Final Report

Humboldt State University Sponsored Programs Foundation
1 Harpst Street
Arcata, CA 95521

To

California Department of Transportation (Caltrans)

Entitled

STANDARDIZING ENVIRONMENTAL DNA METHODOLOGIES FOR COHO SALMON

<table>
<tr><td>**Investigators:**</td><td>**Andrew Kinziger,**
Eric Bjorkstedt
Andre Buchheister
Mark Henderson
Gavin Bandy
Braden Herman
Jason Shaffer</td></tr>
<tr><td>**Period of Performance:**</td><td>**01/02/2020 – 12/31/2023**</td></tr>
</table>

Contractor's Job Reference #65A0762

Date
December 2023

TABLE OF CONTENTS

OVERALL PROJECT SUMMARY

Background:

The California Department of Transportation (Caltrans) is mandated to conduct Federal Endangered Species Act (FESA) Section 7 and California Endangered Species Act (CESA) consultations for projects potentially impacting federally and state-listed species. Coho salmon (*Oncorhynchus kisutch*), listed under both FESA and CESA, require precise data on presence/absence, distribution, and abundance for effective consultation and associated mitigation. California is also home to many other listed aquatic species. Caltrans maintenance and project delivery actions may require mitigation even when fish aren't detected because of the potential for listed species to be illusive, their population numbers are low, and because traditional survey methods are sometimes invasive or inefficient due to methodologies having impact to habitats and right to enter requirements. Surveys can also be time consuming, labor-intensive and costly, and require a permit for capturing and handling listed species. Therefore eDNA methods have the potential to simplify and expedite environmental review and permitting.

Caltrans would benefit by incorporating eDNA into their survey tool kit. The establishment of a database of site occupancy for listed species because it could be used to conduct consultations and as a baseline for future maintenance and project delivery needs. Further, it may allow Caltrans to contribute to establishing population abundance and distribution data in the future, which would advance knowledge conservation of listed species.

Innovation:

Environmental DNA (eDNA) serves as a powerful tool for aquatic species detection and monitoring by capturing genetic material that organisms release into their surrounding environment. eDNA is revolutionizing species monitoring, combining cost-effectiveness with a high sensitivity that surpasses traditional methods, particularly in detecting low-abundance or elusive species. Its non-invasive nature offers a dual advantage: there is minimal disturbance to organisms and their habitats, and detection or absence of species has potential to streamline permitting processes. This will allow Caltrans to make appropriate conservation and management decisions according to environmental laws and regulations.

This project undertook a rigorous examination of eDNA, exploring shedding and decay rates, analyzing eDNA transport in stream ecosystems, and demonstrating the robust capabilities of eDNA estimating species presence and distribution. These detailed studies lay the groundwork for the precise interpretation of eDNA data, affirming eDNA's efficacy in species detection. Moreover, the application of eDNA in this initiative has not only deepened our comprehension of coho salmon ecology but also accentuated the methodology's vital role in modern conservation practices, ecological study, and the proficient governance of natural environments.

Task 1: Laboratory studies on eDNA shedding and decay

- We evaluated the degradation of eDNA derived from two distinct fish taxa, rainbow trout (*Oncorhynchus mykiss*) and white sturgeon (*Acipenser transmontanus*), across a temperature gradient.
    - We identified a two-phase decay pattern: an initial rapid decline in eDNA concentration, followed by a phase where eDNA levels off to a low but stable concentration, persisting at low concentrations for the entire 16 day duration of the experiment.
    - Temperature played a crucial role, with higher temperatures correlating with faster eDNA decay. Interestingly, the decay rates between the two species did not vary significantly, despite the initial eDNA concentration of rainbow trout being significantly higher than that of white sturgeon.

- We estimated eDNA shedding and decay rates in rainbow trout (*Oncorhynchus mykiss*) at two densities.
    - We present a novel approach for estimating eDNA shedding rates, which considers these time varying decay rates.
    - Shedding rate estimates for rainbow trout are provided, illustrating how decay model selection influences these estimates. We did not find clear evidence that fish density significantly impacts decay rates or shedding rates of eDNA.

Task 2: Field studies on eDNA ecology and dynamics

- We developed and implemented a novel Autonomous DNA Introduction Device (ADID) using foreign eDNA (FeDNA) tracers to understand eDNA downstream transport in natural streams.
- This method was designed to calibrate eDNA dynamics and improve the accuracy of species detection and abundance estimates.
- Trials in various streams revealed stream-specific FeDNA transport patterns, emphasizing the importance of FeDNA as a tool for studying eDNA transport.

Task 3: Determine local occupancy using eDNA

- We evaluated the ability of an exogenous DNA tracer to characterize the transport dynamics of naturally occurring DNA over fine spatial scales within a stream.
- The tracer particle allowed us to assess the occupancy of coho salmon and steelhead within 500m study reaches in different streams.

Task 4: eDNA application in field surveys

- We examined the efficacy of eDNA versus underwater visual count (UVC) surveys in detecting juvenile coho salmon in the Smith River, California.
- The study found a high concordance between eDNA and UVC methods, with over 90% detection probability under median habitat conditions for both techniques.
- eDNA was less effective in areas with high discharge, suggesting dilution impacts eDNA detectability more than UVC.

Implementation:

This project's outcomes highlight the transformative potential of eDNA monitoring in the management and conservation of aquatic species. The advanced methods developed through this initiative will considerably enhance Caltrans' capabilities to meet regulatory requirements and support science and data to make decisions regarding project scope and guide decisions regarding compliance with applicable Federal and State Laws.

Looking to the future, the adoption of eDNA monitoring will allow for streamlining comprehensive biodiversity analysis by detecting specific organisms in aquatic environments. Environmental DNA analysis marks a significant step towards advancing data collection and thorough ecosystem assessments, and will allow detection of threatened and endangered species.

TASK 1: CONDUCT LABORATORY STUDIES TO ASSESS EDNA SHEDDING AND DECAY.

Temperature-dependence and long-term persistence demonstrate the value of dynamic decay models for fish environmental DNA

Braden A. Herman 1
Gavin B. Bandy 1
Eric P. Bjorkstedt 2
Andre Buchheister 1
Andrew P. Kinziger 1,*

1 Department of Fisheries Biology, Cal Poly Humboldt, Arcata, CA 95521, USA
2 Southwest Fisheries Science Center, NOAA Fisheries, Trinidad, CA 95570, USA
*Corresponding Author: Andrew.Kinziger@humboldt.edu

Synopsis: This study investigated environmental DNA (eDNA) degradation in two fish species across three temperature treatments. Temperature influenced initial eDNA decay rates, but species showed no significant differences. Trace eDNA persisted for 16 days, requiring further research.

Abstract

We examined the degradation of environmental DNA (eDNA) derived from two distinct fish taxa, rainbow trout (*Oncorhynchus mykiss*) and white sturgeon (*Acipenser transmontanus*), across a temperature gradient. Water samples, containing eDNA from both species and sourced from the Cal Poly Humboldt Fish Hatchery, were incubated at 12, 18, and 24°C, and eDNA concentrations were quantified at intervals of 0, 0.5, 1, 2, 4, 8, 12, and 16 days using quantitative PCR. Our investigation identified a log-linear model with a tail as the most suitable representation of eDNA decay, revealing a two-phase decay pattern marked by an initial rapid decline followed by a very low, residual eDNA concentration that remained relatively constant through the remainder of the experiment. During the initial log-linear decay phase, our results support the significant influence of temperature on eDNA decay rates, as higher temperatures were associated with accelerated decay. We resolved no significant variation in eDNA decay rates between species despite initial concentrations of rainbow trout 16 times greater than white sturgeon. In the second phase of the decay process, we observed that eDNA persisted at low concentrations for at least 16 days (the extent of our experiment) across a broad range of temperatures, indicating a need for further investigations to determine when eDNA concentrations will reach zero.

Introduction

Environmental DNA (eDNA) has emerged as an effective tool for aquatic species monitoring, enabling non-invasive detection of aquatic organisms by capturing traces of genetic material they shed into their surrounding environments (Ficetola et al., 2008; Thomsen et al., 2012; Rodríguez-Ezpeleta et al. 2021). This genetic material, in the form of DNA fragments, holds a wealth of information about the presence and abundance of various species, making it a valuable resource for conservation and management (Dejean et al., 2011; Jerde et al., 2011; Sutter and Kinziger 2019). However, before accurate estimates of distribution and abundance can be made, several aspects of eDNA dynamics need to be thoroughly understood, including, critically, eDNA degradation, or the breakdown of eDNA molecules over time (Goldberg et al., 2013; Harrison et al. 2019).

Research into the degradation of eDNA has uncovered a complex underlying process likely influenced by a variety of physical, chemical, and biological factors, with eDNA decay rates varying significantly based on environmental conditions and time (Jo & Minamoto 2021; Lamb et al. 2022). Temperature has consistently been identified as a critical factor influencing eDNA degradation (Lamb et al., 2022) and elucidating the effects of temperature variations on decay rates is essential, as it reflects the seasonal and geographical dynamics present in aquatic ecosystems. Furthermore, taxonomic differences in eDNA decay dynamics are an intriguing avenue of research, as previous research has produced mixed results regarding the significance of taxon on eDNA decay rates (Sassoubre et al. 2016; Allan et al., 2021; Caza-Allard et al. 2021; Kirtaine et al. 2021).

While log-linear (i.e., exponential decay) models are commonly chosen to characterize the decay dynamics of eDNA (as evident from reviews by Jo & Minamoto 2021; and Lamb et al. 2022), it's worth noting that several studies have demonstrated that alternative models often provide more accurate

representations of eDNA decay. Specifically, Weibull decay models with decreasing decay constants over time (Bylemans et al., 2018; Allan et al., 2020) and models featuring a rapid initial decay phase followed by a slower subsequent decay phase (Eichmiller et al., 2016; Shogren, 2018) have been shown to better capture the complexities of the eDNA decay process compared to simple log-linear models.

In this study, we evaluate the degradation of eDNA from two phylogenetically-divergent lineages of fishes, rainbow trout (Oncorhynchus mykiss) and white sturgeon (Acipenser transmontanus), across temperatures typical in the native ranges of these species. To characterize eDNA decay, we explored various established bacterial survival models, covering a broad range of decay patterns. These models allow for variations in decay rates over time, including those featuring initial periods of no decay (referred to as "shoulders") and models accommodating the persistence of residual eDNA over time (referred to as "tails"; Geeraerd et al., 2005). Our objectives were to (1) characterize temporal patterns of eDNA decay, identifying the most suitable decay model, and (2) evaluate how temperature and taxonomic group influence decay rates.

Methods

Water Collection

Water was collected from the Cal Poly Humboldt Fish Hatchery, which raises rainbow trout and white sturgeon. The hatchery's water supply originates from Fern Lake, a 2500 m^2 reservoir built in 1962 with a maximum depth of 4.5 m and approximate volume of 5000 m^3. Although Fern Lake does not host rainbow trout or white sturgeon, it does support abundant populations of macroinvertebrates, zooplankton and bacteria. The hatchery's recirculating system includes a sand filter, aeration tower, six circular tanks (7,000 L each), two 30 m raceways (40,000 L each), and 16 smaller experimental tanks (300 L each). At the time of sampling (September 2020), the fish populations at the hatchery included ~1,300 rainbow trout weighing ~170 kg and ~750 white sturgeon weighing ~54 kg. The species, number, average total length (mm), and population total weight (kg) of the fish population in Cal Poly Humboldt Fish Hatchery is in Supplemental Table S1.

The hatchery water utilized in these experiments contained eDNA accumulated from multiple recirculation events and was collected at a junction following a fresh pass over all fish in the hatchery. It was filtered through a fine mesh net (<2 mm) to exclude large particulates and pooled into a sanitized 250-L tank. The tank remained continuously mixed while a total of 81 water samples, each measuring ~0.9 L, were individually placed into sterilized plastic wedge-shaped bottles designed to fit within a circular tank. Nine bottles were used to determine the initial eDNA concentration of the target species at the start of the experiment, 63 were used to assess decay, and nine were used as negative controls.

Experimental Design

Bottles filled with hatchery water were systematically arranged in three experimental tanks. The experimental tanks were located adjacent to each other in a sheltered lab (open-sided, but roofed) at the

Cal Poly Humboldt Fish Hatchery. Twenty-four treatment bottles were placed in each of three round tanks, arranged in a circle, and labeled (1…8, repeated, Figure 1). A water bath was maintained in each tank and kept at 12, 18, or 24°C to span the typical temperature range found in northern California streams supporting salmonid populations. Temperature was controlled by thermostat-equipped heaters and chillers, which were calibrated against a NIST-traceable thermometer (Thermco Products Inc ACC10033SFC). Temperature was monitored at 5-minute intervals in each water bath with HOBO temperature loggers and was routinely checked throughout the experiment with the NIST-traceable thermometer. The bottles were covered with aluminum foil to prevent cross contamination and were gently aerated to promote suspension of eDNA and support aerobic activity. Temperature varied by less than 0.5°C during the course of the experiment.

Sample Collection and Filtration

Triplicate water samples (i.e., bottles) were taken from each temperature treatment at each of seven intervals, at time t = 0.5, 1, 2, 4, 8, 12, and 16 days. Bottles were selected systematically, one from each third of the treatment tank, to account for any potential differences in conditions associated with position in the treatment tank. The schedule of collection was designed to track the log-linear decline of eDNA concentration observed in previous eDNA degradation studies (Jo & Minamoto 2021; Lamb et al. 2022) but also attempted to provide information on overall persistence times suggested by previous research (Eichmiller et al. 2016). To evaluate cross contamination among water samples, one control bottle, filled with tap water at the start of incubation, was taken from each temperature treatment on days 1, 2, and 4. Water from each sampled bottle was filtered immediately, after which bottles were refilled with tap water and replaced in the bath to maintain consistent tank volume and contact between the water bath and remaining samples.

Water was filtered on site immediately after the water samples were pulled from the water bath. Each water sample was pulled across a 47 mm diameter 0.45 micron cellulose nitrate (Whatman, GE Healthcare) filter, supported by a filter pad (MilliporeSigma™ AP1004700), mounted in sterile filter funnel and placed under a vacuum using a pneumatic hand pump (EWK EB0103A). The water samples were filtered until dry or for 30 minutes, whichever came first. If the entire sample was not filtered, the total volume filtered was recorded, and eDNA concentration was volume corrected. Filters were immediately placed in 360 µL of cell lysis buffer (Buffer ATL, QIAGEN) and frozen until extraction.

Molecular Methods

Retained eDNA was extracted directly from filters using a QIAGEN DNeasy blood and tissue kit following the manufacturer instructions, except: 40 µL of proteinase k and 360 µL of Buffer ATL was used to ensure filter submersion during lysing; QIAGEN QIAshredder kits were used for lysate homogenization; and samples were eluted into 100 µL of buffer AE. DNA extractions were conducted in a dedicated low-copy laboratory in which sterile conditions are maintained by UVC irradiation when the space is not in use.

Concentrations of eDNA were estimated simultaneously for rainbow trout and white sturgeon by conducting duplexed qPCR reactions. Assays for both species were based on published primers and

probes (Brandl, 2015; Wilcox, 2015). In each qPCR reaction, 7.5 µL of TaqMan™ Environmental Master Mix 2.0 (Applied Biosystems, 4396838), species-specific assay components for white sturgeon (250 nM probe, 900 nM forward primer, and 900 nM reverse primer), rainbow trout (250 nM probe, 300 nM forward primer, and 600 nM reverse primer), 2 µL of DNA extract, and nuclease-free water were combined to achieve a total reaction volume of 15 µL. qPCR reactions were set up in a dedicated low-copy laboratory (as above) and subsequently performed on an Applied Biosystems QuantStudio™ 3 Real-Time PCR System in a dedicated high copy laboratory. Reactions were cycled as follows: incubation at 50°C for 2 min and at 95°C for 10 min, then 45 cycles of 15 s at 95°C and 60 s at 60 °C. All qPCR reactions were run in triplicate.

To quantify the target DNA concentration in each qPCR reaction, a qPCR standard curve was generated. The standard was extracted from tissue using a Qiagen DNeasy Blood & Tissue Kit (69506), and its concentration determined using ddPCR. The standard curve was constructed across at least six serial 1:10 dilutions, with 4-48 replicates per standard concentration. The limit of detection (LOD; the minimum DNA concentration at which positive detections occur at a rate of at least 95%), and the limit of quantification (LOQ; the minimum concentration for which the coefficient of variation among replicates is less than 0.35) were determined as described by Klymus et al. (2020). Samples with zero copies of eDNA per reaction (assumed to occur at Ct>41) were recorded as 0.01 copies per reaction and retained in the analysis.

For the rainbow trout assay, the standard curve analysis yielded a slope of -3.40 and an intercept of 38.63 (qPCR efficiency was 96.7%), with LOD and LOQ of 4.61 copies per reaction and 18 copies per reaction, respectively. For the white sturgeon assay, the standard curve analysis yielded a slope of -3.53 and an intercept of 39.23 (qPCR efficiency was 92.1%), with LOD and LOQ of 5.4 copies per reaction and 9 copies per reaction, respectively.

DNA copies per liter (CN/L) was calculated as

$$CN/L = (CN/PCR) \times (V_e/V_t) * (1/V_s) \quad\quad\quad\quad (Eq. 1)$$

where CN/PCR represents the copy number per qPCR reaction from the standard curve, V_e denotes the elution volume (µL), V_t represents the template volume (µL), and V_s indicates the volume of water filtered (L). Two extreme, outlier values, inconsistent with the other technical replicates, were excluded and assumed to be indicative of a failed or contaminated reaction. Concentrations from the triplicate qPCR technical replicates were averaged into an estimate of sample concentration for analysis.

Statistical Analysis

A suite of nine alternative decay models were fit to each species-temperature dataset using the R package nlsMicrobio (Baty and Delignette-Muller 2014), which includes implementations of most of the bacterial survival (i.e., decay) models described in Geeraerd et al. 2005. The nine decay models fit included: (1) Geeraerd with shoulder and tail (Geeraerd et al., 2000), (2) Geeraerd without a tail (Geeraerd et al., 2000), (3) Geeraerd without a shoulder (Geeraerd et al., 2000), (4) Weibull (Mafart et al., 2002), (5) Weibull with a tail (Albert and Mafart 2005), (6) log-linear with a shoulder (Geeraerd et al., 2000), (7) log-linear with a tail (Geeraerd et al., 2000), (8) a biphasic (Cerf 1977), and (9) log-linear. Estimated eDNA concentrations

(copies/L) were $\log_{10}$ transformed to stabilize the variance. Models were fit using non-linear regression, and compared using Akaike's Information Criterion corrected for small sample sizes (AICc) and AICc differences (delta AICc) (Burnham and Anderson 2002). Any models with nonsensical parameter estimates (e.g., negative shoulder lengths) or whose parameter estimates simplified down to another competing model were excluded from the model comparison.

Model selection identified the log-linear-model-with-tail as the best overall representation of eDNA decay across the six species-temperature trials (see Results: Model Selection). A single unifying model was fit to all the datasets to test for significant species and temperature effects on each of the parameters and to estimate the parameters in unison, leveraging the information across the different trials (e.g., Jo et al., 2019). The log-linear model with a tail is a two-phase model defined as having an exponential (i.e., log-linear) decline followed by a constant eDNA concentration:

$$N_t = N_0 e^{-kt} \text{ if } t \leq t^* \qquad\qquad\qquad (Eq. 2)$$

$$N_t = N_{res} \text{ if } t > t^*$$

where N_t is the eDNA concentration (copies per L) at time t (hours), N_0 is the initial starting eDNA concentration at $t = 0$, k is the decay rate (h^{-1}), and N_{res} is the constant residual eDNA concentration that is maintained after time t^*, which is the time at which the decay changes between the two phases. The breakpoint t^* is calculated as:

$$t^* = \log_e(N_0) - \log_e(N_{res}) / k . \qquad\qquad\qquad (Eq. 3)$$

Each of the three base parameters (k, N_{res}, and N_0) were modeled as functions of species, temperature, and a species-temperature interaction:

$$\theta = b_0 + b_1{}^*S + b_2{}^*T_{18} + b_3{}^*T_{24} + b_4{}^*S{}^*T_{18} + b_5{}^*S{}^*T_{24} \quad (Eq. 4)$$

where θ represents any of the three base model parameters, S is an indicator (i.e., dummy) variable for the rainbow trout species, T_{18} and T_{24} are indicator variables for the 18°C and 24°C temperature (T) trials, and b_{0-5} are estimated coefficients. In this formulation, the b_0 intercept represents the estimate of θ for white sturgeon at the 12°C treatment, and the other coefficients (b_{1-5}) represent deviations from this reference level. Temperature was treated as a categorical variable (instead of continuous) because there were only three treatment levels, and these treatments were effectively constant and distinct throughout the experiment.

The log-linear-model-with-tail (Eq. 1) was fitted with k, N_{res}, and N_0 modeled using Eq. 4. However, N_0 was only allowed to vary by species (i.e., b_{2-5} were set to 0). Also, if the b4 and b5 estimates were not significantly different from zero for k or N_0, then those terms were removed from the model and the model was refitted. All models were $\log_{10}$-transformed to linearize the data and homogenize the variances as implemented in the nlsMicrobio package. Models were fit in R using the nls function (R Core Team 2023), and 95% confidence intervals were estimated by non-parametric bootstrapping using the nlstools package (Baty et al. 2015).

As a sensitivity test, a log-linear model (defined by the first phase of decay in Eq. 2) was refitted using only the first 96 hours (4 days) of data when eDNA concentrations were declining most rapidly. The parameters k and N_0 were modeled using Eq. 4. This simpler model was used to evaluate how sensitive our conclusions were to (1) adjusting eDNA concentrations from 0 to 0.01 copies per reaction for inclusion in the model, and (2) including samples from 8 days and beyond, which were mostly below the LOD.

Results

Model selection

Of the 9 competing decay models fit to each species-temperature data set, the log-linear model with a tail was most consistently the best model with the lowest AICc (Table 1). The log-linear model with a tail was the best model in all 12 and 18°C temperature trials, and it was a competitive model (delta AICc=0.45) for the white sturgeon 24°C trial, but that model was not well supported for the rainbow trout 24°C trial (Table 1). Two models (Geeraerd without a tail, and Log-linear with a shoulder) generated nonsensical estimates (i.e., a negative shoulder length) for all species-temperature data sets and were excluded from consideration.

Best-fit model

Analysis of the full model indicated that the species-temperature interaction terms did not have significant effects on estimates of k or $\log_{10}(N_{res})$ (p>0.3), and a simpler model that excluded the interaction term was preferred under model selection (i.e., had a lower AICc). After removing the species-temperature interactions from the model, the final, log-linear model with a tail, included k and $\log_{10}(N_{res})$ as functions of species and temperature, and $\log_{10}(N_0)$ was a function of species (Table 2). For the final model, overall $R^2 = 0.912$.

Changes in eDNA concentration through time across the species-temperature treatments followed a two-phase decay pattern characterized by initial log-linear decay followed by a very low, residual eDNA concentration that remained relatively constant through the remainder of the experiment (Figure 2). The extremely low residual eDNA concentrations are particularly evident when data are plotted on the raw scale (Figure S1). Initial starting concentrations for rainbow trout were significantly higher (~16 fold) than for white sturgeon (Figure 2). Model fit and parameter estimates indicated that the initial decay of eDNA was not significantly different between species (p=0.65), and that temperature had a significant effect on decay rates (p<<0.001; Figure 2, Table 2). Estimated rates of decay during the initial phase increased linearly with temperature (Figure 3). Faster decay at higher temperatures hastened the transition from initial to the tail phases of the eDNA decay process. Predicted eDNA concentrations reached a tail after approximately 100-200 hours depending on the species and temperature (Figure 2, Table 2).

Diagnostics and fit for the log-linear model with tail indicated nine outliers at low observed concentrations, arising from inclusion of samples with zero eDNA (Figure 2, Figure S2). However, a log-linear model fit only to the data through 96 hours generated identical conclusions; there were no

significant species-temperature interactions for the k and $\log_{10}(N_{res})$ parameters, and the species- and temperature-specific estimates of k from this model (with data ≤96 hrs) were not significantly different from those estimated using the log-linear model with a tail using the full data set. This indicates that our overall conclusions regarding decay during the initial phase of the experiment were not sensitive to (1) outlier values that skewed the normal distribution of the residuals, (2) the inclusion of data beyond 96 hours which had a greater chance of being below the LOD, or (3) our adjustment of eDNA concentrations from 0 to 0.01 copies per reaction to allow for logging and inclusion in model fitting.

Discussion

In our investigation of eDNA decay from rainbow trout and white sturgeon, we identified a distinct two-phase decay pattern characterized by an initial rapid decline followed by a period of indeterminate duration (at least several days) during which no measurable decay was observed. During the initial log-linear decay phase we show a significant impact of temperature on decay rates, as higher temperatures accelerate degradation, but we found no significant differences in eDNA decay rates between these two phylogenetically and ecologically distinct fish species, despite different initial concentrations.

Decay Rate Models

To accurately characterize eDNA degradation, it is essential to identify the model that best describes how eDNA breaks down over time. Our findings provide further evidence that models for eDNA decay, if they are to be useful in understanding eDNA dynamics over a period of any substantial extent, must include dynamic decay rates that change over time. Specifically, we observed an initial phase of rapid degradation followed by a tail without detectable decay, which aligns with similar findings in previous studies (e.g., Eichmiller et al., 2016; Bylemans et al., 2018; Shogren, 2018; Allan et al., 2021). This contrasts with the commonly employed log-linear models for eDNA decay, which assume a constant decay rate over time, as discussed in reviews by Jo and Minamoto (2021) and Lamb et al. (2022).

Accurately characterizing the dynamics of eDNA decay requires careful consideration of the timing and total duration of water sample collection, as these have implications for the resolution of the data and the form of model that can be supported. In our investigation, the majority of eDNA degradation occurred during the initial phase, and a log-linear model was identified as the best fit for describing the initial loss of eDNA. However, our sampling during this initial rapid decay phase was limited. Higher resolution sampling during this period would be required to assess whether eDNA decay follows a more dynamic process than that suggested by a simple log-linear model in this initial phase (e.g., Bylemans et al., 2018; Allan et al., 2020). Our results also suggest that any allocation of effort to high-resolution sampling must be balanced against observations over the longer term, given that detectable eDNA concentrations persisted to our final sampling at 16 days, indicating that longer experiments are needed to robustly estimate persistence times.

Ensuring proper model specification carries profound implications, as it directly shapes our understanding of the eDNA decay process itself, the factors influencing it, and ultimately, our ability to draw inferences

about species detection and abundance through eDNA monitoring techniques. At its core, the decision between models allowing for time-varying decay rates and those relying on a constant rate represents a choice among distinctly contrasting interpretations of the decay process. Models with a constant decay rate are commonly used (reviews by Jo & Minamoto 2021; and Lamb et al. 2022) and are largely based on well-established first-order decay seen in radioactive processes, a fundamental concept in chemistry. Conversely, models that incorporate time-varying decay rates acknowledge the heterogeneous nature of eDNA, which includes various eDNA particle types, including both eDNA surrounded by an intact cell wall (or membrane) and eDNA not surrounded by one (Nagler et al., 2022). Each of these particles can be influenced by a range of physical, chemical, and biological factors, leading to fluctuations in decay rates dependent on environmental conditions and time.

To illustrate the repercussions of using an inappropriate model on our eDNA decay data, we reanalyzed our dataset using a log-linear model (without a tail) as is often used in the literature. The estimates of k were consistently lower (by 37-68%) for the log-linear model (without a tail) compared to the first phase of the log-linear decay with tail model (Table S2, Figure S3). Species and temperature differences in k were not significantly different for the log-linear model (without a tail), in contrast to the log-linear model with a tail, which identified significant temperature-related differences (Figure 2; Table 2). This highlights how using a model that does not accurately represent decay dynamics can lead to vastly different conclusions regarding the factors influencing the process and the estimated decay rates. Some studies (e.g., Jo et al. 2019; Allan et al., 2021) present results of log-linear models that appear to suffer from a similar lack of fit (although this can be hard to discern when eDNA concentrations are not plotted on the log-scale). This example shows the importance of conducting diagnostics of model residuals as a check against interpreting results from models with poor fit.

Log-linear decay phase

In the early stages of rapid eDNA degradation, our analysis yielded decay rate constants that spanned from 0.036 hr^{-1} at 12°C to 0.071 hr^{-1} at 24°C. These values align with the reported range found in existing literature (Jo & Minamoto 2021; Lamb et al. 2022). Our findings are consistent with previous research demonstrating the fundamental role of temperature in describing eDNA degradation during the initial log-linear decay phase (Lamb et al., 2022). Further, we resolved a linear relationship between decay rate constant and temperature, concordant with previous research (Figure 3; Allan et al., 2020). However, further assessments encompassing a wider temperature range are required to ascertain whether this association follows a truly linear relationship or if a more complicated model, such as is commonly used to characterize enzymatic activity relationship with temperature (Berg et al. 2012), would be a better fit.

During the log-linear decay phase, we found no species-specific effects on eDNA decay rates based on our comparison between two phylogenetically distinct fish lineages. These results align with previous research comparing decay rates among marine fishes (Sassoubre et al. 2016; Kirtane et al. 2021) but contrast with Caza-Allard et al. (2021), which identified statistically significant variations in decay rates across fish species.

The biomass of rainbow trout was three times that of white sturgeon in the hatchery, which translated into a 16-fold higher initial eDNA concentration. Previous research has indicated that eDNA decay rates are higher in environments with greater species biomass (Bylemans et al., 2018; Jo et al., 2019), but we did

not observe any differences in decay rates between the two species. A key point to note is that our study examined decay of eDNA originating from two species occurring at different abundances in a common habitat instead of separate pools with varying densities of identical fish species, and the decay process followed the dynamics of each eDNA in mixture with the other (i.e., at a shared, pooled concentration).

Tail phase

In the second phase of the decay process, we observed that eDNA persisted at low concentrations for at least 16 days (the extent of our experiment) across a broad range of temperatures. This suggests a considerably slower rate of eDNA decay relative to the rapid decay observed during the initial phase. We hypothesize that these findings point to the existence of highly stable, decay-resistant eDNA particles. This could be attributed to eDNA particles encased within a double-membrane structure, like mitochondria, or eDNA bound to particles that shield it from degradation by nucleases (Nagler et al. 2021).

Regarding the two studied species, white sturgeon and rainbow trout, different patterns are suggested during this second phase. While white sturgeon observations consistently remained below the species-specific LOD, a large number of rainbow trout observations were above the LOD (Table S3). These differences may potentially be attributed to species-specific factors, the considerably higher initial concentration (16 times higher) of rainbow trout, or variations in qPCR assay efficiencies as estimated by the standard curve method. Despite our attempt to utilize logistic regression for analyzing these distinctions, the available data did not provide sufficient information for model fitting with most models having convergence failures. To gain further insight into this later decay stage, a larger number of samples per time point, an extended sampling duration, and careful consideration of qPCR standard curve methods will be necessary.

Study Design Considerations

In this investigation, we assessed eDNA decay in water obtained from a fish hatchery with recirculating water, including passage through a sand filter. It's important to acknowledge that the sand filter may selectively remove particles, potentially affecting our assessment of eDNA decay dynamics.

We specifically designed this study to ensure the independence of the eDNA decay process in each water sample. Instead of using a large-volume bath with water samples drawn at set time intervals, we created separate water samples from a single initial, well-mixed pool of water. This approach allowed the eDNA decay process to occur independently in each bottle, avoiding interdependencies that can arise when collecting water sequentially (e.g., repeated measures) from a central pool. Consequently, the variability observed in our study's decay process is anticipated to be higher due to minor differences in the initial concentration in each bottle and the lack of autocorrelation from multiple samples taken from a shared pool.

Acknowledgement

Author Information

The study design was conceived by all authors. BAH and GBB performed the experiments and molecular analysis, while AB conducted the statistical analysis. The manuscript was primarily written by APK and BAH, with contributions from all authors.

References

Albert, I., & Mafart, P. (2005). A modified Weibull model for bacterial inactivation. International Journal of Food Microbiology, 100, 197–211. https://doi.org/10.1016/j.ijfoodmicro.2004.10.016

Allan, E. A., Zhang, W. G., Lavery, A. C., & Govindarajan, A. F. (2021). Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. Environmental DNA, 3, 492–514. https://doi.org/10.1002/edn3.141

Baty F., Delignette-Muller, M. (2014). nlsMicrobio: data sets and nonlinear regression models dedicated to predictive microbiology. R package version 0.0-3.

Baty, F., Ritz, C., Charles, S., Brutsche, M., Flandrois, J.P., & Delignette-Muller, M.L. (2015). A Toolbox for Nonlinear Regression in R: The Package nlstools. Journal of Statistical Software, 66(5), 1-21. https://www.jstatsoft.org/article/view/v066i05. doi 10.18637/jss.v066.i05

Berg, J. M., Tymoczko, J. L., Gatto, G. J., & Stryer, L. (2012). Biochemistry (7th ed.). W.H. Freeman and Company.

Brandl, S., Schumer, G., Schreier, B. M., Conrad, J. L., May, B., & Baerwald, M. R. (2015). Ten real-time PCR assays for detection of fish predation at the community level in the San Francisco Estuary-Delta. Molecular Ecology Resources, 15, 278-284.

Burnham, K.P. and Anderson, D.R. (2002) Model Selection and Inference: A Practical Information-Theoretic Approach. 2nd Edition, Springer-Verlag, New York. http://dx.doi.org/10.1007/b97636

Bylemans, J., Furlan, E. M., Gleeson, D. M., Hardy, C. M., & Duncan, R. P. (2018). Does size matter? An experimental evaluation of the relative abundance and decay rates of aquatic environmental DNA. Environmental Science & Technology, 52(11), 6408–6416. Link

Caza-Allard, I., Laporte, M., Côté, G., April, J., & Bernatchez, L. (2021). Effect of biotic and abiotic factors on the production and degradation of fish environmental DNA: An experimental evaluation. Environmental DNA, 4, 453-468.

Cerf, O. (1977). Tailing of survival curves of bacterial spores. Journal of Applied Bacteriology, 42, 1–20.

Dejean, T., Valentini, A., Miquel, C., Taberlet, P., Bellemain, E., & Miaud, C. (2011). Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog Lithobates catesbeianus. Journal of Applied Ecology, 49, 953-959.

Eichmiller, J. J., Best, S. E., & Sorensen, P. W. (2016). Effects of temperature and trophic state on degradation of environmental DNA in lake water. Environmental Science & Technology, 50(4), 1859-1867.

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

Geeraerd, A. H., Herremans, C. H., & Van Impe, J. F. (2000). Structural model requirements to describe microbial inactivation during a mild heat treatment. International Journal of Food Microbiology, 59, 185–209.

Geeraerd, A. H., Valdramidis, V. P., & Van Impe, J. F. (2005). GInaFiT, a freeware tool to assess non-log-linear microbial survivor curves. International Journal of Food Microbiology, 102(1), 95–105.

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., Spear, S. F., McKee, A. M., Oyler-McCance, S. J., Cornman, R. S., Laramie, M. B., Mahon, A. R., Lance, R. F., Pilliod, D. S., Strickler, K. M., Waits, L. P., Fremier, A. K., Takahara, T., Herder, J. E., & Taberlet, P. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. Methods in Ecology and Evolution, 7(11), 1299–1307.

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: Biological Sciences, 286, 20191409.

Jerde, C. L., Mahon, A. R., Chadderton, W. L., & Lodge, D. M. (2011). "Sight-unseen" detection of rare aquatic species using environmental DNA. Conservation Letters, 4(2), 150–157.

Jo, T., Murakami, H., Yamamoto, S., Masuda, R., & Minamoto, T. (2019). Effect of water temperature and fish biomass on environmental DNA shedding, degradation, and size distribution. Ecology and Evolution, 27(1), 112–212.

Jo, T. , & Minamoto, T. (2021). Complex interactions between environmental DNA (eDNA) state and water chemistries on eDNA persistence suggested by meta‑analyses. Molecular Ecology Resources, 21, 1490–1503. 10.1111/1755-0998.13354

Kirtane, A., Wilcox, T. M., Carim, K. J., McKelvey, K. S., Young, M. K., & Schwartz, M. K. (2021). Quantification of Environmental DNA (eDNA) shedding and decay rates for three commercially harvested fish species and comparison between eDNA detection and trawl catches. Environmental DNA, 3, 1142-1155.

Klymus, K. E., Merkes, C. M., Allison, M. J., et al. (2020). Reporting the limits of detection and quantification for environmental DNA assays. Environmental DNA, 2, 271–282. https://doi.org/10.1002/edn3.29

Lamb, P. D., Fonseca, V. G., Maxwell, D. L., & Nnanatu, C. C. (2022). Systematic review and meta-analysis: Water type and temperature affect environmental DNA decay. Molecular Ecology Resources, 22, 2494–2505. https://doi.org/10.1111/1755-0998.13627

Mafart, P., Couvert, O., Gaillard, S., & Leguerinel, I. (2002). On calculating sterility in thermal preservation methods: Application of the Weibull frequency distribution model. International Journal of Food Microbiology, 72, 107–113.

Nagler, M., Podmirseg, S. M., Ascher-Jenull, J., Sint, D., & Traugott, M. (2022). Why eDNA fractions need consideration in biomonitoring. Molecular Ecology Resources, 22, 2458–2470. https://doi.org/10.1111/1755-0998.13658

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rodríguez-Ezpeleta, N., L. Zinger, A.P. Kinziger, H.M. Bik, A. Bonin, E. Coissac, B.C Emerson, C. Martins Lopes, T.A. Pelletier, P. Taberlet, S. Narum. (2021). Biodiversity monitoring using environmental DNA. Molecular Ecology Resources, 21, 1405–1495.

Sassoubre, L. M., Yamahara, K. M., Gardner, L. D., Block, B. A., & Boehm, A. B. (2016). Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. Environmental Science & Technology, 50(19), 10456–10464.

Shogren, A. J., Tank, J. L., Egan, S. P., August, O., Rosi, E. J., Hanrahan, B. R., Renshaw, M. A., Gantz, C. A., & Bolster, D. (2018). Water flow and biofilm cover influence environmental DNA detection in recirculating streams. Environmental Science & Technology, 52(15), 8530–8537.

Sutter, M., & Kinziger, A. P. (2019). Rangewide tidewater goby occupancy survey using environmental DNA. Conservation Genetics, 20(3), 597-613.

Thomsen, P. F., Kielgast, J., Iversen, L. L., Møller, P. R., Rasmussen, M., & Willerslev, E. (2012). Detection of a diverse marine fish fauna using environmental DNA from seawater samples. PLoS One, 7(8), e41732.

Wilcox, T. M., Carim, K. J., McKelvey, K. S., Young, M. K., & Schwartz, M. K. (2015). The dual challenges of generality and specificity when developing environmental DNA markers for species and subspecies of Oncorhynchus. PLOS ONE, 10(11), e0142008.

Tables

Table 1. Delta AICc values for alternative decay models fit to eDNA data of two species at three different temperatures (°C). For each species and temperature, the bolded delta AICc value denotes the best model. Blanks denote models that did not converge or yielded non-sensical parameter estimates. Two models (Geeraerd with a shoulder, and log-linear with a shoulder) were excluded from the table because of poor fits.

Species	Temp	Biphasic	Geeraerd	Geeraerd (no shoulder)	Log-linear	Log-linear (with tail)	Weibull	Weibull (with tail)
Rainbow Trout	12	3.43		0.93	18.92	**0.00**	9.74	3.79
	18		3.64	0.85	33.00	**0.00**	21.76	3.49
	24			4.70	58.16	6.22	26.18	**0.00**
White Sturgeon	12		3.83	2.55	29.39	**0.00**	21.89	3.29
	18	2.91	2.70	0.01	19.06	**0.00**	9.53	2.76
	24	2.32		**0.00**	23.42	0.45	5.70	0.78

Table 2. Parameter estimates (Est.) and standard errors (SE) for the best log-linear (with tail) eDNA decay model including data for two fish species at three temperatures (°C). Parameters include starting concentration (N_0, copies per liter), the residual concentration at the tail (N_{res}), and the exponential (log-linear) decay constant (k, hr^{-1}). t^* is the time (hr) to reach N_{res}. T_{50} is the estimated half-life (hrs) and T_{90} is the time for 90% decline, based on the log-linear decay constant k.

| Species	Temp	$\log_{10}(N_0)$ Est.	SE	$\log_{10}(N_{res})$ Est.	SE	k Est.	SE	t^*	T_{50}	T_{90}
Rainbow Trout	12	5.7933	0.1904	2.7021	0.1407	0.0338	0.0056	210	20.5	68.1
	18	5.7933	0.1904	2.1483	0.1752	0.0558	0.0056	150	12.4	41.3
	24	5.7933	0.1904	2.1551	0.1752	0.0686	0.0056	122	10.1	33.6
White Sturgeon	12	4.5955	0.0798	1.6468	0.1407	0.0364	0.0053	187	19.1	63.3
	18	4.5955	0.0798	1.0930	0.1752	0.0583	0.0053	138	11.9	39.5
	24	4.5955	0.0798	1.0998	0.1752	0.0711	0.0053	113	9.7	32.4

Figures
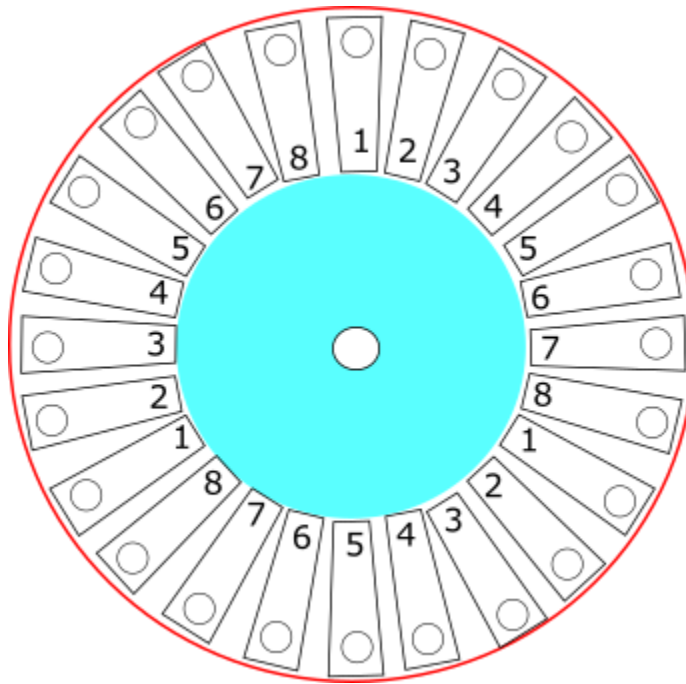
Figure 1. Depiction of the water bath and positioning of the water bottles utilized in the decay experiments. For each temperature treatment, triplicate water samples were collected at intervals of 0.5, 1, 2, 4, 8, 12, and 16 days, denoted as bottles 1-7. Bottle 8 served as a negative control, which was taken from each temperature treatment on days 1, 2, and 4
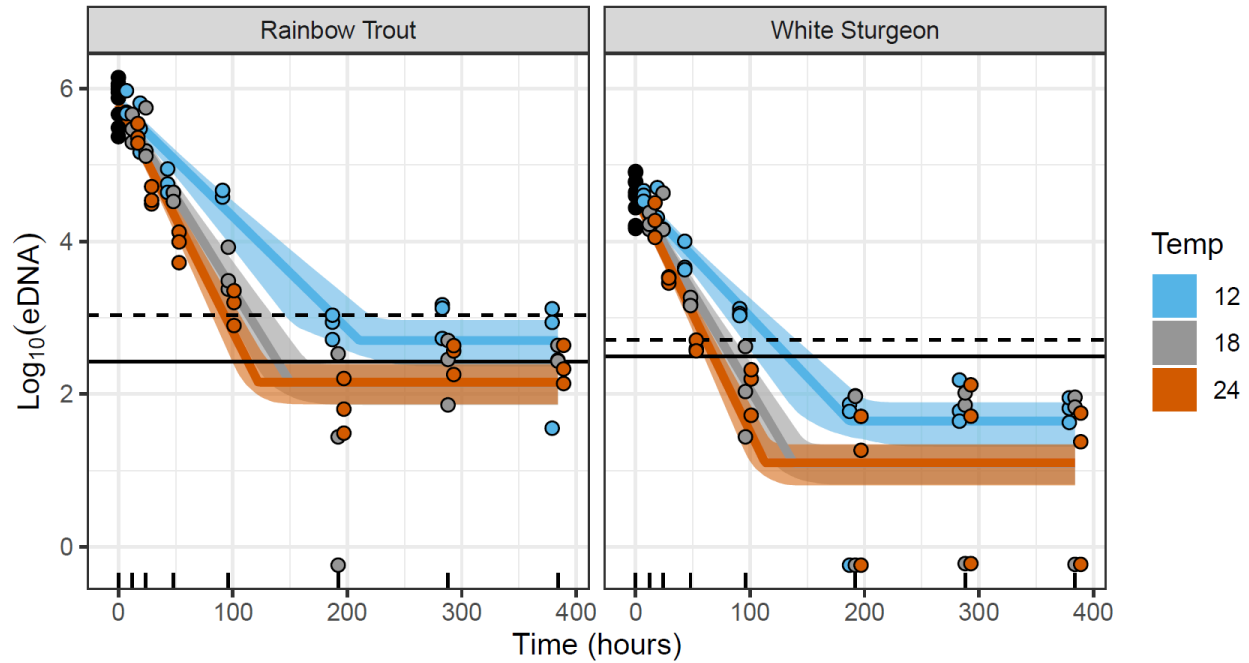
Figure 2. eDNA concentrations through time for rainbow trout and white sturgeon at three different temperatures (C). Black points are the starting concentrations at time=0, colored points represent each temperature treatment, and colored lines are the model predictions (+/- 95% bootstrapped confidence intervals). Points are the means of three technical qPCR replicates, and are offset horizontally for viewing. Horizontal lines are the species-specific limit of detection (solid) and limit of quantification (dashed).
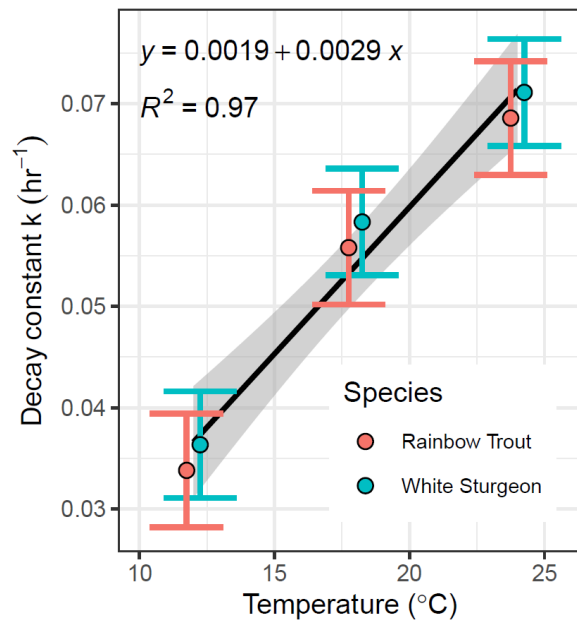
Figure 3. Estimated log-linear decay constants (k, +/-SE) for rainbow trout and white sturgeon at different temperatures. Linear regression (+/- 95% confidence interval) is plotted, along with the regression equation and R^2.

Supplemental Tables

Supplemental Table S1. The species, number, average total length (in millimeters), and population total weight (in kilograms) of the fish population Cal Poly Humboldt Fish Hatchery.

Species	Number	Average TL (mm)	Population Total Weight (kg)
rainbow trout	200	170	10
rainbow trout	300	120	5
rainbow trout	300	120	5
rainbow trout	500	560	150
white sturgeon	750	N/A	54

Supplemental Table S2. Estimates and standard errors (SE) for the exponential decay constant (k, hr^{-1}) for eDNA of two species at three temperatures (°C), based on two different models: a log-linear model with a tail and a simple log-linear model. T_{50} is the estimated half-life (hrs) based on the decay constant k. Both models were fit to the entire dataset and allowed k to vary by species and temperature. For the log-linear model, species and temperature differences in k were not significant ($p < 0.05$).

Species	Temp	log-linear with tail			log-linear		
		k	SE	T_{50}	k	SE	T_{50}
Rainbow Trout	12	0.0338	0.0056	20.5	0.0213	0.0031	32.5
	18	0.0558	0.0056	12.4	0.0242	0.0040	28.6
	24	0.0686	0.0056	10.1	0.0231	0.0040	30.0
White Sturgeon	12	0.0364	0.0053	19.1	0.0207	0.0022	33.5
	18	0.0583	0.0053	11.9	0.0235	0.0034	29.5
	24	0.0711	0.0053	9.7	0.0225	0.0034	30.8

Supplemental Table S3. Proportion of eDNA samples (or total technical replicates) with a concentration greater than the limit of detection (LOD) through time for the different species (Atr=white sturgeon; Omy = rainbow trout) and temperature treatments. Three water samples were taken at each time step (except 9 water samples for time t=0), with 3 technical replicates per water sample.

| Species	Time (hrs)	Prop. of samples (tech. repl.) > LOD		
		12°C	18°C	24°C
Atr	0	1 (1)	1 (1)	1 (1)
Atr	12	1 (1)	1 (1)	1 (1)
Atr	24	1 (1)	1 (1)	1 (1)
Atr	48	1 (1)	1 (1)	1 (0.89)
Atr	96	1 (1)	0.33 (0.22)	0 (0)
Atr	192	0 (0)	0 (0)	0 (0)
Atr	288	0 (0)	0 (0)	0 (0)
Atr	384	0 (0)	0 (0)	0 (0)
Omy	0	1 (1)	1 (1)	1 (1)
Omy	12	1 (1)	1 (1)	1 (1)
Omy	24	1 (1)	1 (1)	1 (1)
Omy	48	1 (1)	1 (1)	1 (1)
Omy	96	1 (1)	1 (1)	1 (1)
Omy	192	1 (1)	0.33 (0.22)	0 (0)
Omy	288	1 (1)	0.67 (0.56)	0.67 (0.44)
Omy	384	0.67 (0.67)	1 (0.78)	0.33 (0.44)

Supplemental Figures
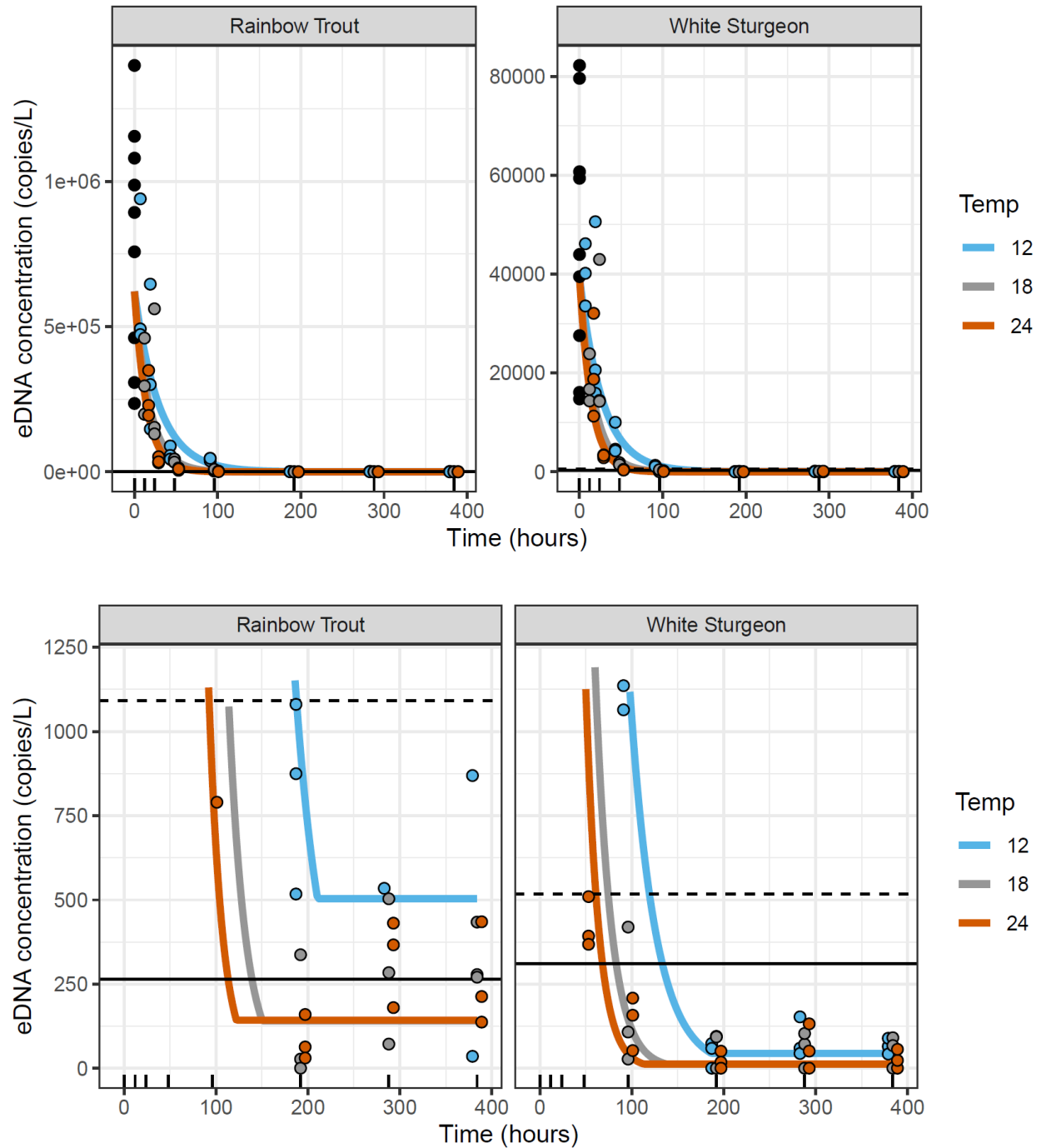
Supplemental Figure S1. eDNA concentrations (copies per liter) through time for rainbow trout and white sturgeon at three different temperatures (C). Black points are the starting concentrations at time=0, colored points represent each temperature treatment, and colored lines are the model predictions. Points are the means of three technical qPCR replicates, and are offset horizontally for viewing. Horizontal lines

are the species-specific limit of detection (solid) and limit of quantification (dashed). Lower panels only display data <1250 copies per liter.
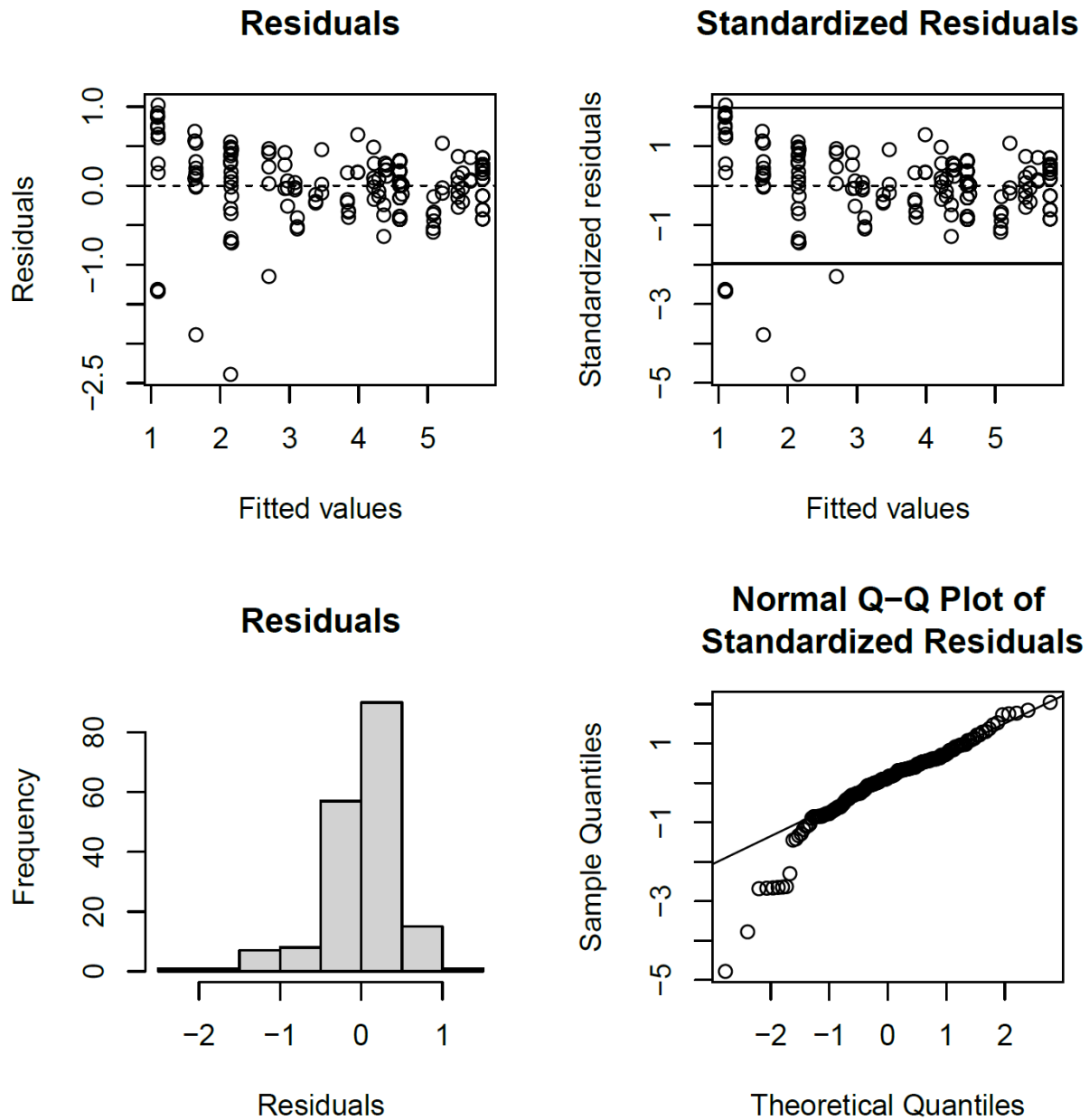
Supplemental Figure S2. Diagnostic plots of residuals for the best eDNA decay model fit to data for all species and temperature treatments.
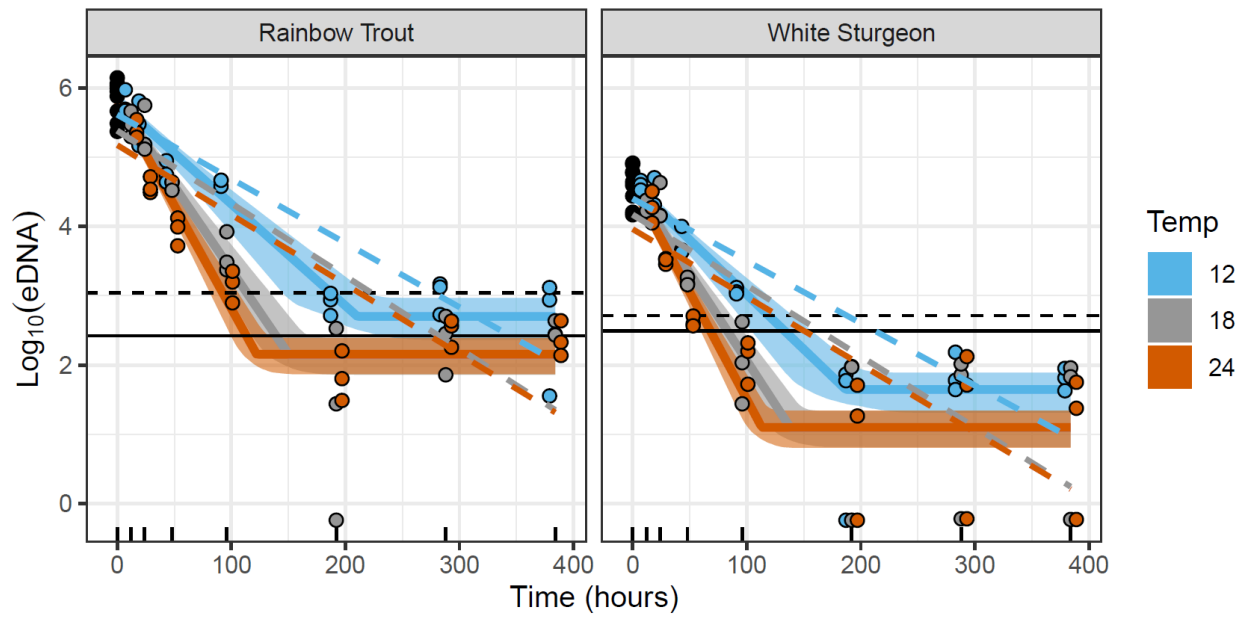
Supplemental Figure S3. eDNA concentrations through time for rainbow trout and white sturgeon at three different temperatures (°C). Black points are the starting concentrations at time=0, colored points represent each temperature treatment, and solid colored lines are the model predictions (+/- 95% bootstrapped confidence intervals). Colored dashed lines are the predictions from a log-linear model. Points are the means of three technical qPCR replicates, and are offset horizontally for viewing. Horizontal lines are the species-specific limit of detection (solid) and limit of quantification (dashed).

Assessing environmental DNA shedding and decay in rainbow trout (*Oncorhynchus mykiss*): an approach for estimating shedding rates that accounts for time-varying decay rates

Andrew P. Kinziger 1,*

Andre Buchheister 1,*

Gavin B. Bandy 1

Braden A. Herman 1

Eric P. Bjorkstedt 2

1 Department of Fisheries Biology, Cal Poly Humboldt, Arcata, CA 95521, USA

2 Southwest Fisheries Science Center, NOAA Fisheries, Trinidad, CA 95570, USA

*Co-first authors and corresponding authors: Andrew.Kinziger@humboldt.edu, Andre.Buchheister@humboldt.edu

Abstract

We estimated environmental DNA (eDNA) shedding and decay rates in rainbow trout (*Oncorhynchus mykiss*) at two densities. We held the trout in tanks at a low density of 5 individuals and a high density of 25, over 96 hours to reach a steady state of eDNA shedding and degradation. After this period, water was siphoned from the tank and placed in separate bags, and periodic water sampling was conducted to measure eDNA levels at the outset and at intervals of 3, 6, 12, 18, 24, 30, 36, 48, 72, and 96 hours post-removal, using quantitative PCR. Our investigation employed a diverse panel of bacterial survival models to characterize eDNA decay, identifying those that account for time-varying decay rates as the most reflective of actual eDNA decay in rainbow trout. Additionally, we present a novel approach for estimating eDNA shedding rates, which considers these time varying decay rates. Shedding rate estimates for rainbow trout are provided, illustrating how decay model selection influences these estimates. We did not find clear evidence that fish density significantly impacts decay rates or shedding rates of eDNA. The research provides insights into the shedding and decay of eDNA in rainbow trout, highlighting the limitations of assuming a log-linear decay in eDNA decay and shedding rate calculations. Accurate estimation of eDNA shedding and decay is required for understanding the dynamics of eDNA in natural settings and consequently characterizing biodiversity.

Introduction

Environmental DNA (eDNA) serves as a powerful tool for aquatic species monitoring by capturing genetic material that organisms release into their surrounding environment (Rodríguez-Ezpeleta et al., 2021). To harness eDNA for species surveillance, it's important to understand the eDNA shedding rate, or the rate that organisms release eDNA. Assessing this rate is pivotal for estimating species abundance and modeling its spatial distribution within an ecosystem (e.g., Carraro et al., 2018). However, accurately determining eDNA shedding rates poses significant challenges, including the measurement of minute quantities shed by individual organisms, the likelihood of pulse release rather than continuous release, and rapid degradation following shedding (Goldberg et al., 2017; Harrison et al., 2019).

Sassoubre et al. (2016, 2017) introduced a method to estimate the eDNA shedding rate in an experimental setting using the conceptual framework of fish tanks as completely mixed batch reactors. The dynamics of eDNA within these tanks was described by a simple differential equation that could be used to calculate shedding rates if information was available on the steady state concentration where the input of eDNA through shedding was equal to the decay. In such conditions, the shedding rate S is the product of the tank's volume, the steady-state eDNA concentration (C_0), and the log-linear decay rate constant (k) which is estimated for a given experiment. However, this model implicitly assumes that all eDNA-bearing particles are homogenous and exhibit a constant, log-linear (i.e., exponential) decay rate.

Although log-linear models are often used in eDNA decay studies (reviewed by Jo and Minamoto, 2021 and Lamb et al., 2022), more recent work has led to emerging consensus that models must account for decay rates which vary over time (Bylemans et al., 2018; Andruszkiewicz Allan et al., 2020; Eichmiller et al., 2016; Shogren, 2018; Herman et al. In Review). For example, including an assumption of

heterogeneity in heat resistance in a model of bacterial mortality leads to a decay model that is not linear on the log-scale (Peleg 1999; Peleg and Cole, 1998, 2000; Mafart et al., 2002). Even though the method by Sassoubre et al. (2017) has been broadly adopted for estimating eDNA shedding rates (e.g, Kirtane et al. 2021, Allan et al., 2021), violation of the log-linear decay assumption can introduce inaccuracies and bias in the shedding rate estimates, and new methods are needed to estimate shedding rates when decay is not log-linear.

In this research we present estimates of eDNA shedding and decay based on experiments with rainbow trout (*Oncorhynchus mykiss*) held at two densities. We first characterize eDNA decay using various bacterial survival models encompassing diverse decay patterns (Geeraerd et al., 2005). Our findings highlight that models accounting for varying decay rates over time offer the most accurate representation of eDNA decay in our study system. We subsequently introduce a simulation approach based on difference equations to estimate shedding rates that consider these time-varying decay rates. We structured our experiments to examine how density, or different numbers of the same size and age fish, affects eDNA shedding rates as previous research has yielded mixed outcomes (Wilder et al., 2023; Nevers et al., 2018). Our study provides shedding and decay rates for rainbow trout, the influence of density on these rates, and illustrates the potential inaccuracies in eDNA shedding rate estimations when assuming a log-linear decay.

Methods

Experimental Design

Experiments were conducted in an indoor wet laboratory at the Cal Poly Humboldt Fish Hatchery, Arcata, California. The experimental design consisted of four circular fiberglass tanks each with 300 L of water (tank dimensions: 58 cm in depth; 89 cm in diameter). Prior to initiating the experiment, all tanks and sampling apparatus were sterilized using a 10% bleach solution, and then rigorously rinsed with chlorinated tap water.

The source of water for the experimental tanks was Fern Lake, a reservoir spanning 2500 m[2,] constructed in 1962. While this reservoir, which has a maximum depth of 4.5 m and an estimated volume of 5000 m[3,] is unpopulated by rainbow trout, it does host communities of macroinvertebrates, zooplankton, and bacteria. Tanks were filled with ⅓ redwood tank water and ⅔ municipal water. We believed sourcing some water from Fern Lake would account for the abiotic and biotic factors common in streams, while dilution with municipal water would minimize filter clogging. Upon filling the tanks with this water, we circulated water using a pump, initiated aeration, and introduced bio balls. We then allowed the system to stabilize for approximately 24 hours prior to adding fish.

Throughout the study duration, we routinely checked water quality parameters, including temperature, pH, dissolved oxygen (mg/l), and ammonia (ppm) to ensure they were within acceptable limits for

rainbow trout. Temperature was monitored at routine intervals in each water bath with HOBO temperature loggers, with temperatures calibrated against a NIST-traceable thermometer (Thermco Products Inc ACC10033SFC). If needed, we would periodically supplement tanks with redwood tank water to maintain consistent 300 L volume.

Rainbow trout sourced from the Cal Poly Humboldt Fish Hatchery were placed into tanks at two densities: a low density of 5 fish and a high density of 25 fish, with duplicate tanks for each density. In the low-density replicates, the fish had a mean mass of 11.3g and 11.32g, with individual mass ranging from 6.9g to 20g and 6.7g to 13.8g, respectively. In contrast, the high-density replicates had a mean mass of 16.2g and 15.6g, spanning from 7g to 29.6g and 8.1g to 30g, respectively. To eliminate potential eDNA contributions from food, fish were kept in a fasting state throughout the experiment but were provided feed prior to its initiation. To prevent unnecessary stress and high levels of eDNA shedding from handling, fish were anesthetized in a bucket using MS222, removed by hand, then weighed, measured and placed into experiment tanks.

The fish were kept in the tanks for a duration of 96 hours, following which 39 1-liter water samples were collected from each tank and stored in sterile Whirl-Pak bags. Water was collected using a pre-established siphon system, set up prior to fish introduction, with the goal of minimizing fish disturbance that may elevate eDNA shedding. After water collection, fish were removed, water was drained, the tanks were mechanically scrubbed, and rinsed with tap water. The tanks were then refilled, the sealed Whirl-Pak bags were reintroduced, water surrounding the bags was circulated with a pump.

Six water samples (e.g., Whirl-Pak bags) were immediately drawn from each tank at time t = 0 to determine steady-state eDNA concentration. For the evaluation of tank-specific eDNA decay rates, triplicate water samples were taken from each tank at intervals of t = 3, 6, 12, 18, 24, 30, 36, 48, 72, and 96 hours. The sampling protocol was designed to include more samples during an expected initial rapid eDNA decline phase, where the bulk of eDNA decays. This enhanced frequency aimed to offer a finer resolution of eDNA decay patterns, which are likely more complex than those depicted by simple log-linear models (Bylemans et al., 2018; Andruszkiewicz Allan et al., 2020).

Upon collection, water samples were immediately filtered under vacuum (pneumatic hand pump (EWK EB0103A)) onto 0.45-micron cellulose nitrate filters (47 mm diameter, Whatman, GE Healthcare) supported by a filter pad (MilliporeSigma™ AP1004700) housed within a sterile funnel . Samples were filtered in their entirety, or until 30 minutes had elapsed. If a sample remained incompletely filtered, the actual filtered volume was recorded for subsequent eDNA concentration volume corrections. Post-filtration, filters were frozen until DNA extraction.

Molecular Methods

eDNA was extracted from filters with a combination of acetone dissolution and QIAGEN DNeasy Blood and Tissue kits. Filter dissolution followed Hallet et al (2012), involving dissolving filters in acetone with the assistance of bead beating and rinsing the sample with 100% anhydrous ethanol to form a pellet of condensed DNA. Pellets were extracted with QIAGEN DNeasy blood and tissue kits following the

manufacturer instructions, with the following modifications: 40 ul of proteinase k and 360 uL of Buffer ATL and samples were eluted into 100ul of buffer AE. DNA extractions were conducted in a dedicated low-copy laboratory in which sterile conditions are maintained by UVC irradiation when the space is not in use.

Concentrations of rainbow trout eDNA were estimated using qPCR reactions, utilizing an assay from published primers and probes (Wilcox, 2015). In each qPCR reaction, 12.5 µL of TaqMan™ Environmental Master Mix 2.0 (Applied Biosystems, 4396838), species-specific assay components for rainbow trout (250 nM probe, 300 nM forward primer, and 600 nM reverse primer), 6 µL of DNA extract, and nuclease-free water were combined to achieve a total reaction volume of 25 µL. qPCR reactions were set up in a dedicated low-copy laboratory (as above) and subsequently performed on an Applied Biosystems QuantStudio™ 3 Real-Time PCR System in a dedicated high copy laboratory. Reactions were cycled as follows: incubation at 95°C for 10 minutes, then 45 cycles of 95°C for 15 seconds and 60 °C for 60 seconds. All qPCR reactions were run in triplicate, and each 96-well reaction plate included negative and positive controls.

To quantify the target DNA concentration in each qPCR reaction, a qPCR standard curve was generated. The standard was extracted from tissue using a Qiagen DNeasy Blood & Tissue Kit (69506), and its concentration determined using ddPCR. The standard curve was constructed across at least six serial 1:10 dilutions, with 4-48 replicates per standard concentration. The limit of detection (LOD; the minimum DNA concentration at which positive detections occur at a rate of at least 95%), and the limit of quantification (LOQ; the minimum concentration for which the coefficient of variation among replicates is less than 0.35) were determined as described by Klymus et al. (2020).

For the rainbow trout assay, the standard curve analysis yielded a slope of -3.40 and an intercept of 38.63 (qPCR efficiency was 96.7%), with LOD and LOQ of 4.61 copies per reaction and 18 copies per reaction, respectively.
DNA copies per liter (copies/L) was calculated as

$$\text{copies/L} = \text{CN/PCR} \times (V_e/V_t) * (1/V_s) \qquad (\text{Eqn. 1})$$

where CN/PCR represents the copy number per qPCR reaction from the standard curve, V_e denotes the elution volume (in µL), V_t represents the template volume (in µL), and V_s indicates the volume of water filtered (in liters). Data from one water sample was excluded because all three technical replicates had quantification cycle or Cq values greater than the intercept. In five instances, we excluded data from a technical replicate because they differed in Ct value by more than three from both of the other technical replicates. Concentrations from the triplicate qPCR technical replicates were averaged into an estimate of sample concentration for analysis. Although we report LOD and LOQ, we did not exclude any samples from the analysis if they fell below these limits, following previous studies (e.g., Pilliod et al., 2014; Maruyama et al., 2014; Minamoto et al., 2017; Bylemans et al., 2018; Jo et al., 2019).

Decay Analysis

Nine alternative decay models were initially fit to each species-temperature dataset using functions in the R package *nlsMicrobio* (Baty and Delignette-Muller, 2014), which includes implementations of most of the bacterial survival (i.e., decay) models described in Geeraerd et al., 2005. The nine decay models fit included: (1) Geeraerd with shoulder and tail (Geeraerd et al., 2000), (2) Geeraerd without a tail (Geeraerd et al., 2000), (3) Geeraerd without a shoulder (Geeraerd et al., 2000), (4) Weibull (Mafart et al., 2002), (5) Weibull with a tail (Albert and Mafart, 2005), (6) log-linear with a shoulder (Geeraerd et al., 2000), (7) log-linear with a tail (Geeraerd et al., 2000), (8) a biphasic (Cerf 1977), and (9) log-linear. Estimated eDNA concentrations (copies/L) were $\log_{10}$ transformed to stabilize the variance. Models were fit using non-linear regression, and compared using Akaike's Information Criterion corrected for small sample sizes (AICc) and AICc differences (delta AICc) (Burnham and Anderson, 2002). Three of these models account for a "shoulder", or period of non-changing eDNA concentration at the start of the experiment; however, these models (Geeraerd with shoulder and tail, Geeraerd without a tail, and log-linear with a shoulder) yielded nonsensical parameter estimates (i.e., negative shoulder lengths), so they were excluded from further consideration and model comparison, leaving six decay models that were considered.

Shedding Analysis

Shedding rate assuming log-linear decay

If decay of eDNA is assumed to be log-linear and the experimental tanks are treated as completely mixed batch reactors at steady state at time t=0, then shedding rates (S) can be calculated as (Sassoubre et al., 2016; Sansom and Sassoubre, 2017):

$$S = kC_0V \qquad\qquad \text{(Eqn. 2)}$$

where S is the shedding rate (copies/h), k is the first-order decay rate (h^{-1}), C_0 is the initial concentration of eDNA at t=0 (copies/L), and V is the volume of the tank (L). Typically, S is calculated using the mean of the empirical C_0 values, and k is estimated via linear regression by modeling $\log_e(C/C_0)$ through time and forcing the y-intercept to be equal to 0 (Sansom and Sassoubre, 2017, Allan et al., 2021). This is equivalent to forcing a regression of $\log_e(C)$ versus time to go through the mean of the empirical C_0 values. Some studies (e.g., Nevers et al. 2018), estimate C_0 as the y-intercept of that linear regression, and this alternative was also explored in the supplemental material.

The standard error of S (δS) is calculated by propagating the errors of each component of the equation (Sansom and Sassoubre, 2017):

$$\frac{\delta S}{S} = \sqrt{\left(\frac{\delta C_0}{C_0}\right)^2 + \left(\frac{\delta k}{k}\right)^2 + \left(\frac{\delta V}{V}\right)^2} \qquad \text{(Eqn. 3)}$$

where δX is the standard error associated with each variable X. For our calculations, we assumed δV was 0. The 95% CI of S for each trial was approximated as $2*\delta S$, and the values were also converted to the $\log_{10}$ scale.

Shedding rate assuming non-log-linear decay

If decay of eDNA deviates from a first-order, log-linear model, then the shedding rate from Equation 2 would be inappropriate. We developed a simulation model to estimate the eDNA shedding rate for fish when decay was not log-linear. We used difference equations to populate a matrix of DNA abundance at age through time:

$$N_{t,\,a=1} = S \times \Delta_t\,, \qquad \text{if } t<t_0 \qquad \text{(Eqn. 4)}$$

$$N_{t+1,\,a+1} = N_{t,\,a} \times f(t) \qquad \text{(Eqn. 5)}$$

where $N_{t,a}$ is a matrix of DNA abundance with time t (hrs) as rows and DNA age a (hrs) as columns, S is the discrete shedding rate (h^{-1}) assumed to be constant over time, and Δ_t is the time-step in the simulation (set to 1 hour). At each timestep, a new "cohort" of DNA is produced via shedding with $a=1$ (Equation 4), and then that cohort or packet of DNA particles is tracked individually through time as it degrades based on the survivorship function f(t) representing some proportion of individuals that survive each time step (Equation 5; Supplemental Figure S1). At some specific time t_0, the input of DNA into the system from shedding is set to zero, simulating the removal of fish from an experimental system and the start of the decay-only phase of an experiment. The total abundance of DNA at any time t (N_t) is the sum of row t in the $N_{t,a}$ matrix. Concentrations (C_t) were calculated as N_t divided by the volume of the theoretical system (300 L in this case). The simulation was run for 96 hrs (4 days) with shedding and then for 96 hours without shedding, with t_0 set equal to 0, consistent with the experimental design that had been implemented.

The generic function, f(t), can represent any decay model (e.g., Geeraerd et al., 2005) when expressed as a survivorship term or a proportion of particles that survive a given time step. For our simulations, we modeled decay using a time-dependent Weibull model (Mafart et al., 2002), which was found to be the overall best model across experimental trials:

$$N_t = N_0 e^{-kt^p} \qquad \text{(Eqn. 6)}$$

where k is a decay rate and p is a shape parameter. Note that if $p=1$, the model is equivalent to the standard first-order, log-linear decay model. Equation 6 can be re-expressed as (Mafart et al., 2002):

$$N_t = N_0 \times 10^{-\left(\frac{t}{\delta}\right)^p} \qquad \text{(Eqn. 7)}$$

where t is time (hours), δ is the time (in hours) it takes for the first decimal reduction in the number of DNA copies (i.e., the time to go from N_0 to $N_0/10$, representing the first 90% loss in eDNA), and p is the same shape parameter. If $p<1$ the survival curve of $\log_{10}(N)$ versus time has an upward concave shape, if $p>1$ that curve has a downward concave shape, and if p=1 the curve is log-linear (Mafart et al. 2002). This formulation of the Weibull model (Equation 7) provides allows us to define:

$$f(t) = 10^{-\left(\frac{t}{\delta}\right)^p} \qquad \text{(Eqn. 8)}$$

35

Equation 8 is substituted into equation 5, to yield:

$$N_{t+1,\,a+1} = N_{t,\,a} \times 10^{-\left(\frac{\Delta_t}{\delta}\right)^p} \qquad \text{(Eqn. 9)}$$

where everything as defined before, except the t in the exponent of Equation 8 is changed to Δ_t (the time-step in the simulation). Although we set Δ_t equal to 1 hour, it would be possible to model minute-by-minute changes in DNA abundance if $\Delta_t = 1/60$.

Our final simulation model (Equations 4 and 9) were thus dependent on three parameters: S, δ, and p. However, given the correlation between δ and p in the decay model, p is often fixed to an empirical measurement (Mafart et al., 2002); we fixed the value of p for each experimental trial to the value estimated by the Weibull model fit for that trial (as described in the Decay Analysis section). S and δ were estimated by fitting to the observed $\log_{10}$ eDNA concentrations through time using maximum likelihood assuming a normal error distribution with standard deviation, σ. Logging was necessary to homogenize the variance. Parameters were estimated using the *optim* function in R (R Core Team, 2023), and the standard errors (SE) of the estimates were calculated as the square root of the inverse of the Hessian matrix. Approximate 95% confidence intervals (CI) for the parameter estimates were calculated on the $\log_{10}$ scale using 2*SE for each parameter; overlapping 95% CI for parameters indicated non-significant differences between trials. Parameters and their 95% CI were back transformed to also present the estimates on the original scale. For comparison, shedding rate was also estimated using the simulation approach by forcing a first-order, log-linear decay process by setting $p=1$, instead of using the empirically estimated p values.

Results

Decay Models and Rate

Six competing decay models were fit to each density-replicate trial and considered in the model selection process (Table 1). The Weibull model was the best fit based on AICc for the two low density replicates, and it was the second best for the high density replicate 1 trial with a $\Delta=0.18$, with the log-linear model slightly better (Table 1). The Weibull (with tail) model was the best for the high density replicate 2 group, with the basic Weibull being the second best ($\Delta=2.62$). Generally, every alternative model received a notable level of support based on AICc, with most Δ values less than 4. However the Weibull model was the best-performing model across all the trials. The log-linear model was the best for high-density replicate 1, but for the high-density replicate 2 data this model was notably less effective, as indicated by a Δ value of 30.47.

Fits of the Weibull model using nonlinear regression (Figure 1) and the corresponding parameter estimates (Table 2) demonstrates a non-log-linear form of eDNA decay across all 4 trials. The

non-linearity is indicated by estimates (and confidence intervals) of the shape parameter p being less than 1 (Table 2), and this was most pronounced for trial 25.2. Estimates of δ (i.e., the first decimal reduction, or 90% decline, in eDNA) ranged from 22.9 to 34.4 hrs for trials 5.1, 5.2, and 25.1, but was only 1.98 hrs for trial 25.2 (Table 2).

There was not a clear effect of fish density on decay rate, but there was high variability between the two trials for the 25-fish treatment. Overall, estimates for p and δ were not significantly different among trials 5.1, 5.2, and 25.1 (based on overlapping 95% CIs; Table 2). However, the δ estimate for trial 25.2 was significantly different from that for 25.1, and nearly significantly different from the estimates for trials 5.1 and 5.2.

Log-linear models did not fit the observed data well (Figure 1). Analysis of model residuals indicated non-linear residual patterns that violated the assumptions of linearity and normality, therefore these model fits are not appropriate for inference. Despite these problems and assumption violations, we include the model results (Table 2) given the prevalence of these models in the eDNA decay literature and for comparative purposes for the shedding calculations.

Shedding rate

The simulation model generated predictions of eDNA concentrations that were extremely similar to the original Weibull model fits across all four trials (Figure 1). The estimates of δ using the simulation model were not significantly different from the original estimates based on the overlapping 95% confidence intervals (Table 2). Similarly, the predicted $\log_{10}(C_0)$ values from the simulation model also fell within the 95% confidence interval of the estimates generated by the original Weibull model (Table 2).

Estimates of shedding rates of eDNA across the four trials ranged from 1,540-10,000 copies/g/h based on the simulation model with Weibull decay dynamics, whereas the log-linear estimates of shedding rate ranged from 2,520-6,070 copies/g/h (Table 3, Figure 2). Within each trial, the two methods of estimating shedding rate were not significantly different based on overlapping confidence intervals. However, the percentage difference in the point estimates was substantial. For example, shedding rates based on the Weibull simulation were 54% and 65% higher (for trials 5.1 and 25.2 respectively) than for the log-linear S estimates. For trials 5.2 and 25.1, the Weibull estimates were 8 and 39% lower than the log-linear approach.

Fish density did not have a significant effect on the shedding rates for either of the estimation methods given the overlap of the 95% confidence intervals (Table 3, Figure 2). The greatest difference in S estimates was between the two replicates of the high density trials, with S from 25.2 being 6.5 times greater than S for 25.1, based on the Weibull simulation.

Discussion

The findings of this study provide insights into the dynamics of eDNA shedding and decay in rainbow trout, and clearly illustrate the inadequacy of log-linear models in accurately depicting and estimating eDNA decay and shedding rates. For eDNA decay, we uncovered a pattern of faster initial decay transitioning to slower decay over time. We introduce a novel simulation approach for estimating shedding rates under non-log-linear decay conditions and illustrate how the choice of decay model impacts the estimation of shedding rates. Lastly, we did not find clear evidence that fish density significantly impacts decay rates or shedding rates of eDNA, but this comparison would have benefitted from having more replicates.

eDNA Decay

The superiority of non-log-linear models over log-linear decay models for describing eDNA dynamics in 3 of 4 trials highlights the importance of exploration of more complex models for describing eDNA decay from aquatic environments. This finding is in line with previous studies that have advocated for the use of more complex models over the conventional log-linear approach for biological decay processes (Eichmiller et al., 2016; Bylemans et al., 2018; Shogren, 2018; Wei et al., 2018; Andruszkiewicz Allan et al., 2020; Herman et al., In Review). Our research reinforces this view, as indicated by the non-linear shape of eDNA decay captured in the Weibull model fits and the inadequacy of log-linear models, which failed to capture the eDNA dynamics in most of our trials. This stands in contrast to the traditional use of log-linear models for eDNA decay that assume a uniform rate of decay throughout the duration, as explored in the reviews by Jo and Minamoto (2021) and Lamb et al. (2022).

Reflecting a growing consensus on decay dynamics, our findings offer further support that eDNA decay more commonly exhibits a rapid initial decay phase after which decay proceeds more slowly (Eichmiller et al., 2016; Bylemans et al., 2018; Shogren, 2018; Wei et al., 2018; Andruszkiewicz Allan et al., 2020; Herman et al., In Review). Notably, our investigation indicates that the Weibull decay model, which permits time-variant decay, most accurately captures the eDNA decay process in our system. This model suggests a decelerating decay rate over time when the shape parameter, p, is below 1, an observation that is consistent with the findings presented by Bylemans et al. (2018).

Overall, the log-linear rates of eDNA decay we estimated align with the reported range found in existing literature (Jo & Minamoto, 2021; Lamb et al., 2022), but we note that the log-linear model was generally not appropriate based on non-linear patterns in the residuals for the trials. Our estimates of time for the first 90% decline in eDNA (i.e., δ), were consistent with the time for a 90% decline based on the log-linear decay rates (Jo & Minamoto, 2021; Lamb et al., 2022). Estimates of eDNA decay in the literature using a Weibull model are rare. Bylemans et al. (2018) reported Weibull shape parameter p estimates for nuclear DNA ranging from approximately 0.02 - 0.08, approximately 1 order of magnitude smaller than our p estimates that ranged from 0.3-0.7; however, these differences could be associated with differences in temporal resolution (they sampled for >300 hours but only had 3 sample events <100

hours) or their results could have been biased by estimated C_0 values that were all higher than the range of observed concentrations measured at t=0 (Bylemans et al., 2018).

We found that the impact of fish density on decay rates of eDNA was unclear in our study, with the differences in δ not significantly different between the high and low density trials. However, one of our higher density trials (25.2) was suggestive of a higher rate of decay that was nearly significant (with 1.98 hrs for the first 90% decline), more consistent with the finding by Jo et al. (2019) that decay rate was positively associated with fish size. Interestingly, the only significant difference in δ was between our two high-density trials, which could be due to stochastic variations in environmental conditions or biological responses within the tanks, underscoring the complexity and unpredictability of eDNA dynamics in natural settings.

Shedding rates

The shedding rate of eDNA, as deduced through the methodology of Sassoubre et al. (2016), is directly related to the estimated rate of eDNA decay, positioning the decay model as a crucial element in the interpretation of shedding rates. This is evident from the variation in shedding rates that arise from different decay models that were fit to the eDNA data. Notably, when the Weibull shape parameter, p, is set to 1, decay rates deduced from the Weibull and the log-linear models are identical. However, deviation from p=1 leads to discrepancies in decay dynamics between these models. The greater the deviation from 1, the more pronounced the divergence from the predictions of the log-linear model is anticipated to be, as most clearly indicated with trial 25.2.

Our study's estimates of shedding rates for rainbow trout were approximately $10^{3.5}$ to 10^4 copies per hour per gram, which is on the low end of published values. The literature reports highly variable eDNA shedding rates ranging from 10^4 to 10^{10} copies per hour per gram, but most estimates tend to range from 10^5 to 10^7 (Maruyama et al., 2014; Klymus et al., 2015; Kirtane et al., 2017; Nevers et al., 2018; Jo et al., 2019; Jo et al., 2020; Wilder et al., 2023). This variability in shedding rates (in copies per hour per gram) can be partially contributed to differences in study designs, given its dependence on species (Allan et al., 2021), fish size (Jo et al., 2019), and life-stage (Maruyama et al., 2014).

The observed differences in shedding rate estimates can ultimately be ascribed to two predominant factors: the initial steady state eDNA concentration (C_0) and the log-linear decay constant (k), given that S = k*C_0*V under the assumption of log-linear decay (Sassoubre et al., 2016; Sansom et al., 2017). In our study, the lower shedding rate is attributable to an appreciably lower C_0, estimated to be around 10^4, which is two to three orders of magnitude less than the 10^6 to 10^7 range more commonly reported in other studies. However, our decay constant (k) values are consistent with those found in previous literature (Jo & Minamoto, 2021; Lamb et al., 2022). The primary difference between our findings and prior research, therefore, resides in the lower C_0 concentration estimates we observed.

We suspect that C_0 eDNA concentrations in our study were relatively low due to our use of a protocol designed to minimize disturbance of fish during water collection. By using a pre-installed siphon system for water collection, we were able to collect water for the decay-phase of the experiment without

disturbing or removing the fish. This approach contrasts with other studies in which handling of fish might induce additional stress—and consequently higher eDNA release (Klymus et al., 2015; Maruyama et al., 2014; Sassoubre et al., 2016; Takahara et al., 2012; Li et al., 2019)—through netting fish out of the tanks, which can result in elevated eDNA levels not representative of a steady state. Maruyama et al. (2014) observed a decline of about half between the initial release of fish and the later stable state concentration reached after several days. Another experimental condition that is critical is feeding practices—where some studies provided food while others did not—and this factor may alter eDNA concentrations by up to 10-fold (Klymus et al., 2015; Nevers et al., 2018; Jo et al., 2019).

Our study detected no significant changes in the shedding rate (copies per hour per gram) with varying numbers of juvenile rainbow trout held in standard-sized tanks, aligning with the findings of Nevers et al. (2018), and the findings of Wilder et al. (2023) who observed no significant differences in shedding rates across different numbers of individuals per tank in larval muskellunge (*Esox masquinongy*). However, Wilder et al. (2023) also noted lower per-gram shedding rates with increased fish density in juvenile muskellunge. Minegishi et al. (2019) reported consistent eDNA shedding (copy numbers per hour) in aquarium groups of 3, 10, and 30 fish but identified markedly higher shedding rates in aquariums with 60 and 120 juveniles. To further clarify the relationship between fish quantity and shedding rate, future research must carefully design experiments with adequate replication and consider density impacts, as increased interaction from closer spacing is expected to augment shedding due to enhanced contact among fish in confined spaces.

Steady State eDNA Concentration (C_0)

The accurate quantification of C_0 is crucial for calculating both shedding and decay rates. The shedding rate is directly proportional to C_0. The significance of C_0 is highlighted by the magnitude of the value that this parameter can taken on, typically 10^6 to 10^7 (Maruyama et al., 2014; Klymus et al., 2015; Kirtane et al., 2017; Jo et al., 2019; Jo et al., 2020; Wilder et al., 2023). Future studies should ensure precise C_0 measurements through increased sampling frequency during the steady state period, as well as considering factors like fish fasting or feeding, water recirculation, and the pulsed eDNA release upon fish removal.

Tank volume (V) and decay constant (k) are also directly proportional to shedding rate, though tank volume is expected to have a minimal impact due to its relatively small magnitude and overall variation in experimental setups, typically ranging from 30 liters to several hundred liters (Kirtane et al., 2021; Allen, 2021). Further, the measurement of tank volume is notably more straightforward than that of eDNA concentration. The decay constant's influence is posited to be more significant than volume but less so than C_0.

C_0 is also pivotal in determining decay rate, which in turn affects shedding rate. C_0 is generally taken as the mean eDNA concentration at the transition point between the initial steady state phase and the decay phase (e.g., the point at which fish are removed). Overestimation skews decay rates, leading to inflated shedding rates and the inverse for underestimation. While using the y-intercept from decay fits offers an alternative for C_0 estimation, we recommend relying on the mean from several steady state measurements, which are less dependent on the specific pattern of eDNA decay.

Study Design Considerations

In contrast to approaches that take sequential samples over time from a common pool, we isolated individual samples at T0. While this does have some implications for how the data are interpreted (as independent realizations of a decay process), we do not expect this to have any meaningful impact on the nature of the results or the conclusions that emerge from this work.

The simulation model developed here provides a flexible approach to estimate shedding rates when decay is not log-linear and a simple analytical equation (e.g., Sassoubre et al., 2016) is not available. To help verify our simulation approach, we constrained the Weibull shape parameter to be 1 (making it equivalent to the log-linear decay model), and we generated shedding estimates that were within 2.5% of standard calculations (Suppl. Figure S4); these small differences were likely due to the discrete, hourly time-step used in our simulation model, which differs from the continuous-time approach. Although we used a Weibull decay model for our simulations, any number of alternative decay models could be implemented through appropriate modification of Equation 5, providing opportunity for further exploration of the influence of the decay model on shedding estimates.

Our Weibull-based shedding estimates did not differ significantly from the standard log-linear estimate for each trial, but we caution against assuming the standard approach provides a reasonable approximation of shedding rate, particularly in cases where the functional form of the decay process deviates substantially from a log-linear relationship between eDNA concentration and time. The lack of significant differences between approaches (Figure 2) probably resulted from the different C_0 values. For the log-linear shedding estimate, C_0 was the empirical mean at t=0 as is common in the literature (e.g., Kirtane et al., 2017; Sansom and Sassoubre, 2017; Andruszkiewicz Allan et al., 2020), but this C_0 was substantially higher than the C_0 values estimated by the simulations which were fit to the entire data set (Supplemental Figure S3). We were not able to constrain the simulation through the empirical mean C_0 (e.g., by modeling $\log_{10}[C/C_0]$) because the shedding rate would no longer be singularly identifiable. When we estimated C_0 as the y-intercept of the log-linear regression, the resulting log-linear shedding rate estimates were significantly lower than our Weibull-based simulation estimates (Supplemental Figure S4), indicating that both the value of C_0 used as well as the decay model could have significant effects on shedding rate estimates.

Conclusions

Ensuring proper model specifications for describing eDNA decay and shedding carries profound implications, as it directly shapes our understanding of the eDNA decay and shedding processes themselves and the factors influencing them. Ultimately this has implications for the use of eDNA as a tool for monitoring aquatic species. eDNA shedding rate is challenging to estimate because of pulsed releases of eDNA (rather than continuous), the quantities being measured are small, and eDNA degradation can be rapid and dynamic. There are now large numbers of studies describing eDNA decay, but fewer have estimated shedding rates and this is an important area of study (Jo & Minamoto, 2021;

Lamb et al., 2022). The discrepancy in decay and shedding rates within a controlled laboratory setting also points to the potential for high variability in eDNA signals in natural water bodies. Our rates may not be reflective of rates in the wild, because we are missing things like UV, microbial degradation, and natural feeding. Future research will need to build on laboratory studies like this to examine decay and shedding in natural settings.

Author Information

The study design was conceived by all authors. GBB performed the experiments and molecular analysis, while AB conducted the statistical analysis. The first draft of the manuscript was primarily written by APK, with contributions from all authors.

References

Albert, I., & Mafart, P. (2005). A modified Weibull model for bacterial inactivation. *International Journal of Food Microbiology*, 100, 197–211. https://doi.org/10.1016/j.ijfoodmicro.2004.10.016

Allan, E. A., Zhang, W. G., Lavery, A. C., & Govindarajan, A. F. (2021). Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. *Environmental DNA*, 3, 492–514. https://doi.org/10.1002/edn3.141

Baty F., Delignette-Muller, M. (2014). nlsMicrobio: data sets and nonlinear regression models dedicated to predictive microbiology. R package version 0.0-3.

Burnham, K.P. and Anderson, D.R. (2002) *Model Selection and Inference: A Practical Information-Theoretic Approach*. 2nd Edition, Springer-Verlag, New York. http://dx.doi.org/10.1007/b97636

Bylemans, J., Furlan, E. M., Gleeson, D. M., Hardy, C. M., & Duncan, R. P. (2018). Does size matter? An experimental evaluation of the relative abundance and decay rates of aquatic environmental DNA. *Environmental Science & Technology*, 52(11), 6408–6416. https://pubs.acs.org/doi/abs/10.1021/acs.est.8b01071

Carraro, L., H. Hartikainen, J. Jokela, E. Bertuzzo, and A. Rinaldo. 2018. Estimating species distribution and abundance in river networks using environmental DNA. *Proceedings of the National Academy of Sciences*, 115(46):11724–11729. https://doi.org/10.1073/pnas.1813843115

Caza-Allard, I., Laporte, M., Côté, G., April, J., & Bernatchez, L. (2021). Effect of biotic and abiotic factors on the production and degradation of fish environmental DNA: An experimental evaluation. *Environmental DNA*, 4, 453-468. https://doi.org/10.1002/edn3.266

Cerf, O. (1977). Tailing of survival curves of bacterial spores. *Journal of Applied Bacteriology*, 42, 1–20. https://doi.org/10.1111/j.1365-2672.1977.tb00665.x

Eichmiller, J. J., Best, S. E., & Sorensen, P. W. (2016). Effects of temperature and trophic state on degradation of environmental DNA in lake water. *Environmental Science & Technology*, 50(4), 1859-1867. https://pubs.acs.org/doi/abs/10.1021/acs.est.5b05672

Geeraerd, A. H., Herremans, C. H., & Van Impe, J. F. (2000). Structural model requirements to describe microbial inactivation during a mild heat treatment. *International Journal of Food Microbiology*, 59, 185–209. https://doi.org/10.1016/s0168-1605(00)00362-7

Geeraerd, A. H., Valdramidis, V. P., & Van Impe, J. F. (2005). GInaFiT, a freeware tool to assess non-log-linear microbial survivor curves. *International Journal of Food Microbiology*, 102(1), 95–105. https://doi.org/10.1016/j.ijfoodmicro.2004.11.038

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., Spear, S. F., McKee, A. M., Oyler-McCance, S. J., Cornman, R. S., Laramie, M. B., Mahon, A. R., Lance, R. F., Pilliod, D. S., Strickler, K. M., Waits, L. P., Fremier, A. K., Takahara, T., Herder, J. E., & Taberlet, P. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution,* 7(11), 1299–1307. 10.1111/2041-210X.12595

Hallett, S. L., Ray, R. A., Hurst, C. N., Holt, R. A., Buckles, G. R., Atkinson, S. D., & Bartholomew, J. L. (2012). Density of the Waterborne Parasite Ceratomyxa shasta and Its Biological Effects on Salmon. *Applied and Environmental Microbiology*, 78(10), 3724–3731. https://doi.org/10.1128/AEM.07801-11

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: Biological Sciences, 286, 20191409. https://doi.org/10.1098/rspb.2019.1409

Herman, B. A., Bandy, G. B., Bjorkstedt, E. P., Buchheister, A., & Kinziger, A. P. (In preparation). Temperature-dependence and long-term persistence demonstrate the value of dynamic decay models for fish environmental DNA.

Jo, T., Murakami, H., Yamamoto, S., Masuda, R., & Minamoto, T. (2019). Effect of water temperature and fish biomass on environmental DNA shedding, degradation, and size distribution. *Ecology and Evolution*, 27(1), 112–212. https://doi.org/10.1002/ece3.4802

Jo, T. , & Minamoto, T. (2021). Complex interactions between environmental DNA (eDNA) state and water chemistries on eDNA persistence suggested by meta‑analyses. *Molecular Ecology Resources*, 21, 1490–1503. 10.1111/1755-0998.13354

Kirtane, A., Wilcox, T. M., Carim, K. J., McKelvey, K. S., Young, M. K., & Schwartz, M. K. (2021). Quantification of Environmental DNA (eDNA) shedding and decay rates for three commercially harvested fish species and comparison between eDNA detection and trawl catches. *Environmental DNA*, 3, 1142-1155. https://doi.org/10.1002/edn3.236

Klymus, K. E., Merkes, C. M., Allison, M. J., et al. (2020). Reporting the limits of detection and quantification for environmental DNA assays. *Environmental DNA*, 2, 271–282. https://doi.org/10.1002/edn3.29

Lamb, P. D., Fonseca, V. G., Maxwell, D. L., & Nnanatu, C. C. (2022). Systematic review and meta-analysis: Water type and temperature affect environmental DNA decay. *Molecular Ecology Resources*, 22, 2494–2505. https://doi.org/10.1111/1755-0998.13627

Li, J., Lawson Handley, L. J., Harper, L. R., et al. (2019). Limited dispersion and quick degradation of environmental DNA in fish ponds inferred by metabarcoding. *Environmental DNA*, 1, 238–250. https://doi.org/10.1002/edn3.24

Mafart, P., Couvert, O., Gaillard, S., & Leguerinel, I. (2002). On calculating sterility in thermal preservation methods: Application of the Weibull frequency distribution model. *International Journal of Food Microbiology*, 72, 107–113. https://doi.org/10.1016/s0168-1605(01)00624-9

Maruyama, A., Nakamura, K., Yamanaka, H., Kondoh, M., & Minamoto, T. (2014). The release rate of environmental DNA from juvenile and adult fish. *PLoS ONE*, 9(12), e114639. https://doi.org/10.1371/journal.pone.0114639

Minamoto, T., Fukuda, M., Katsuhara, K. R., Fujiwara, A., Hidaka, S., Yamamoto, S., Takahashi, K., & Masuda, R. (2017). Environmental DNA reflects spatial and temporal jellyfish distribution. *PLOS ONE*, *12*(2), e0173073. https://doi.org/10.1371/journal.pone.0173073

Minegishi, Y., Wong, M. K., Kanbe, T., Araki, H., Kashiwabara, T., Ijichi, M., Kogure, K., & Hyodo, S. (2019). Spatiotemporal distribution of juvenile chum salmon in Otsuchi Bay, Iwate, Japan, inferred from environmental DNA. *PLOS ONE*, 14(9): e0222052. https://doi.org/10.1371/journal.pone.0222052

Nevers, M. B., Byappanahalli, M. N., Morris, C. C., Shively, D., Przybyla-Kelly, K., Spoljaric, A. M., Dickey, J., & Roseman, E. F. (2018). Environmental DNA (eDNA): A tool for quantifying the abundant but elusive round goby (Neogobius melanostomus). *PLOS ONE*, *13*(1), e0191720. https://doi.org/10.1371/journal.pone.0191720

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., & Waits, L. P. (2014). Factors influencing detection of eDNA from a stream-dwelling amphibian. *Molecular Ecology Resources*, *14*(1), 109–116. https://doi.org/10.1111/1755-0998.12159

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rodríguez-Ezpeleta, N., L. Zinger, A.P. Kinziger, H.M. Bik, A. Bonin, E. Coissac, B.C Emerson, C. Martins Lopes, T.A. Pelletier, P. Taberlet, S. Narum. (2021). Biodiversity monitoring using environmental DNA. *Molecular Ecology Resources*, 21, 1405–1495. https://doi.org/10.1111/1755-0998.13399

Sansom, B. J., & Sassoubre, L. M. (2017). Environmental DNA (eDNA) Shedding and Decay Rates to Model Freshwater Mussel eDNA Transport in a River. *Environmental Science & Technology*, *51*(24), 14244–14253. https://doi.org/10.1021/acs.est.7b05199

Sassoubre, L. M., Yamahara, K. M., Gardner, L. D., Block, B. A., & Boehm, A. B. (2016). Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. *Environmental Science & Technology*, 50(19), 10456–10464. https://pubs.acs.org/doi/abs/10.1021/acs.est.6b03114

Shogren, A. J., Tank, J. L., Egan, S. P., August, O., Rosi, E. J., Hanrahan, B. R., Renshaw, M. A., Gantz, C. A., & Bolster, D. (2018). Water flow and biofilm cover influence environmental DNA detection in recirculating streams. *Environmental Science & Technology*, 52(15), 8530–8537. https://pubs.acs.org/doi/abs/10.1021/acs.est.8b01822

Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., & Kawabata, Z. (2012). Estimation of fish biomass using environmental DNA. *PLoS ONE*, 7, e35868. https://doi.org/10.1371/journal.pone.0035868

Wilcox, T. M., Carim, K. J., McKelvey, K. S., Young, M. K., & Schwartz, M. K. (2015). The dual challenges of generality and specificity when developing environmental DNA markers for species and subspecies of *Oncorhynchus*. *PLOS ONE*, 10(11), e0142008. https://doi.org/10.1371/journal.pone.0142008

Wei, N., Nakajima, F., & Tobino, T. (2018). Effects of treated sample weight and DNA marker length on sediment eDNA based detection of a benthic invertebrate. *Ecological Indicators*, 93, 267–273. https://doi.org/10.1016/j.ecolind.2018.04.063

Wilder, M. L., Farrel, J. M., & Green, H. C. (2023). Estimating eDNA shedding and decay rates for muskellunge in early stages of development. *Environmental DNA*, 5, 251-263. https://doi.org/10.1002/edn3.349

Tables

Table 1. Delta AICc (Δ) values for alternative decay models fit to rainbow trout eDNA data at two different density treatments (5 or 25 fish per tank) and two replicates (rep.). For each density-replicate combination, the bolded values denote models with Δ<1.

Density	Rep.	Biphasic	Geeraerd (no shoulder)	Log-linear	Log-linear (with tail)	Weibull	Weibull (with tail)
5	1	1.53	1.22	3.65	1.72	**0.00**	2.37
	2	4.13	2.06	2.71	2.27	**0.00**	2.90
25	1	2.49	1.67	**0.00**	2.16	**0.18**	2.91
	2	2.96	3.52	30.47	6.07	2.62	**0.00**

Table 2. Parameter estimates for rainbow trout eDNA decay, based on a Weibull model. Models were fit for two density trials (5 vs. 25 fish) with two replicates (Rep.), one of two methods: nonlinear regression or a simulation model. Weibull parameter values for p, delta (hrs), and the $\log_{10}$ of eDNA concentration at time 0 (C_0, copies/L) are given. Lower and upper 95% confidence interval bounds (LCI, UCI) are given if the parameters were statistically estimated.

Method	Density	Rep.	p			delta			$\log_{10}(C_0)$		
			Value	LCI	UCI	Value	LCI	UCI	Value	LCI	UCI
Regression	5	1	0.57	0.30	0.83	22.93	4.79	41.07	3.98	3.62	4.35
	5	2	0.53	0.22	0.83	25.40	0.97	49.83	3.63	3.21	4.05
	25	1	0.66	0.35	0.96	34.34	14.07	54.61	4.41	4.10	4.71
	25	2	0.28	0.18	0.38	1.98	-0.96	4.93	4.27	3.89	4.64
Simulation	5	1	0.57			14.72	10.42	20.78	3.86		
	5	2	0.53			14.41	9.18	22.63	3.51		
	25	1	0.66			25.54	19.03	34.27	4.32		
	25	2	0.28			0.46	0.31	0.67	4.20		

Table 3. Estimates of eDNA shedding rates (S; copies g^{-1} hr^{-1}) for rainbow trout held at different densities of fish. S was calculated for two different replicates (Rep.) based either on simulations using a Weibull decay model or on the standard approach assuming a log-linear decay model. Shedding rates include the estimate (Est.) and the lower and upper 95% confidence intervals (LCI, UCI) on both the raw scale and on a $\log_{10}$ scale.

| Density	Rep.	Model	S (copies g^{-1} hr^{-1})			$\log_{10}$[S (copies g^{-1} hr^{-1})]		
			Est.	LCI	UCI	Est.	LCI	UCI
5	1	Weibull	6.43E+03	2.83E+03	1.46E+04	3.81	3.45	4.16
		Log-linear	4.18E+03	6.97E+02	7.66E+03	3.62	2.84	3.88
	2	Weibull	2.88E+03	1.06E+03	7.81E+03	3.46	3.03	3.89
		Log-linear	3.13E+03	5.09E+02	5.76E+03	3.50	2.71	3.76
25	1	Weibull	1.54E+03	7.85E+02	3.03E+03	3.19	2.89	3.48
		Log-linear	2.52E+03	4.26E+02	4.61E+03	3.40	2.63	3.66
	2	Weibull	1.00E+04	4.40E+03	2.29E+04	4.00	3.64	4.36
		Log-linear	6.07E+03	9.56E+02	1.12E+04	3.78	2.98	4.05

Figures
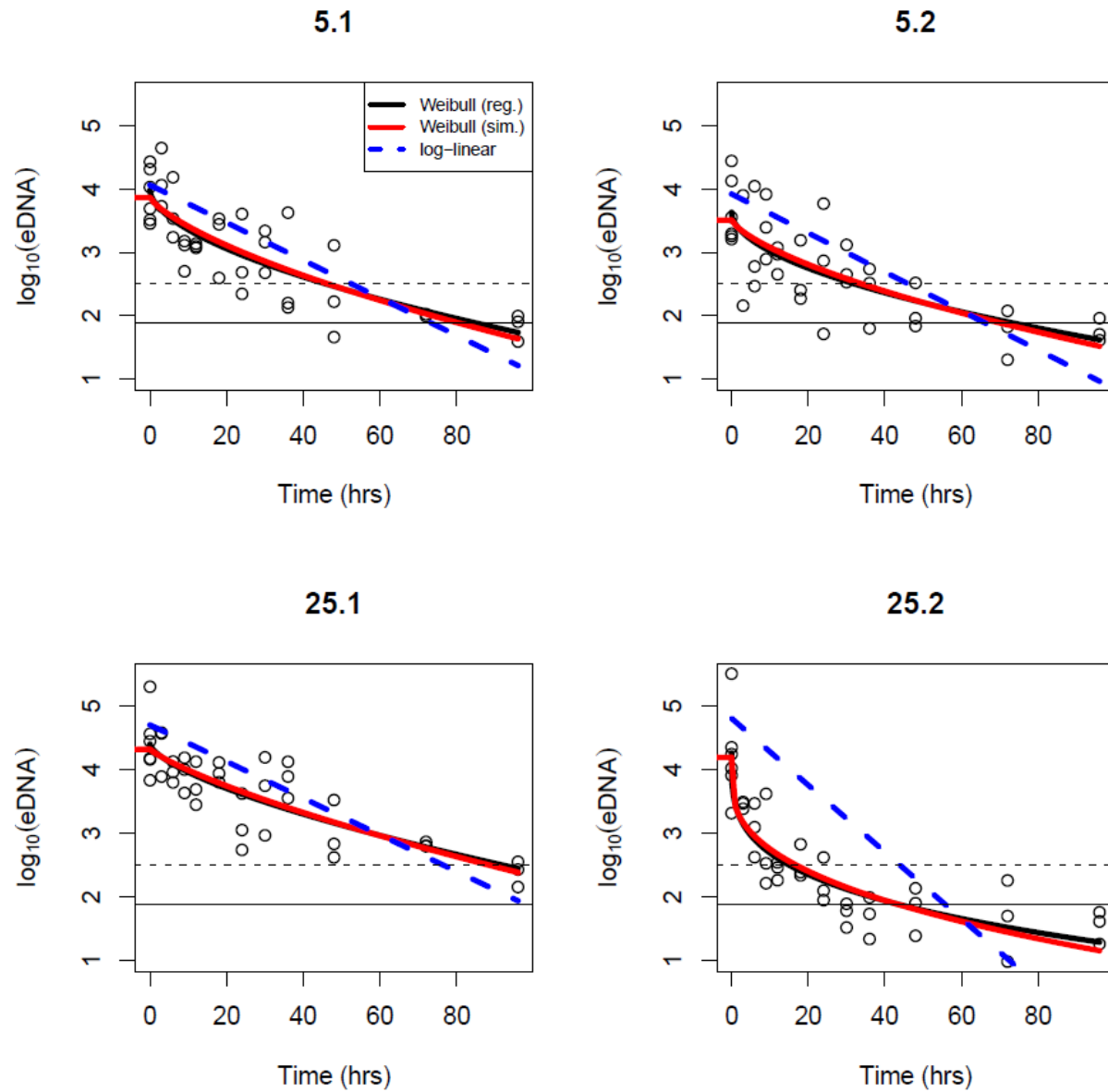
Figure 1. Logged concentrations of eDNA (copies/L) through time for rainbow trout in each of four trials. Each panel represents a trial of fish density (5 or 25) and one of two replicates (1 or 2). Solid lines represent fits of a Weibull decay model fit using either nonlinear regression (black) or the simulation model using maximum likelihood (red). The blue dashed line is the fit of a log-linear decay model with the y-intercept forced to be equal to the mean initial concentration. Points are the means of three technical replicates. Horizontal lines are the limit of detection (solid) and limit of quantification (dashed).
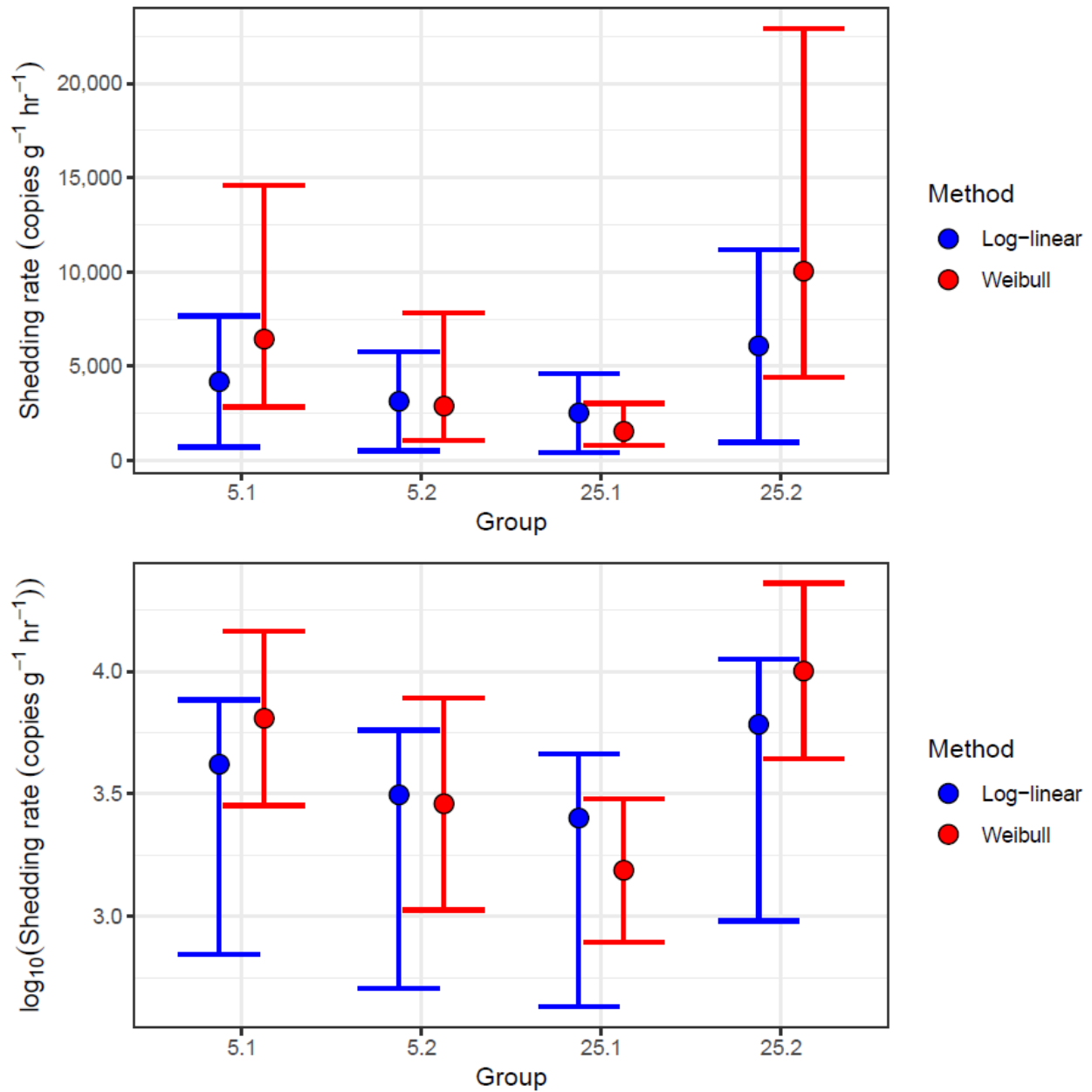
Figure 2. Estimated eDNA shedding rates (+/-95% Confidence Interval) for rainbow trout. Groups were defined by density treatment (5 vs. 25 fish) and replicate number (1 vs. 2). Shedding rates were estimated using either a log-linear (blue) or a Weibull (red) decay model.

Supplemental Figures
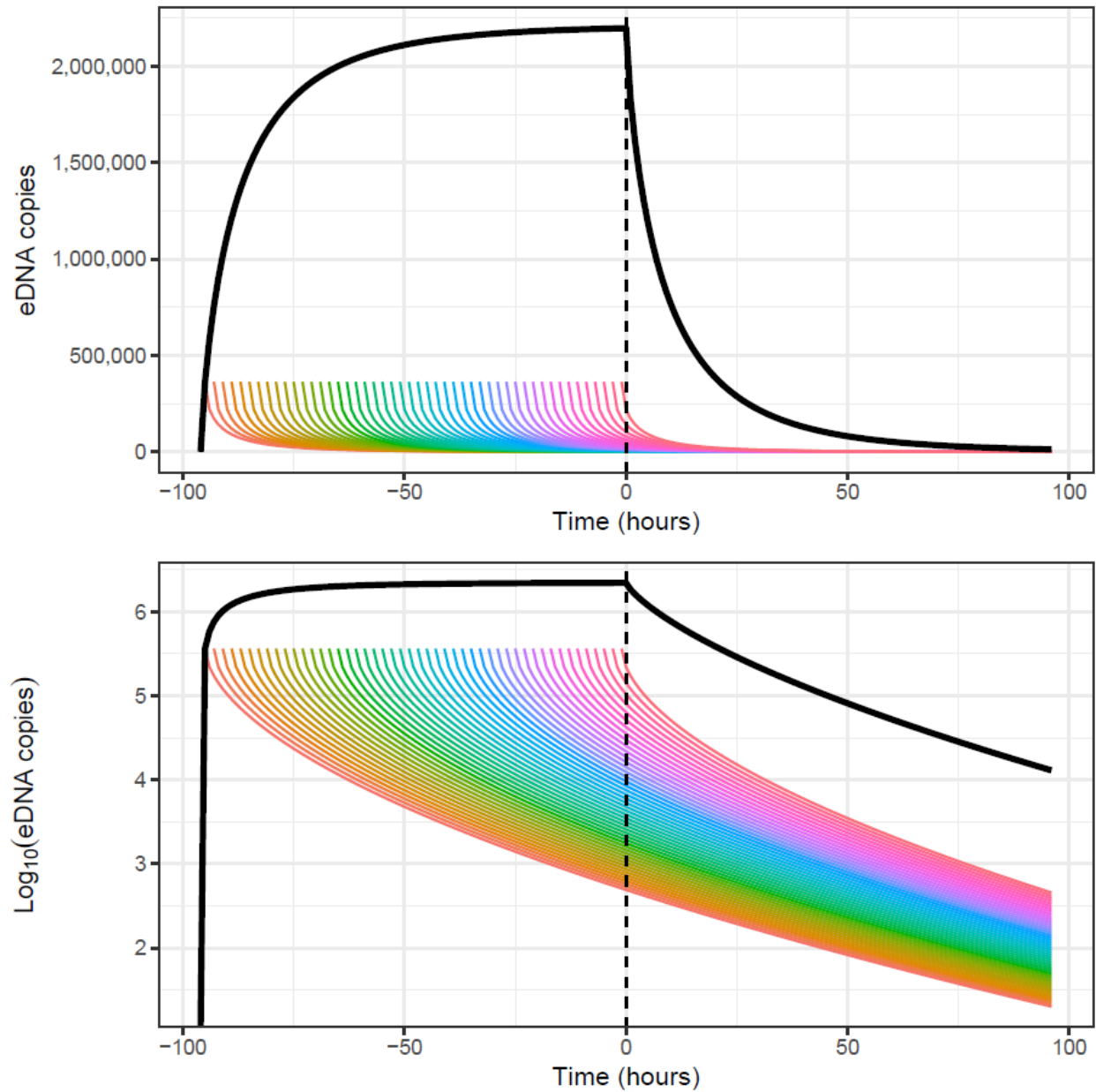
Figure S1. Total number of eDNA copies in a simulated system (black line) in which individual cohorts of eDNA are introduced hourly and decay in a non-log-linear manner (colored lines). Introduction of eDNA via shedding is stopped at time 0 (vertical dashed line). For clarity, individual eDNA cohorts are only shown every two hours.
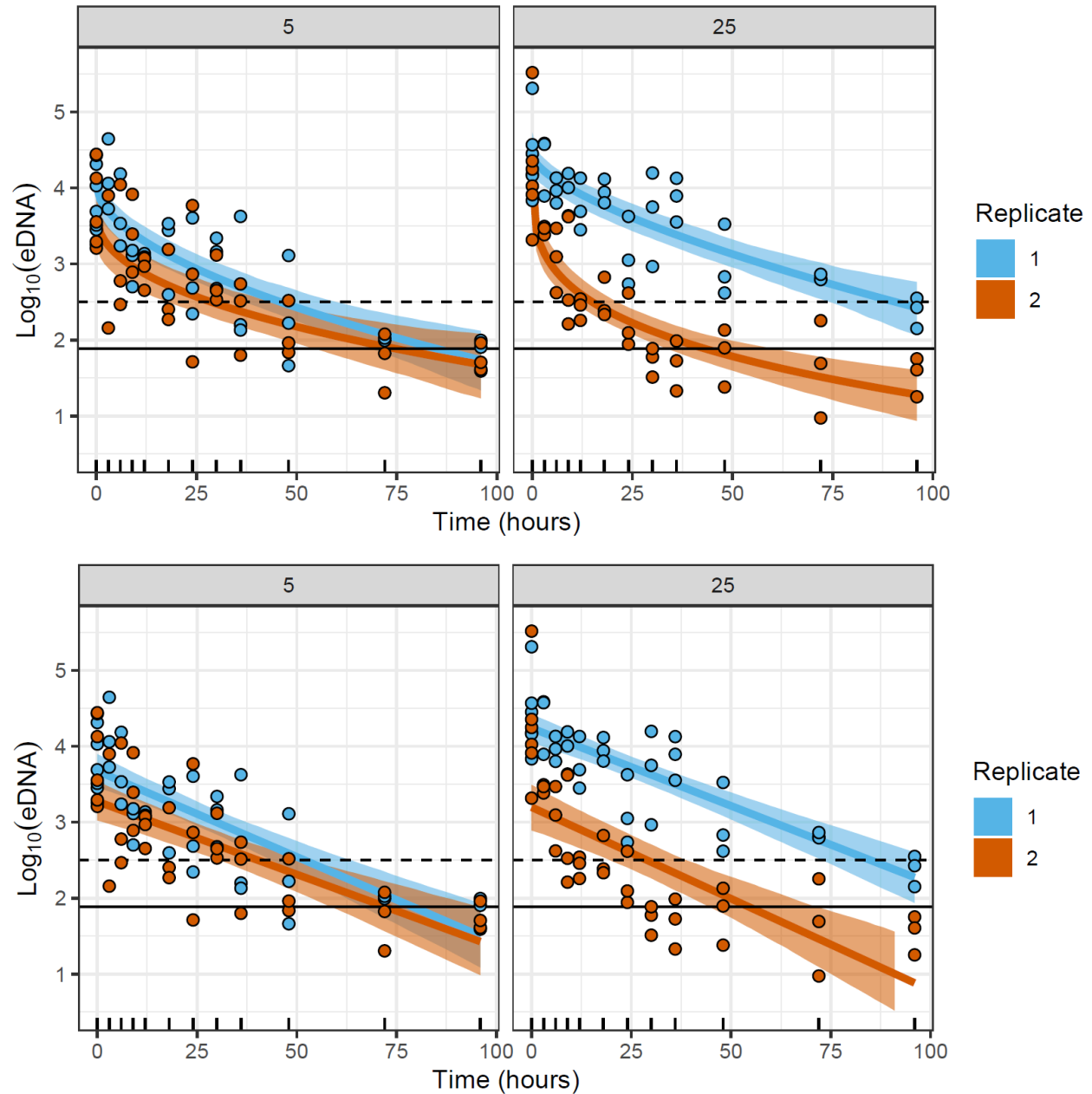
Figure S2. Logged concentrations of eDNA (copies/L) through time for rainbow trout in each of four trials. Each panel represents a trial of fish density (5 or 25) and one of two replicates (1-blue, or 2-orange). Colored lines represent fits of a Weibull decay model fit using nonlinear regression (upper panels) or a log-linear decay model with the y-intercept estimated by the model (lower panels). Shading is the 95% confidence interval approximated from 10,000 bootstrap iterations. Points are the means of three technical replicates. Horizontal lines represent the limit of detection (solid line) and the limit of quantification (dashed line).
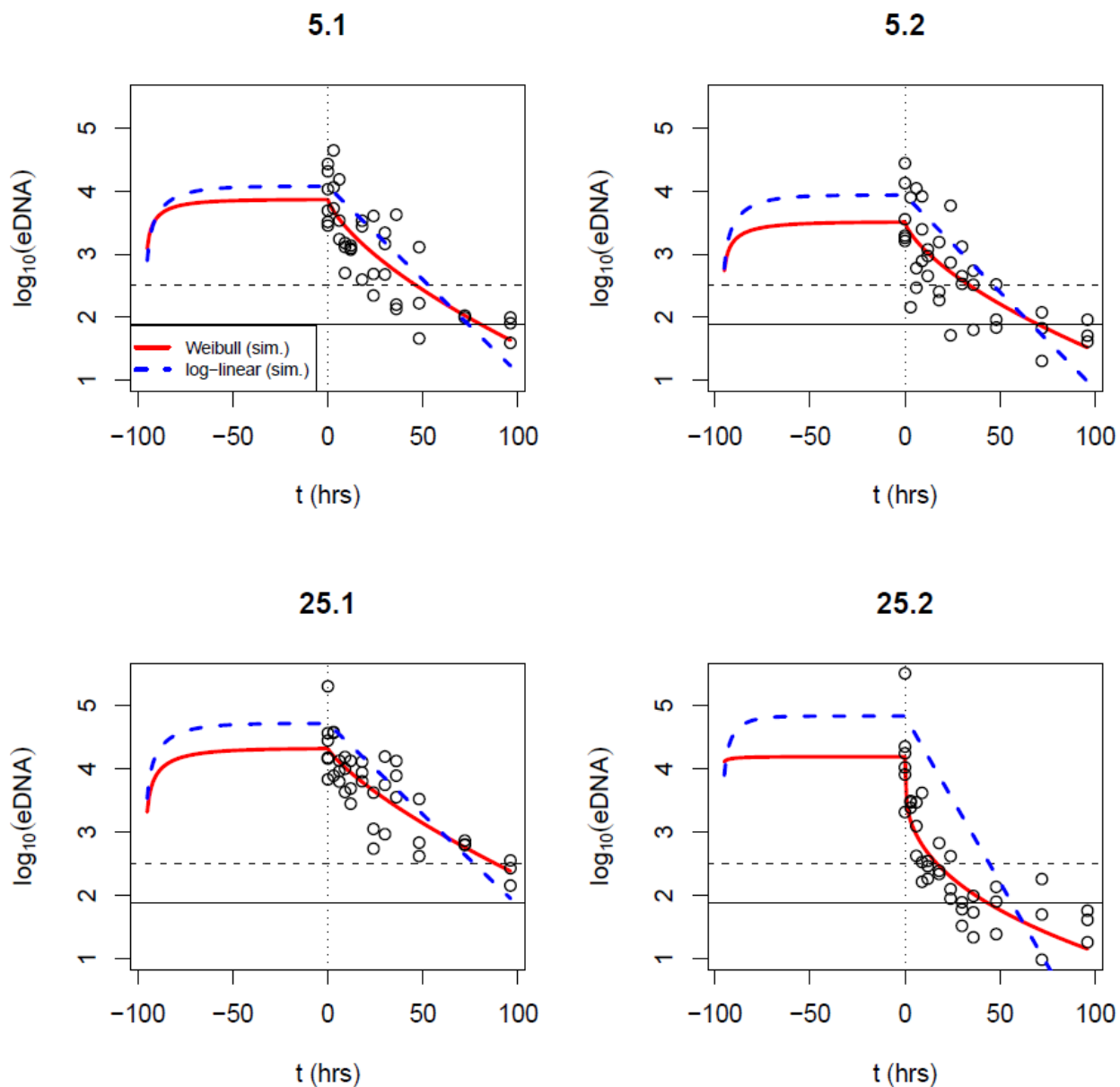
Supplemental Figure S3. Simulated (lines) and observed (points) eDNA concentrations (copies per liter) through time for rainbow trout. Each panel represents a different density treatment (5 vs. 25 fish) and replicate number (1 vs. 2). Red solid line is the best-fit Weibull (i.e., Mafart et al. 2002) model using the simulation approach. Blue dashed lines represent the simulated trajectory for the trials based on shedding and decay rates estimated using first-order log-linear dynamics.
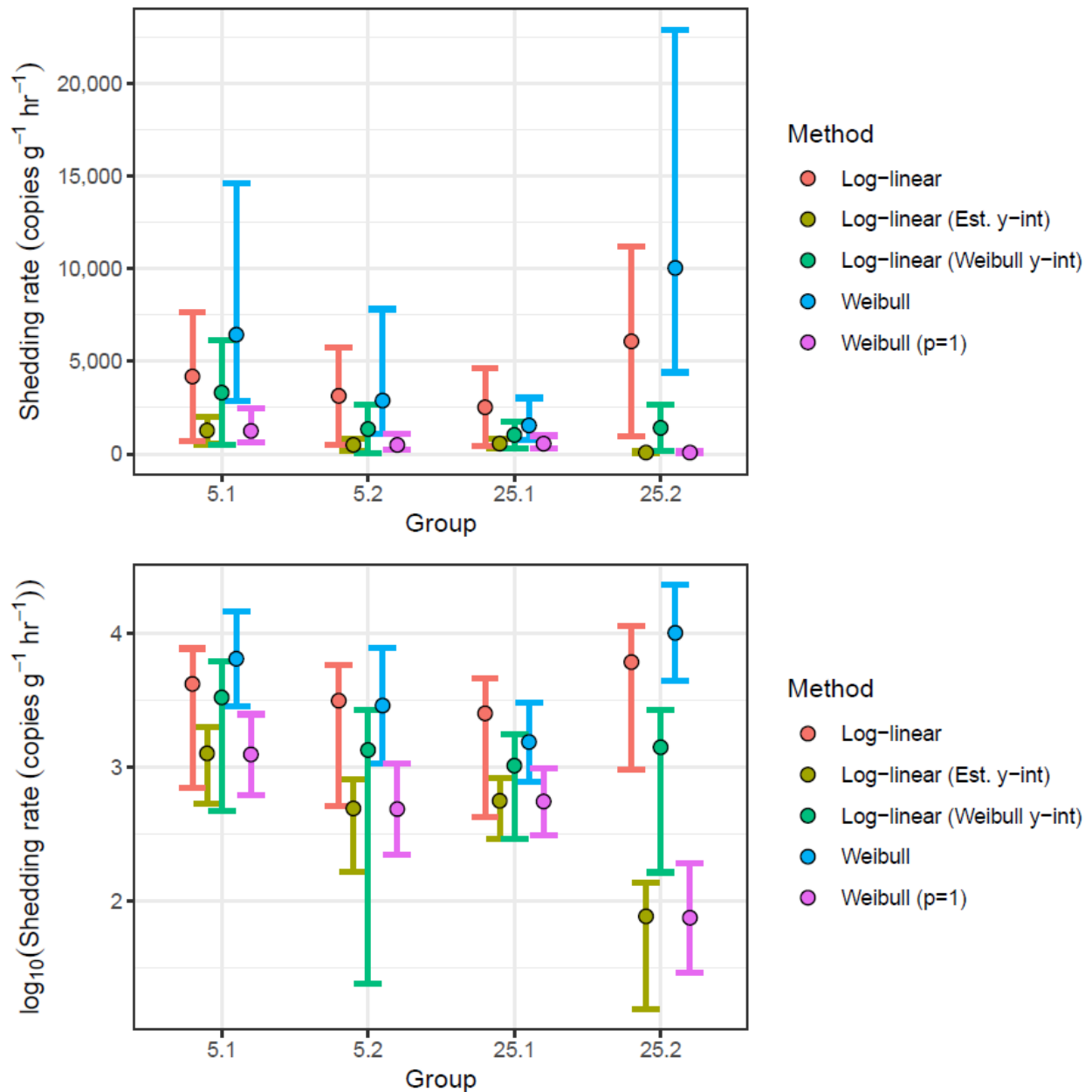
Supplemental Figure S4. Estimated eDNA shedding rates (+/-95% Confidence Interval) for rainbow trout. Groups were defined by density treatment (5 vs. 25 fish) and replicate number (1 vs. 2). Shedding rates were estimated using either a log-linear (blue) or a Weibull (red) decay model.

Supplemental Tables

Supplemental Table S1. Parameter estimates for rainbow trout eDNA decay, based on a first order, log-linear decay model. Models were fit for two densities (5 and 25 fish) with two replicates (Rep.). Estimated log-linear decay rate (k, hr^{-1}) with lower and upper 95% confidence interval bounds (LCI, UCI) are given. The $\log_{10}$ of the mean concentration at time t=0 (LOG10C0) was calculated from the empirical measurements. T.50 is the estimated half-life (hrs) and T.90 is the time for 90% decline, based on the log-linear decay constant k. Note that patterns in the residuals for these decay models indicated that inference from these models and estimates are not appropriate; see text for more description.

Density	Rep.	k			LOG10C0			T.50	T.90
		Value	LCI	UCI	Value	LCI	UCI		
5	1	0.068	0.057	0.080	4.06	3.33	4.32	10.1	33.7
5	2	0.071	0.058	0.084	3.92	3.18	4.18	9.8	32.5
25	1	0.067	0.056	0.077	4.71	3.97	4.97	10.4	34.6
25	2	0.121	0.096	0.147	4.81	4.08	5.07	5.7	19.0

TASK 2: CONDUCT FIELD STUDIES TO EVALUATE EDNA ECOLOGY AND DYNAMICS IN A NATURAL RIVERINE SETTING.

Use of foreign eDNA tracers to resolve site- and time-specific eDNA distributions in natural streams

Braden A. Herman 1
Gavin B. Bandy 1
Andre Buchheister 1
Andrew P. Kinziger 1
Eric P. Bjorkstedt 2,*

1 Department of Fisheries Biology, Cal Poly Humboldt, Arcata, CA 95521, USA
2 Southwest Fisheries Science Center, NOAA Fisheries, Trinidad, CA 95570, USA
*Corresponding Author: Andrew.Kinziger@humboldt.edu

Synopsis: This study explores using foreign eDNA (FeDNA) tracers introduced via an Autonomous DNA Introduction Device (ADID) to understand stream-specific eDNA distribution patterns in small streams. Trials in various streams illustrated within-stream FeDNA patterns, which were often stream-specific, proving its potential in refining eDNA-based estimates of species distribution using eDNA from water samples.

Abstract

Substantial uncertainty in how to interpret eDNA observations motivates a need for a technique to effectively and efficiently account for how well a given sampling scheme characterizes the distribution of eDNA in natural settings. . Here, we demonstrate an approach to calibrating eDNA surveys based on introducing a non-native ("foreign") eDNA (FeDNA) to serve as a calibration tracer for naturally occuring eDNA. To accomplish this, we designed a protocol based on an Autonomous DNA Introduction Device (ADID) to support sustained delivery of FeDNA at constant concentration, and use comparisons among distinct FeDNA tracers to assess the scales over which the combined effects of mixing and attrition establish a steady state distribution of FeDNA against which a NeDNA signal could be compared. Experiments conducted across a suite of streams in coastal California demonstrate that distributions of FeDNA, as perceived by standard sampling techniques, vary among streams, but that within a stream, distributions of distinct FeDNA are consistent with one another. This demonstrates that FeDNA tracers can serve as a basis for interpreting concurrently sampled NeDNA in a manner analogous to mark-recapture studies, and represents a significant advance towards more quantitative use of eDNA in fisheries surveys and monitoring.

INTRODUCTION

Effective management of aquatic species requires information on how the distributions of species vary over space and time, especially to support accurate assessment of the consequences of management actions. Environmental DNA (eDNA) offers an innovative and powerful tool for efficiently detecting the presence of species using a water sample that is generally more cost-effective and requires minimal disturbance of habitats or individuals (Laramie et al. 2015; Schmelzle and Kinziger 2016; Sutter and Kinziger 2019; Ostberg et al. 2019; Duda et al. 2021; Penaluna et al. 2021). Environmental DNA surveys involve collection of environmental samples to recover DNA-bearing particles (e.g., epithelial cells, feces, etc.) shed by organisms, from which DNA is extracted and concentration measured using targeted (e.g., quantitative PCR (qPCR) or droplet digital PCR (ddPCR)) or non-targeted (i.e., eDNA metabarcoding) methods (Thomsen and Willerslev 2015; Rojahn et al. 2021; Dimond et al. 2022). The utility of eDNA for estimating species abundance or for resolving distribution at finer scales is less well developed (Rourke et al. 2022). Improving the value of eDNA as a monitoring tool will require deeper understanding of processes affecting how eDNA is introduced and subsequently distributed relative to its source (Sigsgaard et al. 2020; Burian et al. 2021), and how sampling schemes resolve the resulting distributions.

Multiple factors influence introduction rates from organisms and molecular decay rates of eDNA in rivers. Water temperature, metabolic rate, habitat preference, and activity levels alter the timing, amount, and characteristics of eDNA-bearing particles shed by an organism (Jo et al. 2019; Thalinger et al. 2021; Stoeckle et al. 2021; Ostberg and Chase 2022). Subsequent decay of eDNA depends on several extrinsic abiotic (e.g., temperature, UV exposure, etc.) and biotic (e.g.,

microbial activity) factors (Sassoubre et al. 2016; Andruszkiewicz Allan et al. 2021; Caza‑Allard et al. 2022), as well as intrinsic characteristcs (e.g., size, cellular integrity, etc,) of eDNA-bearing particles (Bylemans et al. 2018; Moushomi et al. 2019; Nagler et al. 2022). Likewise, variation in particle characteristics also affects how particles move in a natural setting (e.g., larger particles tend to settle from the water column more rapidly), and thus how eDNA is distributed and ultimately how it is available for capture in a water sample.

In moving waters, the distribution of eDNA is determined by the location and intensity of eDNA sources relative to flow and stream structure (e.g., eddies, shear zones, large pools, etc.) that affect the movement of particles through the system. . Environmental DNA tends to be transported further at higher flow rates and is more readily resuspended in more turbulent streams, and it tends to settle out slower, deeper areas of stream (Jerde et al. 2016; Bedwell and Goldberg 2020; Thalinger et al. 2020; Laporte et al. 2020; Spence et al. 2021; Curtis et al. 2021; Jo and Yamanaka 2022). Absorption to biofilm (e.g., algae) also varies by substrate type Shogren et al. (2017, 2018). Advancing the potential for eDNA to provide more quantitative information will require accounting for system-specific distributions of eDNA (Bedwell and Goldberg 2020; Carraro et al. 2021). Movement, attrition, and detection probability of eDNA in stream systems has been studied using fish kept in cages (i.e., live cars) to introduce eDNA at a point source (Shogren et al. 2017, 2018; Nukazawa et al. 2018; Robinson et al. 2019; Thalinger et al. 2020; Laporte et al. 2020; Spence et al. 2021). Methods that rely on live fish are not able to closely control the intensity of the introduced eDNA signal or how this signal varies over time. Though effective and representative of a true distribution of particle sizes and properties, these methods incur ecological risks arising from the potential for non-native fish to escape into the natural system, introduction of pathogens, or the handling of listed species. To resolve these uncertainties and alleviate ecological concerns a method to efficiently characterize eDNA transport patterns in streams is needed to increase accuracy of and information gained from eDNA surveys (Burian et al. 2021; Buxton et al. 2021).

Although it is possible to characterize and model transport of various conservative or reactive tracers in relation to stream structure (e.g., Workshop 1990; Miller and Georgian 1992; Payn et al. 2008; Schmadel et al. 2016, Jerde et al. 2016; Carraro et al. 2018, 2021; Harrison et al. 2019, Carrera et al. 2022), such detailed understanding of the intricate dynamics affecting eDNA particles and their distribution may not be necessary or cost-effective for interpreting natural eDNA distributions to address many fisheries management questions. Rather, simply accounting for the pattern that emerges during downstream transport may suffice to put NeDNA signals into context. In this study, we propose and evaluate a straightforward, practical method for doing so through the controlled introduction of a non-native "Foreign" eDNA (FeDNA) tracer that matches the size distribution of eDNA particles, and then documenting the distribution of the FeDNA tracer downstream of the introduction site. The change in the FeDNA signal through space and time would reflect the complex processes governing the distribution of eDNA in flowing water systems, without the need to fully model or understand the specific hydrodynamic and attrition dynamics of a given system, yet provide a basis for interpreting

other, concurrently measured, native eDNA signals, under the assumption that native and introduced eDNA particles are similarly affected by their shared environment. Concurrent use of a second FeDNA tracer would, under the same principles, allow the FeDNA tracers to be calibrated for consistency in how downstream distribution patterns are resolved, thus providing greater confidence in the context for interpreting NeDNA patterns.

<div align="center">Conceptual Model and Assumptionsl</div>

A conceptual model for distribution patterns expected for concurrently introduced FeDNA tracers provides a framework for interpreting data emerging from such an experiment (Figure 1). eDNA concentrations are expected to decline with distance from the eDNA source, and to be mixed and diluted as eDNA is carried downstream (Robinson et al. 2019; Thalinger et al. 2020; Laporte et al. 2020; Spence et al. 2021). Based on previous tracer studies, there are characteristic scales over which mixing is able to distribute tracers across the width of a stream; this pattern has been confirmed to hold for eDNA (Wood et al. 2021).

Once fully mixed across a stream, it is expected that distributions of FeDNA (and any NeDNA) will be shaped by the same flow, mixing, and attrition processes, regardless of differences in the point of introduction, and the ratio between two FeDNA concentrations should remain constant. The expectation is that all FeDNA concentrations should decline with distance downstream and be lower in concentration at subsequent, downstream cross-sections. Importantly, however, any departures from this expectation arising from how sampling locations align with eDNA distributions will also be shared among well mixed FeDNA tracers, preserving the expectation of a constant ration among FeDNA concentrations, regardless of observed patterns measured eDNA concentrations. Note that the constant ratio formed between FeDNA tracers will depend on the introduced concentrations and any losses that occurred prior to fully mixing and so is not required to be equal to 1. This assumption of shared dynamics yielding constant ratios among well mixed eDNA tracers can be used to identify a "well-mixed region" where similar patterns in FeDNA distributions have developed and hold until concentrations diminish below quantifiable or detectable levels.
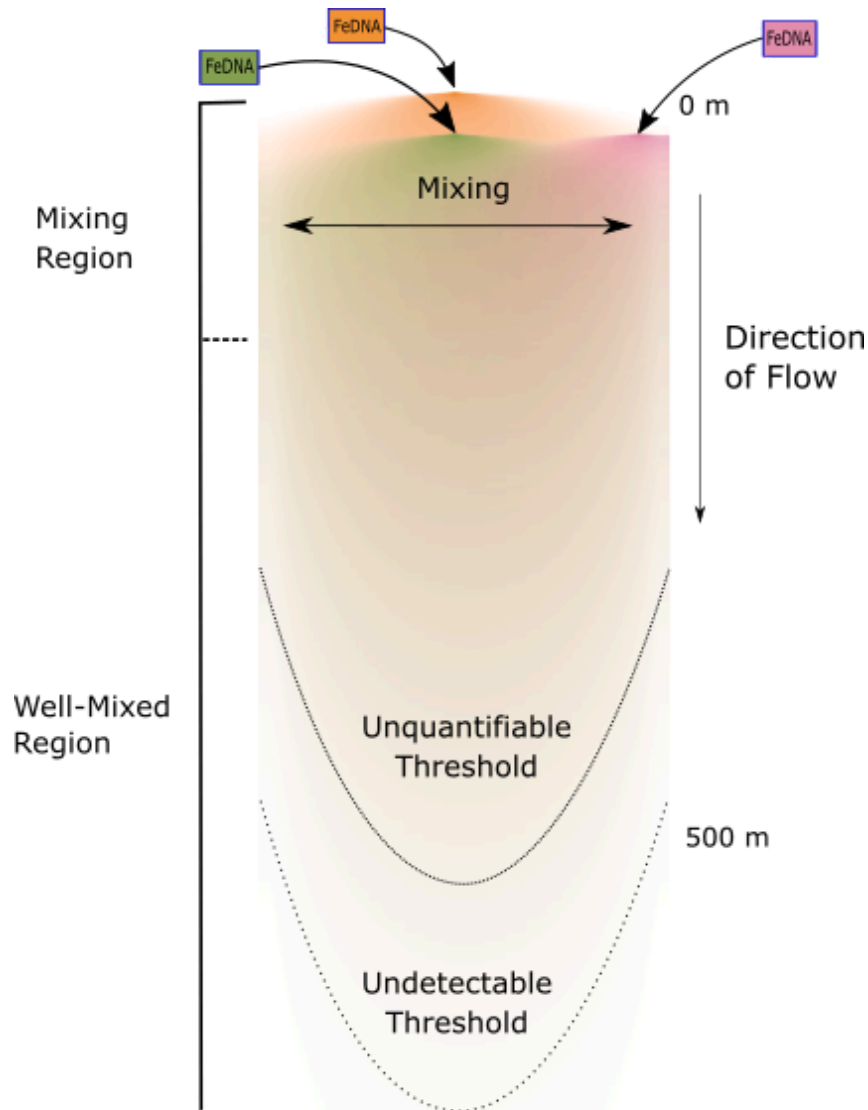
Figure 1. Conceptual model of expected distribution patterns of FeDNA in a stream. The color gradient represents the expected eDNA plume generated from continuous introduction of three different FeDNA tracers at three point sources. Note the shift from multiple plumes to a well-mixed plume below a zone of sufficient extent to allow complete mixing.

Note that such interpretations, including any extension to calibrating signals of NeDNA, rest on assumptions analogous to those for Mark-Recapture analysis (Pine et al. 2003). Specifically, these assumptions are that FeDNA particles from distinct sources are assumed to be of similar size distribution and characteristics, and that FeDNA particles of the same state and size are: 1) subject to and equally affected by the same transport processes, 2) subject to the same

capture probability and processing efficiency regardless the presence of other particles, and 3) lost at the same rates (e.g., through molecular decay or settling out of the water column).

Objectives

This study is a part of a broader research program designed to assess the potential for using FeDNA tracers to calibrate NeDNA surveys. The objectives of this part of this study were to (1) develop a tool to introduce non-native FeDNA at a controlled rate and concentration, (2) to test standard protocols for using this tool to resolve stream-specific patterns of FeDNA distribution, and (3) to evaluate whether patterns observed for distinct FeDNA tracers conform to the assumptions of correlated transport and emergent constant ratios as a basis for identifying scales of mixing and as a proof-of-concept for application to concurrent NeDNA observations.

MATERIALS AND METHODS

Foreign eDNA (FeDNA) Suspension and Autonomous DNA Introduction Device (ADID)

The Autonomous DNA Introduction Device (ADID) is a relatively inexpensive, modular system designed to introduce a suspension of eDNA bearing particles into a stream at a sustained, controlled rate (in this case 25 ml h^{-1} for up to 24 h) . An ADID consists of three main components, an Arduino microchip controller (https://www.arduino.cc/), a peristaltic pump, and a portable battery, all mounted in a waterproof housing (Figure 2; see Appendix B for construction details and instructions). For this study, eDNA suspensions were prepared to (1) closely mimic NeDNA and (2) be distinct from any NeDNA present in the study systems. Three species were selected, *Ictalurus punctatus* (IPU; Channel Catfish), *Micropterus salmoides* (MSA; Largemouth Bass), and *Cyprinus carpio* (CCA; Common Carp), none of which are present in the study systems. All fish were obtained from The Fishery in Galt California, a fish hatchery approved for this use by the California Department of Fish and Wildlife, mechanically euthanized, fileted, skinned, and frozen in vacuum-sealed bags[1]. To prepare FeDNA suspensions, muscle tissue was combined with off-gassed well water at a fixed ratio of 100 g: 1 L and homogenized in a standard household blender. The resulting suspension was passed once through a metal mesh screen, passed three times through double layered, 4 ply, 50 grade cheese cloth to remove coarse particles, then filtered through a 200 μm sieve to truncate particle distribution to the maximum size of NeDNA particles reported in the literature (e.g., Turner et al. 2014b; Wilcox et al. 2015). Instructions for producing the FeDNA suspension are in Appendix A. FeDNA suspension was stored frozen prior to experimental trials. In the field, the (thawed) FeDNA suspension was diluted with stream water and held in a 20 L bucket with a fitted lid that was in turn housed in a cooler of ice to reduce degradation (Figure 3). Buckets were also equipped with aerators to prevent the development of anoxic conditions and to maintain particles in suspension (Figure 3).

---

[1] Tissue collection for this study was approved by Cal Poly Humboldt's Institutional Animal Care and Use Committee (IACUC, No. 2020F74-C).
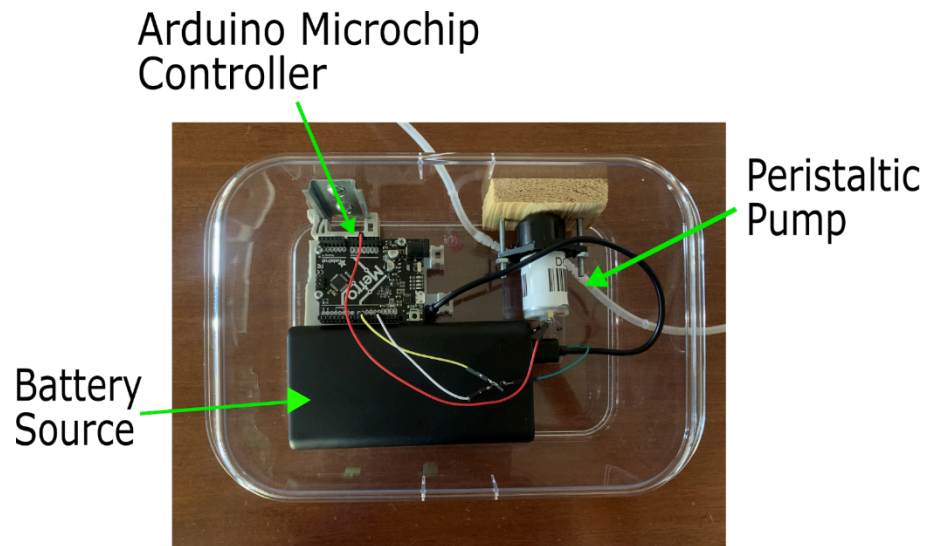
Figure 2. The Autonomous DNA Introduction Device (ADID) including the three main components: Arduino microchip controller, peristaltic pump, and battery source. The entire device was contained in a waterproof housing.

Figure 3. Representative deployment of two ADIDs in the field. Note the bucket containing FeDNA suspension was placed in a cooler on ice. The bucket was closed, and an air pump was attached to induce mixing.

Trials were conducted in several small streams in coastal Northern California: Jacoby Creek, Old Campbell Creek, Little River, and Prairie Creek (Figure 4). These streams were selected based on proximity to Cal Poly Humboldt, the presence and varying abundance of coho salmon *Oncorhynchus kisutch* (to support distinct collaborative research), access, and permission to sample (e.g., State and National Park permits, permissions from private landowners, etc.).
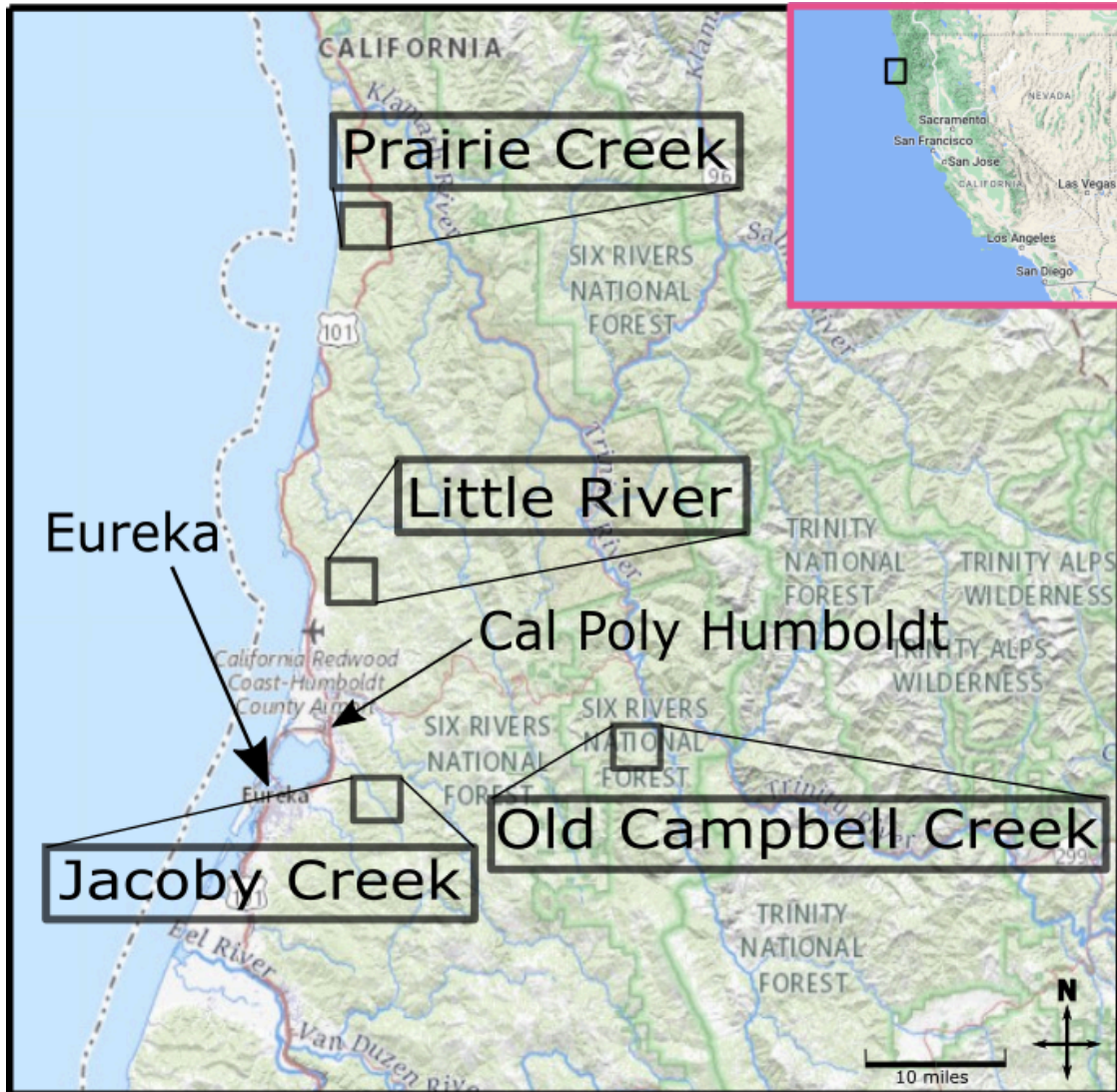
Figure 4. Trial site locations for each stream and location of California Polytechnic University, Humboldt in relation to these sites. The inset box relates the locations of the trial sites to the greater California region.

At each site, one kilometer study reaches were selected based on access to the site and habitat features. Logistical considerations focused on sufficient access to deploy ADIDs in

locations suitable for proper function, to safely transport coolers with ice for the FeDNA suspension, to reach downstream sampling locations, and to be able to accomplish a full trial in one day. Sites were selected to exclude tributaries, slow-moving pools, or stretches of discontinuous surface flow. Any such features that were present are noted in the site descriptions and qualitatively accounted for when interpreting results.

For each site, sampling cross-sections were designated at 50, 100, 150, 200, 300, 500, and 1000 meters (m) downstream from the top of the reach. At each cross-section, stream width was measured. Depth and velocity were recorded at each sampling location. Site-specific discharge was calculated using standard techniques outlined in methods of stream ecology (Hauer and Lamberti 2006). The location of the main flow, where the water was flowing the fastest, was recorded at each introduction site and cross-section. A summary site description was based on averaging conditions at each sampling location and is included in Table 2 of the results.

FeDNA Introduction and Sampling Grid

Based on insights from preliminary trials (Appendix C), ADID were configured to deliver a robust FeDNA signal and sampling schemes were designed to resolve cross-stream distributions of FeDNA originating from several point sources. Each trial consisted of the deployment of three ADIDs, each of which introduced one of the three distinct FeDNA tracers. ADIDS were deployed at locations to support assessment of how distributions of distinct FeDNA as they mixed within the stream. IPU FeDNA was introduced at the head of the reach (0 m) into the main flow of the stream as determined by qualitative assessment of cross-stream structure of stream flow. MSA FeDNA was also introduced into the main flow, but at a point approximately 50 m downstream from the uppermost ADID. In both cases, FeDNA was diluted 1:1 with stream water to produce an initial concentration of approximately 1.5 million copies/ml. CCA FeDNA was diluted 1:2 with stream water (initial concentration approximately 0.75 million copies/ml) and introduced at a point ~50 m downstream from the upper most ADID outside the main flow of the stream (i.e, identified by decreased flow or a distinct eddy separate from the main flow). The side of the stream (i.e. river left or river right) for the CCA introduction was chosen to have the greater difference in flow rate from the main course of the stream.

The ADIDs introduced FeDNA suspension into the stream at a rate of 25 ml/minute for 6 hours prior to collecting water samples (total volume: ~9 L). Based on preliminary trials (Appendix C), 6 hours was sufficient time for FeDNA to have established a steady state distribution extending 1000 m downstream.

Water samples were collected from the most downstream location (1000 m) and proceeding upstream to eliminate the risk of contaminating samples through disturbance of sediments or resuspension of eDNA (Figure 5). At each cross-section, samples were collected at three equidistant points across the stream, labeled left, right, and center. Sampling locations were fixed as a function of stream width and were not adjusted to accommodate any variation in cross-sectional structure or flow patterns. Triplicate water samples were taken at each location,

yielding 9 water samples per cross-section and 63 samples per trial. A field blank was collected for each cross-section (N=7 per trial).
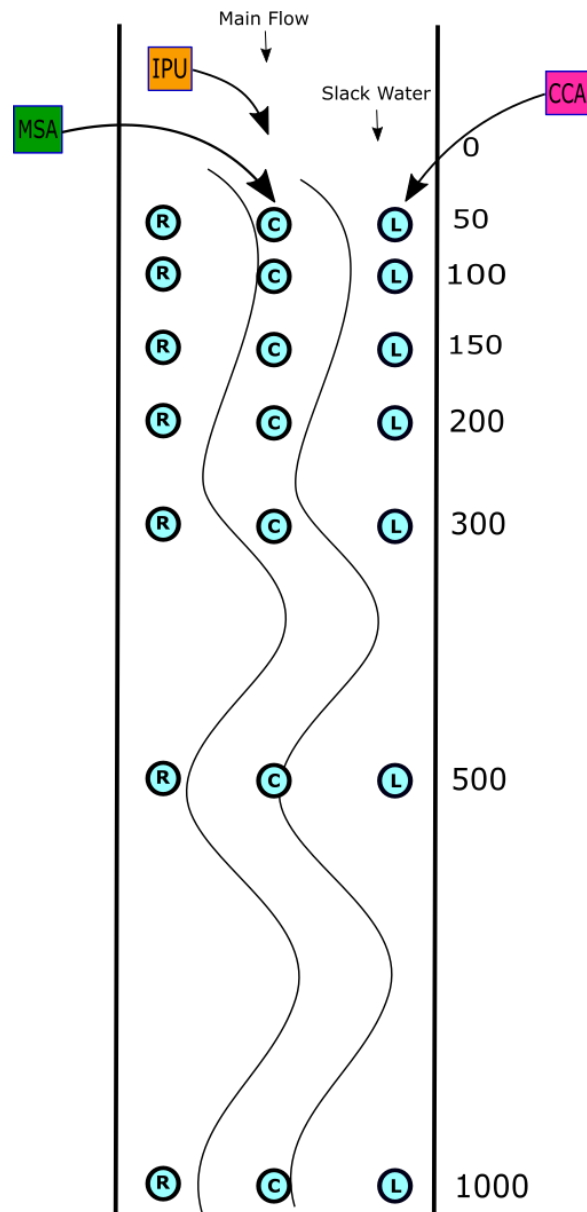

Figure 5. Final sampling design and ADID deployment for field trials. IPU, MSA, and CCA correspond to the FeDNA species used. The color gradient represents the expected eDNA plume generated from continuous introduction of FeDNA at a point source. The numbers on the right represent the cross-section distances (meters) where sampling occurred. The blue circles represent the left, center, and right sampling locations at each cross-section. The main flow and slack water are noted to demonstrate the goal of capturing cross shelf variation using a fixed sampling design.

Characterization of Study Site Characteristics

Environmental DNA Methods

Water collection

Each sample consisted of 1.75 liters of water collected by pulling a sterile single-use Whirl-Pak bag (Nasco) along the water surface, which was then sealed and immediately stored on ice. The water samples were filtered in the lab within 10 hours of collection. Water was filtered under vaccuum onto 47 mm diameter 0.45µm cellulose-nitrate filters (Cytvia; catalog number: 10401170) on a filter support pad (MilliporeSigma™ catalog number: AP1003700) in a sterilized plastic filter funnel. Up to 6 samples were filtered simultaneously using a manifold connected to a pneumatic hand pump (EWK EB0103A). Filters were transferred with sterilized forceps into 2 ml microcentrifuge tubes (Eppendorf catalog number: 022431048) containing 360 µl of cell lysis buffer (QIAGEN buffer ATL, catalog number: 939011) and stored at -20°C until extraction. All filter cups and other water filtering materials were sterilized prior to reuse in a 10% bleach solution. Volume of water filtered was recorded for each sample to support calculations of eDNA copies per liter.

To detect contamination associated with field practices, one field blank was collected for each cross-section. Each field blank was prepared by transferring commercially bottled drinking water into a Whirl-Pak bag in the field, and processed in parallel with concurrently collected environmental water samples to serve as a negative control for contamination (Goldberg et al. 2015).

DNA Extractions

Extraction of samples followed acetone-dissolution methods outlined by Hallet and Bartholomew (2006), with modifications to account for samples not preserved in ethanol as prescribed in the standard protocol. The filters were removed from the Buffer ATL (reserved) and placed into a new microcentrifuge tube and allowed to dry at room temperature for 24 hours. Next, each filter was dissolved by adding 3 mm, sterile, solid glass beads (Fisher Scientific catalog # S80024) and 1.5 ml acetone (Spectrum Chemical Corporation, HPLC Acetone, #CAS 67-64-1) and vigorously agitating the sample on a vortexer. Each sample was centrifuged at 8000 RPM for 1 minute to condense the solids (genetic material and glass beads) into a pellet at the bottom of the tube. The supernatant acetone was removed, clean acetone added, and the agitation and condensing steps were repeated. The pellet of solids formed from this step was rinsed by adding 1.5 mL of 95% ethanol, resuspended by agitation on a vortexer and recondensed by centrifugation at 8000 RPM for 1 minute, after which the ethanol was removed, and the pellet allowed to dry at room temperature. The pellet of solids was recombined with its original 360 µl Buffer ATL, and 40 µl of proteinase k was added. This combination of pellet, buffer, and proteinase k was incubated overnight at 56 C°. Environmental DNA was extracted using

QIAGEN DNeasy Blood and Tissue Kit (69506) following the manufacturer instructions, modified to include 40 µl, 360 µl of Buffer ATL, and a single 100 µl final elution of extracted DNA (Sutter and Kinziger 2019).

<u>DNA Quantification</u>

FeDNA concentrations were estimated by using droplet digital PCR (ddPCR) on the BIORAD QX200 Droplet Digital PCR System (catalog number: 1864001). Assays that target short fragments of the mitochondrial Cytochrome b (less than 150 base pairs)were adapted from the literature for each of our FeDNA tracers:: *M. salmoides* (Yamanaka et al. 2016), *C. carpio* (Eichmiller et al. 2014). and *I. punctatus* (US Forest Service, Rocky Mountain Research Station, unpublished). The assays were procured from Integrated DNA Technologies PrimeTime Probes, which incorporated reporter dyes of FAM or HEX and quenchers of ZEN/IowaBlack FQ. Full base pair sequences and components of the assays are in Appendix D.

Each ddPCR reaction consisted of 900 nM forward primer, 900 nM reverse primer, 250 nM probe, 0.2 µl of 300 mM dithiothreitol (Bio-Rad catalog number: 12012171), 5 µl of ddPCR Multiplex Supermix (Bio-Rad catalog number: 12005911), 12 µl of DNA template, and sufficient volume of water to bring the final reaction volume to 22 µl. An aliquot of 20 µl of the ddPCR reaction mix and 70 µl of Bio-Rad droplet generator oil (Bio-Rad catalog number: 1864006) were placed on a BioRad DG8 droplet generation cartridge (Bio-Rad catalog number: 1864008) secured in a DG8 Cartridge Holder (Bio-Rad catalog number: 1863051), covered with a DG8 Gasket (Bio-Rad catalog number: 1863009), and transferred to the Bio-Rad QX-200 droplet generator (Bio-Rad catalog number: 1864002) that uses microfluidics and surfactants to partition the reaction mix up to 20,000 nanoliter-droplets.

Reactions were thermocycled on an MJ Research PTC-100 Thermal Cycler through a 10-minute enzyme activation phase at 95 ℃ followed by 40 cycles consisting of of a 30-second denaturation stage at 94℃ and a 1-minute annealing/extension phase at 60℃. Cycling was followed by a 10-minute enzyme deactivation phase at 98℃ and a 15-minute droplet stabilization phase at 4℃. The temperature ramp rate was set at 2 °C per second for all steps. Following cycling, a Bio-Rad QX200 droplet reader was used to count PCR-positive and PCR-negative droplets.

Results are reported as the Poisson-corrected number of copies per 20 µl reaction. Each water sample was assayed once except when a sample returned low droplet counts (relative to other reactions run concurrently) or was otherwise strongly anomalous. In these rare instances, the assay was repeated to improve or confirm the results. Each ddPCR plate contained a positive control for each species made from extracted genomic material and a negative control well containing all reagents except DNA template.

The concentration of the DNA in the original water samples (copies per liter),X, was calculated as follows (Knudsen et al. 2019):

$$X = \frac{Y*b}{a*c}$$
(Eq. 1)

Y is the estimate of copies per ddPCR reaction, *b* is the amount of suspended, purified eDNA elute from the extraction (μl), *a* is the volume of the suspended eDNA elute (template) used in the qPCR reaction (μl), and *c* is the total volume filtered (liters).

<u>Limit of Detection and Limit of Quantification</u>

Limits of detection (LOD) and quantification (LOQ) were determined for each assay using a serial dilution of DNA from tissues extracted using a QIAGEN DNeasy Blood & Tissue Kit (69506), following manufacturer guidelines. DNA concentrations were determined using ddPCR, as described above, across serial dilutions designed to range from 1 to 700 copies of the target eDNA per reaction, and 8-16 replicates per concentration. The LOD and LOQ were determined using curve-fitting methods described in Klymus et al. (2020). The LOD was defined as the lowest concentration of target DNA that would result in at least 95% positive detections and the LOQ was defined as the lowest quantifiable concentration of target DNA with a coefficient of variation (CV) below 35% (Klymus et al. 2020). Uncertainty in the estimate of LOD was estimated as 95% confidence interval around LOD. As a measure of uncertainty around LOQ, LOQ for a 30% and 40% CV were calculated.

<div align="center">Data Analysis</div>

Each trial resolves patterns arising from a unique set of conditions, providing limited degrees of freedom for statistical inference about the influence of the drivers of concentration in small streams. Therefore, the analysis that follows is intentionally agnostic with respect to any role or influence that stream characteristics, such as sampling depth, width, or velocity, might have in shaping longitudinal or cross-sectional patterns in observed FeDNA distributions. Instead, this analysis focuses on identifying shared patterns in FeDNA distributions as a foundation for extending the use of FeDNA to calibrate or contextualize other observations of eDNA.

<u>Data Quality Control and Screening</u>

Initial review of the data focused on comparing co-located and nearly concurrent samples to identify cases that represent rare or unique departures from expectations of shared transport, detection, or processing. Samples were excluded from analysis a priori under two circumstances. First samples for which no FeDNA targets were amplified, despite the presence of FeDNA in co-located, nearly synoptic, samples were excluded under the assumption that such cases appear to represent instances in which one or more steps in the filtering, extraction or amplification processes were broken in a way that disrupted all FeDNA assays. Second, any sample for which an assay returned an estimate of FeDNA concentration was at least four orders of magnitude greater than average concentrations observed in other co-located samples was excluded as rare instances of samples including large particles well outside the intended size distribution of eDNA particles (e.g., large particles). Following the removal of starkly idiosyncratic samples, data was

screened against experimentally determined thresholds to produce a data set that excluded samples below the LOD, and one that excluded samples that fell below the LOQ. These data sets represent cases where only sample-specific criteria are used to screen the data.

The data were then further screened based on pairwise correlations between sample specific FeDNA concentrations, using only those samples with concentrations above LOQ (i.e., accurately quantified concentrations). This screening proceeded iteratively. In each step, a linear regression was fit to a set of paired within-sample FeDNA concentrations, and, if necessary the sample that fell furthest from the estimated regression was removed and the model was re-estimated. This algorithm was iterated until the root-mean-square error (RMSE) of the final data set no longer exceeded an (arbitrary) threshold of 0.15. The resulting set of samples represented an internally consistent data set, in which correlation among FeDNA concentrations is an indicator that FeDNA distributions are consistent with the assumption that shared processes shape these distributions. Note that this screen relies on comparisons among FeDNAs across samples, in contrast to criteria based on FeDNA concentrations applied to individual samples as in the LOD and LOQ screens described above. This algorithm was applied to each pairwise comparison of FeDNA tracers within each stream (i.e., IPU with MSA, IPU with CCA, and MSA with CCA). Rather than requiring that a given sample exhibit correlation between all three FeDNA tracers to be retained, samples that were included in at least one comparison with another FeDNA in the correlation-based screen were retained. This more lenient criterion was used to prevent poorly resolved FeDNA distributions (e.g., a FeDNA signal with nearly all concentrations below the LOQ) from eliminating data from other, better resolved, FeDNA distributions.

This algorithm provides a basis for testing whether samples of FeDNA tracers were consistent with one another as would be expected if assumptions are satisfied, noting that these assumptions have not been directly tested in this study. If patterns between FeDNAs are coherent, linear regressions of the FeDNA concentrations were expected to go through the origin (i.e., have zero intercepts on both x- and y- axes of the regression) and to have a slope consistent with a constant ratio between concentrations of the two FeDNAs. Non-linear relationships (or starkly non-zero intercepts arising from linear fits to non-linear data) would be indicative of divergent FeDNA behaviors, such as unequal decay rates or unequal transport, but such patterns were not observed in this study.

When applicable to the analyses that follow, I analyzed data sets emerging from each of the four data screens to assess the sensitivity of results to the choice of how to retain data for analysis. For simplicity, only the comparison of correlation and no screening are shown in the results. Application of only the LOD- and LOQ-based screens yielded intermediate effects on the results and are included in Appendix E.

Data Summary and Effects of Aggregation

Visual comparison of patterns between FeDNA tracers was conducted by plotting the entire data set for each trial by species, sampling location, and cross-section. Averages at each

sampling location and cross-section were used to visualize distributions. Site-specific stream structures thought to be driving the apparent patterns were noted.

<u>Patterns of Detection Probability</u>

Patterns of FeDNA concentration and detection probability were displayed to facilitate direct comparison of these metrics. Averages at each cross-section were used to visualize concentration patterns. Detection probability was defined as the number of samples at a cross-section above the LOD divided by the number of samples collected at a cross-section ($n =$ 9). Confidence intervals for detection estimates are generated using the function 'EnvStats' from package 'ebinom' in R.

<u>Comparison of Ratios</u>

Recall that if patterns between FeDNAs are consistent, they should exist in a constant ratio over space, and thus indicate where distinct FeDNA tracers are thoroughly mixed with one another, i.e., a well-mixed region (WMR). To identify WMR, ratios of concentrations were calculated for all pairwise combinations of FeDNA tracers in each stream and then, log-transformed to homogenize variances across samples. Log-transformed ratios were analyzed using a one-way ANOVA to test for differences by distance and post-hoc tests were used to identify distances that did not have significantly different FeDNA ratios. Non-significant differences in ratios by distance were used to identify a well-mixed region (WMR) of the stream. Estimation of the WMR was based on data that had passed the correlation-based screen and therefore appeared to satisfy assumptions of shared movement, detection, and loss. Similar ratios were calculated for data retained under other screening methods and presented for comparison but were not analyzed statistically due to severe violations of ANOVA model assumptions. Estimates of WMR are calculated for each FeDNA:FeDNA pair at a site for comparison.

Two criteria were established as the basis for determining whether a cross-section was to be included in the WMR. First, for a cross-section to mark the beginning of the WMR, at least one valid ratio was required for each sampling location (left, center, right) to ensure that the FeDNA was well mixed across the stream. . Second, at least three consistent ratios are required at a cross section , regardless of stream position, to extend the WMR. Note that it is possible for a single cross-section to exhibit well-mixed distributions of FeDNAs (indicated by correlated concentrations present at all sampling locations) but for eDNA concentrations to have declined strongly before the next cross section downstream. In this instance, a WMR could not be rigorously defined based on the rule requiring two consecutive similar cross-sections to define he beginning and end of the WMR. However, this does not necessarily mean the FeDNAs were not well-mixed, and a higher resolution of sampling would capture this shorter WMR.

RESULTS

Assay Thresholds

LOD ranged from 2.34 to 3.89 copies per reaction across the three FeDNA (Table 1). LOQ ranged from 3 to 5 times higher than LOD across species, with LOQ for IPU estimated at about half that for the other two species (Table 1). Confidence intervals for LOD included zero and ranged up to a substantial fraction of LOQ.

Table 1. The Limit of Detection (LOD) and Limit of Quantification (LOQ) for each assay (copies per reaction). Lower and upper confidence bounds are included in the parentheses for LOD. Estimates of LOQ and LOQ indicated as measures of uncertainty for LOQ

Target Species	LOD	LOQ
Cyprinus carpio	2.55 (0, 6.96)	15 (13,17)
Micropterus salmoides	3.89 (0, 9.15)	13 (12,15)
Ictalurus punctatus	2.34 (0, 3.16)	6 (5,8)

Field trials yielded 315 water samples and 35 field blanks. None of the field blanks recorded eDNA concentrations above the LOD. Assays for 14 samples failed to amplify any FeDNA, with the number per trial ranging from 0 samples in Jacoby Creek to 7 failed samples in Prairie Creek Low Flow, leaving 301 samples for analysis (Table 3). Table 2 presents a summary of site characteristics during each of the trials. Table 3 presents the number of samples removed from the data set during sequential screening methods.

The number of samples retained based on threshold screening varied by site. Assays for 199 samples were above the LOD, with the number per trial ranging from 34 in Old Campbell Creek to 44 in Prairie Creek High Flow. Assays for 164 samples were above the LOQ, with the number per trial ranging from 30 in Old Campbell Creek to 40 in Prairie Creek High Flow. Note that assays are included in these totals if the sample had at least one pairwise comparison retained following the threshold screening.

Table 2. Characteristics of each stream during each trial in 2021.

Site	Date	Temperature (C°)	Discharge (CMS)	Avg. Depth (m)	Avg. Velocity (m/s)	Avg. Width (m)
Jacoby Creek	July 15th	14	0.0016	0.27	0.018	5.87
Little River	September 22nd	14.2	0.0448	0.34	0.052	7.58
Old Campbell Creek	August 20th	13.5	0.0861	0.32	0.110	6.1
Prairie Creek Low Flow	September 8th	13.5	0.0062	0.20	0.055	3.88
Prairie Creek High Flow	November 11th	12.5	0.1423	0.38	0.424	5.11

Table 3. Summary of FeDNA pairs retained at each sequential step of screening from 63 total pairs and relationship between FeDNAs in the correlation-based screening. Site acronym definitions: JC-Jacoby Creek, LR-Little River, OC-Old Campbell Creek, PCL-Prairie Creek Low Flow, PCH-Prairie Creek High Flow. Valid Pairs is the number of assays without unusual results. Pairs retained LOQ is the number of FeDNA pairs retained after the a priori LOQ screening. Pairs retained 0.15 RMSE is the number of FeDNA pairs correlated to a 0.15 RMSE for the linear model and retained in the final data set. FeDNA relationships are from the pairwise comparison of FeDNAs in the correlation-based screening as described in the methods. Each FeDNA relationship is presented as a linear equation by site. Standard error for the estimates of slope and intercept are included in the parentheses. R-squared for each relationship is included as an indication of goodness of fit within retained data.

Site	FeDNA Comparison	Valid Pairs	Pairs retained LOD	Pairs Retained LOQ	Pairs Retained 0.15 RMSE	FeDNA Relationship	R^2
JC	IPU:CCA	63	36	31	28	IPU = 1.40 (0.11) + 0.78 (0.03)*CCA	0.97
JC	IPU:MSA	63	41	39	28	IPU = 1.45 (0.10) + 0.89 (0.03)*MSA	0.97
JC	CCA:MSA	63	36	31	29	MSA = -0.17 (0.14) + 0.90 (0.04)*CCA	0.96
LR	IPU:CCA	62	41	37	35	IPU = 0.09 (0.14) + 0.90 (0.04)*CCA	0.94
LR	IPU:MSA	62	39	29	21	IPU = -0.08 (0.28) + 1.12 (0.10)*MSA	0.88

Site	FeDNA Comparison	Valid Pairs	Pairs retained LOD	Pairs Retained LOQ	Pairs Retained 0.15 RMSE	FeDNA Relationship	R^2
LR	CCA:MSA	61	39	29	20	MSA = 0.34 (0.28) + 0.74 (0.08)*CCA	0.82
OC	IPU:CCA	59	28	18	11	IPU = 0.88 (0.24) + 0.94 (0.08)*CCA	0.94
OC	IPU:MSA	59	34	30	15	IPU = 0.50 (0.23) + 1.08 (0.08)*MSA	0.94
OC	CCA:MSA	59	28	18	14	MSA = 0.31 (0.19) + 0.87 (0.06) *CCA	0.94
PCH	IPU:CCA	60	22	2	1	not calculated	N/A
PCH	IPU:MSA	60	44	40	30	IPU = 0.78 (0.25) + 0.94 (0.08)*MSA	0.84
PCH	CCA:MSA	60	22	2	2	not calculated	N/A
PCL	IPU:CCA	56	37	32	26	IPU = 0.4 (0.29) + 0.89 (0.07)*CCA	0.92
PCL	IPU:MSA	56	39	33	21	IPU = 0.91 (0.22) + 0.87 (0.07)*MSA	0.90
PCL	CCA:MSA	56	37	32	20	MSA = -0.06 (0.24)+ 0.91 (0.06)*CCA	0.92

Correlation Between FeDNA Sources

  Assay failure rates and rates of inclusion using the correlation-based screening were site dependent (Table 3). The number of retained pairwise comparisons ranged from 2 for CCA:IPU in Prairie Creek High Flow to 35 for CCA:IPU in Little River. The CCA tracer in Prairie Creek High Flow was at a much lower concentration than the MSA and IPU tracers and had few correlated points (Figure 6, Figure 7). Comparisons with MSA in both the Little River and Prairie Creek Low Flow sites yielded fewer retained samples than other comparisons.

  Several patterns are shared across trials, while others illustrate important differences. In Jacoby Creek and Little River, the number of observations retained decreased as distance increased, nearing zero by 500 m (Figure 7). In contrast, for both Prairie Creek trials and Old Campbell Creek, the number of observations retained remained relatively stable out to 1000 m (Figure 7). In Prairie Creek Low Flow and Little River, mixing during downstream transport over a stretch of 150 m was required before and FeDNA distributions become correlated (Figure 6). For the other sites, the number of observations retained increased after mixing over the course of being carried 50-100 meters downstream.

  Each of the relationships between FeDNA sources had r^2 values exceeding 0.8 (Table 3). All intercept estimates were well within the bounds of the LOD, meaning they were effectively zero and went through the origin (Table 3). None of the intercepts indicated a non-linear relationship between FeDNAs.
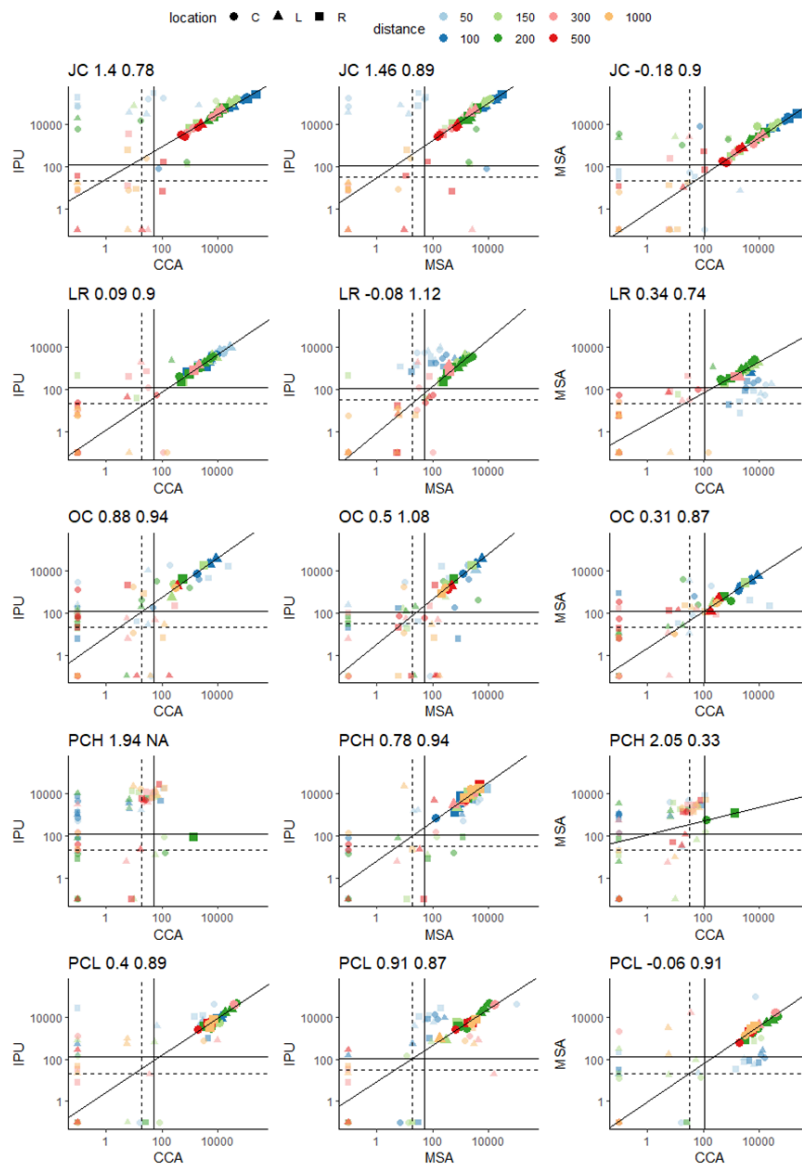
Figure 6. Pairwise correlations between FeDNA tracers by comparison (columns) and trial (rows). The colors in the plot indicate sampled cross-section distances. The shapes indicate the sampling location (C-center, L-left, R-right). Site acronym definitions: JC-Jacoby Creek, LR-Little River, OC-Old Campbell Creek, PCL-Prairie Creek Low Flow, PCH-Prairie Creek High Flow. The dashed line indicates the LOD for the species plotted on that axis. The solid vertical and horizontal lines indicate the LOQ for the species plotted on that axis. The diagonal solid lines in the plot are a linear model fit between the two sources with slope and intercept indicated in the top left corner. Transparent points are points dropped to achieve an RMSE <= 0.15, solid points are points included in the model.
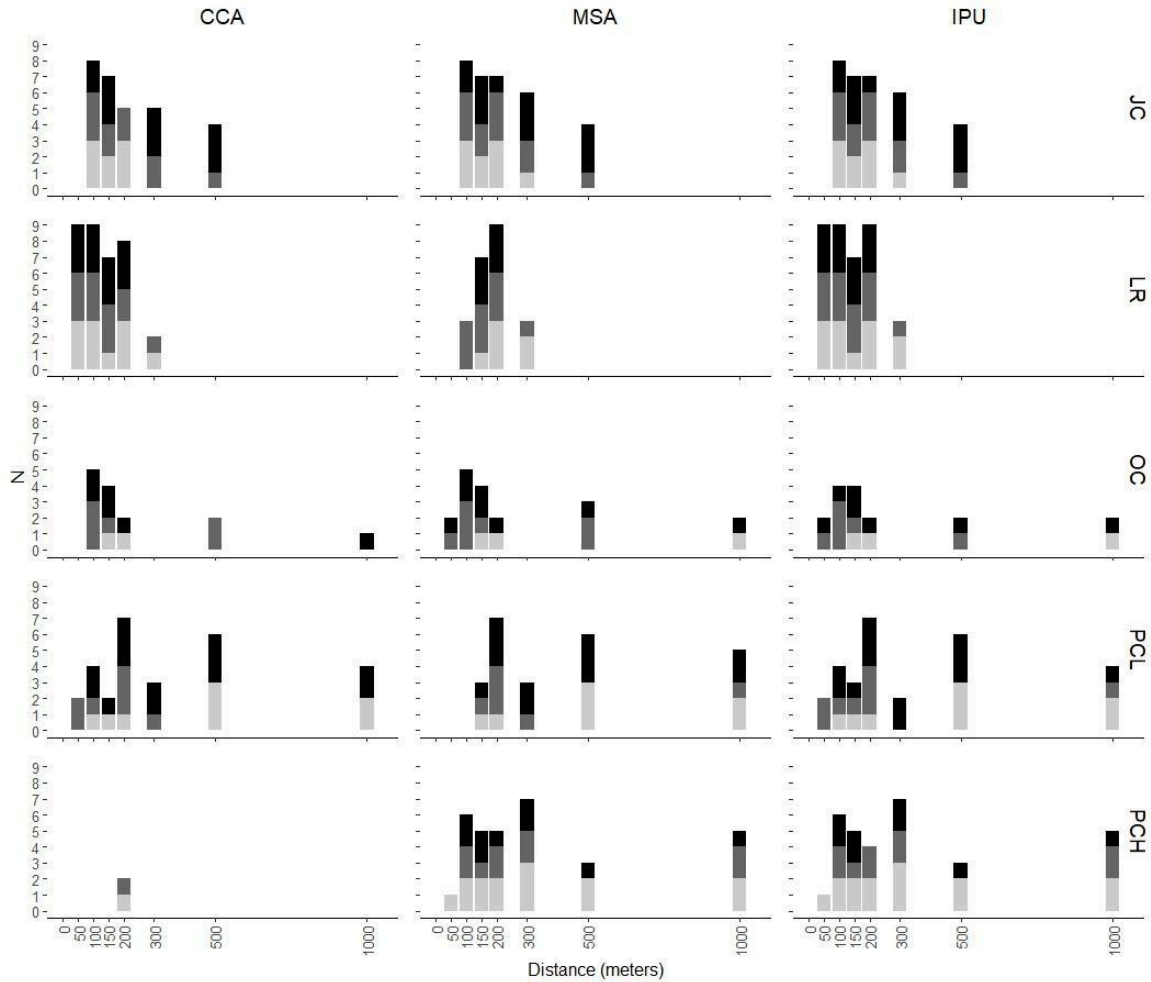
Figure 7. The number of observations retained at each site following pairwise correlationscreening. The site and species for a panel is indicated on the top and right of each column and row. Color denotes the location in the stream with the left being black, center being dark gray, and right being light gray.

Data Summary and Effects of Aggregation

Jacoby Creek

Observed distributions of all three FeDNA sources exhibited similar distributions in both screened and unscreened data sets, whether resolved by cross-stream sampling location, or in aggregate at each cross-section (Figure 8). Coherent steady declines in concentration were observed for all FeDNA sources at the center and left side of the stream while a slight increase in (measured) FeDNA concentrations was observed at the right side at 150 m.
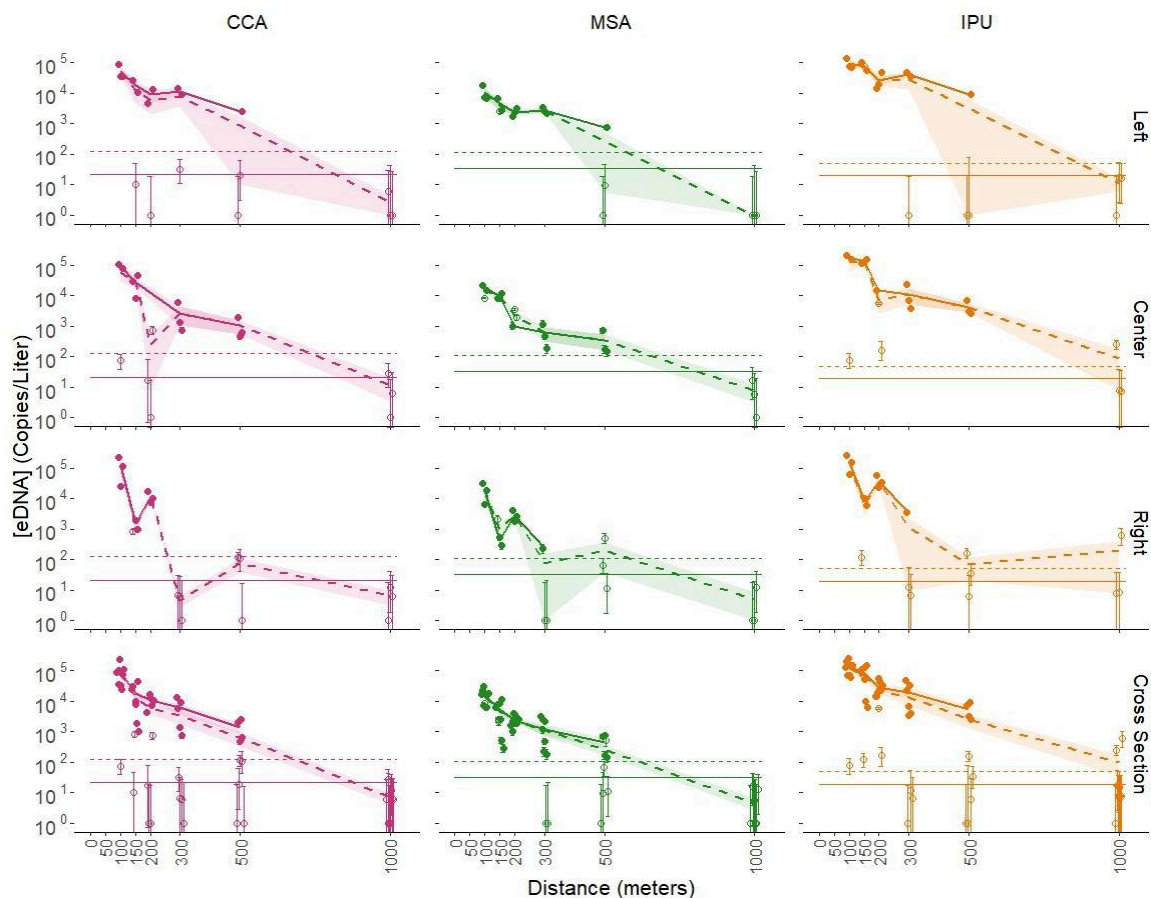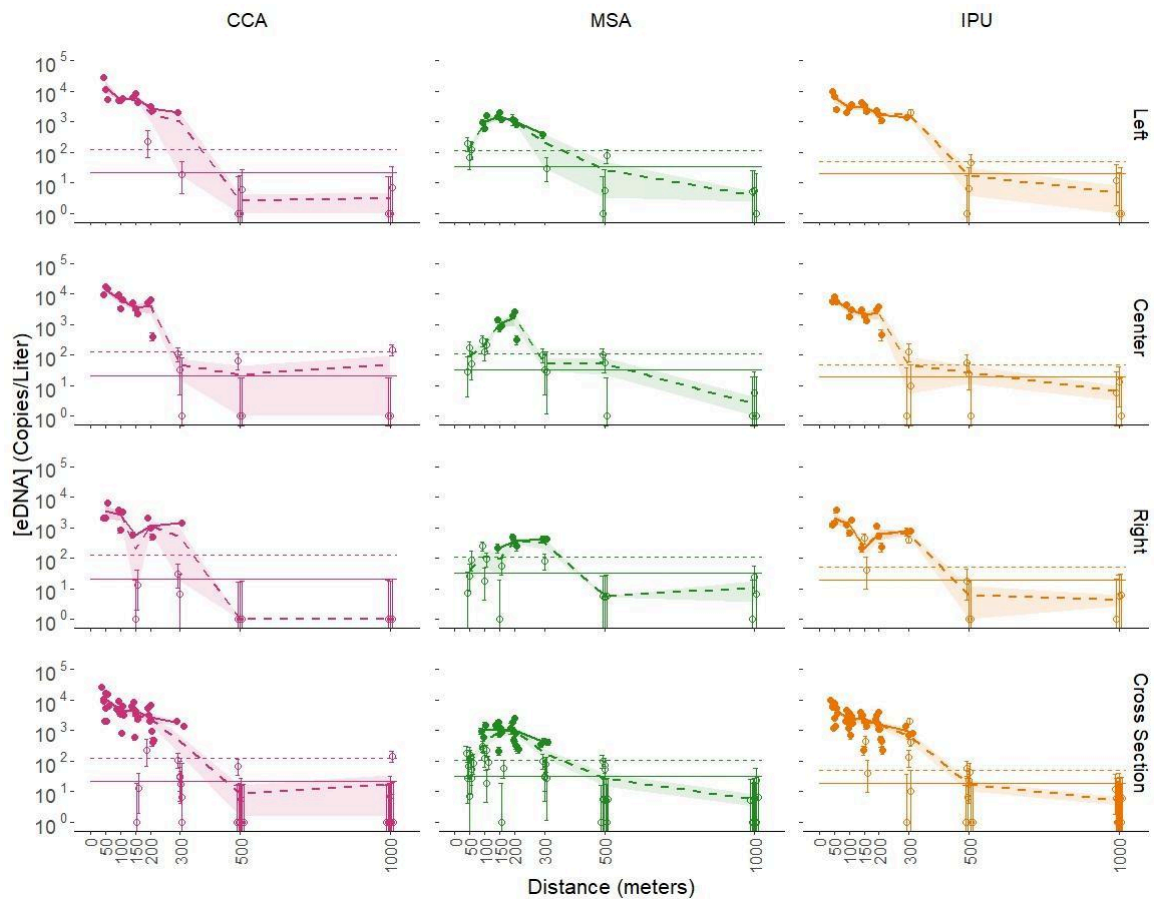
Figure 8. Concentrations (copies/liter) of CCA (left column; pink), MSA (center column; green), and IPU (right column; orange) at steady state for the Jacoby Creek trial. Note log-scaled y-axis. Solid points (average as solid lines) indicate values retained from the correlation-based screen; hollow circles (average including hollow and solid points as dashed lines) indicate values excluded from the correlation-based screen. Error bars for each point are derived from Poisson distributions and are obscured at larger values due to log-scaling. Points are jittered slightly along the X-axis for clarity. Standard errors of the mean are for each average (lines) are indicated by the shaded region bounding the lines. The standard errors are only calculated where replicate samples are included in calculating the average. Horizontal lines indicate FeDNA-specific LOD (solid line) and LOQ (dashed line) in copies per liter.

Little River

Observed distributions of CCA and IPU exhibited coherent patterns when resolved by cross-stream sampling location or in aggregate at each cross-section. This is best illustrated by coherent decline in measured concentration along the left side of the stream and sharp decrease along the right side of the stream (Figure 9). In contrast, the distribution of MSA differed substantially from that of the other two FeDNAs in the first 150 m, below which a more coherent pattern emerged. This appears to reflect a case where mixing processes took longer to fully blend MSA with the other tracers.

All FeDNA distributions generally declined with distance, generally dropping below the LOQ after 300 m, which corresponded with a large pool present between 300 and 500 m in the Little River trial site. No observations , whether above or below LOQ, are retained following the correlation-based screen.
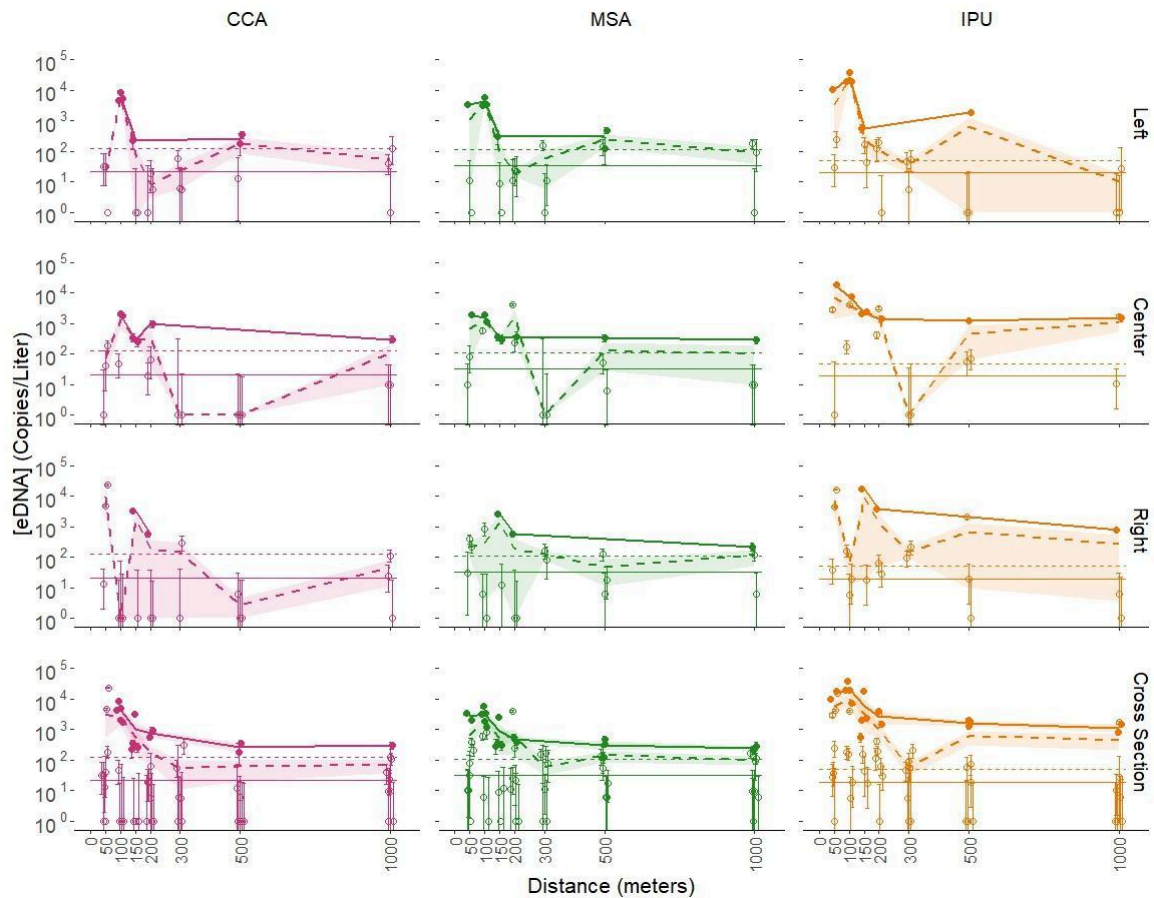

Figure 9. Concentrations (copies/liter) of CCA (left column; pink), MSA (center column; green), and IPU (right column; orange) at steady state for the Little River trial. Note log-scaled y-axis. Solid points (average as solid lines) indicate values retained from the correlation-based screen; hollow circles (average including hollow and solid points as dashed lines) indicate values excluded from the correlation-based screen. Error bars for each point are derived from Poisson distributions and are obscured at larger values due to log-scaling. Points are jittered slightly along the X-axis for clarity. Standard errors of the

mean are for each average (lines) are indicated by the shaded region bounding the lines. The standard errors are only calculated where replicate samples are included in calculating the average. Horizontal lines indicate FeDNA-specific LOD (solid line) and LOQ (dashed line) in copies per liter.

Old Campbell Creek

Complex spatial patterns observed at this site reflect apparent effects of a split channel at 300 m where samples collected contained little FeDNA. The effect of this habitat feature is best observed in the unscreened data where there was a decline in measured concentration at 300 m in the center of the stream in each FeDNA (Figure 10).

Excluding the 300 m sampling location from the correlation-screened data set yielded in more coherent spatial patterns at cross-sections and sampling locations. A substantial decline in measured FeDNA concentration was observed in all sources at 200 m at the left side of the stream (Figure 10). In contrast, there was no substantial decline observed in any of the FeDNA tracers at the center and right side of the stream. A sharp decrease in measured concentration is observed at cross-sections over the first 200 m below which concentrations appear to stabilize.
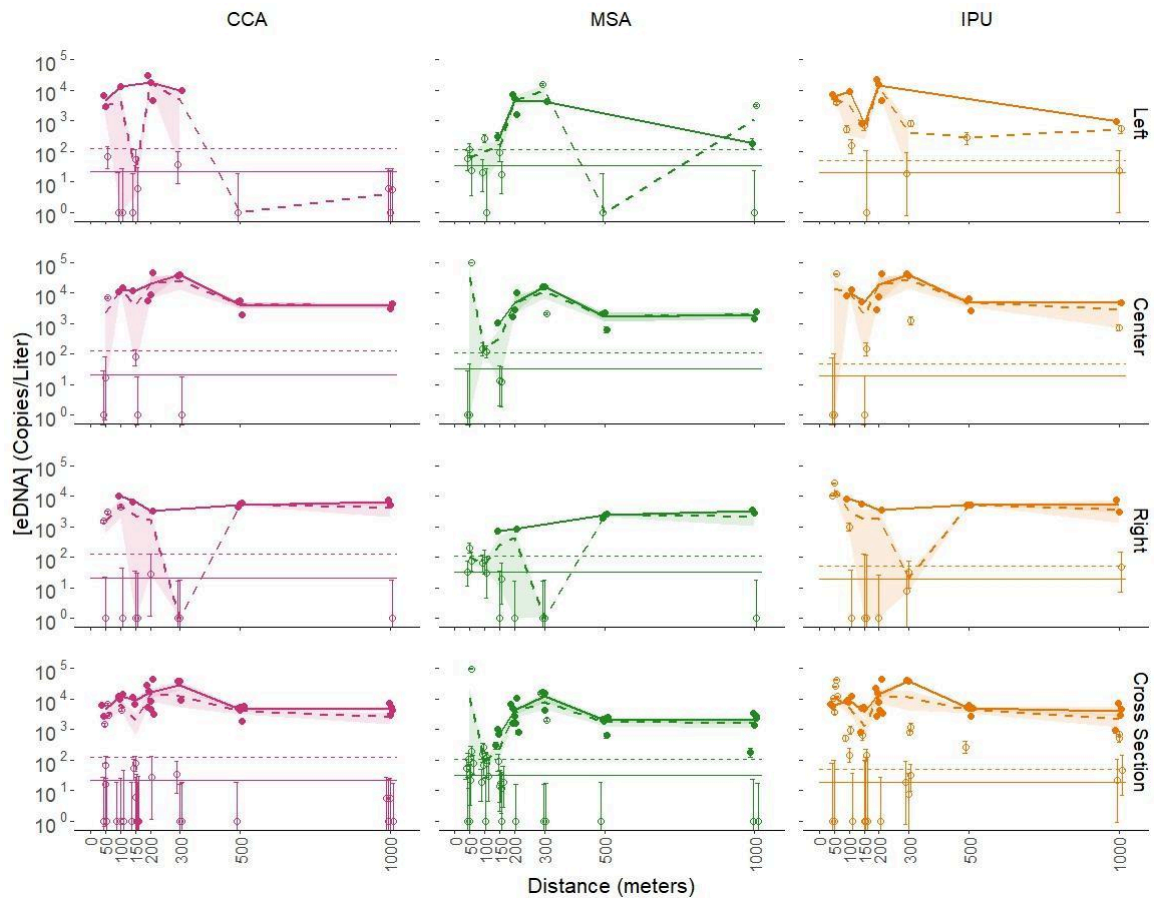

Figure 10. Concentrations (copies/liter) of CCA (left column; pink), MSA (center column; green), and IPU (right column; orange) at steady state for the Old Campbell Creek trial. Note log-scaled y-axis. Solid points (average as solid lines) indicate values retained from the correlation-based screen; hollow circles (average including hollow and solid points as dashed lines) indicate values excluded from the correlation-based screen. Error bars for each point are derived from Poisson distributions and are obscured at larger values due to log-scaling. Points are jittered slightly along the X-axis for clarity. Standard errors of the mean are for each average (lines) are indicated by the shaded region bounding the lines.

The standard errors are only calculated where replicate samples are included in calculating the average. Horizontal lines indicate FeDNA-specific LOD (solid line) and LOQ (dashed line) in copies per liter.

Prairie Creek Low Flow

   Observed distributions of FeDNA exhibited incoherent spatial patterns by cross-section for the first 200 m before becoming similar around 200 m (Figure 11). This is illustrated by a strong increase in all measured FeDNA concentrations at the 300 m cross-sections, before a stable concentration from 500 to 1000 m. These patterns are exhibited in screened and unscreened data and do not exhibit a clear general decline in concentration.

   Observed distributions of all three FeDNA sources exhibited similar spatial patterns resolved along the center and right side of the stream (Figure 11). This is illustrated by similar low concentrations directly downstream of introduction followed by the highest measured concentration at 300 m, and steady concentrations from 100 to 1000 m on the right side of the stream.
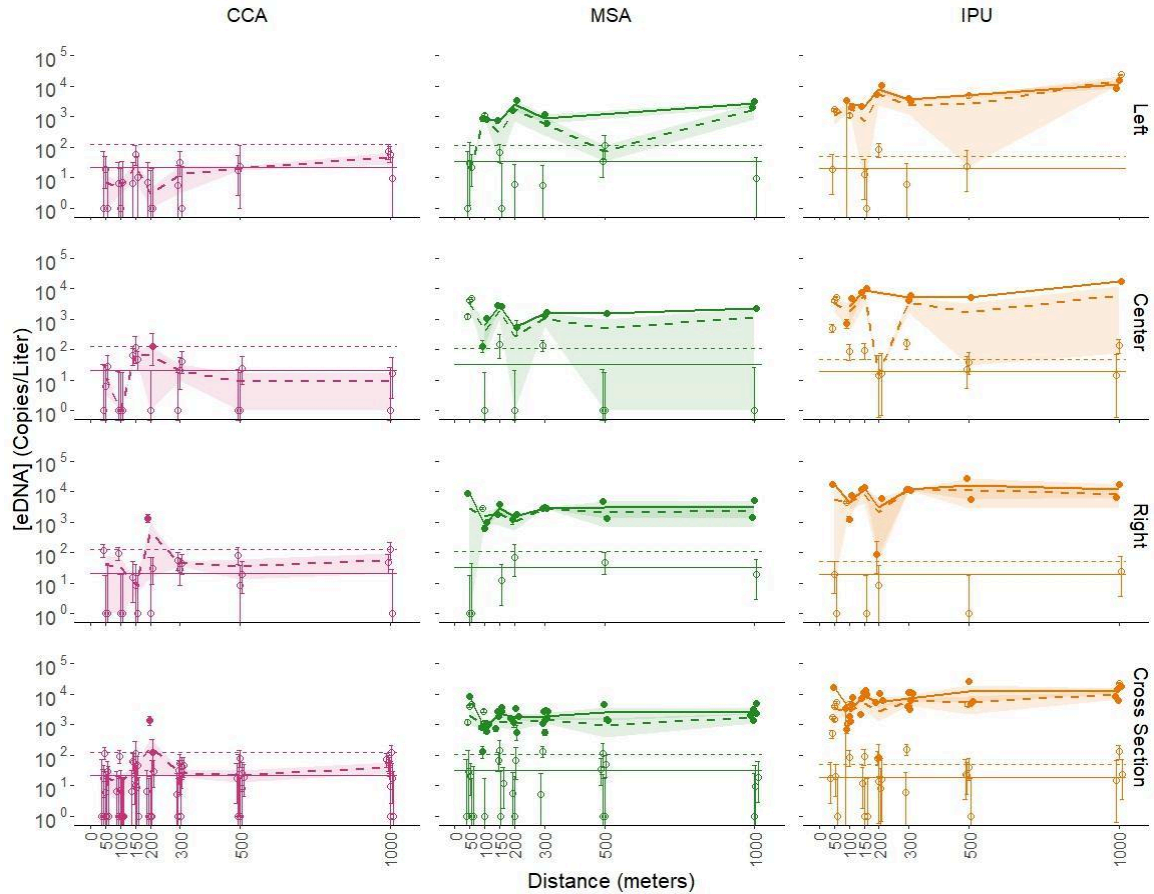

Figure 11. Concentrations (copies/liter) of CCA (left column; pink), MSA (center column; green), and IPU (right column; orange) at steady state for the Prairie Creek Low Flow trial. Note log-scaled y-axis. Solid points (average as solid lines) indicate values retained from the correlation-based screen; hollow circles (average including hollow and solid points as dashed lines) indicate values excluded from the correlation-based screen. Error bars for each point are derived from Poisson distributions and are obscured at larger values due to log-scaling. Points are jittered slightly along the X-axis for clarity. Standard errors of the mean are for each average (lines) are indicated by the shaded region

bounding the lines. The standard errors are only calculated where replicate samples are included in calculating the average. Horizontal lines indicate FeDNA-specific LOD (solid line) and LOQ (dashed line) in copies per liter.

Prairie Creek High Flow

No patterns were noted for CCA as few concentrations were retained after the correlation screening. Observed distributions of IPU and MSA exhibited similar spatial patterns in the correlation-screened dataset whether resolved by sampling location or by cross-section (Figure 12). At all locations and cross-sections of the stream, similar declines and increases in measured concentrations were observed in the first 200 m, below which concentrations remained steady down to 1000 m below the ADIDs. There was not a clear decline in concentration in this trial in any of the FeDNA sources.

Figure 12. Concentrations (copies/liter) of CCA (left column; pink), MSA (center column; green), and IPU (right column; orange) at steady state for the Prairie Creek High Flow trial. Note log-scaled y-axis. Solid points (average as solid lines) indicate values retained from the correlation-based screen; hollow circles (average including hollow and solid points as dashed lines) indicate values excluded from the correlation-based screen. Error bars for each point are derived from Poisson distributions and are obscured at larger values due to log-scaling. Points are jittered slightly along the X-axis for clarity. Standard errors of the mean are for each average (lines) are indicated by the shaded region bounding the lines. The standard errors are only calculated where replicate samples are included in calculating the average. Horizontal lines indicate FeDNA-specific LOD (solid line) and LOQ (dashed line) in copies per liter.

Patterns of Detection Probability and Average Concentration

Detection probability and average concentration exhibited similar patterns among FeDNA tracers within sites but differed across sites (Figure 13). In two of the sites (Jacoby Creek and Little River), detection probability fell below 50 percent, consistent with drops in average concentration below the LOD. In contrast, declines in detection probability were not monotonic with distance downstream in either Prairie Creek trials or Old Campbell Creek. These patterns were generally consistent with longitudinal patterns in average FeDNA concentration (Figure 13).
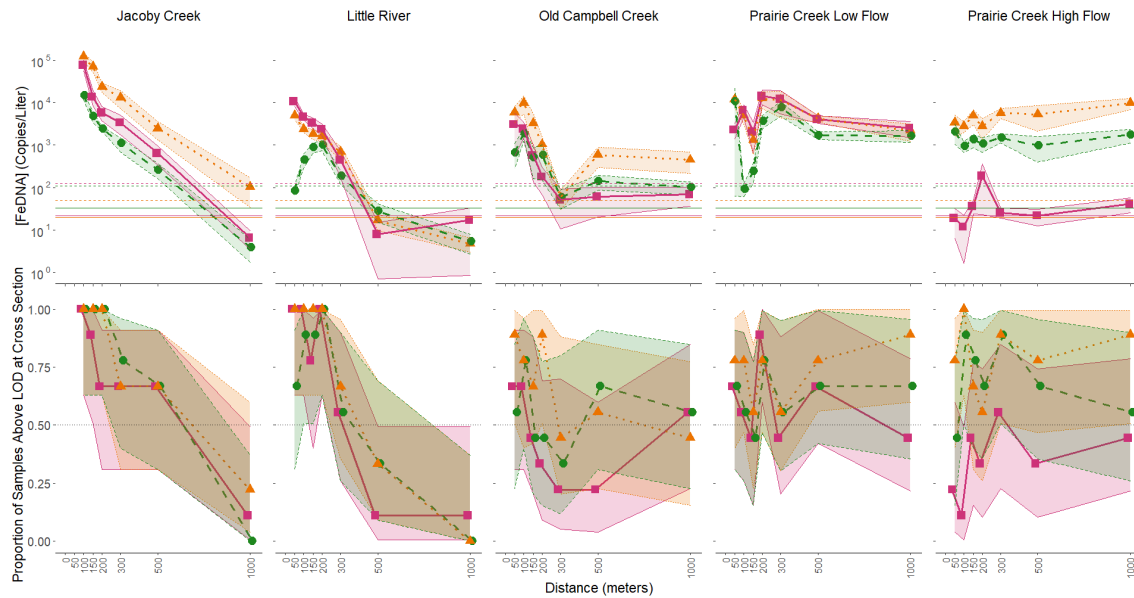
Figure 13. Change in mean concentration across the full data sets (upper panels) and detection probability (lower panels) patterns of FeDNA as a function of distance in each trial. Note log-scaled y-axis in the upper panels. A jitter is applied to points and lines for clarity. FeDNAs are differentiated by color, line type, and point shape: IPU (orange, dotted, triangle), MSA green, dashed, circle) and CCA (pink, solid, square). Detection probability is plotted as the proportion of samples above the LOD at a cross-section. Species specific LOQ (dashed line; upper panel) and LOD (solid line; upper panel) are included for reference. 50 percent detection probability (dotted black line; lower panel) is included for reference.

Comparison of Ratios and Well-Mixed-Regions

Well-Mixed-Regions (WMR) based on constant ratios between sequential cross-sections ranged from as short as 100 m (Little River, CCA:MSA) to as long as 900 m (Prairie Creek Low Flow, CCA:IPU and Prairie Creek High Flow, MSA:IPU) (Figure 14). In most cases, a WMR was established around 100 to 150 m downstream from the point where FeDNAs used in its definition had been introduced. Note that the WMR defined by different pairwise comparisons of FeDNAs were not always identical within streams. For example, WMR defined by the CCA:MSA in Prairie Creek Low Flow began at 300 m and extended to 1000m, while the WMR defined by the CCA:IPU in the same stream began at 100 m and extended to 1000m. Such variability appears to be related to the location of introduction for the FeDNAs, differences in the downstream scale of mixing, which is typically a function of stream width.

FeDNA ratios exhibit much more variable patterns when based on the lightly screened data set, but generally reflect the patterns identified using the correlated screened data (Figure 15). It is in this data set that some trends in mixing can be better identified. Most notably, the MSA signal in Little River and Prairie Creek Low Flow remained incompletely mixed with the other FeDNAs for around 150 m, as indicated by changing ratios. Below this, the MSA signal formed a constant ratio with the other FeDNA tracers downstream.
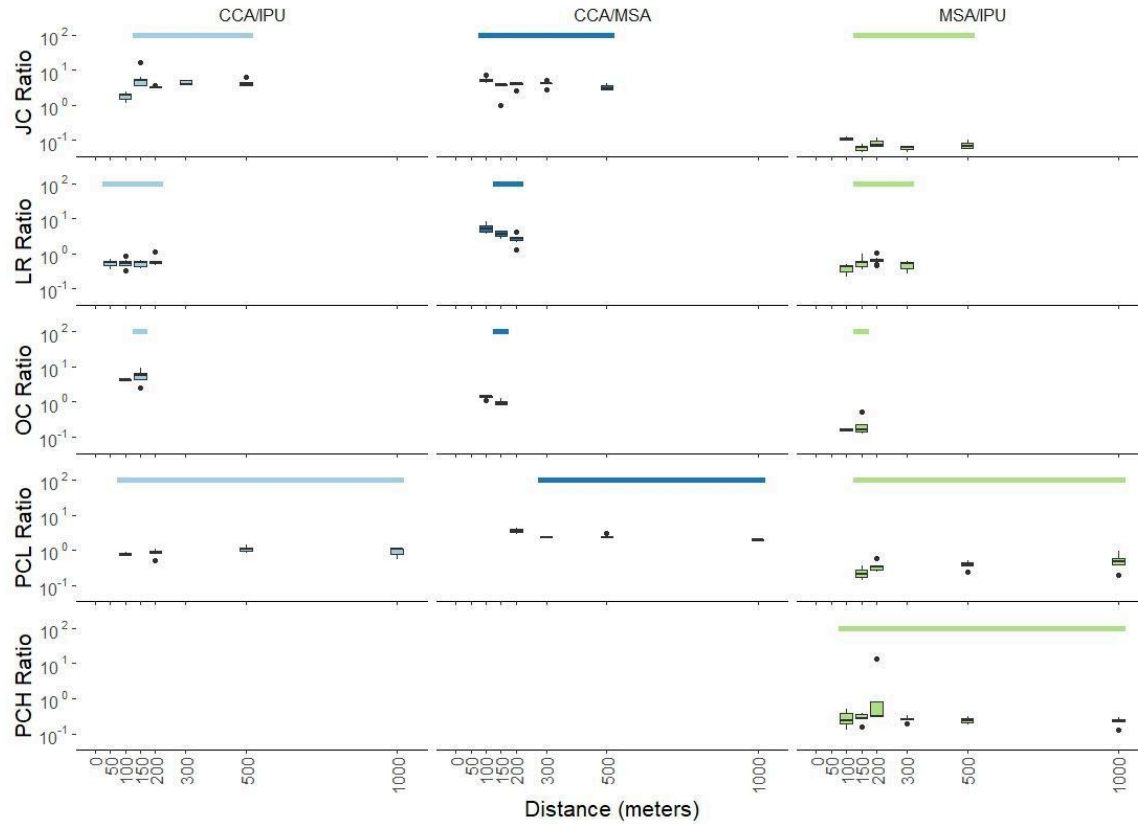
Figure 14. Average FeDNA ratios (CCA:IPU, right column, CCA:MSA, center; MSA:IPU, left) as a function of distance across trials (top to bottom row: JC, LR, OC, PCL, PCH) based on correlation-screened data. Horizontal lines delineate well-mixed region (WMR) based on ANOVA and post-hoc tests used to identify consistency of ratios over distance.
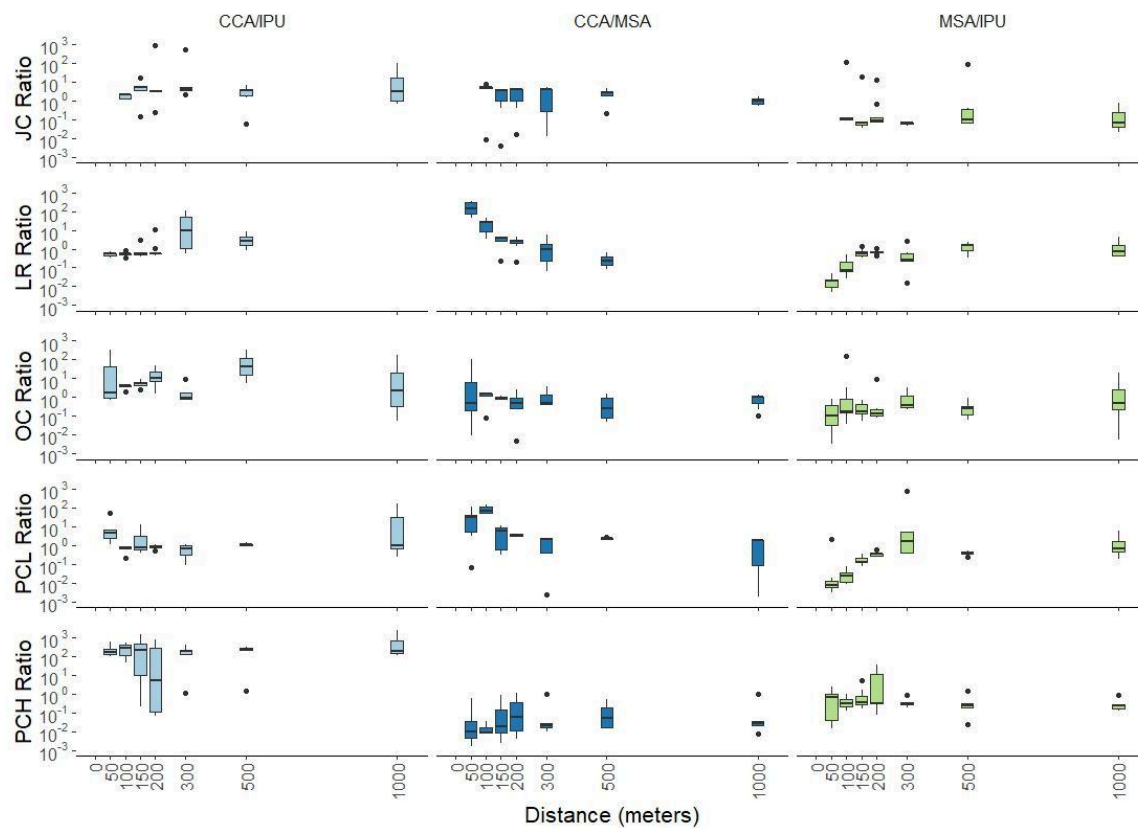
Figure 15. Average FeDNA ratios (CCA:IPU, right column, CCA:MSA, center; MSA:IPU, left) as a function of distance across trials (top to bottom row: JC, LR, OC, PCL, PCH) based on unscreened data. Note the y axis has been expanded compared to Figure 14 to encompass the entire data range.

DISCUSSION

This study demonstrates a method for characterizing shared patterns in FeDNA distributions that reflect the effects of mixing, downstream transport, and attrition (i.e., molecular decay, settling to or retention by the substrate, or other processes that remove eDNA from the water column) and how these distributions are resolved under a particular sampling scheme. Moreover, we show that this approach can resolve such patterns unique to a particular stream reach at a given time. Comparative analysis of the results of this method indicates FeDNA distribution patterns vary across different streams but are similar among distinct tracers within streams. This pattern of stream-specific patterns that are shared by all FeDNA tracers is especially robust given that FeDNA tracers were introduced at different locations in a stream in a way that provided insight into relevant scales of mixing and signal persistence. Correlated spatial patterns among distinct FeDNA tracers indicates the potential for this method to provide informative context for the interpretation of NeDNA patterns, to support calibration of quantitative NeDNA surveys, and to advance our understanding of how eDNA moves and persists through natural systems.

The potential for FeDNA tracers to calibrate NeDNA in stream environments is grounded in assumptions regarding equality of transport, catchability, detection, and loss rates among eDNAs analogous to those that under pin abundance estimates based on mark-recapture techniques. Although this study does not directly test these assumptions, the emergence of coherent patterns in FeDNA concentration within individual streams following sustained introductions of FeDNA is consistent with these assumptions, and reflects shared dynamics as FeDNA tracers mix and are emerging as the signals are carried moved downstream, mixed, and subjected to common transport, mixing, loss, and detection processes.

Despite evidence for common dynamics among FeDNAs within a stream, the scales at which these conditions were satisfied (and detectable) varied across streams. For example, estimates of the WMR were system-specific. They varied in the distance over which full mixing of the FeDNA signals was realized (100 to 300 m) and how far the WMRs could be resolved (500-900 m; Figure 14). These distances are consistent with examples from the literature that report higher and more consistent concentrations of eDNA between 100-1000 m rather than directly below a source (Pont et al. 2018; Stoeckle et al. 2021; Wood et al. 2021; Van Driessche et al. 2023). Ultimately, each data set established the basis for assessing whether FeDNA spatial patterns formed by common drivers were consistent between FeDNA tracers, enabling the resolution of system-specific distribution patterns.

More generally, spatial distributions of (sampled) FeDNA concentrations differed between sites, but were generally consistent within a stream, regardless of where samples were collected along each cross-section, i.e., FeDNA collected at various points along the left side of the stream, or the right, or center (Figures 8-12). In some settings, downstream patterns for a FeDNA tracer (or set of tracers) were observed to differ from one side of the stream to the other (Figures 8-12). Persistence of horizontal gradients and coherent shifts in concentration across

streams are consistent with patterns reported in studies using caged fish (Thalinger et al. 2020; Laporte et al. 2020). Inconsistencies across the stream suggests single bank sampling designs that are sometimes employed in eDNA surveys may not capture the entirety of an eDNA plume (Wood et al. 2021).

Downstream dispersal of FeDNA spanned distances consistent with scales reported elsewhere in the literature (i.e., 500-2000 m; Jo and Yamanaka 2022). Estimates of transport distance developed here add to the collection of site- and time-specific estimates reported in the literature. Any application of these estimates to other systems as anything more than a general rule of thumb should proceed with caution.

In 3 of the 5 trials reported here, concentrations and detection probabilities declined nearly monotonically, as expected under simple conceptual models for transport of eDNA in streams. In the other trials, however, the observed spatial distribution of FeDNA was decidedly non-monotonic. We interpret this as being a consequence of how the sampling frame intersected a distribution of eDNA consistent with a more complex realization of the conceptual model, one in which FeDNA is carried primarily in a core flow that shifts back and forth across the channel as it moves downstream mixes out of this main flow slowly relative to the speed of the main flow, and microscale transport dynamics contribute substantially to variability among parcels of water in a stream (Zhao et al. 2021; Mauvisseau et al. 2021; Jo et al. 2022).

These situations, where sampling does not yield a simple picture of FeDNA distributions, highlight the need for FeDNA tracers, as attempts to make sense of NeDNA patterns absent this context would be at substantial risk of yielding biased conclusions. In this regard, it is also possible that these patterns might reflect variability in the extraction-assay process. Importantly, FeDNA-based comparisons (e.g., ratios) are agnostic to the source of variability among samples, so long as all eDNAs within a sample experience similar transport, loss, and detection processes.

Successfully resolving and comparing FeDNA distributions depends critically on the capacity to sustain introductions of FeDNA at high concentrations over sufficient time for a 'steady state' distribution of FeDNA to develop. Establishing this steady state characterizes stream- and sampling frame-specific patterns, and is an essential source of information for interpreting observations of other FeDNA tracers, as demonstrated here, and by extension, natural eDNA distributions. Interpretations based on an FeDNA distribution that had not yet reached a steady-state distribution would yield biased conclusions, especially in areas near the leading edge of the FeDNA plume and further downstream. Note that calibrating these patterns and characterizing scales of mixing is much more robust when based on comparing independent, distinct FeDNA signals introduced concurrently at different locations in the stream.   .

By demonstrating a capacity to identify scales of mixing and detection that delineate the extent of well-mixed regions, we are able to establish guidelines for where the use of FeDNA tracers is likely to be useful for supporting interpretation of natural eDNA observations. In the cases considered here, such regions ranged in size and location across streams, highlighting the importance of case-by-case assessments. Specifically, the approach taken here rests on the use of internally consistent data based on correlated FeDNA concentrations (i.e., constant

FeDNA:FeDNA ratios) consistent with shared transport and detection processes. These data are amenable to statistical tests (ANOVA and post-hoc comparisons) to assess constancy of FeDNA ratios across sequential cross-sections, which, in conjunction with rules regarding minimum inclusion of sampling locations in a cross-section, provide a rigorous basis for delineating a WMR (Figure 14). Visual inspection of the unscreened data (Figure 15) suggests that a similar result might be achieved by simply discarding obvious outliers from each set of estimated ratios at a cross-section. This approach would, however, risk undermining transparency and repeatability unless specific rules were applied and reported.

The successful proof-of-concept presented here suggests several future applications while identifying several challenges for broader use of this approach. We expect that the ADID approach will prove useful as a tool for calibrating NeDNA surveys for occupancy and (with further work) abundance estimation. In these cases, interpretation will rest on the ability to detect NeDNA:FeDNA ratios that depart from the null hypothesis of constant ratios and thus provide evidence of intervening introduction of NeDNA from organisms in the study reach. ADIDs are also a promising tool for more carefully controlled studies designed to closely measure and model how eDNA moves in natural systems (e.g., Carraro et al. 2018, 2021). With respect to monitoring applications, the use of ADIDs has potential as a positive control for eDNA surveys of rare or cryptic species, and specifically offers the ability to closely mimic how eDNA is introduced to the environment by such species. Important questions that bear upon the broader application of this approach include (1) assessment of the costs and benefits of this approach in terms of materials, the time required to establish the FeDNA steady state necessary for robust results, and how to allocate samples to resolve questions of interest, and (2) whether it is possible to effectively scale this approach to larger systems, and more fundamentally, (3) how well does FeDNA mimic its natural counterparts. Future applications should also take into consideration ways to adapt sampling frames to the structure of the study system, as this will likely pay off in terms of more robust characterization of eDNA distributions (Troth et al., 2021). This might be accomplished by accounting for location of the main flow (typically the deepest and fast-moving section of the river) or by avoiding sampling locations that are likely to be especially difficult to sample representatively (e.g., split channels, large pools, off channel eddies, etc.).

In summary, the use of FeDNA tracers as described here offers a novel new tool for exploring the influence of source location, mixing, and transport dynamics on eDNA concentrations on a case by case basis (per Rourke et al. 2022), and to support improvements in eDNA surveys by directly mitigating the impact of potential sources of error (see Burian et al. 2021), simply by understanding the underlying pattern as context for interpreting natural eDNA signals. In this light, the combination of FeDNA and ADID represents a promising advance in the application of eDNA to fisheries management and conservation.

LITERATURE CITED

Ali, A. S., Z. Zanzinger, D. Debose, and B. Stephens. 2016. Open Source Building Science Sensors (OSBSS): A low-cost Arduino-based platform for long-term indoor environmental data collection. Building and Environment 100:114–126.

Andruszkiewicz Allan, E., W. G. Zhang, A. Lavery, and A. Govindarajan. 2021. Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. Environmental DNA 3(2):492–514.

Bedwell, M. E., and C. S. Goldberg. 2020. Spatial and temporal patterns of environmental DNA detection to inform sampling protocols in lentic and lotic systems. Ecology and Evolution 10(3):1602–1612.

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

Buxton, A., E. Matechou, J. Griffin, A. Diana, and R. A. Griffiths. 2021. Optimising sampling and analysis protocols in environmental DNA studies. Scientific Reports 11(1):11637.

Bylemans, J., E. M. Furlan, D. M. Gleeson, C. M. Hardy, and R. P. Duncan. 2018. Does Size Matter? An Experimental Evaluation of the Relative Abundance and Decay Rates of Aquatic Environmental DNA. Environmental Science & Technology 52(11):6408–6416.

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

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

Carrera, J., M. W. Saaltink, J. Soler-Sagarra, J. Wang, and C. Valhondo. 2022. Reactive Transport: A Review of Basic Concepts with Emphasis on Biochemical Processes. Energies 15(3):925.

Caza‑Allard, I., M. Laporte, G. Côté, J. April, and L. Bernatchez. 2022. Effect of biotic and abiotic factors on the production and degradation of fish environmental DNA: An experimental evaluation. Environmental DNA 4(2):453–468.

Curtis, A. N., J. S. Tiemann, S. A. Douglass, M. A. Davis, and E. R. Larson. 2021. High stream flows dilute environmental DNA (eDNA) concentrations and reduce detectability. Diversity and Distributions 27(10):1918–1931.

Dahlke, H. E., A. G. Williamson, C. Georgakakos, S. Leung, A. N. Sharma, S. W. Lyon, and M. T. Walter. 2015. Using concurrent DNA tracer injections to infer glacial flow pathways: Using Concurrent DNA Tracer Injections to Infer Glacial Flow Pathways. Hydrological Processes 29(25):5257–5274.

Dimond, J. L., B. R. Gathright, J. V. Bouma, H. S. Carson, and K. Sowul. 2022. Detecting endangered pinto abalone (*Haliotis kamtschatkana*) using environmental DNA : Comparison of DDPCR , QPCR , and conventional diver surveys. Environmental DNA:edn3.351.

Duda, J. J., M. S. Hoy, D. M. Chase, G. R. Pess, S. J. Brenkman, M. M. McHenry, and C. O. Ostberg. 2021. Environmental DNA is an effective tool to track recolonizing migratory fish following large‑scale dam removal. Environmental DNA 3(1):121–141.

Eichmiller, J. J., P. G. Bajer, and P. W. Sorensen. 2014. The Relationship between the Distribution of Common Carp and Their Environmental DNA in a Small Lake. PLoS ONE 9(11):e112611.

Goldberg, C. S., K. M. Strickler, and D. S. Pilliod. 2015. Moving environmental DNA methods from concept to practice for monitoring aquatic macroorganisms. Biological Conservation 183:1–3.

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

Hauer, F. R., and G. A. Lamberti, editors. 2006. Methods in stream ecology2nd ed. Academic Press/Elsevier, San Diego, Calif.

Jerde, C. L., B. P. Olds, A. J. Shogren, E. A. Andruszkiewicz, A. R. Mahon, D. Bolster, and J. L. Tank. 2016. Influence of Stream Bottom Substrate on Retention and Transport of Vertebrate Environmental DNA. Environmental Science & Technology 50(16):8770–8779.

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(3):1135–1146.

Jo, T., K. Takao, and T. Minamoto. 2022. Linking the state of environmental DNA to its application for biomonitoring and stock assessment: Targeting mitochondrial/nuclear genes, and different DNA fragment lengths and particle sizes. Environmental DNA 4(2):271–283.

Jo, T., and H. Yamanaka. 2022. Meta‑analyses of environmental DNA downstream transport and deposition in relation to hydrogeography in riverine environments. Freshwater Biology 67(8):1333–1343.

Klymus, K. E., C. M. Merkes, M. J. Allison, C. S. Goldberg, C. C. Helbing, M. E. Hunter, C. A. Jackson, R. F. Lance, A. M. Mangan, E. M. Monroe, A. J. Piaggio, J. P. Stokdyk, C. C. Wilson, and C. A. Richter. 2020. Reporting the limits of detection and quantification for environmental DNA assays. Environmental DNA 2(3):271–282.

Knudsen, S. W., R. B. Ebert, M. Hesselsøe, F. Kuntke, J. Hassingboe, P. B. Mortensen, P. F. Thomsen, E. E. Sigsgaard, B. K. Hansen, E. E. Nielsen, and P. R. Møller. 2019. Species-specific detection and quantification of environmental DNA from marine fishes in the Baltic Sea. Journal of Experimental Marine Biology and Ecology 510:31–45.

Laporte, M., B. Bougas, G. Côté, O. Champoux, Y. Paradis, J. Morin, and L. Bernatchez. 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.

Laramie, M. B., D. S. Pilliod, and C. S. Goldberg. 2015. Characterizing the distribution of an endangered salmonid using environmental DNA analysis. Biological Conservation 183:29–37.

Lemke, D., Z. Liao, T. Wöhling, K. Osenbrück, and O. A. Cirpka. 2013. Concurrent conservative and reactive tracer tests in a stream undergoing hyporheic exchange: Conservative and Reactive Tracer Tests in Streams. Water Resources Research 49(5):3024–3037.

Mauvisseau, Q., L. Harper, M. Sander, R. H. Hanner, H. Kleyer, and K. Deiner. 2021. The multiple states of environmental DNA and what is known about their persistence in aquatic environments. Preprints, preprint.

Miller, J., and T. Georgian. 1992. Estimation of Fine Particulate Transport in Streams Using Pollen as a Seston Analog. Journal of the North American Benthological Society 11(2):172–180.

Moushomi, R., G. Wilgar, G. Carvalho, S. Creer, and M. Seymour. 2019. Environmental DNA size sorting and degradation experiment indicates the state of Daphnia magna mitochondrial and nuclear eDNA is subcellular. Scientific Reports 9(1):12500.

Nagler, M., S. M. Podmirseg, J. Ascher‑Jenull, D. Sint, and M. Traugott. 2022. Why eDNA fractions need consideration in biomonitoring. Molecular Ecology Resources 22(7):2458–2470.

Nukazawa, K., Y. Hamasuna, and Y. Suzuki. 2018. Simulating the Advection and Degradation of the Environmental DNA of Common Carp along a River. Environmental Science & Technology 52(18):10562–10570.

Ostberg, C. O., and D. M. Chase. 2022. Ontogeny of eDNA shedding during early development in Chinook Salmon (*Oncorhynchus tshawytscha*). Environmental DNA 4(2):339–348.

Ostberg, C. O., D. M. Chase, M. S. Hoy, J. J. Duda, M. C. Hayes, J. C. Jolley, G. S. Silver, and C. Cook‑Tabor. 2019. Evaluation of environmental DNA surveys for identifying occupancy and spatial distribution of Pacific Lamprey (*Entosphenus tridentatus*) and *Lampetra* spp. in a Washington coast watershed. Environmental DNA 1(2):131–143.

Payn, R. A., M. N. Gooseff, D. A. Benson, O. A. Cirpka, J. P. Zarnetske, W. B. Bowden, J. P. McNamara, and J. H. Bradford. 2008. Comparison of instantaneous and constant-rate stream tracer experiments through non-parametric analysis of residence time distributions: RESIDENCE TIME DISTRIBUTIONS FROM STREAM TRACER EXPERIMENTS. Water Resources Research 44(6).

Penaluna, B. E., J. M. Allen, I. Arismendi, T. Levi, T. S. Garcia, and J. K. Walter. 2021. Better boundaries: identifying the upper extent of fish distributions in forested streams using eDNA and electrofishing. Ecosphere 12(1).

Pine, W. E., K. H. Pollock, J. E. Hightower, T. J. Kwak, and J. A. Rice. 2003. A Review of Tagging Methods for Estimating Fish Population Size and Components of Mortality. Fisheries 28(10):10–23.

Pont, D., M. Rocle, A. Valentini, R. Civade, P. Jean, A. Maire, N. Roset, M. Schabuss, H. Zornig, and T. Dejean. 2018. Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. Scientific Reports 8(1):10361.

Robinson, A. T., Y. M. Paroz, M. J. Clement, T. W. Franklin, J. C. Dysthe, M. K. Young, K. S. McKelvey, and K. J. Carim. 2019. Environmental DNA Sampling of Small‑Bodied Minnows: Performance Relative to Location, Species, and Traditional Sampling. North American Journal of Fisheries Management 39(5):1073–1085.

Rojahn, J., L. Pearce, D. M. Gleeson, R. P. Duncan, D. M. Gilligan, and J. Bylemans. 2021. The value of quantitative environmental DNA analyses for the management of invasive and endangered native fish. Freshwater Biology 66(8):1619–1629.

Rourke, M. L., A. M. Fowler, J. M. Hughes, M. K. Broadhurst, J. D. DiBattista, S. Fielder, J. Wilkes Walburn, and E. M. Furlan. 2022. Environmental DNA (eDNA) as a tool for

assessing fish biomass: A review of approaches and future considerations for resource surveys. Environmental DNA 4(1):9–33.

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 & Technology 50(19):10456–10464.

Schmadel, N. M., A. S. Ward, M. J. Kurz, J. H. Fleckenstein, J. P. Zarnetske, D. M. Hannah, T. Blume, M. Vieweg, P. J. Blaen, C. Schmidt, J. L. A. Knapp, M. J. Klaar, P. Romeijn, T. Datry, T. Keller, S. Folegot, A. I. M. Arricibita, and S. Krause. 2016. Stream solute tracer timescales changing with discharge and reach length confound process interpretation: SOLUTE TRACER TIMESCALES CONFOUND PROCESS INTERPRETATION. Water Resources Research 52(4):3227–3245.

Schmelzle, M. C., and A. P. Kinziger. 2016. Using occupancy modelling to compare environmental DNA to traditional field methods for regional-scale monitoring of an endangered aquatic species. Molecular Ecology Resources 16(4):895–908.

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

Shogren, A. J., J. L. Tank, S. P. Egan, O. August, E. J. Rosi, B. R. Hanrahan, M. A. Renshaw, C. A. Gantz, and D. Bolster. 2018. Water Flow and Biofilm Cover Influence Environmental DNA Detection in Recirculating Streams. Environmental Science & Technology 52(15):8530–8537.

Sigsgaard, E. E., M. R. Jensen, I. E. Winkelmann, P. R. Møller, M. M. Hansen, and P. F. Thomsen. 2020. Population‑level inferences from environmental DNA—Current status and future perspectives. Evolutionary Applications 13(2):245–262.

Spence, B. C., D. E. Rundio, N. J. Demetras, and M. Sedoryk. 2021. Efficacy of environmental DNA sampling to detect the occurrence of endangered coho salmon (*Oncorhynchus kisutch*) in Mediterranean‑climate streams of California's central coast. Environmental DNA 3(4):727–744.

Stoeckle, B. C., S. Beggel, R. Kuehn, and J. Geist. 2021. Influence of stream characteristics and population size on downstream transport of freshwater mollusk environmental DNA. Freshwater Science 40(1):191–201.

Sutter, M., and A. P. Kinziger. 2019. Rangewide tidewater goby occupancy survey using environmental DNA. Conservation Genetics 20(3):597–613.

Thalinger, B., D. Kirschner, Y. Pütz, C. Moritz, R. Schwarzenberger, J. Wanzenböck, and M. Traugott. 2020. Lateral and longitudinal fish eDNA distribution in dynamic riverine habitats. Ecology, preprint.

Thalinger, B., A. Rieder, A. Teuffenbach, Y. Pütz, T. Schwerte, J. Wanzenböck, and M. Traugott. 2021. The Effect of Activity, Energy Use, and Species Identity on Environmental DNA Shedding of Freshwater Fish. Frontiers in Ecology and Evolution 9:623718.

Thomsen, P. F., and E. Willerslev. 2015. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183:4–18.

Troth, C. R., M. J. Sweet, J. Nightingale, and A. Burian. 2021. Seasonality, DNA degradation and spatial heterogeneity as drivers of eDNA detection dynamics. Science of The Total Environment 768:144466.

Turner, C. R., M. A. Barnes, C. C. Y. Xu, S. E. Jones, C. L. Jerde, and D. M. Lodge. 2014a. Particle size distribution and optimal capture of aqueous macrobial eDNA. Methods in Ecology and Evolution 5(7):676–684.

Turner, C. R., M. A. Barnes, C. C. Y. Xu, S. E. Jones, C. L. Jerde, and D. M. Lodge. 2014b. Particle size distribution and optimal capture of aqueous macrobial EDNA. Methods in Ecology and Evolution 5(7):676–684.

Van Driessche, C., T. Everts, S. Neyrinck, D. Halfmaerten, A. Haegeman, T. Ruttink, D. Bonte, and R. Brys. 2023. Using environmental DNA metabarcoding to monitor fish communities in small rivers and large brooks: Insights on the spatial scale of information. Environmental Research 228:115857.

Wilcox, T. M., K. S. McKelvey, M. K. Young, W. H. Lowe, and M. K. Schwartz. 2015. Environmental DNA particle size distribution from Brook Trout (Salvelinus fontinalis). Conservation Genetics Resources 7(3):639–641.

Wood, Z. T., A. Lacoursière-Roussel, F. LeBlanc, M. Trudel, M. T. Kinnison, C. Garry McBrine, S. A. Pavey, and N. Gagné. 2021. Spatial Heterogeneity of eDNA Transport Improves Stream Assessment of Threatened Salmon Presence, Abundance, and Location. Frontiers in Ecology and Evolution 9:650717.

Workshop, S. S. 1990. Concepts and Methods for Assessing Solute Dynamics in Stream Ecosystems. Journal of the North American Benthological Society 9(2):95–119.

Yamanaka, H., H. Motozawa, S. Tsuji, R. C. Miyazawa, T. Takahara, and T. Minamoto. 2016. On-site filtration of water samples for environmental DNA analysis to avoid DNA degradation during transportation. Ecological Research 31(6):963–967.

Zhao, B., P. M. van Bodegom, and K. Trimbos. 2021. The particle size distribution of environmental DNA varies with species and degradation. Science of The Total Environment 797:149175.

Appendix A

Foreign eDNA (FeDNA) Suspension Standard Operating Procedures

This Appendix expands on the generation of the FeDNA suspension described in the main text and includes motivation, source considerations, methods, recommendations, and discussion of use and further research. Current methods, live cars, hatchery source water, carcass, etc. used to determine eDNA transport and detection probabilities are subject to natural fluctuations in eDNA concentrations and inconsistent decay and particle size characteristics that may influence inferences drawn from eDNA transport studies. The use of a controlled, consistent, and easily reproducible source of eDNA reduces variability, reduces logistical challenges, and allows for replication across landscapes for consistent assessment of eDNA transport dynamics.

Foreign eDNA Source Considerations:

Sources of FeDNA should be selected by the following criteria:

1. The species (or close relatives) are not present in the system or nearby systems (accounting for predator consumption and defecation of DNA tissues)
2. The source of FeDNA could be processed to produce eDNA in a similar form (including intact cells) comparable to the form in which eDNA is naturally shed.
3. The species is resident in freshwater to avoid the potential for differences in particle (cell) breakdown due to differences in osmotic pressure.
4. The availability of a PCR assay for the species used for the FeDNA source.
5. The ability to source a FeDNA species from a regulatory agency certified hatchery/rearing operation that is disease free.
6. The fish used for a FeDNA source must be euthanized without chemical means that may impact the dynamics of a FeDNA introduction technique and to prevent the introduction of harmful chemicals into streams.

Foreign eDNA Design Considerations:

Multiple considerations went into the design of the FeDNA suspension owing to principles of Mark-Recapture surveys. For the FeDNA suspension to be a valid proxy for NeDNA transport in streams, the FeDNA particles and NeDNA particles must be:

1. transported in the same way.
2. subject to the same capture probability regardless of source or presence of other particles.
3. lost to the system permanently whether through cellular decay or deposition permanently or within the sampling time frame.

FeDNA suspension preparation:

1. Prepare fish for addition to the FeDNA suspension by filleting and skinning the fish. Using only muscle tissues results in a more consistent mixture and eases the homogenization process.
2. Obtain an off-gassed water source that mimics the natural mineral composition of the region.
   a. E.g., well water.
3. The suspension will be made using a concentration of 100 grams of muscle tissue into 1 liter of off-gassed water.
4. Place 100 grams of fish tissue into a Cuisinart blender (or similar) and fill to the max with off-gassed water.
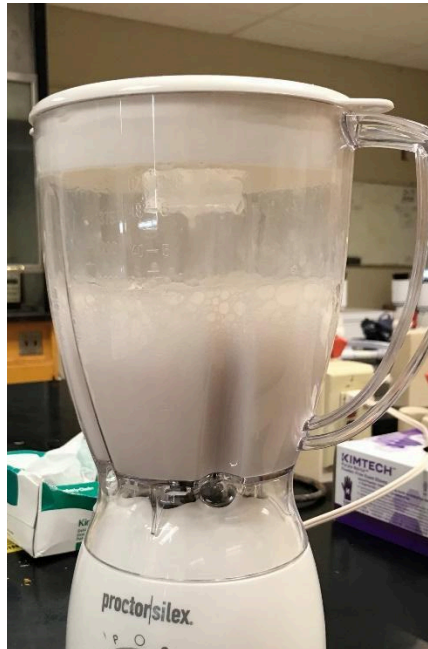
Image B1. A demonstration of the FeDNA suspension in the blender during mechanical homogenization.

   a. If the max of the blender contents is less than 1 full liter, use the proper proportions of fish to off-gassed water. For example, a blender with only 500 ml capacity would necessitate 50 grams of fish and 500 ml off-gassed water.
5. Perform homogenization for roughly 60 seconds or visually consistent. This will provide you with your FeDNA suspension base.
6. Sieve the base suspension through a fine mesh colander to remove the large pieces of fish tissue. This process can be performed multiple times if larger pieces are still present.
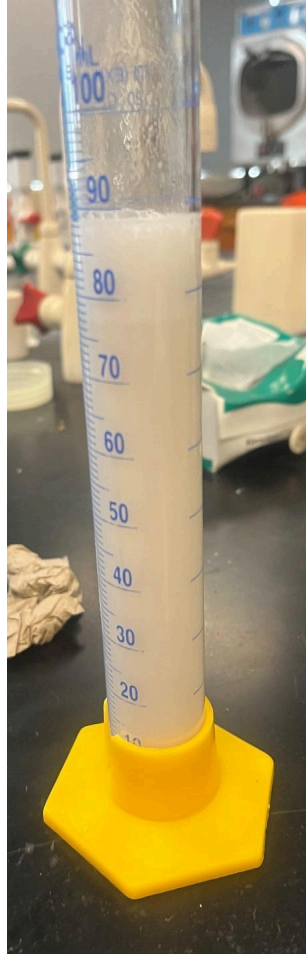
Image B2. An example of the suspension before any filtering.

7. Following removal of large pieces, sieve the suspension 3 times through 5 layers of cheese cloth. This additional step removes larger particles before the finer mesh is used.
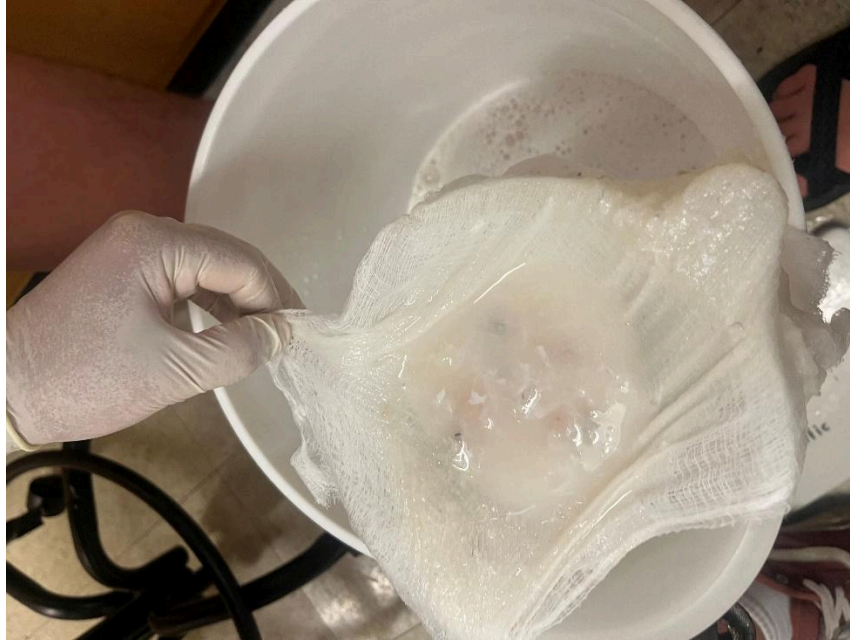
Image B3. An example of the filtering process. Note the larger chunks captured and removed.
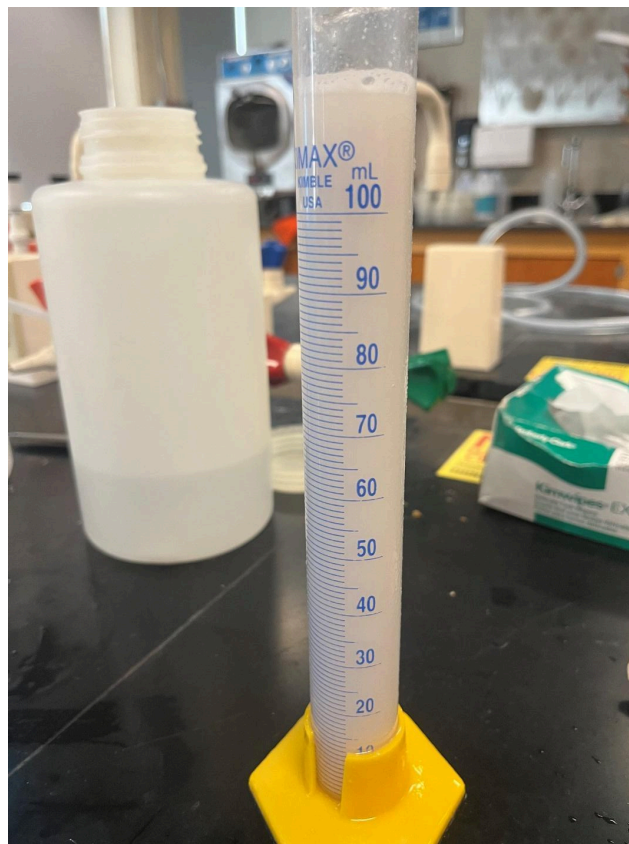

Image B4. An example of the suspension following the first filtering with cheesecloth.

8. Following removal of large chunks, sieve the suspension 1 time through a 200-micron mesh filter. There are paint filters down to very small particle sizes that fit on the top of 5-gallon buckets-the extra surface area makes the process much faster. Following each filtration, clean the 200-micron mesh filter and remove any tissues left on the filter.
9. This further removes particles >200 microns and results in a suspension that will more closely mimic that of NeDNA. (Turner et al. 2014a)

Refer to the text and figure caption for details amount the photograph.
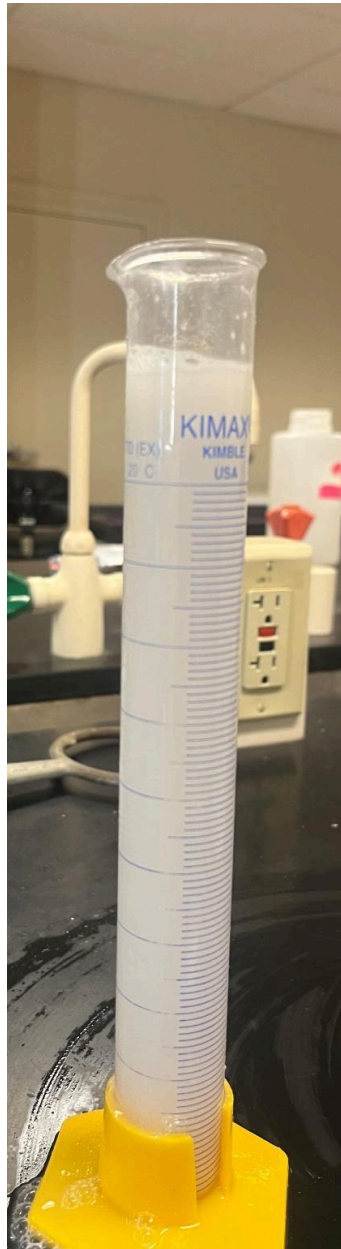
Image B5. An example of the final base suspension after the final sieve through a 200-micron mesh. Note the distinct color change from the first suspension (Image B2) as the particle size was brought to proper distribution.

10. This suspension can now be stored as the base for field applications. If freezing the suspension, re-homogenize using the Cuisinart blender and a single sieve through the 200-micron mesh as the tissue may coagulate during the freezing and thawing process.

11. Before deployment, perform the desired dilution to bring the slurry to the concentration needed for specific applications. Typically, a dilution of 1:2 will bring the suspension to a finer consistency for field deployment and higher dilutions will make the suspension less concentrated if desired. Since the solution is such a high concentration of eDNA, a 1:2 dilution should not drastically affect the concentration of FeDNA.

Figure B1. An example of the ADID and suspension deployed in the field. The suspension is held in a cooler on ice to prevent confounding genetic degradation. An air pump is added to the apparatus to prevent any settling that may occur during the experiment and ensure a homogenous suspension.

Recommendations for Use

It is recommended to dilute the base suspension to make a consistency that can be moved easily through a peristaltic pump. The amount of genetic material (e.g., copies per ml) of tissue may vary among species. In our study, we found carp to have a higher amount of genetic material per weight. This may be due to the composition of cells such as blood, lipid, or water content. If using multiple species, it is recommended to quantify the amount of genetic material per weight of different species.

Discussion and Future Directions

FeDNA suspension derived from different species appear to disperse and deposit in similar manners. Multiple studies need to be conducted to validate the FeDNA suspension as valid proxy for NeDNA. Concurrent studies within this project began this assessment but a full comparison between FeDNA and NeDNA was outside of the scope of this project. Future experiments should include decay, particle size distribution, and deposition velocity. Each of these attributes can vary among NeDNA sources and as such should be assessed concurrently with target species NeDNA and FeDNA used for transport studies.

References

Turner, C. R., M. A. Barnes, C. C. Y. Xu, S. E. Jones, C. L. Jerde, and D. M. Lodge. 2014. Particle size distribution and optimal capture of aqueous macrobial eDNA. Methods in Ecology and Evolution 5(7):676–684.

Autonomous DNA Introduction Device (ADID) Standard Operating Procedures

      This appendix describes the motivation, considerations, materials, methods, and recommendations for building and using the ADID described in the main text. Previous experiments used cost intensive and logistically difficult approaches (e.g., live cars) to assess transport dynamics of eDNA in streams but these methods may not be widely applicable due to limitations such as funding and permitting for using non-native species in streams. Additionally, these methods may not provide clear insight into eDNA transport dynamics in a stream reach due to natural fluctuations in fish shedding rates and low concentrations from naturally shed eDNA. The use of an ADID to introduce eDNA removes confounding factors of fish shedding rates by controlling eDNA concentration at a known rate and concentration. Multiple considerations went into the design of this device. The ADID had to run autonomously for ease of use and potential for multiple remote deployments. The ADID had to be advanced enough to introduce suspension at a known rate and consistent rate. The ADID had to operate for up to 24 hours so that any potential application could be feasible. Constructing the ADID had to be financially feasible and easily replicable for ease of access and use for future collaboration. These considerations guided the following method and standard operating procedures.

      <u>Materials for the ADID: (Description)</u>
- Arduino UNO R3 (A programmable microchip controller, software for Arduino programming available here: https://www.arduino.cc/en/software)
- 12V DIY Peristaltic Pump (Affordable pump for introducing FeDNA with a max introduction rate of 100ml per minute with tubing at 3 mm ID by 5 mm OD)
  - Materials to make a mount for the Peristaltic Pump
    - For ours, we used simple wood and mounting bolts
- 5V Rechargeable battery (Rechargeable battery that can produce up to 5V output, capacity needed depends on length of deployment)
  - The Arduino requires 5V to work, battery must be able to produce at least this much current
  - Can be powered by a generator and 9V input from Arduino if needed
- ADID Housing Unit (Housing unit should be waterproof, ideally clear to assess proper operation without opening the apparatus)
- Rail Mounts (Computer rail mounts ideal for mounting Arduino and battery to housing unit)
- DIN Rail mounting brackets for 35 mm DIN Rails (Battery mounting)
- DIN Rail mount for Arduino
  - This is a specific mounting kit for the Arduino that attaches to the rail mounts
- 3mm inner diameter (ID) by 5 mm outer diameter (OD) pump tubing (lengths needed will depend on an individual's deployment)
- Tubing joint connections 3 mm ID by 5 mm OD

- Breadboard Jumper Wires (serve as a connector for Arduino->Peristaltic Pump)
- NPN Transistor 2N2222
- 1N4001 Diode DO-41 Standard 1A 50V Rectifier Diode

<u>Building Materials/Tools</u>
- Basic electrical soldering kit
- Drill bits
    - Phillips head screwdriver
    - Drill bit that is the same diameter as rail mount screws
    - Drill bit that is the same diameter as tubing
- Silicone sealant
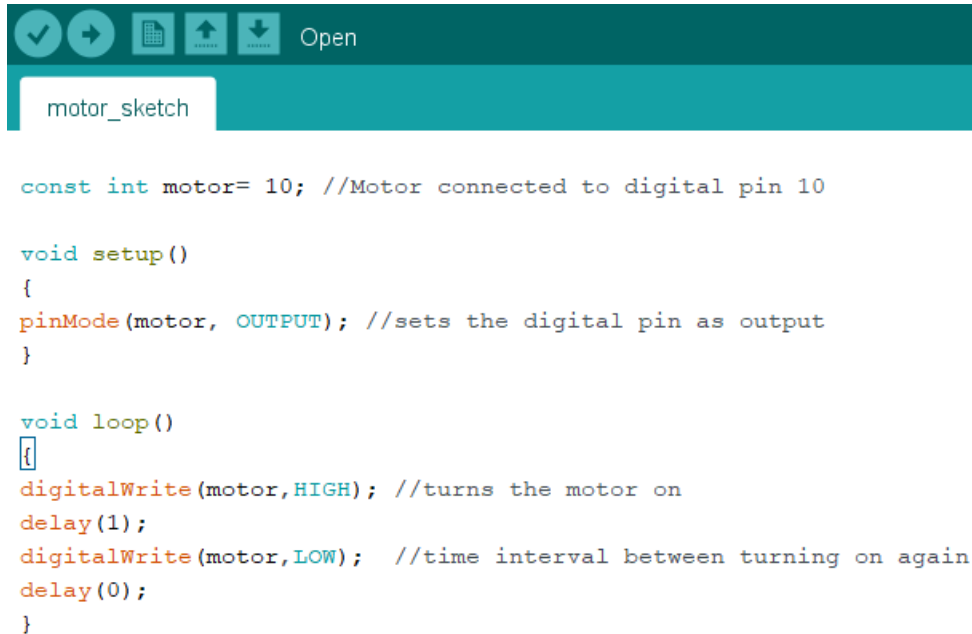- 2 Part quick setting epoxy
- Pliers to cut rail mounts
- Measuring tape/ruler

<u>Production of the ADID:</u>
1. Remove packaging of all products and separate amounts of each part for the number of ADIDs that are to be built.

None of the construction of the individual components (peristaltic pump mounting, electrical soldering, battery mounting, and Arduino mounting) need to be performed in any order. As such, the construction of individual components will be separated into sections to build as you prefer.

<u>Arduino Programming/Mounting</u>
1. Download the Arduino IDE software. (https://www.arduino.cc/en/software)
2. The following basic code (Image A1) is used to power and control the peristaltic:

```
const int motor= 10; //Motor connected to digital pin 10

void setup()
{
pinMode(motor, OUTPUT); //sets the digital pin as output
}

void loop()
{
digitalWrite(motor,HIGH); //turns the motor on
delay(1);
digitalWrite(motor,LOW);  //time interval between turning on again
delay(0);
}
```

Image A1. The code for the Arduino that powers and controls the peristaltic pump.

Code for the Arduino to run the peristaltic pump:
```
const int motor= 10; //Motor connected to digital pin 10
void setup()
{
pinMode(motor, OUTPUT); //sets the digital pin as output
}
void loop()
{
digitalWrite(motor,HIGH); //turns the motor on
delay(1);
digitalWrite(motor,LOW); //time interval between turning on again
delay(0);
}
```

    a. The first line identifies which port is being used on the Arduino. For this application, the pulse width modulator (PWM) number 10 port is used as it can provide very precise concentration outputs.

    b. The void set up indicates that the Arduino should start over and run the following set of code.

    c. The pinMode line sets the PWM as the output for the motor.

    d. A loop function is then generated that will run the Arduino based on electrical input from the PWM port. When the motor is set to HIGH, it will run for the determined amount of time set by the controller. When the motor is set to LOW, it

will not run and will then be delayed for the desired amount of time. A PWM separates a minute into 60000 terms, in other words 5000 would indicate the motor is turned on for 5 seconds, 10000 for 10 seconds, etc. For this code, the motor is turned on and never turned off. If, for example you wanted the motor to run for 30 seconds and then be turned off for 10 seconds, you would set the HIGH delay to 30000, and the LOW delay to 10000.

0. Follow the instructions in the Arduino IDE for downloading the code onto the Arduino. It is very simple and easy to use with a lot of material for trouble shooting online as it is an open-source software.

    a. A common problem that occurs produces an error is that the IDE does not recognize the Arduino is connected. This can be solved easily by making sure the IDE is looking at the correct USB port. This can be accessed by clicking Tools>Port>Choose Arduino port at the top of the IDE.

0. Mount the Arduino to the DIN Rail mount for Arduino.

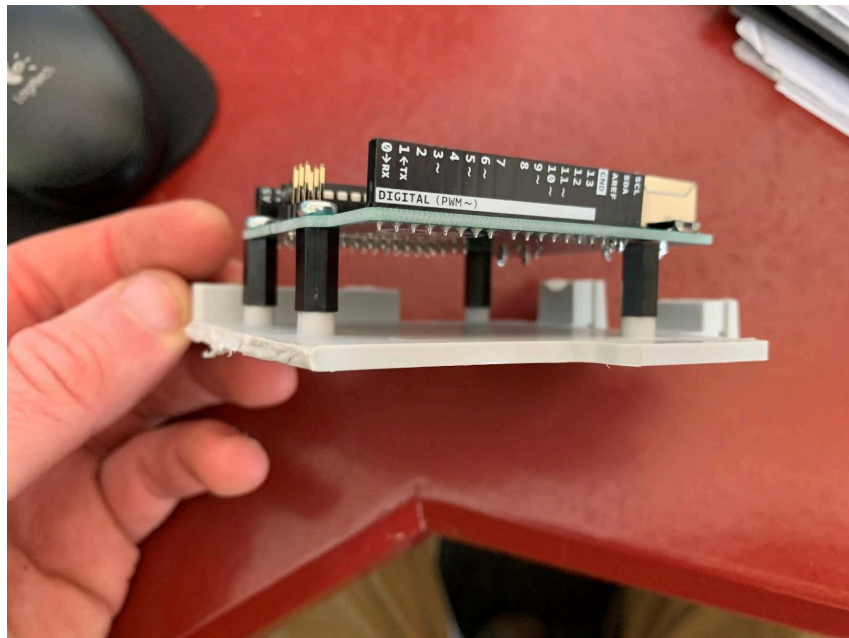   a. Use the long screws to mount the black supports to the DIN Rail mount for Arduino from the bottom.

Image A2. Example of the Arduino mounted using the RIN Rail Mount system.

    b. Once the black supports are secure, attach the Arduino by putting the small screws through the top of the board.

    c. Depending on the model, only 3 of them will insert without obstructing/bending the input boards. Do not try to force them into the board as it could cause structural damage, 3 supports should be rugged enough.
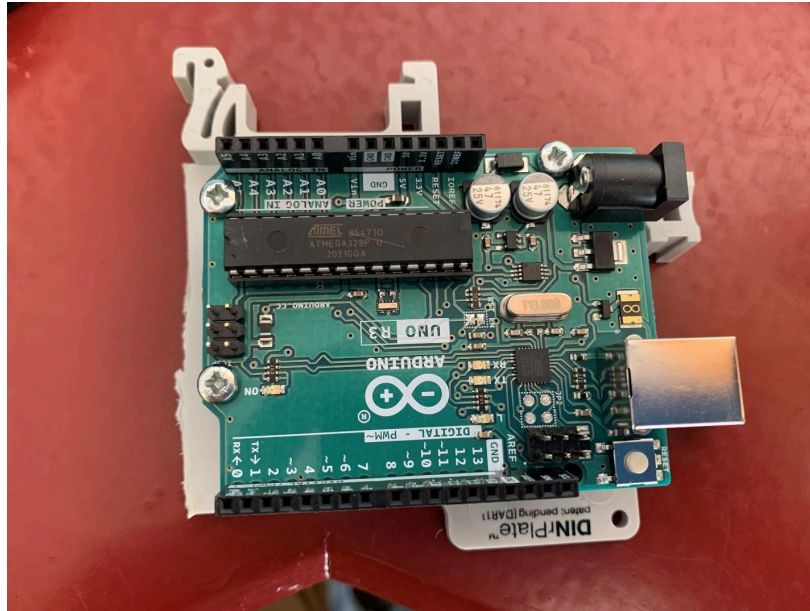
Image A3. Example of the Arduino mounted with 3 screws with this particular model. Some models may be able to accommodate an additional 4th mounting screw.

Peristaltic Pump Mounting and Electrical Soldering

1. Put on proper PPE for using sawing tools including eye and ear protection.
2. Acquire a 1 by 4 piece of lumber, something cheap is fine. Depending on the number of ADIDs, no more than a few feet is required.
3. Using a skill saw or table saw, cut the lumber into 2-inch-wide by 2.5-inch-long pieces.
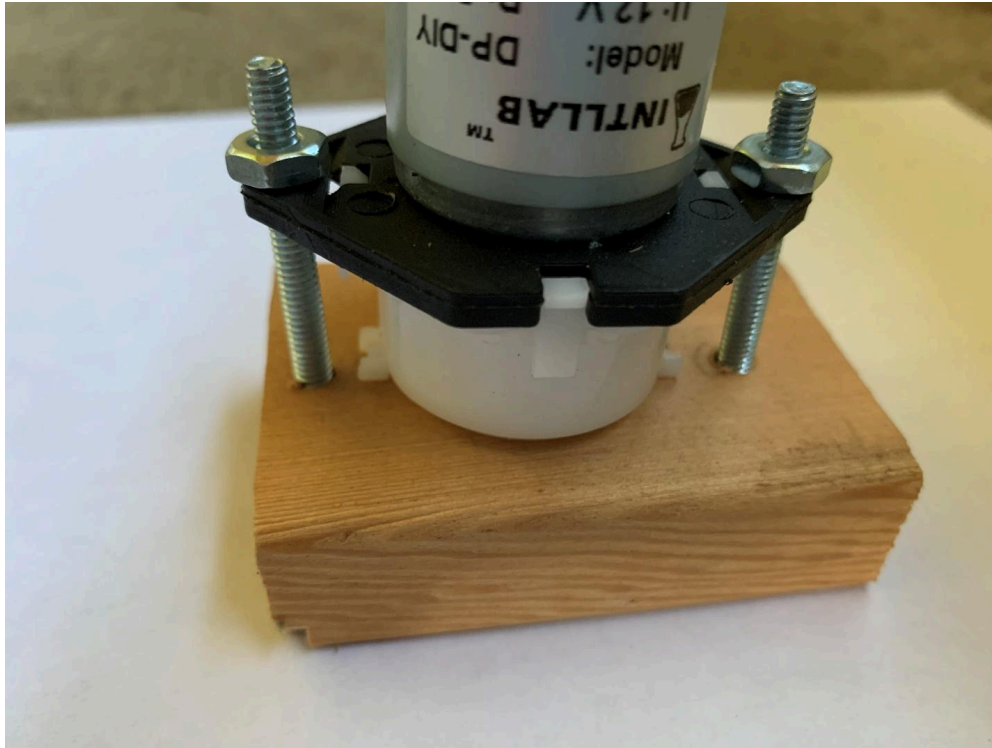4. Drill two holes through them that match the holes in the side mountings of the peristaltic pump.

Image A4. Example of matching the mounting system to the peristaltic pump. Notice the holes
for the pump mount need to match the location of the predrilled holes on the mount itself.

a. An easy way to do this is to use a drill bit that is the same width of the side mountings and marking a starting hole into the wood before drilling them through.

Image A5. Additional angle showing the mounting system on the peristaltic pump.

b. On the back side of the mounting, use a slightly wider drill bit to create a hole for the top of the bolt being used. This allows the bolt to sit snugly within the wood and provides for an even surface to affix to the ADID housing unit.

Image A6. Demonstration of the widened screw mount hole to allow for the mount to sit flatly against the Arduino container.

        c. As an alternative, holes can be drilled into the side of the housing unit and attached to the peristaltic pump that way. Wood was used in this project to reduce the number of holes in the housing unit to keep device watertight.

0.   Purchase bolts that are the same width as the side mountings of the peristaltic pump and are long enough to securely mount the pump.

        a. For this project, the bolts purchased were 2-inch long 8/32-inch round combo machine bolts with nuts

0.   Turn on the soldering kit as it will take a few minutes to warm up.

0.   Soldering fumes are highly toxic, put on proper breathing protection (eg., Respirator) or perform the soldering in a hood.

0.   Collect your materials for soldering the connections to the peristaltic pump.

        a. NPN transistor

        b. 1N4001 Diode DO-41

        c. Breadboard Jumper Wires

           i. It is best to determine colors for each of the specific wires for their various functions. For this project, red was used to connect to the 5V voltage input to the peristaltic pump, green connected the negative side of the pump to

the transistor, black was used for the ground connection, and yellow was used as the transistor switch connection to the PWM input 10 of the Arduino.

0.   For those not experienced in soldering, a great way to ensure good connections is to bend the joints of the diode, jumper wires, or transistors around the other or at opposing 90-degree angles. This makes soldering much easier and ensures good connections.
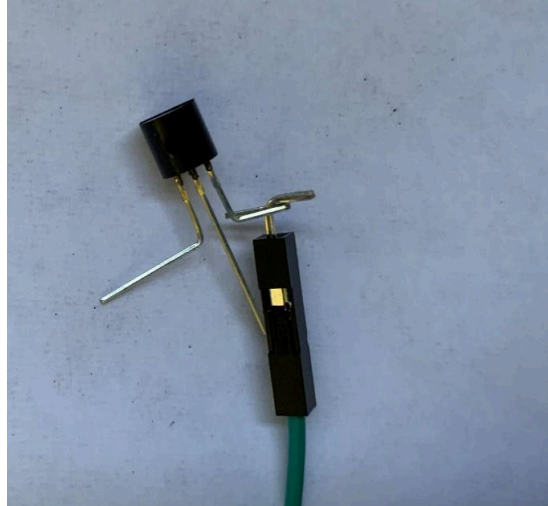

Image A7. Example of bent connections for easier soldering to the NPN transistor.
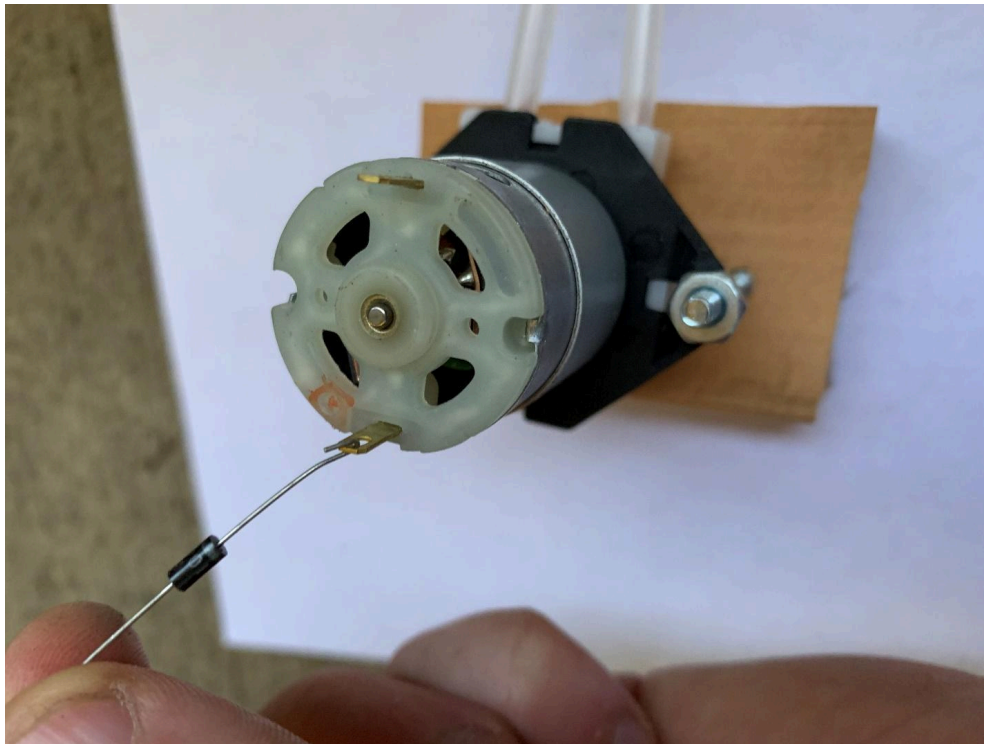

Image A8. Example of bent connections for easier soldering to the peristaltic pump.

0.   Begin by soldering the 1N4001 Diode to the peristaltic pump in parallel reverse bias. Position the cathode at the positive end of the pump, and the anode at the negative end of the pump mountings. The cathode is indicated by the silver line on the round portion of the diode.

       a. The purpose of this is to protect the Arduino microchip controller from back emf that may be produced by the solenoid motor when it turns off. Back emfs can damage microchip controllers and the diode prevents the current from transmitting back into the Arduino.

0.   Solder the red voltage wire to the positive port of the peristaltic pump. The positive port is typically indicated by a red plus sign.

0.   Solder the green transistor connection to the negative port of the peristaltic pump.

0.   The NPN transistor has three prongs. The middle prong is used as the switch for the Arduino and should be kept straight. To make soldering easier, the other two prongs can be bent in a 90-degree angle to connect to the green transistor wire, and the black ground wire. An example is shown below:
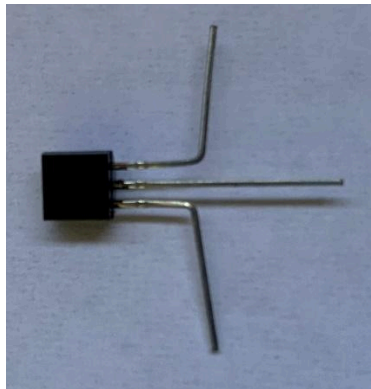

Image A9. Final bent form for the NPN transistor for easier soldering.

       a. The NPN transistor acts as a switch by determining voltage coming in and out of the PWM port 10. Additionally, the transistor amplifies the current enough to power the peristaltic pump.

0.   Solder the green transistor wire to the transistor.

0.   Solder the yellow switch wire to the middle prong of the transistor.

0.   Solder the black ground wire to the remaining transistor prong.

0.   Test the function of the peristaltic pump connections by attaching the red wire to the 5V output on the Arduino, the yellow wire to the 10 port PWD, and the black wire to the ground input labeled GND.

0.   If the pump does not run after the Arduino is turned on, touch the transistor and feel for heat.

       a. WARNING: this can be extremely hot.

0.   If the transistor is hot this indicates that at some point during the construction of the pump, the ports were connected onto the wrong side of the solenoid battery within the pump. A quick fix is to switch the red and green wires on the peristaltic pump ports. In this case, the red voltage wire would be placed on the negative port, and the green transistor wire would be placed on the positive port. This should solve the issue.

0.   Use two-part quick setting JB Weld epoxy to affix the peristaltic pump mountings to the ADID housing unit.

<u>Peristaltic Pump Tubing</u>
1. The peristaltic pumps typically have a pinch point where the tubing exits the peristaltic pump mechanism. This pinch point can cause the tubing to clog and the pump to stop working. Removing the pinch point does not inhibit proper function. You can do this by removing the top portion of the pump by pinching the two clips that hold it on the rest of the pump. You can then remove the additional plastic causing the pinch point.
2. Drill two holes within the ADID housing unit for the inflow of suspension and outflow of suspension at a width of 5 mm wide. Attach tubing joints to the tubing with another piece of 5 mm tubing exiting the housing unit.
   a. Tubing can be added to reach the desired length for field deployment.

<u>Rail installation and Rechargeable Battery Mounting</u>
1. Cut the 6-inch rails into roughly 2-inch and 4-inch pieces.
   1. The 2-inch piece will be for mounting the Arduino, the 4-inch piece for mounting the battery.
2. Using a drill bit that is slightly thinner than the width of the screws provided with the mounting rails, drill holes that will be used to mount the rails.
   a. A slightly smaller diameter is used to assist in waterproofing the screw holes.
   b. An easy way to do this is to place the rail onto the housing unit and mark pilot holes with the drill bit. The configuration below is what you are looking for with the 2-inch piece on the side of the housing unit, and the 4 inch piece on the base of the housing unit.

Image A10. Attachment points of the rail mounts for the Arduino and battery mounts.

0.   Attach the rails with the screws provided.
   a. Be sure to use the smooth end of the cut rails at the top of the Arduino mounts, and the at the middle portion of the battery mount. This will make disassembling much easier.

   b. Before applying the screws, put a small dollop of silicon aquarium sealant onto the screw and hole, this will ensure a waterproof seal where the hole was drilled.

Image A11. Example of the silicon used to waterproof the holes, and the protrusion of the screws outside of the container.

0.   The screws will protrude outside of the housing unit as shown. For this project, a metal grinder was used to smooth out the points on the screws to prevent any unwanted tears in waders, wetsuits, etc. This is optional but recommended.
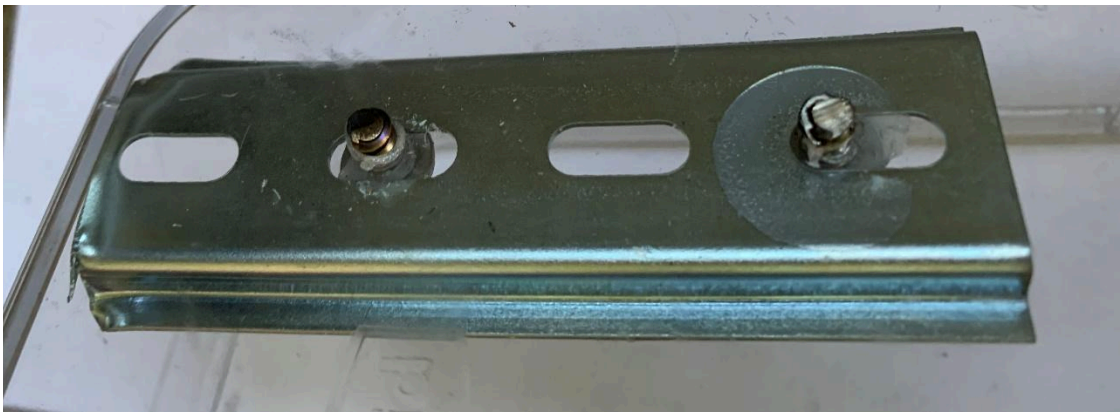

Image A12. Example of the screws shaved down with a metal grinder to remove sharp points.

0.   Attach Command Strips to the green rail mount.
0.   Attach command strips to the rechargeable battery. Depending on the type purchased, be sure to attach them at points where they will not interfere with the mountings for the Arduino and peristaltic pump.

Putting it all Together
1. At this point the mounts for the peristaltic pump, Arduino, and battery source should all be ready for final construction.
   a.   Attach the components to their respective mounts.
2. Insert the wires to the peristaltic pump to their associated Arduino ports.
   a. Black -> ground, Yellow -> 10 port PWM switch, Red 5V port
3. Attach the USB cord from the battery to the Arduino, this should turn the Arduino on and begin the peristaltic pump running.

4. The capacity in the current configuration can transfer a FeDNA suspension at a rate of 25 ml/min.
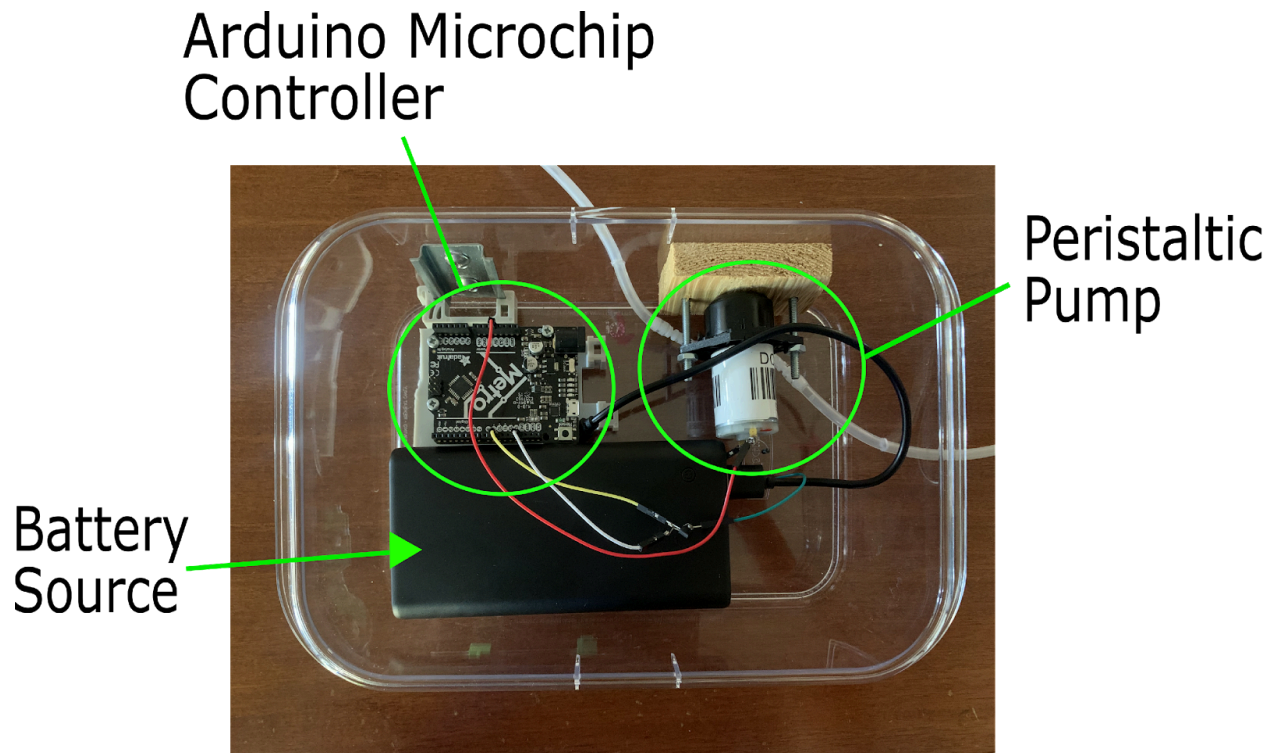

Figure A1. The final configuration of the ADID indicating the Arduino, battery source, and peristaltic pump in a clear waterproof container.

Discussion/Recommendation

The ADID is an acceptable way to introduce FeDNA at a known rate and concentration with some challenges and considerations for use. The streams where this device was used in the described configuration to successfully track FeDNA concentrations ranged in discharge from 0.5 to 5 CFS. When planning FeDNA introduction surveys it is important to scale the use of the ADID's to stream size by either modifying or using multiple ADID's with the same source. It is important to modify the deployment strategy of the ADID dependent on a target species. For example, if calibrating movement of eDNA from a fish species such as juvenile coho salmon the ADID should be deployed in the water column as they reside. If calibrating for a benthic species such as a sculpin or mussels the ADID should be deployed at the bottom of the water column.

After multiple uses the tubing within the peristaltic pump can become worn and cause the rotors on the peristaltic pump to slip. To remove and replace this tubing remove the cap on the peristaltic pump by pinching the two clips holding the top on the peristaltic pump. This should expose the three rotors on the top of the pump and the worn out tubing. This can be replaced with the same tubing used with the other construction of the ADID.

Preliminary Foreign eDNA (FeDNA) trials

This appendix contains descriptions of preliminary field trials using the Autonomous DNA Introduction Device (ADID) to realize instream transport patterns of Natural eDNA (NeDNA) using the introduction of FeDNA. The purpose of these trials was to assess 1) the function of the ADID, 2) concentrations of FeDNA suspension needed to achieve robust signals, and 3) guide sampling and trial design in proof of concept of the method. Each section contains a description of 1) the purpose of the specific trial, 2) methods including introduction rates and sampling design, and 3) impact on final trial design.

<u>Hatchery Trial</u>

A trial in the Cal Poly Humboldt Fish Hatchery was conducted in Summer 2020 to test the function of the ADID and feasibility of collecting, processing, and reliably detecting an introduced FeDNA source generated from fish tissue in water samples using standard protocols.

A consistent laminar flow was generated in a hatchery raceway absent of confounding currents (e.g., eddies) using boards placed across the width of the raceway. FeDNA produced from *Ictalurus punctatus (*IPU) tissues was introduced for 1 hour at a rate of 5 ml/minute at a concentration of 1:4 dilution from base stock. Water samples were taken every 15 minutes for 2 hours 17 m downstream from source due to ease of sampling access in the raceway and water recirculation in the hatchery system.

The results from this trial indicated that the ADID would be a viable method for introducing FeDNA and FeDNA could be reliably detected in water samples using standard protocols. Specifically, FeDNA concentration increases could be tracked during introduction and decreases upon cessation.

<u>Field Trials</u>

Two field trials using the ADID and FeDNA were performed in late summer 2020 to test the method in a natural stream setting. The first trial used the previous introduction rate of FeDNA of 5 ml/minute; this amount of FeDNA was not strong enough to produce a robust signal. For the second trial FeDNA was introduced at a rate of 25 ml/minute for 3 hours. FeDNA suspension produced from IPU tissues was introduced at a concentration of 1:2 dilution from the base suspension.

During introduction, water samples were taken at the bottom of each pool downstream from source for 100 m at hours 1, 2, 3, 4, 6, and 8 to test temporal and spatial changes in FeDNA concentration downstream from source. Water samples were taken in a grid format (e.g., 2 at the top, 3 at the middle, and 2 at the bottom) in the first pool at hour 3 to assess cross stream variations in FeDNA concentrations. All samples were filtered onsite and processed using standard protocols.

The results were interpreted visually. The results from this trial confirmed this method to be viable in natural stream settings. The results yielded other considerations driving trial design including that FeDNA concentrations will vary with 1) habitat structure, 2) distance from source,

and 3) across the width of the stream. The results indicated further trials to characterize a temporal and spatial FeDNA steady state will be needed.

Steady State Trials

A field trial was performed in Spring 2021 to determine the time of introduction and FeDNA concentration needed to reach a temporal and spatial steady state in the study reach. The characterization of a steady state would show an increase in concentration through time that reaches a consistent plateau. The establishment of a consistent steady state accounts for the temporal component of transport in a stream, allowing for more accurate insights into effects of distance, etc. on concentration and distribution patterns of FeDNA. The methods for this trial were derived from traditional solute transport studies that assess transient storage of solutes in streams using constant rate/slug injections and characterization of solute transport and retention in mountain streams.

Three ADIDs introduced three distinct FeDNA suspensions at different concentrations at a rate of 25 ml/min for 12 hours. The three distinct sources were produced from tissues of IPU, *Micropterus salmoides* (MSA), and *Cyprinus carpio* (CCA). The concentrations used were a dilution of 1:2 from base stock, 1:4 from base stock, and 1:10 from base stock. The ADIDs were positioned 25m upstream from the first sample location, designed X=0 to allow for proper cross-sectional mixing and larger particles (greater than 200 microns in length) to settle to the substrate. Water samples were taken at the bottom of the pool roughly 100m from X=0 every hour for 12 hours to assess the temporal development of a steady state. 100m was chosen as this was the extent of snorkel surveys accompanying the final sampling design. At 12 hours, samples were taken at the junction of each habitat unit (e.g., pool and riffle) from 150m to 0m from X=0 to assess the spatial distribution of FeDNA across the steady state. All samples were filtered onsite and processed using standard protocols.

Results from these trials suggested multiple conclusions about the use of FeDNA to determine instream NeDNA transport dynamics. This trial suggested that FeDNA transports in similar patterns as NeDNA given observations of higher and lower concentrations of FeDNA and NeDNA occurring at the same locations (e.g., at the end of pools or riffles). When subject to low-average flows in coastal California streams, a rough FeDNA steady state is achieved in the first 100 m in a few hours.

The stochastic nature of eDNA sampling makes the detailed characterization of a temporal and spatial steady state difficult. This is owed to the fact that a water sample is a mere fraction of the water in a stream and contents of that water is compounded by micro dynamics in streams. Consequently, this nature of eDNA transport in systems would make determining habitat specific effects on a fine scale difficult, likely inconclusive, and require a high density of temporal and spatial sampling.

Lastly, the results from this trial, and the previous trial, suggested that a coarser approach using the ADID to evaluate declines in distance, shared transport patterns, and ideal sampling locations within a study reach may yield more conclusive and relevant results.

Longitudinal Trials

A field trial was conducted to refine eDNA sampling locations. This trial was performed in Freshwater Creek in Northern California (1.5 CFS and 72 square km basin area). Multiple sources of FeDNA were used to cross-calibrate across samples. FeDNA of two species, IPU and MSA, were introduced at a rate of 25 ml/min for 6 hours at the same location. IPU was introduced at a higher concentration than MSA to assess the effect of concentration on detection probability downstream.

A sampling design was generated to realize declines in distance of FeDNA transport and detection probabilities. After 6 hours of FeDNA introduction, samples were taken moving up the stream at 1000m, 800m, 600m, 400m, 300m, 200m, 100m, and 50m from the FeDNA introduction location. Samples were taken in triplicate at cross-sectional sampling locations at right, center, and left facing downstream.

The results from this trial demonstrated a decline in FeDNA concentration with distance consistent with the literature. Samples taken in a cross-section format demonstrated variability across a stream transect. The results from this trial informed the following trial and the final trial design.

A final trial was performed before finalizing the trial design. This trial was performed in Jacoby Creek in Northern California (0.25 CFS, 29 square km). Three sources of FeDNA were introduced at different locations and concentrations to assess the effect of source location on downstream transport and detection. Specifically, the movement of eDNA across the channel and plume effects of eDNA depending on source location.

Multiple sources of FeDNA also enabled cross validation of the model between sources. IPU was introduced 100m upstream from CCA and MSA in the main flow to simulate the influence of eDNA sources upstream from the study reach. MSA was introduced at 100m in the main flow to simulate a group of fishes in the main flow of the stream. CCA was introduced at 100m in a side channel/slack water to simulate a group of fishes in this portion of the stream. CCA was introduced at half the concentration from the other two species.

An intensive sampling design was implemented to determine a robust but efficient sampling scheme (e.g., were there some sampling distances we could remove but still achieve the same quantitative results?). After 6 hours, samples were taken moving up the stream at 1000m, 750m, 500m, 400m, 300m, 200m, 150m, 125m, 100m, and 50m, from the FeDNA introduction site. Samples were taken in triplicate at cross-sectional sampling locations at right, center, and left when facing downstream.

This trial indicated some transects could be removed and still be sufficient. This trial suggested that mixing across the stream becomes roughly uniform (e.g., does not demonstrate strong trends based on source location) after deployment, likely depending on site characteristics. The final trial design followed these characteristics with some slight modifications. The other two sources of FeDNA were introduced just above the 50m sampling location to increase the amount of data collected from a single survey. Samples at 125, 400, and 750 were removed to decrease time for a survey and allocate more samples to additional field trials. See main text for final trial design.

Detailed Assay Information

This appendix contains the detailed assay information for those used in this study including base pair sequences, probe attributes, etc.

Table D1. Detailed assay information for the *Cyprinus carpio* assay used in this study (Eichmiller et al. 2014)

Metric	Value
Make	Integrated DNA Technologies
Mt DNA Locus	Cytochrome B
Base Pairs	149
Probe Attributes	250 nm PrimeTime 5' 6-HEX/ZEN/3' Iowa Black FQ, 18 bases
Probe Sequence	5'-/5HEX/CCCTCTAGT/ZEN/TACACCACC/3IABkFQ/-3'
Forward Primer Sequence	5'-CTAGCACTATTCTCCCCTAACTTAC -3'
Reverse Primer Sequence	5'-ACACCTCCGAGTTTGTTTGGA-3'
LOD	2.55 Copies Per Reaction
LOQ	15 Copies Per Reaction

Table D2. Detailed assay information for the *Micropterus salmoides* assay used in this study (Yamanaka et al. 2016).

Metric	Value
Make	Integrated DNA Technologies
Mt DNA Locus	Cytochrome B
Base Pairs	unreported
Probe Attributes	250 nm PrimeTime 5' 6-FAM/ZEN/3' Iowa Black FQ, 30 bases
Probe Sequence	5'-/56-FAM/CTAACGGTG/ZEN/CATCCTTCTTTTTCATCTGCA/3IABkFQ/-3'
Forward Primer Sequence	5'-GCCCACATTTGTCGTGATGTAA-3'
Reverse Primer Sequence	5'-AGCCCCGGCCGATATG-3'
LOD	2.89 Copies Per Reaction
LOQ	13 Copies Per Reaction

Table D3. Detailed assay information for the *Ictalurus punctatus* assay used in this study
(unpublished).

Metric	Value
Make	Integrated DNA Technologies
Mt DNA Locus	Cytochrome B
Base Pairs	140
Probe Attributes	250 nm PrimeTime 5' 6-FAM/ZEN/3' Iowa Black FQ, 18 bases
Probe Sequence	5'- / 56-FAM/TGGTTCCTC/ZEN/TCTAATCTA/3IABkFQ/-3'
Forward Primer Sequence	5'-TTAGCCCGCGGAATACAAATC-3'
Reverse Primer Sequence	5'-GGACCAATTAAACAGTGCGGTG-3'
LOD	2.34 Copies Per Reaction
LOQ	6 Copies Per Reaction

References

Eichmiller, J. J., P. G. Bajer, and P. W. Sorensen. 2014. The Relationship between the Distribution of Common Carp and Their Environmental DNA in a Small Lake. PLoS ONE 9(11):e112611.

Yamanaka, H., H. Motozawa, S. Tsuji, R. C. Miyazawa, T. Takahara, and T. Minamoto. 2016. On-site filtration of water samples for environmental DNA analysis to avoid DNA degradation during transportation. Ecological Research 31(6):963–967.

Appendix E

Attrition and Detection Probability Patterns

To evaluate the hypothesis that FeDNA concentration will decline with distance, a linear regression using the function 'stats' in the 'lm' package in R was fitted to log transformed concentrations. A statistical test for equivalence of slope was used to test the assumption that FeDNA signals subject to the same processes should decline at similar rates when there is no additional input of FeDNA. The regressions for attrition patterns are analyzed using one-way ANOVA and post hoc comparisons of slope using the function 'emmeans' in package 'emtrends' in R with a 0.95 confidence level.

To evaluate the hypothesis that detection probability will decline over distance, generalized linear models using the function 'stats' in the package 'glm' in R with the family "binomial" were fitted to data from each trial to estimate the effect of downstream distance on

detection probability. Samples in which measured concentrations of FeDNA exceed or fall below FeDNA-specific LOD are designated as "positive" (value 1) or "negative" (value 0), respectively.

Patterns of attrition and detection probability over distance could be estimated only for the three trials (Jacoby Creek, Little River, Old Campbell Creek) that exhibited a monotonic decline in [FeDNA] over distance.

Jacoby Creek

All observed FeDNA concentrations declined monotonically with distance downstream, regardless of how the data were screened prior to analysis (Figure E1, Table E1). Attrition patterns between FeDNA sources are not significantly different when fit with the full or correlation-screened data set.

All observed FeDNA detection probabilities declined monotonically with distance downstream (Figure E2).

Little River

All observed FeDNA concentrations declined monotonically with distance downstream, regardless of how the data were screened prior to analysis (Figure E1, Table E1). Attrition patterns between CCA and IPU are not significantly different when fit with correlation-screened or unscreened data. Attrition patterns between MSA and other FeDNAs were not significantly different when using the correlation-based screen but differed significantly using the unscreened data.

All observed FeDNA Detection probabilities declined monotonically with distance downstream (Figure E2).

Old Campbell Creek

All observed FeDNA concentrations declined slightly with distance downstream, dependent on the data set used (Figure E1, Table E1). Attrition patterns between FeDNA sources in Old Campbell Creek are not significantly different when fit with correlation-screened or unscreened data.

FeDNA detection probabilities exhibited mildly different patterns across sources (Figure E2).Detection Probability of MSA and CCA decline coherently but marginally with distance downstream. Detection probability of IPU exhibited a general decline with distance downstream.

Prairie Creek Low Flow

Observed FeDNA concentrations did not decline monotonically, regardless of how the data were screened prior to the analysis (Figure E1, Table E1). Attrition patterns of FeDNA for IPU and MSA increased slightly with distance (Figure E1, Table E1). Attrition patterns between FeDNA sources are not significantly different when fit with correlation-screened or unscreened data.

Detection probabilities of FeDNA exhibited different patterns across sources (Figure E2). Detection probability of CCA decreased marginally while IPU and MSA increased with distance downstream.

Prairie Creek High Flow

Observed FeDNA concentrations increased slightly, regardless of how the data were screened prior to the analysis (Figure E1, Table E1). Attrition patterns between FeDNA sources are not significantly different when fit with correlation-screened or unscreened data.

Detection probabilities of FeDNA exhibited different patterns across sources (Figure E2). Detection probability of CCA increased with distance downstream, while IPU increased marginally. MSA decreased marginally with distance downstream.
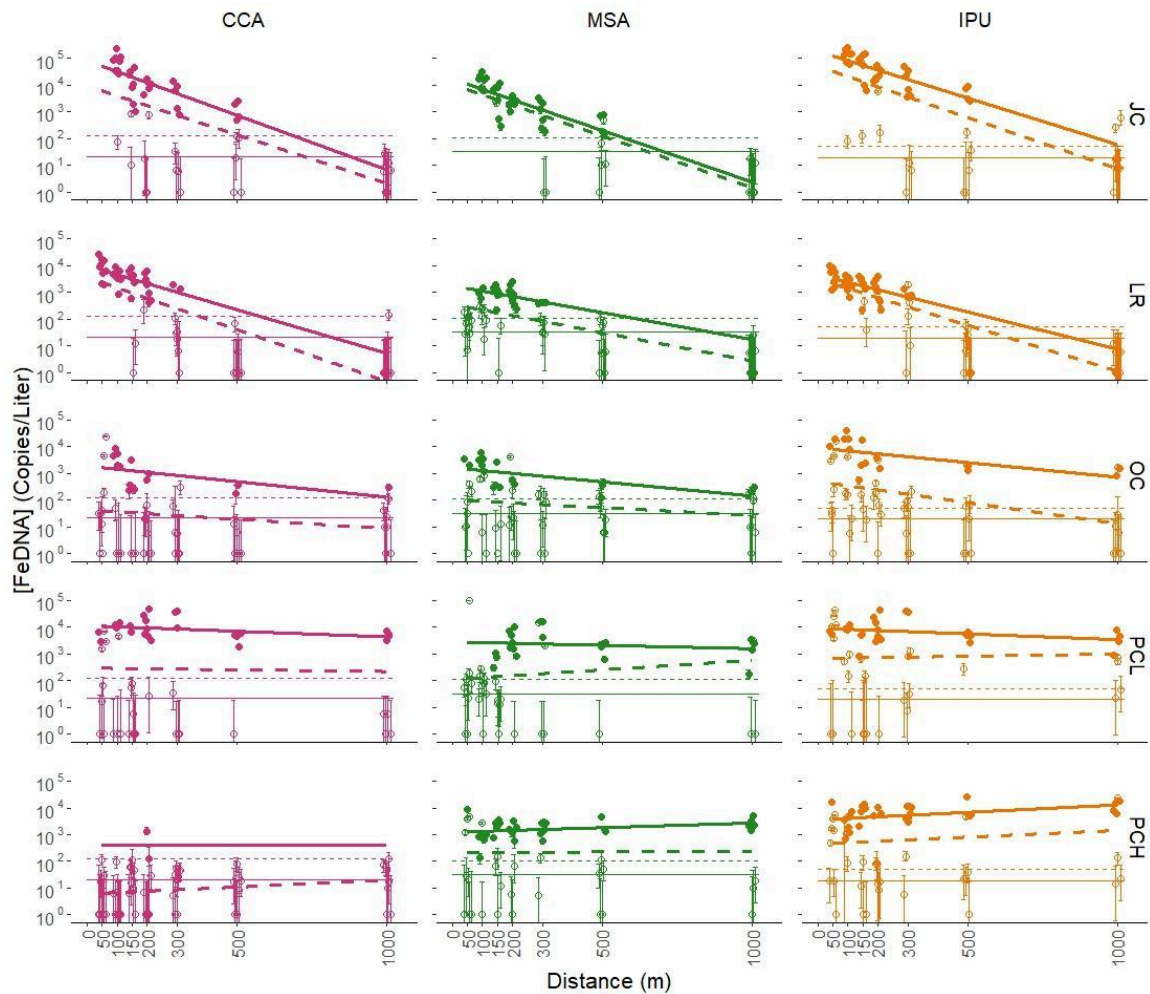
Figure E1. FeDNA attrition patterns of IPU (right column; orange), MSA (center column; green) and CCA (left column; pink) as a function of distance across trials (top to bottom row: JC, LR, OC, PCL, PCH) based on correlation-screened and unscreened data. Solid points (attrition pattern as bolded solid lines) indicate values retained from the correlation-based screen; hollow circles (Attrition pattern as bolded dashed lines) indicate values excluded from the correlation-based screen. Note log-scaled y-axis. Error bars for each point are derived from Poisson distributions and are obscured at larger values due to log-scaling. A jitter is applied to the points for clarity. Horizontal lines indicate FeDNA-specific LOD (solid line) and LOQ (dashed line). Attrition rates and 95% confidence intervals are included in Table 6.

Table F1. Attrition patterns (slopes) for each FeDNA source in each trial when fit with correlation-screened and unscreened data. 95 percent confidence interval is included in parentheses.

| Site	Data	CCA Attrition (95 CI)	MSA Attrition (95 CI)	IPU Attrition (95 CI)
Jacoby Creek	Correlation	-0.00408 (0.00108)	-0.00386 (0.00094)	-0.00349 (0.00076)
Jacoby Creek	Full	-0.00365 (0.00116)	-0.00387 (0.0007)	-0.00384 (0.00112)
Little River	Correlation	-0.00327 (0.00174)	-0.00202 (0.0007)	-0.00278 (0.0008)
Little River	Full	-0.00391 (0.00084)	-0.00212 (0.00062)	-0.00344 (0.00062)
Old Campbell Creek	Correlation	-0.00116 (0.00106)	-0.00106 (0.0008)	-0.00109 (0.00148)
Old Campbell Creek	Full	-0.00069 (0.00106)	-0.00057 (0.00092)	-0.00158 (0.00108)
Prairie Creek Low Flow	Correlation	-0.00042 (0.00092)	-0.000254 (0.00106)	-0.00043 (0.00086)
Prairie Creek Low Flow	Full	-0.00013 (0.00152)	0.00072 (0.00118)	0.00016 (0.00126)
Prairie Creek High Flow	Correlation	N/A	0.00034 (0.00072)	0.00056 (0.00088)
Prairie Creek High Flow	Full	0.00051 (0.00068)	0.00004 (0.0011)	0.0005 (0.0011)
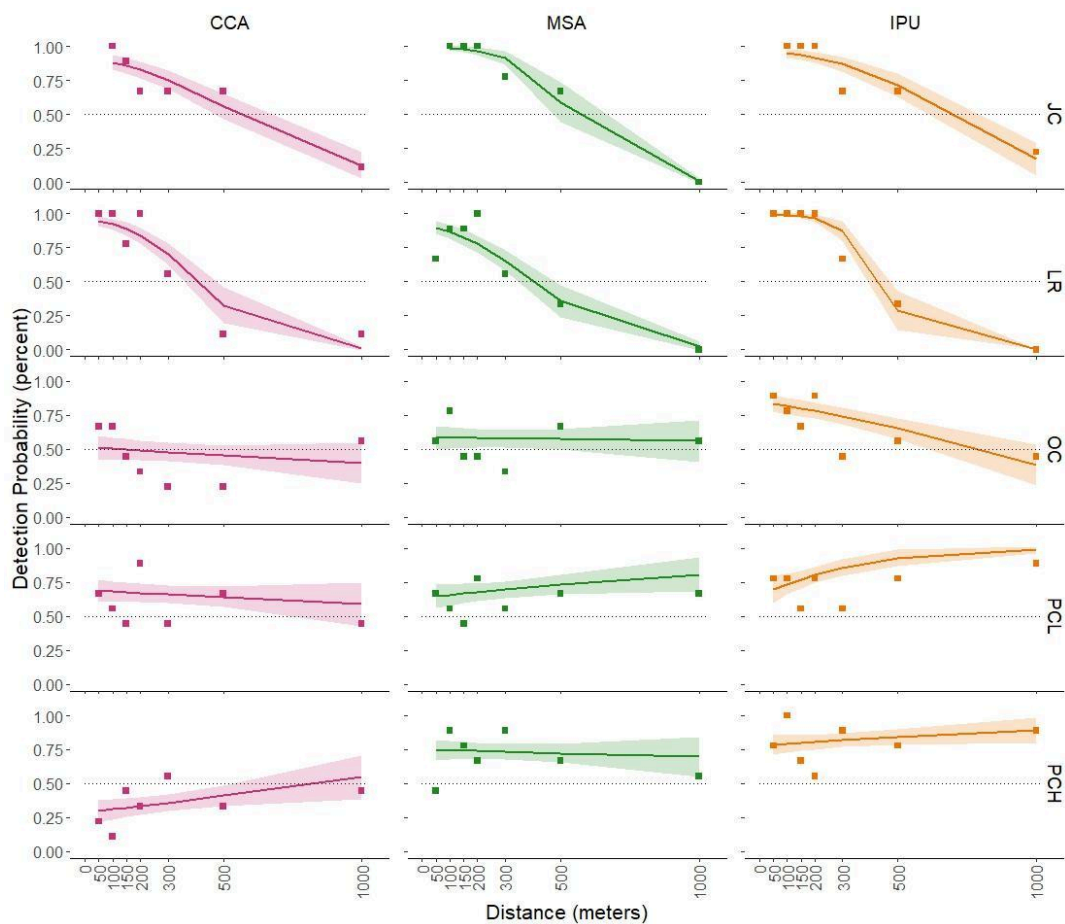
Figure E2. FeDNA detection probability patterns of IPU (right column; orange), MSA (center column; green) and CCA (left column; pink) as a function of distance across trials (top to bottom row: JC, LR, OC, PCL, PCH). Detection probability is plotted as the proportion of samples above the LOD. Points indicate percent of samples above the LOD at that cross-section. Lines and ribbon indicate the results and confidence intervals of the binomial logistic regression. 50 percent detection probability (dotted black line; lower panel) is included for reference.

Appendix F

Supplemental Results Figures and Tables

      This appendix contains supplementary results that demonstrate the comparison of the correlation-based screen to traditional mechanical thresholds used for screening data. Machined based thresholding provides an intermediate of screening between the stringent correlation-based screen and unscreened data. However, these threshold screening still

violate the assumptions of variance for comparison with ANOVA and as such could not be used to examine differences between the ratios at a cross-section. For patterns regarding WMR comparisons refer to the main text. Generally, there were not significant differences between attrition patterns. The only instance where this is not true is MSA in the Little River trials.
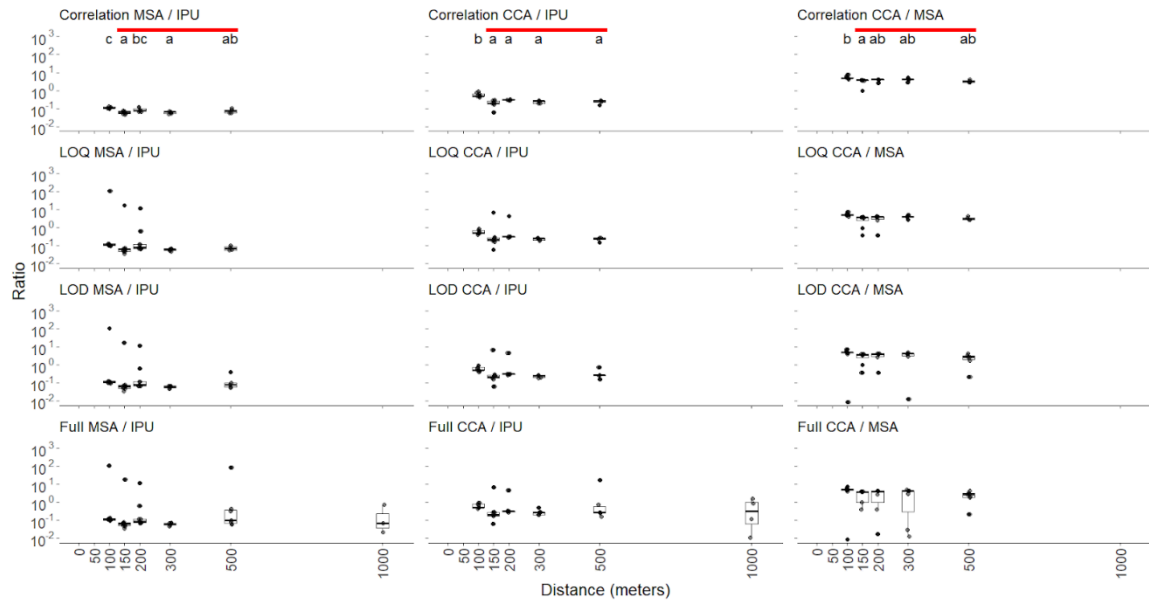
Ratios

Figure F1. Average FeDNA ratios (CCA:MSA, right column, CCA:IPU, center; MSA:IPU, left) as a function of distance across screening methods (top to bottom row: Correlation, LOQ, LOD, Unscreened) in Jacoby Creek. Results from the Tukey test to determine significant difference between ratios is indicated by letters at the top of the correlation-based screening plot. Cross-sections containing the same letter are not statistically different from the other. The Well-Mixed-Region (WMR) is indicated by the red line.
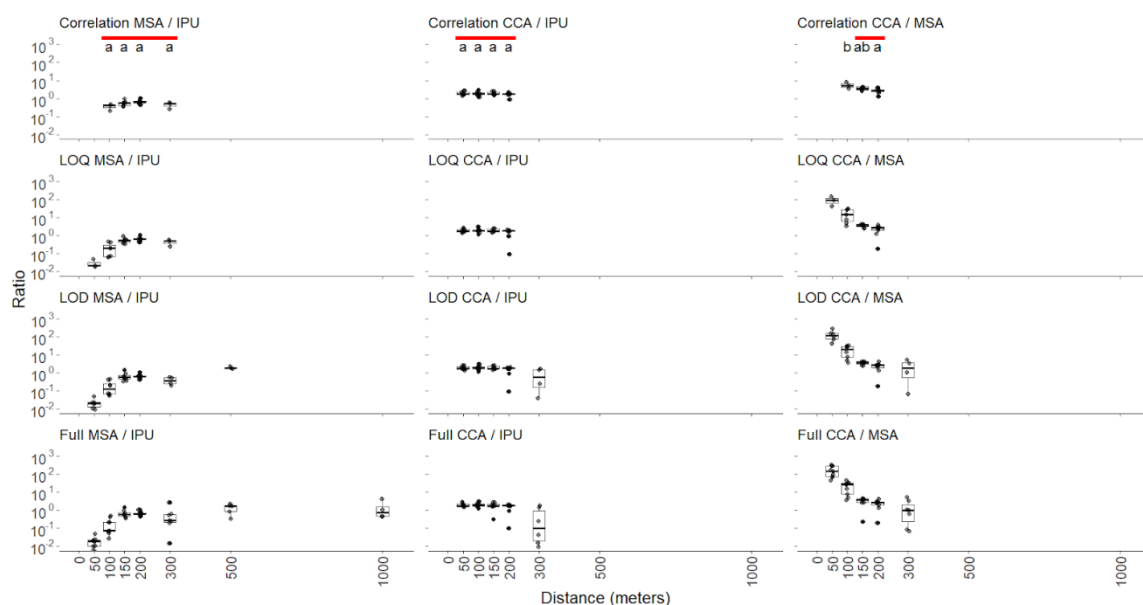
Figure F2. Average FeDNA ratios (CCA:MSA, right column, CCA:IPU, center; MSA:IPU, left) as a function of distance across screening methods (top to bottom row: Correlation, LOQ, LOD, Unscreened) in Little River. Results from the Tukey test to determine significant difference between ratios is indicated by letters at the top of the correlation-based screening plot. Cross-sections containing the same letter are not statistically different from the other. The Well-Mixed-Region (WMR) is indicated by the red line.
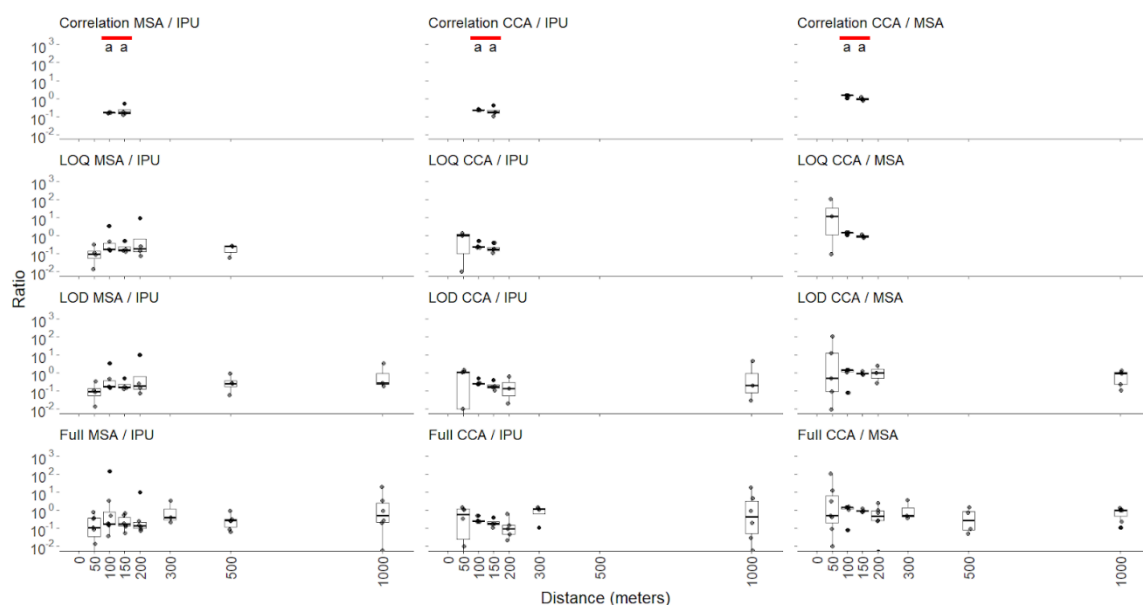
Figure F3. Average FeDNA ratios (CCA:MSA, right column, CCA:IPU, center; MSA:IPU, left) as a function of distance across screening methods (top to bottom row: Correlation, LOQ, LOD, Unscreened) in Old Campbell Creek. Results from the Tukey test to determine significant difference between ratios is indicated by letters at the top of the correlation-based screening plot. Cross-sections containing the same letter are not statistically different from the other. The Well-Mixed-Region (WMR) is indicated by the red line.
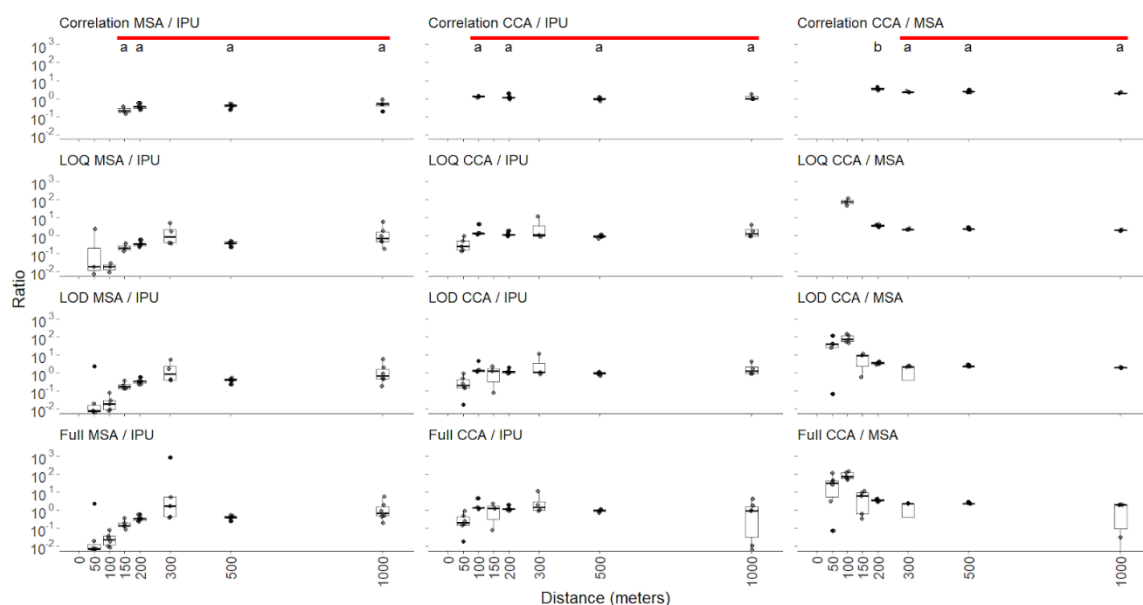
Figure F4. Average FeDNA ratios (CCA:MSA, right column, CCA:IPU, center; MSA:IPU, left) as a function of distance across screening methods (top to bottom row: Correlation, LOQ, LOD, Unscreened) in Prairie Creek Low Flow. Results from the Tukey test to determine significant difference between ratios is indicated by letters at the top of the correlation-based screening plot. Cross-sections containing the same letter are not statistically different from the other. The Well-Mixed-Region (WMR) is indicated by the red line.
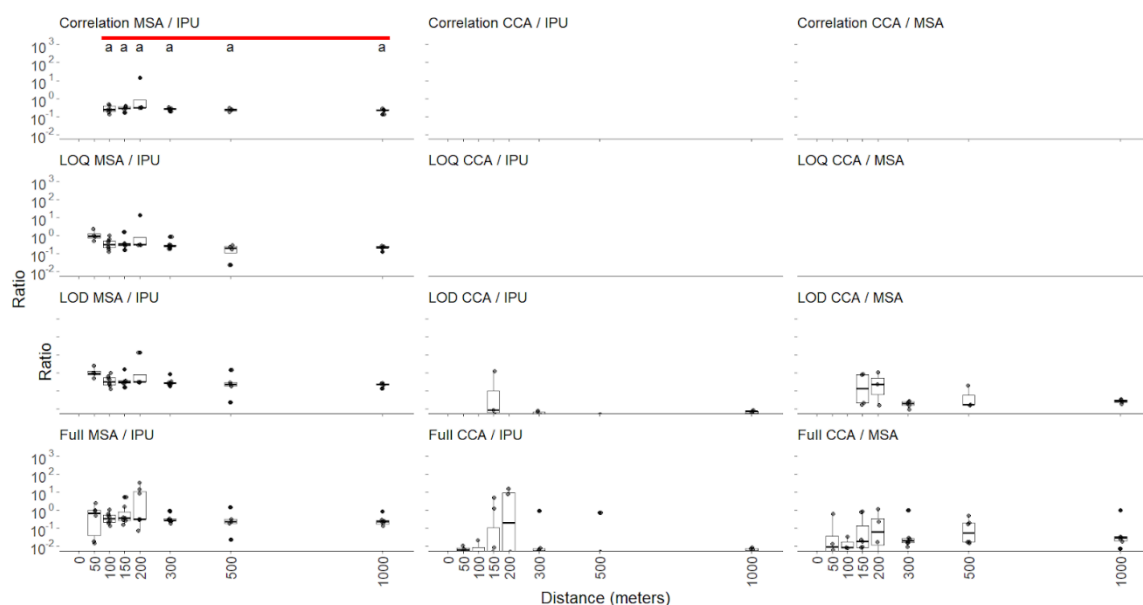
Figure F5. Average FeDNA ratios (CCA:MSA, right column, CCA:IPU, center; MSA:IPU, left) as a function of distance across screening methods (top to bottom row: Correlation, LOQ, LOD, Unscreened) in Prairie Creek High Flow. Results from the Tukey test to determine significant difference between ratios is indicated by letters at the top of the correlation-based screening plot. Cross-sections containing the same letter are not statistically different from the other. The Well-Mixed-Region (WMR) is indicated by the red line.
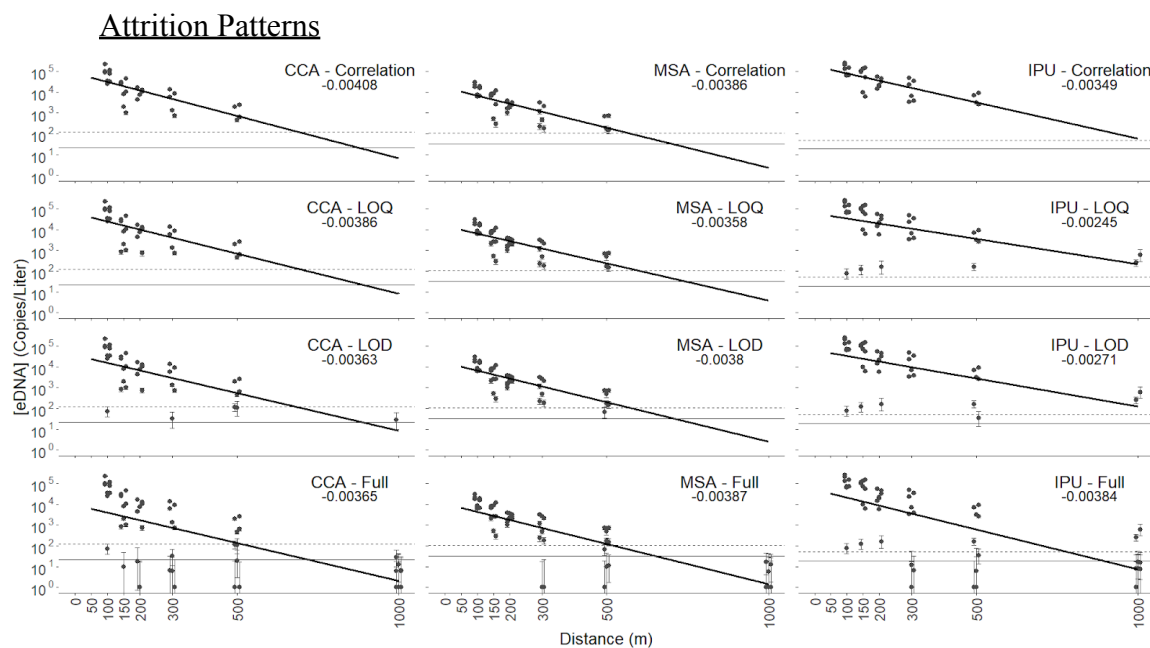
Attrition Patterns

Figure F6. FeDNA attrition patterns of IPU (right column), MSA (center column) and
CCA (left column; pink) as a function of distance across trials (top to bottom row:
Correlation, LOQ, LOD, Unscreened) in Jacoby Creek. Individual concentrations
are plotted as copies/liter on a log transformed y-axis. Attrition patterns are the
fitted line overlaying the concentrations. Attrition patterns, species, and screening
methods are included in the top right corner of each plot. Attrition patterns and
95% confidence intervals from this site are included in Table E1. Error from the
poisson estimate of concentration for each point are included. Some error bars are
hard to discern at higher concentrations due to the log-transformed axis. The
horizontal solid line indicates the species specific LOD and the horizontal dashed
line indicates the species specific LOQ.

Table F1. Attrition rates (slopes) for each FeDNA source in Jacoby Creek using different
methods for data screening.

Data	CCA Attrition	CCA Attrition 95 % CI	MSA Attrition	MSA Attrition 95 % CI	IPU Attrition	IPU Attrition 95 % CI
Correlation	-0.00408	0.00108	-0.00386	0.00094	-0.00349	0.00076

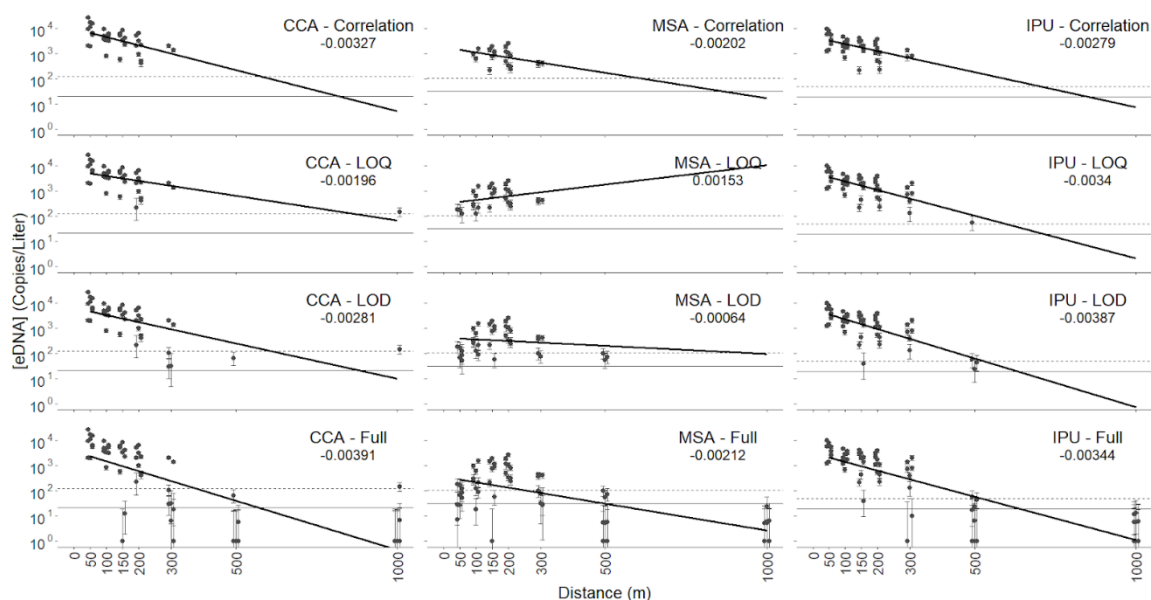| Data	CCA Attrition	CCA Attrition 95 % CI	MSA Attrition	MSA Attrition 95 % CI	IPU Attrition	IPU Attrition 95 % CI
LOQ	-0.00386	0.00166	-0.00358	0.0011	-0.00244	0.00126
LOD	-0.00362	0.00142	-0.00379	0.00102	-0.00271	0.0013
Full	-0.00365	0.00116	-0.00387	0.0007	-0.00384	0.00112

Figure F7. FeDNA attrition patterns of IPU (right column), MSA (center column) and CCA (left column; pink) as a function of distance across trials (top to bottom row: Correlation, LOQ, LOD, Unscreened) in Little River. Individual concentrations are plotted as copies/liter on a log transformed y-axis. Attrition patterns are the fitted line overlaying the concentrations. Attrition patterns, species, and screening methods are included in the top right corner of each plot. Attrition patterns and 95% confidence intervals from this site are included in Table E2. Error from the poisson estimate of concentration for each point are included. Some error bars are hard to discern at higher concentrations due to the log-transformed axis. The horizontal solid line indicates the species specific LOD and the horizontal dashed line indicates the species specific LOQ.

144

Table F2. Attrition rates (slopes) for each FeDNA source in Little River using different methods for data screening.

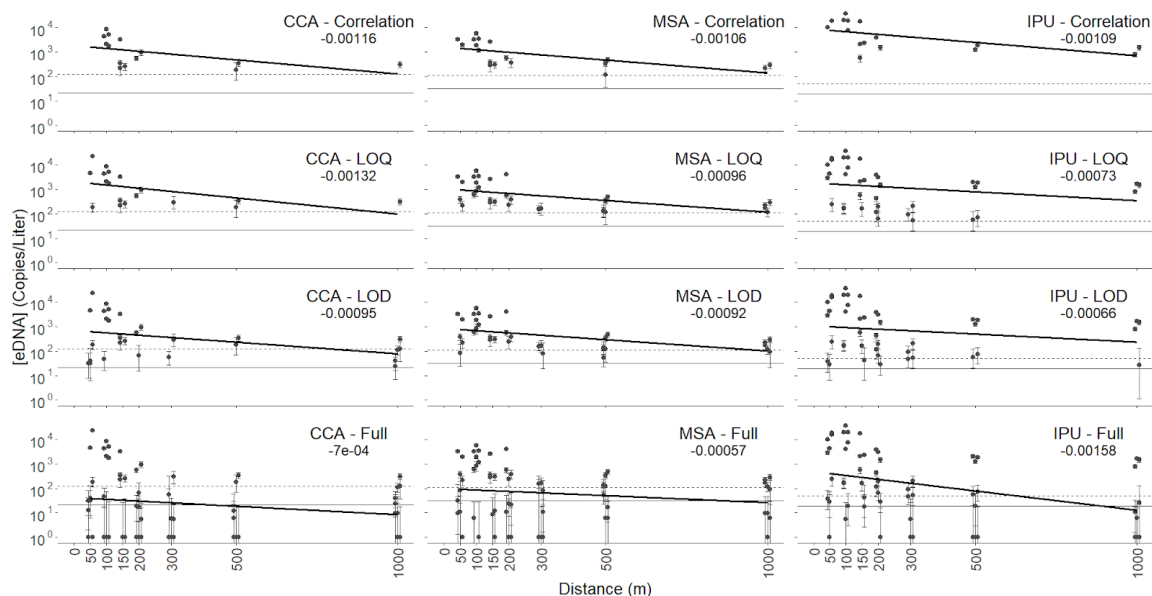| Data	CCA Attrition	CCA Attrition 95 % CI	MSA Attrition	MSA Attrition 95 % CI	IPU Attrition	IPU Attrition 95 % CI
Correlation	-0.00327	0.00174	0.00202	0.0007	-0.00278	0.0008
LOQ	-0.00197	0.00088	0.00153	0.00204	-0.0034	0.00116
LOD	-0.00281	0.0011	-0.00064	0.00138	-0.00387	0.0011
Full	-0.00391	0.00084	-0.00212	0.00062	-0.00344	0.00062

Figure F8. FeDNA attrition patterns of IPU (right column), MSA (center column) and CCA (left column; pink) as a function of distance across trials (top to bottom row: Correlation, LOQ, LOD, Unscreened) in Old Campbell Creek. Individual concentrations are plotted as copies/liter on a log transformed y-axis. Attrition patterns are the fitted line overlaying the concentrations. Attrition patterns, species, and screening methods are included in the top right corner of each plot. Attrition patterns and 95% confidence intervals from this site are included in Table E3. Error from the poisson estimate of concentration for each point are included. Some error bars are hard to discern at higher concentrations due to the

log-transformed axis. The horizontal solid line indicates the species specific LOD, and the horizontal dashed line indicates the species specific LOQ.

Table F3. Attrition rates (slopes) for each FeDNA source in Old Campbell Creek using different methods for data screening.

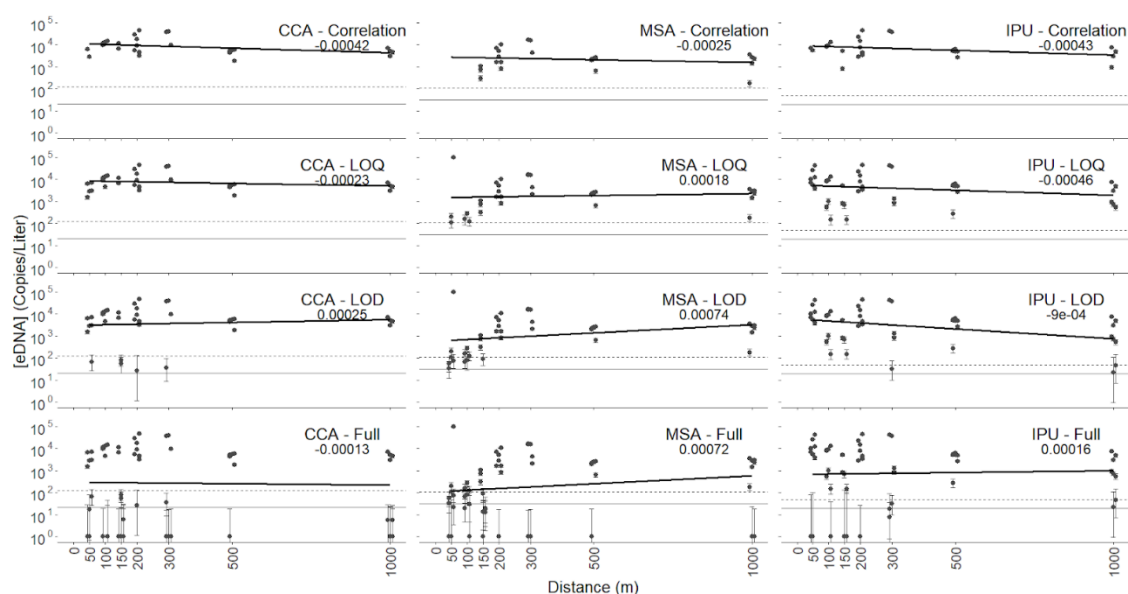| Data	CCA Attrition	CCA Attrition 95 % CI	MSA Attrition	MSA Attrition 95 % CI	IPU Attrition	IPU Attrition 95 % CI
Correlation	-0.00116	0.00106	-0.00106	0.0008	-0.00109	0.00148
LOQ	-0.00132	0.00122	-0.00096	0.00054	-0.00073	0.00108
LOD	-0.00095	0.00088	-0.00092	0.00054	-0.00066	0.00108
Full	-0.00069	0.00106	-0.00057	0.00092	-0.00158	0.00108

Figure F9. FeDNA attrition patterns of IPU (right column), MSA (center column) and CCA (left column; pink) as a function of distance across trials (top to bottom row: Correlation, LOQ, LOD, Unscreened) in Prairie Creek Low Flow. Individual concentrations are plotted as copies/liter on a log transformed y-axis. Attrition patterns are the fitted line overlaying the concentrations. Attrition patterns, species, and screening methods are included in the top right corner of each plot. Attrition patterns and 95% confidence intervals from this site are included in

Table E4. Error from the poisson estimate of concentration for each point are included. Some error bars are hard to discern at higher concentrations due to the log-transformed axis. The horizontal solid line indicates the species specific LOD, and the horizontal dashed line indicates the species specific LOQ.

Table F4. Attrition rates (slopes) for each FeDNA source in Prairie Creek Low Flow using different methods for data screening.

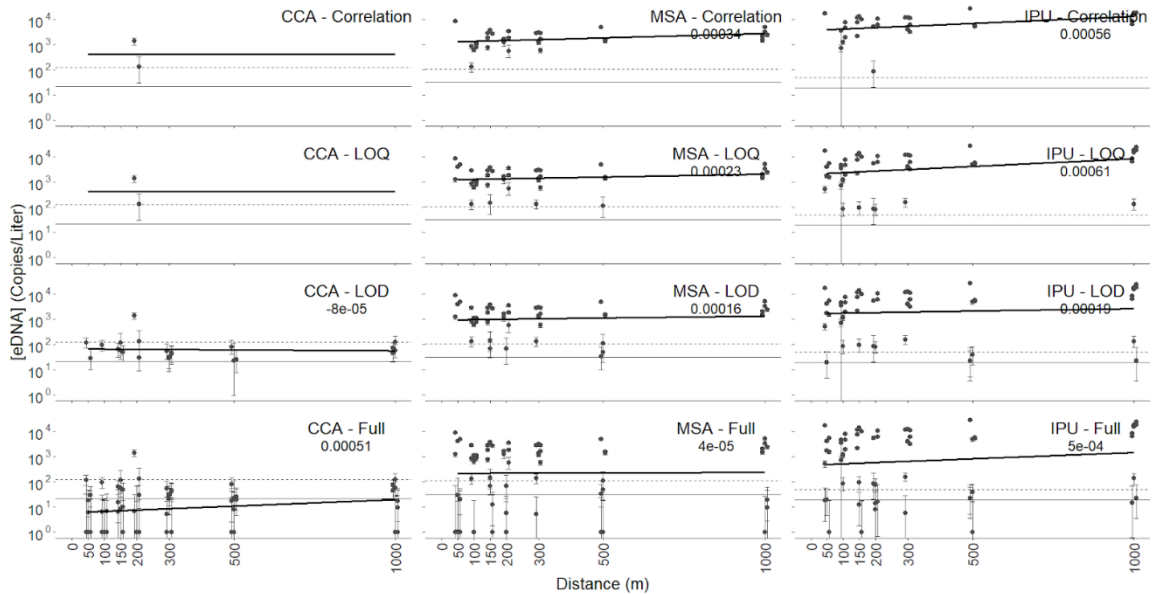| Data	CCA Attrition	CCA Attrition 95 % CI	MSA Attrition	MSA Attrition 95 % CI	IPU Attrition	IPU Attrition 95 % CI
Correlation	-0.00042	0.00092	-0.000254	0.00106	-0.00043	0.00086
LOQ	-0.00023	0.00044	0.00018	0.00076	-0.00046	0.00062
LOD	0.00025	0.00098	0.00074	0.00082	-0.0009	0.00068
Full	-0.00013	0.00152	0.00072	0.00118	0.00016	0.00126

Figure F10. FeDNA attrition patterns of IPU (right column), MSA (center column) and CCA (left column; pink) as a function of distance across trials (top to bottom row: Correlation, LOQ, LOD, Unscreened) in Prairie Creek High Flow. Individual concentrations are plotted as copies/liter on a log transformed y-axis. Attrition patterns are the fitted line overlaying the concentrations. Attrition patterns, species, and screening methods are included in the top right corner of each plot.

Attrition patterns and 95% confidence intervals from this site are included in Table E5. Error from the poisson estimate of concentration for each point are included. Some error bars are hard to discern at higher concentrations due to the log-transformed axis. The horizontal solid line indicates the species specific LOD, and the horizontal dashed line indicates the species specific LOQ.

Table F5. Attrition rates (slopes) for each FeDNA source in Prairie Creek High Flow using different methods for data screening.

Data	CCA Attrition	CCA Attrition 95 % CI	MSA Attrition	MSA Attrition 95 % CI	IPU Attrition	IPU Attrition 95 % CI
Correlation	N/A	N/A	0.00034	0.00072	0.00056	0.00088
LOQ	N/A	N/A	0.00023	0.00048	0.0061	0.00064
LOD	-0.00008	0.00046	0.00016	0.00064	0.00019	0.00082
Full	0.00051	0.00068	0.00004	0.0011	0.0005	0.0011

TASK 3: USE EDNA TO RAPIDLY AND NON-INVASIVELY DETERMINE THE LOCAL-SCALE OCCUPANCY OF COHO SALMON

Using an exogenous DNA tracer to inform native environmental DNA concentrations in small streams

Gavin B. Bandy 1
Braden A. Herman 1
Andre Buchheister 1
Eric P. Bjorkstedt 2
Andrew P. Kinziger 1

1 Department of Fisheries Biology, Cal Poly Humboldt, Arcata, CA 95521, USA
2 Southwest Fisheries Science Center, NOAA Fisheries, Trinidad, CA 95570, USA

Abstract

This study assessed the performance of an exogenous environmental DNA (eDNA) tracer particle in characterizing the distribution patterns of naturally occurring DNA (NeDNA) to gain insight into the local-scale occupancy of target species. Occupancy assessments using eDNA methods were performed at 13 study reaches across 6 small coastal streams under varying conditions across two years. Simultaneous snorkel surveys were conducted to assess the accuracy of eDNA-based occupancy assessments. Results from this study found each stream reach has unique eDNA distributions patterns, emphasizing the need for a calibration tool to accurately assess site occupancy and account for variability arising from sampling processes. The distance between sample locations was identified as a crucial factor influencing the success of occupancy assessments, with increasing distance between sample locations yielding greater levels of detection across all sites. Discharge and the abundance of target species were also found to impact occupancy assessments, requiring consideration of study systems and focal species when conducting eDNA surveys. These findings offer valuable insights into dynamics of eDNA over small spatial scales and provide guidance for future environmental DNA studies.

Introduction

Species monitoring using environmental DNA (eDNA) in water samples is an innovative approach providing high species detection probabilities and causing minimal disruption to the habitat and organisms within it. Utilizing eDNA methods for aquatic species monitoring involves the collection of DNA-bearing particles (mucus, tissue, excrement) released by organisms into their environment (Eichmiller et al. 2014). Collected DNA is then extracted and amplified using targeted and non-targeted methods to estimate the quantity of DNA in a sample (Eichmiller et al. 2014). To date, many studies have implemented this technique in aquatic systems to determine species distribution over broad geographic scales without the need for costly and physically intensive field survey techniques (e.g., Sutter and Kinziger 2019; Ostberg et al. 2019; Duda et al. 2020).

While Environmental DNA has been effectively used to monitor aquatic species in rivers (Wood et al. 2021, Hallet et al. 2012, Everets et al. 2022, Shaffer 2022), the transport of eDNA is still not fully understood (Ostberg and Chase 2022, Thalinger et al. 2021, Harrison 2019). Generally, concentrations of eDNA decline as they are transported away from their source (Spence et al. 2021; Thalinger et al. 2020), however many interacting local factors influence the transport of eDNA rendering the dynamics system-specific. Hydrodynamic processes greatly affect the transport of DNA through mixing, settling, and resuspension processes (Van Driessche 2022, Wood et al. 2012, Spence et al. 2021, Jo and Minamoto 2021) while degradation of eDNA is dependent upon abiotic factors (temperature and UV exposure) and biotic factors (microbial activity) (Snyder et al. 2023, Caza- Allard et al. 2021, Harrison 2019). Crucially, eDNA dynamics are closely tied to the specific characteristics of the river under study, exhibiting distinct variations across geographic locations but also throughout different times (Van Driessche et al. 2022, Troth et al. 2021). To have a comprehensive understanding of eDNA dynamics for a given system would require significant investment, likely exceeding the capabilities of monitoring initiatives and offsetting the advantages of eDNA methods. Developing a system to quickly assess how transport processes influence natural eDNA (NeDNA) signals would be invaluable for the future of eDNA surveys.

Herman (2023) developed an Autonomous eDNA Introduction Device (ADID) that releases an exogenous eDNA tracer, enabling the system-specific study of eDNA transport in small streams. By releasing the tracer from a point source and recapturing it at multiple downstream locations, the ADID facilitates the study of eDNA transport at a specific location. Examining the changes in concentration of the tracer as it is moved away from its source allows for assessment of the vertical and lateral dispersion of eDNA in a stream. Additionally, the eDNA tracer released from the ADID can serve as a comprehensive positive control covering the entirety of the workflow, offering a verification of procedural accuracy starting from the collection of samples through to the implementation of detection techniques like qPCR, ddPCR, or eDNA metabarcoding.

This research used ADIDs and tracers to calibrate eDNA transport within 1000m stream reaches. The tracer patterns were used to inform NeDNA signals with the goal of determining local-scale occupancy of fishes within a focal reach. The tracer is used to characterize how eDNA entering a study reach is influenced by hydrological processes, enabling one to discern whether changes in concentrations result from these processes or via enrichment from a source organism. If changes result from the enrichment of eDNA, it indicates the presence of a target species within a study reach. In this research, the ADID was used for controlled release of exogenous eDNA from a fixed upstream point, and set to run for several hours to establish equilibrium conditions. Multiple sampling locations were designated at varying distances from the release point, allowing for concurrent assessments of NeDNA from a target species and tracer eDNA. This deliberate release of tracer eDNA served as a well-defined system-specific reference point for interpreting native eDNA concentrations.

To assess how the target NeDNA concentration is changing relative to the system-specific transport revealed by the tracer, one can calculate the ratio of the target species' eDNA concentration to the concentration of the tracer at all sampling locations within a study reach. The exogenous DNA tracer is expected to decline with distance, serving as a baseline for how eDNA is transported through a stream reach without input of additional eDNA. Comparing NeDNA concentrations to the tracer using ratios allows one to examine if the two sources of eDNA exhibit similar or divergent patterns, indicating whether or not there is additional input of target eDNA. A constant ratio between the two sources of eDNA across a study area indicates NeNDA is declining proportionally and similarly with the tracer.Constant ratios along the study reach suggest that additional target eDNA is not being introduced and therefore the target species is absent from the reach (Figure 1; Scenario 1 & 2). This results from either; (1) the Target eDNA and tracer declining in proportional to one another across a study reach or (2) the target eDNA remaining undetected across a reach. Conversely, if ratio values increase throughout a reach it indicates that the detected target eDNA exceeds the amount predicted by the tracer either from concentrations of target eDNA surpassing what is expected from the tracer or the first detection of target eDNA occuring within the study reach (Figure 1; Scenario 3 & 4). Both of these cases suggest the presence of the target species within the study reach.

This study was designed to ascertain the local occupancy of a target species over a relatively short distance. The tracer is used to address the potentially misleading influence of target eDNA being transported from upstream of the reach into the focal area and thus affecting conclusions of occupancy. By observing the transport of a tracer, introduced from a point source, I determined if target DNA exhibits similar or divergent patterns. Ultimately, this allows for distinguishing among four occupancy cases regarding species' presence both upstream and within the study reach (Table 1). This approach was developed specifically for scenarios that demand accurate local-scale occupancy data, such as when monitoring impacts to endangered species.

I employed the ADID system to perform a comparative analysis, evaluating the ratio of NeDNA and tracer to assess the dynamics of eDNA, and evaluate presence/absence of coho salmon (*Oncorhynchus kisutch*) and steelhead head (*Oncorhynchus mykiss*) within a 1000-meter stream reach. I used this approach at 13 different stream reaches with varying ecological conditions to ascertain conditions under which the DNA tracer declined with distance and could effectively calibrate NeDNA signals. Snorkel surveys and habitat mapping were conducted at each study reach to evaluate the presence and relative number of coho salmon and steelhead as well as to characterize the environmental conditions at each location.

Methods

Autonomous eDNA Introduction Device (ADID)

Exogenous eDNA tracers were based on two species of freshwater fish that do not occur in the study area, *Ictalurus punctatus* (channel catfish) and *Cyprinus carpio* (common carp), obtained from The Fishery (Galt, California). To produce the tracer, 100 grams of skinned filets were homogenized with 1 liter of reverse osmosis water in a standard household blender. This mixture was passed three times through five-layer cheesecloth, followed by a single pass through a 200-micron (μm) mesh filter that limited the maximum particle size of the tracer to that typical of natural eDNA (Turner et al. 2014). The tracer was generated no more than 24 hours prior to use and stored on wet ice for transport into the field.

To introduce the tracer into a stream, the ADID was set to release tracer at a constant rate. The ADID consists of a programmable Arduino microchip controller, peristaltic pump, and rechargeable lithium battery all housed in a waterproof container. The exogenous eDNA tracer is placed in a chilled container to minimize degradation of DNA and an air pump induces mixing and suspension of particles. The peristaltic pump transferred the tracer from its container into the stream at a rate of 0.025 liters per minute. A full description of the exogenous eDNA tracer and ADID can be found in Herman (2023).

<u>Study Area</u>

This study was conducted in small streams located in the Redwood forests of Humboldt County, California. The study area is in the Coastal California Ecoregion and is characterized by mild, wet winters and cool dry summers, with maximum summer temperature rarely exceeding 25°C. All study streams were home to the threatened Southern Oregon/Northern California Coast coho salmon Evolutionarily Significant Unit (NOAA Fisheries 2023). Systems were sampled under summer low flow conditions when distinct pool and riffle habitats are present.

Study Design

<u>2021 Study Design</u>

In 2021, surveys were conducted on five stream reaches, each spanning 1000 meters. In each reach, two tracers (channel catfish - IPU, common carp - CCA) were introduced as discrete single point sources. The use of multiple tracers allowed for comparison in transport dynamics between sources. The IPU tracer was deployed at the head of the reach in main flow at a 1:1 dilution of tracer to stream water (~1.5 million copies/ml of tracer). The CCA tracer was deployed at 50 meters downstream from the head of the study reach at a 1:2 dilution of tracer to stream water (~500,000 copies/ml of tracer) (Figure 2).

The tracer was allowed to disperse downstream 6 hours to establish a steady state distribution prior to water sample collection (Figure 2 ;Herman 2023). Water samples were then collected beginning at the 1000 m section and proceeding upstream to prevent contamination from resuspension of settled eDNA. Samples were collected at distances of 0 m, 50 m, 100 m, 150 m, 200 m, 300 m, 500 m, and 1 km downstream from the head of the reach. At each cross section, triplicate water samples were collected at evenly spaced locations, specifically at 25 (left), 50 (center), and 75% (right) of the stream's width, for a total of 9 water samples per cross

section. The equidistant placement of sampling locations did not account for flow patterns or stream habitat. Sample collection across all cross sections took 30 - 60 min depending on reach characteristics and access. In total, 63 water samples were collected at each study reach.

At each cross section, stream width (meters) was recorded and at every sample point, water depth (meters) and water velocity (meters per second) were recorded. For each study reach, one discharge measurement was collected. This was done by dividing a cross-section into 10 equally spaced points where depth (m) and velocity (m/s) were measured. Section discharge was then calculated by multiplying the average velocity by the area of a sample point (determined from the depth and the distance between measurement points). Site discharge was then calculated by summing all sectional discharge calculations. A full description of discharge measurements can be found in Hauer and Lamberti (2006).

2022 Study Design

In 2022, surveys were conducted on 8 stream reaches, following an approach similar to that used in 2021 with the following exceptions. Each reach was 500 m in length, and both ADID systems were placed at the head of the reach (0 m) in the main flow (Figure 2). The IPU tracer was introduced at a 1:3 dilution of tracer to stream water (~375,000 copies/ml), and CCA at a 1:7 dilution (~187,500 copies /ml). Water samples were collected at cross sections 100 m, 200 m, 300 m, 400 m and 500 m downstream from the head of the reach.

Snorkel observations

Snorkel surveys were conducted to provide a real-time assessment of coho salmon presence, distribution, and abundance in each study reach. Snorkel surveys took place the day prior to eDNA sampling or immediately following water collection. If snorkel surveys took place the day prior to eDNA sampling, it was assumed that reach occupancy did not change prior to water collection for eDNA analysis. In 2021, a single 100-meter section between the 100-200 cross sections was surveyed by two divers independently. In 2022, the entire study area encompassing 500 meters of stream was surveyed with a single pass. Divers moved upstream from the lower extent of the snorkeled reach to the top, recording the abundance of coho salmon and steelhead (*Oncorhynchus mykiss)*.

Environmental DNA Methods

Water Collection and Filtration

For each sample, 1.75 liters of water was collected by pulling a sterile Whirl-Pak bag along the surface of the stream. Samples were immediately stored on wet ice in a cooler and filtered within 8 hours in a dedicated water filtration laboratory.

Water was filtered under vacuum over 47 mm diameter 0.45 μm cellulose-nitrate-filters placed upon 47 mm diameter filter support pads, and inserted into sterilized plastic filter funnels. Up to 6 samples were filtered at a time using a filter manifold connected to a pneumatic hand pump. The volume of water filtered for each sample was recorded. Filters bearing DNA were

removed from the filter funnels and placed into sterile microcentrifuge tubes using sterilized tweezers. Filters from samples collected in 2021 were stabilized in 360 µl of cell lysis buffer and stored at -20°C due to the long duration of time between collection and extraction. Samples collected in 2022 were frozen without cell lysis buffer due to a short time period between collection and extraction All filter cups, countertops and materials used in filtering were sterilized with a 10% bleach solution and rinsed with reverse osmosis water prior use.

To test for contamination, one field-blank was collected at every cross section. Field blanks consisted of 1 liter of store bought drinking water exposed to the sampling environment then transported, filtered and analyzed concurrently with field samples to serve as comprehensive negative controls.

DNA Extraction

Extraction of eDNA from filters used a combination of acetone dissolution and QIAGEN DNeasy Blood and Tissue kits. Filter dissolution in acetone applied methods adopted from Hallet et al (2012), with a modification to account for samples preserved in cell lysis buffer. Samples collected in 2021 were removed from buffer ATL and placed into new sterile microcentrifuge tubes to dry for 24 hours at room temperature while the buffer was retained. Samples from 2022 were dried at room temperature for 1 hour. Once filters had dried, two 3-mm sterile glass beads and 1.5 mL of acetone was placed into each tube and vortexed until the filters had dissolved ($\sim$ 30 minutes). Samples were then centrifuged to condense the genetic material and glass beads into a pellet. The supernatant was discarded from the tube and an additional acetone dissolution step was conducted. Remaining acetone was removed from the pellet by adding 1.5 ml of 100% anhydrous ethanol, resuspended through vortexing, and concentrated by centrifuging. The ethanol was then removed, and the pellet was air-dried at room temperature overnight. Once any remaining ethanol had evaporated, samples were lysed by adding 360 µl of buffer ATL and 40 µl of proteinase k followed by incubation for 8 hours at 56°C. For samples collected in 2021, the reserved ATL was used in this step. Samples were then extracted using QIAGEN DNeasy Blood and Tissue Kits according to the manufacturer's instructions and eluted to 100 µl of extracted DNA .

DNA Quantification

Assays specific to the target and tracer species were obtained from the literature (coho salmon: Pilliod and Laramie 2016, modified by Spence et al. 2021; steelhead: Wilcox et al. 2015; common carp: Eichmiller et al. 2014; channel catfish: U.S. Forest Service, Rocky Mountain Research Station).

The concentration of DNA in each sample was determined using a Bio-Rad QX200 Droplet Digital PCR system. Each ddPCR reaction included 900 nano molar (nM) of forward primer, 900 nM reverse primer, 250 nM of probe, 0.27 µl of 300 nM Dithiothreitol, 5 µl of ddPCR Multiplex Supermix, 12 µl of extracted DNA, and a sufficient volume of nuclease-free water to bring the final reaction volume to 22 µl. The channel catfish and coho salmon assays

were run in duplex while the common carp assay was run by itself. A total 20 µl of the reaction mix and 70 µl of droplet generator oil were transferred into a Bio-Rad DG8 droplet generation cartridge secured in a Cartridge Holder, covered with a DG8 Gasket, and transferred into a Bio-Rad QX-200 droplet generator where the ddPCR reaction is partitioned into as many as 20,000 nanodroplets. Each sample was transferred into the well of a ddPCR 96-well plate and the plate was sealed with a PCR plate sealer.

Thermocycling was performed on a C1000 Deep Well Touch Thermal Cycler. Thermocycling conditions included a 10-minute enzyme activation at 95℃, followed by 40 cycles of 30-second at 94℃ and 60-seconds at 60℃, followed by incubation at 98℃ for 10 minutes, and then held at 4℃ indefinitely. The ddPCR plate was placed on a QX200 droplet reader to estimate the concentration of DNA in the reaction. If a reaction contained less than 10,000 droplets and a clear distinction between positive and negative droplet fluorescence did not exist, the sample was rerun. Each plate included positive and negative PCR controls. Negative PCR controls consist of ddPCR reaction mix with nuclease-free water in lieu of extracted DNA. Positive controls consisted of individual ddPCR reactions using DNA extracted from the tissue of each assayed species. Bio-Rad software uses a Poisson algorithm based on the count of negative droplets to calculate the copies per 20 µL reaction. Estimated copies per reaction was corrected to copies per liter following Knudson et al. (2019).

Limits of detection and quantification were determined for each assay utilizing curve fitting methods. Limit of detection (LOD) is defined as the lowest concentration of DNA that can be detected with a 95 percent detection rate. Limit of quantification (LOQ) is defined as the lowest concentration of DNA that can be quantified with a coefficient of variation below 35 percent (Klymus et al. 2020). For each assay species, five-fold serial dilutions were made from nuclease free water and tissue extracted using a QIAGEN DNeasy Blood & Tissue Kit following the manufacturer's guidelines. Estimated concentrations across all serial dilutions ranged from 1 to 1 million copies per reaction with 8 - 16 replicates per concentration. LOD and LOQ were determined using the curve-fitting methods presented in Klymus et al. (2020) with the estimated concentrations at each step of serial dilutions.

Limits of detection differed across each assayed species with channel catfish having the lowest, 2.34 copies / reaction, followed by common carp, 2.55 copies / reaction, coho salmon, 4.77 copies / reaction and lastly steelhead 7.67 copies / reaction. The limit of quantification for channel catfish is 6 copies / reaction, common carp is 15 copies /reaction, coho salmon is 18 copies / reaction and steelhead is 23 copies / reaction.

Analysis
Ratios

To evaluate the change in coho salmon and steelhead eDNA concentrations relative to the tracer, sample-specific ratios were calculated by dividing NeDNA concentrations by the IPU tracer concentration within each sample. Ratios reveal patterns between NeDNA sources and a non-enriched source of eDNA from cross section to cross section. Any increase in these ratios

indicates presence of the target species between sampling locations. Ratios also mitigate variability in concentrations stemming from sample processing efficiencies and hydrological processes, which may depress or elevate eDNA concentrations independent of enrichment of eDNA from source organims. If changes in concentration arise from these processes, then the ratio between the two eDNA sources will remain constant. Ratios for 2021 samples were calculated using coho only, whereas ratios for 2022 were calculated for both coho and steelhead. The IPU tracer was selected instead of the CCA tracer due to its more central location of 100m upstream relative to the sample grid (figure 2) and its higher concentration (i.e., lower initial dilution). If DNA for a given target was undetected, an arbitrary concentration of .1 copy / liter was used for ratio calculation to maintain nine values per cross section in the subsequent analysis. If both the NeDNA species and the FeDNA species were undetected, then a ratio value of 0 was assigned to the sample.

Wilcox rank-sum test

Distributions of ratios from cross sections were compared using a two-sample Wilcoxon rank-sum test (function in the core R stats package) to assess if ratio values increased downstream. An increase would suggest the presence of the target species as a source of additional eDNA. The Wilcoxon rank-sum test is nonparametric (i.e., does not assume any distribution for the data) and is suitable for relatively small sample sizes. The analysis was based on a one-tailed test, in which rejection of the null hypothesis (no change in NeDNA:Tracer ratios) would be rejected only if the NeDNA:Tracer distribution increased at the downstream section. Wilcox tests were performed on ratio values from the 200m, 300m, 400m and 500m cross sections against those from the 100m cross sections. Previous work indicated the ADID tracer signals would fully mix within 100m (Herman 2023), therefore 100m cross section values were considered a baseline for NeDNA entering the study reach. This analysis was designed to ascertain the scale at which NeDNA:tracer comparisons could consistently assess occupancy of a target species. Snorkel survey observations will be used to corroborate the results from the wilcox test. Due to differences in the structure of the sampling grid (Figure 2) and the number of NeDNA sources tested across years, the number of tests possible for each distance treatment varies, but ranges from 8 to 13.

Results

Site Descriptions

Over two years of sampling, 13 field trials were conducted across 5 different streams, Mean depth across study reaches ranged from 0.2 m to 0.45 m, mean width of cross sections ranged from 3.29 m to 8.02 meters, and mean water velocity at sample locations collected ranged from 0.018 m/s to 0.427 m/s. Discharges varied between systems and across study reaches from 1.64 L/s to 142.31 L/s (Table 2).

Snorkel Surveys

Snorkel surveys detected coho salmon and steelhead at all sites. Although present at all sites, abundances of target species varied considerably by site. Observed abundances between 100m sections ranged between 8 - 499 coho, and 2 - 117 steelhead. Sites sampled in 2022 had the entire 400m sample range snorkeled, with abundances of coho ranging from 82 to 1797 and steelhead abundances ranging from 25 to 382 (Table 3 & 4).

NeDNA and tracer concentrations

Across every site, eDNA from the target species was detected in at least one sample entering study reaches from upstream at concentrations exceeding the LOD. The detection of NeDNA from upstream confirmed the presence of target species at the reach level and the necessity to use the tracer particle to account for NeDNA originating from upstream.

Across all sites, at least one sample from each cross section detected NeDNA and the tracer at concentrations greater than 0. Within the LR.21 site (Figure 12), all IPU tracer concentrations at the 1000m cross sections were less than the IPU assay's LOD and in the OC.21 site (Figure 13), all coho concentrations at the 100m, 500m and 1000m cross sections were less than the coho assay's LOD. Detections of coho and steelhead across all sites based on eDNA and snorkel surveys revealed that the target species were present within and above the study reaches (Scenario 3 from Table 1 and Figure 1). This means we did not have any sites where the target was absent from the study system (Scenario 2), the target was not present upstream of the study reach but present within it (Scenario 4), and the target was present upstream but not within the study reach (Scenario 1; Table 1, Figure 1).

The transport dynamics of the tracer particle and NeDNA within study reaches exhibited distinct characteristics at each site, reflecting site-specific patterns and sampling effects across eDNA sources. At the LMC4 (Figure 6) site, the tracer displayed a clear decline in measured concentration over the 400m study reach. Concurrently, NeDNA concentrations exhibited an increase between the 100m and 200m cross sections, followed by a gradual decline at a lower rate compared to the tracer. A similar consistent decline in tracer concentration was observed at the JC.21 site (FIgure 11), where NeDNA concentrations remained stable throughout the entire reach. Contrasting trends were captured at the PRC site (Figure 7), where tracer concentrations varied non-monotonically with downstream distance across the 400m study reach, while both NeDNA sources showed an increase. Despite the divergent patterns observed between the tracer and NeDNA at these sites, all scenarios resulted in increasing ratio values across the respective study reaches.

Samples captured consistent patterns across all eDNA sources, indicating strong sampling effects. These effects result in increasing and decreasing concentrations, seemingly independent of distance from its source or enrichment in the case of NeDNA. At the LMC1 site (Figure 3), declines in concentration across all eDNA sources between the 300m and 400m cross sections, followed by an increase between the 400m and 500m. Concurrent patterns between tracer and NeDNA are also captured at LMC2 (Figure 4), where all concentrations increase between the 200 and 300m cross sections. In these instances, ratio values will likely remain

constant as the pattern between sources of eDNA remains constant. In order to characterize divergent patterns between eDNA sources, ratios must be used to remove the effects of sampling.

<u>Wilcox test results</u>

Results from the Wilcox test identified that the frequency of significant increases in ratio values increased with distance downstream, from 5 in 21 (23%, Table 5) cases at a distance of 100m, to 16 of 21 over a separation of 400 m (76%, Table 5). In general, similar results were obtained for each target species, suggesting that the analysis was equally applicable to coho salmon and steelhead. Snorkel surveys found coho and steelhead within all sites and between every cross section, therefore the target species were present at every scale of comparison within the Wilcox test.

Discussion

This study adds to the limited body of research utilizing tracer particles within environmental DNA sampling. The methodology outlined in this paper provides an approach to assess the dynamics of environmental DNA over small spatial scales and apply it to eDNA from a target species to determine their presence. Detections of target species eDNA across all sites, validated through snorkel surveys, confirmed eDNA was able to determine their presence across all reaches. Although tracer concentrations did not decline with distance at every site, it was able to effectively calibrate concentrations of NeDNA and in many cases distinguish between NeDNA introduced from upstream and within a study reach. Results from field trials also highlight that eDNA distribution patterns vary greatly across sites, but share similarities within sites. Examining the distribution patterns in NeDNA revealed that concentration may fluctuate independent of distance from their source or the total abundance of source organisms, signifying the need for a calibration tool when studying eDNA patterns on this scale. The results from this research offer considerations and guidance for further environmental DNA studies utilizing a tracer particle.

<u>Tracer</u>

The utilization of a tracer provided insight into how well we perceive the distribution of NeDNA with the sampling grid used in this study. A key assumption of the tracer is that once it has mixed within a stream, it cannot unmix and is transported in a manner identical to that of any well-mixed NeDNA. This assumption enabled an examination of the distribution of eDNA across a study reach, offering insights into the dispersion of upstream eDNA within a site and the impact of fine-scale hydrodynamic processes on eDNA distribution. The tracer is used to determine how well we are observing NeDNA,and lets us account for the effects of sampling variability in discerning changes in the distribution of NeDNA relative to the tracer.

The way in which the sampling grid is distributed over plumes of eDNA on fine spatial scales highlights the importance of utilizing a tracer. Without a method for identifying and accounting for these processes, conclusions of species presence may be prone to false

conclusions. As seen in the LMC2 site, increasing coho and steelhead eDNA concentrations between 200m and 300m suggest species presence in this range. However, similar patterns captured in the tracer indicate that samples collected at that location resulted in heightened eDNA concentrations across all sources.

Sample locations

The power of tracer-augmented eDNA surveys to detect local occupancy increased with the distance separating sets of eDNA samples, with the most effective scheme being based on samples separated by 300-400m (Table 5). Wood et al. 2021 identified that at velocities of .15 m/s, comparable to those measured across the sites in this study, samples spaced 400 meters apart are ideal for detecting species. The distance scales at which these comparisons were able to effectively determine presence are consistent with those from previous studies conducted in similar systems (Wood et al. 2021, Van Driesche 2022, Wilcox et al. 2015). These studies identified that eDNA moves away from its sources in a dispersing plume where concentrations near the organisms are high but variable across the width of the stream. During transport with distance, eDNA breaks down and mixes across the stream, leading to lower concentrations relative to the source but with greater consistency in a cross section and with higher detection probabilities (as long as eDNA remains above the LOD). This pattern is reflected within this data, where a single outlier ratio is present in one cross section, followed by the elevation of all ratio values in the subsequent cross section (FWC1 or LMC4 figure).

Increased distance between cross sections also allows for the cumulative effects of all fish present between cross sections to be reflected in samples and improves the efficacy of the ADID system to detect NeDNA originating within a study reach. Across all sites sampled, coho and steelhead were found to be distributed throughout a reach from snorkel surveys. Distribution across an entire reach results in more fish being present with increasing distances and the greater chances for NeDNA dispersion plumes from those fish to break down and mix within the streams. With larger distances between cross sections, there are more opportunities to capture plumes of eDNA from more fish and elevate NeDNA signals relative to the tracer.

Discharge and fish abundance

Within this study, both discharge and the abundance of fish affected the success of occupancy assessments. Within sites with the highest recorded discharge (PCH.21, OC.21 & OCC), only 1 out of 14 ranked sum tests found significant increase in ratio values. High discharge sites coincided with low counts of coho, with PCH.21 having only 15 observed over 100 meters and no significant increase in ratio value. The OCC site had the highest observed abundance of steelhead, yet no significant differences in ratio values were found. These results are consistent with previous research showing that at higher discharges, the probability of capturing eDNA decreases due to dilution (Pochardt et al. 2020, Levi et al. 2019). Conversely, sites characterized by low discharge and high coho abundance had increased detections from

ratio comparisons. Sites with the highest coho counts and lowest discharge (FWC1 and JC.21) had significant increases in ratio value across 7 out of 8 rank sum tests.

The PCL.21 site had low discharge as well as low counts of coho, resulting in no significant test results. The relationship between very high and low abundance of fish on test results is consistent with a review article which identified positive relationships between eDNA concentrations and focal species biomass across studies (Rourke et al. 2021).

Sites characterized by observed abundances and discharges at the extremes of the data exhibited general trends consistent with previous studies, however sites with intermediate abundance and discharge values do not seem to exhibit consistent patterns with one another. In one case, two sites with identical discharges and similar coho abundances (LMC2 and LMC3) have diverging occupancy assessment results, with 50% of all rank sum tests having significant differences at LMC2 and only 12.5% of all tests at LMC3. Furthermore, the site with lower test results had three times as many steelhead observed within it. While sites with the highest observations of fish abundance and discharges conformed to expected relationships in some cases, those with intermediate values each resulted in differing estimates of site occupancy. This suggests that while certain factors influence the success of occupancy assessment, stream and site specific abiotic and biotics factors can strongly influence the outcome of these occupancy assessments.

References

Caza-Allard, I., M. Laporte, G. Côté, J. April, and L. Bernatchez. 2022. Effect of biotic and abiotic factors on the production and degradation of fish environmental DNA: An experimental evaluation. Environmental DNA 4(2):453–468.

Eichmiller, J. J., P. G. Bajer, and P. W. Sorensen. 2014. The Relationship between the Distribution of Common Carp and Their Environmental DNA in a Small Lake. PLOS ONE 9(11):e112611.

Everts, T., C. Van Driessche, S. Neyrinck, N. De Regge, S. Descamps, A. De Vocht, H. Jacquemyn, and R. Brys. 2022. Using quantitative eDNA analyses to accurately estimate American bullfrog abundance and to evaluate management efficacy. Environmental DNA 4(5):1052–1064.

Fremier, A. K., K. M. Strickler, J. Parzych, S. Powers, and C. S. Goldberg. 2019. Stream Transport and Retention of Environmental DNA Pulse Releases in Relation to Hydrogeomorphic Scaling Factors. Environmental Science & Technology 53(12):6640–6649.

Hallett, S. L., R. A. Ray, C. N. Hurst, R. A. Holt, G. R. Buckles, S. D. Atkinson, and J. L. Bartholomew. 2012. Density of the Waterborne Parasite Ceratomyxa shasta and Its Biological Effects on Salmon. Applied and Environmental Microbiology 78(10):3724–3731.

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

Hauer, F. R., and G. A. Lamberti. 2006. Methods in Stream Ecology. Elsevier.

Herman, B. 2023. Use of foreign edna tracers to resolve site- and time-specific eDNA distributions in natural streams. Cal Poly Humboldt theses and projects.

Jo, T., and T. Minamoto. 2021. Complex interactions between environmental DNA (eDNA) state and water chemistries on eDNA persistence suggested by meta-analyses. Molecular Ecology Resources 21(5):1490–1503.

Knudsen, S. W., R. B. Ebert, M. Hesselsøe, F. Kuntke, J. Hassingboe, P. B. Mortensen, P. F. Thomsen, E. E. Sigsgaard, B. K. Hansen, E. E. Nielsen, and P. R. Møller. 2019. Species-specific detection and quantification of environmental DNA from marine fishes in the Baltic Sea. Journal of Experimental Marine Biology and Ecology 510:31–45.

NOAA Fisheries. 2023. Southern Oregon/Northern California Coast Coho Salmon | NOAA Fisheries. https://www.fisheries.noaa.gov/west-coast/endangered-species-conservation/southern-oregon-northern-california-coast-coho-salmon.

Ostberg, C. O., and D. M. Chase. 2022. Ontogeny of eDNA shedding during early development in Chinook Salmon (Oncorhynchus tshawytscha). Environmental DNA 4(2):339–348.

Ostberg, C. O., D. M. Chase, M. S. Hoy, J. J. Duda, M. C. Hayes, J. C. Jolley, G. S. Silver, and C. Cook-Tabor. 2019. Evaluation of environmental DNA surveys for identifying occupancy and spatial distribution of Pacific Lamprey (Entosphenus tridentatus) and Lampetra spp. in a Washington coast watershed. Environmental DNA 1(2):131–143.

Pilliod, D. S., and M. B. Laramie. 2016a. Salmon redd identification using environmental DNA (eDNA). Page Open-File Report. U.S. Geological Survey, 2016–1091.

Pilliod, D. S., and M. B. Laramie. 2016b. Salmon redd identification using environmental DNA (eDNA). Page Open-File Report. U.S. Geological Survey, 2016–1091.

Rourke, M. L., A. M. Fowler, J. M. Hughes, M. K. Broadhurst, J. D. DiBattista, S. Fielder, J. Wilkes Walburn, and E. M. Furlan. 2022. Environmental DNA (eDNA) as a tool for assessing fish biomass: A review of approaches and future considerations for resource surveys. Environmental DNA 4(1):9–33.

Schmelzle, M. C., and A. P. Kinziger. 2016. Using occupancy modelling to compare environmental DNA to traditional field methods for regional-scale monitoring of an endangered aquatic species. Molecular Ecology Resources 16(4):895–908.

Shaffer, J. 2022. Comparison of environmental DNA and underwater visual count surveys for detecting juvenile coho salmon (Oncorhynchus kisutch) in rivers. Cal Poly Humboldt theses and projects.

Snyder, E. D., J. L. Tank, P. F. P. Brandão-Dias, K. Bibby, A. J. Shogren, A. W. Bivins, B. Peters, E. M. Curtis, D. Bolster, S. P. Egan, and G. A. Lamberti. 2023. Environmental DNA (eDNA) removal rates in streams differ by particle size under varying substrate and light conditions. Science of The Total Environment 903:166469.

Spence, B. C., D. E. Rundio, N. J. Demetras, and M. Sedoryk. 2021. Efficacy of environmental DNA sampling to detect the occurrence of endangered coho salmon (Oncorhynchus kisutch) in Mediterranean-climate streams of California's central coast. Environmental DNA 3(4):727–744.

Sutter, M., and A. Kinziger. 2019. Rangewide tidewater goby occupancy survey using environmental DNA. Conservation Genetics 20.

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

Thalinger, B., A. Rieder, A. Teuffenbach, Y. Pütz, T. Schwerte, J. Wanzenböck, and M. Traugott. 2021b. The Effect of Activity, Energy Use, and Species Identity on Environmental DNA Shedding of Freshwater Fish. Frontiers in Ecology and Evolution 9.

Troth, C. R., M. J. Sweet, J. Nightingale, and A. Burian. 2021. Seasonality, DNA degradation and spatial heterogeneity as drivers of eDNA detection dynamics. Science of The Total Environment 768:144466.

Turner, C. R., M. A. Barnes, C. C. Y. Xu, S. E. Jones, C. L. Jerde, and D. M. Lodge. 2014. Particle size distribution and optimal capture of aqueous macrobial eDNA. Methods in Ecology and Evolution 5(7):676–684.

Van Driessche, C., T. Everts, S. Neyrinck, and R. Brys. 2022. Experimental assessment of downstream environmental DNA patterns under variable fish biomass and river discharge rates. Environmental DNA 5(1):102–116.

Wilcox, T. M., K. J. Carim, K. S. McKelvey, M. K. Young, and M. K. Schwartz. 2015a. The Dual Challenges of Generality and Specificity When Developing Environmental DNA Markers for Species and Subspecies of Oncorhynchus. PLOS ONE 10(11):e0142008.

Wilcox, T. M., K. S. McKelvey, M. K. Young, W. H. Lowe, and M. K. Schwartz. 2015b. Environmental DNA particle size distribution from Brook Trout (Salvelinus fontinalis). Conservation Genetics Resources 7(3):639–641.

Wood, Z. T., A. Lacoursière-Roussel, F. LeBlanc, M. Trudel, M. T. Kinnison, C. Garry McBrine, S. A. Pavey, and N. Gagné. 2021a. Spatial Heterogeneity of eDNA Transport Improves Stream Assessment of Threatened Salmon Presence, Abundance, and Location. Frontiers in Ecology and Evolution 9.

Wood, Z. T., A. Lacoursière-Roussel, F. LeBlanc, M. Trudel, M. T. Kinnison, C. Garry McBrine, S. A. Pavey, and N. Gagné. 2021b. Spatial Heterogeneity of eDNA Transport Improves Stream Assessment of Threatened Salmon Presence, Abundance, and Location. Frontiers in Ecology and Evolution 9.

Tables

Table 1: *Contingency table illustrating four potential scenarios regarding the presence of a species both upstream and within the study reach. This table corresponds to the scenarios depicted in Figure 1.*

	Target present above study reach	Target absent above study reach
Target absent in study reach	1 Target species eDNA is detected entering the reach and declines synchronously and proportionately with the tracer	2 Target species is not detected entering or within the reach while the tracer decreases
Target present in study reach	3. Target species eDNA is detected entering the reach and does not exhibit the same decline in concentration as the eDNA tracer	4. Target species eDNA is not detected entering the reach but is detected within the reach, increasing in concentration while the eDNA tracer decreases

Table 2: Stream measurements conducted at all study reaches including; stream name, study reach ID, mean width of cross sections (meters), mean depth at all sampling locations across all cross sections (meters), mean velocity at all sampling locations across all cross sections (meters / second), maximum recorded velocity across all sampling location (meters / second) and the single discharge measurement collected at a site (liter/s)

Stream name	Site name	Mean width (m)	Mean depth (m)	Mean velocity (m/s)	Max velocity (m/s)	Discharge (L/s)
Freshwater Creek	FWC1	8.02	0.28	0.02	0.09	4.32
Jacoby Creek	JC.21	5.88	0.28	0.02	0.15	1.64
Jacoby Creek	JC2	4.55	0.22	0.09	0.21	33.48
Lost Man Creek	LMC1	5.80	0.20	0.24	0.55	41.64
Lost Man Creek	LMC2	6.88	0.45	0.21	0.61	21.46
Lost Man Creek	LMC3	5.18	0.29	0.11	0.55	21.46
Lost Man Creek	LMC4	3.29	0.26	0.06	0.30	7.95
Little River	LR.21	7.59	0.34	0.05	0.24	44.80
Old Campbell Creek	OC.21	6.49	0.31	0.11	0.40	86.15
Old Campbell Creek	OCC	6.18	0.34	0.16	0.52	62.36
Prairie Creek	PCH.21	5.13	0.38	0.43	1.01	142.29
Prairie Creek	PCL.21	3.89	0.20	0.06	0.24	6.24
Prairie Creek	PRC	7.78	0.26	0.06	0.27	15.67

Table 3: Visual observations of coho from snorkel surveys conducted on 2021 study reaches.

Site	Distance snorkeled (m)	Number of coho
JC.21	100	97
LR.21	100	58
OC.21	100	38
PCH.21	100	15
PCL.21	100	20

Table 4: Visual observations of coho and steelhead from snorkel surveys conducted on 2022 study reaches.

Site	Distance snorkeled (m)	Number of coho	Number of steelhead
FWC1	400	1797	39
JC2	400	862	77
LMC1	400	696	83
LMC2	400	526	25
LMC3	400	381	80
LMC4	400	972	168
OCC	400	82	382
PRC	400	419	123

Table 5: Results of the Wilcoxon ranked-sum test performed at all distance treatments for the ratios of coho:tracer and steelhead:tracer

Distance between cross sections	Number of tests	Number of significant results	Proportion of significant results
100 m	21	5	23%
200 m	21	8	38%
300 m	16	11	68%
400 m	21	16	76%

Figures
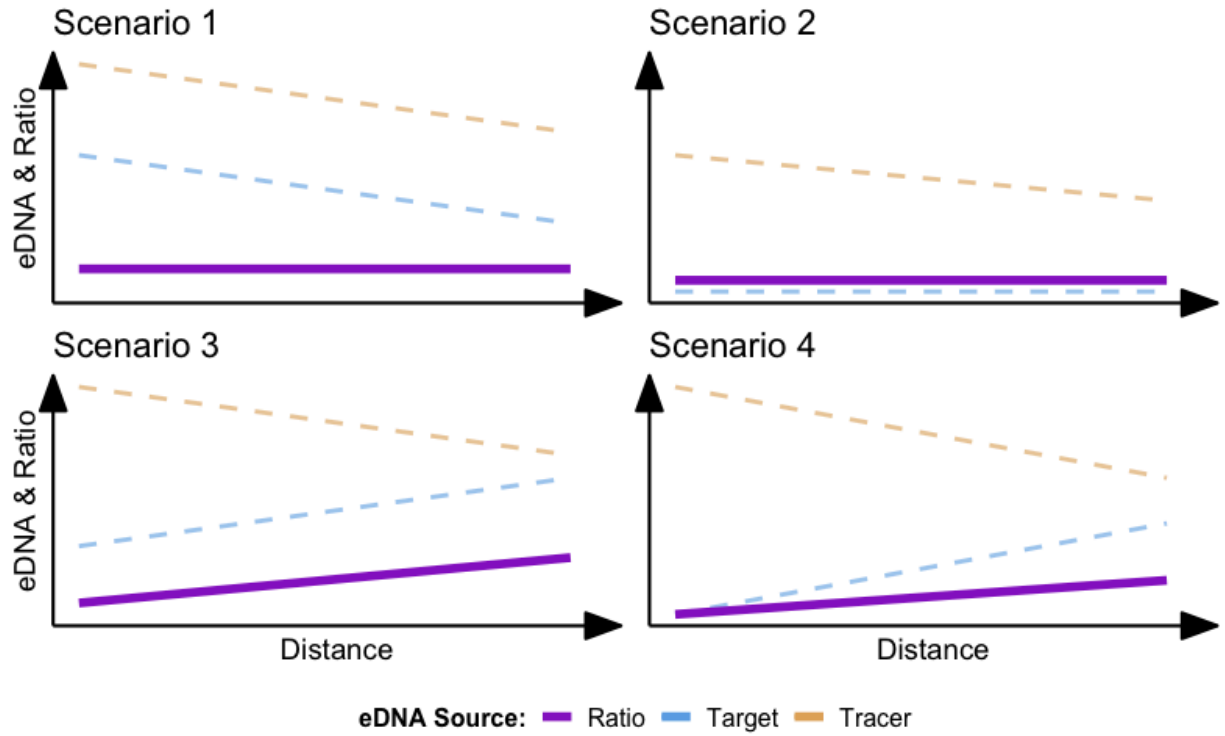
Figure 1: Four occupancy scenarios and the theoretical eDNA concentrations for the tracer (blue), target species (yellow), and the resulting ratio between eDNA sources (purple).
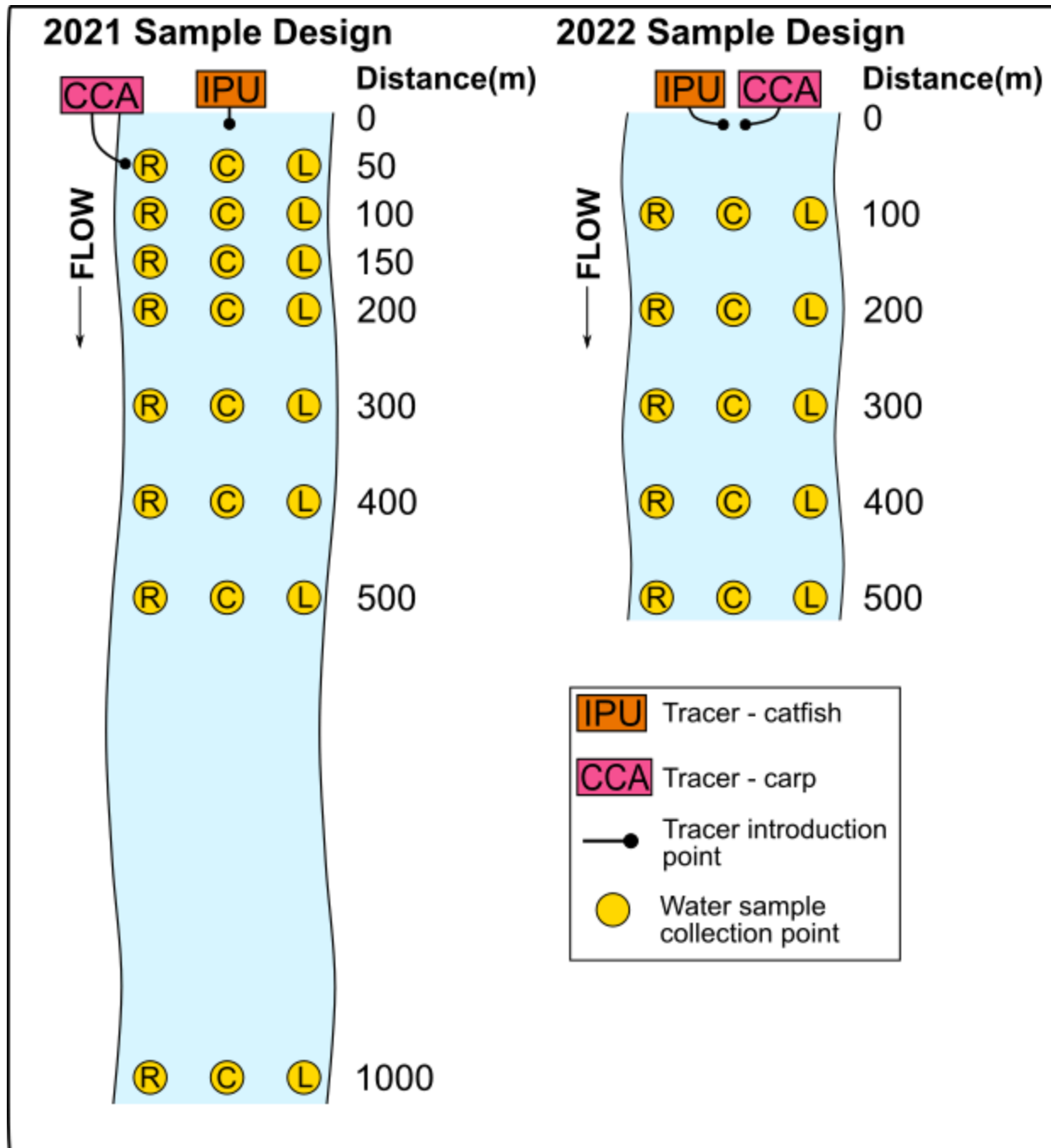
Figure 2: *Diagram of study reaches and sampling grid from 2021 and 2022 showing ADID placement and tracer species (IPU & CCA), distance (in meters) of cross section from the head of the study reach, and locations in the transects where samples were collected (L,C,R).*
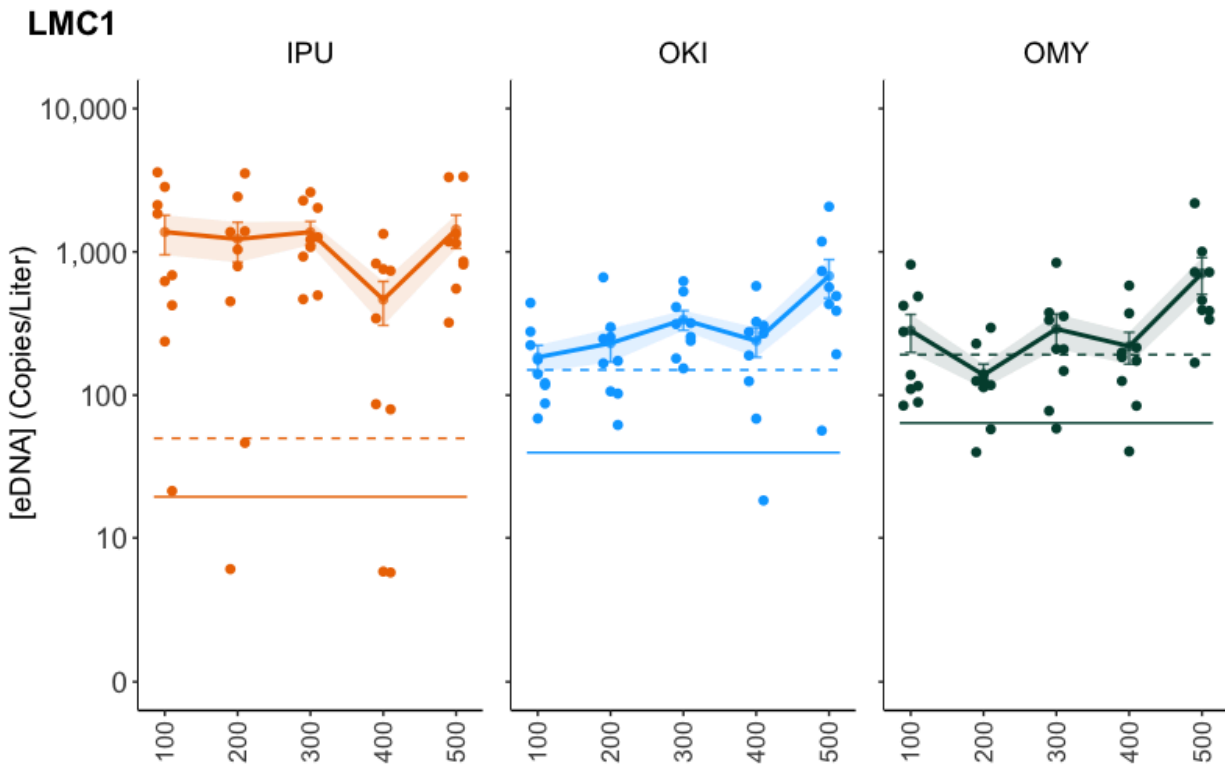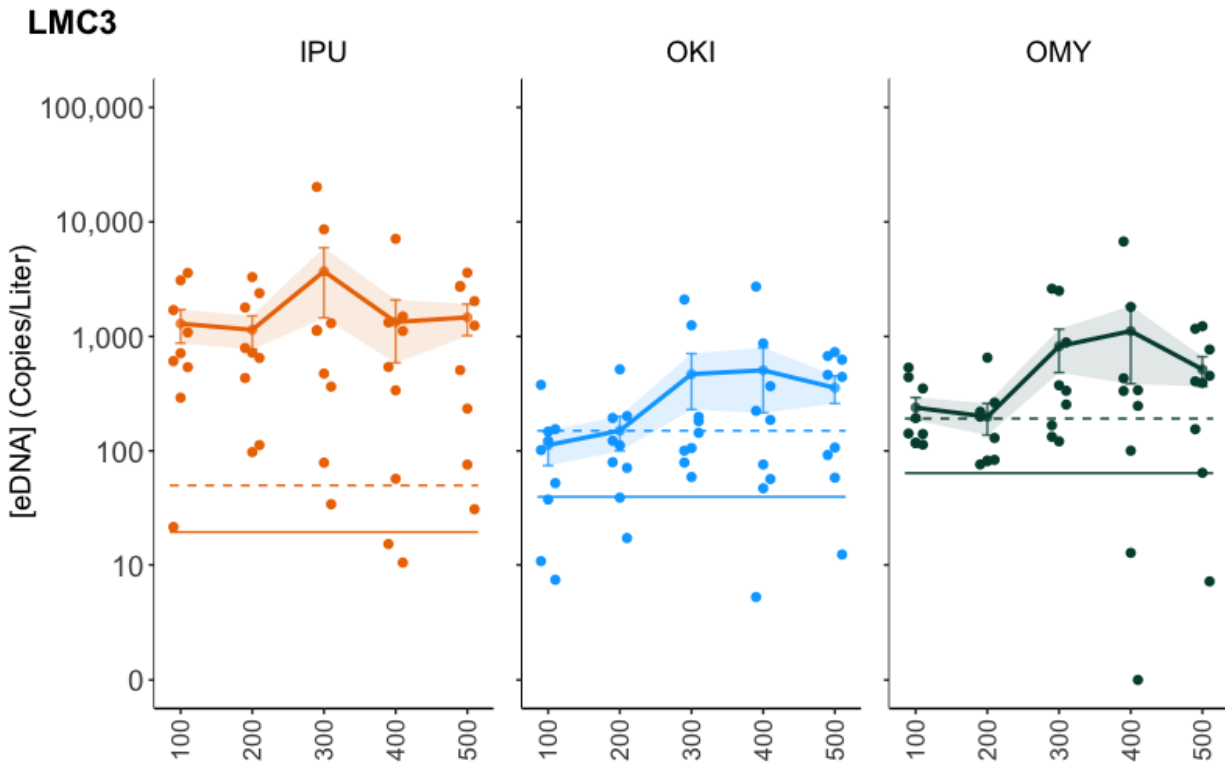
Figure 3: *IPU tracer, Coho and Steelhead eDNA concentrations at the LMC1 site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).*
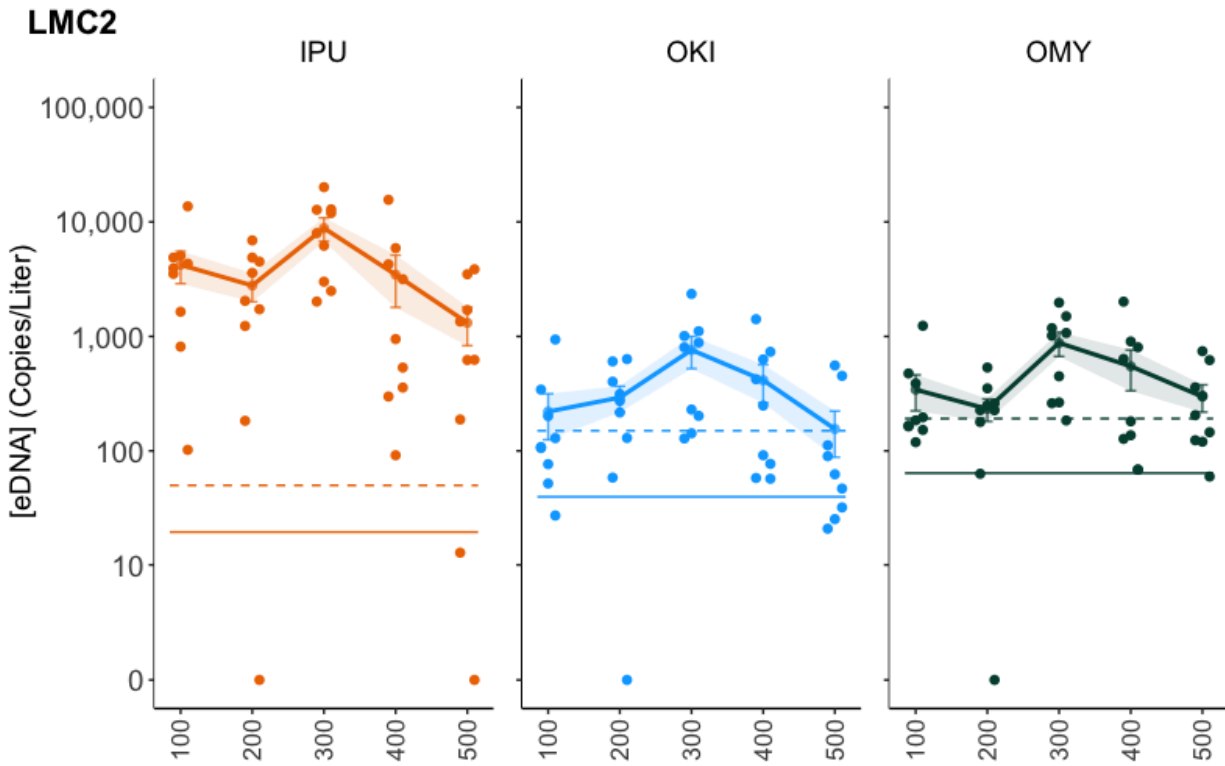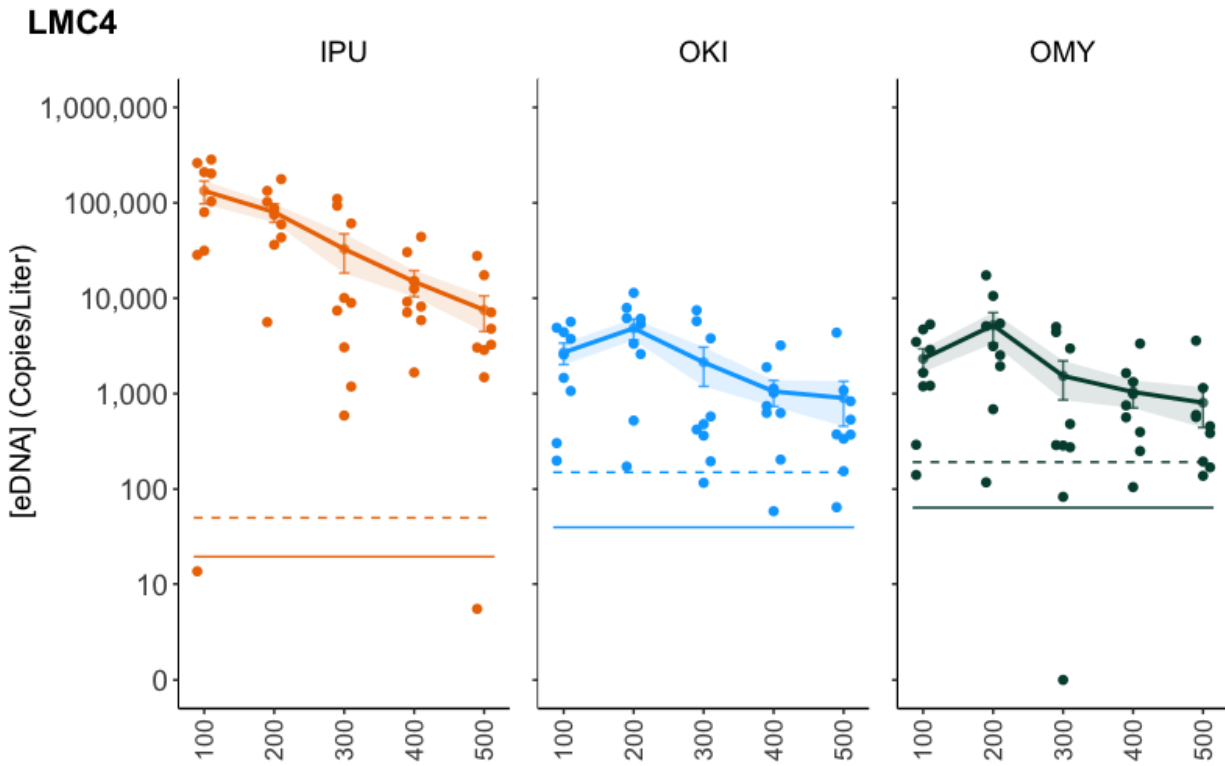
Figure 4: IPU tracer, Coho and Steelhead eDNA concentrations at the LMC2 site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).

Figure 5: *IPU tracer, Coho and Steelhead eDNA concentrations at the LMC3 site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).*
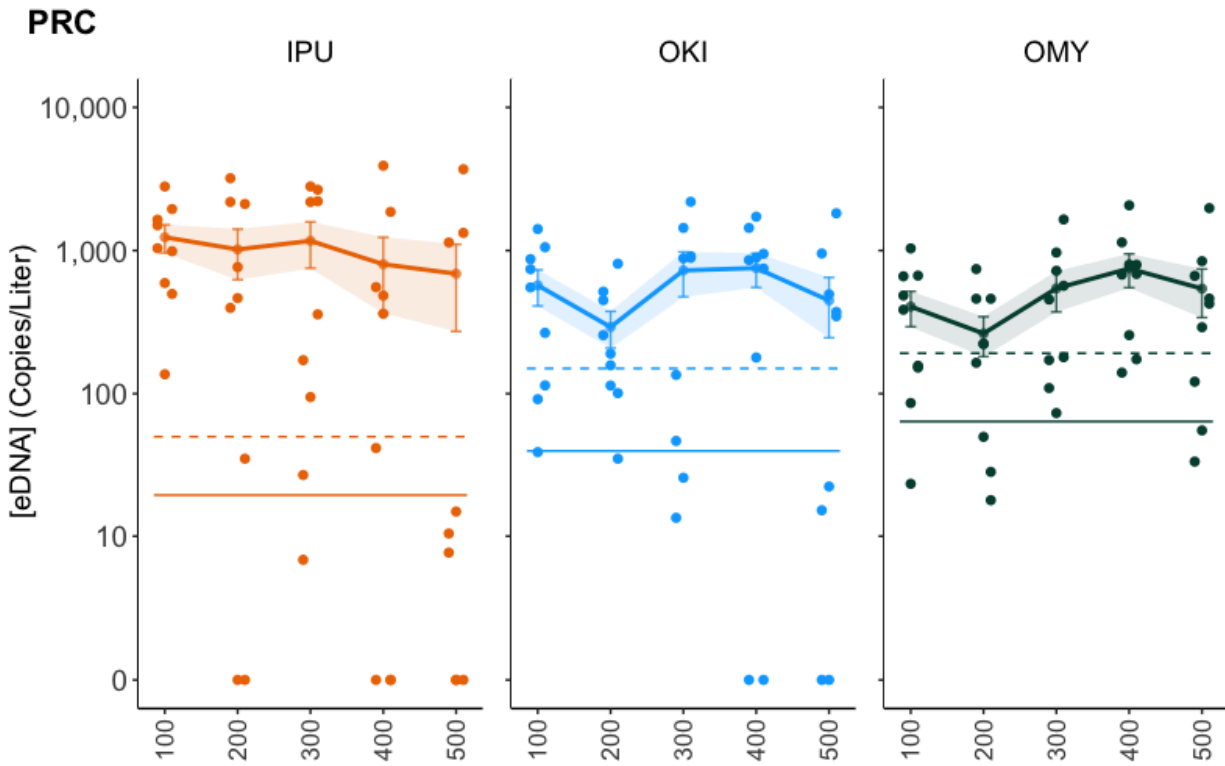
Figure 6: *IPU tracer, Coho and Steelhead eDNA concentrations at the LMC4 site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).*
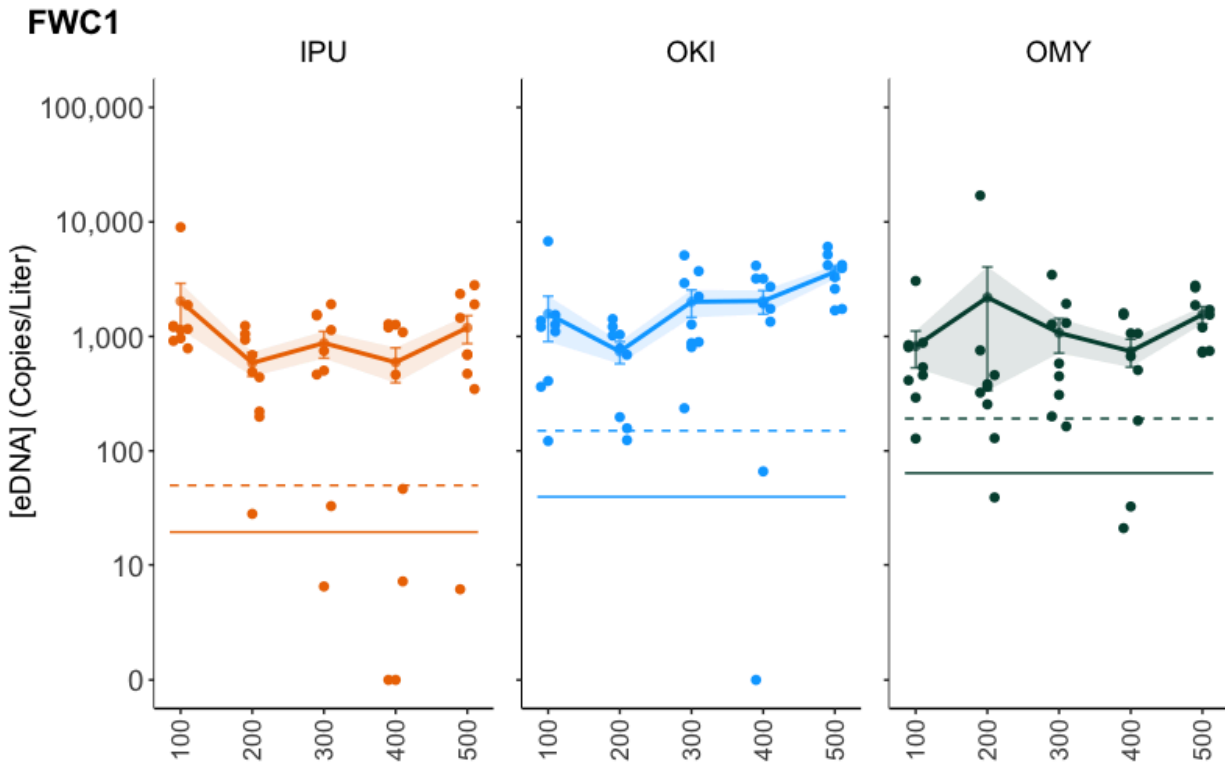
Figure 7: *IPU tracer, Coho and Steelhead eDNA concentrations at the PRC site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).*
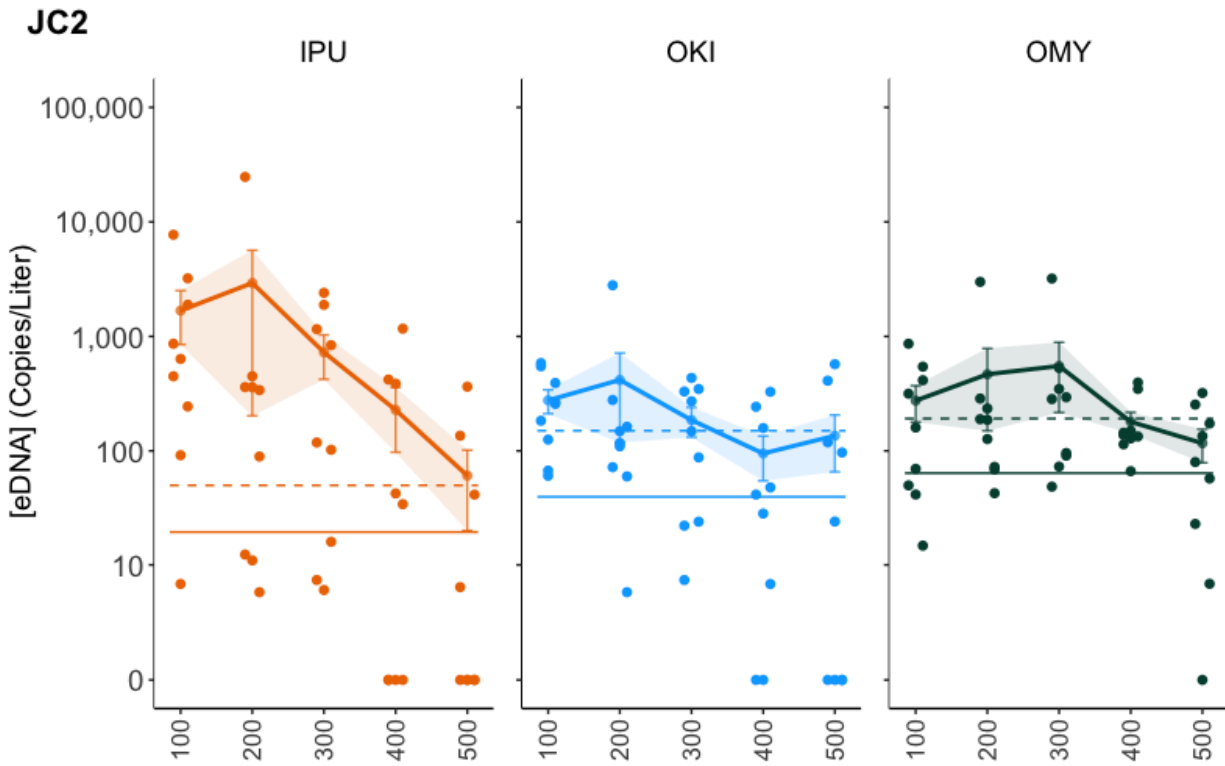
Figure 8: *IPU tracer, Coho and Steelhead eDNA concentrations at the FWC1 site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).*

Figure 9: *IPU tracer, Coho and Steelhead eDNA concentrations at the JC2 site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).*
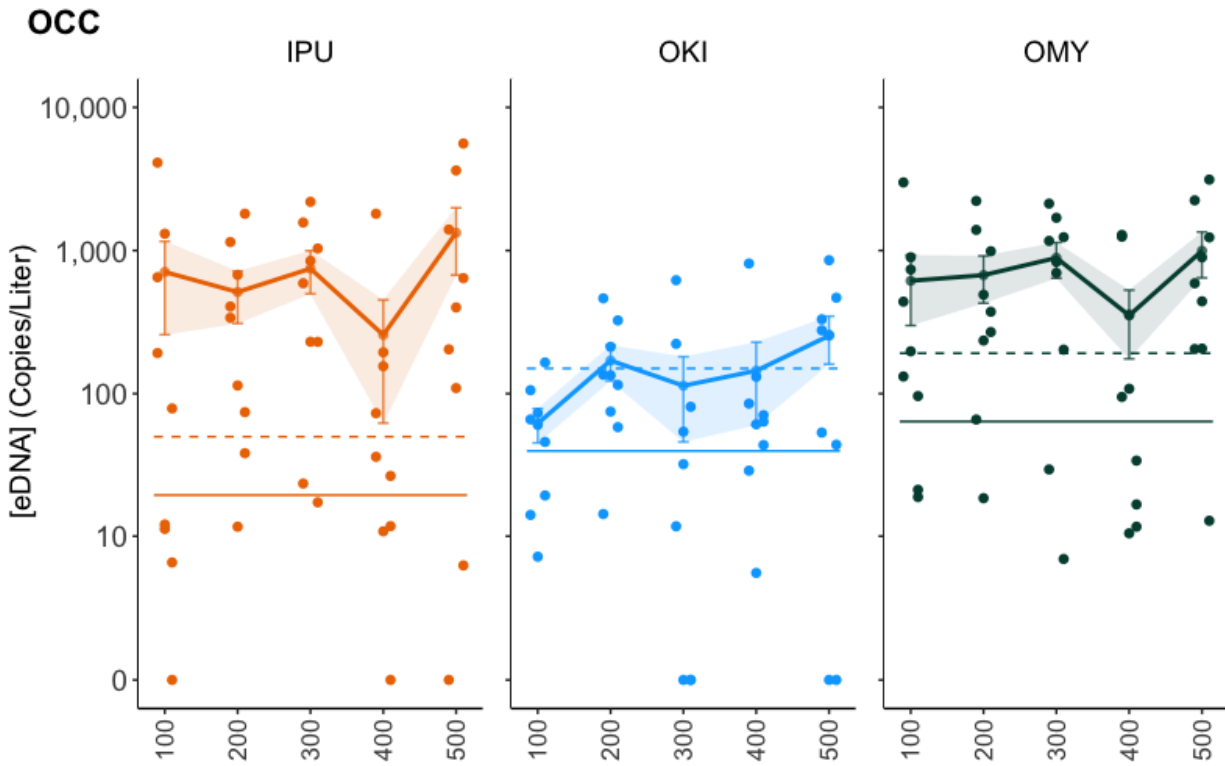
Figure 10: *IPU tracer, Coho and Steelhead eDNA concentrations at the OCC site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).*
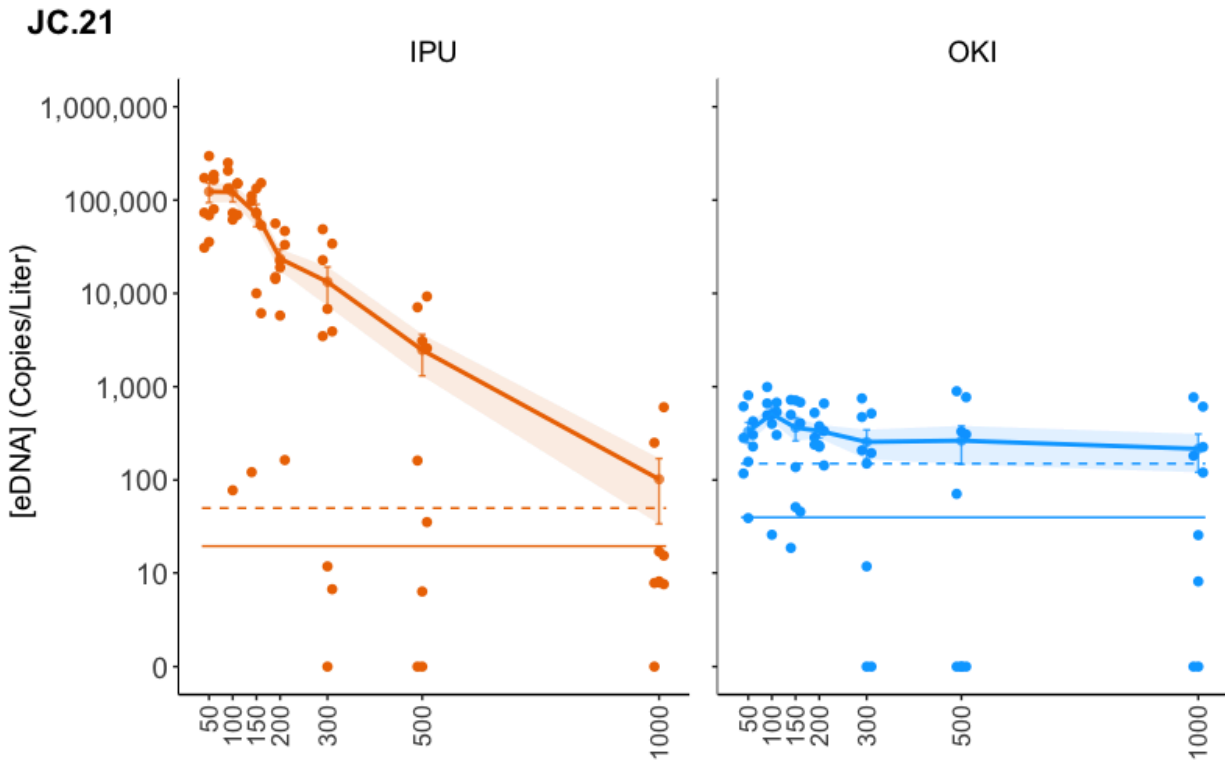
Figure 11: *IPU tracer, Coho and Steelhead eDNA concentrations at the JC.21 site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).*
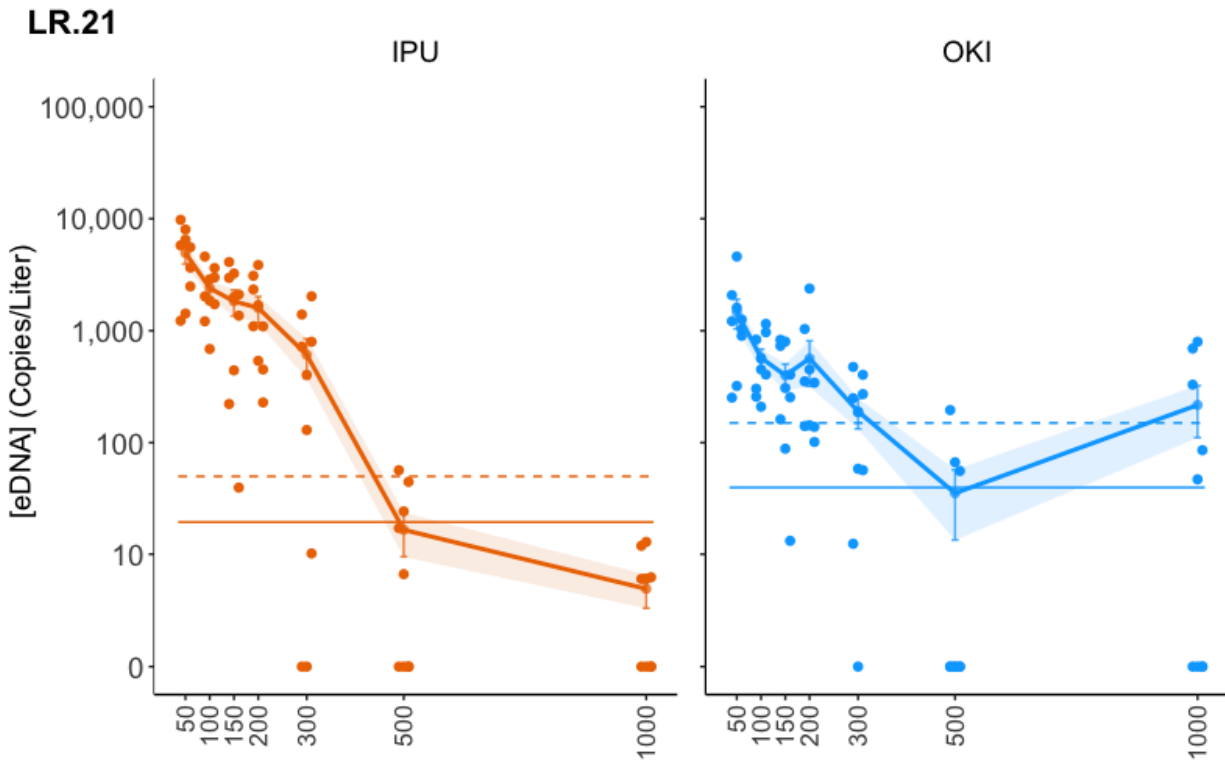
Figure 12: *IPU tracer, Coho and Steelhead eDNA concentrations at the LR.21 site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).*

Figure 13: IPU tracer, Coho and Steelhead eDNA concentrations at the OC.21 site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).
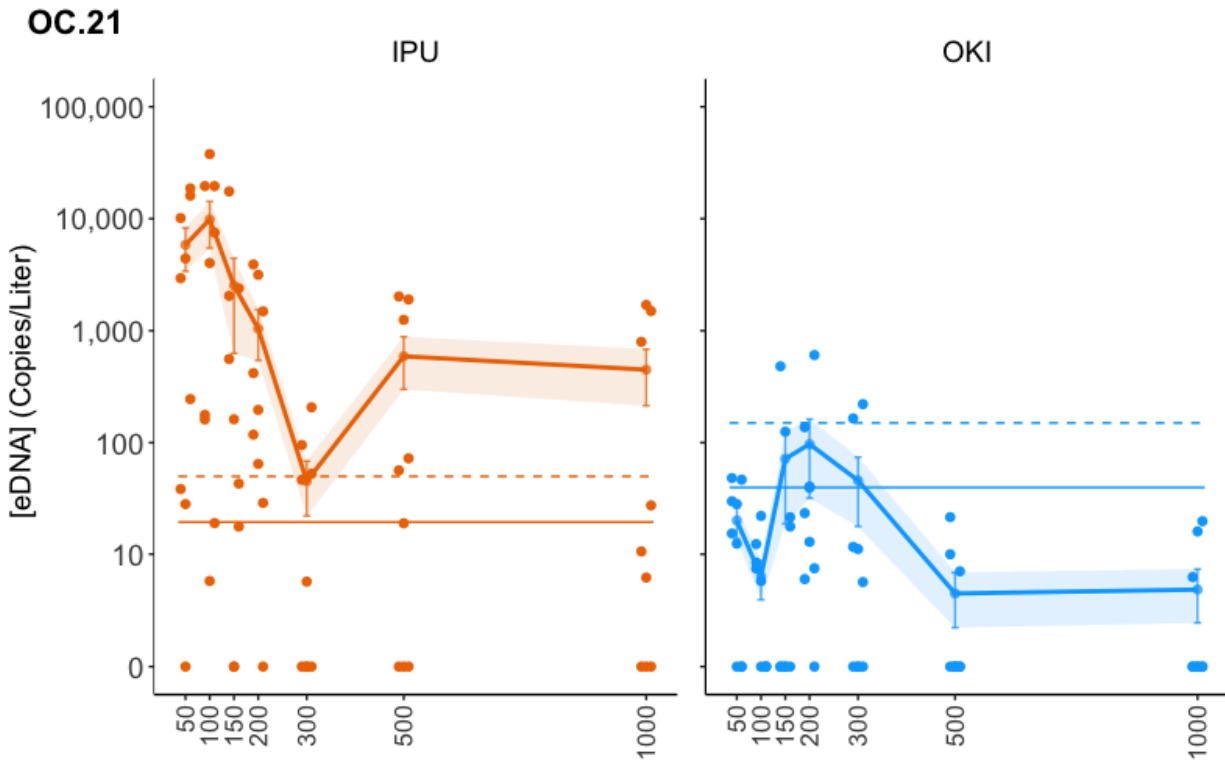
Figure 14: *IPU tracer, Coho and Steelhead eDNA concentrations at the PCH.21 site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).*
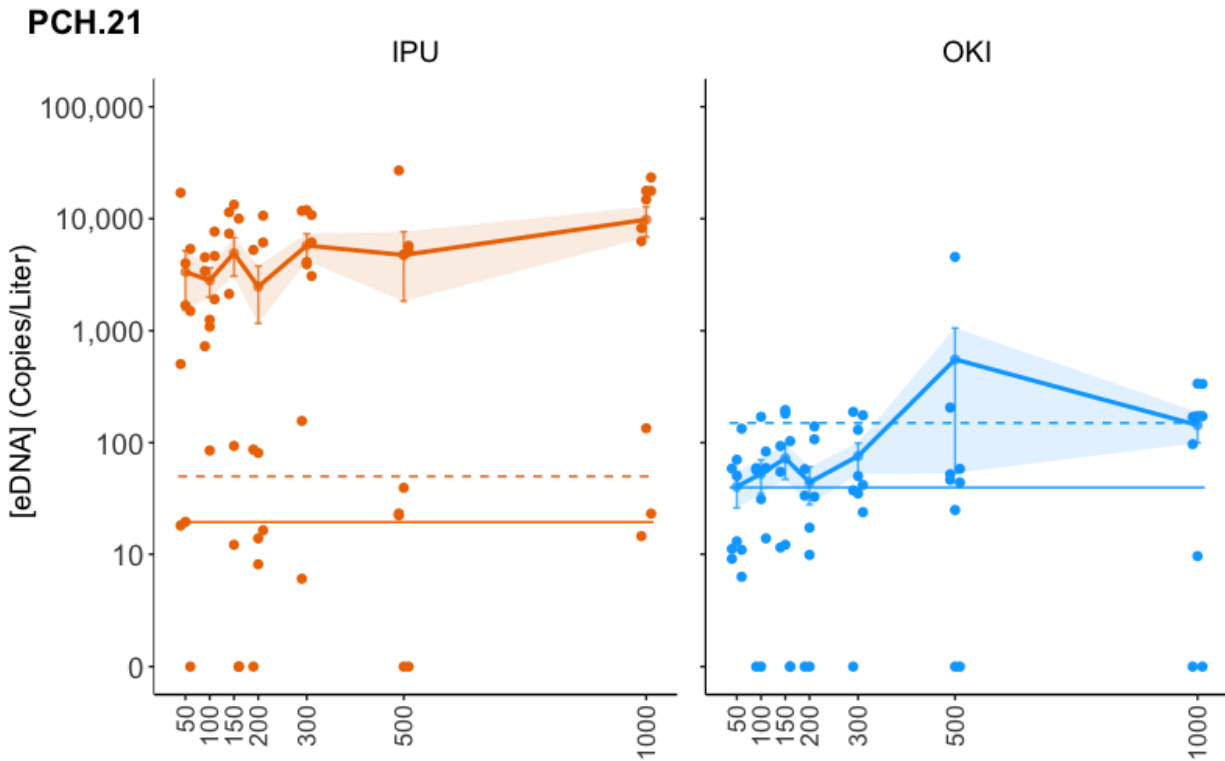
Figure 15: *IPU tracer, Coho and Steelhead eDNA concentrations at the PCL.21 site. Points represent concentrations across the 9 samples at a cross section; light points and error bars are the average concentration +/- Standard error at a cross section. A jitter is applied to the 9 points for clarity. The horizontal lines indicate the species-specific LOD (solid) and LOQ (dashed).*
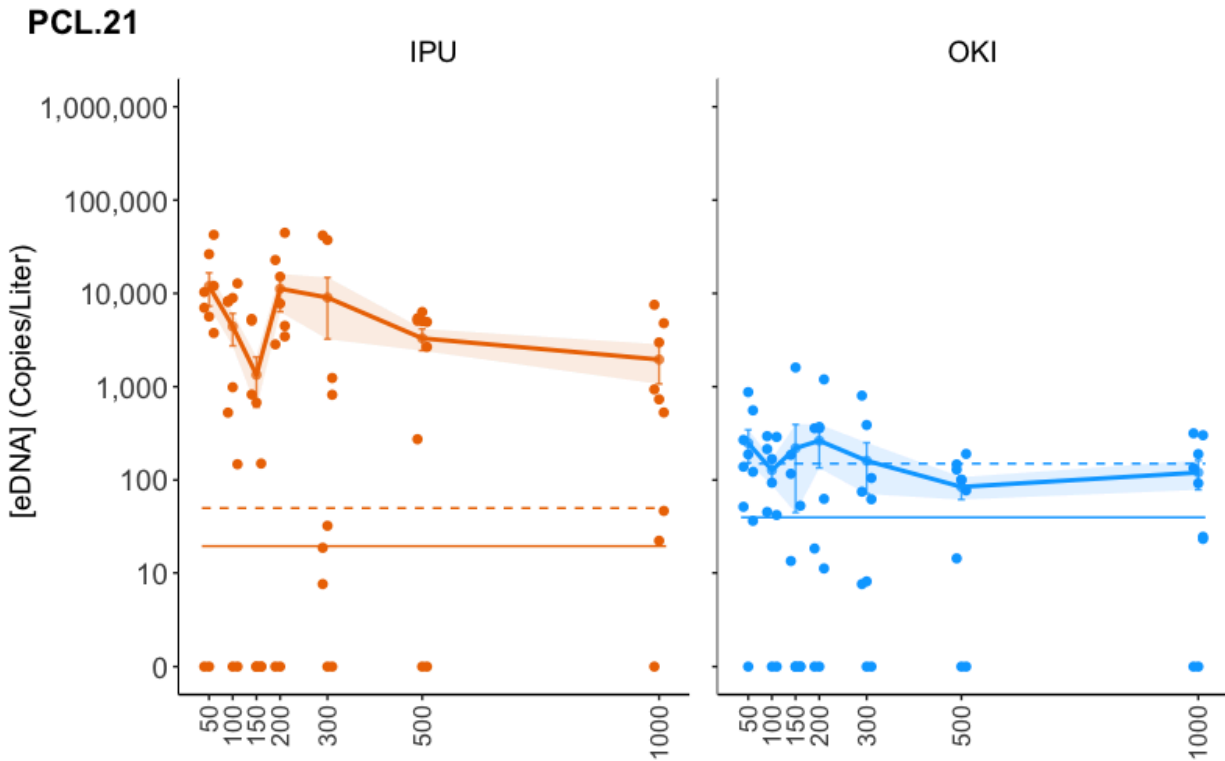
TASK 4: EVALUATE THE UTILITY OF EDNA TO DETERMINE JUVENILE COHO SALMON DISTRIBUTION AT BROADER GEOGRAPHIC SCALES IN COMPARISON TO TRADITIONAL FIELD SURVEYS FROM SNORKEL MONITORING.

Comparison of environmental DNA and underwater visual count surveys for detecting juvenile Coho Salmon in rivers

Jason T. Shaffer[1,*]

Andrew P. Kinziger[1]

Eric P. Bjorkstedt[2]

Andre Buchheister[1]

[1]*Department of Fisheries Biology, Cal Poly Humboldt, Arcata, CA 95521, USA*

[2]*Southwest Fisheries Science Center, NOAA Fisheries, Trinidad, CA 95570, USA*

Corresponding Author: Jason.Shaffer@humboldt.edu

Synopsis: Environmental DNA was a highly effective survey method for detecting juvenile Coho Salmon in streams, yielding detection probabilities >0.90 under most conditions, but eDNA was not a good predictor of relative abundance in small pools.

Abstract

We assessed environmental DNA (eDNA) techniques for species detection within an optimal benchmarking environment, utilizing underwater visual count (UVC) survey methods with established detection probabilities surpassing 0.90. We conducted paired eDNA and UVC surveys via snorkeling to evaluate the ability of eDNA methods to detect and enumerate threatened juvenile Coho Salmon *Oncorhynchus kisutch* in small pools in the Smith River, California. In surveys of 96 pools across 25 stream reaches eDNA and UVC methods showed a high degree of agreement in detecting the presence and the absence of Coho Salmon within a pool (93% agreement) and within a survey reach (80% agreement). Multi-scale occupancy modeling indicated that detection probabilities were similar and high for both methods under median habitat conditions (eDNA=91%, UVC=89%). The contributing basin area to each survey reach (a proxy for discharge) had a strong, negative effect on detection probabilities (p) that was more pronounced for eDNA than for UVC (e.g., at the highest basin areas, $p_{eDNA} = 34\%$ whereas $p_{UVC} = 77\%$), indicating the importance of considering the dilution effects of discharge in eDNA study design. Based on a generalized linear model, eDNA concentration was a poor predictor of Coho Salmon counts in small pools. This study demonstrates that when compared to UVC surveys with high detection probabilities, eDNA methods yielded nearly identical results, and are likely to enhance sensitivity in the likely event that species detection is substantially less than 1. We find that eDNA methods can provide a highly effective approach for determining the distribution of Coho Salmon at broad geographic scales, but additional investigation is required before this approach could be used to determine relative abundance in small pools.

Introduction

Conservation projects typically rely on monitoring programs for assessing population trends and collecting relevant ecological data to make appropriate management decisions or to evaluate the effects of past decisions (Nichols and Williams 2006; Lovett et al. 2007). Monitoring species' geographic distribution through space and time is critical for the conservation of depleted populations (e.g., McElhany et al. 2000), but tracking the spatial structure of populations is challenging when species are hard to observe, have broad spatial distributions, occur at low abundance, or inhabit remote areas (Albanese et al. 2011; MacKenzie et al. 2018). These factors increase the chances of failing to detect a species that is present, and they bias estimates of species distributions if not accounted for. To address these challenges, monitoring programs often require extensive survey efforts and may utilize multiple survey methods to maximize detectability for more accurate quantification of population spatial distribution (Nichols et al. 2008).

Underwater visual count (UVC) surveys are commonly used to monitor the distribution and abundance of aquatic species (Hankin and Reeves 1988; Thurow 1994). Underwater visual surveys via direct (e.g., snorkeling) or indirect (e.g., camera stations) observation are often used in remote areas due to minimal gear requirements or when environmental conditions (e.g., deep water or high conductivity) limit the effectiveness of other methods such as seining or electrofishing (Thurow 1994; Albanese et al. 2011). The minimally invasive nature of UVC surveys makes them well-adapted for surveys of sensitive or imperiled species such as those listed by state or federal agencies. However, UVC surveys are prone to imperfect detection especially when abundance is low, when species are morphologically similar, or when field observations are limited by water clarity, depth, or habitat complexity (Thurow et al. 2012; Staton et al. 2022; see also Gu and Swihart 2004; MacKenzie et al. 2018).

Environmental DNA (eDNA) is a rapidly expanding and promising tool for assessing the distribution and abundance of aquatic species (Thomsen and Willerslev 2015; Rodríguez‑Ezpeleta et al. 2021). Several studies suggest that eDNA methods are more rapid, cost-effective, and sensitive than conventional survey methods, particularly when surveying for rare or endangered species (Laramie et al. 2015; Strickland and Roberts 2019; Sutter and Kinziger 2019; Spence et al. 2021; Yu et al. 2021). Compared to conventional monitoring methods, eDNA surveys offer a notable advantage in species detection due to the broad eDNA plumes emitted by aquatic organisms, resulting in nearly double the detection rates in some studies (Schmelzle and Kinziger 2016; McColl-Gausden et al. 2021; Dougherty et al. 2016; Valdivia-Carrillo et al. 2021). Successful detection of a target species with eDNA methods is influenced by the quantity, distribution, and attrition of eDNA in a system, which in turn depend on the characteristics of the target species (e.g., behavior, abundance, distribution, size, DNA shedding rate, and the form of eDNA produced; Turner et al. 2014; Yates et al. 2019; Castañeda et al. 2020, Andruszkiewicz Allan et al. 2021), and conditions that govern the movement and attrition of eDNA in the system (e.g., discharge, velocity, substrate, temperature, ph, salinity; Baldigo et al. 2017; Mize et al. 2019; Harrison et al. 2019; Spence et al. 2021; Wood et al. 2021; Yates et al. 2021). These factors can influence the efficacy of eDNA methods to a degree that requires eDNA monitoring strategies to be tailored to the location and the species of interest (Goldberg et al. 2016; Yates et al. 2021).

Occupancy modeling frameworks have been successfully applied to eDNA datasets to account for the imperfect detection of eDNA in water samples and for the influence of environmental factors on occupancy and detectability (e.g., Schmelzle and Kinziger 2016; Sutter and Kinziger 2019; Smith and Goldberg 2020; Martel et al. 2021). The hierarchical nature of eDNA surveys yields data that is highly suited to analysis using a multi-scale occupancy framework to estimate occupancy patterns at multiple spatial scales while accounting for environmental covariates and differences in survey method (Nichols et al. 2008; Dorazio and Erickson 2017; MacKenzie et al. 2018). For example, multi-scale occupancy models can be used to estimate and compare the detection probabilities of concurrently applied survey methods,

providing critical information for survey design and method-specific effectiveness (Nichols et al. 2008).

Study designs that utilize eDNA methods in conjunction with other conventional survey methods to compare results are common in the literature, but the formal quantification and comparison of method-specific detection probabilities is limited, particularly for river systems (e.g., Castañeda et al. 2020). Fediajevaite et al. (2021) found that only 18 of 535 studies using eDNA methods (3%) provided a quantitative comparison of eDNA to conventional survey methods via estimation of method-specific detection probabilities. Given the rarity of robust comparative studies, additional comparisons of UVC and eDNA in freshwater river systems are needed to inform management decisions and to better establish the efficacy of eDNA as a potential tool for monitoring.

The goal of this study was to compare eDNA and UVC surveys for monitoring the spatial distribution of naturally spawned juvenile Coho Salmon *Oncorhynchus kisutch* in the Smith River, California. Coho salmon inhabiting the Smith River are part of the Southern Oregon/Northern California Coast Evolutionarily Significant Unit and are currently listed as threatened under the US Endangered Species Act. The California Department of Fish and Wildlife (CDFW) uses UVC methods (i.e., snorkeling) to determine the spatial distribution of juvenile Coho Salmon in the Smith River each summer (Walkley and Garwood 2017), and we integrated eDNA collections into CDFW's pre-existing survey protocols over two survey seasons (2020-2021). The detection capabilities shown in the CDFW snorkeling surveys are very high (>0.9; Walkley and Garwood 2017) and create an optimal environment for precise benchmarking of eDNA methods. This study had two primary objectives: (1) to compare the ability of eDNA and UVC surveys to detect Coho Salmon and assess the influence of environmental covariates on detection probabilities and (2) to evaluate the potential for using eDNA concentrations and habitat covariates to predict the count of Coho Salmon within pools.

Methods

Study site

The Smith River drains a watershed spanning 1,862 km^2 of northern California and southern Oregon (Figure 1). Nearly all of the basin (98%) is within the Klamath and Siskiyou Mountains. The Smith River is the largest free-flowing coastal river in California; there are no dams and anadromous fishes have access to the entire basin (Garwood and Larson 2014; Walkley and Garwood 2017).

Field methods

During July and August of 2020 and 2021, eDNA sampling was conducted in conjunction with the annual UVC survey of the Smith River basin executed as part of the CDFW Coastal Salmonid Monitoring Program (Adams et al. 2011; Garwood and Ricker 2016; Walkley and Garwood 2017). For the UVC survey, sampling units of approximately 1-3 km in length (hereafter reaches) are randomly selected out of 166 total sampling units that represent all juvenile salmonid-rearing habitat in the Smith River basin (Figure 1; Garwood and Larson 2014, Adams et al. 2011). The eDNA samples were collected in all non-mainstem UVC survey reaches in 2020 and in a subset of the UVC reaches in 2021, yielding a total of 26 reaches (19 in 2020 and 7 in 2021). One reach was surveyed in both years but was treated as two independent samples because Coho Salmon typically spend only one summer rearing in streams (Brown et al. 1994). In 2021, eDNA sampling occurred selectively in reaches with historically higher observed counts of juvenile Coho Salmon to increase the sample size of occupied reaches and to increase the range of observed fish counts in the data set, which were relatively low in 2020.

Systematic sampling was used to select survey pools within a survey reach. Only pools that met minimum habitat criteria, as described in Garwood and Ricker (2016), were included in the survey. A coin flip decided which of the first two pools was the starting point of the survey, after which every alternate pool was surveyed. Water samples for eDNA analysis were collected at the deepest point at the downstream end of all double-pass survey pools with some exceptions (see below). All water samples were collected before divers entered the water, by personnel dedicated to water sample collection (i.e., not divers). When reach lengths were greater than 2 km or when many (>30) pools were expected water samples were collected at every other double-pass survey pool.

Divers surveyed pools by conducting two independent census counts of the first sampled pool and then every alternating upstream pool was surveyed systematically with the next three surveyed pools only getting a single pass; this sequence (i.e., 2-0-1-0-1-0-1-0) was repeated for the remainder of the reach. For each double pass pool, two divers independently and sequentially surveyed the pool, allowing approximately five minutes between dives for disturbed sediment in the pool to settle. When surveying a pool, divers proceeded upstream, examined the entire width of the pool, and recorded the number of juvenile Coho Salmon present. Divers also recorded the number of large woody debris (LWD; >30 cm in diameter), the residual pool depth (RD; the difference in height between the deepest point in a pool and the downstream riffle crest; Lisle 1987), the total pool length, and the average pool width. Additionally, the contributing basin area (BA) to each survey reach was used as a proxy for river discharge as the two measures were assumed to scale geometrically. The BA values were obtained using the StreamStats application (U.S Geological Survey 2016) and assumed to be constant for a given survey reach. On average, reaches were 2.1 km in length and there were four double-pass pools per survey reach (max of 11, min of 1). Within a reach, individual survey pools were, on average, 17 m long, 6 m wide, 157 m apart, and the double-pass pools were 536 m apart. LWD counts ranged from 0-11

structures per pool, RD ranged from 1-320 cm per pool, and BA ranged from 0.26-155 km^2 per reach (Figure S1).

At each double-pass survey pool, three 1-liter water samples were collected using single-use Whirl-Pak bags (Nasco). Water samples were obtained by drawing a Whirl-Pak bag along the water's surface. All samples were filtered in the field immediately after collection, across 0.45-micron cellulose nitrate filters (Cytvia), held in filter funnels (Thermo Scientific™ Nalgene™ Single-Use Analytical Filter Funnels). Filter funnels were held in a filtration manifold that allowed up to four samples to be filtered simultaneously using a manual vacuum pump. Filter support pads (MilliporeSigma™) were used to ensure equal filtration across the surface of the filter. A field blank was collected at least once per survey day by filtering 1 liter of store-bought drinking water. Field blanks were processed the same as the other samples and served as comprehensive contamination controls. After filtration, filters were placed into 2 ml microcentrifuge tubes (Eppendorf) containing 360 µL of cell lysis buffer (QIAGEN, buffer ATL). Samples remained unfrozen for a maximum of three days post-filtration due to the remote nature of some survey locations but were stored at -20℃ upon returning from the field.

[B]*Molecular methods.*— The DNA was extracted directly from filters using the QIAGEN DNeasy Blood and Tissue Kits following the manufacturer's instructions with three exceptions: 1) we used 360 µl of buffer ATL and 40 µl proteinase K (Schmelzle and Kinziger 2016), 2) QIAshredders (QIAGEN) were used to ensure lysate homogenization, and 3) during the final elution step, 100 µl of elution buffer was used to increase the final DNA concentration of the elution. All extractions were completed within three months of field collection and extracted DNA was stored at -20 ℃.

The concentration of eDNA in a sample was estimated using droplet digital PCR (ddPCR) with the Bio-Rad QX200 Droplet Digital PCR System. Each ddPCR reaction was run in duplex using assays for Coho Salmon (developed by Pilliod and Laramie (2016), modified by Spence et al. 2021) and Chinook salmon (*Oncorhynchus tshawytscha;* [Unpublished, U.S. Forest Service National Genomics Center for Wildlife and Fish Conservation at the Rocky Mountain Research Station, Missoula, Montana]). Assay specificity and sensitivity were verified by testing against co-occurring species: Coho Salmon, Chinook salmon, steelhead *Oncorhynchus mykiss,* and Coastal Cutthroat Trout *Oncorhynchus clarkii clarkii*. Fluorescence plots indicated signal interference in the Chinook salmon channel when the Coho Salmon assay was present, leading to the exclusion of the Chinook data from any subsequent analysis or consideration.

Each ddPCR reaction mix was comprised of 900 nM forward primer, 900 nM of reverse primer, 250 nM probe, 5 µl of ddPCR Multiplex Supermix, 0.2 µl of 300 mM dithiothreitol, 15 µl of DNA template to maximize the probability of target DNA presence in the analyte (Rees et al. 2014; Doi et al. 2015), and DNA-free water to bring the total volume to 22 µl. Each reaction mix contained equal amounts of primers and probes for both Coho Salmon and Chinook salmon. Then, for each sample, 20 µl of the total reaction mix and 70 µl of Bio-Rad droplet generator oil

were placed into individual wells of a Bio-Rad DG8 cartridge and placed into the Bio-Rad QX-200 droplet generator which partitions the reaction mix into as many as 20,000 droplets for PCR amplification. Partitioned samples were transferred to 96 well ddPCR plates, sealed with a PX1 PCR plate sealer and placed in an MJ Research PTC-100 Thermal Cycler. Thermocycling conditions were 10-minutes at 95℃ followed by 40 cycles of 30-seconds at 94℃ and 60-seconds at 60℃, and concluded with 10-minutes at 98℃, and holding at 4℃. The temperature ramp rate was set to 2°C between all steps. The number of positive and negative droplets in each reaction were determined using a QX200 droplet reader. Concentrations are reported as the Poisson-corrected copies per reaction as estimated by the Bio-Rad QX Manager Software Version 1.2. Each ddPCR plate run contained at least one positive control (genomic DNA extracted from the tissue of the target species) and at least one negative control (containing all reagents except DNA template, which was replaced with DNA-free water). Each water sample was analyzed only a single time (i.e., single technical replicate) unless the results showed signs of anomalous fluorescence patterns or low droplet counts. When this occurred, the sample was re-run, and the updated results were used.

All DNA extractions and ddPCR setups were conducted in a dedicated low-copy eDNA laboratory. All work surfaces and extraction tools (i.e., benches, centrifuges, and racks) were sterilized with UV light and researchers could not enter if they had entered the separate dedicated lab designated for any high concentrations of DNA (e.g., from running PCR reactions).

Limits of detection (LOD) and quantification (LOQ) were determined using serially diluted genomic DNA extracted from fin clips using Qiagen DNeasy Blood and Tissue Kits. The LOD, defined as the lowest concentration of DNA resulting in at least 95% positive detections, and LOQ, defined as the lowest concentration with a coefficient of variation below 35%, were determined using curve fitting methods. Positive detections were indicated by samples for which measured concentrations of Coho Salmon eDNA exceeded the LOD (7 DNA copies per reaction). Estimates of eDNA concentration that exceeded the LOQ (47 DNA copies per reaction) were considered to be accurate measures.

Occupancy analysis

Multi-scale occupancy models were used to estimate and compare the method-specific detection probabilities for eDNA and UVC. Models were fitted in a maximum-likelihood framework using the software PRESENCE (version 2.13.10; Hines 2006). The occupancy model included three parameters at different hierarchical levels: Psi (Ψ) is the probability of species occurrence in a river reach; theta (θ_t) describes the probability of the species occurrence in any given pool t of the larger survey reach which is conditional on the species being present within the reach; and $p_{m,t}$ describes the probability of the species being detected by survey method m in pool t of the survey reach, conditional upon the species being present in both the reach and the pool. This analysis only included data from the double-pass survey pools with both eDNA and

UVC observations. A detection history for a survey pool would have five digits, with the first two numbers representing detections from each of the two dive passes and the last three numbers representing detections from the three eDNA samples. For example, a detection history of 11101 in a survey pool would indicate that Coho Salmon were detected in the pool by both divers and in eDNA water samples one and three, and a detection history of 00000 would indicate that neither method detected Coho Salmon. Note that this parameterization differs from some other applications of hierarchical modeling of eDNA in which p for eDNA is defined as the probability of detecting Coho Salmon DNA in a replicate quantitative PCR (qPCR) run within a single water sample and θ is the probability that the water sample contains Coho Salmon DNA (Schmidt et al. 2013; Schmelzle and Kinziger 2016; Dorazio and Erickson 2017; Spence et al. 2021).

The analysis was structured to assess the influence of several covariates that potentially affect detection of Coho Salmon by either or both survey methods. We hypothesized that UVC detection probability would be reduced by increasing RD and LWD due to difficulties in observing individuals in deeper water or with visual obstructions (e.g., Thurow et al. 2006; Staton et al. 2022) and that eDNA detection probability would be reduced by increasing BA due to the dilution of eDNA particles (Baldigo et al. 2017). For this analysis, both RD and BA were $\log_{10}$ transformed. Finally, a covariate for year was included to account for possible differences in detection probabilities between years. Turbidity was not included as a covariate because it was consistently very low, and it was not expected to have inhibited a diver's ability to detect the target species. Count was also considered as a covariate but was not included in the final analysis, due to model conversion issues. Note that we did not assess the influence of covariates on Ψ or θ because our focus was to compare detection probabilities rather than determine species occupancy patterns and because site selection for eDNA methods was not random in 2021.

Model selection was done using Akaike's information criterion (AIC) with a suite of 23 models defined by nearly all possible combinations of the covariates (method, BA, RD, LWD, year). Aside from a null model, where the detection probabilities were constant, 'method' was included in every model fit. Covariates were included individually and with interactions by method (with the exception of a year-method interaction). Occupancy models were ranked using AIC, AIC differences (ΔAIC), and AIC weights (Burnam and Anderson 2002). All models within a ΔAIC of 2 were considered as competing models supported by the data.

For the best model, we generated response plots to assess the influence of each covariate on detection probabilities. Predictions were made over the observed range of values for a covariate while all other covariates were held at their median values. A Monte Carlo approach was used to approximate the standard error (SE) of the estimated detection probabilities. This was done by taking 1000 random samples of coefficients from a multivariate normal distribution defined by the estimated coefficients and their variance-covariance matrix using the MASS package in R version 4.0.5 (Venables and Ripley 2002; R Core Team 2021). Each set of

randomly drawn coefficients was used to generate a response curve for each covariate (while holding the other covariates at their medians). The SE for the response plots was approximated using the distribution of 1000 Monte Carlo predictions generated for each covariate value.

The cumulative probability of detection (p^*) was calculated as $p^* = 1 - (1-p_t)^n$, where n is the number of replicate water samples taken from a pool that contained Coho Salmon DNA. Cumulative detection probabilities were calculated for the highest, median, and lowest estimates of p_t for the sampled pools based on the observed covariates and the best occupancy model. Plots of p^* as a function of sample size were analyzed to estimate the sampling effort required to detect Coho Salmon DNA with 95% cumulative probability under the observed values of covariates (McArdle 1990).

Concentration-count analysis

To assess the potential for estimating abundance of Coho Salmon in the absence of visual counts, we fit a generalized linear model (GLM) that related the average of the two counts from the double-pass survey pools (rounded to the nearest integer), to the natural log (ln) of the average of the three eDNA concentrations [ln(copies/reaction)] and the three habitat covariates: LWD [count], RD [log10(cm)] and BA [log10(km2)]. To account for variation in pool size, an offset of the natural log of pool area was also included, where pool area was calculated using the product of the maximum pool length and the average pool width. This analysis excluded data from pools in which no fish were observed. Initial analysis of models including all variables indicated that a zero-truncated, generalized Poisson error distribution yielded superior fits compared to zero-truncated Poisson or Negative Binomial distributions based on AICc and residual variance diagnostics (Zuur et al. 2009). A total of 32 models with all possible combinations of the covariates and the area offset were fit and compared by AICc using packages *glmmTMB* (Brooks et al. 2017) and *MuMIn* (Bartoń 2020). Model diagnostics (e.g., residual variance, overdispersion, data distribution) were assessed using package *performance* (Lüdecke et al. 2021). The best model was used to estimate the effects of the covariates on mean fish count while holding all other covariates at their median values.

Results

Survey results

A total of 114 double-pass pools and 318 single-pass pools distributed among 25 reaches were surveyed in 2020 and 2021. Of the double-pass pools, 96 were surveyed with both eDNA and UVC methods. Diver counts of juvenile Coho Salmon ranged from 0-210 fish per pool (mean = 32; Figure S1). The difference between the two independent dive counts ranged from 0 to 81 (mean = 4), but there was 100% agreement between divers regarding whether a pool was

occupied or not. None of the field blanks or negative PCR controls tested positive for Coho Salmon eDNA. All of the positive controls were positive for Coho Salmon eDNA.

Survey methods yielded identical results with respect to the detection and non-detection of Coho Salmon in 93% (89 of 96) of pools in which both surveys were conducted (27 presences; 62 absences; Table 1), indicating strong agreement between the two methods (Pearson's Chi-squared test; X^2=63.00, p<<0.001). Similarly, the two methods agreed at 80% (20 of 25) of survey reaches (7 presences; 13 absences; Table 1) and the relationship was statistically significant (Pearson's Chi-squared test with simulated p-value based on 2000 replicates; X^2=7.77, p=0.013). Coho salmon eDNA was detected in three reaches where none were observed, and Coho Salmon were observed in three (different) reaches where no eDNA was detected.

Occupancy modeling

Five of the 23 occupancy models examined had a ΔAIC less than two (Table 2) and were far more supported than the null model with no covariates for p (ΔAIC of 19.96). The top model (AIC Weight = 0.25) included survey method, RD, BA, and a method-BA interaction as covariates for detection probability. Survey method, RD, and BA were included in all five models while all other covariates were less consistent (Table 2). For plotting the response, we focused on the model with the highest model weight for simplicity.

At median environmental conditions, the estimated detection probability of eDNA (p_{eDNA}=0.91; 95% CI: 0.80 - 0.96), was very similar to the estimated detection probability of UVC (p_{UVC}=0.89; 95% CI: 0.79 - 0.94; Figure 2A). Estimated detection probabilities for both methods increased with RD; p_{UVC} ranged from $0.61 - 0.94$ and p_{eDNA} from $0.68 - 0.96$ (Figure 2B). Although the large standard errors of predictions at low RD values were strongly influenced by a single data point, the predicted response was unaffected by exclusion of this observation. Increasing BA had a strong negative effect on both methods, but the effect was more pronounced for eDNA; p_{eDNA} declined rapidly from 0.99 to 0.34 after values of $\log_{10}$(BA) greater than one (i.e., BA > 10 km^2), while p_{UVC} declined more gradually from 0.98 to 0.77 (Figure 2B). Across the surveyed pools, the range of the site-specific estimates of detection probability for p_{eDNA} values (0.99 to 0.13) varied more widely than p_{UVC} (0.97 to 0.42).

Under our sampling protocol, triplicate water samples were sufficient to achieve cumulative detection probabilities >95% at nearly all (98%) of the surveyed pools. Given the presence of Coho Salmon DNA in a pool with median values of RD and BA, it would be detected with high probability (p*= 0.91, 95% CI: 0.79 - 0.95) in a single water sample, and with >95% probability in two samples (Figure 3). Under conditions that strongly suppress the detection of eDNA (i.e., high BA or low RD), twenty or more samples would be required to achieve >95% cumulative probability of detection.

Concentration-count analysis

GLM models explained observed fish counts in pools well, but eDNA was not an important predictor in the best model. Competing GLMs (with ΔAIC<5) consistently included the pool-area offset and basin area as covariates, whereas LWD, eDNA, and RD only occurred in models with ΔAICc > 2 (Table 3). The best model had a model weight of 47% (Table 3) and had a strong correlation between observed and predicted fish counts in pools (R^2=0.83). Fish count was predominantly driven by a positive relationship with pool area that accounted for larger pools being able to hold more fish at a given fish density (all else being equal), but there was also strong evidence of lower counts in reaches with larger basin areas (Figure 4), which is consistent with juvenile coho habitat preferences for small streams (Brown et al. 1994).

Discussion

In this study, we demonstrate that eDNA methods produce very similar results to UVC surveys known to have exceptionally high detection probabilities for juvenile Coho Salmon in our research system. Our findings suggest that these eDNA methods have the potential to substantially enhance sensitivity in settings where species detection rates are much less than 1, which is common in most field surveys (Fediajevaite et al. 2021). We found that p_{eDNA} was effectively equivalent to p_{UVC} in reaches with contributing basins < 36 km^2, but p_{eDNA} declined sharply in reaches with larger basin areas and with presumably higher discharge. Our results also corroborate other studies, demonstrating that eDNA is a highly effective method for detecting rare species (McKelvey et al. 2016; Rice et al. 2018; Strickland and Roberts 2019; Sutter and Kinziger 2019, Fediajevaite et al. 2021, Spence et al. 2021). However, additional investigations are necessary to determine whether eDNA methods could provide a reliable tool for predicting juvenile Coho Salmon abundance in small pools.

This study demonstrated that eDNA methods have the capability to yield similar estimates of Coho Salmon spatial distribution with less overall sampling effort. Specifically, UVC and eDNA methods agreed on the presence or absence of Coho Salmon in 80% of surveyed reaches, despite differences in sampling intensity (432 pool surveys with UVC; 96 with eDNA). These results align with earlier investigations conducted by Evans et al. (2017) and Yu et al. (2021) that showed eDNA methods require less sampling effort than some conventional survey methods.

The best occupancy model from our analysis indicates that both methods had a high probability of detecting Coho Salmon in a pool (~90% in typical settings), but those detection probabilities were influenced strongly by BA and weakly by RD. In general, the eDNA detection probabilities were high and greater than that of UVC until the median BA (18 km^2), at which point eDNA detection probabilities began to decline more rapidly than those of UVC. These results are consistent with the hypothesis that increasing discharge, for which BA is a proxy, would decrease the probability of capturing and detecting rare organismal DNA due to the

dilution of particles (Levi et al. 2019, Pochardt et al. 2020). However, the effect of discharge on observed eDNA concentrations can be complex and dependent upon numerous factors that may vary across time and space (e.g., eDNA plume dynamics, increased turbidity and presence of inhibitors, eDNA particle settling and resuspension dynamics; Jane et al. 2015, Wilcox et al. 2016; Matter et al. 2018, Wood et al. 2021, Van Driessche et al. 2023). In this study, the negative BA effect for eDNA may have been more substantial because the eDNA sampling effort per pool was fixed (three samples per pool), whereas the sampling effort for UVC (in terms of area surveyed) was commensurate with pool size. The slight decrease in method-specific detection probabilities at lower values of RD may have resulted from faster water velocities in survey pools with low values of RD which could have hindered UVC divers and diluted available eDNA as these sites were more similar to runs or riffles; for example, Wood et al. (2021) found that the amount of available eDNA in the midstream water column was generally lower and more variable in areas with high velocity than where velocities were low.

In river systems, eDNA is carried from the source organism and can be detected downstream if eDNA concentrations remain sufficiently high (Goldberg et al. 2016). Although this eDNA transport can enhance the detectability of a target species throughout an extended spatial area, it undermines the spatial independence of samples assumed within occupancy models (MacKenzie et al. 2018). In the present study, however, the spacing between pools surveyed for eDNA consistently exceeded scales of eDNA transport reported for streams of similar size (survey spacing: 536 m; 95% C.I.: 438 - 635 m, detection ranges of ~200 m [Spence et al. 2021; see also Jo and Yamanaka 2022]), supporting our modeling assumption of independence across sampling locations. It is possible that some eDNA detections could have resulted from fish that were upstream of an unoccupied pool and thus not available for detection via UVC. However, the strong method-specific agreement in pool-level detections suggests this was not a common occurrence nor a substantial source of bias in the occupancy models. Of the few cases where eDNA detected Coho Salmon and UVC did not, fish were not typically observed by divers in the single-pass pools immediately upstream of the eDNA sample pool; in the singular case where this happened, Coho Salmon were detected a substantial distance upstream (300 m) in low numbers (n=3).

Although our capacity to use eDNA in water samples to estimate abundance is still developing, a growing body of research demonstrates a positive correlation between eDNA concentration and biomass or abundance (Rourke et al. 2022, Takahara et al. 2012, Schmelzle and Kinziger 2016, Tillotson et al. 2018, Capo et al. 2020, Shelton et al. 2022). In the present study, however, we found that eDNA concentration was not a good predictor of the average count of Coho Salmon in a pool. We suspect that the limited variability in abundance and biomass of juvenile Coho Salmon across sampling units (2-210 individuals per pool [2-orders of magnitude]) contributed to our inability to resolve such a relationship. Studies that have identified strong relationships between eDNA concentration and abundance indices have done so when differences in abundance or biomass (ranging over 3 to 6 orders of magnitude) varied

substantially more than in our system (Tillotson et al. 2018; Yates et al. 2019; Pochardt et al. 2020; Sepulveda et al. 2021; Shelton et al. 2022).

This study and others suggest that eDNA surveys could be a suitable alternative or complement to UVC surveys, but more work is needed to develop robust and optimal sampling designs. Our protocol of collecting triplicate 1-liter water samples was sufficient to achieve a 95% probability of detecting Coho Salmon DNA in a pool (if present) in basins up to 70 km^2, but more sampling effort would be required to achieve similar confidence levels for detection of eDNA in sites with larger contributing basins. Instead of filtering substantially larger volumes of water, which can be logistically difficult in some systems (e.g., filter clogging, increased presence of inhibitors; Capo et al. 2020), future eDNA monitoring efforts should consider increasing the number of water samples commensurate with basin size to maintain high cumulative detection probabilities, or they could limit eDNA surveys to areas with acceptably high performance given sampling constraints. Monitoring programs that use eDNA also need to evaluate appropriate levels of spatial sampling effort (e.g., number of pools sampled within a reach) to achieve desired objectives, and this will depend on the pool-level occupancy rate and detection probability for the target species (see equations in Spence et al. 2021). Based on pool-level occupancy rates of juvenile Coho Salmon in the Smith River (0.47 ± 0.02 SE; Walkley and Garwood 2017) and our median estimated eDNA detection probabilities sampling five pools per reach (using three water samples and 1 ddPCR replicate) was sufficient to yield >95% cumulative probabilities of encountering and detecting Coho Salmon at the reach scale. In other California streams, Spence et al. (2021) found that sampling two locations per survey reach (using 3 water samples and 1-2 PCR replicates) resulted in an overall detection probability of 0.74 - 0.99, depending on Coho Salmon densities. Overall, eDNA can be an effective tool for detecting and monitoring fishes in rivers and streams but monitoring programs using this method should be designed and optimized to achieve desired objectives given case-specific detection and occupancy rates, environmental conditions, and sampling constraints.

Acknowledgements

References

Adams, P. B., L. B. Boydstun, S. P. Gallagher, M. K. Lacy, and T. McDonald. 2011. The California coastal salmonid monitoring plan: overall strategy, design, and methods. Pages 1–82. California Department of Fish and Wildlife, Technical Report.

Albanese, B., K. A. Owers, D. A. Weiler, and W. Pruitt. 2011. Estimating Occupancy of Rare Fishes Using Visual Surveys, with a Comparison to Backpack Electrofishing. Southeastern Naturalist 10(3):423–442.

Andruszkiewicz Allan, E., W. G. Zhang, A. Lavery, and A. Govindarajan. 2021. Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. Environmental DNA 3(2):492–514.

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(1):99–111.

Bartoń, K. 2020, April 15. MuMIn: Multi-Model Inference.

Brown, L. R., P. B. Moyle, and R. M. Yoshiyama. 1994. Historical Decline and Current Status of Coho Salmon in California. North American Journal of Fisheries Management 14(2):237–261.

Brooks, M. E., K. Kristensen, K. J. van Bentham, A. Magnusson, C. W. Berg, A. Nielsen, H. J. Skaug, M. Maechler, and B. M. Bolker. 2017. glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. The R Journal 9(2):378–400.

Burnam, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd edition. Springer-Verlag, New York.

Capo, E., G. Spong, H. Königsson, and P. Byström. 2020. Effects of filtration methods and water volume on the quantification of brown trout (*Salmo trutta*) and Arctic char (*Salvelinus alpinus*) eDNA concentrations via droplet digital PCR. Environmental DNA 2(2):152–160.

Castañeda, R. A., A. Van Nynatten, S. Crookes, B. R. Ellender, D. D. Heath, H. J. MacIsaac, N. E. Mandrak, and O. L. F. Weyl. 2020. Detecting Native Freshwater Fishes Using Novel Non-invasive Methods. Frontiers in Environmental Science 8:29.

Doi, H., T. Takahara, T. Minamoto, S. Matsuhashi, K. Uchii, and H. Yamanaka. 2015. Droplet Digital Polymerase Chain Reaction (PCR) Outperforms Real-Time PCR in the Detection of Environmental DNA from an Invasive Fish Species. Environmental Science & Technology 49(9):5601–5608.

Dorazio, R. M., and R. A. Erickson. 2017. eDNAoccupancy: Multi-scale Occupancy Modeling of Environmental DNA data 18(2):368–380.

Dougherty, M. M., E. R. Larson, M. A. Renshaw, C. A. Gantz, S. P. Egan, D. M. Erickson, and D. M. Lodge. 2016. Environmental DNA (eDNA) detects the invasive rusty crayfish Orconectes rusticus at low abundances. Journal of Applied Ecology 53(3):722–732.

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(2):90–99.

Fediajevaite, J., V. Priestley, R. Arnold, and V. Savolainen. 2021. Meta-analysis shows that environmental DNA outperforms traditional surveys, but warrants better reporting standards. Ecology and Evolution 11:4803–4815.

Garwood, J. M., and M. D. Larson. 2014. Reconnaissance of salmonid redd abundance and juvenile salmonid spatial structure in the Smith River with emphasis on Coho Salmon (*Oncorhynchus kisutch*). California Department of Fish and Wildlife, 91251.

Garwood, J., and S. Ricker. 2016. Salmonid Spatial Structure Monitoring Survey Protocol: Summer Snorkel Survey Methods. California Department of Fish and Wildlife.

Goldberg, C. S., C. R. Turner, K. Deiner, K. E. Klymus, P. F. Thomsen, M. A. Murphy, S. F. Spear, A. McKee, S. J. Oyler‑McCance, R. S. Cornman, M. B. Laramie, A. R. Mahon, R. F. Lance, D. S. Pilliod, K. M. Strickler, L. P. Waits, A. K. Fremier, T. Takahara, J. E. Herder, and P. Taberlet. 2016. Critical considerations for the application of environmental DNA methods to detect aquatic species. Methods in Ecology and Evolution 7(11):1299–1307.

Gu, W., and R. K. Swihart. 2004. Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. Biological Conservation 116(2):195–203.

Hankin, D. G., and G. H. Reeves. 1988. Estimating Total Fish Abundance and Total Habitat Area in Small Streams Based on Visual Estimation Methods. Canadian Journal of Fisheries and Aquatic Sciences 45(5):834–844.

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

Hines, J. E. 2006. PRESENCE- Software to estimate patch occupancy and related parameters. USGS-PWRC.

Jane, S. F., T. M. Wilcox, K. S. McKelvey, M. K. Young, M. K. Schwartz, W. H. Lowe, B. H. Letcher, and A. R. Whiteley. 2015. Distance, flow and PCR inhibition: eDNA dynamics in two headwater streams. Molecular Ecology Resources 15(1):216–227.

Jo, T., and H. Yamanaka. 2022. Meta‑analyses of environmental DNA downstream transport and deposition in relation to hydrogeography in riverine environments. Freshwater Biology 67(8):1333–1343.

Laramie, M. B., D. S. Pilliod, and C. S. Goldberg. 2015. Characterizing the distribution of an endangered salmonid using environmental DNA analysis. Biological Conservation 183:29–37.

Levi, T., J. M. Allen, J. Joyce, J. R. Russell, D. A. Tallmon, S. C. Vulstek, C. Yang, and D. W. Yu. 2019. Environmental DNA for the enumeration and management of Pacific salmon. Molecular Ecology Resources 19(3):597–608.

Lisle, T. E. 1987. Using "Residual Depths" to Monitor Pool Depths Independently of Discharge. Page PSW-RN-394. U.S. Department of Agriculture, Forest Service, PSW-RN-394, Pacific Southwest Forest and Range Experiment Station.

Lovett, G. M., D. A. Burns, C. T. Driscoll, J. C. Jenkins, M. J. Mitchell, L. Rustad, J. B. Shanley, G. E. Likens, and R. Haeuber. 2007. Who needs environmental monitoring? Frontiers in Ecology and the Environment 5(5):253–260.

Lüdecke, D., M. Ben-Shachar, I. Patil, P. Waggoner, and D. Makowski. 2021. performance: An R Package for Assessment, Comparison and Testing of Statistical Models. Journal of Open Source Software 6(60):3139.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2018. Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence, 2nd edition. Elsevier.

Martel, C. M., M. Sutter, R. M. Dorazio, and A. P. Kinziger. 2021. Using environmental DNA and occupancy modelling to estimate rangewide metapopulation dynamics. Molecular Ecology 30(13):3340–3354.

Matter, A. N., J. A. Falke, J. A. López, and J. W. Savereide. 2018. A Rapid-Assessment Method to Estimate the Distribution of Juvenile Chinook Salmon in Tributary Habitats Using eDNA and Occupancy Estimation. North American Journal of Fisheries Management 38(1):223–236.

McArdle, B. H. 1990. When Are Rare Species Not There? Oikos 57(2):276–277.

McColl‑Gausden, E. F., A. R. Weeks, R. A. Coleman, K. L. Robinson, S. Song, T. A. Raadik, and R. Tingley. 2021. Multispecies models reveal that eDNA metabarcoding is more sensitive than backpack electrofishing for conducting fish surveys in freshwater streams. Molecular Ecology 30(13):3111–3126.

McElhany, P., M. H. Ruckelshaus, M. J. Ford, T. C. Wainwright, and E. P. Bjorkstedt. 2000. Viable salmonid populations and the recovery of evolutionarily significant units. Page 156. NMFS-NWFSC-42.

McKelvey, K. S., M. K. Young, W. L. Knotek, K. J. Carim, T. M. Wilcox, T. M. Padgett‑Stewart, and M. K. Schwartz. 2016. Sampling large geographic areas for rare species using environmental DNA : a study of bull trout *Salvelinus confluentus* occupancy in western Montana. Journal of Fish Biology 88(3):1215–1222.

Mize, E. L., R. A. Erickson, C. M. Merkes, N. Berndt, K. Bockrath, J. Credico, N. Grueneis, J. Merry, K. Mosel, M. Tuttle‑Lau, K. Von Ruden, Z. Woiak, J. J. Amberg, K. Baerwaldt, S. Finney, and E. Monroe. 2019. Refinement of eDNA as an early monitoring tool at the landscape‑level: study design considerations. Ecological Applications 29(6).

Nichols, J. D., L. L. Bailey, A. F. O'Connell Jr., N. W. Talancy, E. H. Campbell Grant, A. T. Gilbert, E. M. Annand, T. P. Husband, and J. E. Hines. 2008. Multi-scale occupancy estimation and modeling using multiple detection methods. Journal of Applied Ecology 45(5):1321–1329.

Nichols, J., and B. Williams. 2006. Monitoring for conservation. Trends in Ecology & Evolution 21(12):668–673.

Pilliod, D. S., and M. B. Laramie. 2016. Salmon Redd Identification Using Environmental DNA (eDNA). USGS.

Pochardt, M., J. M. Allen, T. Hart, S. D. L. Miller, D. W. Yu, and T. Levi. 2020. Environmental DNA facilitates accurate, inexpensive, and multiyear population estimates of millions of anadromous fish. Molecular Ecology Resources 20(2):457–467.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rees, H. C., K. Bishop, D. J. Middleditch, J. R. M. Patmore, B. C. Maddison, and K. C. Gough. 2014. The application of eDNA for monitoring of the Great Crested Newt in the UK. Ecology and Evolution 4(21):4023–4032.

Rice, C. J., E. R. Larson, and C. A. Taylor. 2018. Environmental DNA detects a rare large river crayfish but with little relation to local abundance. Freshwater Biology 63(5):443–455.

Rodríguez‑Ezpeleta, N., L. Zinger, A. Kinziger, H. M. Bik, A. Bonin, E. Coissac, B. C. Emerson, C. M. Lopes, T. A. Pelletier, P. Taberlet, and S. Narum. 2021. Biodiversity monitoring using environmental DNA. Molecular Ecology Resources 21(5):1405–1409.

Rourke, M. L., A. M. Fowler, J. M. Hughes, M. K. Broadhurst, J. D. DiBattista, S. Fielder, J. Wilkes Walburn, and E. M. Furlan. 2022. Environmental DNA (eDNA) as a tool for assessing fish biomass: A review of approaches and future considerations for resource surveys. Environmental DNA 4(1):9–33.

Schmelzle, M. C., and A. P. Kinziger. 2016. Using occupancy modeling to compare environmental DNA to traditional field methods for regional-scale monitoring of an endangered aquatic species. Molecular Ecology Resources 16(4):895–908.

Schmidt, B. R., M. Kéry, S. Ursenbacher, O. J. Hyman, and J. P. Collins. 2013. Site occupancy models in the analysis of environmental DNA presence/absence surveys: a case study of an emerging amphibian pathogen. Methods in Ecology and Evolution 4(7):646–653.

Sepulveda, A. J., R. Al-Chokhachy, M. B. Laramie, K. Crapster, L. Knotek, B. Miller, A. V. Zale, and D. S. Pilliod. 2021. It's complicated … environmental DNA as a predictor of trout and char abundance in streams. Canadian Journal of Fisheries and Aquatic Sciences 78(4):422–432.

Shelton, A. O., A. Ramón-Laca, A. Wells, J. Clemons, D. Chu, B. E. Feist, R. P. Kelly, S. L. Parker-Stetter, R. Thomas, K. M. Nichols, and L. Park. 2022. Environmental DNA provides quantitative estimates of Pacific hake abundance and distribution in the open ocean. Proceedings of the Royal Society B: Biological Sciences 289(1971):20212613.

Smith, M. M., and C. S. Goldberg. 2020. Occupancy in dynamic systems: accounting for multiple scales and false positives using environmental DNA to inform monitoring. Ecography 43(3):376–386.

Spence, B. C., D. E. Rundio, N. J. Demetras, and M. Sedoryk. 2021. Efficacy of environmental DNA sampling to detect the occurrence of endangered Coho Salmon *Oncorhynchus kisutch* in Mediterranean‑climate streams of California's central coast. Environmental DNA 3(4):727–744.

Staton, B. A., C. Justice, S. White, E. R. Sedell, L. A. Burns, and M. J. Kaylor. 2022. Accounting for uncertainty when estimating drivers of imperfect detection: An integrated approach illustrated with snorkel surveys for riverine fishes. Fisheries Research 249:106209.

Strickland, G. J., and J. H. Roberts. 2019. Utility of eDNA and occupancy models for monitoring an endangered fish across diverse riverine habitats. Hydrobiologia 826(1):129–144.

Sutter, M., and A. P. Kinziger. 2019. Rangewide tidewater goby occupancy survey using environmental DNA. Conservation Genetics 20(3):597–613.

Takahara, T., T. Minamoto, H. Yamanaka, H. Doi, and Z. Kawabata. 2012. Estimation of Fish Biomass Using Environmental DNA. PLoS ONE 7(4):e35868.

Thomsen, P. F., and E. Willerslev. 2015. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183:4–18.

Thurow, R. F. 1994. Underwater methods for study of salmonids in the intermountain West. USDA, Intermountain Research Station (Ogden, Utah).

Thurow, R. F., A. C. Dolloff, and E. J. Marsden. 2012. Visual Observation of Fishes and Aquatic Habitat. Pages 781–817 Fisheries Techniques, 3rd edition. American Fisheries Society, Bethesda, MD.

Thurow, R. F., J. T. Peterson, and J. W. Guzevich. 2006. Utility and Validation of Day and Night Snorkel Counts for Estimating Bull Trout Abundance in First- to Third-Order Streams. North American Journal of Fisheries Management 26(1):217–232.

Tillotson, M. D., R. P. Kelly, J. 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.

Turner, C. R., M. A. Barnes, C. C. Y. Xu, S. E. Jones, C. L. Jerde, and D. M. Lodge. 2014. Particle size distribution and optimal capture of aqueous macrobial eDNA. Methods in Ecology and Evolution 5(7):676–684.

U.S Geological Survey. 2016. The StreamStats program. U.S Geological Survey.

Valdivia‑Carrillo, T., Rocha‑Olivares, A., Reyes‑Bonilla, H., Domínguez‑Contreras, J. F., & Munguia‑Vega, A. (2021). Integrating eDNA metabarcoding and simultaneous underwater visual surveys to describe complex fish communities in a marine biodiversity hotspot. Molecular Ecology Resources, 21(5), 1558-1574.

Van Driessche, C., T. Everts, S. Neyrinck, and R. Brys. 2023. Experimental assessment of downstream environmental DNA patterns under variable fish biomass and river discharge rates. Environmental DNA 5(1):102–116.

Venables, W. N., and B. D. Ripley. 2002. Modern Applied Statistics with S, 4th edition. Springer, New York.

Walkley, J., and J. M. Garwood. 2017. 2011-2016 Salmonid Redd Abundance and Juvenile Salmonid Spatial Structure in the Smith River Basin, California and Oregon. Page 88. California Department of Fish and Wildlife.

Wilcox, T. M., K. S. McKelvey, M. K. Young, A. J. Sepulveda, B. B. Shepard, S. F. Jane, A. R. Whiteley, W. H. Lowe, and M. K. Schwartz. 2016. Understanding environmental DNA detection probabilities: A case study using a stream-dwelling char *Salvelinus fontinalis*. Biological Conservation 194:209–216.

Wood, Z. T., A. Lacoursière-Roussel, F. LeBlanc, M. Trudel, M. T. Kinnison, C. Garry McBrine, S. A. Pavey, and N. Gagné. 2021. Spatial Heterogeneity of eDNA Transport Improves Stream Assessment of Threatened Salmon Presence, Abundance, and Location. Frontiers in Ecology and Evolution 9:650717.

Yates, M., C., D. J. Fraser, and A. Derry M. 2019. Meta‑analysis supports further refinement of eDNA for monitoring aquatic species‑specific abundance in nature (1):5–13.

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(24):6531–6550.

Yu, D., Z. Shen, T. Chang, S. Li, and H. Liu. 2021. Using environmental DNA methods to improve detectability in an endangered sturgeon (*Acipenser sinensis*) monitoring program. BMC Ecology and Evolution 21(1):216.

Zuur, A., E. Ieno, N. Walker, A. Saveliev, and G. Smith. 2009. Mixed effects models and extensions in ecology with R, 3rd edition. Springer-Verlag.

Tables

TABLE 1. The percentage (and number) of survey pools and reaches in which coho salmon were detected (+) or not detected (-) by each survey method (i.e., eDNA and UVC). Numbers in parentheses indicate the number of pools or reaches. Pool comparisons are based on the 96 double-pass pools that were surveyed using both methods. Reach comparisons were calculated for the 25 reaches using an additional 318 pools where only UVC observations occurred, which reflects the current survey protocol.

eDNA Detection	UVC Detection	Pools	Reaches
+	+	28% (27)	28% (7)
-	-	65% (62)	52% (13)
+	-	5% (5)	8% (2)
-	+	2% (2)	12% (3)

TABLE 2. Top five occupancy models (with ΔAICc < 2) for juvenile Coho Salmon from the multi-method occupancy analysis. Covariates that were included (+) in each model are identified. Detection probability at the pool-level was modeled as a function of survey method (m), count of large woody debris (LWD), the $\log_{10}$ of residual pool depth (RD), the $\log_{10}$ of the contributing basin area (BA), Year (Yr), and interactions between the habitat covariates and method (e.g., BA * m). K represents the number of estimated parameters in the model. Differences in AICc values relative to the top-ranked model (ΔAICc) and model weights (W) are provided for all models.

Model Rank	m	LWD	RD	BA	Yr	LWD*m	RD*m	BA*m	K	ΔAICc	W
1	+		+	+				+	7	0	0.25
2	+		+	+					6	1.48	0.12
3	+	+	+	+		+	+	+	10	1.84	0.10
4	+		+	+			+	+	8	1.89	0.10
5	+		+	+	+			+	8	1.95	0.10

Table 3. Model selection table of generalized linear models (with ΔAICc < 5) for predicting non-zero counts of juvenile Coho Salmon in a pool. Models are ranked according to AICc, ΔAICc, and model weight. Covariates that were included (+) in each model are identified. Covariates included log10 of basin area (BA), large woody debris (LWD), the mean of the natural log transformed eDNA concentrations (eDNA [copies/reaction]), the residual pool depth (RD), and the natural log of pool area (PA) as an offset.

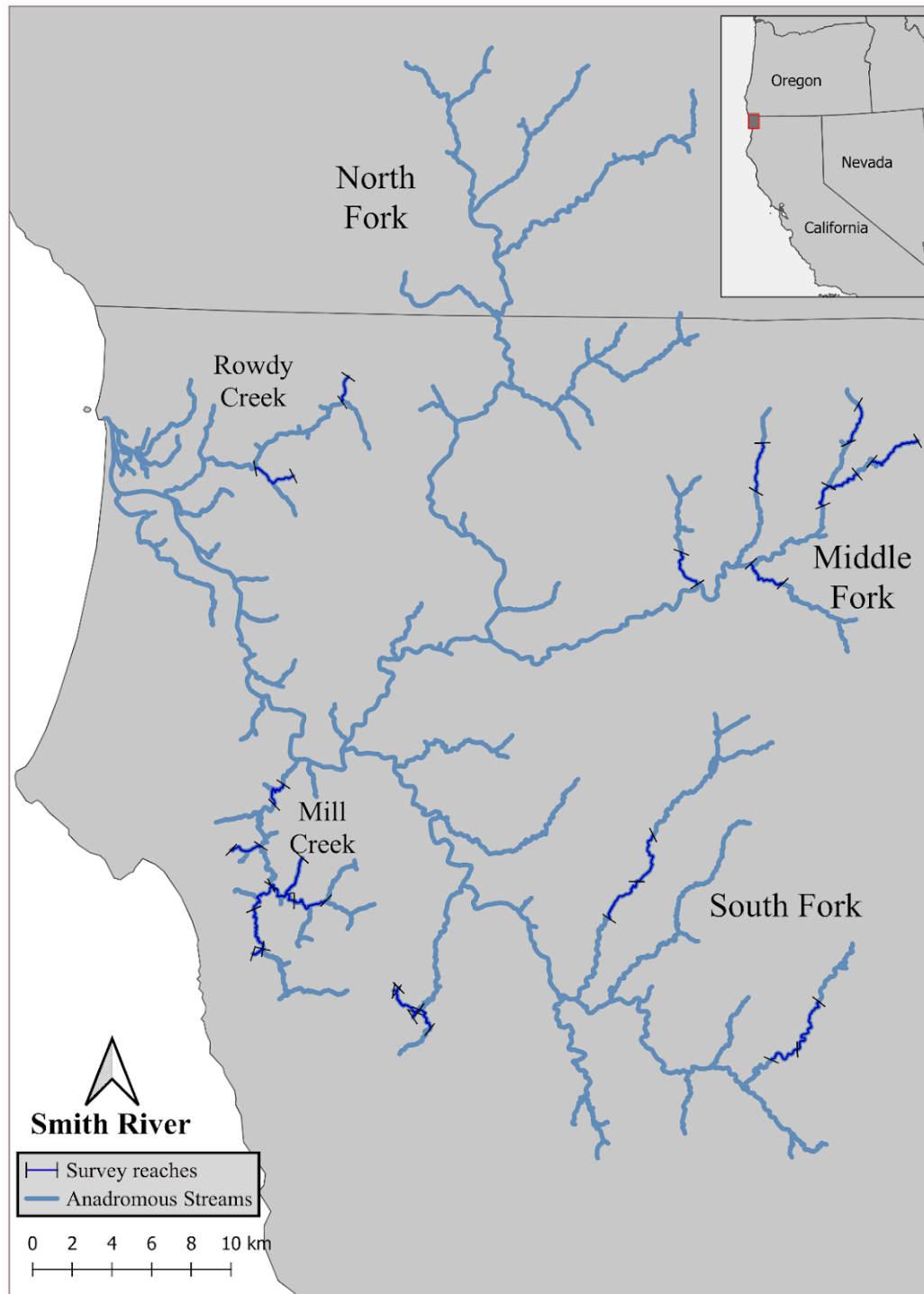| Model	BA	LWD	eDNA	RD	PA	df	AICc	ΔAICc	W
1	+				+	3	253.66	0.00	0.474
2	+	+			+	4	256.14	2.48	0.137
3	+		+		+	4	256.30	2.64	0.127
4	+			+	+	4	256.36	2.70	0.123

FIGURE 1. The anadromous rearing habitat (light blue lines) of the Smith River (California, USA) and the locations of the 25 surveyed stream reaches (dark blue lines) to compare the ability of eDNA and UVC surveys to detect Coho Salmon.
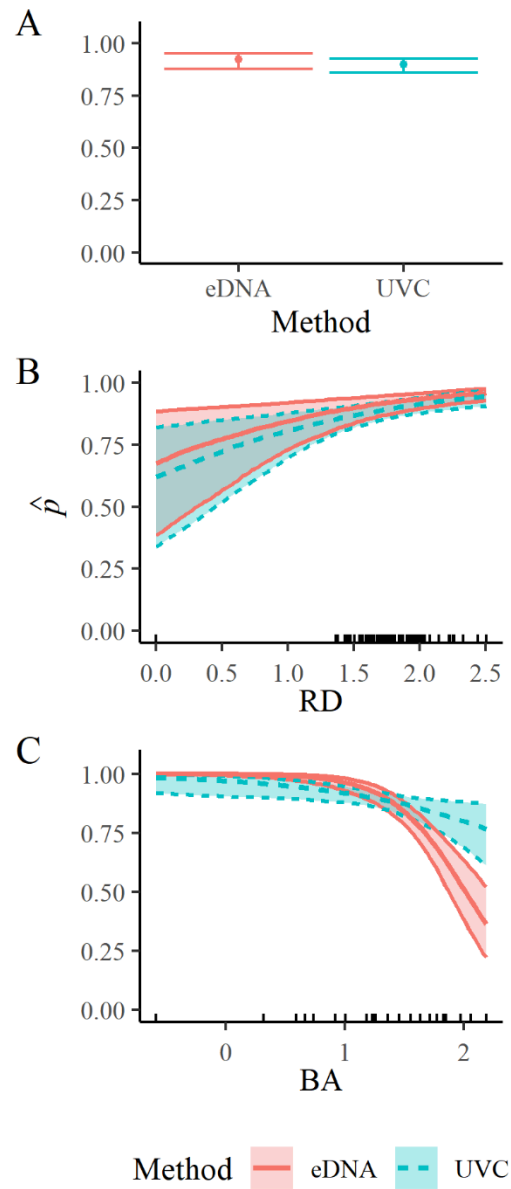
FIGURE 2. Predicted effects of **(A)** survey method, **(B)** $\log_{10}$ residual pool depth (RD), and **(C)** $\log_{10}$ of basin area (BA) on the detection probabilities (p) for eDNA (thick, solid, red line) and UVC (thick, dashed blue line) with the associated standard error (thin lines of the same color). Effect sizes were calculated over the observed range of values for the covariate, shown as ticks (i.e., rug), while all other covariates were held at their median values
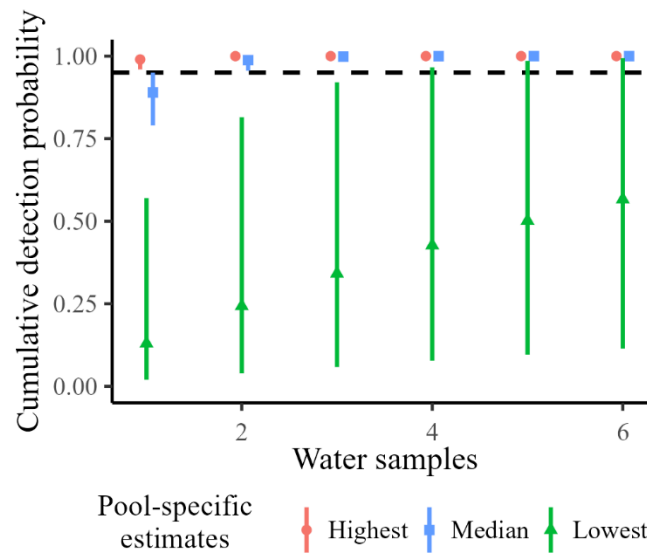
FIGURE 3. The cumulative detection probability as a function of the number of replicate water samples, calculated using the highest (p = 0.99), median (p = 0.89), and lowest (p = 0.13) estimated pool-specific detection probabilities. The vertical bars represent the 95% confidence interval. The horizontal dashed line represents the 95% cumulative detection probability.
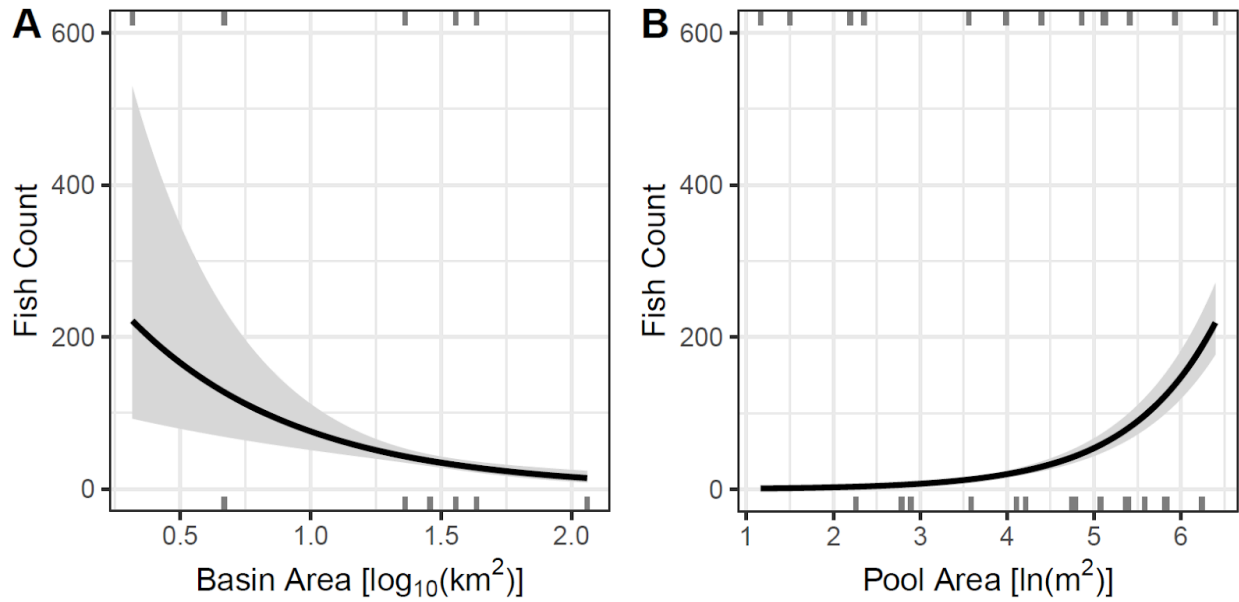
FIGURE 4. Predicted effects of (**A**) basin area and (**B**) an offset for pool area on the predicted mean count of Coho Salmon in a survey pool based on the top generalized linear model. Solid lines are the predicted effect of the variable (with 95% confidence interval) when all other model covariates are held at their median value. Ticks (i.e., rug) on horizontal axes denote location of positive and negative partial residuals.
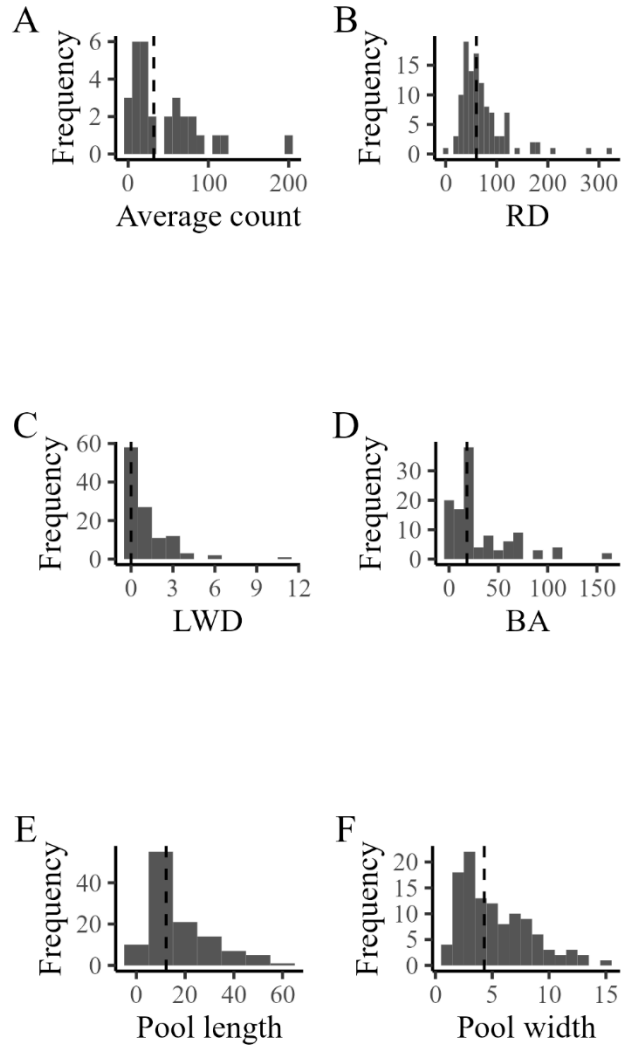
FIGURE S1. Frequency distributions of **(A)** the average count per occupied reaches of coho salmon, **(B)** the Residual Depth (RD in cm), **(C)** Large Woody Debris (LWD), **(D)** Basin Areas (BA in km^2), **(E)** Pool length (m), and (**F**) Pool width (m) observed in the 2020-2021 survey seasons. Median values are shown (vertical dashed lines).

Estimating fish abundance and eDNA production using the eDITH model

Introduction

Environmental DNA (eDNA), loose fragments of cellular material collected from environmental samples, is a promising new tool for determining species presence and, in some cases, abundance (Thomsen and Willerslev 2015; Yates et al. 2019; Rodríguez‑Ezpeleta et al. 2021; Rourke et al. 2022). In river systems, eDNA methods can have higher sensitivity and detection probability than conventional survey methods, especially when targeting rare or endangered species (Strickland and Roberts 2019; Spence et al. 2021; Yu et al. 2021). Furthermore, studies have found that eDNA concentrations can be strongly correlated with fish abundance and used to obtain estimates of biomass and density that are comparable to estimates derived from conventional survey techniques (see reviews in Yates et al. [2019] and Rourke et al. [2022]). Both abiotic and biotic factors influence the concentration and distribution of eDNA particles within a system, but recent studies suggest that the abiotic processes more strongly influence observed eDNA concentrations (Jerde et al. 2016, Shogren et al. 2017).

Recently, Carraro et al. (2018) developed the eDNA Integrating Transport and Hydrology model (hereafter eDITH model) that uses a Bayesian modeling framework and incorporates eDNA point measurements and hydro-morphological data collected throughout a river basin to estimate spatial distribution and relative density of a parasite (*Tetracapsuloides bryosalmonae*) and its primary host (the freshwater bryozoan *Fredericella sultana*). This modeling framework divides a dendritic system into a series of hierarchically organized "nodes" or river reaches that have homogeneous hydrologic conditions (e.g., velocity, depth, stream width) and target species density. The model then uses eDNA point observations throughout the system to develop a species distribution model and to predict species density throughout the basin. Estimates of species density are made based on the rate at which DNA attenuates once it is released from the source organism, and the model estimates a baseline eDNA production value which is a function of several environmental covariates to predict the density of organisms required to produce the observed eDNA concentration (Jerde et al. 2016; Carraro et al. 2018). Here, the eDNA attenuation rate is assumed to be the exponential loss of eDNA due to hydrologic transport (e.g., dilution, settling, absorption) and the physical breakdown or decay of genetic material due to UV radiation, temperature, enzyme activity; see Shogren et al. (2017) and Andruszkiewicz Allan et al. (2021) for more thorough discussion of these topics. For further discussion of model assumptions see Carraro et al. (2018).

We applied the eDITH modeling framework to field measurements of eDNA concentration at 15 locations throughout the Mill Creek basin, California, to obtain a basin-wide "snapshot" of threatened Coho Salmon (*Oncorhynchus kisutch*) distribution and abundance. Our goals were to 1) compare predicted eDNA concentrations and occupancy to our observed eDNA

data, and 2) examine the sensitivity of model predictions to changes in model parameters as these may vary considerably between systems.

Methods

Study area

The Mill Creek basin, a subbasin of the Smith River, is located in Del Norte County, California (Figure 1). It encompasses 95.8 km^2, ranges from 84 ft to 2,348 ft above sea level, and is known to have the highest densities of juvenile Coho Salmon in the entire Smith River basin (Walkley and Garwood 2017).

River network extraction

Following the methods of Carraro et al. (2018), the Mill Creek river network was extracted from a 1 m resolution digital elevation model of the region using a TauDEM implementation of the D8 method (O'Callaghan and Mark 1984; U.S. Geological Survey 2020). We selected a minimum threshold drainage area of 25 km^2 which resulted in a total river channel length of 43.5 km. The river network was then partitioned into discrete reaches or river sections. We tested five maximum reach lengths, ranging from 100 to 1000 m, to determine which would best reproduce the observed eDNA concentrations.

Field collection methods

Fifteen sampling locations were identified for a broad spatial distribution of the major areas of confluence where tributaries joined the main creek, and these sites were sampled over the course of three days in June 2021. At each sampling location, we collected a total of six 1-liter water samples using single-use Whirl-Pak bags (Nasco); three samples from the tributary, approximately 10 m above where it joined the main flow, and three samples from the main flow approximately 10 m upstream of the tributary junction. Water grabs were taken by drawing the bag along the surface, and water was filtered immediately in the field across 0.45-micron cellulose nitrate filters held in filter funnels. Filter funnels were held in a filtration manifold which allowed up to four samples to be filtered simultaneously using a manual vacuum pump. Filter support pads were used to ensure equal filtration across the surface of the filter. A field blank was collected at least once per survey day by filtering 1 liter of store-bought drinking water. Field blanks were processed the same as the other samples and served as comprehensive contamination controls. After filtration, filters were folded with sterilized forceps and placed into 2 ml microcentrifuge tubes containing 360 µL of cell lysis buffer. Samples remained unfrozen for a maximum of three days post-filtration due to the remote nature of some survey locations but were stored at -20℃ upon returning from the field. To prevent contamination, forceps were sterilized in a 10% bleach solution and new disposable gloves were worn when placing filters into storage vials.

Hydrologic data (e.g., width, depth, discharge) was collected at all eDNA collection sites and later used to develop a power-law relationship with drainage area as described in Carraro et al. (2018). This relationship was then used to extrapolate hydrologic data for all reaches throughout the basin based on drainage area. Following the methods of Carraro et al. (2018), water velocity was calculated for all reaches assuming rectangular river cross-sections (Equation 1) as:

$$v=q/(wd) \qquad\qquad \text{(Equation 1)}$$

where v is velocity (m/s), q is discharge (m^3/s, w is width (m), and d is depth(m).

Cal Poly Humboldt's Institutional Animal Care and Use Committee (IACUC, No. 2020F57E) approved all field protocols for this study.
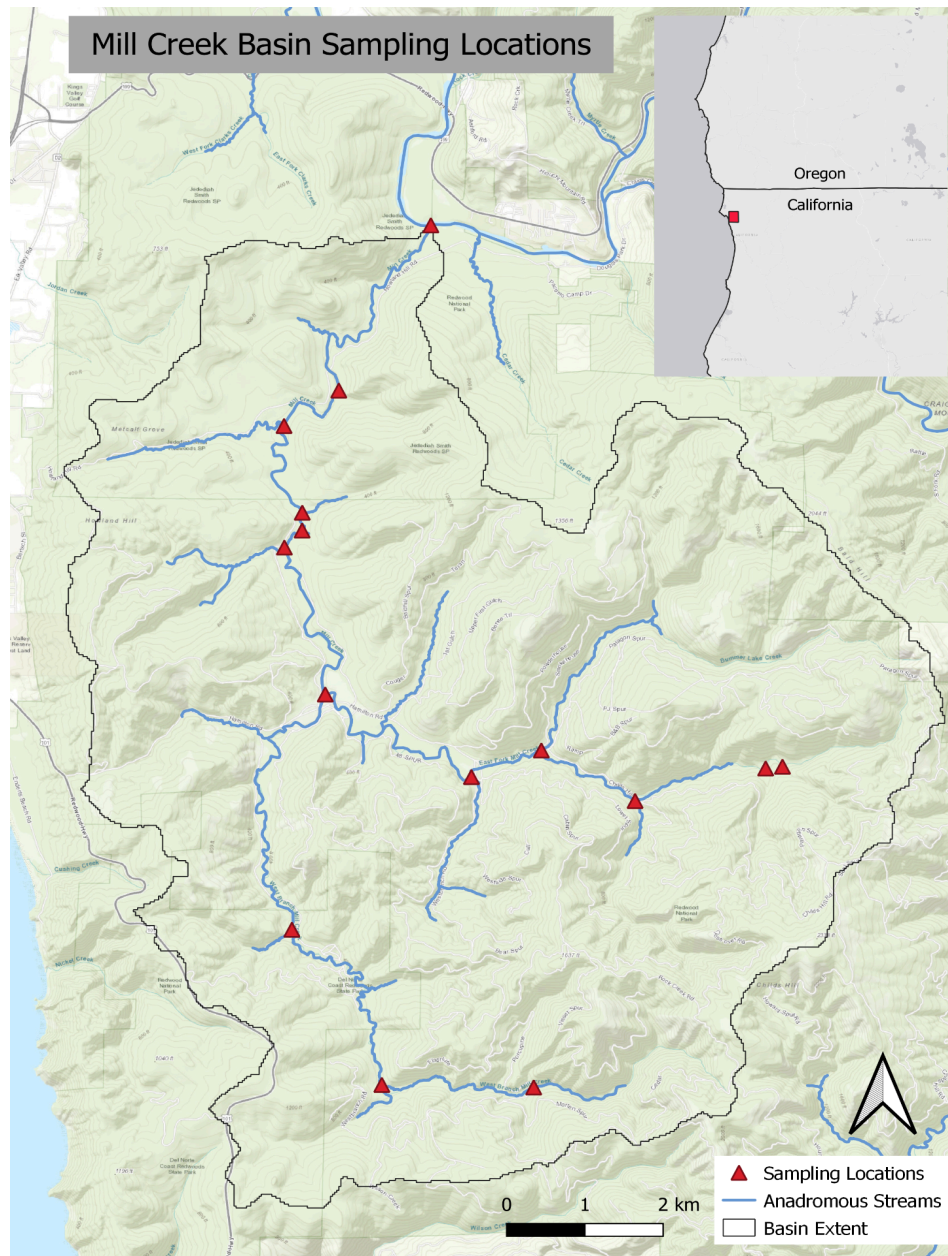

Figure 1. The 15 eDNA collection sites (triangles) visited in June 2021. The boundaries of the Mill Creek basin (black line), and the known anadromous salmonid habitat (blue line; based on CDFW survey reaches).

Molecular Methods

All molecular methods and limits of detection and quantification are the same as those described in the occupancy analysis (see Task 4 section).

eDITH Model

The eDITH model utilizes the continuous downstream transport of eDNA and predicts the eDNA concentration ($\overline{C}_j$ [NL^{-3}]) at node j as an additive function of all i nodes upstream of j (Carraro et al. 2018):

$$\overline{C}_j = \frac{1}{\overline{Q}_j} \sum_{i \in \gamma(j)} A_{S,i} \, exp(- \frac{L_{ij}}{\overline{v}_{ij} \tau}) p_i \qquad \text{(Equation 2)}$$

where $\overline{Q}_j$ [L^3 T^{-1}] is the model-derived discharge at node j; $\gamma(j)$ is the index of all connected nodes upstream of j; $A_{S,i}$ [L^2] is the contributing source area of the river network upstream of j; L_{ij} [L] is the stream distance between i and j; τ is the inverse of the DNA attenuation rate (i.e., a characteristic decay time); $\overline{v}_{ij}$ [LT^{-1}] is the average water velocity between i and j; finally, p_i is the derived eDNA production at node i which is assumed to be proportional to the target species density at node i. The eDNA production rates were derived from a Poisson generalized linear model:

$$p_i = p_0 \, exp(\beta^T X_{(i)}) \qquad \text{(Equation 3)}$$

where $X_{(i)}$ is a vector of environmental covariates, β^T is a transposed vector of parameters of the predicted effect of the environmental covariates, and p_0 is the base eDNA production rate. Here, we incorporated five hydrologic covariates derived from a digital elevation model of the catchment: drainage area, stream order, slope of survey reach, average upstream slope, and average elevation of the survey reach.

The inverse of the attenuation rate or attenuation time (τ), is the time required for an initial input of concentrated DNA to decrease past the point of detectability. Due to the difficulty in distinguishing between decay processes (e.g., UV exposure, thermal degradation, digestion, etc.) and particle deposition (e.g., absorption into sediment), we followed the methods of Carraro et al. (2018) and did not differentiate the two processes.

Due to the considerable variation in estimates of eDNA transport in the literature (e.g., Pont et al. 2018; Jo and Yamanaka 2022), we conducted a sensitivity analysis to examine how

eDITH model predictions were influenced by variation in eDNA attenuation times. We compared eDITH model predictions made using the DNA attenuation time (τ = 5 h) and reach length (1000 m) reported by Carraro et al. (2018), with values more similar to attenuation distances observed in other local eDNA studies (e.g., 450 m; Shaffer et al. 2022, Herman et al. 2023). Using an attenuation distance of 450 m and dividing by an assumed constant water velocity of 0.1 m/s (derived from field estimates of discharge), we obtained a local estimate of attenuation time (τ = 1). Using the two values for τ, and five reach lengths (e.g., 100 m, 250m, 500m, 750m, 1000m), we compared 10 models to the observed eDNA concentrations to identify the combination of parameters that best reproduced the observed data.

The posterior distributions of the unknown parameters were found by using the DREAM$_{ZS}$ algorithm from the *BayesianTools* R package (Hartig et al. 2023). A normal prior distribution with a null mean and standard deviation of 3 was used for all β coefficients; p_0 had a uniform prior distribution between 0 and 1; τ had a log-normal prior distribution with a median of one and mode of 0.5 based on the values derived from local field observations or a median of five and a mode of four as implemented by Carraro et al. (2018). For each parameter, three independent Markov chains were run with a chain length of $5 \cdot 10^6$, a burn-in of $1 \cdot 10^6$, and a thinning rate of 3.

Results

eDNA survey results

A total of 78 water samples for eDNA analysis were collected from 15 sites throughout the Mill Creek basin but only eight of the intended 15 tributaries were sampled due to a lack of in-stream flow (Figure 1, 2). The average observed eDNA concentrations from the sampled locations ranged from 0 – 8.8 copies per m^3 (Figure 2). The observed cross-sectional stream widths ranged from 1.5 – 13 m and measures of stream depth ranged from 0.5 – 52 cm. Field estimates of discharge ranged from 0.01 – 0.67 m^3/s. Within the West fork of Mill Creek, surveyors encountered an approximately 300 m section of river where flow became hyporheic before reemerging lower in the system. As the hyporheic transport of eDNA particles is currently not well understood, all model predictions from the West Fork were removed from consideration during model comparison.
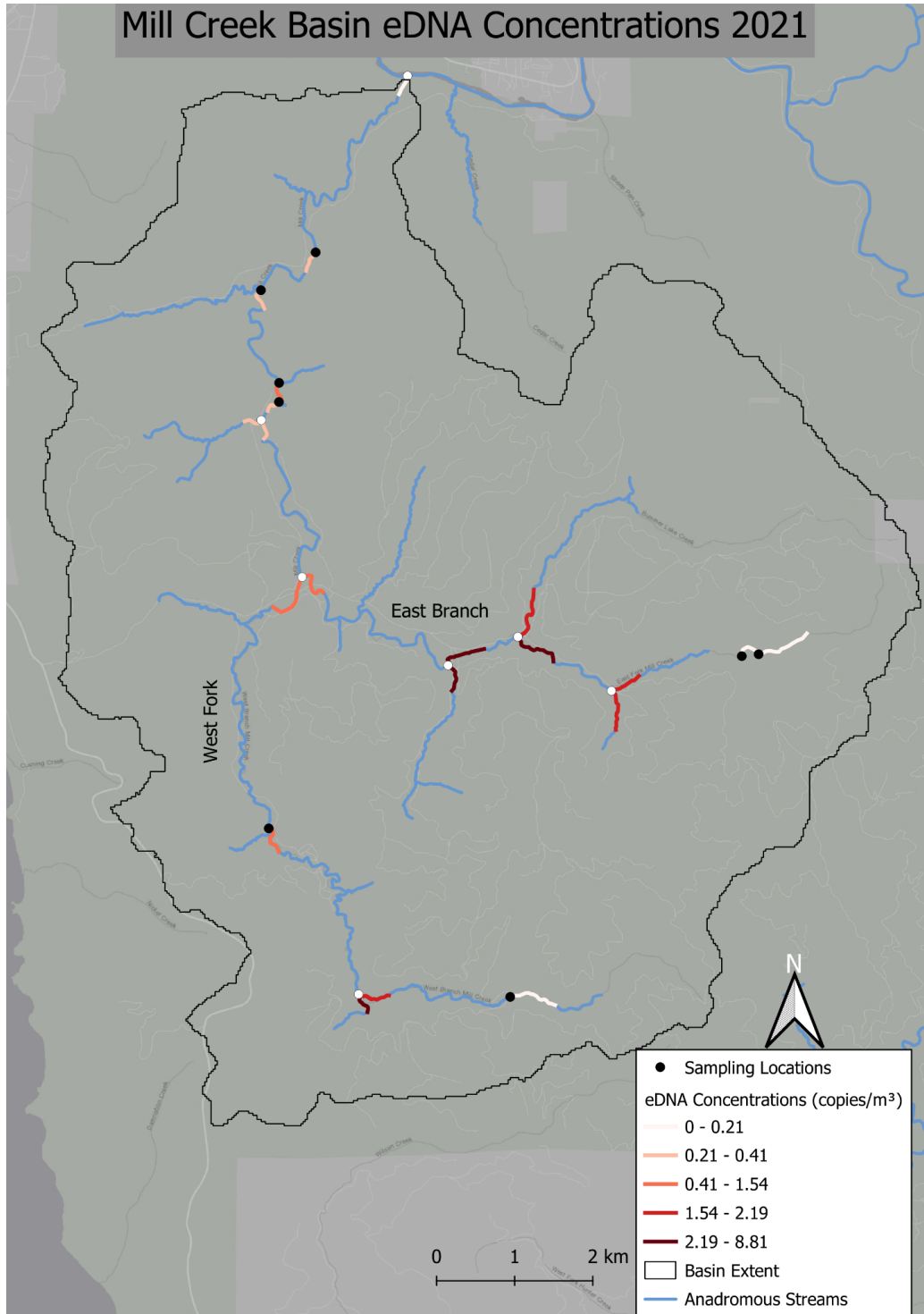
Figure 2. The gradient of DNA concentrations in the water samples collected from the Mill Creek basin in June of 2021. Filled points indicate the sampling locations where tributaries were dry. Estimated DNA concentrations are in copies/m³. The length of the lines is not equivalent to the scale of eDNA inference.

eDITH model fits

The eDITH model parameters that produced the highest correlation between the observed and predicted eDNA concentrations had a τ of one hour and reach lengths of 250 m (Pearson's correlation coefficient of 0.61 [Figure 3]). For comparison, correlations for the Carraro et al. (2018) model reached 0.84. When τ was one hour, the eDITH model failed to converge as reach lengths exceeded 250m (Table 1).
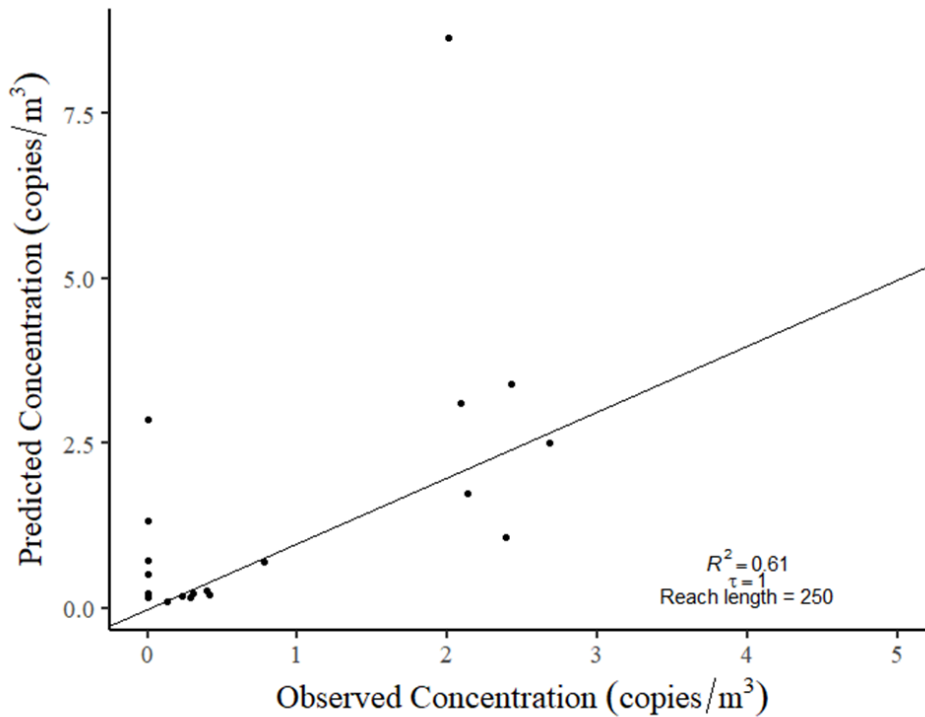

Figure 3. The highest correlation between the observed and predicted coho salmon DNA concentrations from the eDITH model. Each circle represents a river reach (excludes the West Fork). The Pearson's correlation coefficient (R) and model parameters, τ and reach length, that were used to achieve this result are shown.

Table 1. Model parameter combinations and model fit for each of the tested initial parameters.

Model parameter combination	Pearson's correlation coefficient
$\tau\ =\ 1\ ;\ Reach\ length\ =\ 100\ m$	0.56
$\tau\ =\ 1\ ;\ Reach\ length\ =\ 250\ m$	0.61
$\tau\ =\ 1\ ;\ Reach\ length\ =\ 500\ m$	Failed to converge
$\tau\ =\ 1\ ;\ Reach\ length\ =\ 750\ m$	Failed to converge
$\tau\ =\ 1\ ;\ Reach\ length\ =\ 1000\ m$	Failed to converge
$\tau\ =\ 5\ ;\ Reach\ length\ =\ 100\ m$	0.56
$\tau\ =\ 5\ ;\ Reach\ length\ =\ 250\ m$	0.58
$\tau\ =\ 5\ ;\ Reach\ length\ =\ 500\ m$	0.46
$\tau\ =\ 5\ ;\ Reach\ length\ =\ 750\ m$	0.15
$\tau\ =\ 5\ ;\ Reach\ length\ =\ 1000\ m$	-0.07

Initial exploration of the modeling process revealed that eDITH model predictions are extremely sensitive to the size of the river network. When the river network was extracted with a minimum tributary catchment size of 2.5 km^2, as used by Carraro et al. (2018), the resulting river network included many headwater streams, and juvenile Coho Salmon were predicted to be patchily distributed throughout the mainstem and largest tributaries (Figure 4A). However, this river network was unrealistic as many of the small, headwater streams were dry during the summer baseflow conditions when our survey was conducted. Therefore, we increased the minimum catchment size to 25 km^2 to more closely resemble the extent of the river network during summer baseflow conditions. With this change, the extent of the river network was reduced, and Coho Salmon were predicted to be uniformly distributed throughout the basin except for some of the highest headwaters (Figure 4B). Given the low eDNA concentrations observed in the mainstem and near the mouth, it is unlikely that these estimates match the true distribution of Coho Salmon in the Mill Creek basin (Walkley and Garwood 2017).
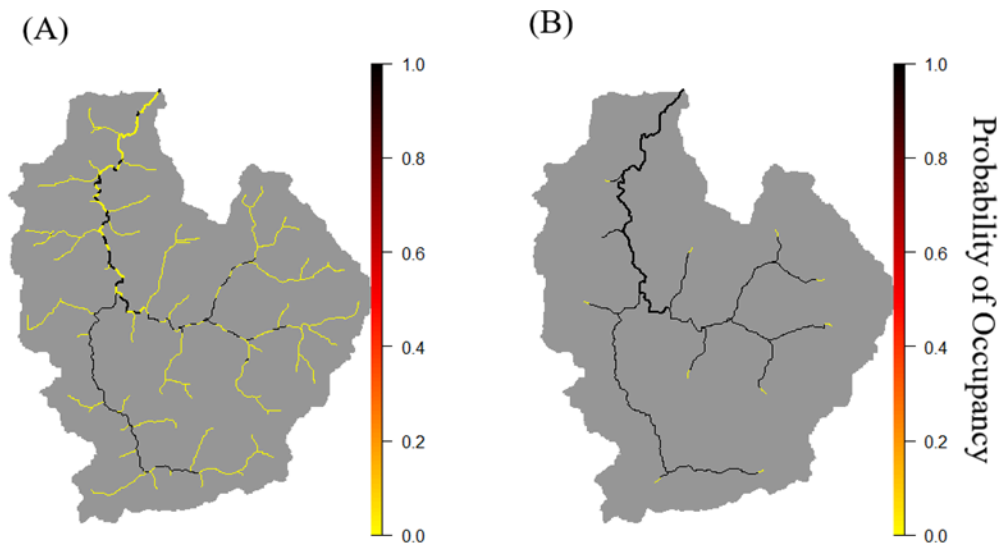
Figure 4. The predicted distribution of coho salmon when the minimum catchment size threshold is 2.5 km^2 (A) and 25 km^2 (B)

Discussion

   After developing an eDITH model for the Mill Creek basin, we have concluded that application of this approach for estimating juvenile Coho Salmon distribution and abundance in this system would require more sampling and research, and it may not be well-suited to rapid, short-term assessments of low-abundance organisms like juvenile Coho Salmon. We attempted to obtain a "snapshot" in time of the distribution and abundance of juvenile Coho Salmon throughout the basin with sampling protocols implemented over a short period of time. In their original study, Carraro et al. (2018) collected a total of 21 500-ml water samples from each of their 15 sampling sites at approximately biweekly intervals over a year. Their sampling protocol encompassed the full lifecycle of their target species and increased the likelihood of encountering periods with large amounts of available target eDNA, which other studies have shown to be important for accurate quantification of species abundance (Yates et al. 2019). Furthermore, a large number of replicate water samples provides more data for informing the eDITH model, helping minimize problems of model convergence due to a lack of data. Given that our best-fit model still fell well short of the fits achieved by Carraro et al. (2018), this modeling procedure is likely unsuitable for these rapid "snapshots" of distribution and abundance of rare Coho Salmon during summer baseflow conditions.

The extremely low eDNA concentrations observed in our study also contributed to a poor model fit with high levels of uncertainty. The Carraro et al. (2018) study had a median estimated eDNA concentration of 90,330 copies/m^3 while our system had an observed average of 1 copy/m^3, which was likely attributable to differences in target species biomass and number of eDNA sources. When present, *F. sultana* colonies form 1 to 2 cm thick mats and reach a total biomass of up to 1.5 kg per m^2 across the wetted channel (Raddum and Johnsen 1983). If infected, bryozoans release the parasitic spores of *T. bryosalmonae* which infect brown trout (*Salmo trutta*) and further disperse *T. bryosalmonae* eDNA as spores released in the fish's urine. Therefore, there were multiple sources of DNA for *T. bryosalmonae* and likely high biomass of *F. sultana* if we assume that, when present, *F. sultana* colonized the entire riverbed channel. The assumption of a uniform contributing area across the river channel is built into the eDITH model, where $A_{S.i}$ is equal to the riverbed area of an occupied river stretch. Coho salmon are generally patchily distributed in deep pools, so the assumption of equal contributing area is unlikely to hold (Walkley and Garwood 2017; Shaffer et al. 2022). This may contribute to the lack of fit observed in our models and was likely compounded by differences in eDNA particle decay and transport distances between the two studies.

We found that reach lengths of 250 m and a decay time (τ) of one hour best reproduced our observed data yet these values vary considerably from those of Carraro et al. (2018) and failed to produce similar model fits. A growing body of literature suggests that distances of eDNA particle transport and decay times are system and species-specific (Spence et al. 2021; Jo and Yamanaka 2022; Shaffer et al. 2022; Herman et al. 2023). We note that the present study and modeling effort were conducted prior to our estimation of eDNA decay constants for rainbow trout, a species closely related to Coho salmon (this report, Task 1). Those decay rates indicate that decay time may be larger by up to one order of magnitude; but these higher decay times lead to unusually large transport distances (up to 5km) that are inconsistent with the transport distances (~450 m) observed in the field (Task 2 and Task 3). These conflicting findings highlight the challenges of accurately parameterizing the eDITH model.

We have found that the uncertainties associated with this modeling effort hinder the reliable application of the eDITH model to assess juvenile Coho Salmon distribution and abundance in the Mill Creek system. Given that we were unable to predict Coho Salmon abundance even within small pools (Task 4; Shaffer et al. 2022), basin-scale efforts such as this are likely better suited to basins where species' abundance or biomass is much greater than observed here. Accuracy of the eDITH model would be improved by increased sample sizes and using system-specific information on eDNA decay and transport which could be collected using the doser system developed and applied under tasks 2 and 3. Continued research on system and species-specific eDNA particle decay, transport distances, and the current hydrologic conditions within the system of interest is essential before this modeling framework could be strongly recommended.

References

Andruszkiewicz Allan, E., W. G. Zhang, A. Lavery, and A. Govindarajan. 2021. Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. Environmental DNA 3(2):492–514.

Barnes, M. A., C. R. Turner, C. L. Jerde, M. A. Renshaw, W. L. Chadderton, and D. M. Lodge. 2014. Environmental Conditions Influence eDNA Persistence in Aquatic Systems. Environmental Science & Technology 48(3):1819–1827.

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

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

Hartig, F., Minunno, F., Paul, S. 2023. BayesianTools: General-Purpose MCMC and SMC Samplers and Tools for Bayesian Statistics. R package version 0.1.8., https://github.com/florianhartig/BayesianTools.

Herman, B. A. 2023, January. USE OF FOREIGN EDNA TRACERS TO RESOLVE SITE- AND TIME-SPECIFIC EDNA DISTRIBUTIONS IN NATURAL STREAMS. California State Polytechnic University Humboldt.

Garwood, J. M., and M. D. Larson. 2014. Reconnaissance of Salmonid Redd Abundance and Juvenile Salmonid Spatial Structure in the Smith River with Emphasis on Coho Salmon (Oncorhynchus kisutch). The Smith River Alliance and the California Department of Fish and Wildlife Anadromous Fisheries Resource and Monitoring Program, Final Report 91251.

Jerde, C. L., B. P. Olds, A. J. Shogren, E. A. Andruszkiewicz, A. R. Mahon, D. Bolster, and J. L. Tank. 2016. Influence of Stream Bottom Substrate on Retention and Transport of Vertebrate Environmental DNA. Environmental Science & Technology 50(16):8770–8779.

Jo, T., and H. Yamanaka. 2022. Meta‑analyses of environmental DNA downstream transport and deposition in relation to hydrogeography in riverine environments. Freshwater Biology 67(8):1333–1343.

Levi, T., J. M. Allen, D. Bell, J. Joyce, J. R. Russell, D. A. Tallmon, S. C. Vulstek, C. Yang, and D. W. Yu. 2019. Environmental DNA for the enumeration and management of Pacific salmon. Molecular Ecology Resources 19(3).

Pilliod, D. S., and M. B. Laramie. 2016. Salmon Redd Identification Using Environmental DNA (eDNA). USGS.

O'Callaghan, J. F., and D. M. Mark. 1984. The Extraction of Drainage Networks from Digital Elevation Data. Computer Vision, Graphics, and Image Processing 28:323–344.

Pont, D., M. Rocle, A. Valentini, R. Civade, P. Jean, A. Maire, N. Roset, M. Schabuss, H. Zornig, and T. Dejean. 2018. Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. Scientific Reports 8(1):10361.

Raddum, G. G., and T. M. Johnsen. 1983. Growth and feeding of Fredericella sultana (bryozoa) in the outlet of a humic acid lake. Hydrobiologia 101:115–120.

Rourke, M. L., A. M. Fowler, J. M. Hughes, M. K. Broadhurst, J. D. DiBattista, S. Fielder, J. Wilkes Walburn, and E. M. Furlan. 2022. Environmental DNA (eDNA) as a tool for assessing fish biomass: A review of approaches and future considerations for resource surveys. Environmental DNA 4(1):9–33.

Schmelzle, M. C., and A. P. Kinziger. 2016. Using occupancy modelling to compare environmental DNA to traditional field methods for regional-scale monitoring of an endangered aquatic species. Molecular Ecology Resources 16(4):895–908.

Shaffer, J. T. 2022, December. COMPARISON OF ENVIRONMENTAL DNA AND UNDERWATER VISUAL COUNT SURVEYS FOR DETECTING JUVENILE COHO SALMON (ONCORHYNCHUS KISUTCH) IN RIVERS. California State Polytechnic University Humboldt.

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

Spence, B. C., D. E. Rundio, N. J. Demetras, and M. Sedoryk. 2021. Efficacy of environmental DNA sampling to detect the occurrence of endangered coho salmon (*Oncorhynchus kisutch*) in Mediterranean‑climate streams of California's central coast. Environmental DNA 3(4):727–744.

Thomsen, P. F., and E. Willerslev. 2015. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183:4–18.

U.S Geological Survey. 2016. The StreamStats program. U.S Geological Survey.

Wilcox, T. M., K. S. McKelvey, M. K. Young, A. J. Sepulveda, B. B. Shepard, S. F. Jane, A. R. Whiteley, W. H. Lowe, and M. K. Schwartz. 2016. Understanding environmental DNA detection probabilities: A case study using a stream-dwelling char *Salvelinus fontinalis*. Biological Conservation 194:209–216.

Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society 73(1):3–36.

Yates, M., C., D. J. Fraser, and A. Derry M. 2019. Meta‑analysis supports further refinement of eDNA for monitoring aquatic species‑specific abundance in nature (1):5–13.