*Environmental DNA*

WILEY

Environmental DNA
Open Access
Dedicated to the study and use of environmental DNA for basic and applied sciences

PERSPECTIVE OPEN ACCESS

# A Framework to Unify the Relationship Between Numerical Abundance, Biomass, and Environmental DNA

Matthew C. Yates[1] | Taylor M. Wilcox[2] | Shannon Kay[3] | Pedro Peres-Neto[4] | Daniel D. Heath[1]

[1]Great Lakes Institute for Environmental Research, University of Windsor, Windsor, Ontario, Canada | [2]National Genomics Center for Wildlife and Fish Conservation, Rocky Mountain Research Station, Missoula, Montana, USA | [3]Rocky Mountain Research Station, Fort Collins, Colorado, USA | [4]Department of Biology, Concordia University, Montreal, Quebec, Canada

**Correspondence:** Matthew C. Yates (matthew.yates@outlook.com)

**Received:** 10 August 2024 | **Revised:** 6 January 2025 | **Accepted:** 6 February 2025

**Funding:** This work was supported by Natural Sciences and Engineering Research Council of Canada.

**Keywords:** abundance | allometric scaling | allometry | biomass | eDNA | fishes | metabolic theory | *N* | population size

## ABSTRACT

Does environmental DNA (eDNA) concentration correlate with numerical abundance (*N*) or biomass in aquatic organisms? We hypothesize that eDNA can be adjusted to simultaneously reflect both. Building on frameworks developed from the Metabolic Theory of Ecology, we derive two equations to adjust eDNA data to simultaneously reflect both *N* and biomass using population size structure data and allometric scaling coefficients. We also demonstrate that these equations share model parameters, necessitating the joint estimation of regressions between adjusted eDNA, *N*, and biomass. Furthermore, our framework can be extended to model how other variables (temperature, taxa, diet, trophic level, etc.) might impact relationships between eDNA, *N*, and biomass in natural ecosystems. We applied our framework to data from two previously published studies correlating eDNA to Brook Trout (*Salvelinus fontinalis*) *N* and biomass. In both case studies, point estimates of the scaling coefficient (*b*) reflected allometric processes ($b = 0.51$ and 0.37 for Case Study 1 and 2, respectively), with credible intervals indicating that b likely differed from zero (i.e., eDNA scales with *N*) and one (i.e., eDNA scales with biomass). Directly estimating the value of b improved estimates of *N* and biomass relative to assuming b equals 0, which particularly affected the capacity to estimate biomass. However, models assuming eDNA production scaled with biomass (i.e., $b = 1$) were largely similar to estimating *b*, implying that assuming eDNA scales linearly with biomass might be a sufficient approximation for some systems. Nevertheless, the framework demonstrates that correlating eDNA directly with either *N* or biomass (as is commonly done in many studies) inherently necessitates an adjustment to infer the other metric if populations exhibit size structure variation. Collectively, we demonstrate that quantitative eDNA data is unlikely to correspond exactly to either population *N* or biomass but can be adjusted to simultaneously reflect both.

**JEL Classification:** BiologicalSciences, Ecology

## 1 | Introduction

Environmental DNA ("eDNA," referring to DNA found in environmental mediums like air, soil, or water) sampling is a cost-efficient and highly sensitive method for detecting and identifying aquatic species, often outperforming traditional sampling techniques (Boivin-Delisle et al. 2021; Evans et al. 2017; Sard et al. 2019). The extent to which quantitative eDNA data (e.g., concentration in the environment) can provide inference regarding the "unseen" numerical abundance (*N*) or biomass of organisms within an environment is less clear and remains an intensive line of research (Rourke et al. 2022;

Yates, Cristescu, et al. 2021). Empirical studies typically show positive correlations between the concentration of organisms' DNA in an environment and their abundance (which we define as either $N$ or biomass), but the extent to which observed eDNA concentrations will more closely reflect $N$ or biomass remains a major outstanding question (Yates et al. 2019).

Numerical abundance and biomass represent two linked, but distinct parameters. Although they are often both highly correlated with eDNA concentration data (Yates et al. 2019), they can also exhibit subtle variations in their relationships with eDNA (Baldigo et al. 2017; Lacoursière-Roussel et al. 2016; Pilliod et al. 2013; Stoeckle et al. 2021; Yates, Glaser, et al. 2021). This variation likely arises because eDNA production may not directly correlate with either $N$ or biomass. Instead, it may scale allometrically; that is, mass-specific eDNA production rates may tend to decline as organism mass increases (Maruyama et al. 2014; Stoeckle et al. 2021; Takeuchi et al. 2019; Yates, Glaser, et al. 2021). As such, improving the associations between $N$, biomass, and quantitative eDNA data should be achieved by modeling eDNA production using classic allometric frameworks derived from the Metabolic Theory of Ecology (MTE) (Brown et al. 2004). Specifically, eDNA production might be better represented as a function of the individual organism's mass raised to the power of an allometric scaling coefficient ("$b$") multiplied by a normalization constant (Brown et al. 2004; Yates, Glaser, et al. 2021).

Numerous studies have demonstrated that integrating allometry into metrics of organism numerical abundance can strengthen correlations between quantitative eDNA concentration data and population/species abundance in natural ecosystems (Skelton et al. 2023; Stoeckle et al. 2021; Yates et al. 2023; Yates, Glaser, et al. 2021). It should be noted, however, that these studies have incorporated allometry by applying data adjustments to population-level organism abundance variables. Specifically, they express organism abundance as the sum of individual mass values within a population raised to the power of an allometric scaling coefficient $b$, with the value of $b$ depending on the rate at which eDNA production scales with individual body size (Skelton et al. 2023; Stoeckle et al. 2021; Yates et al. 2023; Yates, Glaser, et al. 2021).

Metabolic scaling transformations reflect the "direction" of functional relationships between eDNA and organism abundance (e.g., eDNA is a product of allometrically adjusted biomass). However, to improve the estimation of unobserved (unmeasured) $N$ and biomass from quantitative eDNA data, such adjustments must instead be applied to eDNA data (Levi et al. 2019). We hypothesize that the allometric scaling parameter represents the fundamental parameter that links population-level eDNA production to both $N$ and biomass by accounting for the effect of size structure on eDNA production. Therefore, we propose that predicting both $N$ and biomass from quantitative eDNA data necessitates the derivation of a framework that simultaneously models both $N$ and biomass from quantitative eDNA data (e.g., $N = f(\text{eDNA})$ and $\text{biomass} = g(\text{eDNA})$). This framework can then be used to adjust quantitative eDNA data based on species or population size structure data, allowing the simultaneous inference of both $N$ and biomass as distinct yet linked (by size structure) parameters, thus "unifying" the inference of $N$ and biomass from eDNA.

We propose the framework described in Figure 1, derived from the assumption that the individual-level eDNA production rate follows the same allometric relationship with body mass as other key metabolically linked rates (Brown et al. 2004; Vanni and McIntyre 2016) and body surface area (O'Shea et al. 2006) in animals:

$$I = I_0 \times M^b \tag{1}$$

where $I$ = individual eDNA production rate per unit of time, $M$ = the individual body mass of an organism, $I_0$ = an eDNA production rate normalization constant, and $b$ = an allometric scaling coefficient.

If eDNA production can be approximated using Equation (1), an aggregate "population"-level of eDNA production per unit of time can be expressed using the following formula:

$$\text{eDNA}_i = I_0 \times \sum_j^{N_i} x_{ij}^b \tag{2a}$$

where $\text{eDNA}_i$ is the aggregate eDNA production rate for the $i$th population, $\sum_j^{N_i} x_{ij}^b$ is equal to the sum of the $j$th individual mass values raised to the power of the allometric scaling coefficient, $b$, for the $i$th population (Yates, Glaser, et al. 2021).

In limnology, typical mass-balance approaches to modeling particle concentrations in lentic ecosystems integrate particle input, particle transport, and particle decay (Kalff 2001):

$$P = \frac{M/V}{\frac{Q}{V} + \sigma}$$

where $P$ equals the particle concentration, $M$ equals the quantity input of the particle per unit time, $V$ equals the waterbody volume, $Q$ equals the transportation rate per unit time, and $\sigma$ equals the particle decay rate per unit time. We note that the

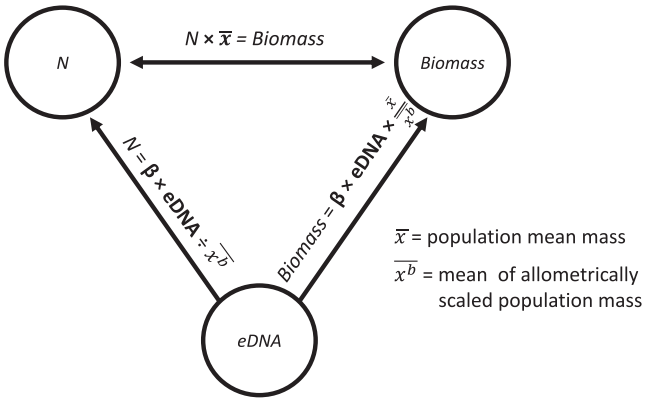


**FIGURE 1** | Conceptual diagram outlining the derived framework. $N$ and biomass can each be individually derived by adjusting quantitative eDNA data according to the equations displayed. However, derivations of $N$ and biomass should satisfy the condition that biomass is equal to $N$ multiplied by the mean mass of individuals within a population.

model can be further parameterized to account for waterbody volume, eDNA particle transportation, and eDNA particle decay (if data are available), as has been parameterized in several studies (e.g., Jo et al. 2019; Sassoubre et al. 2016; Beaulieu et al. 2024). Although these factors can be integrated into eDNA/abundance/biomass models, our primary focus is on unifying the relationships between eDNA concentrations, *N*, and biomass. Therefore, in this model description, we assume that system-specific parameters for DNA in the environment (*V*, *Q*, and $\sigma$) remain constant across populations, and we concentrate on the input parameter, *M*. With this simplification, eDNA steady-state concentration are directly proportional to eDNA production and, assuming a constant waterbody volume, our model can be simplified to:

$$\mathrm{eDNA}_i = J_0 \times \sum_j^{N_i} x_{ij}^b \tag{2b}$$

where $\mathrm{eDNA}_i$ represents the steady-state concentration of eDNA in population *i*, and $J_0$ represents the eDNA production rate constant $(I_0)$ divided by transportation and decay rates. The simplified Equation (2b) can alternatively be expressed as:

$$\mathrm{eDNA}_i = J_0 \times N_i \times \overline{x_i^b} \tag{3}$$

where $N_i$ is the numerical abundance of the *ith* "population" and $\overline{x_i^b}$ represents the mean of allometrically scaled individual mass values for the *ith* population. Equation (3) can be reformulated to express population *N* as a function of $\mathrm{eDNA}_i$ divided by $\overline{x_i^b}$ of the *i*th population as:

$$N = f(\mathrm{eDNA}): N_i = \frac{1}{J_0}\, eDNA_i \div \overline{x_i^b} \tag{4}$$

Similarly, Equation (5) can be used to estimate population biomass as a function of eDNA, population mean mass ($\bar{x}_i$), and an allometric scaling coefficient. Given that:

$$\mathrm{Biomass}_i = N_i \times \bar{x}_i \tag{5}$$

we can then substitute Equation (4) for $N_i$ and rearrange it to obtain:

$$\mathrm{Biomass} = g(\mathrm{eDNA}): \mathrm{Biomass}_i = \frac{1}{J_0}\, eDNA_i \times \frac{\bar{x}_i}{\overline{x_i^b}} \tag{6}$$

From a mathematical perspective, note that *N* and biomass are correlated metrics, as established in Equation (5). Therefore, any quantitative framework estimating or deriving both metrics from quantitative eDNA data must satisfy this correlation condition. Given that the parameter estimates for $J_0$ and *b* are shared between Equations (4) and (6), they must be estimated *jointly* across both regressions to ensure consistency.

Below, we describe the application of the above framework to two previously published datasets studying Brook Trout (*Salvelinus fontinalis*) inhabiting lentic systems (see below), including a discussion of the results and their implications for the application of eDNA to monitor abundance and biomass. However, metabolic theory and the literature on eDNA dynamics represent a rich foundation on which the above framework can be expanded to model such relationships. Before demonstrating the empirical application of the framework on Case Study datasets, we present a series of testable predictions based on our framework for future research. If validated, the predictions offer a quantitative, conceptual, and empirical roadmap to advance the future application of eDNA in monitoring both *N* and biomass across taxonomic groups and trophic levels in natural ecosystems.

## 2 | Expanding the Framework: Predictions From Metabolic Theory

The equations herein to unify the relationships between eDNA, *N*, and biomass were derived from frameworks originally developed to model ontogenetic changes in physiological and metabolic processes and form the basis of the broader Metabolic Theory of Ecology (Brown et al. 2004). While this work focused on unifying relationships between eDNA, *N*, and biomass, metabolic theory represents a rich and well-developed body of work that can be readily extended to model the impact of variables beyond allometric scaling parameters on relationships between eDNA and abundance metrics. Below, we highlight key predictions that can be made as extensions of broader metabolic theory and eDNA dynamics to model eDNA production; our goal is to move the study and interpretation of quantitative eDNA signals forward by providing a set of testable predictions to validate the extension of this framework to the interpretation of quantitative eDNA signals.

### 2.1 | Bioenergetics: Integrating Temperature Dependency Functions ($f(T)$) Into Model Predictions

Classic bioenergetics frameworks that model key physiological rates like consumption (shown below), excretion, and egestion often take the form of:

$$\mathrm{Consumption} = \mathrm{I}_0 \times \mathrm{M}^b \times f(T)$$

where $f(T)$ equals a temperature dependency function (Hanson et al. 1997). Temperature is a key parameter influencing eDNA pseudo-steady-state concentrations (Caza-Allard et al. 2021; Jo et al. 2019); such frameworks could likely be extended to model the broad effect of temperature on quantitative eDNA data recovered from natural ecosystems (Yates, Cristescu, et al. 2021; Beaulieu et al. 2024). The $J_0$ term represents a parameter which describes, for a given size of fish, the expected pseudo-steady-state concentration of eDNA (e.g., how much eDNA a single 100 g fish would produce under given conditions) and would vary across ecosystems depending on the value of different parameters affecting eDNA production. The $f(T)$ would traditionally apply a transformation to $I_0$ or $J_0$ (Stewart and Ibarra 1991) and, when applied through our framework, the associated adjustment would represent the reciprocal of $f(T)$ and should be a shared parameter across both the *N* and biomass models.

No studies, to date, have examined eDNA production across a wide enough temperature gradient to characterize the shape of the eDNA $f(T)$. However, we would expect it to follow a similar form as other temperature-dependency functions associated with physiological processes related to eDNA production, such as consumption (Stewart and Ibarra 1991), with observed eDNA concentrations decreasing at lower temperatures (Caza-Allard et al. 2021; Jo et al. 2019). In the absence of controlled experiments, it may also be possible to broadly derive $f(T)$ in natural ecosystems if data from a sufficient natural gradient in temperature are collected.

### 2.2 | Breakdown of Allometric Relationship During Spawning

Most Teleost fish and many other aquatic taxonomic organisms (e.g., most Anurans) exhibit external fertilization, in which gametes are broadly released and fertilization occurs outside of the female organism's body. The release of gametes (including ovarian and seminal fluid) is likely to result in large quantities of eDNA in the water column (Collins et al. 2022; Tsuji and Shibata 2021). Additionally, increased activity can result in additional shedding of eDNA (Thalinger et al. 2021); increased activity or aggressive behavior associated with mate competition could result in increased eDNA production. For semelparous species, eDNA is also likely to increase substantially as a result of the decay of corpses following spawning (Tillotson et al. 2018). Collectively, spawning is likely to result in an increase in mass-specific eDNA production rates for reproducing individuals, which are typically larger bodied than nonreproductive individuals. As a result, the value of $b$ is likely to increase during reproductive periods, potentially even surpassing 1 (e.g., larger reproductive individuals have higher mass-specific eDNA production rates than smaller immature individuals).

### 2.3 | Expanding to InterSpecific Relationships: $I_0$ (and Thus eDNA/Biomass Regression Slopes) Will Likely Reflect Metabolic Rates and Surface Area Characteristics Across Macroscopic Taxa

Metabolic theory can be applied to model allometry in physiological and metabolic processes on both an *intra-* and *inter-specific* basis (Allegier et al. 2015; Jerde et al. 2019; Vanni and McIntyre 2016). We note that taxonomic variation in the value of $b$ exists for key physiological rates likely related to eDNA production (Allegier et al. 2015; Jerde et al. 2019; Vanni and McIntyre 2016). However, based on metabolic theory, we predict that we would also observe consistent differences in the regression slopes (moderated by $J_0$) for eDNA production between taxonomic groups based on both a combination of levels of metabolic activity and external surface area characteristics (Figure 2).

Turtles, for example, are expected to have low mass-specific eDNA production rates due to their low metabolic activity and hard external surfaces (Adams et al. 2019); this would manifest in a lower $I_0$ and/or $J_0$ (and thus higher $\beta$ slope values, given the slope is equivalent to $\frac{1}{J_0}$) relative to other large-bodied

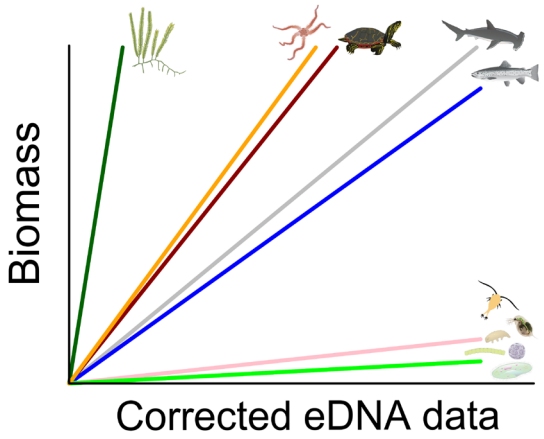


**FIGURE 2** | Conceptual figure describing predicted trends across taxonomic groups. We would expect slope values between allometrically adjusted eDNA–biomass relationships to vary based on metabolic activity and trophic level. Plants would be expected to produce very little eDNA relative to biomass. Metazoans with low metabolic and activity rates would also likely have high slope values relative to active taxonomic groups like fishes. Small multicellular and single-celled organisms, which are captured whole body on eDNA filters, would have very low slope values. Illustrations by Tracy Saxby, Emily Nastase, Kim Kraeer, and Jane Thomas; Integration and Application Network, University of Maryland Center for Environmental Science (https://ian.umces.edu/media-library/).

macroorganisms (Figure 2). Depending on the relative importance of excretion/egestion and surface shedding for the production of eDNA, chitinous organisms may similarly produce less eDNA relative to nonchitinous groups due to possessing hard external surfaces versus an epidermal or mucous layer (Stewart 2019). In marine systems, for example, echinoderms (a metabolically slow taxonomic group with hard external surfaces) may also exhibit low values of $I_0$. In a study on eDNA derived from Antarctic water, researchers found poor recovery of echinoderm reads in metagenomic data despite echinoderms representing a diverse and high-abundance taxa (Cowart et al. 2018). While the authors note that this may be due to a general lack of reference genomes for this taxa, echinoderms are also characterized by low activity and resting metabolic rates relative to other metazoans (e.g., fish; Seibel and Drazen 2007). Aquatic plants and macroalgae, which are notoriously difficult to detect using eDNA and for whom weak relationships between eDNA and biomass are often found (Jo 2023; Kuehne et al. 2020), represent an even more extreme example. Identifying the taxonomic level at which $b$ and $I_0/J_0$ coefficients may vary across macroorganisms is a critical next step, with additional potential application to detection probability models.

### 2.4 | $I_0$ and "$b$" May Be Affected by Diet and Trophic Position Among Macroscopic Taxa

Feeding rate has been observed to have a positive effect on eDNA production (Klymus et al. 2015), and differences in the volume of fecal production among trophic groups have long been observed. Carnivores, for example, tend to eat higher quality food and have higher assimilation efficiencies; as a

result, they tend to produce smaller quantities of feces compared to herbivores and detritivores, which consume and egest much larger volumes of biomass as a proportion of their body mass (Mill et al. 2007; Wotton and Malmqvist 2001). eDNA production rates may thus be affected by diet composition and trophic level.

### 2.5 | Interactions Between eDNA Capture Methodology and Organism Body Size Will Moderate $I_0$ and "$b$" Across Multicellular Macro- and Microorganisms

Mechanisms of eDNA capture may have significant impacts on the distribution of $I_0$ and "$b$" among taxonomic groups and trophic levels, with particularly important implications for eDNA from plankton and microorganisms. We note that this framework is agnostic as to the source of the DNA found in environmental samples, that is, it is flexible enough to apply to both extraorganismal DNA from macroorganisms as well as eDNA derived from whole-body multicellular and unicellular microorganisms captured in environmental samples. For the purposes of this discussion, we will therefore broadly adopt the definition of eDNA as DNA from environmental samples derived from both organismal and extraorganismal sources, as defined in Pawlowski et al. (2020).

Scaling coefficient values for extraorganismal eDNA produced by macroorganisms (i.e., organisms too large to be captured directly on eDNA filtering media) would be expected to reflect metabolic/surface area allometry, as discussed above. However, for multicellular plankton (e.g., zooplankton) and multicellular microorganisms, most eDNA in samples would likely derive from captured whole organisms. Assuming that the cell density of an organism scales approximately linearly with organism biomass, eDNA derived from these organisms will thus likely scale approximately linearly with their biomass; we thus predict that the expected value of the scaling coefficient for these organisms would likely be close to 1 (Figure 3).

Similarly, multicellular microorganisms (e.g., plankton) captured whole-body in eDNA samples will likely produce substantial quantities of DNA, given that high-quality DNA from their entire body will be lysed and extracted during sample processing (as opposed to trace quantities of potentially partially degraded extraorganismal DNA from macroorganisms). As a result, we would predict substantially lower slope values for biomass/eDNA relationships (moderated by $I_0$) for microorganisms captured whole-body in eDNA samples (Figure 2). Identifying if, and at what approximate gradient of body size, eDNA shifts from being generated primarily by the excretion and/or shedding of extraorganismal DNA to predominantly whole-body DNA from captured organisms will be important to broadly interpret quantitative eDNA signals across trophic levels.

### 2.6 | $b = 0$ for Unicellular Organisms, and eDNA/Biomass Slope Values Will Vary Between Prokaryotes and Single-Celled Eukaryotes due to Differences in Mean Cell Size

Unicellular organisms, similar to multicellular microorganisms, are captured whole body/cell. However, as many microbial metagenomics studies have noted, eDNA copies should roughly scale with the number of individuals of a given species present in the sample (Nayfach and Pollard 2016), rather than with their biomass. For these organisms, we would thus expect that the eDNA present in the environmental sample would scale with a $b$-value close to 0 (i.e., with numerical abundance), rather than 1 (i.e., with biomass; Figure 3). Reads assigned to prokaryotes and single-celled eukaryotes, which are both captured "whole organism" in environmental samples, dominate metagenomic shotgun sequencing data, reflecting their high abundance in eDNA samples (Cowart et al. 2018; Stat et al. 2017). As with multicellular microorganisms, this would manifest in lower slope values for adjusted eDNA/biomass relationships (moderated by $I_0$) relative to macroorganisms (Figure 2). However, we also predict that slope values between adjusted eDNA data and biomass should be higher among multicellular plankton/microorganisms and protists relative to prokaryotes, as differences in

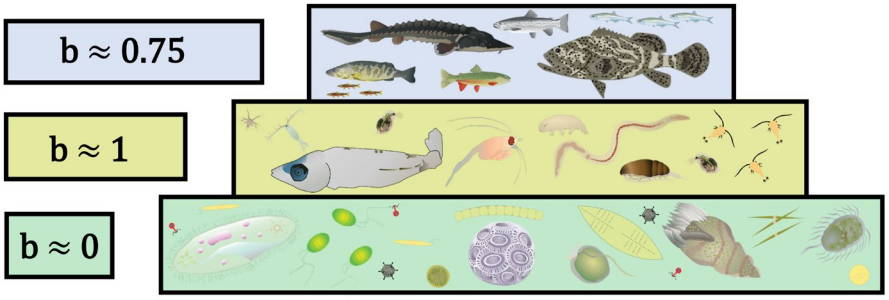


**FIGURE 3** | Allometric scaling coefficient is likely to vary based on the combination of the source of eDNA and trophic level. Environmental DNA captured in filters from large metazoans (e.g., fish, amphibians, large invertebrates, etc.) is likely to be extraorganismal, and thus scale with values similar to physiological rates and surface area (e.g., ~0.75; note that intraspecific metabolic scaling coefficients may be closer to 1, see Jerde et al. (2019)). Environmental DNA from small multicellular organisms may be predominantly derived from the capture of whole-bodied organisms on water filters, and thus eDNA is likely to scale with biomass (e.g., $b \sim 1$). Environmental DNA collected from single-celled organisms is likely to also be predominantly derived from whole-bodied organisms on water filters, but will scale with gene copies per cell (functionally numerical abundance, e.g., $b \sim 0$). Note that we define "environmental DNA" as any DNA (extraorganismal or organismal) captured in an environmental sample. Illustrations by Tracy Saxby, Emily Nastase, Kim Kraeer, and Jane Thomas; Integration and Application Network, University of Maryland Center for Environmental Science (https://ian.umces.edu/media-library/).

slopes between these groups may be largely influenced by mean cell size (Figure 2).

## 3 | Applying the Unified Framework to Empirical Datasets

Herein, we apply the quantitative framework described above in Equations (4), (5), and (6) to two previously published Brook Trout (*Salvelinus fontinalis*) datasets in freshwater lake systems: One where population size structure varied by an order of magnitude across populations (Yates, Glaser, et al. 2021) and another where it varied fourfold (Gaudet-Boulay et al. 2023). We used a Bayesian approach in which we jointly estimated models relating quantitative eDNA data to $N$ and biomass in these two case studies based on Equations (4) and (6), which share $\overline{x_i^b}$ and beta coefficient ($\frac{1}{J_0}$) parameters. While these two datasets lack data to empirically test further expansions of the framework described above, they serve as illustrative examples of how the framework can be applied to empirically estimate abundance and biomass, simultaneously, from quantitative eDNA data in conjunction with size–structure data.

## 4 | Materials and Methods

### 4.1 | Case Study 1: Brook Trout Abundance and eDNA in a System With Substantial Size Structure Variation Among Populations

Yates, Glaser, et al. (2021) examined the relationship between eDNA concentration and Brook Trout abundance in nine study populations in the Rocky Mountains, Canada (see methods described therein for a full description). In brief, Brook Trout abundance was assessed using fyke net capture-mark-recapture data, and population body size structure was assessed from standardized mixed-mesh index netting. The mean concentration of eDNA in each lake was assessed by collecting four 1 L littoral and four 1 L pelagic water samples systematically distributed throughout the study lakes. Water samples were filtered onsite, and qPCR (Wilcox et al. 2016) was used to assess brook trout eDNA concentration in each sample. Mean lake eDNA concentration was then estimated via the weighted average eDNA concentrations observed in samples from the littoral and pelagic zones; samples from each zone were assigned weights based on the respective proportions of these two zones in each lake. Notably, size structure among studied populations varied substantially in this dataset; the mean mass of individuals in the largest-bodied population was approximately an order of magnitude greater than the mean mass of individuals in the smallest-bodied population (mean mass range = 43.1 g–404.8 g/individual). As a result, it is a useful case study to explore the impact of a significant gradient in size structure on observed eDNA concentrations.

### 4.2 | Case Study 2: Brook Trout Abundance and eDNA in a System With Little Size Structure Variation Among Populations

For full methodological details, refer to Gaudet-Boulay et al. (2022). In brief, eDNA samples were collected from 15 lakes in 2020. Twenty 2 L samples were collected from each lake (depth = 0.5–2 m, 5–10 m from the shoreline). Samples were collected at systematic even-intervals throughout the perimeter of each lake; equal distancing between samples was calculated by dividing the perimeter by the total sample number (20). Quantitative PCR (Wilcox et al. 2016) was used to assess Brook Trout eDNA concentrations, with mean lake concentrations estimated from the average concentrations observed across all samples collected from a lake. We also only used 2020 in Case Study 2 due to unusually low correlations between capture effort and eDNA concentration in 2019, likely due to year-specific, seasonal effects on brook trout distributions (Gaudet-Boulay et al. 2022).

This study differed from the Yates, Glaser, et al. (2021) dataset used in Case Study 1 in two important ways; first, population size structure varied less than fourfold across studied populations (mean mass range = 83.4 g–296.6 g/individual). Second, abundance and biomass were directly estimated in Yates, Glaser, et al. (2021); eDNA concentrations in Gaudet-Boulay et al. (2023), were instead compared to abundance indices based on catch estimates obtained from anglers: Fish harvested per hectare (density/ha) and biomass harvested per hectare (biomass/ha). However, we lacked individual size data. As a result, it was only possible to calculate the mean individual mass for each population (mean mass = biomass ÷ $N$). Given that only mean population body mass data were available, several simplifying assumptions were made in our allometric eDNA adjustment for this empirical example. All individual members of a population were assumed to have a body mass equal to the mean body mass of that population. When all individuals of a population are assumed to have identical body mass, the mean of allometrically scaled individual mass values in the $i$th population is equal to the allometrically scaled mean mass of the $i$th population ($\overline{x_i^b} = \overline{x}_i^b$). In this case, Equation (4) simplifies to:

$$N_i = \frac{1}{J_0}\,\mathrm{eDNA}_i \div \overline{x}_i^b \tag{7}$$

Similarly, Equation (6) simplifies to:

$$\mathrm{Biomass}_i = \frac{1}{J_0}\,\mathrm{eDNA}_i \times \overline{x}_i^{1-b} \tag{8}$$

where $\mathrm{eDNA}_i$ denotes aggregate eDNA production for the $i$th population, $J_0$ denotes a normalization constant representing the equilibrium between production, degradation, and dispersion, $N_i$ is the numerical abundance of the $i$th population, $\overline{x}_i$ is the mean individual body mass of the $i$th population, and $b$ is the allometric scaling coefficient.

For Case Study 2, we applied eDNA adjustments for abundance metrics by adjusting eDNA data or abundance using the mean mass of a population raised to the power of $b$ ($\overline{x}_i^b$) rather than the mean of allometrically scaled individual mass values of a given population ($\overline{x_i^b}$) as in our original formulation (see Equations (7) and (8) above). Note, again, that this was only applied in the case of this second empirical example. However, this scenario may represent common datasets, and Equation (7) and (8) allow for the extension of our modeling framework when data include fewer details regarding population size structure.

### 4.3 | Integrating Allometry Into eDNA Adjustments in Natural Ecosystems to Estimate Population *N* and Biomass

We estimated the value of $b$ that jointly optimized the $b$ and $I_0$ values between "corrected" eDNA data and our two abundance metrics. We modeled $N$ as a function of the adjusted eDNA concentration using a negative binomial distribution with an identity link and modeled biomass as a function of adjusted eDNA concentration using a truncated (below by zero) Gaussian distribution for biomass:

$$N_i \sim NB(p_i, r)$$

$$\text{Biomass}_i \sim TN\left(\mu_i, \tau_i / \left(\text{eDNA}_i \frac{\overline{x_i}}{\overline{x_i^b}}\right)^a\right)$$

$$\lambda_i = \beta \frac{\text{eDNA}_i}{\overline{x_i^b}}$$

$$p_i = r/(r + \lambda_i),$$

$$\mu_i = \beta \text{eDNA}_i \frac{\overline{x_i}}{\overline{x_i^b}}$$

where $p_i$ is the probability parameter derived from the mean abundance predictor, $\lambda_i$, $r$ is the size parameter for the negative binomial model ($NB$), $\mu_i$ is the mean of the truncated Gaussian (TN) model, $\overline{x_i}$ is the mean mass of the $i$th population, and $\overline{x_i^b}$ is the mean of the allometrically scaled size structure distribution of the $i$th population. Note that the regression coefficient for the eDNA predictor term, $\beta$, is shared between the abundance and biomass models, and is equivalent to $1/J_0$ in Equation (4) and (6) above. The precision term $\left(\tau / \left(\text{eDNA}_i \frac{\overline{x_i}}{\overline{x_i^b}}\right)^a\right)$ allows variance to increase heteroscedastically with adjusted eDNA. We used a global Gaussian prior truncated below by zero for the $\beta$ parameter (i.e., the slope of the relationship between eDNA, $N$, and biomass) to constrain the model to an increasing function with eDNA, a gamma prior for $\tau$ to allow variance in biomass to increase with adjusted eDNA, and uniform priors for $b$, $a$, and size terms:

$$\beta \sim TN(0, 0.001)$$

$$r \sim \text{Uniform}(0, 50)$$

$$\tau \sim \text{gamma}(1, 0.01)$$

$$a \sim \text{Uniform}(0, 1)$$

$$b \sim \text{Uniform}(0, 2)$$

We fitted an origin-intercept model to estimate both $N$ and biomass. Under our framework, a joint model that simultaneously estimates $N$ and biomass must satisfy all the conditions established in Figure 1. Therefore, any estimates of population $N$ and biomass must meet the condition described by Equation (5): population $N$ multiplied by the mean mass of organisms within a population must equal biomass. Fitting origin-intercept models with a shared beta coefficient (representing the reciprocal of the $J_0$ parameter) estimated across both regressions satisfies this constraint. Since our framework assumes eDNA is produced solely by organisms and, in the absence of those organisms, no eDNA will persist in a stable steady state, an origin-intercept model is theoretically appropriate for our data and allows us to satisfy the conditions established in Figure 1.

We assessed model fit using Bayesian $p$-values based on the mean squared error (MSE) between observed and simulated values, visual assessment of traceplots (see Appendix S1), and Gelman–Rubin convergence statistics for each parameter ($\hat{R} < 1.05$) on $10^6$ iterations, three chains, a 100-thinning rate, and a burn-in of $5 \times 10^4$. All analyses for both case studies were performed within the statistical software *R* (Core R Team 2019) and *JAGS* (Plummer 2003a) using the packages *rjags* (Plummer 2003b), *jagsUI* (Kellner 2021), and *MCMCvis* (Youngflesh 2018). Graphs were created using base R as well as the package ggplot (Wickham 2016).

## 5 | Results

### 5.1 | Case Study 1

Our joint modeling approach demonstrated high positive correlations between $N$, biomass, and size structure adjusted eDNA data when $b$ was directly estimated from the data ($b = 0.51$, CI = 0.17–0.84; Table 1, Figure 4b). The credible interval estimates for $b$ did not overlap with either 0 or 1, indicating that eDNA production likely scaled allometrically within this system, rather than with $N$ ($b = 0$) or biomass ($b = 1$). Directly estimating the value of $b$ represented an improvement over modeling approaches that fix

**TABLE 1** | Fitted scaling coefficient values ($b$) for the relationship between size–structure-adjusted eDNA data and $N$, biomass using data from Yates, Glaser, et al. (2021; Case Study 1) and Gaudet-Boulay et al. (2022; Case Study 2).

| Yates, Glaser, et al. (2021) | | Gaudet-Boulay et al. (2022) | |
|---|---|---|---|
| $b$ | $r^2$ - predicted versus observed ($N$, Biomass) | $b$ | $r^2$ - predicted versus observed ($N$, Biomass) |
| 0[a] | 0.64, 0.53 | 0[a] | 0.71, 0.56 |
| 0.51 (0.17–0.84) | 0.87, 0.72 | 0.37 (0.03–0.83) | 0.74, 0.66 |
| 1[b] | 0.92, 0.69 | 1[b] | 0.74, 0.67 |

*Note:* Results are shown for the model in which $b$ was estimated and for models assuming $b = 0$ (i.e., eDNA production scales with numerical abundance) and $b = 1$ (i.e., assuming eDNA production scales linearly with biomass). Values in brackets represent 95% credible intervals. The $r^2$ corresponds to squared Pearson correlation coefficients between predicted and observed $N$ and biomass values, respectively.

[a]Value fixed in model, assumes eDNA production scales with numerical abundance.

[b]Value fixed in model, assumes eDNA production scales with biomass.

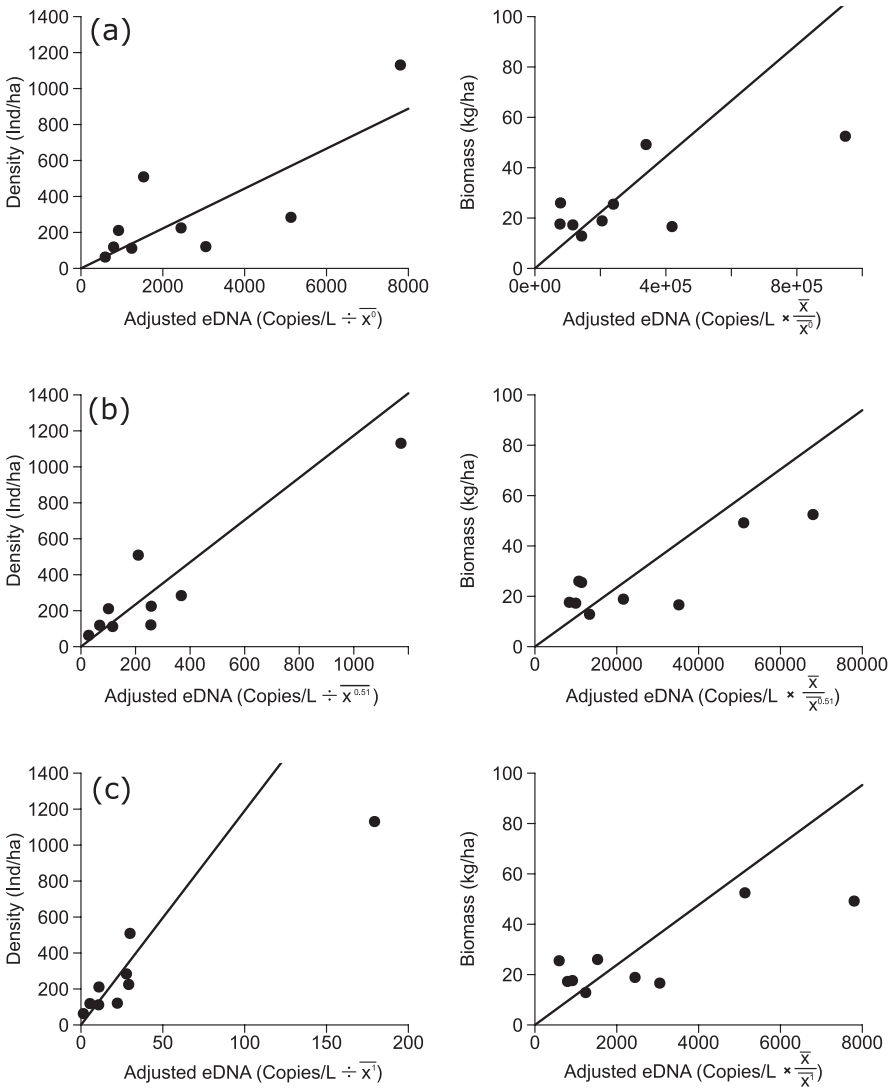


**FIGURE 4** | Linear regressions for Brook Trout eDNA concentration versus *N* (left column) and biomass (right column) for Case Study 1. Environmental DNA data has been adjusted for different values of the scaling coefficient: (a) fixed $b = 0$; (b) estimated $b = 0.51$; and (c) fixed $b = 1.00$ based on Equations (4) and (6). "x" refers body size (mass) of a population. Note scales on the x-axis differ. Data from Yates, Glaser, et al. (2021).

(assume) *b* to 0.00 (i.e., eDNA production scales linearly with *N*; Table 1, Figure 4a). However, estimating the value of *b* represented comparatively marginal model improvement over fixing *b* at 1.00 (Table 1, Figure 4c), although credible intervals for population estimates tended to be narrower when *b* was directly estimated. The squared Pearson correlation ($r^2$) values between predicted and observed *N* and biomass when *b* was estimated (0.51) versus fixed at 1.00 (i.e., eDNA production scaled with biomass) were broadly comparable, and both methods produced credible intervals that permitted differentiation between high and low *N* populations (Figure 5b,c). Fixing *b* at 0.00 (i.e., eDNA production scales with *N*) resulted in poorer relationships between predicted and observed *N* and biomass values, with credible intervals overlapping for all population pairs (Figures 4a and 5a).

## 5.2 | Case Study 2

After adjusting eDNA data using Equation (7) (due to unknown individual size structure), our joint modeling approach demonstrated high positive correlations between population *N*, biomass, and adjusted eDNA data ($b = 0.37$, CI = 0.03–0.83; Table 1, Figure 6b), and credible intervals overlapped with the point estimate of *b* from Case Study 1. Directly estimating the value of *b* represented an improvement over modeling approaches fixing *b* to 0.00 (i.e., eDNA production scales linearly with *N*; Table 1, Figure 6a). However, as with Case Study 1, estimating *b* also represented comparatively marginal modeling gains relative to fixing *b* at 1.00 (i.e., eDNA production scales with biomass; Table 1, Figure 6c). The $r^2$ values between predicted and observed *N* and biomass values when *b* was estimated (0.37) and fixed at 1.00 were comparable, and both methods produced credible intervals that permitted differentiation between high- and low-*N* and biomass populations (Figure 7b,c). However, assuming a *b* value of 0.00 resulted in poorer relationships between predicted and observed *N* and biomass values (Table 1). Contrary to Case Study 1, some high- and low-abundance and/or biomass populations were distinguishable based on eDNA data when *b* was fixed at 0.00. However, credible intervals for point estimates of *N* and biomass overlapped

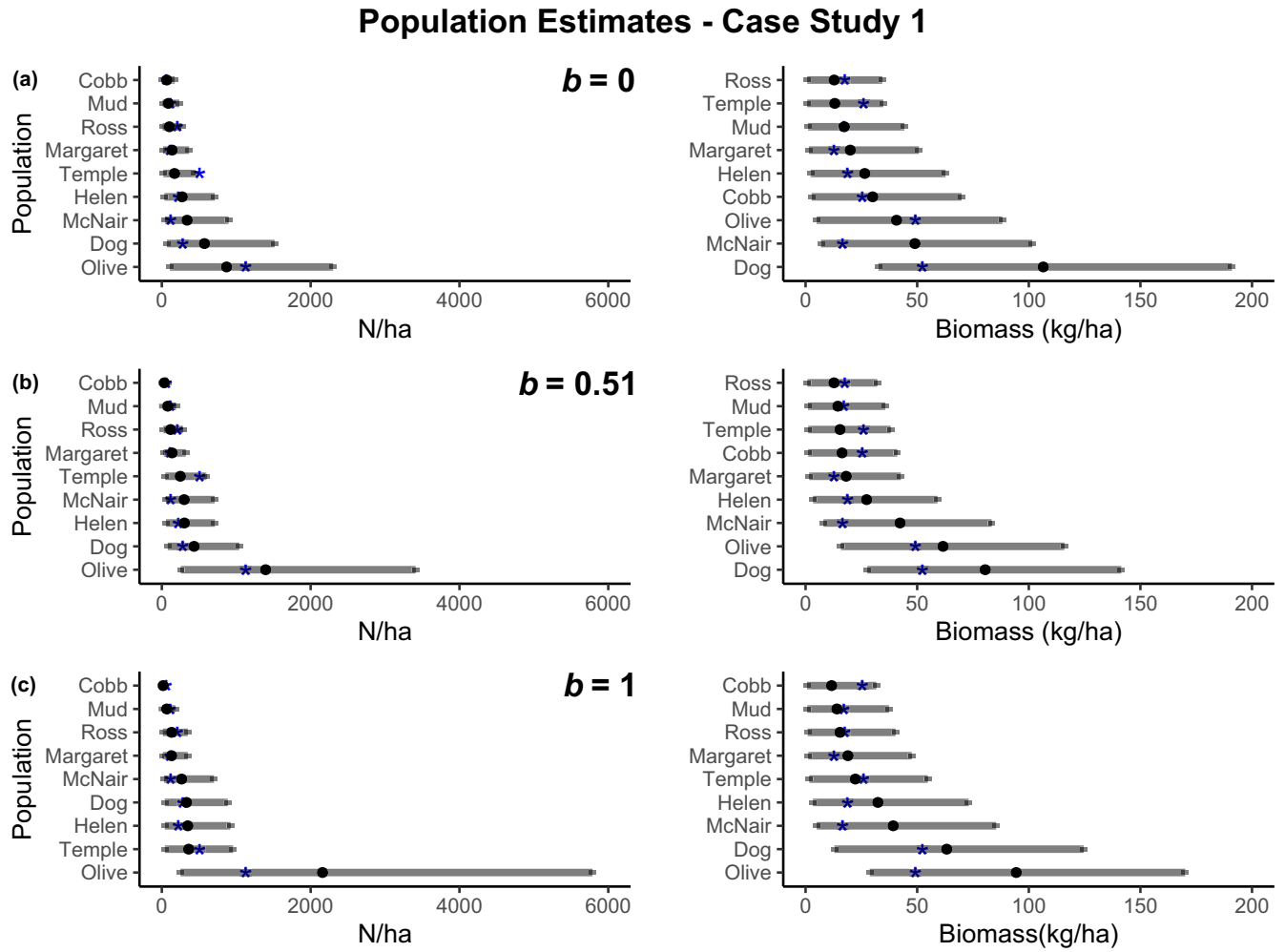


**FIGURE 5** | Brook Trout $N$ and biomass point-estimates (black circle) and 95% credible intervals based on eDNA data adjusted using Equations (4) and (5) for Case Study 1. Environmental DNA data has been adjusted for different values of the scaling coefficient: (a) fixed $b = 0$; (b) estimated $b = 0.51$; and (c) fixed $b = 1.00$. Blue asterisks represent the true $N$ or biomass per population (estimated from mark-recapture). Data from Yates, Glaser, et al. (2021).

for more populations when $b$ was fixed at 0.00 than when $b$ was estimated or fixed at 1 (Figure 7a–c).

## 6 | Discussion

### 6.1 | Unifying the Inference of *N* and Biomass From Quantitative eDNA Data

Building from metabolic frameworks describing eDNA production at the level of individual organisms, we derived and validated a framework to adjust for size structure to infer unobserved variation in population abundance ($N$) and biomass among populations using quantitative eDNA data. This framework resolves the long-standing debate about whether eDNA should be more closely related to $N$ or biomass; we demonstrate that quantitative eDNA data can be adjusted to simultaneously reflect *both*, depending on the assumed or inferred value of the scaling coefficient, $b$, associated with eDNA production. The allometric scaling parameter $b$ and the eDNA production normalization constant $I_0$ thus represent the critical links that unify the relationships between eDNA, $N$, and biomass by accounting for the effect of size structure on eDNA production. We also note that the relationships described by Equations (4) and (6) share the parameters $b$ and $I_0$. Jointly estimating $b$ and the eDNA production constant $I_0/J_0$ (i.e., the shared beta across eDNA/N and eDNA/biomass regressions) using an origin-intercept model is thus crucial to satisfying the conditions defined in Figure 1, namely that estimated biomass is equal to $N$ multiplied by the mean mass of a population.

Our framework demonstrated that directly correlating quantitative eDNA data to $N$ or biomass requires inherent assumptions regarding the physiology of eDNA production. Previous studies seeking to investigate relationships between eDNA, $N$, and biomass typically model observed eDNA data as a direct function of either metric (or vice versa). However, correlating population $N$ directly with quantitative eDNA data presumes eDNA production is invariable with individual body mass (i.e., $b = 0$); similarly, directly correlating biomass with quantitative eDNA data assumes that eDNA production scales linearly with individual body mass (i.e., $b = 1$). Under the assumption that $b$ is equal to 0, Equation (4) shows that untransformed quantitative eDNA data should be directly equivalent to $N$ (i.e., $\overline{x_i^b} = 1$, since

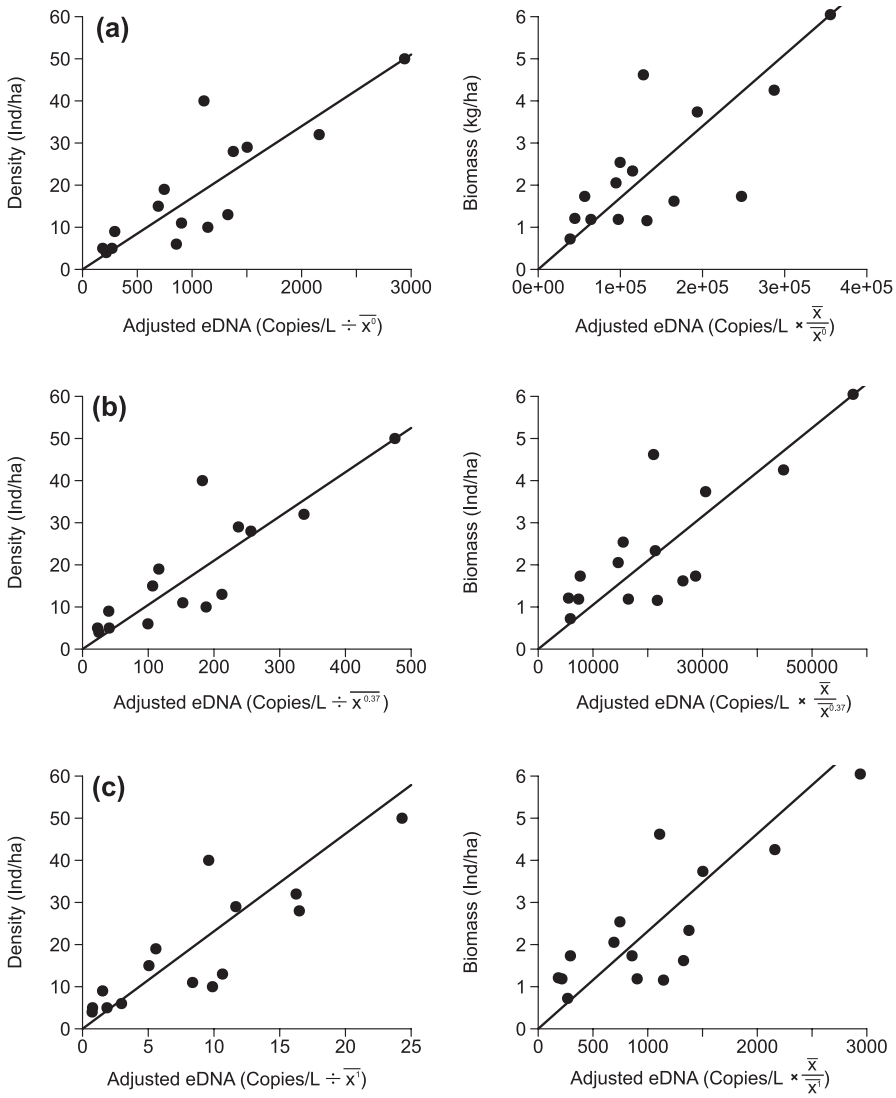


**FIGURE 6** | Linear regressions for Brook trout eDNA concentration versus $N$ (left column) and biomass (right column) for Case Study 2. Environmental DNA data have been adjusted for different values of the scaling coefficient: (a) fixed $b=0$; (b) estimated $b=0.37$; and (c) fixed $b=1$ based on Equations (4) and (6). "$x$" refers body size (mass) of a population. Note scales on the x-axis differ. Data from Gaudet-Boulay et al. (2022).

$b=0$), and quantitative eDNA data can be adjusted to reflect biomass using Equation (6) by multiplying quantitative eDNA data by the mean mass of individuals in a population (i.e., $\frac{\overline{x_i}}{x_i^b}=\overline{x_i}$, since $b=0$). Conversely, under the assumption that $b$ equals 1, Equation (6) shows that untransformed quantitative eDNA data should be equivalent to population biomass (i.e., $\frac{\overline{x_i}}{x_i^b}=1$, since $b=1$), and quantitative eDNA data can be adjusted to reflect $N$ using Equation (4) by dividing quantitative eDNA data by the mean mass of individuals in a population (i.e., $x_i^b=\overline{x_i}$, since $b=1$). The framework herein, however, provides a more generalizable eDNA adjustment to accurately reflect both $N$ and biomass. It is flexible enough to model the above scenarios (e.g., $b=1$ or 0) and any alternate values.

Even if further studies do not directly estimate the value of $b$, future studies correlating eDNA to an abundance metric (e.g., $N$ or biomass) where study populations exhibit considerable size structure variation should present data on correlations between adjusted eDNA concentrations and the other metric, for example, if correlating eDNA with biomass, report relationships between $N$ and eDNA divided by mean organism mass (and vice versa). Equations (4), (5), and (6) are essentially tautological; if the purpose of a study is to examine relationships between eDNA and one of either $N$ or biomass, both metrics should reasonably correlate with eDNA after appropriate adjustment. A substantial weakening of the relationship between eDNA and one metric relative to the other could indicate that population size structure could be significantly impacting eDNA production and that $b$ may deviate from 0 or 1. In these circumstances, we would recommend that researchers consider directly estimating the value of $b$.

## 6.2 | To "$b$," or Not to "$b$"?

For Case Studies 1 and 2, estimates of $b$ were lower than predicted based on general metabolic theory ($b=0.75$) and the estimated brook trout metabolic scaling coefficient (0.73) (Hartman

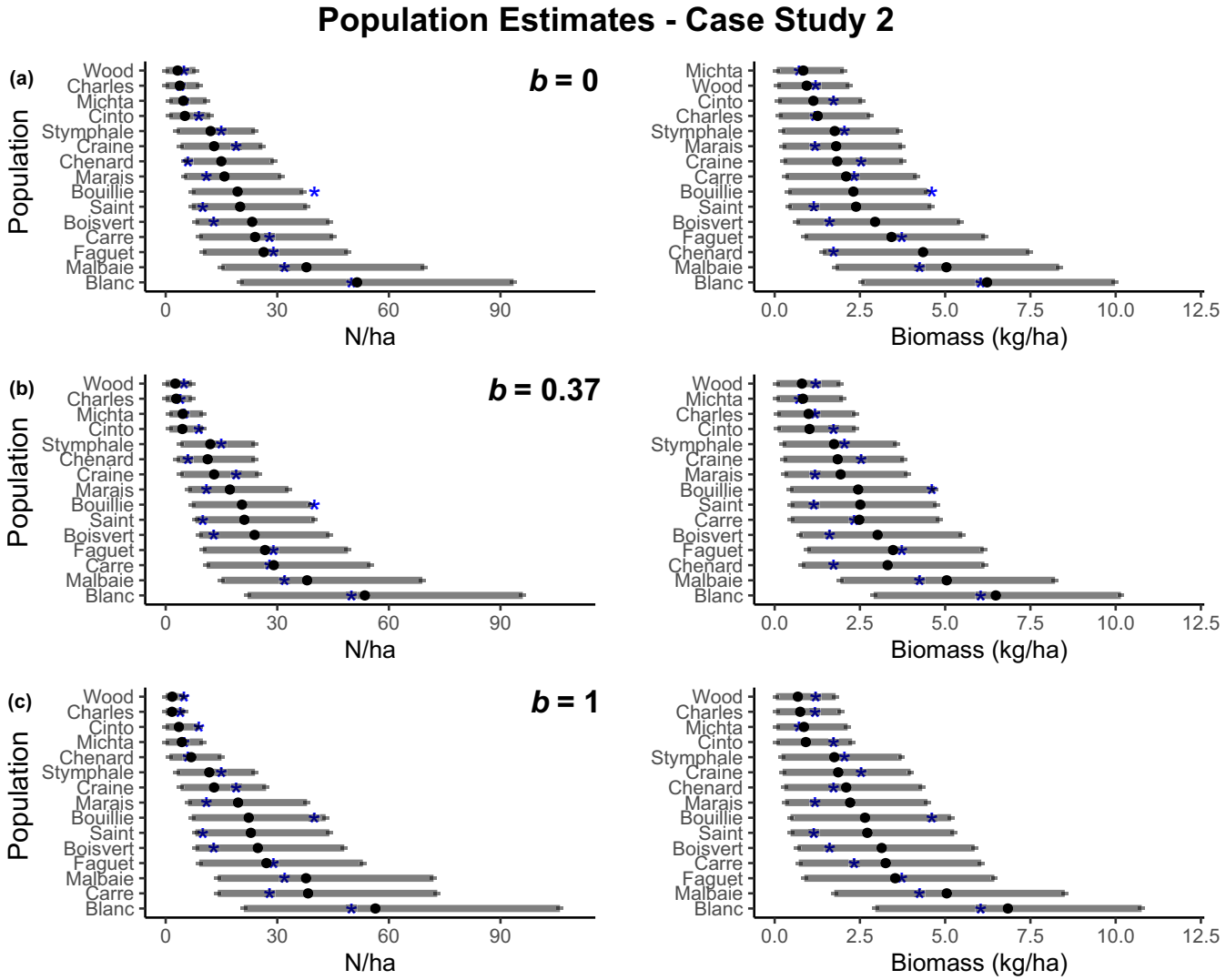


**FIGURE 7** | Brook Trout *N* and biomass point-estimates (black circle) and 95% credible intervals based on eDNA data adjusted using Equations (4) and (5) for Case Study 2. Environmental DNA data has been adjusted for different values of the scaling coefficient: (a) fixed $b = 0$; (b) estimated $b = 0.37$; and (c) fixed $b = 1.00$. Blue asterisks represent the true *N* or biomass per population (estimated from fisheries data). Data from Gaudet-Boulay et al. (2023).

and Cox 2008), although credible intervals for both estimates of *b* were wide and overlapped these values. We do, however, note that credible intervals for the estimates of *b* across both case studies did not overlap with either 0.00 or 1.00, indicating that eDNA production likely scales allometrically within these systems. Directly estimating *b*, therefore, had relevant outcomes for differentiating population *N* and biomass from quantitative eDNA data.

For Case Study 1, assuming a *b* value of 0.00 produced credible intervals for *N* and biomass that substantially overlapped across all population pairs, indicating a potentially limited capacity to differentiate population abundance and biomass based on eDNA data. Conversely, when estimating *b*, credible intervals for *N* did not overlap in populations with high and low *N*. We also note that this model was estimated from only nine datapoints; with additional data, the predictive utility of such models could improve. In Case Study 2, similar patterns were observed, although credible intervals for the estimate of *b* were wider and the point estimate lower. The utility of directly estimating *b* was also less pronounced, with some high- and low-*N*/biomass populations in Case Study 2 distinguishable based on eDNA data even assuming a *b* value of 0.00 (Figure 7a). We suspect that this is likely due to lower gradients in body mass across study populations compared to Case Study 1, where the smallest bodied population was an order of magnitude smaller than the largest bodied population; in Case Study 2, these extremes differed only fourfold. Nevertheless, when estimating *b* directly in these Case Study systems, *N* and biomass were distinguishable for more population pairs than if *b* was assumed to be 0.00.

However, directly estimating *b* yielded only modest improvement compared to assuming eDNA scaled with biomass (i.e., fixing $b = 1$; Figures 5 and 7). Population pairs that were distinguishable when *b* was directly estimated typically also did not have overlapping credible intervals when *b* was fixed at 1.00, although credible intervals for population estimates tended to be narrower when *b* was directly estimated. The extent to which directly estimating the value of *b* versus broadly assuming a value

of 1.00 (i.e., eDNA correlates to biomass) generally improves modeling efforts remains to be determined, but in similar systems, assuming eDNA scales linearly with biomass may be a sufficient approximation from a practical monitoring standpoint.

Nevertheless, our framework shows that assuming a value of $b$ equal to 0.00 (eDNA scales with $N$) or 1.00 (eDNA scales with biomass) inherently necessitates adjustments to infer the other metric if populations or experimental units exhibit substantial body size variation. While body size tends to vary more across species than across populations within a species (Blanck and Lamouroux 2007), freshwater fishes can still exhibit substantial interpopulation variation in body size (Arranz et al. 2016). Migratory salmonids, for example, can exhibit body sizes that differ by multiple orders of magnitude (Hutchings and Myers 1988; Sepulveda et al. 2021). However, accounting for differences in body mass may still be important for species that do not exhibit such extreme body size variation. The gradient of body mass across populations in Case Study 2 was substantially lower than in Case Study 1, but was comparable to (and often less than) gradients observed across many common freshwater species (Arranz et al. 2016). Our modeling also demonstrated that size structure variation was still important to account for when determining $N$, even if eDNA was assumed to scale linearly with biomass (i.e., $b = 1$). We also note that our two case studies cover a single species in similar ecological conditions; the extent to which eDNA production scales allometrically may vary between and across taxonomic groups.

### 6.3 | Quantitative Ecological Inferences From eDNA

Setting realistic expectations is critical for enhancing the utility of eDNA in quantifying abundance. Even when integrating the "ecology of eDNA," it is likely that it can only provide coarse, yet cost-effective, abundance inferences (Yates, Cristescu, et al. 2021; but see Beaulieu et al. 2024). After applying our framework, we were able to differentiate between high- and low-$N$ and biomass populations based on a combination of quantitative eDNA data and population size structure data, lending support to the assertion that eDNA, after accounting for size structure, may be able to provide abundance inferences regarding critical conservation thresholds (Spear et al. 2021). However, differentiating moderate-sized populations remained challenging. Additionally, the case studies demonstrated that the utility of inferences based on eDNA for differentiating biomass may be more limited relative to $N$. More population pairs in both case studies, for example, were distinguishable for $N$ relative to biomass. However, the case study lakes were almost solely inhabited by brook trout, representing ecosystems dominated by a single fish species that are apex predators within these ecosystems. As a result, these populations were likely close to their maximum carrying capacity; biomass thus only varied by a factor of four across populations in Case Study 1 and by a factor of eight in Case Study 2. How that biomass was distributed across individuals, however, varied substantially among populations, with the lowest- and highest-$N$ populations differing by more than an order of magnitude in both Case Studies. Under similar ecological conditions, populations may be more likely to exhibit a reduced gradient of biomass relative to $N$, which could affect the effectiveness of eDNA to differentiate biomass relative to $N$. On a broader species-distribution level, all of the study lakes in Case Studies 1 and 2 were also likely relatively high biomass close to carrying capacity due to a lack of predators or competitors. We suspect that it would be possible to identify ecosystems where Brook Trout have substantially lower biomass based on quantitative eDNA data signals from standardized surveys, for example, in mixed-species ecosystems with other predatory/competitor fish species (Browne and Rasmussen 2009; Myers et al. 2014).

While model accuracy may be improved with size structure data, these results imply that eDNA data alone (at least in our case study systems) may potentially be useful for differentiating very low and very high biomass populations, given the general overlap in performance between models with an estimated $b$ versus assuming a value of $b$ fixed at 1. However, $N$ is a potentially important population parameter in-and-of itself, particularly for the long-term genetic health and resilience of populations of conservation concern (Lande 1993; Reed et al. 2003; Willi et al. 2006) and the propagule success of invasive species (Cassey et al. 2018). It can be common for fish to exhibit small body size in high-density conditions due to a variety of factors (competition, resource limitations, etc.; Arranz et al. 2016). Conversely, populations can exhibit substantial variation in size structure at low-to-moderate densities. In both case study systems, for example, many low-density populations are dominated by individuals with high body mass, leading to moderate-to-high levels of total population biomass. These patterns can be observed in both datasets, with populations exhibiting small mean body sizes at higher densities and variable body sizes at low-to-moderate densities (Figure 8). Understanding how biomass is distributed across individuals within such populations may therefore be critical to using eDNA to assess potential differences in $N$ across populations and/or determining if populations are potentially above $N$ conservation thresholds (e.g., Spear et al. 2021). To determine whether biomass is distributed across many small individuals or fewer large ones, complementary data on population size structure would be necessary alongside eDNA analyses. Given the larger gradients in $N$ across populations compared to biomass in both case studies, eDNA may therefore have strong potential for differentiating low- and high-$N$ populations when paired with size structure data. As a result, eDNA surveys may be well suited to pair with index netting efforts commonly undertaken by conservation agencies.

Finally, it is important to note that, as applied in the current models, our framework only accounts for the effect of abundance and size structure on eDNA pseudo-steady-state concentrations. Other factors affect production and degradation in natural ecosystems, and when unaccounted for, such factors are integrated into models as residual noise. The accuracy of models predicting $N$ and biomass from eDNA concentration data could thus likely be substantially improved by collecting and integrating basic (and easy-to-collect) data on environmental factors we know impact eDNA concentrations (e.g., temperature [Caza-Allard et al. 2021; Jo et al. 2019]). Comprehensive eDNA-based models that integrate across

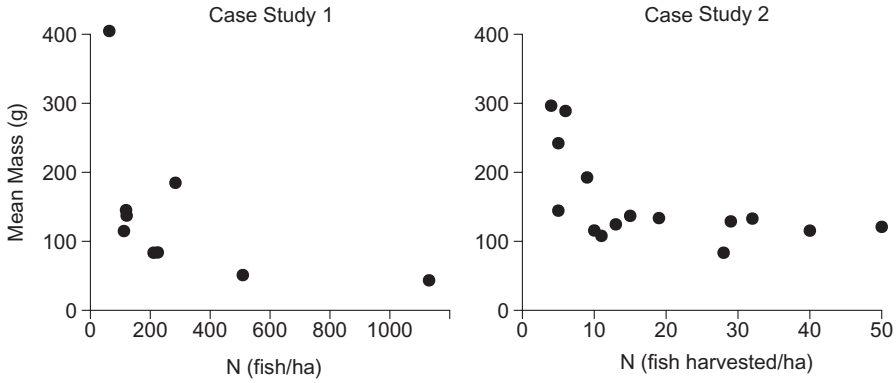


**FIGURE 8** | Relationship between mean body mass and density across study populations in Case Studies 1 and 2.

critical parameters could potentially rival conventional methods of estimating abundance (e.g., mark-recapture) in their precision and accuracy (Beaulieu et al. 2024).

### 6.4 | Limitations and Caveats

To evaluate the potential general utility of our proposed framework, it is crucial to first determine: (i) how prevalent an effect allometry has on observed concentrations of eDNA; and (ii) if the value of the scaling coefficient $b$ tends to be consistent across systems, or if these values tend to be specific to individual systems, species, or populations. Both case studies provided an illustrative example for the importance of accounting for body size structure when applying our framework to make inferences regarding $N$ or biomass, with models improving after considering population size structure. For Case Study 2, the relative benefit of accounting for size structure was less impactful compared to Case Study 1, which we suspect was a result of Case Study 2 populations exhibiting substantially less variation in size structure. We anticipate that the modeling improvement gained from accounting for size structure will be positively correlated with body size variation among study populations or species because our framework suggests that allometric adjustments approximate multiplication or division by a constant when body size variation is absent. Additionally, whether biomass or numerical abundance will exhibit a stronger relationship with eDNA concentrations will likely depend on how close the value of the allometric scaling coefficient is to zero or one for the species or taxonomic group being investigated. However, results for Case Study 2 imply that this relationship may not be linear, as the value of $b$ in Case Study 2 was ~0.4, yet biomass exhibited a slightly stronger relationship with eDNA. These observations could be an emergent result of the negative relationship observed between population body mass and population density (Figure 8) across both Case Studies. However, further studies are needed to explore the conditions under which eDNA is likely to best predict either abundance or biomass, and how correlations between variables and variable gradients can impact model performance when estimating abundance and biomass.

Additionally, it is worthwhile to mention that eDNA is also likely produced, in some measure, by the shedding of cells from the surface of an organism (Stewart 2019). While this is likely related to activity (and thus metabolism) to some extent (Thalinger et al. 2021), it also is likely related to the surface area of an organism, which itself scales allometrically with a scaling coefficient of ~2/3 (O'Shea et al. 2006). We speculate that metabolic processes are likely responsible for the majority of eDNA released into an environment. However, in some circumstances, cells shed from the surface of an organism may be the predominant source of DNA recovered in environmental samples, for example, from eDNA samples collected from blood and/or discharge water collected from fisheries catches (Urban et al. 2023). Nevertheless, we emphasize that $b$-values of 0.75 (expected under metabolic theory) and ~0.65 (surface area to body mass allometry) are broadly similar, and at the level of granularity observed in our case studies such differences would likely have little impact on the accuracy of model predictions.

We also want to (re)emphasize that integrating size structure represents only one aspect of understanding the relationship between quantitative eDNA data and organism abundance, which can also be impacted by environmental factors (transportation, temperature, pH, activity, etc.; Caza-Allard et al. 2021; Jo et al. 2019; Klymus et al. 2015; Levi et al. 2019; Shogren et al. 2017; Strickler et al. 2015; Thalinger et al. 2021; Wilcox et al. 2016, Beaulieu et al. 2024) and methodological factors (inhibitors, extraction method, qPCR/ddPCR vs. metabarcoding, etc.; Capo et al. 2021; Uchii et al. 2019; Yates, Cristescu, et al. 2021).

## 7 | Conclusion

Our framework unifies relationships between eDNA, $N$, and biomass and requires jointly estimating scaling and beta coefficients shared across regressions. We want to reemphasize that correlating unadjusted eDNA data to abundance metrics (either $N$ or biomass) makes inherent assumptions about eDNA production processes. Our framework demonstrates that, in the presence of substantial size structure variation across populations, adjustments should be made to model at least one of $N$ or biomass. If assuming eDNA scales linearly with $N$, eDNA should also reasonably correlate with biomass after multiplying by $\bar{x}_i$. Similarly, if assuming eDNA scales linearly with biomass, eDNA should correlate with $N$ after dividing by $\bar{x}_i$. If either of those adjustments result in a poor correlation, researchers should seek to directly estimate the value of $b$ within their systems, ideally using a joint modeling approach.

We hypothesize that our framework (and metabolic theory more broadly) can be applied to make predictions regarding relationships between eDNA, *N*, and biomass broadly across ecosystems and trophic levels, and can be readily integrated into hydrological modeling in lotic systems or mass-balance frameworks commonly used to model particle dynamics in lentic systems. The framework described herein is likely flexible enough to model dynamics associated with eDNA sources (i.e., extraorganismal or organismal eDNA, metabolic activity, trophic level, temperature, etc.) and thus lays the foundation for broadly interpreting quantitative eDNA signals across ecosystems.


### Acknowledgments

M.C.Y. was supported by a CIGLR and NSERC postdoctoral fellowships. Comments from M. Young, A. Sepulveda, and C. Jerde greatly improved previous versions of this manuscript. We thank A. Derry, M. Cristescu, and D. Fraser for their instrumental role in generating the data from Yates, Glaser, et al. (2021), and we thank L. Bernatchez for allowing us to use data from Gaudet-Boulay et al. (2023).

This manuscript is dedicated to the memory of Prof. Louis Bernatchez, who was a significant mentor and collaborator for several of the coauthors, a foundational innovator in the study of eDNA, and supportive and enthusiastic about the research covered by this manuscript.


### Conflicts of Interest

The authors declare no conflicts of interest.

### Data Availability Statement

No new data were generated for this manuscript; data from Case Studies 1 and 2 can be found in Yates, Glaser, et al. (2021) and Gaudet-Boulay et al. (2023), respectively. Model code for the joint framework can be found on Dryad.

### References


Adams, C. I. M., L. A. Hoekstra, M. R. Muell, and F. J. Janzen. 2019. "A Brief Review of Non-avian Reptile Environmental DNA (eDNA), with a Case Study of Painted Turtle (*Chrysemys picta*) eDNA Under Field Conditions." *Diversity* 11, no. 4: 50. https://doi.org/10.3390/d11040050.

Allegier, J. E., S. J. Wenger, A. D. Rosemond, D. E. Schindler, and C. A. Layman. 2015. "Metabolic Theory and Taxonomic Identity Predict Nutrient Recycling in a Diverse Food Web." *Proceedings of the National Academy of Sciences of the United States of America* 112, no. 20: E2640-7. https://doi.org/10.1073/pnas.1420819112.

Arranz, I., T. Mehner, L. Benejam, et al. 2016. "Density-Dependent Effects as Key Drivers of Intraspecific Size Structure of Six Abundant Fish Species in Lakes Across Europe." *Canadian Journal of Fisheries and Aquatic Sciences* 73, no. 4: 519–534. https://doi.org/10.1139/cjfas-2014-0508.

Baldigo, B. P., L. A. Sporn, S. D. George, and J. A. Ball. 2017. "Efficacy of Environmental DNA to Detect and Quantify Brook Trout Populations in Headwater Streams of the Adirondack Mountains, New York." *Transactions of the American Fisheries Society* 146, no. 1: 99–111. https://doi.org/10.1080/00028487.2016.1243578.

Beaulieu, J., M. C. Yates, D. J. Fraser, M. E. Cristescu, and A. M. Derry. 2024. "eDNA Provides Accurate and Precise Estimates of Abundance by Integrating Bioenergetics and Particle Mass-Balance." *bioRxiv*: 2024.12.23.629923. https://doi.org/10.1101/2024.12.23.629923.

Blanck, A., and N. Lamouroux. 2007. "Large-Scale Intraspecific Variation in Life-History Traits of European Freshwater Fish." *Journal of Biogeography* 34, no. 5: 862–875. https://doi.org/10.1111/j.1365-2699.2006.01654.x.

Boivin-Delisle, D., M. Laporte, F. Burton, R. Dion, E. Normandeau, and L. Bernatchez. 2021. "Using Environmental DNA for Biomonitoring of Freshwater Fish Communities: Comparison With Established Gillnet Surveys in a Boreal Hydroelectric Impoundment." *Environmental DNA* 3: 105–120. https://doi.org/10.1002/edn3.135.

Brown, J. H., J. F. Gillooly, A. P. Allen, V. Savage, and G. B. West. 2004. "Toward a Metabolic Theory of Ecology." *Ecology* 85, no. 7: 1771–1789.

Browne, D. R., and J. B. Rasmussen. 2009. "Shifts in the Trophic Ecology of Brook Trout Resulting From Interactions With Yellow Perch: An Intraguild Predator-Prey Interaction." *Transactions of the American Fisheries Society* 138, no. 5: 1109–1122. https://doi.org/10.1577/t08-113.1.

Capo, E., S. Koizumi, I. Puts, and F. Olajos. 2021. "Droplet Digital PCR Applied to Environmental DNA, a Promising Method to Estimate Fish Population Abundance From Humic-Rich Aquatic Ecosystems." *Environmental DNA* 3: 343–352. https://doi.org/10.1002/edn3.115.

Cassey, P., S. Delean, J. L. Lockwood, J. S. Sadowski, and T. M. Blackburn. 2018. "Dissecting the Null Model for Biological Invasions: A Meta-Analysis of the Propagule Pressure Effect." *PLoS Biology* 16, no. 4: e2005987. https://doi.org/10.1371/journal.pbio.2005987.

Caza-Allard, I., M. Laporte, G. Cote, J. April, and L. Bernatchez. 2021. "Effect of Biotic and Abiotic Factors on the Production and Degradation of Fish Environmental DNA: An Experimental Evaluation." *Environmental DNA* 4: 453–468.

Collins, R. A., C. Baillie, N. C. Halliday, et al. 2022. "Reproduction Influences Seasonal eDNA Variation in a Temperate Marine Fish Community." *Limnology and Oceanography Letters* 7, no. 5: 443–449. https://doi.org/10.1002/lol2.10271.

Cowart, D. A., K. R. Murphy, and C. H. C. Cheng. 2018. "Metagenomic Sequencing of Environmental DNA Reveals Marine Faunal Assemblages From the West Antarctic Peninsula." *Marine Genomics* 37: 148–160. https://doi.org/10.1016/j.margen.2017.11.003.

Evans, N. T., P. D. Shirey, J. G. Wieringa, A. R. Mahon, and G. A. Lamberti. 2017. "Comparative Cost and Effort of Fish Distribution Detection via Environmental DNA Analysis and Electrofishing." *Fisheries* 42, no. 2: 90–99. https://doi.org/10.1080/03632415.2017.1276329.

Gaudet-Boulay, M., E. García-Machado, M. Laporte, et al. 2023. "Relationship Between Brook Charr (*Salvelinus fontinalis*) eDNA Concentration and Angling Data in Structured Wildlife Areas." *Environmental DNA* 5, no. 5: 861–879. https://doi.org/10.1002/edn3.341.

Hanson, P., T. Johnson, D. Schindler, and J. Kitchell. 1997. *Fish Bioenergetics 3.0 Manual*. University of Wisconsin Center for Limnology.

Hartman, K. J., and M. Cox. 2008. "Refinement and Testing of a Brook Trout Bioenergetics Model." *Transactions of the American Fisheries Society* 137: 357–363. https://doi.org/10.1577/T05-243.1.

Hutchings, J. A., and R. A. Myers. 1988. "Mating Success of Alternative Maturation Phenotypes in Male Atlantic Salmon, *Salmo salar*." *Oecologia* 75, no. 2: 169–174. https://doi.org/10.1007/BF00378593.

Jerde, C. L., K. Kraskura, E. J. Eliason, S. R. Csik, A. C. Stier, and M. L. Taper. 2019. "Strong Evidence for an Intraspecific Metabolic Scaling Coefficient Near 0.89 in Fish." *Frontiers in Physiology* 10: 1166. https://doi.org/10.3389/fphys.2019.01166.

Jo, T., H. Murakami, S. Yamamoto, R. Masuda, and T. Minamoto. 2019. "Effect of Water Temperature and Fish Biomass on Environmental DNA Shedding, Degradation, and Size Distribution." *Ecology and Evolution* 9: 1135–1146. https://doi.org/10.1002/ece3.4802.

Jo, T. S. 2023. "Synthesizing the Relationships Between Environmental DNA Concentration and Freshwater Macrophyte Abundance: A Systematic Review and Meta-Analysis." *Hydrobiologia* 851, no. 7: 1697–1710. https://doi.org/10.1007/s10750-023-05409-x.

Kalff, J. 2001. *Limnology: Inland Water Ecosystems*. 1st ed. Prentice Hall.

Kellner, K. 2021. "jagsUI: A Wrapper Around 'rjags' to Streamline 'JAGS' Analyses." R Package Version 1.5.2. https://cran.r-project.org/package=jagsU.

Klymus, K. E., C. A. Richter, D. C. Chapman, and C. Paukert. 2015. "Quantification of eDNA Shedding Rates From Invasive Bighead Carp *Hypophthalmichthys nobilis* and Silver Carp *Hypophthalmichthys molitrix*." *Biological Conservation* 183: 77–84. https://doi.org/10.1016/j.biocon.2014.11.020.

Kuehne, L. M., C. O. Ostberg, D. M. Chase, J. J. Duda, and J. D. Olden. 2020. "Use of Environmental DNA to Detect the Invasive Aquatic Plants *Myriophyllum spicatum* and *Egeria Densa* in Lakes." *Freshwater Science* 39, no. 3: 521–533. https://doi.org/10.1086/710106.

Lacoursière-Roussel, A., M. Rosabal, and L. Bernatchez. 2016. "Estimating Fish Abundance and Biomass From eDNA Concentrations: Variability Among Capture Methods and Environmental Conditions." *Molecular Ecology Resources* 16: 1401–1414. https://doi.org/10.1111/1755-0998.12522.

Lande, R. 1993. "Risks of Population Extinction From Demographic and Environmental Stochasticity and Random Catastrophes." *American Naturalist* 142, no. 6: 911–927.

Levi, T., J. M. Allen, D. Bell, et al. 2019. "Environmental DNA for the Enumeration and Management of Pacific Salmon." *Molecular Ecology Resources* 19: 597–608.

Maruyama, A., K. Nakamura, H. Yamanaka, M. Kondoh, and T. Minamoto. 2014. "The Release Rate of Environmental DNA From Juvenile and Adult Fish." *PLoS One* 9, no. 12: e114639. https://doi.org/10.1371/journal.pone.0114639.

Mill, A. C., J. K. Pinnegar, and N. V. C. Polunin. 2007. "Explaining Isotope Trophic-Step Fractionation: Why Herbivorous Fish Are Different." *Functional Ecology* 21, no. 6: 1137–1145. https://doi.org/10.1111/j.1365-2435.2007.01330.x.

Myers, B. J. E., C. A. Dolloff, and A. L. Rypel. 2014. "Rainbow Trout Versus Brook Trout Biomass and Production Under Varied Climate Regimes in Small Southern Appalachian Streams." Wild Trout XI Symposium, 1, 127–135.

Nayfach, S., and K. S. Pollard. 2016. "Toward Accurate and Quantitative Comparative Metagenomics." *Cell* 166, no. 5: 1103–1116. https://doi.org/10.1016/j.cell.2016.08.007.

O'Shea, B., A. J. Mordue-Luntz, R. J. Fryer, C. C. Pert, and I. R. Bricknell. 2006. "Determination of the Surface Area of a Fish." *Journal of Fish Diseases* 29, no. 7: 437–440. https://doi.org/10.1111/j.1365-2761.2006.00728.x.

Pawlowski, J., L. Apothéloz-Perret-Gentil, and F. Altermatt. 2020. "Environmental DNA: What's Behind the Term? Clarifying the Terminology and Recommendations for Its Future Use in Biomonitoring." *Molecular Ecology* 29, no. 22: 4258–4264. https://doi.org/10.1111/mec.15643.

Pilliod, D. S., C. S. Goldberg, R. S. Arkle, and L. P. Waits. 2013. "Estimating Occupancy and Abundance of Stream Amphibians Using Environmental DNA From Filtered Water Samples." *Canadian Journal of Fisheries and Aquatic Sciences* 70: 1123–1130.

Plummer, M. 2003a. "JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling JAGS: Just Another Gibbs Sampler." DSC 2003 Working Papers, DSC.

Plummer, M. 2003b. "rjags: Bayesian Graphical Models Using MCMC." R Package Version 4.13. https://cran.r-project.org/package=rjags.

R Core Team. 2019. *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. http://www.r-project.org.

Reed, D. H., E. H. Lowe, D. A. Briscoe, and R. Frankham. 2003. "Fitness and Adaptation in a Novel Environment: Effect of Inbreeding, Prior Environment, and Lineage." *Evolution* 57, no. 8: 1822–1828.

Rourke, M. L., A. M. Fowler, J. M. Hughes, et al. 2022. "Environmental DNA (eDNA) as a Tool for Assessing Fish Biomass: A Review of Approaches and Future Considerations for Resource Surveys." *Environmental DNA* 4: 9–33. https://doi.org/10.1002/edn3.185.

Sard, N. M., S. J. Herbst, L. Nathan, et al. 2019. "Comparison of Fish Detections, Community Diversity, and Relative Abundance Using Environmental DNA Metabarcoding and Traditional Gears." *Environmental DNA* 1, no. 4: 368–384. https://doi.org/10.1002/edn3.38.

Sassoubre, L. M., K. M. Yamahara, L. D. Gardner, B. A. Block, and A. B. Boehm. 2016. "Quantification of Environmental DNA (eDNA) Shedding and Decay Rates for Three Marine Fish." *Environmental Science and Technology* 50, no. 19: 10456–10464. https://doi.org/10.1021/acs.est.6b03114.

Seibel, B. A., and J. C. Drazen. 2007. "The Rate of Metabolism in Marine Animals: Environmental Constraints, Ecological Demands and Energetic Opportunities." *Philosophical Transactions of the Royal Society B: Biological Sciences* 362, no. 1487: 2061–2078. https://doi.org/10.1098/rstb.2007.2101.

Sepulveda, A. J., R. L. Al-Chokhachy, M. B, et al. 2021. "It's Complicated... Environmental DNA as a Predictor of Trout and Char Abundance in Streams." *CJFAS* 78, no. 4: 422–432.

Shogren, A. J., J. L. Tank, E. Andruszkiewicz, et al. 2017. "Controls on eDNA Movement in Streams: Transport, Retention, and Resuspension." *Scientific Reports* 7: 5065. https://doi.org/10.1038/s41598-017-05223-1.

Skelton, J., A. Cauvin, and M. E. Hunter. 2023. "Environmental DNA Metabarcoding Read Numbers and Their Variability Predict Species Abundance, but Weakly in Non-Dominant Species." *Environmental DNA* 5, no. 5: 1092–1104. https://doi.org/10.1002/edn3.355.

Spear, M. J., H. S. Embke, P. J. Krysan, and M. J. Vander. 2021. "Application of eDNA as a Tool for Assessing Fish Population Abundance." *Environmental DNA* 3, no. 1: 83–91. https://doi.org/10.1002/edn3.94.

Stat, M., M. J. Huggett, R. Bernasconi, et al. 2017. "Ecosystem Biomonitoring With eDNA: Metabarcoding Across the Tree of Life in a Tropical Marine Environment." *Scientific Reports* 7, no. 1: 12240. https://doi.org/10.1038/s41598-017-12501-5.

Stewart, D. J., and M. Ibarra. 1991. "Predation and Production by Salmonine Fishes in Lake Michigan 1978–88." *Canadian Journal of Fisheries and Aquatic Sciences* 48, no. 5: 909–922. https://doi.org/10.1139/f91-107.

Stewart, K. A. 2019. "Understanding the Effects of Biotic and Abiotic Factors on Sources of Aquatic Environmental DNA." *Biodiversity and Conservation* 28: 983–1001. https://doi.org/10.1007/s10531-019-01709-8.

Stoeckle, M. Y., J. Adolf, Z. Charlop-powers, K. J. Dunton, G. Hinks, and S. M. Vanmorter. 2021. "Trawl and eDNA Assessment of Marine Fish Diversity, Seasonality, and Relative Abundance in Coastal New Jersey, USA." *ICES Journal of Marine Science* 78, no. 1: 293–304. https://doi.org/10.1093/icesjms/fsaa225.

Strickler, K. M., A. K. Fremier, and C. S. Goldberg. 2015. "Quantifying Effects of UV-B, Temperature, and pH on eDNA Degradation in Aquatic Microcosms." *Biological Conservation* 183: 85–92. https://doi.org/10.1016/j.biocon.2014.11.038.

Takeuchi, A., T. Iijima, W. Kakuzen, S. Watanab, and Y. Yamada. 2019. "Release of eDNA by Different Life History Stages and During Spawning Activities of Laboratory-Reared Japanese Eels for Interpretation of Oceanic Survey Data." *Scientific Reports* 9: 6074. https://doi.org/10.1038/s41598-019-42641-9.

Thalinger, B., A. Rieder, A. Teuffenbach, et al. 2021. "The Effect of Activity, Energy Use, and Species Identity on Environmental DNA Shedding of Freshwater Fish." *Frontiers in Ecology and Evolution* 9: 623718. https://doi.org/10.3389/fevo.2021.623718.

Tillotson, M. D., R. P. Kelly, J. Duda, M. Hoy, J. Kralj, and T. P. Quinn. 2018. "Concentrations of Environmental DNA (eDNA) Reflect Spawning

Salmon Abundance at Fine Spatial and Temporal Scales." *Biological Conservation* 220: 1–11. https://doi.org/10.1016/j.biocon.2018.01.030.

Tsuji, S., and N. Shibata. 2021. "Identifying Spawning Events in Fish by Observing a Spike in Environmental DNA Concentration After Spawning." *Environmental DNA* 3, no. 1: 190–199. https://doi.org/10.1002/edn3.153.

Uchii, K., H. Doi, T. Okahashi, et al. 2019. "Comparison of Inhibition Resistance Among PCR Reagents for Detection and Quantification of Environmental DNA." *Environmental DNA* 1, no. 4: 359–367. https://doi.org/10.1002/edn3.37.

Urban, P., D. Bekkevold, H. Degel, et al. 2023. "Scaling From eDNA to Biomass: Controlling Allometric Relationships Improves Precision in Bycatch Estimation." *ICES Journal of Marine Science* 80, no. 4: 1066–1078. https://doi.org/10.1093/icesjms/fsad027.

Vanni, M. J., and P. B. McIntyre. 2016. "Predicting Nutrient Excretion of Aquatic Animals With Metabolic Ecology and Ecological Stoichiometry: A Global Synthesis." *Ecology* 97, no. 12: 3460–3471. https://doi.org/10.1002/ecy.1582.

Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wilcox, T. M., K. S. Mckelvey, M. K. Young, et al. 2016. "Understanding Environmental DNA Detection Probabilities: A Case Study Using a Stream-Dwelling Char *Salvelinus fontinalis*." *Biological Conservation* 194: 209–216. https://doi.org/10.1016/j.biocon.2015.12.023.

Willi, Y., J. Van Buskirk, and A. A. Hoffmann. 2006. "Limits to the Adaptive Potential of Small Populations." *Annual Review of Ecology, Evolution, and Systematics* 37, no. 1: 433–458. https://doi.org/10.1146/annurev.ecolsys.37.091305.110145.

Wotton, R. S., and B. Malmqvist. 2001. "Feces in Aquatic Ecosystems." *Bioscience* 51, no. 7: 537–544. https://doi.org/10.1641/0006-3568(2001)051[0537:FIAE]2.0.CO;2.

Yates, M. C., M. E. Cristescu, and A. M. Derry. 2021. "Integrating Physiology and Environmental Dynamics to Operationalize Environmental DNA (eDNA) as a Means to Monitor Freshwater Macro-Organism Abundance." *Molecular Ecology* 30, no. 24: 6531–6550.

Yates, M. C., D. J. Fraser, and A. M. Derry. 2019. "Meta-Analysis Supports Further Refinement of eDNA for Monitoring Aquatic Species-Specific Abundance in Nature." *Environmental DNA* 1: 5–13. https://doi.org/10.1002/edn3.7.

Yates, M. C., D. M. Glaser, J. R. Post, M. E. Cristescu, D. J. Fraser, and A. M. Derry. 2021. "The Relationship Between eDNA Particle Concentration and Organism Abundance in Nature Is Strengthened by Allometric Scaling." *Molecular Ecology* 30: 3068–3082. https://doi.org/10.1111/mec.15543.

Yates, M. C., T. M. Wilcox, M. Y. Stoeckle, and D. D. Heath. 2023. "Interspecific Allometric Scaling in eDNA Production Among Northwestern Atlantic Bony Fishes Reflects Physiological Allometric Scaling." *Environmental DNA* 5, no. 5: 1105–1115. https://doi.org/10.1002/edn3.381.

Youngflesh, C. 2018. "MCMCvis: Tools to Visualize, Manipulate, and Summarize MCMC Output." *Journal of Open Source Software* 3: 640. https://doi.org/10.21105/joss.00640.

**Supporting Information**

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