obtained from tissue of *T. cristatus*, *L. vulgaris*, Alpine newt (*Ichthyosaura alpestris*), common toad (*Bufo bufo*), and common frog (*Rana temporaria*) to supplement the amphibian database (Supporting Information: Appendix 1). The complete reference databases compiled in GenBank format were deposited in a GitHub repository and permanently archived (https://doi.org/10.5281/zenodo.1188710) by Harper et al. (2018).

169 **2.3 Primer validation**

170

171 Reference databases were combined for *in silico* validation of published 12S rRNA primers
172 12S-V5-F (5'-ACTGGGATTAGATACCCC-3') and 12S-V5-R (5'-TAGAACAGGCTCCTCTAG-3') (Riaz
173 et al., 2011) using ecoPCR software (Ficetola et al., 2010). Set parameters allowed a 50-250
174 bp fragment and three mismatches between each primer and reference sequence. Primers
175 were validated *in vitro* for UK fish by Hänfling et al. (2016) and by Harper et al. (2018) for six
176 UK amphibian species (Fig. S2).

177

178 **2.4 eDNA metabarcoding**

179

180 We used the taxonomically assigned sequence reads generated with vertebrate eDNA
181 metabarcoding by Harper et al. (2018). The eDNA metabarcoding workflow is fully described
182 in Harper et al. (2018) and Supporting Information: Appendix 1. eDNA was first amplified
183 with the aforementioned primers, where PCR positive controls (six per PCR plate; *n* = 114)
184 were cichlid (*Rhamphochromis esox*) DNA (0.284 ng/µL) and PCR negative controls (six per
185 PCR plate; *n* = 114) were sterile molecular grade water (Fisher Scientific UK Ltd, UK). PCR
186 products were individually purified using E.Z.N.A® Cycle Pure V-Spin Clean-Up Kits (Omega
187 Bio-tek, GA, USA) following the manufacturer's protocol. The second PCR bound Multiplex
188 Identification tags to the purified products. PCR products were individually purified using
189 magnetic bead clean-up and quantified with a Quant-IT™ PicoGreen™ dsDNA Assay
190 (Invitrogen, UK). Samples were normalised, pooled, and libraries quantified using a Qubit™
191 dsDNA HS Assay (Invitrogen, UK). Libraries were sequenced on an Illumina® MiSeq using 2 x
192 300 bp V3 chemistry (Illumina, Inc, CA, USA) and raw sequence reads processed using

193 metaBEAT (metaBarcoding and Environmental Analysis Tool) v0.97.7
194 (https://github.com/HullUni-bioinformatics/metaBEAT). After quality filtering, trimming,
195 merging, chimera detection, and clustering, non-redundant query sequences were
196 compared against our reference database using BLAST (Zhang et al., 2000). Putative
197 taxonomic identity was assigned using a lowest common ancestor (LCA) approach based on
198 the top 10% BLAST matches for any query matching with at least 98% identity to a reference
199 sequence across more than 80% of its length. Unassigned sequences were subjected to a
200 separate BLAST against the complete NCBI nucleotide (nt) database at 98% identity to
201 determine the source via LCA as described above. The bioinformatic analysis was archived
202 (https://doi.org/10.5281/zenodo.1188710) by Harper et al. (2018) for reproducibility.
203
204 **2.5 Data analysis**
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 at: https://doi.org/10.5281/zenodo.2634427. Assignments from different
209 databases were merged, and spurious assignments (i.e. non-UK species, invertebrates and
210 bacteria) removed from the dataset. The family Cichlidae was reassigned to
211 *Rhamphochromis esox*. The green-winged teal (*Anas carolinenisis*) was reassigned to *Anas*
212 (Dabbling ducks) because this species is a rare migrant and available reference sequences
213 were identical across our metabarcode to those for mallard (*Anas platyrhynchos*) and
214 Eurasian teal (*Anas crecca*), which are widely distributed across the UK. Scottish wildcat
215 (*Felis silvestris*) does not occur at the sampling localities (Kent, Lincolnshire and Cheshire)
216 and was therefore reassigned to domestic cat (*Felis catus*). Wild boar (*Sus scrofa*) and grey

wolf (*Canis lupus*) were reassigned to domestic pig (*Sus scrofa domesticus*) and domestic dog (*Canis lupus familiaris*) given the restricted distribution of *S. scrofa* and absence of *C. lupus* in the UK. The genus *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 49227, 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, our 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 species presence-absence matrix. Hypotheses tested relating to biotic and abiotic determinants of *T. cristatus* occupancy as well as the umbrella status of *T. cristatus* are summarised in Table 1. We employed

241 Generalized Linear Mixed-effects Models (GLMMs) using the package lme4 v1.1-12 (Bates et
242 al., 2015) for hypothesis testing. First, we investigated the influence of vertebrate group
243 species richness on *T. cristatus* occupancy using a binomial GLMM with the logit link
244 function that included species richness of other amphibians, fish, waterfowl, terrestrial
245 birds, and mammals as fixed effects and pond as a random effect ($N$ = 532). We performed a
246 preliminary analysis using the package cooccur v1.3 (Griffith, Veech & Marsh, 2016) to
247 identify species associations between *T. cristatus* and other vertebrates ($N$ = 532). Identified
248 associations in conjunction with the existing *T. cristatus* literature informed candidate biotic
249 variables to be modelled against *T. cristatus* occupancy ($n$ = 504). The existing *T. cristatus*
250 literature informed candidate abiotic variables to be modelled against *T. cristatus*
251 occupancy ($n$ = 504).

252 Selection of a suitable set of explanatory variables and modelling framework is fully
253 described in Supporting Information: Appendix 1. Briefly, candidate biotic and abiotic
254 explanatory variables were assessed for collinearity, relative importance, and non-linearity.
255 We constructed separate binomial GLMMs with the logit link function for biotic and abiotic
256 explanatory variables that included sample as a random effect. We modelled HSI score
257 (fixed effect) and sample (random effect) separately to prevent HSI score masking variation
258 caused by the individual biotic and abiotic variables it encompasses. Using a Poisson GLMM
259 with sample as a random effect, we tested the umbrella species status of *T. cristatus* by
260 modelling vertebrate species richness against *T. cristatus* presence-absence and the *T.*
261 *cristatus* HSI score.

262 For each GLMM, we employed an information-theoretic approach using Akaike
263 Information Criterion (AIC) to determine the most parsimonious approximating model to
264 make predictions (Akaike, 1973). Biotic and abiotic models considered respectively were

nested thus the best models were chosen using stepwise backward deletion of terms based on Likelihood Ratio Tests (LRTs). The HSI score and vertebrate species richness models were compared to null GLMMs. 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 v0.2-4 (Lele, Keim & Solymos, 2016), quantile-quantile plots, and partial residual plots (Zuur et al., 2009). Model predictions were obtained using the predictSE function in the package AICcmodavg v2.0-3 (Mazerolle, 2016) and upper and lower 95% CIs were calculated from the standard error of the predictions. Results were plotted using the package ggplot2 v2.1.0 (Wickham, 2016).

## 3. Results

### 3.1 eDNA metabarcoding

A total of 532 eDNA samples and 228 PCR controls were processed across two sequencing runs. The runs generated raw sequence read counts of 36,236,862 and 32,900,914 respectively. After quality filtering, trimming, and merging of paired-end reads, 26,294,906 and 26,451,564 sequences remained. Following removal of chimeras and redundancy via clustering, the libraries contained 14,141,237 and 14,081,939 sequences (average read counts of 36,826 and 36,671 per sample respectively), of which 13,126,148 and 13,113,143 sequences were taxonomically assigned. The final dataset (assignments corrected, thresholds applied, and assignments removed) contained 53 vertebrate species (Table S5),

289 including six amphibians, 14 fish, 17 birds, and 16 mammals (Fig. 1).

290

291 **3.2 Pondscape biodiversity**

292

293 All native amphibians were found as well as the non-native marsh frog (*Pelophylax*
294 *ridibundus*). *T. cristatus* (*n* = 148), *L. vulgaris* (*n* = 151) and *R. temporaria* (*n* = 122) were
295 widespread, but *B. bufo* (*n* = 42), palmate newt (*Lissotriton helveticus* [*n* = 5]) and *P.*
296 *ridibundus* were uncommon (*n* = 1). The threatened European eel (*Anguilla anguilla* [*n* =
297 15]), European bullhead (*Cottus gobio* [*n* = 14]), and *C. carassius* (*n* = 2) were detected
298 alongside native fishes, such as pike (*Esox Lucius* [n = 17]) and roach (*Rutilus rutilus* [*n* = 71]),
299 but also introduced species, including *C. carpio* (*n* = 40), ruffe (*Gymnocephalus cernua* [*n* =
300 1]), and rainbow trout (*Oncorhynchus mykiss* [*n* = 3]). Some identified waterfowl were
301 ubiquitous, such as common moorhen (*Gallinula chloropus* [*n* = 211]), whereas others were
302 less common, e.g. grey heron (*Ardea cinerea* [*n* = 1]) and Eurasian oystercatcher
303 (*Haematopus ostralegus* [*n* = 1]). Terrestrial fauna were often detected in fewer than five
304 ponds (Figs. 1c, d). Buzzard (*Buteo buteo* [*n* = 4]), Eurasian jay (*Garrulus glandarius* [*n* = 7]),
305 dunnock (*Prunella modularis* [*n* = 4]), and starling (*Sturnus vulgaris* [*n* = 4]) were the most
306 frequently detected terrestrial birds. Introduced mammals (Mathews et al., 2018), such as
307 grey squirrel (*Sciurus carolinensis* [*n* = 57]) and Reeve's muntjac (*Muntiacus reevesi* [*n* = 3]),
308 outweighed native mammals. Nonetheless, we detected several mammals with Biodiversity
309 Actions Plans and/or of conservation concern (Mathews et al., 2018), including otter (*Lutra*
310 *lutra* [*n* = 1]), water vole (*Arvicola amphibious* [*n* = 16]), European polecat (*Mustela putorius*
311 [*n* = 1]), brown hare (*Lepus europaeus* [*n* = 1]) and water shrew (*Neomys fodiens* [*n* = 8]).
312 Notably, the invasive American mink (*Neovison vison*) was absent despite widespread UK

313 distribution (Mathews et al., 2018). All species and their detection frequencies are listed in
314 Table S5.
315
316 **3.3 Biotic determinants of *T. cristatus* occupancy**
317
318 *T. cristatus* occupancy was positively influenced by amphibian and waterfowl species
319 richness, yet negatively influenced by mammal species richness. *T. cristatus* occupancy was
320 reduced as fish and terrestrial bird species richness increased, but these trends were not
321 significant (Fig. 2; GLMM: overdispersion $\theta$ = 0.994, $\chi^2_{525}$ = 521.778, $P$ = 0.532; fit $\chi^2_8$ =
322 48.537, $P$ < 0.001, $R^2$ = 10.91%). *T. cristatus* had significant ($P$ < 0.05) positive associations
323 with three species (Fig. S5), including *L. vulgaris*, common coot (*Fulica atra*), and *G.*
324 *chloropus*. However, *T. cristatus* had significant ($P$ < 0.05) negative associations with six
325 species (Fig. S5), including *B. bufo*, *C. carpio*, *G. aculeatus*, *P. pungitius*, common pheasant
326 (*Phasianus colchicus*), and *S. carolinensis*. Only presence-absence of *L. vulgaris*, *B. bufo*, *C.*
327 *carpio*, *G. aculeatus*, and *G. chloropus* were retained by model selection as explanatory
328 variables for the biotic GLMM of *T. cristatus* occupancy (Figs. 3a-e; GLMM: overdispersion $\theta$
329 = 1.001, $\chi^2_{495}$ = 495.297, $P$ = 0.488; fit $\chi^2_8$ = 101.820, $P$ < 0.001, $R^2$ = 27.80%). Waterfowl
330 presence-absence was also retained in the biotic GLMM (Fig. 3f). Results of analyses are
331 summarised and compared to previously reported determinants in Table 1.
332 *T. cristatus* individuals were more likely to occupy ponds with more amphibian
333 species (Fig. 2a). *T. cristatus* was detected in 51.66% of ponds ($n$ = 151) containing *L.*
334 *vulgaris*, but in only 11.91% of ponds ($n$ = 42) with *B. bufo* (Fig. 1a). Probability of *T. cristatus*
335 occupancy was lower in ponds with more fish species, and *T. cristatus* was absent from
336 ponds with more than four fish species (Fig. 2b). *T. cristatus* was only found in 15.00% ($n$ =

337 40), 14.55% (*n* = 55) and 6.67% (*n* = 15) of ponds inhabited by *C. carpio*, *G. aculeatus* and *P.*
338 *pungitius* respectively (Fig. 1b). In contrast, *T. cristatus* individuals were more likely to occur
339 in ponds with more waterfowl species (Fig. 2c). *T. cristatus* occupied 41.67% (*n* = 48) and
340 36.02% (*n* = 211) of ponds with *F. atra* and *G. chloropus* respectively (Fig. 1c). *T. cristatus*
341 occupancy was negatively influenced by higher terrestrial bird species richness, but not
342 significantly so (Fig. 2d). However, *T. cristatus* was negatively associated with *P. colchicus*,
343 being found in only 12.00% (*n* = 25) ponds with *P. colchicus* records (Fig. 1c). *T. cristatus*
344 were less likely to occupy ponds with more mammal species (Figs. 2e). Specifically, *T.*
345 *cristatus* was negatively associated with *S. carolinensis* and found in only 15.79% (*n* = 57) of
346 ponds with *S. carolinensis* records (Fig. 1d).
347
348 **3.4 Abiotic determinants of *T. cristatus* occupancy**
349
350 Three explanatory variables were retained in the abiotic GLMM explaining *T. cristatus*
351 occupancy (GLMM: overdispersion $\theta$ = 1.004, $\chi^2_{499}$ = 500.995, *P* = 0.467; fit $\chi^2_8$ = 14.409, *P* =
352 0.072, $R^2$ = 9.73%). The probability of *T. cristatus* occupancy decreased in ponds with inflow
353 present, larger area, and higher percentage of shading (Table 1, Figs. 3g-i).
354
355 **3.5 *T. cristatus* HSI and umbrella status**
356
357 HSI score positively correlated with *T. cristatus* occupancy (GLMM: overdispersion $\theta$ = 1.012,
358 $\chi^2_{501}$ = 506.893, *P* = 0.418; fit $\chi^2_8$ = 7.270, *P* = 0.508, $R^2$ = 5.45%), where *T. cristatus*
359 individuals were more likely to occupy ponds with a higher HSI score (Table 1, Fig. 3j).
360 Vertebrate species richness was positively associated with *T. cristatus* occupancy (GLMM:

361 overdispersion θ = 0.977, $\chi^2_{500}$ = 488.687, *P* = 0.633; fit $\chi^2_8$ = -158.03, *P* = 1.000, $R^2$ =
362 13.63%), with more species detected in ponds occupied by *T. cristatus* (Fig. 4a; $\chi^2_1$ = 40.985,
363 *P* < 0.001). However, vertebrate species richness did not significantly increase with the *T.*
364 *cristatus* HSI score (Fig. 4b; $\chi^2_1$ = 1.207, *P* = 0.272).
365
366
367 **4. Discussion**
368
369 We have validated eDNA metabarcoding for ecological hypothesis testing using the
370 community data generated by this tool in combination with environmental metadata for
371 ponds. We tested biotic and abiotic determinants of *T. cristatus* occupancy, whether the HSI
372 can be applied to *T. cristatus* eDNA survey, and whether *T. cristatus* is truly an umbrella
373 species for pond conservation. *T. cristatus* occupancy was higher in ponds containing *L.*
374 *vulgaris* and *G. chloropus*, and ponds without *B. bufo*, *C. carpio*, and *G. aculeatus*. *T.*
375 *cristatus* individuals were also more likely to occupy ponds where waterfowl occurred.
376 Ponds inhabited by *T. cristatus* were typically small, absent of inflow, and not excessively
377 shaded. The *T. cristatus* HSI was appropriate for predicting *T. cristatus* occupancy, but not
378 vertebrate species richness. Nonetheless, more vertebrates were present in ponds occupied
379 by *T. cristatus* thus presence of this amphibian may indicate good quality habitat for other
380 vertebrates. Our findings demonstrate the power of eDNA metabarcoding to enhance
381 freshwater monitoring and research by providing biodiversity data *en masse* at low cost.
382

383 **4.1 Pondscape biodiversity**

384

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

394

395 **4.2 Biotic determinants of *T. cristatus* occupancy**

396

397 *T. cristatus* were more likely to occupy ponds with higher amphibian species richness -
398 particularly ponds containing *L. vulgaris* and absent of *B. bufo*. *T. cristatus* and *L. vulgaris*
399 have similar habitat requirements and tend to breed in the same ponds (Skei et al., 2006;
400 Rannap et al., 2009a; Denoël et al., 2013; Cayuela et al., 2018), with >60% overlap reported
401 (Rannap & Briggs, 2006). However, *L. vulgaris* can inhabit a broader range of habitat
402 (Rannap & Briggs, 2006; Skei et al., 2006) than *T. cristatus*, which depends on larger, deeper
403 ponds with abundant macrophytes and no fish located in open, semi-rural landscapes
404 (Denoël et al., 2013). *B. bufo* can inhabit fish-containing ponds (Manenti & Pennati, 2016)
405 and *T. cristatus* may predate *B. bufo* eggs and larvae (Langton et al., 2001). This may explain
406 the negative association between *B. bufo* and *T. cristatus* as opposed to the positively

407 associated *T. cristatus* and *L. vulgaris*.

408 *T. cristatus* occupancy marginally decreased with higher fish species richness, and *T.*
409 *cristatus* was negatively associated with *C. carpio*, *G. aculeatus*, and *P. pungitius*. These
410 fishes are common in and typical of ponds. All *T. cristatus* life stages may be predated by
411 fishes (Langton et al., 2001) and negative effects of fish presence-absence on *T. cristatus*
412 occupancy, distribution, and abundance are repeatedly reported (Joly et al., 2001; Rannap &
413 Briggs, 2006; Skei et al., 2006; Denoël & Ficetola, 2008; Rannap et al., 2009a, b; Hartel et al.,
414 2010; Denoël et al., 2013). *G. aculeatus* predates *T. cristatus* eggs and larvae (McLee &
415 Scaife, 1992; Jarvis, 2010), and has non-consumptive effects on *T. cristatus* embryos (Jarvis,
416 2010). *T. cristatus* larvae were also found to alter their behaviour when exposed to
417 predatory *G. aculeatus* but not non-predatory *C. carassius* (Jarvis, 2012), another fish
418 characteristic of ponds.

419 In our study, we detected *T. cristatus* in 50% of ponds inhabited by *C. carassius*, but
420 <20% of ponds containing large and/or predatory fishes, e.g. *C. carpio*, *G. aculeatus*, *E.*
421 *lucius*. Although fewer ponds contained *C. carassius* than *C. carpio*, *G. aculeatus* or *E. lucius*,
422 previous research also indicates large and/or predatory fish are more detrimental to *T.*
423 *cristatus* occurrence (Skei et al., 2006; Hartel et al., 2010; Chan, 2011). *C. carassius* does not
424 hinder *T. cristatus* oviposition, larval behaviour, or recruitment success (Chan, 2011; Jarvis,
425 2012), or pond invertebrate and macrophyte diversity (Stefanoudis et al., 2017). In contrast,
426 *C. carpio* foraging reduces invertebrate density and macrophyte cover (Maceda-Veiga, López
427 & Green, 2017), which lowers *T. cristatus* reproductive and foraging success and heightens
428 predator exposure (Rannap & Briggs, 2006; Gustafson et al., 2006; Chan, 2011). *C. carassius*
429 and *C. carpio* are both included among fish species assumed to negatively impact *T. cristatus*
430 and whose presence-absence is assessed for the *T. cristatus* HSI (ARG-UK, 2010). However, it

431 is evident that *C. carassius* does not directly predate *T. cristatus* or indirectly alter its
432 behaviour, reproductive success, or habitat. Therefore, we advocate a systematic re-
433 evaluation of problematic fish species for *T. cristatus* conservation.

434 *T. cristatus* was positively associated with waterfowl species richness, namely
435 presence of *F. atra* and *G. chloropus*. These waterfowl species share macrophytes and
436 macroinvertebrates as resources with amphibians, feeding on both directly (Perrow et al.,
437 1997; Paillisson & Marion, 2001; Wallau et al., 2010). *F. atra* and *G. chloropus* crop
438 emergent macrophytes to search for invertebrate prey (Paillisson & Marion, 2001; Wallau et
439 al., 2010), which may indirectly benefit *T. cristatus* foraging. Although *Fulica* spp. can also
440 pull up submerged vegetation and damage vegetation banks (Lauridsen, Jeppesen &
441 Andersen, 1993), diet is macrophyte-dominated in late summer and autumn (Perrow et al.,
442 1997) and unlikely to impact *T. cristatus* breeding in spring (Langton et al., 2001). The
443 positive association identified here between *T. cristatus* and these waterfowl most likely
444 reflects a shared preference for macrophyte-rich ponds.

445 *T. cristatus* were less likely to occupy ponds with higher mammal species richness.
446 Our preliminary cooccur analysis indicated *T. cristatus* had negative associations with *P.*
447 *colchicus* and *S. carolinensis*. The terrestrial associations identified are most likely indirect
448 and a reflection of land-use rather than direct as a result of predation or competition, but
449 further investigation would be worthwhile.

450

451 **4.3 Abiotic determinants of *T. cristatus* occupancy**

452

453 *T. cristatus* was less likely to inhabit large ponds with inflow present and a greater
454 percentage of shading. Although our results indicate *T. cristatus* prefers smaller ponds, pond

455 area does not always influence occupancy (Maletzky, Kyek & Goldschmid, 2007; Denoël &
456 Ficetola, 2008; Gustafson et al., 2011) and was deemed a poor predictor of reproductive
457 success (Vuorio et al., 2013). *T. cristatus* has been found to utilise small and large ponds
458 (Rannap & Briggs, 2006; Skei et al., 2006); however, very small ponds (<124 $m^2$) may be
459 unable to support all life stages, and larger ponds may contain fish and experience
460 eutrophication due to agricultural or polluted run-off (Rannap & Briggs, 2006). Inflow to
461 ponds may exacerbate these problems by facilitating entry of agricultural or polluted run-off
462 and connections to streams and rivers containing large, predatory fish (Freshwater Habitats
463 Trust, 2015). Our results corroborate existing research where viable *T. cristatus* populations
464 were unlikely in ponds that were shaded (Vuorio et al., 2013) or had dense macrophyte
465 cover (Rannap & Briggs, 2006; Skei et al., 2006; Hartel et al., 2010).

466 In our study, most environmental metadata available were qualitative, preventing
467 detailed analyses on pond properties and terrestrial habitat in relation to *T. cristatus*
468 occupancy. Better understanding of *T. cristatus* occupancy in relation to species interactions
469 and habitat quality could be achieved with quantitative data on pond properties (e.g. water
470 chemistry), terrestrial habitat (e.g. type, density, distance to ponds), and aquatic and
471 terrestrial habitat usage by different vertebrate species. Furthermore, given the
472 metapopulation dynamics of *T. cristatus*, future research should investigate spatial drivers
473 (e.g. land cover, pond density, climate variables, roads, rivers, elevation) of *T. cristatus*
474 occupancy using innovative modelling approaches, such as individual-based models
475 (Messager & Olden, 2018). However, acquiring this data to perform these models is a
476 phenomenal task for large numbers of ponds across a vast landscape (Denoël & Ficetola,
477 2008).
478

### 4.4 *T. cristatus* HSI and umbrella status

We found the HSI can predict eDNA-based *T. cristatus* occupancy at the UK pondscape. This contradicts conventional studies which deemed the index inappropriate for predicting *T. cristatus* occupancy or survival probabilities (Unglaub et al., 2015). We detected more vertebrates in ponds containing *T. cristatus*, which may support its status as an umbrella species for pond biodiversity and conservation (Gustafson et al., 2006). We also observed a non-significant increase in vertebrate species richness with increasing *T. cristatus* HSI score. An adapted HSI, designed to predict species richness, could help select areas for management and enhancement of aquatic and terrestrial biodiversity. Until then, presence of *T. cristatus* and its HSI may confer protection to broader biodiversity by identifying optimal habitat for pond creation and restoration to encourage populations of this threatened amphibian. The HSI is not without issue due to qualitative data used for score calculation and subjective estimation of indices (Oldham et al., 2000). For future application of this index in *T. cristatus* eDNA survey, we recommend metabarcoding to quantify some qualitatively assessed indices (e.g. water quality via macroinvertebrate diversity, fish and waterfowl presence) alongside *T. cristatus* detection. Provided rigorous spatial and temporal sampling are undertaken, eDNA metabarcoding can also generate site occupancy data to estimate relative species abundance (Valentini et al., 2016; Hänfling et al., 2016).

### 4.5 Prospects of eDNA metabarcoding for freshwater conservation, management, and research

We have demonstrated the effectiveness of eDNA metabarcoding for landscape-scale

503 biodiversity monitoring and ecological hypothesis testing. We combined metabarcoding
504 with environmental metadata to revisit hypotheses relating to biotic and abiotic
505 determinants of a threatened amphibian at the UK pondscape. Our findings will guide *T.*
506 *cristatus* conservation in the face of increasing land-use and habitat fragmentation - a
507 poignant issue as protective legislation for this species in the UK is changing. Whilst
508 conservation of threatened species and their habitat should be a priority, the bigger picture
509 should not be ignored. eDNA metabarcoding could enhance our understanding of
510 freshwater networks, particularly pondscapes, to enable more effective monitoring,
511 protection, and management of aquatic and terrestrial biodiversity. We are only now
512 beginning to realise and explore these opportunities.
513
514

515 **Acknowledgements**

516

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

523
524

## Data Availability

The taxonomically assigned sequence reads used in this study were generated by Harper et al. (2018). The raw sequence reads were archived on the NCBI Sequence Read Archive (Bioproject: PRJNA417951; SRA accessions: SRR6285413 - SRR6285678). The bioinformatics analysis was deposited in a GitHub repository and permanently archived (https://doi.org/10.5281/zenodo.1188710). R scripts and corresponding data for this study have been deposited in a separate GitHub repository which has been permanently archived (https://doi.org/10.5281/zenodo.2634427).

## 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

Cayuela, H., Grolet, O., & Joly, P. (2018). Context-dependent dispersal, public information, and heterospecific attraction in newts. *Oecologia*, **188**, 1069–1080.

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

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.

590 Evans, N. T., Li, Y., Renshaw, M. A., Olds, B. P., Deiner, K., Turner, C. R., Jerde, C. L., Lodge, D.
591 M., Lamberti, G. A., & Pfrender, M. E. (2017). Fish community assessment with eDNA
592 metabarcoding: effects of sampling design and bioinformatic filtering. *Canadian Journal*
593 *of Fisheries and Aquatic Sciences*, **74**, 1362-1374. https://doi.org/10.1139/cjfas-2016-
594 0306
595 Ficetola, G. F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., Taberlet, P., &
596 Pompanon, F. (2010). An *In silico* approach for the evaluation of DNA barcodes. *BMC*
597 *Genomics*, **11**, 434. https://doi.org/10.1186/1471-2164-11-434
598 Freshwater Habitats Trust (2015). Pond Habitat Survey Method Booklet. Freshwater
599 Habitats Trust, Oxford.
600 Griffith, D., Veech, J., & Marsh, C. (2016). cooccur: Probabilistic Species Co-Occurrence
601 Analysis in R. *Journal of Statistical Software*, **69**, 1–17.
602 https://doi.org/10.18637/jss.v069.c02
603 Gustafson, D. H., Pettersson, C. J., & Malmgren, J. C. (2006). Great crested newts (*Triturus*
604 *cristatus*) as indicators of aquatic plant diversity. *The Herpetological Journal*, **16**, 347–
605 352.
606 Gustafson, D. H., Malmgren, J. C., & Mikusiński, G. (2011). Terrestrial Habitat Predicts use of
607 Aquatic Habitat for Breeding Purposes — A Study on the Great Crested Newt (*Triturus*
608 *cristatus*). *Annales Zoologici Fennici*, **48**, 295–307. https://doi.org/10.5735/086.048.0504
609 Hänfling, B., Lawson Handley, L., Read, D. S., Hahn, C., Li, J., Nichols, P., Blackman, R. C.,
610 Oliver, A., & Winfield, I. J. (2016). Environmental DNA metabarcoding of lake fish
611 communities reflects long-term data from established survey methods. *Molecular*
612 *Ecology*, **25**, 3101–3119. https://doi.org/10.1111/mec.13660
613 Harper, L. R., Lawson Handley, L., Hahn, C., Boonham, N., Rees, H. C., Gough, K. C., Lewis, E.,
614 Adams, I. P., Brotherton, P., Phillips, S., & Hänfling, B. (2018). Needle in a haystack? A
615 comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested
616 newt (*Triturus cristatus*). *Ecology and Evolution*, *8*, 6330–6341.
617 https://doi.org/10.1002/ece3.4013
618 Harper, L. R., Buxton, A. S., Rees, H. C., Bruce, K., Brys, R., Halfmaerten, D., Read, D. S.,
619 Watson, H. V., Sayer, C. D., Jones, E. P., Priestley, V., Mächler, E., Múrria, C., Garcés-
620 Pastor, S., Medupin, C., Burgess, K., Benson, G., Boonham, N., Griffiths, R. A., Lawson
621 Handley, L., & Hänfling, B. (2019). Prospects and challenges of environmental DNA
622 (eDNA) monitoring in freshwater ponds. *Hydrobiologia*, **826**, 25–41.
623 https://doi.org/10.1007/s10750-018-3750-5
624 Hartel, T., Nemes, S., & Oellerer, K. (2010). Using connectivity metrics and niche modelling
625 to explore the occurrence of the northern crested newt *Triturus cristatus* (Amphibia,
626 Caudata) in a traditionally managed landscape. *The Environmentalist*, **37**, 195–200.
627 https://doi.org/10.1017/S037689291000055X
628 Hering, D., Borja, A., Jones, J. I., Pont, D., Boets, P., Bouchez, A., Bruce, K., Drakare, S.,
629 Hänfling, B., Kahlert, M., Leese, F., Meissner, K., Mergen, P., Reyjol, Y., Segurado, P.,
630 Vogler, A., & Kelly, M. (2018). Implementation options for DNA-based identification into
631 ecological status assessment under the European Water Framework Directive. *Water*
632 *Research*, **138**, 192–205. https://doi.org/10.1016/j.watres.2018.03.003
633 Hill, M. J., Hassall, C., Oertli, B., Fahrig, L., Robson, B. J., Biggs, J., Samways, M. J., Usio, N.,
634 Takamura, N., Krishnaswamy, J., & Wood, P. J. (2018). New policy directions for global
635 pond conservation. *Conservation Letters*, **e12447**. https://doi.org/10.1111/conl.12447

636 Jarvis, L. E. (2010). Non-consumptive effects of predatory three-spined sticklebacks
637 (*Gasterosteus aculeatus*) on great crested newt (*Triturus cristatus*) embryos. *The*
638 *Herpetological Journal*, **20**, 271–275.
639 Jarvis, L. E. (2012). Microhabitat preferences of the great crested newt *Triturus cristatus* in a
640 woodland area. PhD Thesis. Open University.
641 Joly, P., Miaud, C., Lehmann, A., & Grolet, O. (2001). Habitat Matrix Effects on Pond
642 Occupancy in Newts. *Conservation Biology*, **15**, 239–248. https://doi.org/10.1111/j.1523-
643 1739.2001.99200.x
644 Kelly, R. P., O'Donnell, J. L., Lowell, N. C., Shelton, A. O., Samhouri, J. F., Hennessey, S. M.,
645 Feist, B. E., & Williams, G. D. (2016). Genetic signatures of ecological diversity along an
646 urbanization gradient. *PeerJ*, **4**, e2444. https://doi.org/10.7717/peerj.2444
647 Klymus, K. E., Richter, C. A., Thompson, N., & Hinck, J. E. (2017). Metabarcoding of
648 Environmental DNA Samples to Explore the Use of Uranium Mine Containment Ponds as
649 a Water Source for Wildlife. *Diversity*, **9**, 1-18. https://doi.org/10.3390/d9040054
650 Langton, T., Beckett, C., & Foster, J. (2001). Great crested newt conservation handbook.
651 Froglife, Halesworth.
652 Lauridsen, T. L., Jeppesen, E., & Andersen, F. Ø. (1993). Colonization of submerged
653 macrophytes in shallow fish manipulated Lake Væng: impact of sediment composition
654 and waterfowl grazing. *Aquatic Botany*, **46**, 1–15. https://doi.org/10.1016/0304-
655 3770(93)90061-Z
656 Lawson Handley, L. (2015). How will the "molecular revolution" contribute to biological
657 recording? *Biological Journal of the Linnean Society,* **115**, 750–766.
658 https://doi.org/10.1111/bij.12516
659 Lele, S. R., Keim, J. L., & Solymos, P. (2016). ResourceSelection: Resource Selection
660 (Probability) Functions for Use-Availability Data.
661 Li, Y., Evans, N. T., Renshaw, M. A., Jerde, C. L., Olds, B. P., Shogren, A. J., Deiner, K., Lodge,
662 D. M., Lamberti, G. A., & Pfrender, M. E. (2018a). Estimating fish alpha- and beta-diversity
663 along a small stream with environmental DNA metabarcoding. *Metabarcoding and*
664 *Metagenomics*, **2**, e24262. https://doi.org/10.3897/mbmg.2.24262
665 Li, F., Peng, Y., Fang, W., Altermatt, F., Xie, Y., Yang, J., & Zhang, X. (2018b). Application of
666 environmental DNA metabarcoding for predicting anthropogenic pollution in rivers.
667 *Environmental Science & Technology*, **52**, 11708–11719.
668 https://doi.org/10.1021/acs.est.8b03869
669 Maceda-Veiga, A., López, R., & Green, A. J. (2017). Dramatic impact of alien carp *Cyprinus*
670 *carpio* on globally threatened diving ducks and other waterbirds in Mediterranean
671 shallow lakes. *Biological Conservation*, **212**, 74–85.
672 https://doi.org/10.1016/j.biocon.2017.06.002
673 Macher, J. -N., Vivancos, A., Piggott, J. J., Centeno, F. C., Matthaei, C. D., & Leese, F. (2018).
674 Comparison of environmental DNA and bulk-sample metabarcoding using highly
675 degenerate cytochrome c oxidase I primers. *Molecular Ecology Resources*, **18**, 1456–
676 1468. https://doi.org/10.1111/1755-0998.12940
677 Maletzky, A., Kyek, M., & Goldschmid, A. (2007). Monitoring status, habitat features and
678 amphibian species richness of crested newt (*Triturus cristatus* superspecies) ponds at the
679 edge of the species range (Salzburg, Austria). *Annales de Limnologie - International*
680 *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

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

728 Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers:
729 inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic*
730 *Acids Research*, **39**, e145. https://doi.org/10.1093/nar/gkr732 Port
731 Skei, J. K., Dolmen, D., Rønning, L., & Ringsby, T. H. (2006). Habitat use during the aquatic
732 phase of the newts *Triturus vulgaris* (L.) and *T. cristatus* (Laurenti) in central Norway:
733 proposition for a conservation and monitoring area. *Amphibia-Reptilia*, **27**, 309–324.
734 https://doi.org/10.1163/156853806778189972
735 Stefanoudis, P. V., Sayer, C. D., Greaves, H. M., Davidson, T. A., Robson, H., Almeida, D., &
736 Smith, E. (2017). Consequences of fish for cladoceran, water beetle and macrophyte
737 communities in a farmland pond landscape: implications for conservation. *Fundamental*
738 *and Applied Limnology*, **190**, 141–156. https://doi.org/10.1127/fal/2017/1004
739 Unglaub, B., Steinfartz, S., Drechsler, A., & Schmidt, B. R. (2015). Linking habitat suitability to
740 demography in a pond-breeding amphibian. *Frontiers in Zoology*, **12**.
741 https://doi.org/10.1186/s12983-015-0103-3
742 Ushio, M., Fukuda, H., Inoue, T., Makoto, K., Kishida, O., Sato, K., Murata, K., Nikaido, M.,
743 Sado, T., Sato, Y., Takeshita, M., Iwasaki, W., Yamanaka, H., Kondoh, M. & Miya, M.
744 (2017). Environmental DNA enables detection of terrestrial mammals from forest pond
745 water. *Molecular Ecology Resources*, **17**, e63-e65. https://doi.org/10.1111/1755-
746 0998.12690
747 Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E.,
748 Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G.
749 H., Geniez, P., Pont, D., Argillier, C., Baudoin, J. -M., Peroux, T., Crivelli, A. J., Olivier, A.,
750 Acqueberge, M., Le Brun, M., Møller, P. R., Willerslev, E. & Dejean, T. (2016). Next-
751 generation monitoring of aquatic biodiversity using environmental DNA metabarcoding.
752 *Molecular Ecology*, **25**, 929–942. https://doi.org/10.1111/mec.13428
753 Vuorio, V., Heikkinen, R. K., & Tikkanen, O. -P. (2013). Breeding Success of the Threatened
754 Great Crested Newt in Boreal Forest Ponds. *Annales Zoologici Fennici*, **50**, 158–169.
755 https://doi.org/10.5735/086.050.0303
756 Wallau, G. L., Della-Flora, F., Bueno, A. S., Corso, J., Ortiz, M. F., & Cáceres, N. C. (2010).
757 Behaviour of the Common Moorhen in Rio Grande do Sul, Brazil. *Acta Ethologica*, **13**,
758 127–135. https://doi.org/10.1007/s10211-010-0082-5
759 Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer New York, USA.
760 Wilcox, T. M., Young, M. K., McKelvey, K. S., Isaak, D. J., Horan, D. L. & Schwartz, M. K.
761 (2018). Fine-scale environmental DNA sampling reveals climate-mediated interactions
762 between native and invasive trout species. *Ecosphere*, **9**, e02500.
763 https://doi.org/10.1002/ecs2.2500
764 Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A greedy algorithm for aligning DNA
765 sequences. *Journal of Computational Biology*, **7**, 203–214.
766 https://doi.org/10.1089/10665270050081478
767 Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects*
768 *models and extensions in ecology with R*. Springer New York, USA.
769

770 **Table 1.** Summary of biotic and abiotic determinants of *T. cristatus* occupancy. Reported
771 effects on *T. cristatus* occupancy in the literature and hypothesised effects on eDNA-based
772 *T. cristatus* occupancy are given for each determinant. Any determinants not reported in the
773 literature are listed as UNK. Direction of observed effects on eDNA-based *T. cristatus*
774 occupancy determined by each analysis (GLMM assessing species richness in each
775 vertebrate group, *N* = 532; preliminary cooccur analysis, *N* = 532; GLMM for biotic factors, *n*
776 = 504; GLMM for abiotic factors *n* = 504; and GLMM assessing HSI score, *n* = 504) are given.
777 No, negative and positive effects are listed as 0, - and + respectively. For categorical
778 variables with more than one level, effect size and standard error (SE) are only given for
779 levels reported in the model summary. Test statistic is for LRT used and significant P-values
780 (<0.05) are in bold. Variables included for model selection but not retained in the final
781 model are listed as NR. The preliminary cooccur analysis was not applicable (NA) to abiotic
782 factors.
783

| Determinant | Effect reported | Hypothesised effect | Analysis | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Cooccur | | GLMM | | | |
| | | | Effect | *P* | DF | Effect size (SE) | $\chi^2$ | *P* |
| **Amphibians** | | | | | | | | |
| Species richness | UNK | | | | 1 | 0.514 (0.148) | 13.875 | **<0.001** |
| *L. vulgaris* | + | + | + | **<0.001** | 1 | 1.624 (0.302) | 50.899 | **<0.001** |
| *B. bufo* | UNK | | - | **0.010** | 1 | -1.799 (0.659) | 11.150 | **0.001** |
| **Fish** | | | | | | | | |
| Species richness | UNK | | | | 1 | -0.186 (0.123) | 2.415 | **0.120** |
| *G. aculeatus* | - | - | - | **0.012** | 1 | -1.053 (0.469) | 6.302 | **0.012** |
| *C. carpio* | - | - | - | **0.040** | 1 | -1.463 (0.603) | 7.841 | **0.005** |
| *P. pungitius* | - | - | - | **0.049** | NR | | | |
| *C. carassius* | - | - | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Waterfowl** | | | | | | | | |
| Species richness | UNK | | | | 1 | 0.475 (0.141) | 12.715 | **<0.001** |
| *G. chloropus* | UNK | | + | **0.001** | 1 | 0.897 (0.251) | 15.705 | **<0.001** |
| *F. atra* | UNK | | + | **0.022** | | | NR | |
| Waterfowl presence | **-** | **-** | | | 2 | | 6.352 | **0.042** |
| Minor | | | | | | 0.534 (0.244) | | |
| Major | | | | | | 0.812 (0.506) | | |
| **Terrestrial birds** | | | | | | | | |
| Species richness | UNK | | | | 1 | -0.046 (0.302) | 0.024 | 0.878 |
| *P. colchicus* | UNK | | - | **0.050** | | | | |
| **Terrestrial mammals** | | | | | | | | |
| Species richness | UNK | | | | 1 | -0.519 (0.193) | 8.599 | **0.007** |
| *S. carolinensis* | UNK | | - | **0.020** | | | | |
| **Pond area** | -/+ | - | NA | NA | 1 | -0.0003 (0.0002) | 4.108 | **0.043** |
| **Pond density** | + | + | NA | NA | | | NR | |
| **Pond depth** | + | + | NA | NA | | | NR | |
| **Pond substrate** | + | + | NA | NA | | | NR | |
| **Pond permanence** | + | + | NA | NA | | | NR | |
| **Water quality** | + | + | NA | NA | | | NR | |
| **Inflow** | - | - | NA | NA | 1 | | 15.066 | **<0.001** |
| Present | | | | | | -0.887 (0.249) | | |
| **Outflow** | UNK | | NA | NA | | | NR | |
| **Macrophyte cover** | -/+ | - | NA | NA | | | NR | |
| **Shading** | -/+ | - | NA | NA | 1 | -0.010 (0.003) | 12.133 | **<0.001** |
| **Woodland** | + | + | NA | NA | | | NR | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Scrub/hedge** | + | + | NA | NA | | | NR | |
| **Ruderals** | UNK | | NA | NA | | | NR | |
| **HSI score** | 0/+ | + | NA | NA | 1 | 3.164 (0.798) | 17.039 | **<0.001** |

784

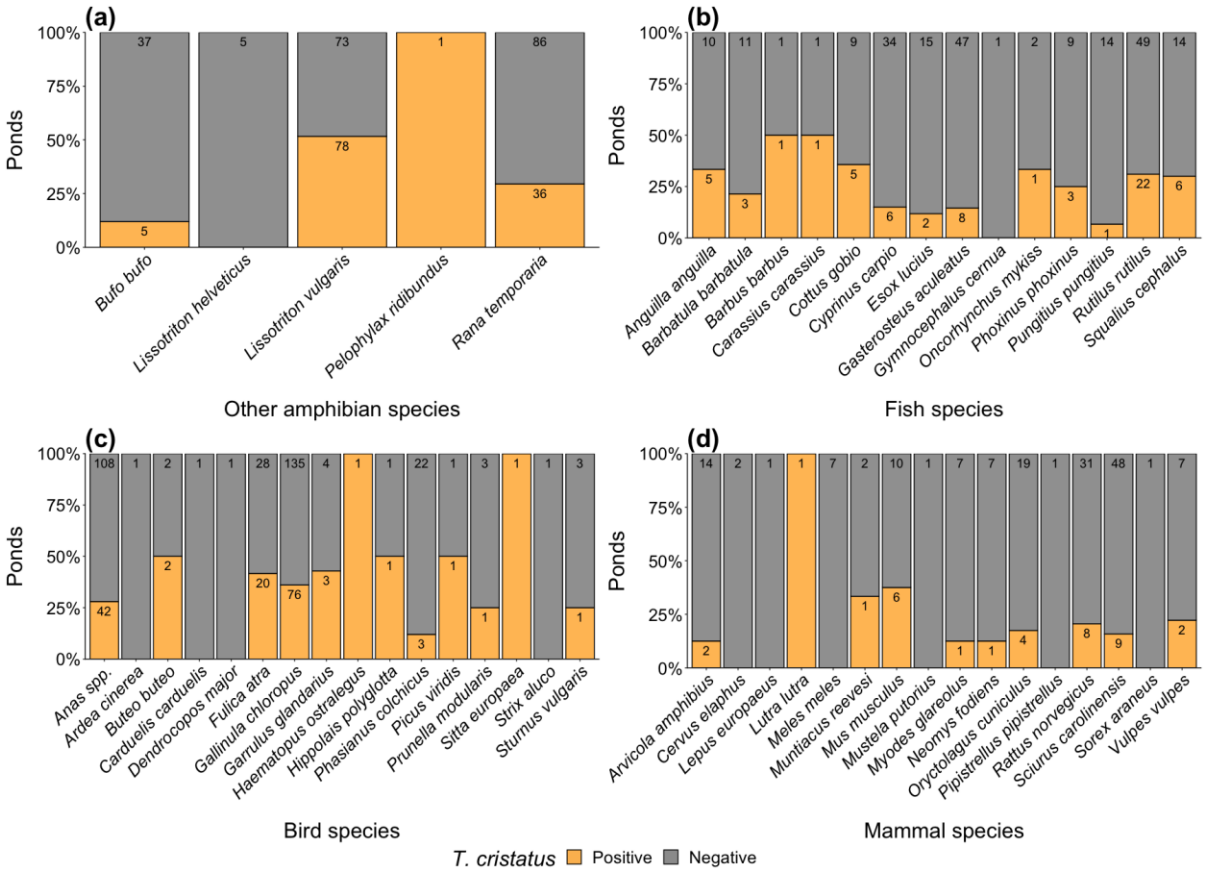
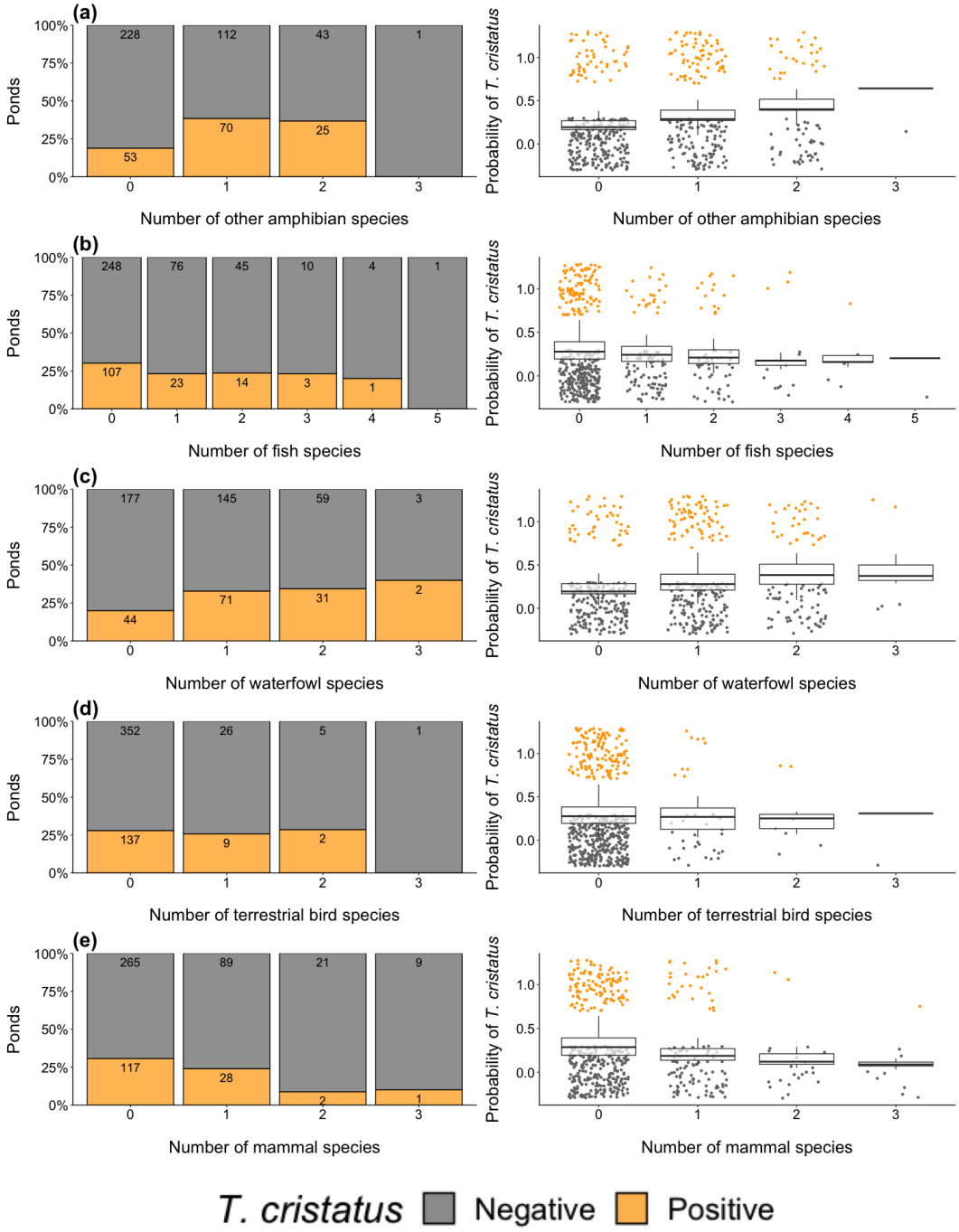



**Figure 1.** eDNA metabarcoding detection of *T. cristatus* in relation to other vertebrate species (*N* = 532 ponds): **(a)** other amphibians, **(b)** fish, **(c)** birds, and **(d)** mammals. Numbers on each bar are the number of ponds with (orange) and without (grey) *T. cristatus* in which a vertebrate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




791

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

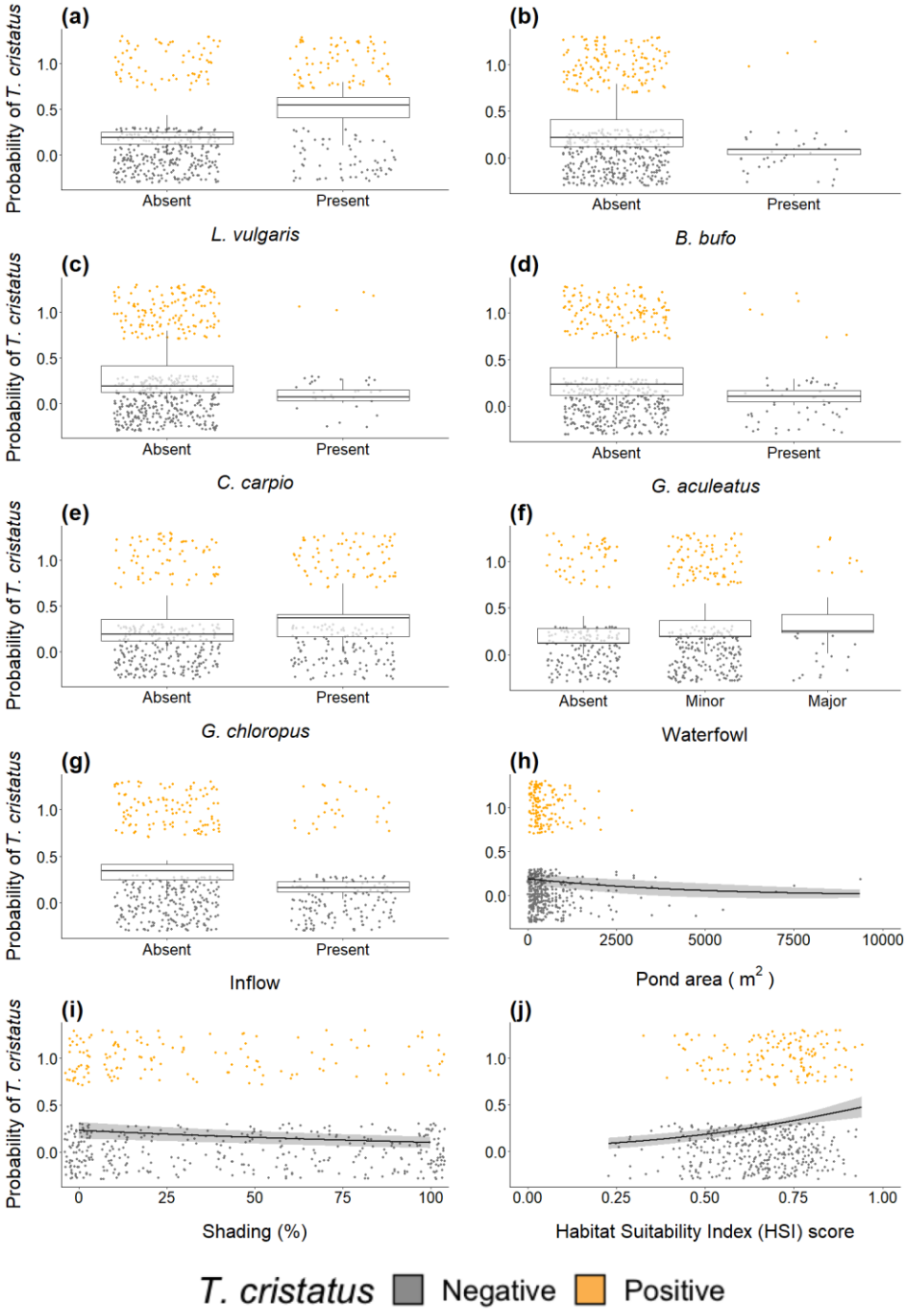



**Figure 3.** Biotic and abiotic determinants of *T. cristatus* occupancy (*n* = 504 ponds): **(a)** *L. vulgaris* occupancy, **(b)** *B. bufo* occupancy, **(c)** *C. carpio* occupancy, **(d)** *G. aculeatus* occupancy, **(e)** *G. chloropus* occupancy, **(f)** extent of waterfowl presence, **(g)** presence 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* presence (orange) and absence (grey) data are displayed as points (jittered around 0 and 1 to clarify variation in point density) against the predicted relationships (boxes/lines).

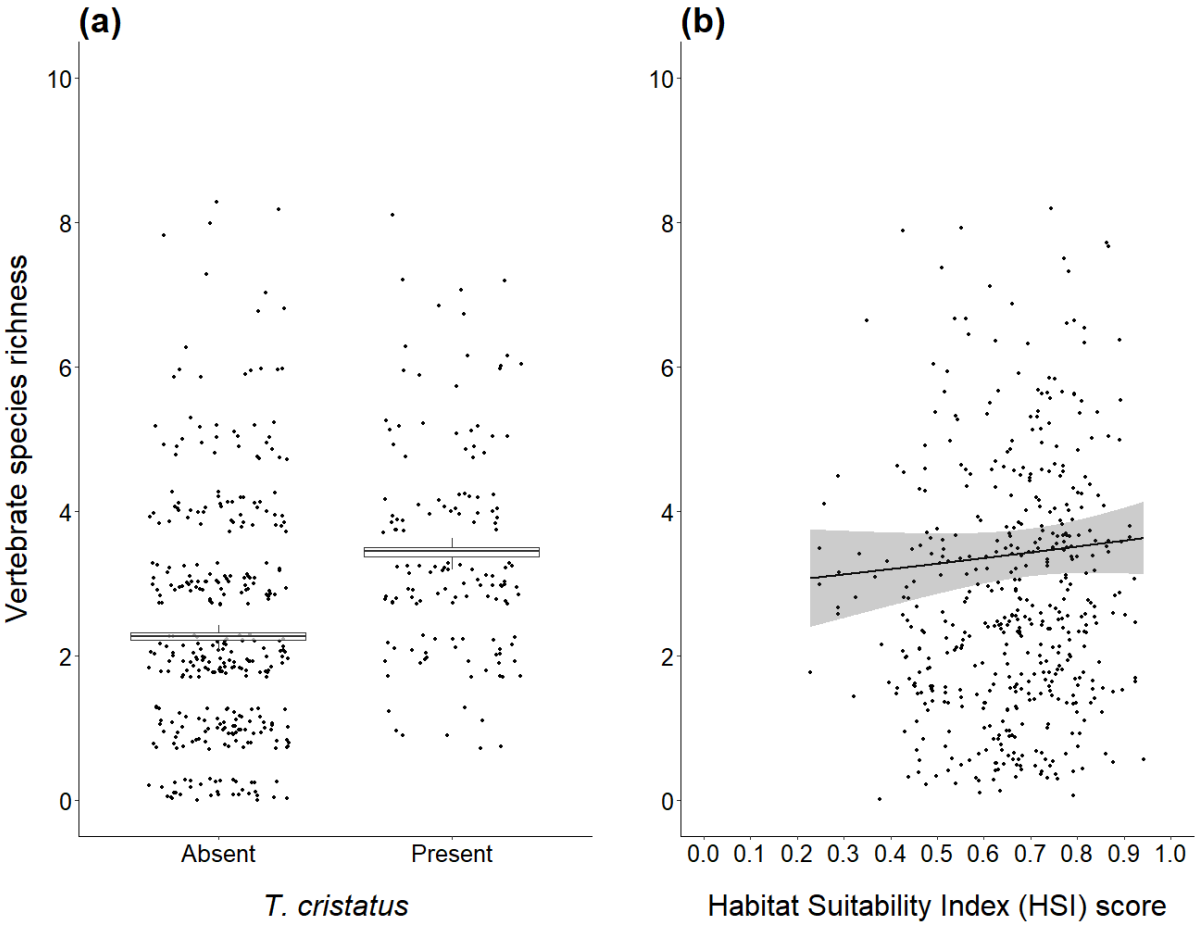



**Figure 4.** Vertebrate species richness in ponds (*n* = 504) in relation to: **(a)** *T. cristatus* occupancy and **(b)** the *T. cristatus* HSI score. The 95% CIs, as calculated using the predicted species richness values and standard error for these predictions, are given for each relationship. The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