PERSPECTIVE

WILEY

Quantifying biodiversity using eDNA from water bodies: General principles and recommendations for sampling designs

Florian Altermatt[1,2] | Luca Carraro[1,2] | Manuel Antonetti[3] | Camille Albouy[4,5] | Yan Zhang[6] | Arnaud Lyet[7] | Xiaowei Zhang[6] | Loïc Pellissier[4,5]

[1]Department of Evolutionary Biology and Environmental Studies, University of Zürich, Zürich, Switzerland

[2]Department of Aquatic Ecology, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

[3]ZHAW School of Life Sciences and Facility Management, Institute of Natural Resource Sciences, Wädenswil, Switzerland

[4]Ecosystems and Landscape Evolution, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland

[5]Unit of Land Change Science, Swiss Federal Research Institute for Forest, Snow and Landscape Research (WSL), Birmensdorf, Switzerland

[6]State Key Laboratory of Pollution Control & Resource Reuse, School of the Environment, Nanjing University, Nanjing, China

[7]World Wildlife Fund, Wildlife Conservation Team, District of Columbia, Washington, USA

Correspondence
Florian Altermatt, Department of Evolutionary Biology and Environmental Studies, University of Zürich, Zürich 8057, Switzerland.
Email: florian.altermatt@ieu.uzh.ch

Xiaowei Zhang, State Key Laboratory of Pollution Control & Resource Reuse, School of the Environment, Nanjing University, Nanjing 210023, China.
Email: zhangxw@nju.edu.cn

Loïc Pellissier, Ecosystems and Landscape Evolution, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zürich, Zürich 8092, Switzerland.
Email: loic.pellissier@usys.ethz.ch

Abstract

Reliable and comparable estimates of biodiversity are the foundation for understanding ecological systems and informing policy and decision-making, especially in an era of massive anthropogenic impacts on biodiversity. Environmental DNA (eDNA) metabarcoding is at the forefront of technological advances in biodiversity monitoring, and the last few years have seen major progress and solutions to technical challenges from the laboratory to bioinformatics. Water eDNA has been shown to allow the fast and efficient recovery of biodiversity signals, but the rapid pace of technological development has meant that some important principles regarding sampling design, which are well established in traditional biodiversity inventories, have been neglected. Using a spatially explicit river flow model, we illustrate how sampling must be adjusted to the size of the watercourse to increase the quality of the biodiversity signal recovered. We additionally investigate the effect of sampling parameters (volume, number of sites, sequencing depth) on detection probability in an empirical data set. Based on traditional sampling principles, we propose that aquatic eDNA sampling replication and volume must be scaled to match the organisms' and ecosystems' properties to provide reliable biodiversity estimates. We present a generalizable conceptual equation describing sampling features as a function of the size of the ecosystem monitored, the abundance of target organisms, and the properties of the sequencing procedure. The aim of this formalization is to enhance the standardization of critical steps in the design of biodiversity inventory studies using eDNA. More robust sampling standards will generate more comparable biodiversity data from eDNA, which is necessary for the method's long-term plausibility and comparability.

1 | INTRODUCTION

The fast pace of the global erosion of biodiversity calls for more efficient ways to identify how anthropogenic activities are responsible for species loss (Schmeller et al., 2017), determine the effectiveness of protected areas (Leverington et al., 2010), and monitor whether restoration efforts lead to biodiversity recovery (Barral et al., 2015). Technological advances in molecular biology are revolutionizing biodiversity sciences, and new molecular techniques are replacing classical methods for estimating species composition, species richness, and other biodiversity components based on species composition, such as functional and phylogenetic diversity. Environmental DNA (eDNA) metabarcoding, in which signals from macro- and microorganisms in the environment are recovered from simple samples of soil, air or water (Deiner et al., 2017; Pawlowski, Apothéloz-Perret-Gentil, & Altermatt, 2020; Taberlet et al., 2012), is currently increasing the speed and spatial scale of biodiversity monitoring programmes (Ruppert et al., 2019). An increasing number of studies have demonstrated that eDNA metabarcoding can detect species equally well or better than traditional visual survey methods (Keck et al., 2022; Polanco Fernández, Marques, et al., 2021) and that it can extend biodiversity studies into previously overlooked taxonomic groups (Nester et al., 2020; Pellissier et al., 2014). In particular, the aquatic environment is appropriate for collecting information on organisms found in large areas, including both aquatic (freshwater and marine; e.g., Li et al., 2020; Polanco Fernández, Mutis Martinezguerra, et al., 2021) and terrestrial organisms, such as mammals (e.g., Lyet et al., 2021), since water may integrate biodiversity information across entire landscapes and seascapes (Altermatt et al., 2020; Deiner et al., 2016; Zhang et al., 2023).

The ambition of eDNA research is to complement or even replace existing methods to monitor biodiversity and estimate species composition and all biodiversity components (Deiner et al., 2017; Lawson Handley, 2015). The entire set of eDNA metabarcoding operations should therefore be based on protocols that are entirely reproducible (Dickie et al., 2018), even while the methodology has some inherent limitations (Cristescu & Hebert, 2018). An increasing body of eDNA research has been devoted to defining common standards in laboratory protocols (e.g., Coutant et al., 2021; Spens et al., 2017) and has resulted in general guidelines and best practices (e.g., Bruce et al., 2021; Burian et al., 2021; Pawlowski, Apothéloz-Perret-Gentil, Mächler, & Altermatt, 2020). However, the protocols and established general practices for sample collection in the field have received less attention. Furthermore, studies examining eDNA sampling protocols are few compared with those on traditional sampling methods (Dickie et al., 2018), including traditional plant quadrants, plankton netting, and electrofishing.

The distribution and dynamics of eDNA in water bodies are complex and should be accounted for in the sampling procedure (Carraro et al., 2018). For example, DNA from whole organisms, tissue fragments or loose molecules is subject to advection, diffusion and selective decay in the water medium (Barnes et al., 2014; Barnes & Turner, 2016; Harrison et al., 2019; Shogren et al., 2017), and these processes crucially affect sampled DNA concentrations. eDNA studies have hitherto generally assumed that particles are well mixed in aquatic ecosystems, and that a small volume is thus sufficient to represent the entire system. However, there are concerns among ecologists that the DNA fragments of target organisms in commonly used sampling volumes – typically less than 1 L – reach only "homeopathic levels", hence offering poor species detectability and representation in the target aquatic system. Species accumulation curves show saturation at larger volumes than those sampled in most eDNA studies (Bessey et al., 2020; Broadhurst et al., 2021; Macher et al., 2021). As a result, suboptimal sampling replicates or volumes can lead to stochasticity in the recovered signals of species diversity (Bessey et al., 2020; Stauffer et al., 2021). Moreover, eDNA metabarcoding studies target different aquatic ecosystems, from ponds to rivers to marine systems, and the inferences made can vary widely across spatial, temporal and taxonomic scales (e.g., Altermatt et al., 2020; Carraro et al., 2022; Deiner & Altermatt, 2014).

To reach the full potential of biodiversity monitoring with eDNA, the sampling methods should be adjusted based on the properties of the study system (e.g., a regional area, a catchment, a river), including the expected diversity of the species pool (Gotelli & Colwell, 2001) and the level of environmental heterogeneity (Barbour, 1999). This may seem obvious when considering traditional sampling methods: for example, the sampling design for vascular plants in a meadow differs from that for forest trees, amphibians are sampled differently than marine mammals, and microbe surveys have different sampling designs than fish surveys. The commonality of all these approaches, however, is that the sampling has been optimized with respect to the system and organismal group covered, inspired and guided by decades of traditional sampling experience. As such, best practices and principles for eDNA sampling designs could be developed based on lessons learned from the optimization of traditional sampling methods. For example, estimating the regional species pool requires a more intense sampling effort in tropical than in temperate forests because tropical areas have a higher total diversity and higher levels of rarity (Figure 1a). This principle is also valid for eDNA sampling (Figure 1b), where marine coral reefs require more sampling effort than cold Atlantic waters.

Computer simulations can be used to determine the best sampling strategy in terrestrial, riverine or marine systems. In this context, the virtual ecologist approach combines simulated data and observer models to mimic species distributions and how they would be observed (Zurell et al., 2010). The virtual data are then modeled, and the results are compared with the true data, allowing an assessment of sampling protocols and an improvement of system-specific sampling methods (Hirzel & Guisan, 2002). For example, researchers have used the virtual ecologist approach to compare different sampling strategies and investigate the impact of eDNA transport on freshwater ecosystems (Carraro et al., 2021).
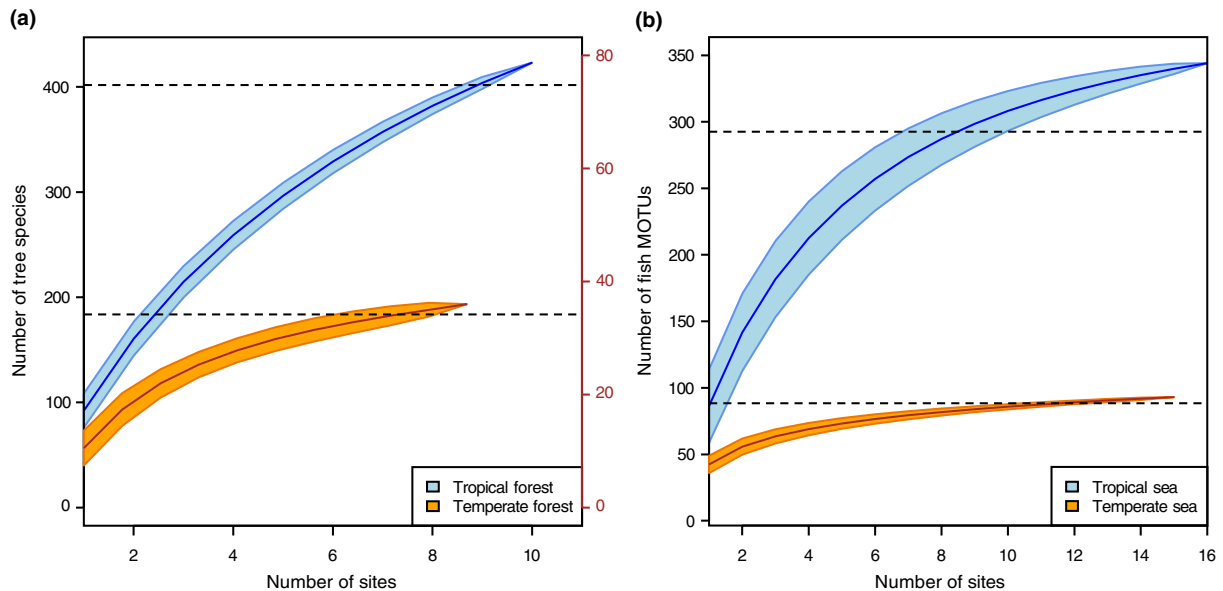
FIGURE 1 Results of ecological sampling displayed as accumulation curves of species richness and molecular operational taxonomic units (MOTUs). (a) Sampling in forest communities using traditional methods. (b) Sampling in marine fish communities using eDNA methods in tropical and temperate environments. The curves in (a) are based on a public data set (https://forestplots.net/) for 10 forest plots in Bolivia (tropical) and on data from 10 forest plots in western Virginia, USA (temperate) that are part of the US forest inventory and analysis database. We show that species saturation can be reached with fewer samples at higher latitudes than at lower latitudes, where ecosystems are more complex and contain more rare species. The tropical curve in (b) is based on the data set of Polanco Fernández, Marques, et al. (2021) on two tropical regions of Colombia, i.e. the island of Providencia (10 stations) and the Tayrona National Natural Park (6 stations). Two filtration replicates of 30 L were collected per station. The temperate curve in (b) is based on a data set collected during the EVHOE bottom trawl research survey in 2019, where two filtration replicates of 22.5 L were collected per station using Niskin bottles available on a circular rosette. eDNA results are similar to forest plot results, in that a more intense sampling effort is necessary to represent the species pool in tropical regions compared with in temperate regions.

Here, we present the challenge of optimizing the sampling design for eDNA studies from aquatic environments, considering the interaction between hydrological conditions, the volume of water filtered, the number of replicates, and the sampling locations. To evaluate and illustrate these aspects, we used an advanced river hydrological simulation coupled with a virtual ecologist approach to gain insight on the difference in sampling strategy (a large single sample vs. multiple small samples) required between different river types (narrow alpine streams vs. wide lowland rivers). We quantified the effect of field sampling (sampling volume, number of spatial replicates) and laboratory (sequencing depth) variables on species detection probability and recovered biodiversity. Based on both simulations and empirical data, we propose a set of criteria for the development of unbiased biodiversity detection when planning an eDNA sampling campaign for conservation and management. We focus on eDNA in the wider sense, that is, including both intra- and extracellular DNA, yet acknowledge the ongoing discussion about how the targeted DNA potentially reflects a wide range of organisms and possible states (see Pawlowski et al., 2021; Pawlowski, Apothéloz-Perret-Gentil, & Altermatt, 2020; Rodriguez-Ezpeleta et al., 2021). Our general conclusions hold true irrespective of the different states of DNA, yet the specific recommendations for sampling may vary.

2 | CONTRASTING VIRTUAL ENVIRONMENTS FOR eDNA SAMPLING DESIGN

Traditional sampling schemes generally include a clear spatial design and sampling efforts that optimize the detection of target species in assemblages (Lengyel et al., 2018; Stewart et al., 2018). The most common sampling strategy used in environmental science is the stratified random sampling approach, with the objective to equally survey the varied ecological conditions of a considered area (Carvalho et al., 2016). Optimized biodiversity monitoring networks based on traditional vegetation surveys (Grabherr et al., 2000), insect surveys (Kery & Plattner, 2007), plankton surveys (Cermeño et al., 2014; Rodriguez-Ramos et al., 2014), or freshwater and marine fish surveys (e.g., Alexander & Seehausen, 2021) have emerged from systematic methodological trials to improve species detection and diversity assessments locally and across large areas. In contrast, aquatic eDNA studies have so far mainly been exploratory or opportunistic (i.e. coupled with another sampling engine), as suggested by the absence of clear designs. To reach the robustness of traditional sampling, the location of samples of aquatic eDNA within the target area must be determined, with the objective of representativeness of the study system, and the sampling design should follow rules that

4 | **WILEY**—Environmental DNA
Dedicated to the study and use of environmental DNA for basic and applied sciences
Open Access

ALTERMATT ET AL.

generate unbiased regional biodiversity measurements (Carraro et al., 2021). According to our literature review (see Appendix S1 for details), one to three replicates of about 1 L of water were collected in most studies, but this sampling effort was, in most cases, not adjusted according to the characteristics of the target organisms and systems. In this review, we defined replicate as "a sample taken at a given site and time point", which is, to our understanding, the level of replication that is most fundamental to many studies involving biodiversity assessments (e.g., to document diversity across space at a given time; Gotelli & Colwell, 2001). All in all, decades of research have optimized different traditional monitoring methodologies, accounting for the properties of the local system and the target species, and a similar optimization should be done for eDNA sampling as well.

To assess the impact of various sampling designs on eDNA signal recovery, we conducted an in silico analysis comparing the efficiency of different eDNA sampling techniques using a river process model in two vastly different river systems: a typical alpine headwater stream and a major lowland river. We used the BASEMENT v3 model (VAW-ETHZ; Vanzo et al., 2021), which can simulate the advection and diffusion of passive tracers with different concentrations, here representing the eDNA of distinct species, in river reaches on a two-dimensional (2D) computational mesh. We defined both river reaches with a rectangular cross section and modeled them under uniform flow conditions. The alpine stream reach had a width of 2.5 m and length of 1 km and was simulated using a computational mesh with 77,895 cells with a mean cell size of $0.03\,\text{m}^2$. A constant discharge of $1.3\,\text{m}^3\,\text{s}^{-1}$, a longitudinal slope of 0.05, and a friction (Strickler) coefficient of $20\,\text{m}^{1/3}\,\text{s}^{-1}$ led to a mean simulated water depth of 0.27 m and a mean flow velocity of $1.9\,\text{m s}^{-1}$. The major river had a width of 500 m and a length of 20 km and was discretized with a computational mesh with ca 7.6×10^6 cells (mean cell size $1.3\,\text{m}^2$). A constant discharge of $28{,}000\,\text{m}^3\,\text{s}^{-1}$, a slope of 0.0001, and a Strickler friction coefficient of $20\,\text{m}^{1/3}\,\text{s}^{-1}$ led to water depth of 19.4 m and a mean flow velocity of $2.9\,\text{m s}^{-1}$. We defined a total of five DNA-shedding species (i.e. five distinguished taxa) as stationary local tracer sources, equidistantly located along the upstream cross section of each model domain. Each source was characterized by a different, time-constant shedding rate (equal to 1000, 2000, 3000, 4000 or 5000 copies s^{-1}). These values were randomly attributed to the five equidistant source locations along the upstream cross sections of the rivers. We simulated the transport dynamics of these tracers, assuming a constant diffusion coefficient (K_c) of $0.25\,\text{m}^2\,\text{s}^{-1}$ (Figure 2a,b). At the downstream end of each river reach, we simulated water sampling with various sampling volumes V_s under two different strategies: (i) collection of a single sample of volume V_s at a random point along the cross section (single sample strategy, hereafter SSS); (ii) collection of five samples, each of volume $V_s/5$, at five random points along the cross section (multiple sample strategy, hereafter MSS). For each river reach, we tested 100 different V_s values spanning river-specific ranges, given that much larger volumes need to be sampled in a larger river. For the alpine stream, V_s ranged from 0.1 to 10 L, with increments of 0.1 L; for the lowland river, V_s

ranged from 1 to $100\,\text{m}^3$, with increments of $1\,\text{m}^3$. For each river, sampling strategy and V_s value, we performed 10,000 sampling replicates, to account for uncertainty resulting from the random positioning of the transversal coordinates of the sampling points.

We found that, while complete mixing of eDNA along the transversal direction occurred a few meters downstream of the source location for the alpine stream (Figure 2a, consistent with Thalinger et al., 2021), in the lowland river complete transversal mixing did not occur even 20 km downstream of the source site (Figure 2b – see also Laporte et al., 2020). As a result, for the alpine stream no differences between the sampling strategies emerged with respect to the amount of eDNA sampled for a given species (Figure 2c), the total amount of DNA sampled pooled across the five species (Figure 2e), or the total number of species detected (Figure 2g; evaluated by assuming a limit of detection of two copies for each species). Conversely, limited transversal mixing in the lowland river led to an uneven DNA concentration distribution across the sampling cross section (Figure 2b), which resulted in DNA amounts being generally greater for MSS than for SSS (Figure 2d – results for the other tracers are shown in Figures S1–S4). The median of the total DNA quantity sampled (pooled over the five species) did not depend on the sampling strategy, although MSS had a lower uncertainty than SSS (Figure 2f). MSS clearly outperformed SSS in terms of the number of species detected (Figure 2h), as no simulation under SSS was able to detect all five species with a sampling volume as large as $100\,\text{m}^3$. SSS performed better than MSS only for very small sampling volumes ($< 20\,\text{m}^3$), which were in any case insufficient to detect more than two out of five species. Importantly, empirical assessments and comparisons of eDNA distributions in natural settings, specifically considering hydrological dynamics, are needed. So far, this has been done in only a few studies (e.g., Laporte et al., 2020; Thalinger et al., 2021), with results highlighting the importance of hydrological conditions for the lateral and longitudinal distribution of eDNA and thus for the interpretation of eDNA-based data. Our model can be considered a first formalization that can be used in generalizable comparisons.

This in silico experiment shows that a small sampling volume does not capture the same level of biodiversity in these two systems (note the different x-axis scales in the two columns of Figure 2), and that a large volume is necessary to detect most species in the large, lowland river. Moreover, the strategy of taking multiple samples (MSS) involves a higher likelihood of species detection, and therefore better representativeness of the system, than solely relying on a large volume (SSS). Our illustration shows the need to sample at multiple locations across the river and to use the appropriate volume to reach the eDNA concentration that allows detection. Note that the BASEMENT model reproduces the transport dynamics of conservative tracers; hence, our proof-of-concept analysis neglects the role of eDNA decay and sedimentation, and is thus an idealized approximation. In fact, the inclusion of decay dynamics would lead to concentration profiles across the sampling cross sections of the two rivers that would be qualitatively equal to those in Figure 2a,b, i.e. flat for the alpine stream and unimodal for the lowland river. In our 2D model, factors such as sedimentation were not included,
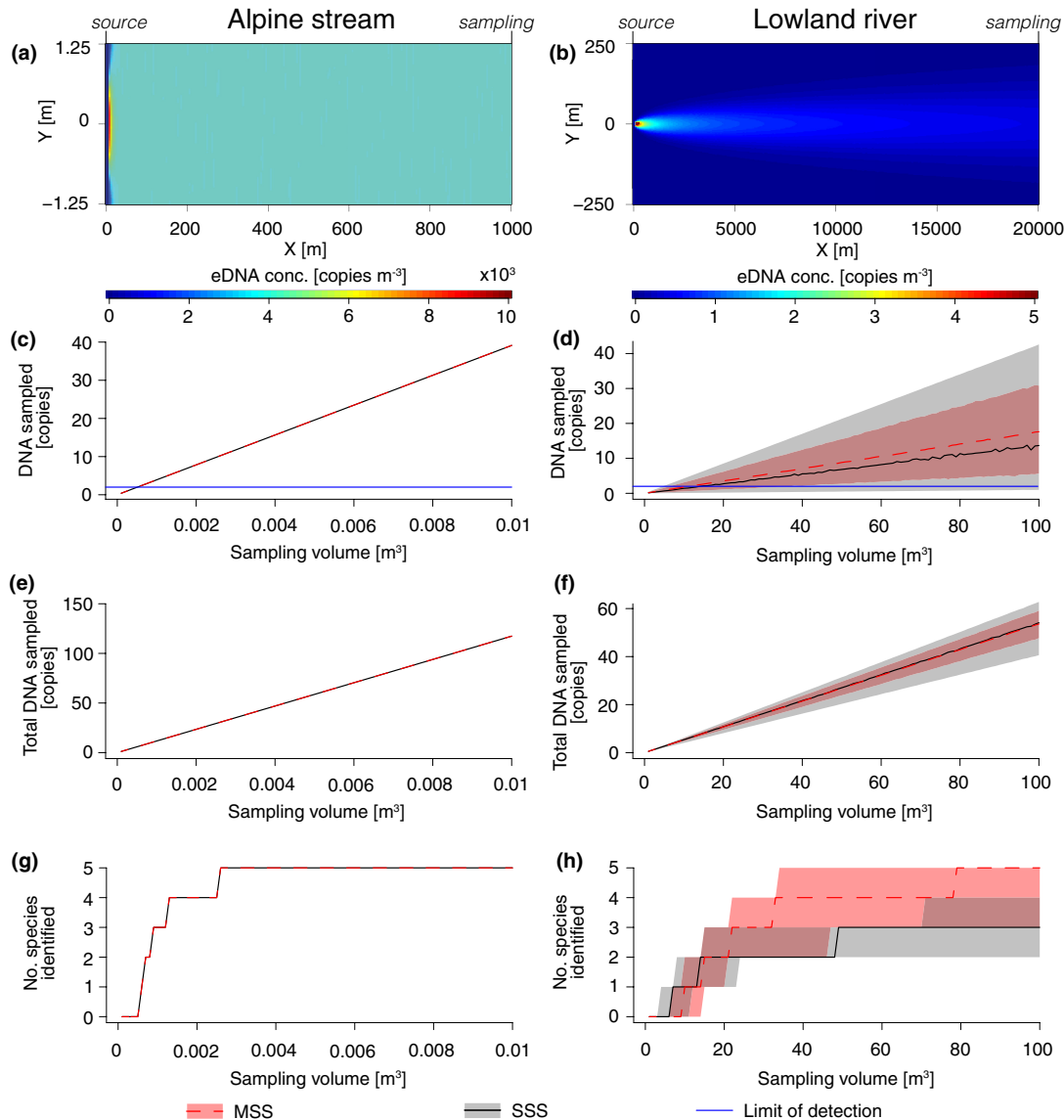
FIGURE 2 Effect of sampling volume and sampling intensity on eDNA retrieval and species detection. The left and right columns refer to an alpine stream and a lowland river, respectively (note that scales differ between columns). (a, b) Steady-state eDNA concentration profile for one of the five species (i.e. the one located at $Y=0$). (c, d) Amount of eDNA collected for the species of panels (a) and (b), respectively, as a function of sampling volume. Red lines and shading correspond to the multiple sampling strategy (MSS) and black to the single (SSS) sampling strategy. Shaded areas correspond to the 2.5th to 97.5th percentiles; lines correspond to median values. The blue line is the limit of detection (assumed to be two DNA copies). (e, f) Total eDNA sampled for the five species pooled together as a function of V_s. (g, h) Number of species detected as a function of V_s, where a species is detected if the eDNA present in the collected sample(s) exceeds the limit of detection.

but since eDNA binds to small sediments, sediment concentration, transport and precipitation should be considered further (Laporte et al., 2020).

3 | QUANTIFICATION OF DETECTION PROBABILITIES USING AN ILLUSTRATIVE DATASET

Detecting all species comprehensively in a target ecosystem requires that the sampling design be optimized according to the properties of

the ecosystem and the target organisms. eDNA sampling campaigns should be designed with parameters that enable the detection of all species of interest and thus ensure that rare species, or those more difficult to detect, are also recovered in the survey (Roberts et al., 2016; Zhang et al., 2014). eDNA sampling has proven particularly efficient for aquatic microorganisms (Cordier et al., 2022; Mansfeldt et al., 2020), which are small and abundant enough to be consistently found in the sample (Pawlowski, Apothéloz-Perret-Gentil, & Altermatt, 2020), and are mostly sampled in their cellular stage, leading to large DNA quantities in the samples. When targeting large species (not entirely sampled, i.e. extracellular DNA

6 | **WILEY**— Environmental DNA
Dedicated to the study and use of environmental DNA for basic and applied sciences
Open Access

ALTERMATT ET AL.

samples) living directly in the sampled habitat, e.g., fishes in freshwater, smaller quantities of DNA may occur in the environment, and attention should be paid to volume and duration of filtration to make the sampling integrative. Possibly more challenging is the detection of terrestrial organisms, which have only limited contact with the aquatic environment. In this case, direct contact between water and the organism (or its feces, urine, hair, etc.) needs to occur for DNA to be released in the aquatic system. This happens, for example, when animals drink or bathe, and the water serves as a DNA collector (Rodgers & Mock, 2015; Ushio et al., 2017). For these species, good detection is reached with larger sampling volumes (Lyet et al., 2021). Complementing the information gained from simulations, comprehensive eDNA metabarcoding data sets can be used to evaluate the role of sampling parameters on the detection probability of rare or elusive species.

Detecting terrestrial vertebrates by sampling the river reaches of a catchment has been shown to be challenging and require sensitive eDNA sampling protocols (Lyet et al., 2021). While their DNA traces still diffuse into the aquatic system, the shedding rate and transmission to the water is less predictable, such that eDNA sampling would require larger volumes and more replication to successfuly detect those species. Consequently, eDNA sampling can offer an efficient illustration of the effect of sampling parameters on the probability of detection of species. We used a dataset collected in British Columbia (Lyet et al., 2021) to evaluate the effect of sampled volume (within a given site), number of spatially replicated sites across the catchment, and sequencing depth on the probability of detection and species richness of terrestrial mammals. To evaluate the spatial replication, we used 50 eDNA samples collected from 42 sites at different geographic positions within 2 juxtaposed catchments. Moreover, to evaluate the effect of sampled volume, we used 36 samples collected at 2 focal sites at the outlet of the 2 catchments, with the duration of filtration proportionally increasing with filtered water volume (25–30 L over 30 min; 40–45 L over 45 min; 50–60 L over 2 h; 70 – 80 L over 5 h). At each of these 2 sites, we collected replicate samples for each setting. We used a universal mammal 12S mitochondrial rDNA primer (Mamm01) for PCR amplifications, and we prepared the libraries for sequencing to reach a theoretical sequencing depth of 300,000 reads per sample (Lyet et al., 2021). Using this data set, we quantified the probability of detection of mammal species as a function of sampling volume, number of samples, and sequencing depth. We used a re-sampling approach to assess the effect of each of these parameters on detection probabilities and number of species detected while keeping all the other parameters constant. We selected four species for illustrative purposes, but considered all species in the quantification.

The detection probability increased with sequencing depth, number of sampled sites, and filtered water volume (Figure 3). Increasing the number of reads from 1000 to 20,000 led to a marked increase in detection probability, from 0.22 to 0.55 across all mammal species detected (from a mean number of 7 mammal species detected to a mean number of 10 species per sample), but
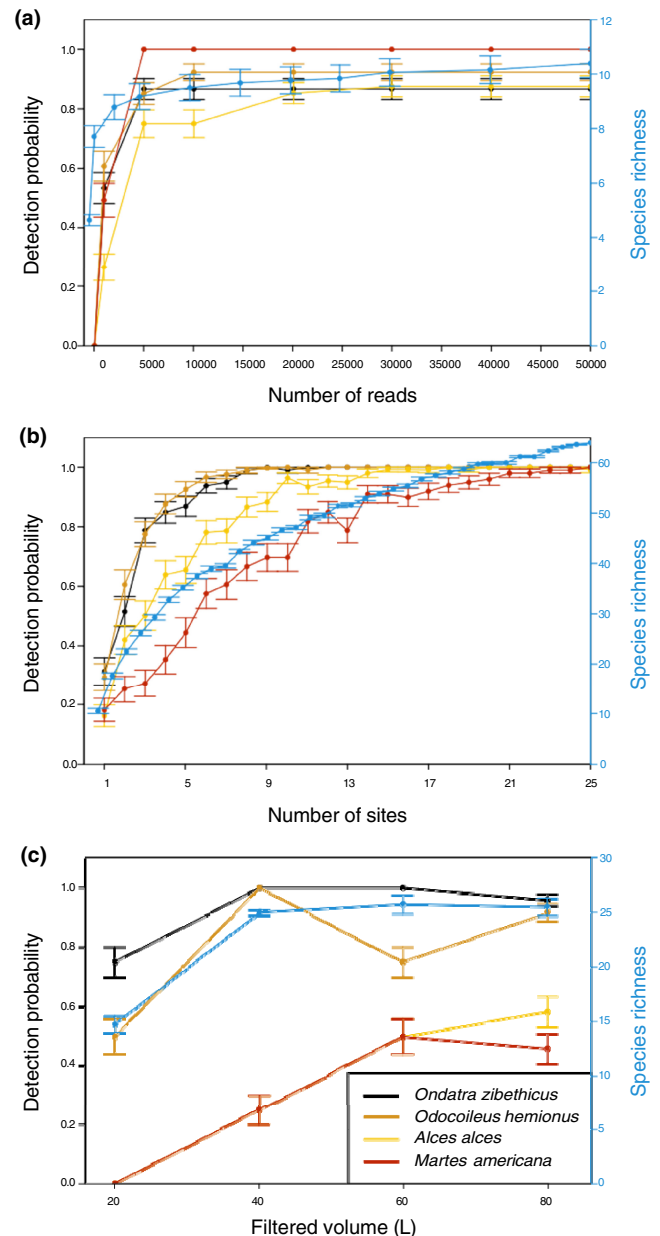

FIGURE 3 Estimation of the probability of detection of four example mammal species, as well as the species richness recovered, as a function of (a) the number of reads considered per sample, (b) the number of different sites where samples were taken within the catchments, and (c) the volume of water filtered at one site. The detection probability of species increases as a function of these three parameters. The eDNA samples used to assess the effects of the number of reads and the number of sites are different from those used to assess the effect of filtered volume. The figure illustrates how the detection of rarer species requires more field replicates or a larger volume than needed for the detection of more common species.

a large number of reads made it possible to detect the maximum number of species. Increasing the number of samples across the catchments increased the detection probability of the species: an increase from 1 site to 5 sites raised the detection probability from 0.16 to 0.41 on average across all mammal species detected

(corresponding to a mean number of 10 versus a mean number of 29 mammal species detected when cumulating samples). As shown for marine ecosystems (Stauffer et al., 2021), this analysis illustrates the importance of field replicates to increase species detection at the regional level. We further found that increasing the sampled volume (associated with longer filtering duration) from 25–35 L over 30 min to 50–60 L over 2 h increased the detection probability from 0.22 to 0.38 on average across all mammal species detected (from a mean number of 15 mammal species detected to a mean number of 25 mammal species detected per sample). These results demonstrate the critical role of sampling parameters on species detection, showing how sampling optimization is required for eDNA-based, as previously done for traditional sampling designs. For example, in camera trapping surveys (O'Connell et al., 2011), the probability of detection of an animal in an area increases if the camera is left in the field for a longer period, if a larger number of cameras are used, and if the cameras cover a larger area of detection (i.e. wide camera viewshed). Hence, similar to how the duration of sampling and the volume and number of samples influence the detection of terrestrial mammals via eDNA, the combined effect of these parameters determines the likelihood of detecting a rare event, e.g., an animal crossing the camera field of view. Therefore, considering these sampling parameter combinations is critical for terrestrial species, whose eDNA is more unevenly released into water compared with that of aquatic taxa, but is also likely to impact other organisms beyond terrestrial mammals.

4 | A MECHANISTIC EQUATION TO DETERMINE SAMPLING VOLUME

Multiple factors need to be accounted for to estimate the volume required for eDNA sampling (see Box 1 for a detailed discussion). We formalized the conceptual equation Equation (1) under a number of simplifying assumptions. Specifically, we assumed a linear channel with spatially and temporally constant water discharge Q [m^3 s^{-1}]. Additionally, we assumed that biomass M [g] of a (sessile) eDNA source is concentrated at a given point in the channel sheds eDNA at a rate S [copies s^{-1} g^{-1}]. We further assumed instantaneous, complete mixing of eDNA in the river cross section where the eDNA source is located. The input eDNA flux is then SM, while the downstream flux is CQ, where C [copies m^{-3}] is the eDNA concentration in water immediately downstream of the eDNA source. Hence, the mass balance of such a river cross section is $SM = CQ$. By assuming first-order decay of eDNA in stream water, the eDNA concentration at a distance L downstream of the source site is equal to:

$$C_L = C \exp\left(-k \frac{LA}{Q} \right) = \frac{SM}{Q} \exp\left(-k \frac{LA}{Q} \right)$$

where A [m^2] is the river cross-section area and k [s^{-1}] is the decay rate.

If a water volume V_s is sampled at distance L from the source site, then the number N of copies in the sample after filtration can be expressed as:

$$N = \frac{SMV_s}{Q} \exp\left(-k \frac{LA}{Q} \right)$$

which assumes complete mixing of eDNA, i.e. that the eDNA concentration in the sample is equal to C_L. However, due to primer bias and sequencing depth issues, the estimated number of copies from a PCR analysis N_e is only a fraction of the true N. We can express this as $N_e = K_{pb} K_{sd} N$, where K_{pb} and K_{sd} are two coefficients ranging from 0 to 1, indicating how read number estimation is affected by primer bias and sequencing depth, respectively. In particular, it may be reasonable to assume that:

$$K_{pb} = \exp\left(-k_{pb} P_M \right); \qquad K_{sd} = 1 - \exp\left(-k_{sd} D_S \right)$$

where P_M is the number of mismatched pairs between the primer and the template; D_s is the sequencing depth; and k_{pb} and k_{sd} are two positive, dimensionless coefficients.

By assuming that the minimum number of copies required for detection is 1, and by coupling the previous relationships, one can determine the minimum sampling volume required for species detection:

$$V_{S,min} = \frac{Q \exp(k_{pb} P_M)}{SM[1 - \exp(-k_{sd} D_S)]} \exp\left(k \frac{LA}{Q} \right) \tag{2}$$

As an example, we considered an eDNA source that sheds $SM = 5000$ copies s^{-1} (see Fukaya et al., 2021), a decay rate of $(1/4 \times 3600)$ s^{-1} (i.e. a decay time set to 4 h), $P_B = 10$ base pairs, $SD = 100,000$ reads, $k_{pb} = 0.07$, and $k_{sd} = 7 \times 10^{-6}$ (and hence $K_{pb} = K_{sd} \approx 0.5$, i.e. in order to detect one copy in the sample there must be at least four copies, owing to primer bias and finite sequencing depth). We assumed eDNA sampling at $L = 1$ km from the source in a river with $Q = 10$ m^3 s^{-1} and $A = 8$ m^2 (hence the mean water velocity is 1.25 m s^{-1}). The resulting minimum sampling volume based on Equation (2) is $V_{S,min} = 8.46$ L. According to Equation (2), minimum values of the required sampling volume change dramatically as a function of physical (water discharge Q, distance from the shedding source L) and sequencing parameters (sequencing depth S_D, number of mismatched pairs P_B; Figure 4).

Although not all parameters of this equation can be easily quantifiable (most notably, this is the case for coefficients K_{pb} and K_{sd}), Equation (2) can be used as a step towards a better understanding of the factors affecting the detectability of eDNA traces in a sample. Importantly, this equation considers the main parameters to be representative of a local sample. As illustrated with the river model of Figure 2, in the case of a heterogeneous distribution of eDNA, multiple spatial samples are necessary to be representative of a system. Recent studies based on an eDNA transport model at the catchment scale (Carraro et al., 2018, 2020, 2021)

BOX 1 A theoretical equation to plan eDNA sampling

In most previous eDNA studies, the decisions on sampling parameters reflected individual research groups' best practices and experiences, and were often based on rules of thumb or values that worked in other studies, without formal corroboration. Consequently, it has been found – and has become general practice – that a few hundred milliliters of water can be enough to detect single amphibian species in ponds (Biggs et al., 2015; Ficetola et al., 2008), while small headwater streams require 0.5 L to a few liters, and large rivers, lakes and marine systems often need 30 L or more to reach saturation (Hänfling et al., 2016; Mächler et al., 2021; Stauffer et al., 2021). While there is some discussion on sampling volumes (e.g., Wilcox et al., 2018), approaches often differ and the choice of the sampling volume is often not (yet) corroborated or justified.

Ecoinformatic tools and models support the planning of traditional sampling and can thus also help to optimize eDNA sampling plans for a specific system (see also Carraro et al., 2021). Similar to statistical power analysis, we propose that eDNA sampling be guided by a conceptual equation to determine the most important factors for evaluating the water volume that should be sampled, and that either a single integrative sampling event or separate sample replicates be used. Such an equation should contain relevant aspects known to affect detection probabilities, such that it is possible to optimize sampling volume based on the values of different parameters. We identified five key parameters that can be quantified and modified to optimize sampling efficacy. First, total volume of the sampled system: the larger the system, the greater the volume that needs to be sampled. For example, a larger volume of water must be sampled and more DNA must be extracted from a lake versus a pond, or from a large river versus a headwater stream. Second, the minimum expected concentration of the focal DNA: the lower the concentration, the larger the volume that must be sampled to collect the minimum detectable amount of target DNA in the sample. Third, sequencing depth: the more sequences that result from a sample in the sequencing process (i.e. greater sequencing depth/read numbers), the greater the likelihood of recovering the targeted DNA. Sequencing depth is especially relevant with respect to thresholds applied to exclude low-read signals, and increasing sequencing depth can often be an efficient way to better detect target organisms. Fourth, the average abundance distribution of the target organisms: the more equally distributed the target organisms are with respect to their rank-abundance distribution, the smaller the volume that needs to be sampled. This is again related to the total sequencing depth, as a more balanced distribution of taxa and their representation in the sample gives more reliable estimates about presences and absences compared with a situation with unevenly distributed taxa. Last, the taxonomic specificity of the primer used, i.e. the primer breadth and fit to the target organisms: the more specific a primer is, the better it can be expected to fit the target organisms and the more likely it is to detect them in small sampling volumes. In contrast, less specific primers or the presence of non-target DNA competing with the target DNA for a given primer lower the latter's detectability.

We summarize these five major parameters in the following conceptual equation:

$$\text{Sample volume} \approx \frac{\text{Total volume} \times \text{Minimum eDNA conc.} \times \text{Sequencing depth}}{\text{Abundance} \times \text{Primer specificity}} \tag{1}$$

Each of these parameters has, if not considered adequately, the potential to cause false-absences and thus incorrect inference about species' occurrences. While Equation (1) is conceptual, each of these (as well as possible further) parameters can be quantified. In designing studies, researchers should at least consider these five parameters conceptually and provide some quantitative support for the values chosen. Further, Equation (1) allows a semi-quantitative optimization with respect to sampling procedure. Assuming linear relationships, these different parameters can be gauged in relation to each other each other. Consequently, these parameters can be adjusted individually, even without knowledge of their exact values.

Several of these parameters have an inherent minimum value. For example, if the target DNA is found in concentrations of less than one copy per 1 L, then a smaller sampling volume will result, by pure probability, in false absences. In this case, it is likely that a much larger volume needs to be sampled, as the lowest detectability, i.e. the smallest number of eDNA copies that must be present in the sample to detect a signal, has not yet been quantified and may be much greater than only one. An eDNA signal can only be found if DNA of the target organism is present, and values below the detection limit are simply homeopathic. Other parameters, such as primer specificity, also depend on the presence of non-target DNA, which may be amplified by a more-or-less specific primer, and thus minimum values are not well defined.

While the exact values of these parameters, yet to be estimated, may be specific to the DNA amplified for different organismal groups, the relationship in Equation (1) would not be fundamentally different. For example, microbial DNA extracted from an environmental sample may largely contain whole organisms (e.g., bacteria, protists), but DNA of macroorganisms (e.g., fish) may be largely composed of extracellular DNA. Consequently, organismal abundance or DNA concentrations of the former are often found to be much higher, yet the extraction and amplification process generally does not differentiate between different states of DNA. We argue that parameter values are informative within – yet not necessarily across – these sample types. Besides these five relevant parameters, we acknowledge that further aspects, such as seasonality, life history, or other facets of organisms' ecology may also need to be considered, and we recommend that additional empirical studies on such aspects be conducted.
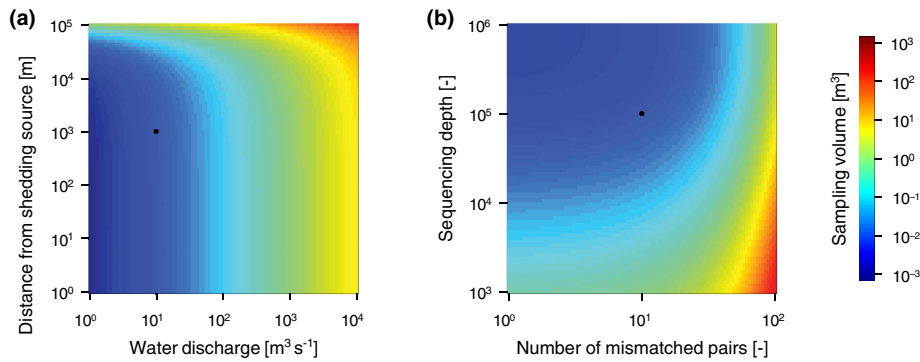
FIGURE 4 Minimum sampling volume required, expressed as a function of relevant parameters according to Equation (2). (a) Water discharge Q and distance from shedding source L. (b) Number of mismatched pairs P_M and sequencing depth D_S. The cross-section area A is dependent on water discharge Q according to the scaling relationship of Leopold and Maddock (1953): $A = aQ^{0.9}$, where Q is in $m^3 s^{-1}$, A is in m^2, and it is assumed that $a = 1$. Default parameters (corresponding to the black dot) are: $Q = 10\,m^3 s^{-1}$, $L = 1000\,m$, $k = 1/4/3600\,s^{-1}$, $SM = 500$ copies s^{-1}, $D_S = 10^5$, $P_M = 10$, $k_{sd} = 7 \times 10^{-6}$, and $k_{pb} = 0.07$.

have demonstrated the importance of considering both the spatial distribution of the shedding species and the spatial location of the sampling sites. In particular, Carraro et al. (2021) showed that best practices for the selection of sampling sites in a river network depend on multiple factors, such as the species' spatial distribution, the total number of sites that can be sampled, and the knowledge about eDNA decay rates. Local and regional parameters could additionally be integrated into a model to guide sampling in future eDNA studies.

5 | DISCUSSION AND RECOMMENDATIONS

eDNA studies have generally involved the assumption that DNA particles are well mixed in most aquatic ecosystems and that a small volume is sufficient to represent the entire system. However, both intra- and extracellular DNA fragments are distributed in the water medium, and the particle dynamics are determined by the hydrological conditions of the system (Carraro et al., 2021). Here, we used both a virtual ecologist approach and illustrative data sets to improve our understanding of eDNA sampling. Using an empirical data set, we showed how sampling parameters, including both the number of spatial replicates and the sampled volume, influence the species detection probability and the recovered number of species, suggesting that these axes of sampling parameters can help us to optimize local and/or regional detection of species.

Subjective methods and insufficient reporting of sampling procedures and strategies are a general issue in ecological science (Haddaway & Verhoeven, 2015; Smith et al., 2017), hindering generalizations and comparability. Given the recent emergence of the method, many existing eDNA studies have been exploratory, and often focused on a proof of concept. The number of studies using eDNA is rapidly increasing, however, and the method is starting to be used and applied by ecologists and biodiversity scientists in general, extending beyond the manageable number of laboratories

that pioneered the field. This widespread application requires clear sampling plans that are well suited to the study objectives, as well as transferable sampling methods. These methodological aspects are becoming critical, as a few national and global initiatives were recently launched with the aim to use eDNA for long-term biodiversity monitoring. As the number of research labs and commercial companies offering eDNA-based surveys is increasing, and given that these surveys will eventually feed into openly accessible databases for global analyses (Berry et al., 2021), we need protocols designed for the best quality, reliability and reproducibility of results.

We proposed criteria and justification with respect to the number of sample replicates and the total sampling volume. We strongly suggest that these sampling aspects be coherently justified, quantified and reported as much as possible. By formulating a specific equation that integrates these components, the factors affecting reproducibility can be considered more effectively. Unfortunately, such information has not been consistently reported in existing studies. In comparisons of diversity assessments conducted with traditional and with DNA-based methods, the heterogeneity in the completeness of the methodological details reported is still very high, often prohibiting a coherent meta-analysis across studies (Keck et al., 2022). Hence, whenever new sampling methodologies are developed or applied in an ecosystem for the first time, we recommend that a set of easily applied tests be run to demonstrate the methodological efficiency. These should include species accumulation curves to illustrate species detection at different sampling volumes and species recovered in relation to sequencing depth.

The wide range of approaches currently used in eDNA studies calls for the development of distinct standards for eDNA collection and analysis for different objectives and systems, but these standards should be properly justified and documented. The measurement and verification of biodiversity change, with the aim to inform policies and new green investments, must be based on reliable and repeatable measurements to support robust decisions. Only by ensuring that sufficiently replicated and representative samples are collected (requiring e.g., sufficient sampling depth, sample

10 | **WILEY**— Environmental DNA
Dedicated to the study and use of environmental DNA for basic and applied sciences
Open Access

ALTERMATT ET AL.

replication, and sample stratification) can eDNA be seen as a robust method for biodiversity estimates, capable of withstanding critical evaluation.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The data set of Polanco Fernández, Marques, et al. (2021) is open access and available with that publication. The data from the EVHOE bottom trawl research survey is published under the link https://doi.org/10.6084/m9.figshare.22592932.v1. The data for British Colombia is available with the publication by Lyet et al. (2021).

ORCID

Florian Altermatt https://orcid.org/0000-0002-4831-6958
Luca Carraro https://orcid.org/0000-0002-3933-1144
Camille Albouy https://orcid.org/0000-0003-1629-2389
Yan Zhang https://orcid.org/0000-0003-0688-5359
Xiaowei Zhang https://orcid.org/0000-0001-8974-9963
Loïc Pellissier https://orcid.org/0000-0002-2289-8259

REFERENCES

Alexander, T., & Seehausen, O. (2021). *Diversity, distribution and community composition of fish in perialpine lakes. "Projet lac" synthesis report*. Das Wasserforschungsinstitut des ETH-Bereichs.

Altermatt, F., Little, C. J., Mächler, E., Wang, S., Zhang, X., & Blackman, R. C. (2020). Uncovering the complete biodiversity structure in spatial networks: The example of riverine systems. *Oikos*, *129*(5), 607–618.

Barbour, M. T. (1999). *Rapid bioassessment protocols for use in wadeable streams and rivers: Periphyton, benthic macroinvertebrates and fish*. US Environmental Protection Agency, Office of Water.

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, *17*(1), 1–17.

Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., & Lodge, D. M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science and Technology*, *48*(3), 1819–1827.

Barral, M. P., Rey Benayas, J. M., Meli, P., & Maceira, N. O. (2015). Quantifying the impacts of ecological restoration on biodiversity and ecosystem services in agroecosystems: A global meta-analysis. *Agriculture, Ecosystems and Environment*, *202*, 223–231.

Berry, O., Jarman, S., Bissett, A., Hope, M., Paeper, C., Bessey, C., Schwartz, M. K., Hale, J., & Bunce, M. (2021). Making environmental DNA (eDNA) biodiversity records globally accessible. *Environmental DNA*, *3*(4), 699–705.

Bessey, C., Jarman, S. N., Berry, O., Olsen, Y. S., Bunce, M., Simpson, T., Power, M., McLaughlin, J., Edgar, G. J., & Keesing, J. (2020). Maximizing fish detection with eDNA metabarcoding. *Environmental DNA*, *2*(4), 493–504.

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.

Broadhurst, H. A., Gregory, L. M., Bleakley, E. K., Perkins, J. C., Lavin, J. V., Bolton, P., Browett, S. S., Howe, C. V., Singleton, N., Tansley, D., Sales, N. G., & McDevitt, A. D. (2021). Mapping differences in mammalian distributions and diversity using environmental DNA from rivers. *Science of the Total Environment*, *801*, 149724.

Bruce, K., Blackman, R. C., Bourlat, S. J., Hellström, M., Bakker, J., Bista, I., Bohmann, K., Bouchez, A., Brys, R., Clark, K., Elbrecht, V., Fazi, S., Fonseca, V. G., Hänfling, B., Leese, F., Mächler, E., Mahon, A. R., Meissner, K., Panksep, K., … Deiner, K. (2021). *A practical guide to DNA-based methods for biodiversity assessment*. Pensoft Publishers.

Burian, A., Mauvisseau, Q., Bulling, M., Domisch, S., Qian, S., & Sweet, M. (2021). Improving the reliability of eDNA data interpretation. *Molecular Ecology Resources*, *21*(5), 1422–1433.

Carraro, L., Blackman, R. C., & Altermatt, F. (2022). Modelling eDNA transport in river networks reveals highly resolved spatiotemporal patterns of freshwater biodiversity. *bioRxiv*. https://doi.org/10.1101/2022.01.25.475970

Carraro, L., Hartikainen, H., Jokela, J., Bertuzzo, E., & Rinaldo, A. (2018). Estimating species distribution and abundance in river networks using environmental DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(46), 11724–11729.

Carraro, L., Mächler, E., Wüthrich, R., & Altermatt, F. (2020). Environmental DNA allows upscaling spatial patterns of biodiversity in freshwater ecosystems. *Nature Communications*, *11*(1), 3585.

Carraro, L., Stauffer, J. B., & Altermatt, F. (2021). How to design optimal eDNA sampling strategies for biomonitoring in river networks. *Environmental DNA*, *3*(1), 157–172.

Carvalho, S. B., Gonçalves, J., Guisan, A., & Honrado, J. P. (2016). Systematic site selection for multispecies monitoring networks. *Journal of Applied Ecology*, *53*(5), 1305–1316.

Cermeño, P., Teixeira, I. G., Branco, M., Figueiras, F. G., & Marañón, E. (2014). Sampling the limits of species richness in marine phytoplankton communities. *Journal of Plankton Research*, *36*(4), 1135–1139.

Cordier, T., Angeles, I. B., Henry, N., Lejzerowicz, F., Berney, C., Morard, R., Brandt, A., Cambon-Bonavita, M. A., Guidi, L., Lombard, F., Arbizu, P. M., Massana, R., Orejas, C., Poulain, J., Smith, C. R., Wincker, P., Arnaud-Haond, S., Gooday, A. J., de Vargas, C., & Pawlowski, J. (2022). Patterns of eukaryotic diversity from the surface to the deep-ocean sediment. *Science Advances*, *8*(5), eabj9309.

Coutant, O., Cantera, I., Cilleros, K., Dejean, T., Valentini, A., Murienne, J., & Brosse, S. (2021). Detecting fish assemblages with environmental DNA: Does protocol matter? Testing eDNA metabarcoding method robustness. *Environmental DNA*, *3*(3), 619–630.

Cristescu, M. E., & Hebert, P. D. (2018). Uses and misuses of environmental DNA in biodiversity science and conservation. *Annual Review of Ecology, Evolution, and Systematics*, *49*, 209–230.

Deiner, K., & Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. *PLoS One*, *9*(2), e88786.

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(21), 5872–5895.

Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J.-C., & Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nature Communications*, 7(1), 1–9.

Dickie, I. A., Boyer, S., Buckley, H. L., Duncan, R. P., Gardner, P. P., Hogg, I. D., Holdaway, R. J., Lear, G., Makiola, A., Morales, S. E., Powell, J. R., & Weaver, L. (2018). Towards robust and repeatable sampling methods in eDNA-based studies. *Molecular Ecology Resources*, 18(5), 940–952.

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

Fukaya, K., Murakami, H., Yoon, S., Minami, K., Osada, Y., Yamamoto, S., Masuda, R., Kasai, A., Miyashita, K., Minamoto, T., & Kondoh, M. (2021). Estimating fish population abundance by integrating quantitative data on environmental DNA and hydrodynamic modelling. *Molecular Ecology*, 30(13), 3057–3067.

Gotelli, N. J., & Colwell, R. K. (2001). Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 4(4), 379–391.

Grabherr, G., Gottfried, M., & Pauli, H. (2000). GLORIA: A global observation research initiative in alpine environments. *Mountain Research and Development*, 20(2), 190–191.

Haddaway, N. R., & Verhoeven, J. T. (2015). Poor methodological detail precludes experimental repeatability and hampers synthesis in ecology. *Ecology and Evolution*, 5(19), 4451–4454.

Hänfling, B., Lawson Handley, L., Read, D. S., Hahn, C., Li, J., Nichols, P., Blackman, R. C., Oliver, A., & Winfield, I. J. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, 25(13), 3101–3119.

Harrison, J. B., Sunday, J. M., & Rogers, S. M. (2019). Predicting the fate of eDNA in the environment and implications for studying biodiversity. *Proceedings of the Royal Society B*, 286(1915), 20191409.

Hirzel, A., & Guisan, A. (2002). Which is the optimal sampling strategy for habitat suitability modelling? *Ecological Modelling*, 157(2–3), 331–341.

Keck, F., Blackman, R. C., Bossart, R., Brantschen, J., Couton, M., Hürlemann, S., Kirschner, D., Locher, N., Zhang, H., & Altermatt, F. (2022). Meta-analysis shows both congruence and complementarity of DNA and eDNA metabarcoding to traditional methods for biological community assessment. *Molecular Ecology*, 31(6), 1820–1835.

Kery, M., & Plattner, M. (2007). Species richness estimation and determinants of species detectability in butterfly monitoring programmes. *Ecological Entomology*, 32(1), 53–61.

Laporte, M., Bougas, B., Côté, G., Champoux, O., Paradis, Y., Morin, J., & Bernatchez, L. (2020). Caged fish experiment and hydrodynamic bidimensional modeling highlight the importance to consider 2D dispersion in fluvial environmental DNA studies. *Environmental DNA*, 2(3), 362–372.

Lawson Handley, L. (2015). How will the 'molecular revolution' contribute to biological recording? *Biological Journal of the Linnean Society*, 115(3), 750–766.

Lengyel, S., Kosztyi, B., Schmeller, D. S., Henry, P.-Y., Kotarac, M., Lin, Y.-P., & Henle, K. (2018). Evaluating and benchmarking biodiversity monitoring: Metadata-based indicators for sampling design, sampling effort and data analysis. *Ecological Indicators*, 85, 624–633.

Leopold, L. B., & Maddock, T. (1953). *The hydraulic geometry of stream channels and some physiographic implications*. Technical Report Geological Survey Professional Paper 252, Washington DC.

Leverington, F., Costa, K. L., Pavese, H., Lisle, A., & Hockings, M. (2010). A global analysis of protected area management effectiveness. *Environmental Management*, 46(5), 685–698.

Li, F., Altermatt, F., Yang, J., An, S., Li, A., & Zhang, X. (2020). Human activities' fingerprint on multitrophic biodiversity and ecosystem functions across a major river catchment in China. *Global Change Biology*, 26(12), 6867–6879.

Lyet, A., Pellissier, L., Valentini, A., Dejean, T., Hehmeyer, A., & Naidoo, R. (2021). eDNA sampled from stream networks correlates with camera trap detection rates of terrestrial mammals. *Scientific Reports*, 11(1), 11362.

Macher, T., Schütz, R., Arle, J., Beermann, A. J., Koschorreck, J., & Leese, F. (2021). Beyond fish eDNA metabarcoding: Field replicates disproportionately improve the detection of stream associated vertebrate species. *Metabarcoding and Metagenomics*, 5, 59–71.

Mächler, E., Walser, J.-C., & Altermatt, F. (2021). Decision making and best practices for taxonomy-free eDNA metabarcoding in biomonitoring using hill numbers. *Molecular Ecology*, 30(13), 3326–3339.

Mansfeldt, C., Deiner, K., Mächler, E., Fenner, K., Eggen, R. I. L., Stamm, C., Schönenberger, U., Walser, J. C., & Altermatt, F. (2020). Microbial community shifts in streams receiving treated wastewater effluent. *Science of the Total Environment*, 709, 135727.

Nester, G. M., De Brauwer, M., Koziol, A., West, K. M., DiBattista, J. D., White, N. E., Power, M., Heydenrych, M. J., Harvey, E., & Bunce, M. (2020). Development and evaluation of fish eDNA metabarcoding assays facilitate the detection of cryptic seahorse taxa (family: Syngnathidae). *Environmental DNA*, 2(4), 614–626.

O'Connell, A. F., Nichols, J. D., & Karanth, K. U. (2011). *Camera traps in animal ecology: Methods and analyses* (Vol. 271). Springer.

Pawlowski, J., Apothéloz-Perret-Gentil, L., & Altermatt, F. (2020). Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. *Molecular Ecology*, 29(22), 4258–4264.

Pawlowski, J., Apothéloz-Perret-Gentil, L., & Altermatt, F. (2021). Environmental versus extra-organismal dna. *Molecular Ecology*, 30(19), 4606–4607.

Pawlowski, J., Apothéloz-Perret-Gentil, L., Mächler, E., & Altermatt, F. (2020). *Environmental DNA applications for biomonitoring and bioassessment in aquatic ecosystems*. Federal Office for the Environment (FOEN/BAFU), Environmental studies. https://doi.org/10.5167/uzh-187800

Pellissier, L., Niculita-Hirzel, H., Dubuis, A., Pagni, M., Guex, N., Ndiribe, C., Salamin, N., Xenarios, I., Goudet, J., Sanders, I., & Guisan, A. (2014). Soil fungal communities of grasslands are environmentally structured at a regional scale in the Alps. *Molecular Ecology*, 23(17), 4274–4290.

Polanco Fernández, A., Marques, V., Fopp, F., Juhel, J.-B., Borrero-Pérez, G. H., Cheutin, M.-C., Dejean, T., González Corredor, J. D., Acosta-Chaparro, A., Hocdé, R., Eme, D., Maire, E., Spescha, M., Valentini, A., Manel, S., Mouillot, D., Albouy, C., & Pellissier, L. (2021). Comparing environmental DNA metabarcoding and underwater visual census to monitor tropical reef fishes. *Environmental DNA*, 3(1), 142–156.

Polanco Fernández, A., Mutis Martinezguerra, M., Marques, V., Villa-Navarro, F., Borrero Perez, G. H., Cheutin, M.-C., Dejean, T., Hocdé, R., Juhel, J.-B., Maire, E., Manel, S., Spescha, M., Valentini, A., Mouillot, D., Albouy, C., & Pellissier, L. (2021). Detecting aquatic and terrestrial biodiversity in a tropical estuary using environmental DNA. *Biotropica*, 53(6), 1606–1619.

Roberts, T. E., Bridge, T. C., Caley, M. J., & Baird, A. H. (2016). The point count transect method for estimates of biodiversity on coral reefs: Improving the sampling of rare species. *PLoS One*, 11(3), e0152335.

Rodgers, T. W., & Mock, K. E. (2015). Drinking water as a source of environmental DNA for the detection of terrestrial wildlife species. *Conservation Genetics Resources*, 7(3), 693–696.

Rodriguez-Ramos, T., Dornelas, M., Marañón, E., & Cermeño, P. (2014). Conventional sampling methods severely underestimate phytoplankton species richness. *Journal of Plankton Research*, 36(2), 334–343.

12 | **WILEY**—Environmental DNA
Dedicated to the study and use of environmental DNA for basic and applied sciences
Open Access

ALTERMATT ET AL.

Rodriguez-Ezpeleta, N., Morissette, O., Bean, C. W., Manu, S., Banerjee, P., Lacoursière-Roussel, A., Beng, K. C., Alter, S. E., Roger, F., Holman, L. E., Stewart, K. A., Monaghan, M. T., Mauvisseau, Q., Mirimin, L., Wangensteen, O. S., Antognazza, C. M., Helyar, S. J., de Boer, H., Monchamp, M.-E., … Deiner, K. (2021). Trade-offs between reducing complex terminology and producing accurate interpretations from environmental dna: Comment on "environmental dna: What's behind the term?" by Pawlowski et al. *Molecular Ecology*, 30(19), 4601–4605.

Ruppert, K. M., Kline, R. J., & Rahman, M. S. (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.

Schmeller, D. S., Böhm, M., Arvanitidis, C., Barber-Meyer, S., Brummitt, N., Chandler, M., Chatzinikolaou, E., Costello, M. J., Ding, H., García-Moreno, J., Gill, M., Haase, P., Jones, M., Juillard, R., Magnusson, W. E., Martin, C. S., McGeoch, M., Mihoub, J. B., Pettorelli, N., … Belnap, J. (2017). Building capacity in biodiversity monitoring at the global scale. *Biodiversity and Conservation*, 26(12), 2765–2790.

Shogren, A. J., Tank, J. L., Andruszkiewicz, E., Olds, B., Mahon, A. R., Jerde, C. L., & Bolster, D. (2017). Controls on eDNA movement in streams: Transport, retention, and resuspension. *Scientific Reports*, 7(1), 5065.

Smith, A. N. H., Anderson, M. J., & Pawley, M. D. M. (2017). Could ecologists be more random? Straightforward alternatives to haphazard spatial sampling. *Ecography*, 40(11), 1251–1255.

Spens, J., Evans, A. R., Halfmaerten, D., Knudsen, S. W., Sengupta, M. E., Mak, S. S. T., Sigsgaard, E. E., & Hellström, M. (2017). Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: Advantage of enclosed filter. *Methods in Ecology and Evolution*, 8(5), 635–645.

Stauffer, S., Jucker, M., Keggin, T., Marques, V., Andrello, M., Bessudo, S., Cheutin, M. C., Borrero-Pérez, G. H., Richards, E., Dejean, T., Hocdé, R., Juhel, J. B., Ladino, F., Letessier, T. B., Loiseau, N., Maire, E., Mouillot, D., Mutis Martinezguerra, M., Manel, S., … Waldock, C. (2021). How many replicates to accurately estimate fish biodiversity using environmental DNA on coral reefs? *Ecology and Evolution*, 11(21), 14630–14643.

Stewart, L., Simonsen, C. E., Svenning, J. C., Schmidt, N. M., & Pellissier, L. (2018). Forecasted homogenization of high Arctic vegetation communities under climate change. *Journal of Biogeography*, 45(11), 2576–2587.

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, 21(8), 1789–1793.

Thalinger, B., Kirschner, D., Pütz, Y., Moritz, C., Schwarzenberger, R., Wanzenböck, J., & Traugott, M. (2021). Lateral and longitudinal fish environmental DNA distribution in dynamic riverine habitats. *Environmental DNA*, 3(1), 305–318.

Ushio, M., Fukuda, H., Inoue, T., Makoto, K., Kishida, O., Sato, K., Murata, K., Nikaido, M., Sado, T., Sato, Y., Takeshita, M., Iwasaki, W., Yamanaka, H., Kondoh, M., & Miya, M. (2017). Environmental DNA enables detection of terrestrial mammals from forest pond water. *Molecular Ecology Resources*, 17(6), e63–e75.

Vanzo, D., Peter, S., Vonwiller, L., Bürgler, M., Weberndorfer, M., Siviglia, A., Conde, D., & Vetsch, D. F. (2021). BASEMENT v3: A modular freeware for river process modelling over multiple computational backends. *Environmental Modelling and Software*, 143, 105102.

Wilcox, T. M., Carim, K. J., Young, M. K., McKelvey, K. S., Franklin, T. W., & Schwartz, M. K. (2018). Comment: The importance of sound methodology in environmental DNA sampling. *North American Journal of Fisheries Management*, 38(3), 592–596.

Zhang, H., Mächler, E., Morsdorf, F., Niklaus, P. A., Schaepman, M. E., & Altermatt, F. (2023). A spatial fingerprint of land-water linkage of biodiversity uncovered by remote sensing and environmental dna. *Science of the Total Environment*, 867, 161365.

Zhang, J., Nielsen, S. E., Grainger, T. N., Kohler, M., Chipchar, T., & Farr, D. R. (2014). Sampling plant diversity and rarity at landscape scales: Importance of sampling time in species detectability. *PLoS One*, 9(4), e95334.

Zurell, D., Berger, U., Cabral, J. S., Jeltsch, F., Meynard, C. N., Münkemüller, T., Nehrbass, N., Pagel, J., Reineking, B., Schröder, B., & Grimm, V. (2010). The virtual ecologist approach: Simulating data and observers. *Oikos*, 119(4), 622–635.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.