Molecular Ecology Resources

WILEY

MOLECULAR ECOLOGY RESOURCES

RESOURCE ARTICLE OPEN ACCESS

Navigating Methodological Trade-Offs in eDNA Metabarcoding Biodiversity Monitoring: Insights From a Mediterranean Watershed

Joana Veríssimo[1,2,3] | Manuel Lopes-Lima[1,2] | Fábio Amaral[1,2] | Cátia Chaves[1,2] | Vasco Fernandes[1,2] | Mutaleni Kemanja[4] | Amílcar Teixeira[5] | Filipa M. S. Martins[1,2] | Pedro Beja[1,2,6]

[1]CIBIO, Centro de Investigação Em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, Vairão, Portugal | [2]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, Vairão, Portugal | [3]Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Porto, Portugal | [4]Namibia University of Science and Technology, Windhoek, Namibia | [5]CIMO, LA SusTEC, Instituto Politécnico de Bragança, Campus de Santa Apolónia, Bragança, Portugal | [6]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Instituto Superior de Agronomia, Tapada da Ajuda, Universidade de Lisboa, Lisboa, Portugal

Correspondence: Joana Veríssimo (j_verissimo@cibio.up.pt)

Received: 12 July 2024 | **Revised:** 19 December 2024 | **Accepted:** 27 January 2025

Handling Editor: Andrew P. Kinziger

Funding: This work was supported by Fundação para a Ciência e a Tecnologia, 2020.03608.CEECIND, COVID/BD/152600/2022, SFRH/BD/133159/2017, UIDP/50027/2020, LA/P/0048/2020 and European Regional Development Fund (ERDF), NORTE-01-0246-FEDER-000063.

Keywords: biodiversity monitoring | environmental DNA | freshwater ecosystems | high-throughput sequencing | metabarcoding | replication

ABSTRACT

Environmental DNA (eDNA) metabarcoding technologies promise significant advances in biodiversity monitoring, yet their application requires extensive optimisation and standardisation. Recent research demonstrated that increased sampling and analytical efforts are needed to improve biodiversity estimates, though fully optimising study designs is often hindered by resource constraints. Consequently, researchers must carefully navigate methodological trade-offs to design effective eDNA metabarcoding monitoring studies. We conducted a water eDNA survey of vertebrates in a Mediterranean watershed to identify key methodological factors influencing species richness and composition estimates. We examined the impacts of using high- versus low-capacity filtration capsules, varying levels of biological and technical replication, and the pooling of PCR replicates before indexing. The primary sources of variation identified were capsule filtration capacity and site replication across the watershed. While biological replication within sites and PCR replication also improved biodiversity estimates, their effects were comparatively smaller. Pooling PCR replicates before indexing performed more poorly than analysing them independently. Methodological impacts were stronger on terrestrial than on aquatic species. Based on these results, we recommend that priority should be given to high-capacity filtration and sampling across multiple sites. Site-level replication deserves lower priority, especially when filtering large water volumes. PCR replication is crucial for detecting rare species but should be balanced with increased site sampling and eventually site-level replication. Avoiding the pooling of PCR replicates is important to enhance sensitivity for rare species. Overall, we stress the importance of balancing methodological choices with resource constraints and monitoring goals, and we emphasise the need for research assessing methodological trade-offs in different study systems.

1 | Introduction

Monitoring is critical for assessing the state and trends in biodiversity and ecosystem health (Pereira et al. 2010), and it is key for the implementation of the Kunming-Montreal Global Biodiversity Framework (CBD 2022). Yet, monitoring needs to be done over large geographic areas, across borders, and with high temporal and spatial resolution, thereby requiring the development of standardised methods that should be relatively simple, rapid, and cost-effective (Moersberger et al. 2024). Environmental DNA (eDNA) metabarcoding has emerged over the past decade as highly promising tools, capable of fulfilling most, if not all, of these requirements (Blackman et al. 2024; Fediajevaite et al. 2021; Keck et al. 2022; Ruppert et al. 2019; Schenekar 2023). These technologies offer significant advantages over conventional techniques, particularly in consistently detecting rare or elusive species (Duarte et al. 2023; Neice and McRae 2021) and in characterising entire communities of diverse organisms, especially where taxonomic expertise is limited (Elbrecht et al. 2017; Pereira et al. 2021). To fulfil these promises, substantial efforts have been undertaken to optimise and standardise metabarcoding-based monitoring protocols (Blackman et al. 2024; Bohmann et al. 2021; Schenekar 2023; Shu et al. 2020; Zinger et al. 2019). However, it remains far from trivial to identify appropriate methods for practical monitoring applications, particularly where they are constrained by logistical challenges and limited human and financial resources (e.g., Lopes-Lima et al. 2024).

Monitoring based on eDNA captured from water provides a case in point, where gaps persist between methodological development and practical monitoring applications (Bruce et al. 2021; Fonseca et al. 2023; Schenekar 2023). Many studies have examined how methodological options in metabarcoding protocols affect species detections and the estimation of community descriptors (e.g., diversity, composition), from field sampling effort, through lab analysis, to bioinformatic processing (Piggott 2016; Stauffer et al. 2021; Xiong et al. 2022; Yao et al. 2022). The general conclusion of these studies is that 'the more, the better', with increases in sampling and analysis efforts producing better results, though often involving significant methodological trade-offs that must be carefully considered during study design. For instance, earlier studies captured aquatic eDNA by filtering small volumes of water (< 2 L) in the field, using simple methods that could be easily replicated across vast areas with the collaboration of citizen scientists (Agersnap et al. 2022; Burian et al. 2023; Clarke et al. 2023). However, there is at present strong evidence that filtering such small volumes hinders the detection of all but the most abundant species (Peixoto et al. 2023; Polanco et al. 2022; Stauffer et al. 2021), with recent studies recommending the filtration of tens or even hundreds of litres (Cantera et al. 2019; Schabacker et al. 2020; Sepulveda et al. 2019). Such high volumes require the use of high-capacity capsules and other filtration equipment that is more expensive and needs to be operated by specialised personnel, thereby potentially limiting the number of sites that can realistically be sampled (Govindarajan et al. 2022; Schabacker et al. 2020).

Deciding on the levels of biological and technical replication is also a key component of monitoring design, with increasing replication enhancing the accuracy of biodiversity estimates, albeit at the expense of higher costs. At the biological level, spatial replication may involve sampling multiple sites within a region, which can help detect more species by covering larger areas and thus produce more accurate estimates of regional species diversity and composition (Buxton et al. 2021; Cantera et al. 2019; Ficetola et al. 2016; Stauffer et al. 2021; Yao et al. 2022). However, this entails higher investment in travel and operation time and costs, which may be unwarranted if eDNA in the water is well-mixed (e.g., through river flow), in which case just a few strategically located sampling sites may be sufficient to capture regional biodiversity (Cantera et al. 2022; Deiner et al. 2016; Pont et al. 2018). Biological replication may also involve taking multiple samples from each sampling site, which can increase the accuracy of local biodiversity estimates and has the added value of permitting the estimation of species detection probabilities (e.g., Beentjes et al. 2019; Peixoto et al. 2023; Zhang et al. 2020). Yet, this effort may be unnecessary if sample completeness is increased by processing large volumes of water through a single filter per site (Bessey et al. 2020; Cantera et al. 2019; Schabacker et al. 2020). Regarding technical replication, many studies recommend the use of DNA extraction or PCR replicates, or both, to enhance the reliability of biodiversity estimates. For instance, up to 12 PCR replicates or even more may be needed to detect rare species (Ficetola et al. 2015), though this can greatly increase lab costs. To reduce such costs, PCR replicates may be pooled before the addition of indexes (Coghlan et al. 2021; Sakata et al. 2022). Sato et al. (2017) provided insights into the impact of pooling biological or sampling replicates in lentic environments on fish species detections. However, they did not assess the implications of pooling multiple technical PCR replicates in comparison to the same number of individual PCRs sequenced independently.

These and other methodological trade-offs can have strong impacts on the results of eDNA monitoring of biodiversity, as maximising all aspects of sampling and analysis efforts simultaneously is often impractical. Therefore, understanding which components of the study design most impact biodiversity estimates is critical to optimising resource allocation for the greatest effect. Here we investigate these issues through an eDNA survey of vertebrates across a Mediterranean watershed. Specifically, we implemented sampling and analytic designs to determine sources of variability in biodiversity estimates, focusing on species richness and community composition of aquatic and terrestrial vertebrates. Our approach examined the effects of (a) utilising high- versus low-capacity filtration capsules; (b) sampling multiple sites within a catchment; (c) taking multiple replicates at each site; (d) conducting multiple PCR replicates per sample; and (e) pooling PCR replicates before indexing. Our findings provide practical guidelines for addressing methodological trade-offs in eDNA biodiversity monitoring.

2 | Materials and Methods

2.1 | Study Design

The effects of capsule capacity, biological and technical replication, and PCR replicate pooling on biodiversity estimates were assessed using two separate experimental designs. Experiment A involved sampling for eDNA at 28 sites within the Sabor watershed (Figure 1, Table S1), with four field sampling replicates

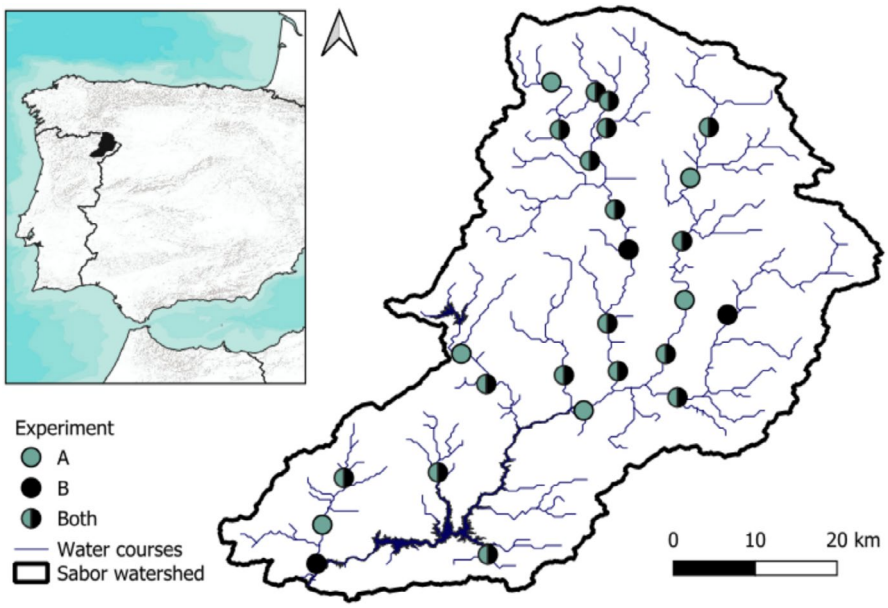

FIGURE 1 | Map of Sabor river watershed in northeastern Portugal, showing the location of sampling sites used solely for Experiment A (green points), Experiment B (black points), and for both Experiments.

collected per site using high- and low-capacity capsules. eDNA from each sample was extracted in the laboratory and then amplified using three separate PCRs. Each PCR product was then analysed following the procedures described below. This experiment was used to partition the variability in estimates of species richness and composition resulting from sampling multiple sites, multiple replicates within sites using capsules with high and low filtration capacity, and multiple PCRs within replicate samples. Experiment B involved a subsample of 20 sites used in Experiment A (Figure 1), where we randomly selected one high-capacity capsule per site. eDNA from each capsule was extracted in the lab and amplified using 12 separate PCR replicates. These replicates were then analysed using two different approaches: (i) each replicate was individually indexed, purified, quantified, normalised, and pooled at equimolar concentrations; and (ii) the 12 replicates were pooled, and then a procedure similar to (i) was followed. This experiment was used to compare species richness estimated using pooled versus separate PCRs at similar levels of sequencing coverage. We also estimate whether detection of individual species in the pool was affected by their prevalence in the individual PCR replicates.

2.2 | eDNA Sampling

Sampling for eDNA was undertaken at sites distributed within the watershed of the Sabor River (NE Portugal), where previous studies have thoroughly described the vertebrate fauna using conventional monitoring methods (e.g., Ferreira et al. 2016; Mota-Ferreira and Beja 2020; Quaglietta et al. 2018). At each site, we filtered water using both high-capacity capsules (VigiDNA 0.45 μM filtration kit; K-0101; SPYGEN, France) and low-capacity capsules (Sterivex 0.45 μM; SVHVL10RC; Merck Millipore, Germany). We collected two field replicates of each capsule at each sampling point. Water filtration was aided with a peristaltic pump (Vampire sampler, Germany) with disposable tubing until the filter clogged. The capsules were then emptied and filled with CL1 preservation buffer (SPYGEN) to minimise eDNA degradation. Capsules were stored at room temperature until eDNA extraction. No field negatives were included in this study, as all the materials used for sample collection were disposable.

2.3 | DNA Extraction

DNA extraction was performed at CIBIO molecular labs in a dedicated room equipped with positive pressure and UV treatment, following strict protocols that include disposable laboratory wear. All benches were decontaminated with a 10% commercial bleach solution before and after each manipulation and exposed to UV light for at least 10 min. Before DNA extraction, capsules were agitated in a rocker at 50 rpm for 30 min. For high-capacity capsules, the CL1 buffer was transferred to a 50 mL tube, centrifuged at 15,000 *g* for 15 min, and the supernatant discarded, leaving 15 mL. This step was repeated until all CL1 buffer was processed (each capsule contained 80 mL of the CL1 buffer), as a maximum volume of 50 mL can be processed by most common centrifuges available in laboratories. DNA was then precipitated by adding 33 mL of ethanol and 1.5 mL of 3 M sodium acetate to each sample and stored at −20°C overnight. For low-capacity capsules, the initial centrifugation step was not necessary as CL1 buffer volume was small enough to allow for direct DNA precipitation. The volumes of ethanol and 3 M sodium acetate were adjusted on a sample-by-sample basis to account for the volume of CL1 buffer present in each capsule (on average 1.8 mL ± 0.3 of CL1). Samples were then centrifuged at 15,000 *g* for 15 min (high-capacity capsules) or 4000 *g* for 30 min (low-capacity capsules) at 6°C, and the supernatant discarded, leaving a pellet. To this, 1 mL of PW1 solution from the DNeasy PowerWater protocol (Qiagen, Germany) and 40 μL of proteinase K (NZYtech, Portugal) were added and incubated at 56°C for 3 h. Samples were centrifuged at 4000 *g* for 1 min. From this step, DNA extraction was performed using the DNeasy PowerWater

protocol from step 8 according to the manufacturer's instructions. Eluted DNA was diluted 1:2 prior to further processing. A negative control (i.e., a sample tube with reagents and no DNA template) was included in each extraction batch of 12 samples.

2.4 | PCR Amplification, Library Preparation, and Sequencing

To assess vertebrate diversity, we selected the primer pair Vert01 (also commonly referred to as 12S-V5; Riaz et al. 2011; Taberlet et al. 2018), which amplifies a short portion (average amplicon length of 97 bps) of the 12S rRNA. This primer has been successfully used to assess vertebrate diversity in different aquatic ecosystems worldwide (Ando et al. 2020; Lopes-Lima et al. 2024; Polanco Fernández et al. 2021; Schneider et al. 2021; Xiong et al. 2022). Primers were ordered with 5′ adaptor sequences compatible with a two-step protocol for Illumina sequencing as described in Egeter et al. (2022). Additionally, we included a 7-bp heterogeneity spacer (Fadrosh et al. 2014) to enhance sequencing quality. For experiment A, all samples were amplified in triplicate. As experiment B contains samples used in experiment A, 9 additional PCR replicates (totalling 12) were performed for these samples. PCRs were performed with a final volume of 11 μL, containing 5 μL of 2× Qiagen Multiplex PCR Master Mix (Qiagen, Germany), 0.4 μL of each primer (10 μM), 2.7 μL of ddH_2O, and 2.5 μL of eDNA on a T100 Thermal Cycler (BioRad, USA). PCR conditions began with an initial denaturation at 95°C for 15 min, followed by 40 cycles of denaturation at 95°C for 30 s, annealing at 49°C for 30 s, and extension at 72°C for 15 s, followed by a final extension at 60°C for 10 min. One negative control (i.e., sample reaction with PCR reagents and no DNA template) was included per PCR replicate plate. Amplification success was confirmed by electrophoresis in 2% agarose gels stained with GelRed (Biotium, USA). All samples (including those with low yields) and negative controls were considered for sequencing.

Illumina P5 and P7 adapters with unique 7 bp index combinations were incorporated through a 2nd PCR (see Appendix S1 for details). For Experiment A, all PCR replicates were individually indexed. For Experiment B, two approaches were implemented: (i) For each sample, all 12 PCR replicates were individually indexed as in Experiment A (hereafter referred to as 'IndRep'); (ii) For each sample, we pooled 5 μL from each of the 12 PCR replicates, and then each pooled sample was indexed (hereafter referred to as 'Pool'). See Appendix S1 for additional details on pooling procedures. For each experiment (i.e., Experiment A, Experiment B—IndRep, and Experiment B—Pool), all indexed samples were purified using AMPure XP beads (Beckman Coulter, USA) at a 0.9x ratio, quantified using a BioTek Epoch Spectrophotometer (Agilent Technologies, USA), and pooled at equimolar concentrations. Each pool was then validated using the High Sensitivity D5000 Screen Tape Assay (TapeStation system). An additional double-sided purification with AMPure Beads XP (0.5×–1× ratio) was performed for the removal of unspecific larger fragments. The final pools were quantified by qPCR (KAPA Library Quantification Kit qPCR Mix for Illumina platforms) and sequenced on a HiseqX lane targeting the equal coverage of 500,000 reads per sample (i.e., capsule) by an external service provider (Macrogen, South Korea).

2.5 | Bioinformatics Processing

Demultiplexed sequencing data was provided based on each index combination. We used flash2 (v2.2; Magoč and Salzberg 2011) to align paired-end reads using default settings except for –max-overlap = 100. Merged reads were filtered for quality using *fastq-filter* from VSEARCH (v2.15; Rognes et al. 2016), and primers were removed using Cutadapt (v4.5; Martin 2011). All subsequent steps were performed using VSEARCH. Singletons and sequences shorter than 80 and longer than 130 bp were removed. Denoising was performed with *cluster_unoise* with unoise-alpha set to 5, and chimera sequences were removed with uchime3_denovo. Sequences were clustered to molecular operational taxonomic units (OTUs) at 99% identity.

For taxonomic assignment, we first constructed a local DNA reference database for the list of vertebrate species present in Iberia (Tables S2 and S3). To retrieve published data, we used the NSDPY (v1.0; Hebert and Meglécz 2022) by searching the NCBI for 'mitochondrion[Title] AND complete[Title] OR mitochondrial[Title] AND complete[Title] OR 12S[Title]' on 23 January 2024. Unpublished sequences from the Iberian ichthyofauna, eulipotyphla, and rodentia were also included to cover database gaps. We then aligned each OTU to the reference database using *blastn* v2.12 with the following settings: perc_identity 80, max_hsp 1, and qcov_hsp_perc 80. Taxonomy was assigned using a lowest common ancestor approach with metabinkit (v0.2.3 Egeter et al. 2022) with the following settings: S 98 (minimum sequence similarity for species-level assignment), G 95 (for genus-level), F 92 (for family-level), A 85 (for above family-level), and TopSpecies 1, to retain all hits to be considered for species-level assignments within 1% of the best score. In addition, we used the option to filter out certain accession numbers as they are likely to be misidentified in NCBI (Table S4). Assignments were checked considering species occurrence in the geographic range and the inability to assign species due to mitochondrial introgression (e.g., all *Lepus* detections were joined as *Lepus* sp.; Melo-Ferreira et al. 2005). To filter out contaminants, OTUs present in negative controls were subtracted from all samples, and to remove spurious detections (i.e., potential false positives), we filtered out all OTUs with a read count below 0.1% of each replicate.

2.6 | Statistical Analysis

Only OTUs identified to species level were considered for analysis in both experiments. In Experiment A, sampling sites with $< 2/3$ of PCR replicates with species detections were removed from the statistical analysis. All statistical tests were performed in R v4.3.2 (R Core Team 2023), and all graphics were produced with ggplot2 (Wickham 2016).

2.6.1 | Experiment A

We built species accumulation curves using iNEXT (v3.0.0; Hsieh et al. 2016) to infer how the number of species detected at the watershed and site scales increased as a function of the number of replicates analysed for high- and low-capacity capsules. Observed and Chao2 estimates were calculated at the watershed

and site scales for both capsule types (using the ChaoRichness function from iNEXT). The number of species detected in each replicate was calculated using phyloseq (v1.46; McMurdie and Holmes 2013) and kept separate. To estimate variation in richness estimates associated with the different replication levels, we used a generalised linear mixed-effects model (GLMM) with Poisson errors and a log link, relating the total number of vertebrate species detected in a PCR replicate to capsule type and observed number of reads (fixed effects), with sampling site, field replicate, and PCR replicate as nested random factors. The observed number of reads was log-transformed and scaled before being included as a fixed effect to control for variation in sequencing depth amongst samples (Martins et al. 2021). Two additional models were built considering only either aquatic or terrestrial species. Semi-aquatic vertebrates, such as the otter, were included under the aquatic category. All models were built using the glmmadmb function from the package glmmADMB (v0.8.5; Fournier et al. 2012; Skaug et al. 2016).

To estimate variation in species composition associated with capsule types and replication levels, we performed nonmetric multidimensional scaling (NMDS) based on Jaccard distances using 1000 iterations ('metaMDS' function in the VEGAN package v2.6-4; Oksanen et al. 2022). We then performed PERMANOVA using the adonis2 function in VEGAN, with 1000 permutations, to estimate the nested contributions to variation in species composition of sampling sites, capsule type within sites, field replicates with each capsule type within sites, and PCR replicates within field replicates within capsule types within sites. Separate analyses were performed for aquatic and terrestrial species.

2.6.2 | Experiment B

To assess the impact of PCR pooling on richness estimates at the watershed scale, we first tallied the number of species detected in either the set of individual PCRs (12) or in the pool at each site, based on one randomly selected high-capacity capsule. Then, we built species accumulation curves as a function of the number of sites and estimated the total observed and Chao2 (calculated using the ChaoRichness function from iNEXT) richness for either the individual replicates or the pool. To analyse the impact of pooling at the site scale, we used a paired *t*-test to compare the site-level richness estimates obtained with either the 12 individual replicates or the pool. Because the differences observed could be due to differences in sequencing coverage, we used a paired t-test to compare across sites the scaled and log-transformed number of reads obtained for the set of 12 PCR replicates and the pool at each site. To evaluate how site-level richness estimates using the pool compared to different numbers of individual replicates, we first computed species accumulation curves per site as a function of the number of PCR replicates. We then built a mean accumulation curve by averaging estimates for each number of PCR replicates across the sampling sites, and we compared it with the mean richness estimates obtained from the pools. Finally, to estimate factors affecting the probability of detecting a species in a pool, we used a GLMM with binomial errors to relate the presence/absence of a taxon in a pool to its prevalence in the corresponding set of 12 PCR replicates, to whether the species was aquatic or terrestrial, and to the scaled and log-transformed number of reads in the pool. The model included the sampling site and species as random effects. Models were built using the glmer function of the lme4 package (v1.1-35; Bates et al. 2014). The GLMM was used to predict detection curves for a given taxon as a function of its prevalence in the corresponding set of PCR replicates, considering a coverage of either 500,000 or 2,000,000 reads.

3 | Results

The water volume filtered per capsule was, as expected, much higher for high-capacity (24.6 ± 9.0 L) than low-capacity capsules (0.9 ± 0.6 L). We obtained a total of 168,221,875 raw PE reads. Sequencing depth averages were similar between field replicates, for both Experiment A (1,154,549 ± 451,014 reads) and Experiment B (pool: 1,062,744 ± 600,427 reads; IndRep: 1,141,708 ± 240,228 reads). After all filtering steps, we retained 26,979,469 reads for Experiment A and 10,961,070 reads for Experiment B.

3.1 | General Diversity and Composition Patterns

We detected 65 and 71 vertebrate species in the Sabor watershed for experiments A and B, respectively. Both experiments detected 60 species in common, comprising 23 aquatic and 37 terrestrial species. These included 9 fish, 11 amphibians, 1 reptile, 18 birds, and 21 mammals. Additionally, in Experiment A we detected 5 terrestrial species (3 birds, 1 reptile, and 1 mammal) that were not detected in Experiment B, while unique to the latter we detected 3 aquatic species (2 fishes and 1 amphibian) and 8 terrestrial species (4 birds and 4 mammals). The most common species detected across experiments were wild/domestic boar *Sus scrofa*, fire salamander *Salamandra salamandra*, Pumpkinseed *Lepomis gibbosus*, wolf/dog *Canis lupus*, cattle *Bos taurus*, Iberian Midwife Toad *Alytes cisternasii*, and Perez's frog *Pelophylax perezi*.

3.2 | Experiment A

Twenty-three sampling sites were included in the analysis, encompassing 275 PCR replicates (46 field replicates of each capsule type). Of these, 49.8% corresponded to high- and 50.2% to low-capacity capsules. At the watershed scale, the observed and estimated (Chao2) richness of vertebrate species was higher using high-capacity (63; Chao2 = 80.9 ± 14.3) than low-capacity (48; Chao2 = 52.5 ± 4.8) capsules (Table 1; Figure 2). Likewise, the mean observed number of species per site was higher when using high-capacity (16.9 ± 0.9) than low-capacity (11.4 ± 0.8) capsules (Table 1). The GLMMs revealed a significant positive effect of capsule capacity on the number of species detected after correcting for the significantly positive effect of sequencing depth (Table 2). Variance associated with random effects was mostly accounted for by differences in richness across sites (38%) and by sampling replicates within sites (39%), while the amount of variance associated with differences between PCR replicates within sampling replicates was negligible (<< 0.1%). In the case of aquatic species, the variance associated with sampling replicates within sites was also negligible (<< 0.1%; Table S5). For

TABLE 1 | Observed and Chao2 estimated (mean ± SE) richness of vertebrate species from water eDNA captured with either high- (HIGH) or low-capacity (LOW) filtration capsules, at the watershed and sampling site scales. Results are provided for the whole vertebrate community, and separately for aquatic and terrestrial species. At site level, the values provided are mean ± SEM (standard error of mean estimates) across sites.

	Observed		Estimated (Chao)	
	LOW	HIGH	LOW	HIGH
Watershed				
All Species	48	63	52.5 ± 4.8	80.9 ± 14.3
Aquatic	21	23	21 ± 0.4	23 ± 0.59
Terrestrial	27	40	33 ± 6.4	55 ± 12.4
Site				
All Species	11.4 ± 0.6	16.9 ± 0.9	15.7 ± 1.0	22.9 ± 1.8
Aquatic	5.7 ± 0.4	7.8 ± 0.4	6.8 ± 0.6	9.4 ± 0.7
Terrestrial	5.7 ± 0.4	9.1 ± 0.8	8.3 ± 0.9	12.4 ± 1.2

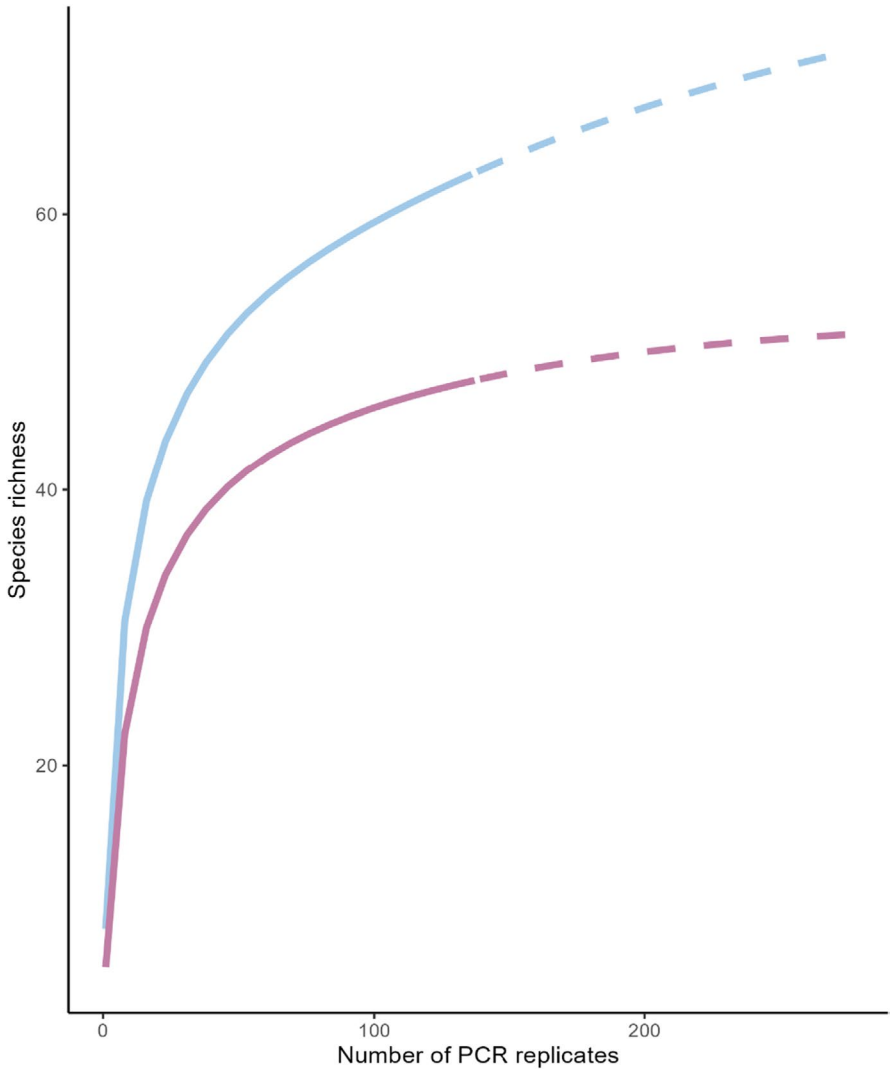

FIGURE 2 | Rarefaction curves showing the observed (solid line) and estimated species richness (dashed line) at the watershed scale, as a function of the number of PCR replicates analysed. Curves were produced using PCR replicates from Experiment A, differentiating high (blue) and low-capacity (pink) capsule types.

terrestrial species, there were higher shares of variance corresponding to differences in richness across sites and amongst field replicates within sites (Table S6).

Regarding vertebrate community composition (Table 3), the PERMANOVA revealed significant effects of sampling site ($R^2 = 0.50$; p-value = < 0.01). In contrast, the nested effects of capsule type, capsule replicates, and PCR replicates were not significant. Accordingly, the proportion of variance (MSS) associated with differences between sampling sites was much higher (72%) than that of any other component (< 15%) (Table 3). Still, there was some variation associated with sampling different capsule types within a site (14%), different replicates of a capsule type in a site (10%), and different PCR replicates of a capsule replicate in a site (5%) (Table 3). In an analysis considering only aquatic species, results were qualitatively similar, but only the effect of sampling site was statistically significant (Table S7). In analysis using only terrestrial species, the amount of variance associated with sampling site decreases (56%), while it increased for capsule type within the site (21%) and different field replicates of a capsule type in a site (14%) (Table S8).

TABLE 2 | Summary results of a Generalised Linear Mixed Model (GLMM) with Poisson errors and a log link, relating the total number of species detected in a given PCR replicate to capsule type (Capsule; HIGH versus LOW capacity, using HIGH as baseline) and the scaled and log-10 transformed number of reads (*Reads*). Random effects included the sampling site (Site), and the nested effects of capsule replicate within site (*CapRep/Site)* and PCR replicate within capsule replicate within site (*PCRRep/CapRep/Site*).

Fixed effect	**Estimate**	**SE**	***Z***	***p***
Intercept	2.0332	0.059	34.03	< 2E−16
Capsule-LOW	−0.3754	0.065	−5.76	8.6E−09
Reads	0.0653	0.029	2.28	0.022
Random effect	**Variance**	**SD**		
Site	0.038	0.1949		
CapRep/Site	0.039	0.1981		
PCRRep/CapRep/Site	1.29E-06	0.001		

3.3 | Experiment B

At the watershed scale, the observed and estimated (Chao2) richness were much higher for the sets of individual replicates (72; Chao2 = 94.2 ± 13.2) versus the PCR pools (46; Chao2 = 60.0 ± 10.4; Figure 3). Likewise, richness estimates per site were significantly higher (paired sample t-test = 11.55, df = 19, p-value < 0.001) for the sets of individual replicates (mean ± SD: 21 ± 6.32) than for the pools (mean ± SD: 9.5 ± 4.55). The total number of species observed per site increased steadily with the number of PCR replicates, without reaching a clear plateau (Figure 4a), while the Chao2 richness increased rapidly up to about 3–4 PCR replicates and then largely levelled off (Figure 4b). Surprisingly, site-level richness estimates obtained from the pool of 12 PCR replicates were largely similar to those obtained from a single PCR replicate. The GLMM indicated that after controlling for the positive effect of sequencing depth, the probability of detecting a species in a pool was positively affected by its prevalence in the 12 individual PCRs and not significantly affected by whether it was aquatic or terrestrial (Table 4). The detection curves predicted from this model indicated that the probability of detecting a species in a pool at a coverage of 500,000 reads is virtually nil when its prevalence is low (< 5 PCR replicates), increasing steadily up to about 8 PCR replicates, and levelling off thereafter. A similar curve was obtained for 2,000,000 reads, though with higher probabilities of detection for very rare species and reaching the plateau of high detection at about 5–6 PCR replicates (Figure 5).

4 | Discussion

Our study reinforces the principle that in eDNA monitoring, "the more, the better," with increased sampling and analysis efforts leading to more accurate estimates of species richness and community composition (Mathieu et al. 2020; Mauvisseau et al. 2019; Stauffer et al. 2021; Zhang et al. 2020). However, our findings also underscore that certain aspects of study design have a more pronounced influence on biodiversity estimates

TABLE 3 | Summary results of PerMANOVA estimating the contribution to overall variation in vertebrate species composition estimated from water eDNA, of the differences amongst sampling sites (*Site*), capsule types within sites (*Capsule/Site*), capsule replicates (*CapRep*) within capsule types within sites, and PCR replicates (*PCRRep*) within capsule replicates within capsule types within sites. For each component we provide the number of degrees of freedom (df), sum of squares (SS), mean squares (MS), the coefficient of determination (R^2), and the F- and p-values.

Coefficient	**df**	**SS**	**MSS**	R^2	***F***	***p***
Site	22	50.126	2.278	0.503	8.932	0.001
Capsule/site	23	10.007	0.435	0.100	1.705	0.085
CapRep/Capsule/Site	46	13.734	0.299	0.138	1.170	0.293
PCRRep/CapRep/Capsule/Site	182	25.547	0.140	0.256	0.550	1
Residual	1	0.255				
Total	274	99.669				

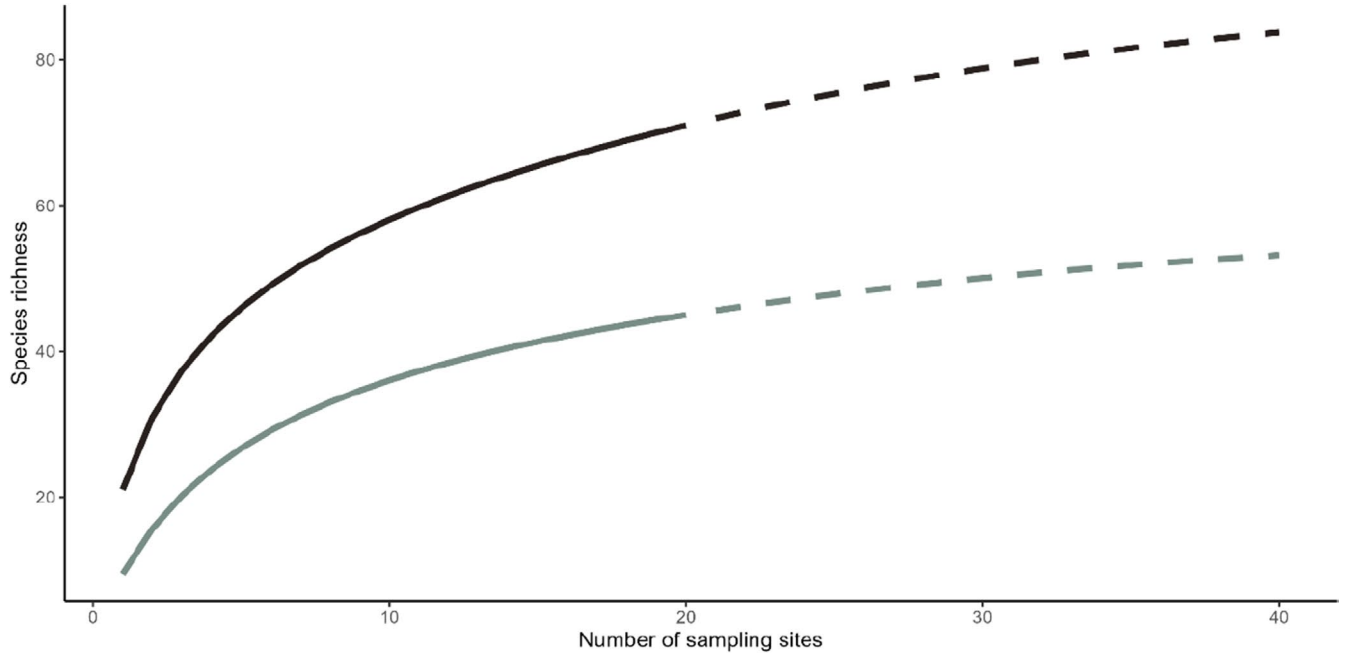

FIGURE 3 | Rarefaction curves showing the observed (solid line) and estimated (dashed line) richness of vertebrate species in the Sabor watershed, as a function of the number of sites sampled. At each site, estimates are based on the design of Experiment B, considering either the species detected in the set of individually indexed replicates (black line) or in the set of pooled replicates (grey line).

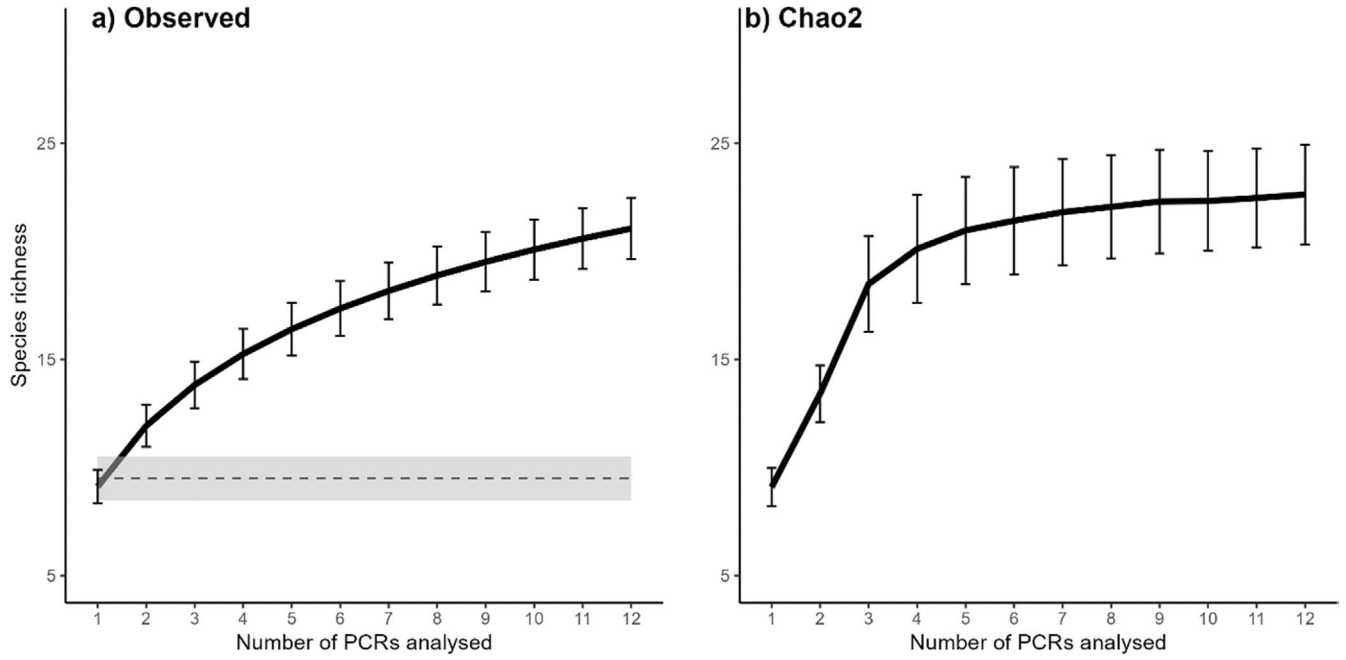

FIGURE 4 | Rarefaction curves for the number of (a) observed and (b) estimated (Chao2) species richness detected per sampling site, when varying the number of PCR replicates. The curves show averages across 20 sampling sites analysed, and error bars are the standard errors of mean estimates. In (a) the average of observed species per sampling site for the pooled replicates is presented as the dashed horizontal line and shading represents the standard errors of mean estimates.

than others. First, we found a very strong effect of capsule capacity, with high-capacity capsules consistently detecting more species at the watershed and site scales than low-capacity capsules. Second, irrespective of capsule type, we found that spatial replication (i.e., number of sampling sites) has a pivotal role in estimating community richness and composition at the watershed scale, significantly more so than any other level of biological or technical replication, particularly for terrestrial species. Third, although less impactful than spatial replication, replication at the site level (i.e., filtering different capsules per site) and conducting multiple PCR replicates per sample also affected biodiversity estimates. Fourth, increasing the number of individual PCR replicates up to at least 12 significantly enhances the number of species detected per sample. Finally, and

perhaps surprisingly, pooling 12 PCR replicates before indexing results in richness estimates largely similar to those obtained from a single PCR replicate, primarily detecting the most common species. These results contribute to identifying the most critical factors affecting biodiversity estimates, offering strategic insights to help prioritise efforts and enhance the accuracy of eDNA monitoring.

Despite some limitations, it is unlikely that our key results and conclusions are affected to any significant extent by artefacts or shortcomings. First, while our study was carried out in a single watershed, it encompassed highly diverse communities as well as a high ecological variability (Ferreira et al. 2016; Mota-Ferreira and Beja 2020), thereby suggesting that our results can be applicable beyond our study system. However, in megadiverse communities, such as the ones found in the tropics, where the dynamics between common and rare species may differ, further testing may still be needed to ensure accurate species detections. Second, our study was based on a single metabarcoding marker, and while it is unlikely that using other markers would significantly affect our results, exploring other markers or targeting different groups could provide valuable insights towards protocol standardisation. Nonetheless, the marker used in this study has been widely employed to survey vertebrate communities from water eDNA (Haderlé et al. 2024; Lopes-Lima et al. 2024; Marques et al. 2021; Reji Chacko et al. 2023). Third, during the bioinformatic analysis, we validated the presence of a species if it represented at least 0.1% of the total PCR replicate reads, which is consistent with other studies (García-Machado et al. 2023), including for highly diverse fish communities (Doble et al. 2020). It is possible that using a lower threshold would have allowed the detection of more species in the pool, but this approach may also have an impact on the estimates derived from the individually indexed PCR replicates. Moreover, reducing the threshold could have undesired consequences by increasing the number of false positives.

TABLE 4 | Summary results of a generalised linear mixed model (GLMM) with binomial errors and logit link, relating the probability of a taxa occurring in a pool to its frequency of occurrence in the 12 individual PCR replicates (*Prevalence*), its categorisation as terrestrial or aquatic, and the scaled and log-10 transformed number of reads of the pool. Sampling site (*Site*) and species (Species) were included as random effects.

Fixed effect	**Estimate**	**SE**	***Z***	***p***
Intercept	−10.189	2.1421	−4.757	1.97E−06
Prevalence	1.767	0.3461	5.107	3.28E−07
Category-Terrestrial	1.778	1.1867	1.498	0.1341
Reads	2.0262	0.5597	3.620	0.00029
Random Effect	**Variance**	**SD**		
Site	2.779	1.667		
Species	13.046	3.612		

A key result of our study is the strong impact of capsule capacity on vertebrate community estimates, with about 30% more species detected with high-capacity than with low-capacity capsules, at both the watershed and the site scales. This is probably related to the much larger volume of water filtered by high-capacity capsules, which was about 25 times higher than that filtered by low-capacity capsules. This result is in line with previous studies

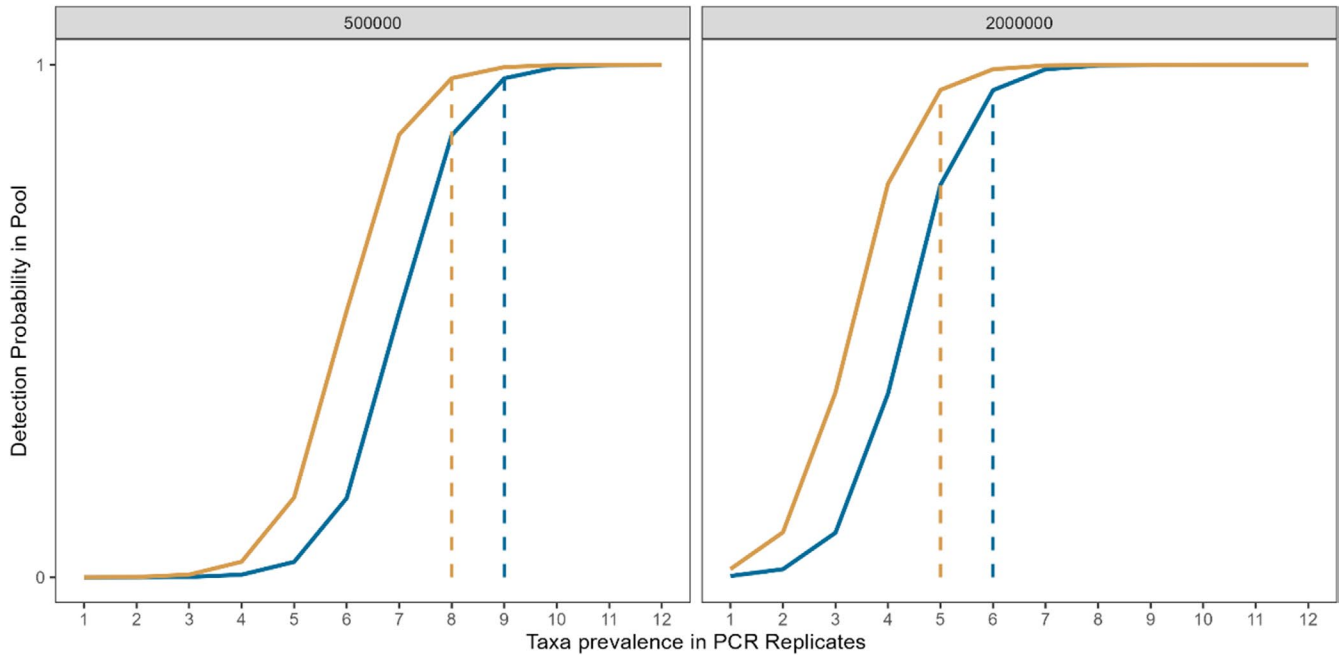

FIGURE 5 | Detection curves predicted from a GLMM relating the probabilities of detecting a given species in a pool as a function of the number of PCR replicates where it occurs, for aquatic (blue) and terrestrial (yellow) species. Detection curves are provided for two sequencing depths: 500,000 reads (left panel) and 2,000,000 reads (right panel). Dashed lines indicate the minimum number of PCR replicates needed to obtain a detection probability in the pool higher than 90%.

showing that species detection probabilities greatly increase with the volume of water filtered, with small volumes only detecting the most abundant species (Bessey et al. 2020; Blackman et al. 2024; Peixoto et al. 2023). In fact, increasing the volume of water filtered is needed to increase the likelihood of capturing the DNA from rare species, which typically occur at low concentrations in the water (Cantera et al. 2019; Sepulveda et al. 2019). Exploring the relationship between the number of species detected and intermediate water volumes may provide a more detailed understanding of the sampling effort required to optimise species detection, as our study only evaluated two distinctly different filtration capacities. It is worth noting, however, that our low-capacity capsules filtered volumes (~1 L) were in the range of recommended volumes (1–5 L; Bruce et al. 2021) and were much higher than those used in earlier studies (e.g., < 100 mL; Deiner et al. 2016), which may be inadequate for accurate eDNA biodiversity surveys. Clearly, large water volumes comparable to those filtered by our high-capacity capsules (~25 L) are required to minimise false absences (Cantera et al. 2019; Govindarajan et al. 2022; Stauffer et al. 2021), though even larger volumes may be needed where DNA occurs at very low concentrations, such as terrestrial species, or to detect very rare species (Schabacker et al. 2020; Sepulveda et al. 2019).

Our results suggest that spatial replication can play an important role in estimating vertebrate species richness at the watershed scale, particularly in an ecologically diverse system such as the Sabor watershed. We observed a strong increase in the number of species detected when increasing the number of sampling sites, which was related to the strong variation in the set of species detected by eDNA across sites. This pattern likely reflects the heterogeneous distribution of species across the Sabor watershed, as shown by conventional methods (Ferreira et al. 2016; Mota-Ferreira and Beja 2020; Quaglietta et al. 2018). These results may seem surprising under the "conveyor belt" hypothesis (Deiner et al. 2016), which suggests that sampling eDNA at a single river point along the river network would be sufficient to provide integrative information on aquatic and terrestrial biodiversity in the upstream catchment. However, our results are consistent with recent studies showing that eDNA can travel up to just a few kilometres (Cantera et al. 2022; Troth et al. 2021; Van Driessche et al. 2023) and that eDNA transport can be affected by river topography, substrate, and other abiotic factors leading to a patchy distribution (Jerde et al. 2016; Sgarlatta et al. 2023; Stewart 2019). This suggests that each site may have a different pool of species eDNA, thereby requiring multiple sampling sites to obtain a robust estimate of richness and composition at the watershed scale. Imperfect detection of rare species may also play a role in the observed pattern, as increasing the number of sites sampled also increases the volume of water analysed, and thus the likelihood of detecting species with low DNA concentrations in the water. These two processes likely explain why the effect of sampling multiple sites was particularly strong for terrestrial species, as their DNA is likely to be present at lower concentrations and have a more patchy distribution than that of aquatic species (Harper et al. 2019; Lyet et al. 2021; Macher et al. 2021). This is because the detection of terrestrial species (birds and mammals) from aquatic eDNA requires individuals to interact with the river water, which may be a relatively rare event for many species, even if they are abundant in the watershed (Sales et al. 2020).

In contrast to some previous studies (Cantera et al. 2019; Stauffer et al. 2021), we found that local biological replication had low to moderate effects on species richness and composition estimates, particularly for aquatic species. In general, high variation between local replicates is expected when DNA is at low concentration and distributed heterogeneously, there are many rare species, and sampling effort per replicate is low (Cantera et al. 2019; Stauffer et al. 2021). In our study, this probably did not happen for aquatic species because the Sabor watershed was characterised by local fish and amphibian communities with relatively low richness and high abundance (Ferreira et al. 2016; Mota-Ferreira and Beja 2020), which likely resulted in high concentrations of DNA well-mixed in the water column. Moreover, there were a few rare aquatic species (but see Quaglietta et al. 2018), and sampling effort was high, particularly for high-capacity capsules. In contrast, there was some compositional difference in terrestrial vertebrate species between local replicates, probably because their DNA occurred at low concentration and was patchily distributed in the river section sampled.

Our results also supported the importance of PCR replication to increase the accuracy of biodiversity surveys (Alberdi et al. 2018; Blackman et al. 2024; Vourka et al. 2023). Although the number of species detected was largely similar across replicates, there were some compositional differences, which explain the steady increment in the number of species detected with the number of replicates analysed. Actually, the accumulation curves suggested that the number of species detected continues to increase up to 12 PCR replicates and beyond, which is a number far higher than the three PCR replicates used in many DNA metabarcoding studies (Alberdi et al. 2018; Doi et al. 2019; Ficetola et al. 2015; Ushio et al. 2022). This is likely related to the inherent stochasticity of the PCR process, which causes rare species (i.e., with low abundance of DNA template) to occur in some replicates but not in others (Alberdi et al. 2018; Shirazi et al. 2021; Vourka et al. 2023). The use of small PCR reaction volumes, as in our study, could also contribute to this stochasticity by increasing the sensitivity of amplification conditions to small variations in template DNA concentration. However, similar patterns have been reported in studies using larger PCR volumes (Shirazi et al. 2021). Because of this, species with a very low abundance of DNA template have a very small likelihood of being amplified in any given PCR replicate, thereby requiring a very large number of replicates to increase their probabilities of detection in a sample. In our study, this problem appeared to affect much more the terrestrial than the aquatic species, likely due to their very low DNA concentration in the water.

A notable finding of our study is the limited performance of pooling PCR replicates, with pools detecting almost as many species as a single PCR replicate. This is potentially problematic, as pooling of PCR replicates prior to indexing is used to reduce the cost of library preparation (de Muinck et al. 2017). These results are consistent with a study of dietary metabarcoding, which found that pooling bat pellets provided estimates of the number of prey species ingested similar to those obtained by analysing a single pellet (Mata et al. 2019). The reasons for these patterns are not entirely clear, but it appears that pooling concentrates DNA from common species that occur in most PCR replicates while further diluting DNA from rare taxa that occur in only a few replicates, thus reducing the likelihood of indexing the latter. This was

supported by our model of the probability of species detection in pools, which showed that this probability is very low for species that occur in only a few PCR replicates. The model further suggests that this problem can be compensated for to some extent by greatly increasing sequencing depth, although taxa occurring in very few replicates remain unlikely to be detected. Nevertheless, the effect of pooling different numbers of PCR replicates should be further evaluated to assess whether pools created from fewer PCR replicates or using alternative methodological approaches can mitigate the observed differences.

Our findings have significant implications for designing eDNA biodiversity surveys in river systems and potentially other ecosystems. Ideally, with unlimited resources, studies could maximise efforts in all aspects of sampling and analysis, including filtering the largest possible volumes of water, sampling multiple sites, taking multiple replicates per site, and analysing multiple PCR replicates independently (Govindarajan et al. 2022; Lyet et al. 2021; Stauffer et al. 2021). However, in reality, researchers must carefully balance various methodological trade-offs against the study's objectives and constraints, such as human, financial, and logistical limitations. Based on our results, we recommend prioritising the use of high-capacity filtration capsules capable of processing large water volumes (30 L or more), without which only the most common species are consistently retrieved (Altermatt et al. 2023; Cantera et al. 2019; Lopes et al. 2017). This appears to be particularly important for detecting terrestrial species, where higher capture efforts are needed due to low DNA concentrations (Harper et al. 2019; Kamoroff et al. 2020; Lyet et al. 2021). Although high-capacity capsules were more expensive (~€80) than low-capacity capsules (~€4.5), more affordable equivalent high-capacity capsules are (~lt; €20), which may improve accessibility with further testing. Additionally, critical to produce accurate biodiversity estimates at the regional scale is the sampling of multiple sites, encompassing the variety of habitats, and accounting for the patchy distribution of aquatic eDNA from different species (Feng et al. 2020; Troth et al. 2021). Again, this seems to be particularly important when surveying terrestrial species, as the DNA of aquatic species occurs at higher concentration and thus is probably detectable farther from their sources along the river network (Macher et al. 2021; Sales et al. 2020). Site-level replication appeared less influential and can be deprioritised if resources are limited, especially when filtering large volumes of water. Multiple replicates per site might, however, be needed for studies focusing on local (versus regional) estimates of biodiversity or employing occupancy models to assess species detectability (Dorazio and Erickson 2018; Peixoto et al. 2023; Pont et al. 2018). Conducting multiple PCR replicates is important for detecting rare species within samples, but the number of replicates should be balanced against the need to increase the number of sites sampled and site replication. Importantly, pooling PCR replicates before indexing should be avoided, as it significantly diminishes the ability to detect rare species, even if sequencing coverage is greatly increased.

Overall, the guidelines emerging from our study should assist researchers in designing effective eDNA biodiversity studies within the constraints of available resources. However, we emphasise that researchers should bear in mind the "the more, the better" principle and encourage careful allocation of resources across different study design components, aligning efforts with the specific monitoring objectives. These considerations should include the target scale of the survey (e.g., watershed versus local biodiversity), the types of species targeted (e.g., terrestrial versus aquatic), and their expected abundance (e.g., common versus rare). Furthermore, additional research is required to evaluate the generalizability of our findings across various ecosystems. Understanding how to balance trade-offs amongst different aspects of eDNA sampling and analysis in diverse environments will be crucial for refining these guidelines and enhancing the effectiveness of eDNA biodiversity monitoring in different contexts.

Author Contributions

All authors contributed to the research experiment, and they reviewed and approved the final manuscript. The research was designed by J.V., F.M.S.M., M.L.-L., and P.B. The sampling was conducted M.L.-L., F.A., V.F., M.K., and A.T. The laboratory was conducted by C.C. and J.V. The bioinformatic analyses and data analysis were conducted by J.V., F.M.S.M., M.L.-L., and P.B. The manuscript was drafted by J.V. and P.B., with contributions from all coauthors.

Acknowledgements

This work was supported by National Funds through FCT—Fundação para a Ciência e a Tecnologia in the scope of the project UIDP/50027/2020 and LA/P/0048/2020. FCT also supported J.V. through Ph.D. grants SFRH/BD/133159/2017 and COVID/BD/152600/2022, and M.L.L. through the researcher contract 2020.03608.CEECIND. F.M.S.M. was supported by BIOPOLIS CCDRN: Work co-funded by the project NORTE-01-0246-FEDER-000063, supported by Norte Portugal Region Operational Programme (NORTE2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All sequencing raw data is available on NCBI under BioProject PRJNA1232667. All metadata needed to replicate the analyses reported in this study (scripts and metadata) are available at https://doi.org/10.5281/zenodo.15002675.

Benefit-Sharing Statement

Benefits from this research accrue from the sharing of our data and results on public databases as described above, as well, as guidelines for water eDNA studies.

References

Agersnap, S., E. E. Sigsgaard, M. R. Jensen, et al. 2022. "A National Scale "BioBlitz" Using Citizen Science and eDNA Metabarcoding for Monitoring Coastal Marine Fish." *Frontiers in Marine Science* 9: 24100. https://doi.org/10.3389/fmars.2022.824100.

Alberdi, A., O. Aizpurua, M. T. P. Gilbert, and K. Bohmann. 2018. "Scrutinizing Key Steps for Reliable Metabarcoding of Environmental Samples." *Methods in Ecology and Evolution* 9, no. 1: 134–147. https://doi.org/10.1111/2041-210X.12849.

Altermatt, F., L. Carraro, M. Antonetti, et al. 2023. "Quantifying Biodiversity Using eDNA From Water Bodies: General Principles and Recommendations for Sampling Designs." *Environmental DNA* 5, no. 4: 671–682. https://doi.org/10.1002/edn3.430.

Ando, H., H. Mukai, T. Komura, T. Dewi, M. Ando, and Y. Isagi. 2020. "Methodological Trends and Perspectives of Animal Dietary Studies by Noninvasive Fecal DNA Metabarcoding." *Environmental DNA* 2, no. 4: 391–406. https://doi.org/10.1002/edn3.117.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2014. "Fitting Linear Mixed-Effects Models Using Lme4." *arXiv*:1406. https://doi.org/10.18637/jss.v067.i01.

Beentjes, K. K., A. G. C. L. Speksnijder, M. Schilthuizen, M. Hoogeveen, and B. B. van der Hoorn. 2019. "The Effects of Spatial and Temporal Replicate Sampling on eDNA Metabarcoding." *PeerJ* 7: e7335. https://doi.org/10.7717/peerj.7335.

Bessey, C., S. N. Jarman, O. Berry, et al. 2020. "Maximizing Fish Detection With eDNA Metabarcoding." *Environmental DNA* 2, no. 4: 493–504. https://doi.org/10.1002/edn3.74.

Blackman, R., M. Couton, F. Keck, et al. 2024. "Environmental DNA: The Next Chapter." *Molecular Ecology* 33, no. 11: e17355. https://doi.org/10.1111/mec.17355.

Bohmann, K., P. Chua, L. E. Holman, and C. Lynggaard. 2021. "DNAqua-Net Conference Unites Participants From Around the World With the Quest to Standardize and Implement DNA-Based Aquatic Biomonitoring." *Environmental DNA* 3: 884–888. https://doi.org/10.1002/edn3.207.

Bruce, K., R. Blackman, S. J. Bourlat, et al. 2021. "A Practical Guide to DNA-Based Methods for Biodiversity Assessment." *Advanced Books* 1: e68634. https://doi.org/10.3897/ab.e68634.

Burian, A., K. Bruce, E. Tovela, et al. 2023. "Merging Two eDNA Metabarcoding Approaches and Citizen-Science-Based Sampling to Facilitate Fish Community Monitoring Along Vast Sub-Saharan Coastlines." *Molecular Ecology Resources* 23, no. 7: 1641–1655. https://doi.org/10.1111/1755-0998.13839.

Buxton, A., E. Matechou, J. Griffin, A. Diana, and R. A. Griffiths. 2021. "Optimising Sampling and Analysis Protocols in Environmental DNA Studies." *Scientific Reports* 11, no. 1: 11637. https://doi.org/10.1038/s41598-021-91166-7.

Cantera, I., K. Cilleros, A. Valentini, et al. 2019. "Optimizing Environmental DNA Sampling Effort for Fish Inventories in Tropical Streams and Rivers." *Scientific Reports* 9, no. 1: 3085. https://doi.org/10.1038/s41598-019-39399-5.

Cantera, I., J.-B. Decotte, T. Dejean, et al. 2022. "Characterizing the Spatial Signal of Environmental DNA in River Systems Using a Community Ecology Approach." *Molecular Ecology Resources* 22, no. 4: 1274–1283. https://doi.org/10.1111/1755-0998.13544.

CBD. 2022. "Kunming-Montreal Global Biodiversity Framework." Draft decision submitted by the President. CBD/COP/15/L.25. https://www.cbd.int/doc/c/e6d3/cd1d/daf663719a03902a9b116c34/cop-15-l-25-en.pdf.

Clarke, S. J., E. Long, J. Biggs, et al. 2023. "Co-Design of a Citizen Science Study: Unlocking the Potential of eDNA for Volunteer Freshwater Monitoring." *Ecological Solutions and Evidence* 4, no. 3: e12273. https://doi.org/10.1002/2688-8319.12273.

Coghlan, S. A., A. B. A. Shafer, and J. R. Freeland. 2021. "Development of an Environmental DNA Metabarcoding Assay for Aquatic Vascular Plant Communities." *Environmental DNA* 3, no. 2: 372–387. https://doi.org/10.1002/edn3.120.

de Muinck, E. J., P. Trosvik, G. D. Gilfillan, J. R. Hov, and A. Y. M. Sundaram. 2017. "A Novel Ultra High-Throughput 16S rRNA Gene Amplicon Sequencing Library Preparation Method for the Illumina HiSeq Platform." *Microbiome* 5, no. 1: 68. https://doi.org/10.1186/s40168-017-0279-1.

Deiner, K., E. A. Fronhofer, E. Mächler, J.-C. Walser, and F. Altermatt. 2016. "Environmental DNA Reveals That Rivers Are Conveyer Belts of Biodiversity Information." *Nature Communications* 7, no. 1: 12544. https://doi.org/10.1038/ncomms12544.

Doble, C. J., H. Hipperson, W. Salzburger, et al. 2020. "Testing the Performance of Environmental DNA Metabarcoding for Surveying Highly Diverse Tropical Fish Communities: A Case Study From Lake Tanganyika." *Environmental DNA* 2, no. 1: 24–41. https://doi.org/10.1002/edn3.43.

Doi, H., K. Fukaya, S. Oka, K. Sato, M. Kondoh, and M. Miya. 2019. "Evaluation of Detection Probabilities at the Water-Filtering and Initial PCR Steps in Environmental DNA Metabarcoding Using a Multispecies Site Occupancy Model." *Scientific Reports* 9, no. 1: 3581. https://doi.org/10.1038/s41598-019-40233-1.

Dorazio, R. M., and R. A. Erickson. 2018. "Ednaoccupancy: An r Package for Multiscale Occupancy Modelling of Environmental DNA Data." *Molecular Ecology Resources* 18, no. 2: 368–380. https://doi.org/10.1111/1755-0998.12735.

Duarte, S., L. Simões, and F. O. Costa. 2023. "Current Status and Topical Issues on the Use of eDNA-Based Targeted Detection of Rare Animal Species." *Science of the Total Environment* 904: 166675. https://doi.org/10.1016/j.scitotenv.2023.166675.

Egeter, B., J. Veríssimo, M. Lopes-Lima, et al. 2022. "Speeding Up the Detection of Invasive Bivalve Species Using Environmental DNA: A Nanopore and Illumina Sequencing Comparison." *Molecular Ecology Resources* 22, no. 6: 2232–2247. https://doi.org/10.1111/1755-0998.13610.

Elbrecht, V., E. E. Vamos, K. Meissner, J. Aroviita, and F. Leese. 2017. "Assessing Strengths and Weaknesses of DNA Metabarcoding-Based Macroinvertebrate Identification for Routine Stream Monitoring." *Methods in Ecology and Evolution* 8, no. 10: 1265–1275. https://doi.org/10.1111/2041-210X.12789.

Fadrosh, D. W., B. Ma, P. Gajer, et al. 2014. "An Improved Dual-Indexing Approach for Multiplexed 16S rRNA Gene Sequencing on the Illumina MiSeq Platform." *Microbiome* 2: 6. https://doi.org/10.1186/2049-2618-2-6.

Fediajevaite, J., V. Priestley, R. Arnold, and V. Savolainen. 2021. "Meta-Analysis Shows That Environmental DNA Outperforms Traditional Surveys, but Warrants Better Reporting Standards." *Ecology and Evolution* 11, no. 9: 4803–4815. https://doi.org/10.1002/ece3.7382.

Feng, W., G. Bulté, and S. C. Lougheed. 2020. "Environmental DNA Surveys Help to Identify Winter Hibernacula of a Temperate Freshwater Turtle." *Environmental DNA* 2: 200–209. https://doi.org/10.1002/edn3.58.

Ferreira, M., A. F. Filipe, D. C. Bardos, M. F. Magalhães, and P. Beja. 2016. "Modeling Stream Fish Distributions Using Interval-Censored Detection Times." *Ecology and Evolution* 6, no. 15: 5530–5541. https://doi.org/10.1002/ece3.2295.

Ficetola, G. F., J. Pansu, A. Bonin, et al. 2015. "Replication Levels, False Presences and the Estimation of the Presence/Absence From eDNA Metabarcoding Data." *Molecular Ecology Resources* 15, no. 3: 543–556. https://doi.org/10.1111/1755-0998.12338.

Ficetola, G. F., P. Taberlet, and E. Coissac. 2016. "How to Limit False Positives in Environmental DNA and Metabarcoding?" *Molecular Ecology Resources* 16, no. 3: 604–607. https://doi.org/10.1111/1755-0998.12508.

Fonseca, V. G., P. I. Davison, V. Creach, D. Stone, D. Bass, and H. J. Tidbury. 2023. "The Application of eDNA for Monitoring Aquatic Non-Indigenous Species: Practical and Policy Considerations." *Diversity (Basel)* 15, no. 5: 631. https://doi.org/10.3390/d15050631.

Fournier, D. A., H. J. Skaug, J. Ancheta, et al. 2012. "AD Model Builder: Using Automatic Differentiation for Statistical Inference of Highly Parameterized Complex Nonlinear Models." *Optimization Methods and Software* 27, no. 2: 233–249. https://doi.org/10.1080/10556788.2011.597854.

García-Machado, E., E. Normandeau, G. Côté, and L. Bernatchez. 2023. "How eDNA Data Filtration, Sequence Coverage, and Primer Selection Influence Assessment of Fish Communities in Northern Temperate

Lakes." *Environmental DNA* 5, no. 6: 1216–1233. https://doi.org/10.1002/edn3.444.

Govindarajan, A. F., L. McCartin, A. Adams, et al. 2022. "Improved Biodiversity Detection Using a Large-Volume Environmental DNA Sampler With In Situ Filtration and Implications for Marine eDNA Sampling Strategies." *Deep Sea Research Part I: Oceanographic Research Papers* 189: 103871. https://doi.org/10.1016/j.dsr.2022.103871.

Haderlé, R., L. Bouveret, J. Chazal, et al. 2024. "eDNA-Based Survey of the Marine Vertebrate Biodiversity off the West Coast of Guadeloupe (French West Indies)." *Biodiversity Data Journal* 12: e125348. https://doi.org/10.3897/BDJ.12.e125348.

Harper, L. R., A. S. Buxton, H. C. Rees, et al. 2019. "Prospects and Challenges of Environmental DNA (eDNA) Monitoring in Freshwater Ponds." *Hydrobiologia* 826, no. 1: 25–41. https://doi.org/10.1007/s10750-018-3750-5.

Hebert, R., and E. Meglécz. 2022. "NSDPY: A Python Package to Download DNA Sequences From NCBI." *SoftwareX* 18: 101038. https://doi.org/10.1016/j.softx.2022.101038.

Hsieh, T. C., K. H. Ma, and A. Chao. 2016. "iNEXT: An R Package for Rarefaction and Extrapolation of Species Diversity (Hill Numbers)." *Methods in Ecology and Evolution* 7, no. 12: 1451–1456. https://doi.org/10.1111/2041-210X.12613.

Jerde, C. L., B. P. Olds, A. J. Shogren, et al. 2016. "Influence of Stream Bottom Substrate on Retention and Transport of Vertebrate Environmental DNA." *Environmental Science & Technology* 50, no. 16: 8770–8779. https://doi.org/10.1021/acs.est.6b01761.

Kamoroff, C., N. Daniele, R. L. Grasso, R. Rising, T. Espinoza, and C. S. Goldberg. 2020. "Effective Removal of the American Bullfrog (*Lithobates catesbeianus*) on a Landscape Level: Long Term Monitoring and Removal Efforts in Yosemite Valley, Yosemite National Park." *Biological Invasions* 22, no. 2: 617–626. https://doi.org/10.1007/s10530-019-02116-4.

Keck, F., R. C. Blackman, R. Bossart, et al. 2022. "Meta-Analysis Shows Both Congruence and Complementarity of DNA and eDNA Metabarcoding to Traditional Methods for Biological Community Assessment." *Molecular Ecology* 31, no. 6: 1820–1835. https://doi.org/10.1111/mec.16364.

Lopes, C. M., T. Sasso, A. Valentini, et al. 2017. "eDNA Metabarcoding: A Promising Method for Anuran Surveys in Highly Diverse Tropical Forests." *Molecular Ecology Resources* 17, no. 5: 904–914. https://doi.org/10.1111/1755-0998.12643.

Lopes-Lima, M., V. Prié, M. Camará, et al. 2024. "Rapid eDNA Survey Reveals a Unique Biodiversity Hotspot: The Corubal River, West Africa." *Bioscience* 74, no. 6: biae036. https://doi.org/10.1093/biosci/biae036.

Lyet, A., L. Pellissier, A. Valentini, T. Dejean, A. Hehmeyer, and R. Naidoo. 2021. "eDNA Sampled From Stream Networks Correlates With Camera Trap Detection Rates of Terrestrial Mammals." *Scientific Reports* 11, no. 1: 11362. https://doi.org/10.1038/s41598-021-90598-5.

Macher, T.-H., R. Schütz, J. Arle, A. J. Beermann, J. Koschorreck, and F. Leese. 2021. "Beyond Fish eDNA Metabarcoding: Field Replicates Disproportionately Improve the Detection of Stream Associated Vertebrate Species." *Metabarcoding and Metagenomics* 5: e66557. https://doi.org/10.3897/mbmg.5.66557.

Magoč, T., and S. L. Salzberg. 2011. "FLASH: Fast Length Adjustment of Short Reads to Improve Genome Assemblies." *Bioinformatics* 27, no. 21: 2957–2963. https://doi.org/10.1093/bioinformatics/btr507.

Marques, V., P. Castagné, A. Polanco Fernández, et al. 2021. "Use of Environmental DNA in Assessment of Fish Functional and Phylogenetic Diversity." *Conservation Biology* 35, no. 6: 1944–1956. https://doi.org/10.1111/cobi.13802.

Martin, M. 2011. "Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads." *EMBnet.Journal* 17, no. 1: 10–12. https://doi.org/10.14806/ej.17.1.200.

Martins, F. M. S., M. Porto, M. J. Feio, et al. 2021. "Modelling Technical and Biological Biases in Macroinvertebrate Community Assessment From Bulk Preservative Using Multiple Metabarcoding Markers." *Molecular Ecology* 30, no. 13: 3221–3238. https://doi.org/10.1111/mec.15620.

Mata, V. A., H. Rebelo, F. Amorim, G. F. McCracken, S. Jarman, and P. Beja. 2019. "How Much Is Enough? Effects of Technical and Biological Replication on Metabarcoding Dietary Analysis." *Molecular Ecology* 28, no. 2: 165–175. https://doi.org/10.1111/mec.14779.

Mathieu, C., S. M. Hermans, G. Lear, T. R. Buckley, K. C. Lee, and H. L. Buckley. 2020. "A Systematic Review of Sources of Variability and Uncertainty in eDNA Data for Environmental Monitoring." *Frontiers in Ecology and Evolution* 8: 135.

Mauvisseau, Q., A. Burian, C. Gibson, R. Brys, A. Ramsey, and M. Sweet. 2019. "Influence of Accuracy, Repeatability and Detection Probability in the Reliability of Species-Specific eDNA Based Approaches." *Scientific Reports* 9, no. 1: 580. https://doi.org/10.1038/s41598-018-37001-y.

McMurdie, P. J., and S. Holmes. 2013. "Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data." *PLoS One* 8, no. 4: e61217. https://doi.org/10.1371/journal.pone.0061217.

Melo-Ferreira, J., P. Boursot, F. Suchentrunk, N. Ferrand, and P. C. Alves. 2005. "Invasion From the Cold Past: Extensive Introgression of Mountain Hare (*Lepus timidus*) Mitochondrial DNA Into Three Other Hare Species in Northern Iberia." *Molecular Ecology* 14, no. 8: 2459–2464. https://doi.org/10.1111/j.1365-294X.2005.02599.x.

Moersberger, H., J. Valdez, J. G. C. Martin, et al. 2024. "Biodiversity Monitoring in Europe: User and Policy Needs." *Conservation Letters* 17, no. 5: e13038. https://doi.org/10.1111/conl.13038.

Mota-Ferreira, M., and P. Beja. 2020. "Combining Geostatistical and Biotic Interaction Model to Predict Amphibian Refuges Under Crayfish Invasion Across Dendritic Stream Networks." *Diversity and Distributions* 26, no. 6: 699–714. https://doi.org/10.1111/ddi.13047.

Neice, A. A., and S. B. McRae. 2021. "An eDNA Diagnostic Test to Detect a Rare, Secretive Marsh Bird." *Global Ecology and Conservation* 27: e01529. https://doi.org/10.1016/j.gecco.2021.e01529.

Oksanen, J., G. L. Simpson, F. G. Blanchet, et al. 2022. "vegan: Community Ecology Package." https://CRAN.R-project.org/package=vegan.

Peixoto, S., M. Mota-Ferreira, C. Chaves, G. Velo-Antón, P. Beja, and B. Egeter. 2023. "Multi-Species Occupancy Modeling Reveals Methodological and Environmental Effects on eDNA Detection of Amphibians in Temporary Ponds." *Environmental DNA* 5, no. 4: 796–811. https://doi.org/10.1002/edn3.443.

Pereira, C. L., M. T. P. Gilbert, M. B. Araújo, and M. G. Matias. 2021. "Fine-Tuning Biodiversity Assessments: A Framework to Pair eDNA Metabarcoding and Morphological Approaches." *Methods in Ecology and Evolution* 12, no. 12: 2397–2409. https://doi.org/10.1111/2041-210X.13718.

Pereira, H. M., J. Belnap, N. Brummitt, et al. 2010. "Global Biodiversity Monitoring." *Frontiers in Ecology and the Environment* 8, no. 9: 459–460. https://doi.org/10.1890/10.WB.23.

Piggott, M. P. 2016. "Evaluating the Effects of Laboratory Protocols on eDNA Detection Probability for an Endangered Freshwater Fish." *Ecology and Evolution* 6, no. 9: 2739–2750. https://doi.org/10.1002/ece3.2083.

Polanco, A., C. Waldock, T. Keggin, et al. 2022. "Ecological Indices From Environmental DNA to Contrast Coastal Reefs Under Different Anthropogenic Pressures." *Ecology and Evolution* 12, no. 8: e9212. https://doi.org/10.1002/ece3.9212.

Polanco Fernández, A., V. Marques, F. Fopp, et al. 2021. "Comparing Environmental DNA Metabarcoding and Underwater Visual Census to Monitor Tropical Reef Fishes." *Environmental DNA* 3, no. 1: 142–156. https://doi.org/10.1002/edn3.140.

Pont, D., M. Rocle, A. Valentini, et al. 2018. "Environmental DNA Reveals Quantitative Patterns of Fish Biodiversity in Large Rivers Despite Its Downstream Transportation." *Scientific Reports* 8, no. 1: 10361. https://doi.org/10.1038/s41598-018-28424-8.

Quaglietta, L., J. Paupério, F. M. S. Martins, P. C. Alves, and P. Beja. 2018. "Recent Range Contractions in the Globally Threatened Pyrenean Desman Highlight the Importance of Stream Headwater Refugia." *Animal Conservation* 21, no. 6: 515–525. https://doi.org/10.1111/acv.12422.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Reji Chacko, M., F. Altermatt, F. Fopp, et al. 2023. "Catchment-Based Sampling of River eDNA Integrates Terrestrial and Aquatic Biodiversity of Alpine Landscapes." *Oecologia* 202, no. 4: 699–713. https://doi.org/10.1007/s00442-023-05428-4.

Riaz, T., W. Shehzad, A. Viari, F. Pompanon, P. Taberlet, and E. Coissac. 2011. "ecoPrimers: Inference of New DNA Barcode Markers From Whole Genome Sequence Analysis." *Nucleic Acids Research* 39, no. 21: e145. https://doi.org/10.1093/nar/gkr732.

Rognes, T., T. Flouri, B. Nichols, C. Quince, and F. Mahé. 2016. "VSEARCH: A Versatile Open Source Tool for Metagenomics." *PeerJ* 4: e2584. https://doi.org/10.7717/peerj.2584.

Ruppert, K. M., R. J. Kline, and M. S. Rahman. 2019. "Past, Present, and Future Perspectives of Environmental DNA (eDNA) Metabarcoding: A Systematic Review in Methods, Monitoring, and Applications of Global eDNA." *Global Ecology and Conservation* 17: e00547. https://doi.org/10.1016/j.gecco.2019.e00547.

Sakata, M. K., M. U. Kawata, A. Kurabayashi, et al. 2022. "Development and Evaluation of PCR Primers for Environmental DNA (eDNA) Metabarcoding of Amphibia." *Metabarcoding and Metagenomics* 6: e76534. https://doi.org/10.3897/mbmg.6.76534.

Sales, N. G., M. B. McKenzie, J. Drake, et al. 2020. "Fishing for Mammals: Landscape-Level Monitoring of Terrestrial and Semi-Aquatic Communities Using eDNA From Riverine Systems." *Journal of Applied Ecology* 57, no. 4: 707–716. https://doi.org/10.1111/1365-2664.13592.

Sato, H., Y. Sogo, H. Doi, and H. Yamanaka. 2017. "Usefulness and Limitations of Sample Pooling for Environmental DNA Metabarcoding of Freshwater Fish Communities." *Scientific Reports* 7, no. 1: 14860. https://doi.org/10.1038/s41598-017-14978-6.

Schabacker, J. C., S. J. Amish, B. K. Ellis, et al. 2020. "Increased eDNA Detection Sensitivity Using a Novel High-Volume Water Sampling Method." *Environmental DNA* 2, no. 2: 244–251. https://doi.org/10.1002/edn3.63.

Schenekar, T. 2023. "The Current State of eDNA Research in Freshwater Ecosystems: Are We Shifting From the Developmental Phase to Standard Application in Biomonitoring?" *Hydrobiologia* 850, no. 6: 1263–1282. https://doi.org/10.1007/s10750-022-04891-z.

Schneider, J., E. Mas-Carrió, C. Jan, et al. 2021. "Comprehensive Coverage of Human Last Meal Components Revealed by a Forensic DNA Metabarcoding Approach." *Scientific Reports* 11, no. 1: 8876. https://doi.org/10.1038/s41598-021-88418-x.

Sepulveda, A. J., J. Schabacker, S. Smith, R. Al-Chokhachy, G. Luikart, and S. J. Amish. 2019. "Improved Detection of Rare, Endangered and Invasive Trout in Using a New Large-Volume Sampling Method for eDNA Capture." *Environmental DNA* 1, no. 3: 227–237. https://doi.org/10.1002/edn3.23.

Sgarlatta, M. P., A. Ramírez-Valdez, L. B. Ladah, and L. E. Calderón-Aguilera. 2023. "Fish Functional Diversity Is Modulated by Small-Scale Habitat Complexity in a Temperate Ecosystem." *Hydrobiologia* 850, no. 4: 747–759. https://doi.org/10.1007/s10750-022-05061-x.

Shirazi, S., R. S. Meyer, and B. Shapiro. 2021. "Revisiting the Effect of PCR Replication and Sequencing Depth on Biodiversity Metrics in Environmental DNA Metabarcoding." *Ecology and Evolution* 11, no. 22: 15766–15779. https://doi.org/10.1002/ece3.8239.

Shu, L., A. Ludwig, and Z. Peng. 2020. "Standards for Methods Utilizing Environmental DNA for Detection of Fish Species." *Genes* 11, no. 3: 296. https://doi.org/10.3390/genes11030296.

Skaug, H., D. Fournier, B. Bolker, A. Magnusson, and A. Nielsen. 2016. "glmmADMB: Generalized Linear Mixed Models Using AD Model Builder." In *R package version 0* (Vol. 5).

Stauffer, S., M. Jucker, T. Keggin, et al. 2021. "How Many Replicates to Accurately Estimate Fish Biodiversity Using Environmental DNA on Coral Reefs?" *Ecology and Evolution* 11, no. 21: 14630–14643. https://doi.org/10.1002/ece3.8150.

Stewart, K. A. 2019. "Understanding the Effects of Biotic and Abiotic Factors on Sources of Aquatic Environmental DNA." *Biodiversity and Conservation* 28, no. 5: 983–1001. https://doi.org/10.1007/s10531-019-01709-8.

Taberlet, P., A. Bonin, L. Zinger, and E. Coissac. 2018. *Environmental DNA: For Biodiversity Research and Monitoring*. Oxford University Press.

Troth, C. R., M. J. Sweet, J. Nightingale, and A. Burian. 2021. "Seasonality, DNA Degradation and Spatial Heterogeneity as Drivers of eDNA Detection Dynamics." *Science of the Total Environment* 768: 144466. https://doi.org/10.1016/j.scitotenv.2020.144466.

Ushio, M., S. Furukawa, H. Murakami, R. Masuda, and A. J. Nagano. 2022. "An Efficient Early-Pooling Protocol for Environmental DNA Metabarcoding." *Environmental DNA* 4, no. 6: 1212–1228. https://doi.org/10.1002/edn3.337.

Van Driessche, C., T. Everts, S. Neyrinck, et al. 2023. "Using Environmental DNA Metabarcoding to Monitor Fish Communities in Small Rivers and Large Brooks: Insights on the Spatial Scale of Information." *Environmental Research* 228: 115857. https://doi.org/10.1016/j.envres.2023.115857.

Vourka, A., I. Karaouzas, and A. Parmakelis. 2023. "River Benthic Macroinvertebrates and Environmental DNA Metabarcoding: A Scoping Review of eDNA Sampling, Extraction, Amplification and Sequencing Methods." *Biodiversity and Conservation* 32, no. 13: 4221–4238. https://doi.org/10.1007/s10531-023-02710-y.

Wickham, H. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.

Xiong, F., L. Shu, H. Zeng, X. Gan, S. He, and Z. Peng. 2022. "Methodology for Fish Biodiversity Monitoring With Environmental DNA Metabarcoding: The Primers, Databases and Bioinformatic Pipelines." *Water Biology and Security* 1: 100007. https://doi.org/10.1016/j.watbs.2022.100007.

Yao, M., S. Zhang, Q. Lu, et al. 2022. "Fishing for Fish Environmental DNA: Ecological Applications, Methodological Considerations, Surveying Designs, and Ways Forward." *Molecular Ecology* 31, no. 20: 5132–5164. https://doi.org/10.1111/mec.16659.

Zhang, S., Q. Lu, Y. Wang, X. Wang, J. Zhao, and M. Yao. 2020. "Assessment of Fish Communities Using Environmental DNA: Effect of Spatial Sampling Design in Lentic Systems of Different Sizes." *Molecular Ecology Resources* 20, no. 1: 242–255. https://doi.org/10.1111/1755-0998.13105.

Zinger, L., A. Bonin, I. G. Alsos, et al. 2019. "DNA Metabarcoding—Need for Robust Experimental Designs to Draw Sound Ecological Conclusions." *Molecular Ecology* 28, no. 8: 1857–1862. https://doi.org/10.1111/mec.15060.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.