Quantifying Impacts of an Environmental Intervention Using Environmental DNA

Elizabeth Andruszkiewicz Allan[1*‡], Ryan P. Kelly[1*], Erin R. D'Agnese[1,3], Maya N. Garber-Yonts[1], Megan R. Shaffer[1], Zachary J. Gold[2†], Andrew O. Shelton[2]

[1] University of Washington, School of Marine and Environmental Affairs, 3737 Brooklyn Ave NE, Seattle, WA 98105, U.S.A.

[2] Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Blvd. E, Seattle, WA 98112, U.S.A.

[3] Wild EcoHealth, Tacoma WA, 98465, USA

[*] Authors contributed equally to this work.

[†] Currently at NOAA Pacific Marine Environmental Laboratory, Seattle, WA, USA.

[‡] Corresponding author: eallan@uw.edu

For submission to: *Ecological Applications*

Manuscript type: Article

Open Research Statement: Data are already published and publicly available, with those publications properly cited in this submission. The repository for code can be found at: Andruszkiewicz Allan et al. 2023. https://doi.org/10.5281/zenodo.8029271

Keywords: environmental DNA, quantitative metabarcoding, environmental impact assessments, salmon, culvert

Abstract

Environmental laws around the world require some version of an environmental impact assessment surrounding construction projects and other discrete instances of human development. Information requirements for these assessments vary by jurisdiction, but nearly all require an analysis of biological elements of ecosystems. Amplicon-sequencing - also called metabarcoding - of environmental DNA (eDNA) has made it possible to sample and amplify the genetic material of many species present in those environments, providing a tractable, powerful, and increasingly common way of doing environmental impact analysis for development projects. Here, we analyze a 18-month time-series of water samples taken before, during, and after two culvert removals in a salmonid-bearing freshwater stream. We also sampled multiple control streams to develop a robust background expectation against which to evaluate the impact of this discrete environmental intervention in the treatment stream. We generate calibrated, quantitative metabarcoding data from amplifying the 12s MiFish mtDNA locus and complementary species-specific quantitative PCR data to yield multi-species estimates of absolute eDNA concentrations across time, creeks, and sampling stations. We then use a linear mixed-effects model to reveal patterns of eDNA concentrations over time, and to estimate the effects of the culvert removal on salmonids in the treatment creek. We focus our analysis on four common salmonid species: cutthroat trout (*Oncorhynchus clarkii*), coho salmon (*O. kisutch*), rainbow trout (*O. mykiss*), and sockeye salmon (*O. nerka*). We find that one culvert in the treatment creek seemed to have no impact while the second culvert had a large impact on fish passage. The construction itself seemed to have only transient effects on salmonid species during the two construction events. In the context of billions of dollars of court-mandated road culvert replacements taking place in Washington State, USA, our results suggest that culvert replacement can be conducted with only minimal impact of construction to key species of management concern. Furthermore, eDNA methods can be an effective and efficient approach for monitoring hundreds of culverts to prioritize culverts that are required to be replaced. More broadly, we demonstrate a rigorous, quantitative method for environmental impact reporting using eDNA that is widely applicable in environments worldwide.

Introduction

At present, it remains difficult to comprehensively measure the environmental impacts of discrete human activities, despite such assessment often being required by law. Within the United States, both state and federal laws often require a form of environmental-impact assessment for medium- to large-scale projects (i.e., those that might have a significant impact on the environment) (Morgan 2012). Outside the US, many nations have their own versions of these same laws. Specifically when measuring impacts on aquatic ecosystems, assessments generally are based on literature reviews or field measurements of key species selected beforehand (Rubin et al. 2017). These traditional methods are often expensive, rely on just a few species, and are limited in spatial and temporal coverage (Martin et al. 2012). Moreover, they often lack pre-project monitoring and any or sufficient post-project monitoring, given that the goals of a development project normally focus on construction itself and funding is often extremely limited. For example, a recent literature review of stream restoration projects cited that more than half of projects evaluated (62%) had no pre-project monitoring and only sampled once per year (for before, during, and post-project sampling) (Rubin et al. 2017). Thus, current assessment efforts relying on traditional survey methods often fall short in documenting and quantifying environmental impacts.

A key difficulty in conducting ecosystem assessments is that there is no one way to survey the world and just "see what is there." All methods of environmental sampling are biased as they capture a selective portion of the biodiversity present (Rubin et al. 2017). Net samples for fish, for example, fail to capture species too small or too large to be caught in the net. Environmental DNA (eDNA), however, comes as close to this goal as any method yet developed although not without bias (see below): a sample of water, soil, or even air, contains the genetic traces of many thousands of species, from microbes to whales. Sequencing eDNA is therefore a means of surveying many species in a consistent and scalable way (Taberlet et al. 2012, Thomsen and Willerslev 2015). Environmental assessments have begun to make use of eDNA for such work around the world (Muha et al. 2017, Duda et al. 2021, Klein et al. 2022, Maasri et al. 2022, Moss et al. 2022), but are not yet common practice. Surveying the world by eDNA has long been commonplace in microbial ecology (Ogram et al. 1987, Rondon et al. 2000, Turnbaugh et al. 2007) but has recently become popular for characterizing eukaryotic communities (Taberlet et al. 2012, Kelly et al. 2014, De Vargas et al. 2015, Port et al. 2015, Valentini et al. 2016, Stat et al. 2017). Techniques generally include an amplification step such as quantitative PCR, digital or digital-droplet PCR, or traditional PCR from mixed templates followed by high-throughput sequencing (Ruppert et al. 2019). This last technique is known as eDNA metabarcoding.

In a metabarcoding approach, broad-spectrum PCR primers identify hundreds or thousands of taxa across a very wide diversity of the tree of the life (e.g., Leray et al. (2013)). Nevertheless the absence of a taxon from a

sequenced sample does not indicate the absence of that taxon from the environment but rather that the taxon failed to amplify (Shelton et al. 2016, Kelly et al. 2019, Buxton et al. 2021, Gold et al. 2023). In virtually all comparisons, metabarcoding recovers far more taxa than any other sampling method (Port et al. 2015, Kelly et al. 2017, Seymour et al. 2021). However, we expect results from metabarcoding to differ dramatically from non-PCR based sampling methods due to the fundamental differences in sampling residual genetic material as opposed to whole organisms. Furthermore, eDNA analyses rely on several laboratory processes, including PCR amplification, all of which contribute to complicating the interpretation of results (see Shelton et al. (2016) and Kelly et al. (2019)). Specifically, PCR amplification is an exponential process for which the efficiency varies across species and primer set (Gloor et al. 2016). By understanding these differences, we can correct for taxon-specific biases to yield quantitative estimates of the community composition prior to PCR (McLaren et al. 2019, Shelton et al. 2022). Other ways to conduct quantitative metabarcoding include using qSeq (Hoshino et al. 2021), a process in which a random tag is added to target sequences before PCR. However, if different species amplify at different rates during PCR, these quantifications would reflect not just the starting concentration but also the amplification efficiency.

After correcting for amplification biases, the resulting dataset is compositional, revealing the proportions of each species' DNA present in each sample, but importantly, contains no information about the absolute abundance of DNA present (Gloor et al. 2016, McLaren et al. 2019, Silverman et al. 2021, Shelton et al. 2022). We can tie these proportional estimates to absolute abundances using additional data such as a quantitative PCR (qPCR) assay for one of the taxa present. Thus, a single qPCR assay and a single metabarcoding assay can together provide quantitative estimates of many species as opposed to running as many qPCR assays as species of interest (see also (Pont et al. 2022)). Together, we can use these data to assess changes in eDNA concentrations of species over time, and due to environmental impacts, such as replacing a culvert under a road.

As a result of a ruling in a federal court (Martinez 2013), Washington State is under a mandate to replace hundreds of culverts that allow water to pass under roads and highways, costing billions of dollars. Improperly designed culverts can lead to many negative consequences for fish, especially anadromous salmon, including habitat fragmentation, loss of accessibility to spawning and rearing habitat, and genetic isolation (Price et al. 2010, Frankiewicz et al. 2021). These impacts on salmonids furthermore violate the sovereign treaty rights of the region's indigenous tribes (Martinez 2013). Salmonid species are of cultural and economic importance to the indigenous peoples of the region, and without restoration of historic salmon-rearing habitat, the continued decline of salmonids can lead to not only ecological destruction, but the loss of cultural and economic viability for many indigenous tribes (Schmidhauser 1976, Lackey 2003, Long and Lake 2018).

Presently, the prioritization process for replacing culverts preventing fish passage conducted by the Washington Department of Transportation is a protocol provided by the Washington Department of Fish and Wildlife, which includes factors such as the amount of habitat blocked by the barrier, the types of species blocked by the barrier, and estimated cost of repair, among other things (Washington Department of Fish and Wildlife 2019). However, data on fish presence upstream of barriers (i.e., culverts blocking fish passage) are rare and often not included in these assessments. Using eDNA as a proxy for fish presence could provide important data for project prioritization and have the potential to be more cost effective.

Once a culvert has been designated as in need of repair, the intention is to improve conditions for biota, including migrating fish, but the construction itself might have a short-term negative effect before the longer-term improvements are realized. Specifically in culvert replacements, studies have cited the negative impacts of construction to include sediment accumulation, removal of vegetation, and blocking flow and stranding fish (Wellman et al. 2000, Washington Department of Fish and Wildlife 2019). However, it is unclear how long these effects might last and if the long-term benefits of the culvert replacement justify the short-term costs of the construction. These disruptions also underscore the importance of both properly assessing culverts to determine if they are blocking fish passage and monitoring after construction to ensure the replacement actually improved fish passage.

Many studies have attempted to quantify when culverts are barriers to fish passage and how effective culvert replacements are for fish passage, either by measuring physical parameters of the culvert and stream after replacement (Price et al. 2010), or by measuring biological parameters, including electrofishing (Ogren and Huckins 2015) or genetic differentiation from fish tissues (Wood et al. 2018, Nathan et al. 2018). In some cases, culverts deemed blockages did not prove to block fish passage (MacPherson et al. 2012), while in others, blockages that were replaced were not found to improve fish passage (Price et al. 2010) or improve overal biotic integrity (Ogren and Huckins 2015). Sampling water for eDNA analysis before, during, and post-restoration can provide valuable information on if the restoration is needed, how the restoration negatively impacts communities during construction, and if the restoration efforts did in fact correct the blockage.

Here, we report the results of an approximately 18-month eDNA sampling effort before, during, and after the replacement of two culverts (one small and one large) in a creek, assessing the impact of these projects on the salmonid species present. We do so using a combination of metabarcoding (12s mtDNA) and qPCR to yield estimates of the concentrations of DNA present at each time point, and we use parallel samples from four control creeks to develop a causal analysis of changes in these concentrations. A clear opportunity for policy-relevant eDNA work is in using its power to survey many species at a time to improve the way we assess the impacts of human activities. Here, we demonstrate the utility of eDNA for policy-relevant

environmental assessments by surveying many species simultaneously and improving the way we assess the impacts of human activities.

Methods

Site and Species Selection

We selected sampling locations based loosely on an asymmetrical BACI (Before-After-Control-Impact) study design (Underwood 1992, 1994, Benedetti-Cecchi 2001) to measure the environmental impact of a culvert replacement using eDNA. We sampled four control creeks where construction was not occurring (Figure 1) at monthly intervals, both upstream and downstream of each creek's culvert. The two culverts in the treatment creek (Padden) were suspected to be partially impassible and thus were removed and replaced during the course of the study. The four control creeks ranged from preventing fish passage (Barnes and Chuckanut), partially passable (Squalicum), to allowing fish passage (Portage; see Appendix S1) (Washington Department of Fish and Wildlife 2019). These creeks were chosen due to their comparable size, flow, watersheds, and species presumed to be present to constrain as many ecological variables as possible.

The first culvert replacement (SR-11) in Padden Creek occurred over about two months and included the "de-watering" of the creek, removal of the existing culvert, installation of the new culvert, and then the "re-watering" of the creek from late August 2021 to early October 2021 (Appendix S1: Figure S4). The second culvert replacement (I-5) in Padden Creek was a much larger construction project, including daylighting the creek and building a bridge under a large, five-lane interstate. In-water work for the I-5 culvert replacement began in late June 2022 and was completed in September 2022. By sampling before, during, and after both construction events, we were then able to isolate the effect of the culvert replacement itself – controlling for temporal trends, background environmental variability, and sampling variability – using a linear mixed effects model of eDNA abundances across creeks, time points, sampling stations, and species.

Because salmonids are the primary species of management concern in these creeks, we focus the present analysis on the four salmonid species most common in our data: cutthroat trout (*Oncorhynchus clarkii)*, coho salmon (*O. kisutch*), rainbow/steelhead trout (*O. mykiss*), and sockeye/kokanee salmon (*O. nerka*). Not all four salmonids are expected to be found in all five of the creeks sampled. As documented by WA Department of Fish and Wildlife SalmonScape (http://apps.wdfw.wa.gov/salmonscape/map.html), all creeks contain cutthroat trout, steelhead trout, and coho salmon. Barnes Creek is the only creek documented to have kokanee salmon (a freshwater sub-type of sockeye salmon). However, local spawner surveys conducted by the City of Bellingham from 2015-2020 in Padden Creek documented kokanee salmon, as well as the

other three species (City of Bellingham 2015). The four salmonid species in this study have different life histories and behaviors that would impact when fish (and therefore eDNA concentrations) occur in the creeks. Furthermore, three of the four species in this study have both freshwater resident and saltwater migrating behavior. For the fish exhibiting migratory behavior, the run timings vary for each species in the study area (see Discussion and Appendix S1: Figure S4). Therefore, our eDNA concentrations might reflect contributions from both migrating and non-migrating individuals at any given time point in the dataset.

Water Sampling

From March 2021 to February 2022, all five creeks were sampled monthly (n=12). Monthly sampling continued in Portage Creek, Padden Creek, and Squalicum Creek through August 2022, with one additional sampling point in December 2022 (n=19). At each sampling station (upstream and downstream of a culvert) at each creek, we collected three 2-liter water samples. Samples were collected using an eDNA Backpack (Smith Root; Thomas et al. (2018)), a portable pumping-and-filtering device set to filter at 1 L/min at 82.7 kPa (12 psi). In some months, less than 2 L of water was filtered due to clogging. Water samples were filtered using single-use inlet tubes through 5μm self-preserving filters (Smith Root, Vancouver, WA), which were then dried and kept at room temperature until DNA extraction within 1 month of collection (Thomas et al. 2019).

Over the course of the sampling, water discharge varied from very low to no flow in summer months to high flow in winter months (Figure 2). We account for this dilution by converting eDNA concentration [copies/μL] to an eDNA mass flow rate [copies/s] by multiplying eDNA concentrations by discharge [L/s] (Tillotson et al. 2018, Thalinger et al. 2019). Flow gauges maintained by the United States Geological Survey (USGS) were used for Padden Creek (USGS Gauge 12201905), Chuckanut Creek (USGS Gauge 12201700), and Squalicum Creek (USGS Gauge 12204010; https://maps.waterdata.usgs.gov/mapper/index.html; U. S. Geological Survey (1994); Appendix S1: Figure S1). Over the course of sampling, the flow gauges at Chuckanut Creek and Squalicum Creek became inoperable after a major flooding event. To find discharge rates for Chuckanut and Squalicum Creeks, five years of historical data (2015-2020) were used to generate a monthly averaged correction factor based on Padden Creek (Appendix S1, Appendix S1: Figure S3). No discharge data was available for Portage Creek or Barnes Creek. Based on field sampling conditions, the discharge from Padden Creek was used as a proxy for both Portage and Barnes as they are in similarly sized watershed areas and land-cover characteristics.

DNA Extraction, Amplification, Sequencing

All molecular work prior to sequencing was performed at the University of Washington. Details of the molecular work can be found in Appendix S1. Briefly, DNA was extracted off filters using a Qiashredder column (Qiagen, USA) and the DNeasy Blood and Tissue Kit (Qiagen, USA) with an overnight incubation (Appendix S1, Thomas et al. (2019)). Extracts were stored at -20°C until PCR amplification within 2 months of extraction.

For metabarcoding, we targeted a ~186 bp hypervariable region of the mitochondrial DNA 12S rRNA gene for PCR amplification (MiFish; Miya et al. 2015), but using modified primer sequences as given in Praebel and Wangensteen (unpublished; via personal communication). The primer sequences, final reaction recipe, and cycling conditions can be found in Appendix S1. Each month of samples was amplified on a single plate with the addition of a no template control (NTC; molecular grade water in lieu of template) and a positive control (genomic DNA from kangaroo, a species not present in the environment). PCR products were visualized, size-selected, and diluted iteratively if inhibited. After cleaning, a second PCR amplification added unique indices to each sample using Nextera indices (Illumina, USA) to allow pooling multiple samples onto the same sequencing run (See Appendix S1 for details). Indexed PCR products were also size-selected and visualized before pooling for sequencing. Samples were randomized in 3-month blocks and each block split across 3 sequencing runs to avoid run effects, for a total of 14 sequencing runs. The loading concentration of each library was 4-8 pM and 5-20% PhiX was included depending on the composition of the run. Sequencing was conducted using an Illumina Miseq with v3 2x300 chemistry at the NOAA Northwest Fisheries Science Center and the University of Washington's Northwest Genomics Center.

Here, we used mock communities to determine the species-specific amplification efficiencies for each salmonid in the study. We constructed five communities with known proportions of starting DNA from different species (total DNA as measured by Qubit). The communities ranged from having a total of 12 to 20 species, but six salmonid species were included in all five mock communities (Appendix S3: Table S2). We sequenced these communities using the same metabarcoding primers and thermocycling conditions above and then determined the species-specific amplification rates given the discrepancy between the known starting proportion and the proportion of reads after sequencing. The mock community data were then used to correct the sequencing reads from the environmental samples to estimate the starting DNA proportions of each species in environmental samples, which is the metric of interest (Figure 3, green boxes). This is the first application of the model to correct eDNA data from water samples with mock community data as described in Shelton et al. (2022) (see Appendix S2 for more information).

Bioinformatics

After sequencing, bioinformatic analyses were conducted in R (R Core Team 2017). A more detailed description of the bioinformatics pipeline is included in Appendix S1. Briefly, primer sequences were removed using *Cutadapt* (Version 1.18) (Martin 2011) before *dada2* (Callahan et al. 2016) trimmed, filtered, merged paired end reads, and generated amplicon sequence variants (ASVs). Taxonomic assignment was conducted via the *insect* package (Wilkinson et al. 2018) using a tree generated by the developers for the MiFish primers that was last updated in November 2018. Only species level assignments from *insect* were retained and ASVs not annotated or not annotated to species level were then checked against the NCBI nucleotide database using BLAST+ (Camacho et al. 2009). Query sequences that matched a single species at >95% identity were retained.

Quantitative PCR and Inhibition Testing

We quantified cutthroat trout (*O. clarkii*) DNA in each sample, targeting a 114 bp fragment of the cytochrome b gene with a qPCR assay (Duda et al. 2021). The primer/probe sequences, final recipe, and thermocycling conditions can be found in Appendix S1. Each DNA sample was run in triplicate and was checked for inhibition using the EXO-IPC assay (Applied Biosystems). The majority of environmental samples (60%) were inhibited and accordingly diluted for analysis. In 80% of inhibited samples, a 1:10 dilution or less remedied the inhibition, but some samples (11%) required dilution by a factor of up to 1000. Each plate included a 8-point standard curve created using synthetic DNA (gBlocks) ranging from 1 to 100,000 copies/μL and six no template controls (NTCs) were included on each plate with molecular grade water instead of template. All qPCRs were conducted on an Applied Biosystems StepOnePlus thermocycler.

All qPCR data was processed in R using Stan (Stan Development Team 2022), relating environmental samples to the standard curve via a linear model (Figure 3, blue boxes; Appendix S2: Figure S1). We amended the standard linear regression model to more realistically capture the behavior of qPCR observations, accommodating non-detections as a function of underlying DNA concentration, and letting the standard deviation vary with the mean (lower-concentration samples had more uncertainty). See McCall et al. (2014) and Shelton et al. (2019) for similar models; see Appendix S2 for full statistical details. Subsequent analysis corrected for sample-specific dilution if found inhibited and corrected for any variation in water-volume filtered during sample collection. Samples with standard deviations between technical replicates larger than 1.5 Ct values were removed from analyses.

Quantitative Metabarcoding

The intercalibration of the mock community samples demonstrated the rank order of amplification efficiencies for salmonids (Appendix S1: Figure S10 and Appendix S1: Figure S11). Cutthroat trout (*O. clarkii*) and sockeye/kokanee salmon (*O. nerka*) had similar amplification efficiencies, both of which were higher than rainbow/steelhead trout (*O. mykiss*) and coho salmon (*O. kisutch*), which had the lowest amplification efficiency. Calibrated metabarcoding analysis yielded quantitative estimates of the proportions of species' DNA in environmental samples prior to PCR. We then converted these proportions into absolute abundances by expansion, using the qPCR results for our reference species, cutthroat trout (*O. clarkii*). We estimated the total amplifiable salmonid DNA in environmental sample i as $\mathrm{C}_{\mathrm{amplifiable}_i} = \frac{\mathrm{C}_{\mathrm{qPCR\ reference}_i}}{\mathrm{Proportion}_{\mathrm{reference}_i}}$, where C has units of [DNA copies/uL] and then expanded species' proportions into absolute concentrations by multiplying these sample-specific total concentrations by individual species' proportions, such that for species j in sample i, $\mathrm{C}_{i,j} = \mathrm{C}_{\mathrm{amplifiable}_i} * \mathrm{Proportion}_{i,j}$. Here, we combine the modeled output of the qPCR model for cutthroat trout (Figure 3, dashed blue box) and modeled proportions of salmonid DNA from metabarcoding (Figure 3, dashed green box). Although in the future this could be used as a joint model, here the precision of our modeled estimates were very high such that we used the mean of the posterior estimates from each model to move forward as input to the time series model (Figure 3, dashed purple box; see Appendix S2 for more details). Finally, due to the range of water discharge over the course of the year, we converted from DNA concentration [copies/L] to a mass flow rate [copies/s] after multiplying by the discharge of each creek [m^3/s] (Figure 3, solid purple boxes).

Estimating the Effects of Culvert Replacement and of Culverts Themselves

We sampled four control creeks as context against which to compare the observations in Padden Creek, our treatment creek where the two culverts were being replaced. At a given station in a given creek, some DNA concentration exists for each species. For simplicity, we focus on a single species and a single station (downstream or upstream) for the moment. Our observations of the (log) DNA concentration in creek i at time t are distributed as $Y_{i,t} \sim \mathcal{N}(\mu_{i,t}, \sigma^2)$. More complex versions of the model may let σ vary across creeks, time points, species, or with environmental covariates of interest.

We are interested in how the DNA concentration changes over time, so we assert that the expected value of DNA in a creek at time t, $\mu_{i,t}$, depends upon its time point, in some way. We considered three ways of modeling the salmonid eDNA data, each in a Bayesian framework, but each treating non-independence among time points somewhat differently:

- A linear auto-regressive (AR(1)) model, written in `stan`. For each species in each creek, the expected

concentration of eDNA of each month is a linear function of the expected value from the previous month. Within a species, the monthly autoregressive parameters are shared across creeks.

- A generalized additive model (GAM), written in `brms` (which itself writes a `stan` model). For each species in each creek, an independent set of spline (weighting) parameters describes the temporal trends in expected eDNA concentration; the number of spline knots is shared across species and creeks.
- A linear mixed-effects (LME) model, written in `rstanarm`. For each species in each creek, sampling month is treated as a random effect. Each species-creek-month effect is treated as an independent draw from a common distribution.

Ultimately, the three models yielded very similar results (Appendix S2), and the LME model proved simplest and most flexible insofar as it could easily handle datasets with uneven sets of observations – for example, cases in which a species was detected downstream of a barrier, but not upstream.

In `R` code using `rstanarm`, this model is coded as

```
stan_glmer(log(observed) ~ (1 + time_idx|creek:species) + (1|station:creek:species:time_idx)
```

Results

Metabarcoding and Quantitative PCR

In total, we generated ~52 million reads across all environmental samples and 27 mock community samples (3 communities x 9 replicates [6 even, 3 skewed proportions]) for calibration (see below). After quality-filtering and merging all runs, ~45 million reads remained from ~24,000 amplicon sequence variants (ASVs) in the environmental samples, of which ~83% of reads were annotated to species level (per sample: mean = 82%, median = 93%, min = 0%, max = 99.99% of reads annotated). We only focus on the metabarcoding data from four salmonids for the remainder of this paper. The four salmonids represent ~54% of all reads and ~64% of the annotated reads found in environmental samples.

In the mock community samples, 98.7% of the ~5 million reads after quality filtering were annotated to species level. Importantly, the target salmonid ASVs in the mock communities were found in environmental samples, unambiguously linking the taxa in calibration samples with those in environmental samples. The most common salmonid species found in the environmental samples was cutthroat trout (*O. clarkii*), which was found in ~85% of samples, followed by coho salmon (*O. kisutch*) found in ~62% of samples, then rainbow/steelhead trout (*O. mykiss*) found in ~40% of samples, and finally sockeye/kokanee salmon (*O. nerka*) found in ~10% of samples. Over half of the samples across all times, creeks, and stations had at least

50% of reads assigned to cutthroat trout.

After calibrating metabarcoding data using mock communities (See Appendix S1 and Appendix S2), we estimated the salmonid composition across time points, creeks, and stations (Figures 4 and 5). The culvert in one control creek (Barnes) appeared to be nearly a total barrier to salmonid passage, with salmonid eDNA detected upstream of the culvert at only three time points, in contrast to being detected at every time point in the downstream station of the same creek. The other four creeks had no such pattern associated with the culverts, suggesting that fish passage may have been possible through the culverts, or that there were resident populations upstream of the culverts.

All environmental samples were quantified for absolute concentrations of cutthroat trout DNA across 32 qPCR plates, resulting in ~630 samples (~60%) with a positive detection in at least 1 of 3 technical replicates. The modeled output of cutthroat trout DNA concentrations, ranged from 50 copies/L to 1.4 x 10^6 copies/L, with a mean value of ~47,000 copies/L (Figure 6).

We combined compositional information from metabarcoding with absolute concentrations from qPCR for our reference species, cutthroat trout (*O. clarkii*), to estimate the total concentration of DNA for each species (See Appendix S2). These quantitative data for all four target species were then used in the linear mixed-effects model to assess salmonid trends over time (Figure 7).

Effects of Culverts

Before considering the effect of construction, the difference in trends between upstream and downstream stations (Figure 7) demonstrates that the culverts themselves have some effect, but generally not a large effect on the salmonid species surveyed. A notable exception was Barnes Creek, as the culvert was so clearly a barrier as most time points had no salmonid DNA upstream. Padden Creek upstream of I-5 also was more clearly a barrier to fish passage, while the culvert across SR-11 seemed to be less of a barrier to fish passage. In other cases, salmonid DNA is found upstream but not downstream, indicating that the culvert is likely not a barrier and there are resident individuals upstream of the culvert.

Summarizing over all species and the four creeks used in the time series model, the culvert effect was minimal (Figure 8); the average log-fold change between upstream and downstream sites was not significantly different from zero. Individual species' patterns were similar, indicating that there is not a species-specific effect where culverts block the passage of some salmon but not others (Appendix S1: Figure S12). The maximum positive log-fold change (i.e., upstream having a higher mass flow rate) was 2.78 in Squalicum Creek for coho salmon (*O. kisutch*) in August 2021, while the maximum negative log-fold change (i.e., downstream having a

higher mass flow rate) was -1.11 found in Squalicum Creek for cutthroat trout (*O. clarkii*) in December 2021 (Appendix S1: Figure S12). Of all species, creeks, and time points, 23 of the 151 observations were within a log-fold change of -0.1 to 0.1, which corresponds with eDNA mass flow rates upstream within 10% of mass flow rates downstream.

We also considered how the log-fold change in eDNA mass flow rates were impacted by the flow rate or discharge of the creek itself (Figure 9). We found that at months of the lowest flow (summer months), the log-fold changes between mass flow rates were the highest, while in winter months with highest discharge the log-fold changes were lower and downstream sites often had higher eDNA mass flow rates than upstream sites (Appendix S1: Figure S12).

Effects of Culvert Replacement

By comparing the difference in upstream and downstream mass flow rates before and after construction in Padden Creek, we can assess how large of an impact the two culvert replacements had on salmonid species (Figure 10). The effects of the culvert replacement operations appeared to have been transient and fairly minor for the four salmonid species surveyed. We saw very minor fluctuations in the difference between upstream and downstream salmonid DNA mass flow rates, and did not see an increase in this difference due to the culvert removal as the log-fold changes in Padden Creek were similar to those in the control creeks at the same time points (Figure 10, grey points vs. black points in areas of yellow shading).

Discussion

Environmental DNA can provide quantitative measurements of environmental impacts

Here, we used both eDNA metabarcoding and a single species-specific qPCR assay to rigorously quantify both the effect of culverts and the impact of two culvert replacements on salmonids in the same creek. We observed a clear seasonal pattern in the DNA concentrations of four salmonid species detected in the study. The sampling design and the linear mixed effects model leveraged information across treatment and control creeks to integrate the change in eDNA concentrations due to time, whether a sample was collected below or above a barrier (i.e., culvert), and whether or not there was construction occurring. Thus, we could isolate the changes in eDNA concentrations as a result of the intervention (i.e., construction) while accounting for the variance due to time and station (i.e., season and culvert).

A few other studies have used eDNA to measure environmental impacts in rivers and streams. Duda et al. (2021) used 11 species-specific qPCR assays to document the distribution of resident and migratory fish

after a large dam removal project (Elwha River near Port Angeles, Washington). No eDNA sampling was conducted before the dam removal, but the study provided a wealth of information about species returning after the dam removal, providing a very important dataset to use eDNA to monitor ecological changes due to human intervention. Similarly, Muha et al. (2017) sampled three locations upstream and three locations downstream before and after the removal of a weir that was thought to be a barrier to salmonid migrations. The authors only sampled once before and twice after the removal, spanning about a year, and used eDNA metabarcoding to look at the presence/absence of species detected. They found that in fact the before sample demonstrated that the weir was not preventing fish passage (similar to the results found in this study) and furthermore documented a slight increase in alpha diversity in the first time point after the barrier removal and then a return to a similar alpha diversity in the second time point after the removal (similar results found in this study using eDNA concentrations rather than diversity). Finally Yamanaka and Minamoto (2016) sampled along a river with three barriers, finding some fish able to cross barriers and some not, suggesting that the eDNA can indicated habitat connectivity for fishes across barriers.

Importantly, our study demonstrates the value of combining a single qPCR assay with metabarcoding data to generate quantitative estimates of eDNA concentrations of many species without requiring n qPCR assays for n species of interest. Here, we ultimately only quantified the impacts of four species, but importantly, we did not know *a priori* how many species of interest there might be and we reduced our efforts two fold by only conducting two assays (one species-specific qPCR and one metabarcoding assay) as opposed to four assays (four species-specific qPCR assays). This can also be particularly helpful for taxa that don't have a previously published qPCR assay, but are detected using universal metabarcoding assays. Metabarcoding data alone only gives compositional data, which cannot be used in a time series to quantify environmental impacts because there is no information about absolute eDNA concentrations. However, by anchoring or grounding proportions using a single qPCR assay, the proportional data can be turned into quantitative data. The species for which to run the qPCR assay can be determined after the metabarcoding is completed; the most commonly found species with a robust qPCR assay should be used to glean the most information.

Fish life histories and expected patterns

The four salmonid species in this study have different life histories and behaviors that would impact when fish (and therefore eDNA concentrations) occur in the creeks. Three of the four species in this study have both freshwater and anadromous populations. Cutthroat trout (*O. clarkii*) encompasses both non-migrating, resident trout in the creeks and coastal run cutthroat that migrate into Padden Creek from saltwater (Bellingham Bay). Similarly, *O. nerka* includes both anadromous sockeye salmon and freshwater resident

kokanee salmon and *O. mykiss* includes both anadromous steelhead trout and non-migrating rainbow trout. Using eDNA, we cannot distinguish between the migrating and non-migrating subspecies of *O. clarkii*, *O. nerka*, and *O. mykiss*. Therefore, our eDNA concentrations might reflect contributions from both migrating and non-migrating individuals at any given time point in the dataset.

For these four anadromous salmonids, the run timings for the migrating populations vary for each species in the study area (Bellingham, WA). Adult coastal cutthroat (*O. clarkii*) are documented to run throughout the entire year, whereas coho salmon (*O. kisutch*) run from September to December, sockeye salmon (*O. nerka*) run from October to December, and steelhead trout (*O. mykiss*) run from November to June. For migrating coho (*O. kisutch*) and steelhead trout (*O. mykiss*), juveniles may be present in the creeks year-round (Appendix S1: Figure S4). eDNA methods at present cannot distinguish adults versus juveniles from DNA found in a water sample.

Despite the mix of migrating and non-migrating populations and various run timings, our metabarcoding data demonstrate that in Padden Creek, there was a clear signal of sockeye/kokanee salmon (*O. nerka*) both upstream and downstream only in November 2021 - February 2022 and again in December 2022. This signal corresponds well with the documented run timing of October to December and the presence of out-migrating juveniles in early spring. In contrast, cutthroat trout (*O. clarkii*) and coho salmon (*O. kisutch*) were found nearly year-round in Padden Creek. The persistent signal from *O. clarkii* could be explained by resident cutthroat trout. However, *O. kisutch* does not have a resident subspecies and the run timing is only documented from September to December. This could potentially be due to juveniles maturing and residing in the creeks for 1-2 years after hatching while adults migrate into the creeks only during the run time to spawn. Visual surveys (e.g., snorkel surveys, electrofishing, smolt traps) are conducted infrequently to determine adult and juvenile salmonid abundances. Though *O. kisutch* eDNA was found year round, the highest concentrations were found near the expected run timing and the life history of *O. kisutch* includes rearing year-round in freshwater. Finally, though the lowest concentrations on average, rainbow/steelhead trout (*O. mykiss*) was also found nearly year-round in Padden Creek, which could be contributions from migrating steelhead (November to June), juveniles maturing and migrating, or from resident rainbow trout. Though the *O. mykiss* signal is found year-round, the highest concentrations do seem to correspond with the steelhead run timing.

Interpreting eDNA with respect to fish abundance and flow

By capturing residual eDNA from water samples, we are measuring a different signal than counting how many fish are in the creek at each time of sampling. We should not expect the eDNA concentration to directly

correlate to the number of fish in the creek at the time of sampling. Shelton et al. (2019) used a paired eDNA sampling and seine netting analysis to demonstrate that eDNA concentrations provide a smoothed biological signal over space and time. We acknowledge this smoothing effect and emphasize that in the context of using eDNA for environmental impact assessments, it is preferable to use a survey technique such as eDNA that integrates signal across a larger spatial and temporal scale.

Many previous papers have commented on the "ecology" of eDNA and the various processes that contribute to eDNA concentrations in environmental samples (e.g., shedding rates, decay rates, transport) (Barnes and Turner 2015). For example, higher concentrations of eDNA could be the result of a greater number (or biomass) of fish present, or increased shedding rates, or decreased decay. Many review papers document the nuances of interpreting eDNA data and we recommend reviewing them for a deeper understanding (see Andruszkiewicz Allan et al. (2020) for a review on shedding and decay rates and Harrison et al. (2019) for a review on transport). Other studies have also documented the relative importance of eDNA transport in streams. Most notably, Tillotson et al. (2018) measured eDNA at four sites with similar discharge rates to the creeks in this study and specifically addressed spatial and temporal resolutions, finding that eDNA concentrations reflect short time- (and therefore length-) scales by comparing peaks in eDNA concentrations to counts of salmon and accumulation by measuring both upstream and downstream sites. The authors found that the sampling site furthest downstream did not accumulate eDNA and that two tributaries feeding into a main channel were additive (Tillotson et al. 2018). For more general models and empirical data documenting transport distances in streams, see Wilcox et al. (2016), Jane et al. (2014), Jerde et al. (2016), Shogren et al. (2016), and Civade et al. (2016). Certainly eDNA concentrations can arise from different scenarios and future work should continue to investigate how to tease apart the nuances of relating eDNA concentrations to fish abundance.

In this study, to assess the impact of a culvert on fish passage, we compare eDNA concentrations upstream and downstream at the same time point in a given creek. The distance between the upstream and downstream sampling was minimal (~60-300 m, average distance of ~150 m). Therefore, we assume that the small differences in spatial and temporal scale between sampling locations is minimal such that the impacts of these various processes will affect the downstream and upstream concentrations equally. That is, in the upstream station, some amount of eDNA is coming from upstream of that location into the sampling station and leaving at the same time – in the same way that eDNA would be both entering and exiting the downstream station. Additionally, at almost every single time point for all creeks and species, the upstream DNA concentration is higher than the downstream DNA concentration. Based on that alone, we do not expect that downstream accumulation of salmonid DNA is occurring to bias our results of whether fish can pass through these culverts.

Not all culverts are barriers to salmonids

By measuring DNA concentrations of salmonid species above and below culverts on a small spatial scale, we were able to determine how much of a barrier each culvert was to fish passage. Barnes Creek was clearly a very large barrier to fish passage as we only found salmonid DNA in three months of the twelve months of sampling, and those three months had very low concentration of salmonid DNA relative to the other creeks. Within the treatment creek (Padden Creek), the SR-11 culvert did not seem to be a large barrier, while the I-5 culvert clearly was a barrier, demonstrated by the difference in salmonid composition and eDNA mass flow rates over the course of sampling.

Here, we find instances where culverts designated as barriers were likely not blocking fish passage, while others (Padden I-5 and Barnes Creek) were barriers to fish passage. Importantly, this demonstrates that collecting water samples for eDNA analysis might help to prioritize restoration of culverts suspected to be barriers to salmonids and provide a new method for post-restoration monitoring to confirm that the barrier has been corrected and allows for fish passage. Given the large amount of spending and effort required to replace culverts, this finding is important and emphasizes the potential for new tools for environmental impact assessments. We note that our sampling occurred only over a short temporal scale and future work could monitor culverts for longer time periods, different species, and different environmental conditions.

Salmonids can quickly recover from a short-term intervention in a creek

The construction had remarkably minimal effects on salmonid DNA concentrations. The disruption of disconnecting the creek, demolition of the old culvert, installation of the new culvert, and the reconnecting of the creek during both culvert replacement events showed almost no change in the difference in eDNA concentrations between downstream and upstream sampling sites. The differences in the control creeks between upstream and downstream were often higher than the treatment creek. The post-construction sampling point of the I-5 culvert replacement (only one time point), does show that the composition of salmonid DNA after replacement is now very similar to the two downstream stations, whereas before construction compositions were very different (because the culvert was a barrier). However, we lack the quantitative analysis as the site upstream of SR-11 and downstream of I-5 had no quantifiable cutthroat DNA. More time points would help demonstrate the effect of the culvert replacement. Here we found that one culvert had very minimal effect on salmonid passage while the other culvert had a large effect on salmonid passage. We note that these findings are likely not universal and certainly projects need to monitor comprehensively and quantitatively in order to assess the passability of culverts and impacts of construction.

Conclusion

It is notoriously difficult to quantify the environmental impact of discrete human impacts on ecosystems and species. Surveying species and communities by eDNA provides an opportunity for monitoring before, during, and after impacts in a scaleable and cost-effective way. Here, we demonstrate that monthly eDNA sampling before, during, and after an intervention alongside control sites can quantify the environmental impact of replacing a culvert. We found that in our treatment creek and control sites, four of the six barriers did not prohibit salmonid passage. We found that of the two culvert replacements in the treatment creek, one was a barrier and one was not, but both had minimal impacts on the four salmonid species monitored over the course of construction. We also provide a framework in which compositional metabarcoding data can be linked with qPCR data to obtain quantitative estimates of eDNA concentrations of many species. This provides a practical way to utilize the large amount of information from metabarcoding data without needing a unique qPCR assay for every species of interest. Environmental DNA is moving into practice and this study demonstrates how eDNA can be broadly used for environmental impact assessments for a wide range of species and environments.

Acknowledgements

This work was made possible by a grant from OceanKind to Ryan Kelly. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Figure 3 was created with BioRender.com. We thank Tammy Schmidt and Susan Kanzler from Washington Department of Transportation for facilitating access to field sites and providing helpful feedback throughout the project. We also thank Jenna McLaughlin, Joe Duprey and Ally Im for help field sampling and Dr. Ramon Gallego, Dr. Kim Parsons, and the University of Washington's Northwest Genomics Center for sequencing support. Dr. Braeden Van Deynze, Dr. Sunny Jardine, and Dr. Julian Olden provided helpful insight into culverts and salmonid life histories. We thank Katherine Pearson Maslenikov at the Burke Museum of Natural History and Culture for providing voucher specimens for Sanger sequencing. Thanks to Dr. Jameal Samhouri and Dr. Chris Sergeant for reviewing the manuscript.

Conflict of Interest Statement

The authors declare there are no conflicts of interest.

References

Andruszkiewicz Allan, E., W. G. Zhang, A. Lavery, and A. Govindarajan. 2020. Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. Environmental DNA:edn3.141.

Barnes, M. A., and C. R. Turner. 2015. The ecology of environmental DNA and implications for conservation genetics. Conservation Genetics 17:117.

Benedetti-Cecchi, L. 2001. Beyond Baci: Optimization of Environmental Sampling Designs Through Monitoring and Simulation. Ecological Applications 11:783–799.

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

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: High resolution sample inference from illumina amplicon data. Nature methods 13:581–583.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden. 2009. BLAST+: Architecture and applications. BMC Bioinformatics 10:4219.

City of Bellingham. 2015. Urban spawner surveys.

Civade, R. l., T. Dejean, A. Valentini, N. Roset, J.-C. Raymond, A. Bonin, P. Taberlet, and D. Pont. 2016. Spatial representativeness of environmental DNA metabarcoding signal for fish biodiversity assessment in a natural freshwater system. PLOS ONE 11:e015736619.

De Vargas, C., S. Audie, N. Henry, J. Decelle, F. Mahe, R. Logares, E. Lara, C. Berney, N. Le Bescot, I. Probert, M. Carmichael, J. Poulain, S. Romac, S. Colin, J.-M. Aury, L. Bittner, S. Chaffron, M. Dunthorn, S. Engelen, O. Flegontova, L. Guidi, A. Horak, O. Jaillon, G. Lima-Mendez, J. Lukes, S. Malviya, R. Morard, M. Mulot, E. Scalco, R. Siano, F. Vincent, A. Zingone, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, T. O. coordinators, S. G. Acinas, P. Bork, C. Bowler, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, F. Not, H. Ogata, S. Pesant, J. Raes, M. E. Sieracki, S. Speich, L. Stemmann, S. Sunagawa, J. Weissenbach, P. Wincker, and E. Karsenti. 2015. Eukaryotic plankton diversity in the sunlit ocean. Science 348:112.

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 3:121–141.

Frankiewicz, P., A. Radecki-Pawlik, A. Wałęga, M. Łapińska, and A. Wojtal-Frankiewicz. 2021. Small hydraulic structures, big environmental problems: is it possible to mitigate the negative impacts of culverts on stream biota? Environmental Reviews 29:510–528.

Gloor, G. B., J. M. Macklaim, M. Vu, and A. D. Fernandes. 2016. Compositional uncertainty should not be

ignored in high-throughput sequencing data analysis. Austrian Journal of Statistics 45:73–87.

Gold, Z., A. O. Shelton, H. R. Casendino, J. Duprey, R. Gallego, A. V. Cise, M. Fisher, A. J. Jensen, E. D'Agnese, E. A. Allan, A. Ramón-Laca, M. Garber-Yonts, M. Labare, K. M. Parsons, and R. P. Kelly. 2023. Signal and noise in metabarcoding data. PLOS ONE 18:e0285674.

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 286:20191409.

Hoshino, T., R. Nakao, H. Doi, and T. Minamoto. 2021. Simultaneous absolute quantification and sequencing of fish environmental DNA in a mesocosm by quantitative sequencing technique. Scientific Reports 11:4372.

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. 2014. Distance, flow and PCR inhibition: eDNA dynamics in two headwater streams. Molecular Ecology Resources 15:216227.

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 50:87708779.

Kelly, R. P., C. J. Closek, J. L. O'Donnell, J. E. Kralj, A. O. Shelton, and J. F. Samhouri. 2017. Genetic and manual survey methods yield different and complementary views of an ecosystem. Frontiers in Marine Science 3:73511.

Kelly, R. P., J. A. Port, K. M. Yamahara, and L. B. Crowder. 2014. Using environmental DNA to census marine fishes in a large mesocosm. PLOS ONE 9:e8617511.

Kelly, R. P., A. O. Shelton, and R. Gallego. 2019. Understanding PCR Processes to Draw Meaningful Conclusions from Environmental DNA Studies. Scientific Reports 9:12133.

Klein, S. G., N. R. Geraldi, A. Anton, S. Schmidt-Roach, M. Ziegler, M. J. Cziesielski, C. Martin, N. Rädecker, T. L. Frölicher, P. J. Mumby, J. M. Pandolfi, D. J. Suggett, C. R. Voolstra, M. Aranda, and Carlos. M. Duarte. 2022. Projecting coral responses to intensifying marine heatwaves under ocean acidification. Global Change Biology 28:1753–1765.

Lackey, R. 2003. Pacific Northwest Salmon: Forecasting Their Status in 2100. Reviews in Fisheries Science 11:35–88.

Leray, M., J. Y. Yang, C. P. Meyer, S. C. Mills, N. Agudelo, V. Ranwez, J. T. Boehm, and R. J. Machida. 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. Frontiers in Zoology 10:114.

Long, J. W., and F. K. Lake. 2018. Escaping social-ecological traps through tribal stewardship on national

forest lands in the pacific northwest, united states of america. Ecology and Society 23.

Maasri, A., S. C. Jähnig, M. C. Adamescu, R. Adrian, C. Baigun, D. J. Baird, A. Batista-Morales, N. Bonada, L. E. Brown, Q. Cai, J. V. Campos-Silva, V. Clausnitzer, T. Contreras-MacBeath, S. J. Cooke, T. Datry, G. Delacámara, L. De Meester, K.-D. B. Dijkstra, V. T. Do, S. Domisch, D. Dudgeon, T. Erös, H. Freitag, J. Freyhof, J. Friedrich, M. Friedrichs-Manthey, J. Geist, M. O. Gessner, P. Goethals, M. Gollock, C. Gordon, H.-P. Grossart, G. Gulemvuga, P. E. Gutiérrez-Fonseca, P. Haase, D. Hering, H. J. Hahn, C. P. Hawkins, F. He, J. Heino, V. Hermoso, Z. Hogan, F. Hölker, J. M. Jeschke, M. Jiang, R. K. Johnson, G. Kalinkat, B. K. Karimov, A. Kasangaki, I. A. Kimirei, B. Kohlmann, M. Kuemmerlen, J. J. Kuiper, B. Kupilas, S. D. Langhans, R. Lansdown, F. Leese, F. S. Magbanua, S. S. Matsuzaki, M. T. Monaghan, L. Mumladze, J. Muzon, P. A. Mvogo Ndongo, J. C. Nejstgaard, O. Nikitina, C. Ochs, O. N. Odume, J. J. Opperman, H. Patricio, S. U. Pauls, R. Raghavan, A. Ramírez, B. Rashni, V. Ross-Gillespie, M. J. Samways, R. B. Schäfer, A. Schmidt-Kloiber, O. Seehausen, D. N. Shah, S. Sharma, J. Soininen, N. Sommerwerk, J. D. Stockwell, F. Suhling, R. D. Tachamo Shah, R. E. Tharme, J. H. Thorp, D. Tickner, K. Tockner, J. D. Tonkin, M. Valle, J. Vitule, M. Volk, D. Wang, C. Wolter, and S. Worischka. 2022. A global agenda for advancing freshwater biodiversity research. Ecology Letters 25:255–263.

MacPherson, L. M., M. G. Sullivan, A. Lee Foote, and C. E. Stevens. 2012. Effects of Culverts on Stream Fish Assemblages in the Alberta Foothills. North American Journal of Fisheries Management 32:480–490.

Martin, C. J. B., B. J. Allen, and C. G. Lowe. 2012. Environmental impact assessment: Detecting changes in fish community structure in response to disturbance with an asymmetric multivariate BACI sampling design. Bulletin, Southern California Academy of Sciences 111:119–131.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10.

Martinez, R. 2013. United states v. washington.

McCall, M. N., H. R. McMurray, H. Land, and A. Almudevar. 2014. On non-detects in qPCR data. Bioinformatics 30:2310–2316.

McLaren, M. R., A. D. Willis, and B. J. Callahan. 2019. Consistent and correctable bias in metagenomic sequencing experiments. eLife 8:e46923.

Morgan, R. K. 2012. Environmental impact assessment: The state of the art. Impact Assessment and Project Appraisal 30:5–14.

Moss, W. E., L. R. Harper, M. A. Davis, C. S. Goldberg, M. M. Smith, and P. T. J. Johnson. 2022. Navigating the trade-offs between environmental DNA and conventional field surveys for improved amphibian monitoring. Ecosphere 13:e3941.

Muha, T. P., M. Rodríguez-Rey, M. Rolla, and E. Tricarico. 2017. Using environmental DNA to improve

626 species distribution models for freshwater invaders. Frontiers in Ecology and Evolution 5:143957.
627 Nathan, L. R., A. A. Smith, A. B. Welsh, and J. C. Vokoun. 2018. Are culvert assessment scores an indicator
628 of Brook Trout Salvelinus fontinalis population fragmentation? Ecological Indicators 84:208–217.
629 Ogram, A., G. Sayler, and T. Barkay. 1987. The extraction and purification of microbial DNA from sediments.
630 Journal of Microbiological Methods 7:5766.
631 Ogren, S. A., and C. J. Huckins. 2015. Culvert replacements: improvement of stream biotic integrity?
632 Restoration Ecology 23:821–828.
633 Pont, D., P. Meulenbroek, V. Bammer, T. Dejean, T. Erős, P. Jean, M. Lenhardt, C. Nagel, L. Pekarik,
634 M. Schabuss, B. C. Stoeckle, E. Stoica, H. Zornig, A. Weigand, and A. Valentini. 2022. Quantitative
635 monitoring of diverse fish communities on a large scale combining eDNA metabarcoding and qPCR.
636 Molecular Ecology Resources n/a.
637 Port, J. A., J. L. O'Donnell, O. C. Romero-Maraccini, P. R. Leary, S. Y. Litvin, K. J. Nickols, K. M.
638 Yamahara, and R. P. Kelly. 2015. Assessing vertebrate biodiversity in a kelp forest ecosystem using
639 environmental DNA. Molecular Ecology 25:527541.
640 Price, D. M., T. Quinn, and R. J. Barnard. 2010. Fish Passage Effectiveness of Recently Constructed Road
641 Crossing Culverts in the Puget Sound Region of Washington State. North American Journal of Fisheries
642 Management 30:1110–1125.
643 R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical
644 Computing.
645 Rondon, M. R., P. R. August, A. D. Bettermann, S. F. Brady, T. H. Grossman, M. R. Liles, K. A. Loiacono, B.
646 A. Lynch, I. A. MacNeil, C. Minor, C. L. Tiong, M. Gilman, M. S. Osburne, J. Clardy, J. Handelsman, and
647 R. M. Goodman. 2000. Cloning the soil metagenome: a strategy for accessing the genetic and functional
648 diversity of uncultured microorganisms. Applied and Environmental Microbiology 66:2541–2547.
649 Rubin, Z., G. M. Kondolf, and B. Rios-Touma. 2017. Evaluating Stream Restoration Projects: What Do We
650 Learn from Monitoring? Water 9:174.
651 Ruppert, K. M., R. J. Kline, and M. S. Rahman. 2019. Past, present, and future perspectives of environmental
652 DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global
653 eDNA. Global Ecology and Conservation 17:e00547.
654 Schmidhauser, J. R. 1976. Struggles for cultural survival: The fishing rights of the treaty tribes of the Pacific
655 Northwest. Notre Dame Law 52:30–40.
656 Seymour, M., F. K. Edwards, B. J. Cosby, I. Bista, P. M. Scarlett, F. L. Brailsford, H. C. Glanville, M. de
657 Bruyn, G. R. Carvalho, and S. Creer. 2021. Environmental DNA provides higher resolution assessment of
658 riverine biodiversity and ecosystem function via spatio-temporal nestedness and turnover partitioning.

Communications Biology 4:1–12.

Shelton, A. O., Z. J. Gold, A. J. Jensen, E. D'Agnese, E. Andruszkiewicz Allan, A. Van Cise, R. Gallego, A. Ramón-Laca, M. Garber-Yonts, K. Parsons, and R. P. Kelly. 2022. Toward quantitative metabarcoding. Ecology n/a:e3906.

Shelton, A. O., R. P. Kelly, J. L. O'Donnell, L. Park, P. Schwenke, C. Greene, R. A. Henderson, and E. M. Beamer. 2019. Environmental DNA provides quantitative estimates of a threatened salmon species. Biological Conservation 237:383–391.

Shelton, A. O., J. L. O'Donnell, J. F. Samhouri, N. C. Lowell, G. D. Williams, and R. P. Kelly. 2016. A framework for inferring biological communities from environmental DNA:115.

Shogren, A. J., J. L. Tank, E. A. Andruszkiewicz, B. P. Olds, C. L. Jerde, and D. Bolster. 2016. Modelling the transport of environmental DNA through a porous substrate using continuous flow-through column experiments. Journal of The Royal Society Interface 13:2016029011.

Silverman, J. D., R. J. Bloom, S. Jiang, H. K. Durand, E. Dallow, S. Mukherjee, and L. A. David. 2021. Measuring and mitigating PCR bias in microbiota datasets. PLoS Computational Biology 17:e1009113.

Stan Development Team. 2022. RStan: The r interface to stan.

Stat, M., M. J. Huggett, R. Bernasconi, J. D. DiBattista, T. E. Berry, S. J. Newman, E. S. Harvey, and M. Bunce. 2017. Ecosystem biomonitoring with eDNA: Metabarcoding across the tree of life in a tropical marine environment. Scientific Reports:111.

Taberlet, P., E. Coissac, M. Hajibabaei, and L. H. Rieseberg. 2012. Environmental DNA. Molecular Ecology 21:17891793.

Thalinger, B., E. Wolf, M. Traugott, and J. Wanzenböck. 2019. Monitoring spawning migrations of potamodromous fish species via eDNA. Scientific Reports 9:15388.

Thomas, A. C., J. Howard, P. L. Nguyen, T. A. Seimon, and C. S. Goldberg. 2018. ANDe ™: A fully integrated environmental DNA sampling system. Methods in Ecology and Evolution 9:13791385.

Thomas, A. C., P. L. Nguyen, J. Howard, and C. S. Goldberg. 2019. A self-preserving, partially biodegradable eDNA filter. Methods in Ecology and Evolution 10:1136–1141.

Thomsen, P. F., and E. Willerslev. 2015. Environmental DNA: An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183:418.

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 220:111.

Turnbaugh, P. J., R. E. Ley, M. Hamady, C. M. Fraser-Liggett, R. Knight, and J. I. Gordon. 2007. The Human Microbiome Project. Nature 449:804–810.

U. S. Geological Survey. 1994. USGS water data for the nation. Retrieved from https://waterdata.usgs.gov/nwis <11/30/2022>.

Underwood, A. J. 1992. Beyond BACI: the detection of environmental impacts on populations in the real, but variable, world. Journal of Experimental Marine Biology and Ecology 161:145–178.

Underwood, A. J. 1994. On Beyond BACI: Sampling Designs that Might Reliably Detect Environmental Disturbances. Ecological Applications 4:3–15.

Valentini, A., P. Taberlet, C. Miaud, R. l. Civade, J. E. Herder, P. F. Thomsen, E. Bellemain, A. Besnard, E. Coissac, F. Boyer, C. Gaboriaud, P. Jean, N. Poulet, N. Roset, G. H. Copp, P. Geniez, D. Pont, C. Argillier, J.-M. Baudoin, T. Peroux, A. J. Crivelli, A. Olivier, M. Acqueberge, M. Le Brun, P. R. Moller, E. Willerslev, and T. Dejean. 2016. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. Molecular Ecology 25:929942.

Washington Department of Fish and Wildlife. 2019. Fish passage inventory, assessment, and prioritization manual.

Wellman, J., D. Combs, and S. B. Cook. 2000. Long-Term Impacts of Bridge and Culvert Construction or Replacement on Fish Communities and Sediment Characteristics of Streams. Journal of Freshwater Ecology 15:317–328.

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 194:209216.

Wilkinson, S. P., S. K. Davy, M. Bunce, and M. Stat. 2018. Taxonomic identification of environmental DNA with informatic sequence classification trees.

Wood, D. M., A. B. Welsh, and J. Todd Petty. 2018. Genetic Assignment of Brook Trout Reveals Rapid Success of Culvert Restoration in Headwater Streams. North American Journal of Fisheries Management 38:991–1003.

Yamanaka, H., and T. Minamoto. 2016. The use of environmental DNA of fishes as an efficient method of determining habitat connectivity. Ecological Indicators 62:147153.

Figures

Figure 1. Map of sampling locations near Bellingham, Washington. In the red subset, triangles designate the downstream sampling location and circles designate the upstream sampling location. Padden Creek (the treatment creek where the two culverts were replaced) is shown in the blue subset where downstream is a triangle, upstream of the first culvert (SR-11) is a circle, and upstream of the second culvert (I-5) is a square.

Figure 2. Discharge (m^3/s) in Padden (USGS Gauge 12201905), Chuckanut (USGS Gauge 12201700), and Squalicum (USGS Gauge 12204010) Creeks over the course of sampling. Gauges at Chuckanut and Squalicum Creek went offline in November 2021 after a major storm event. Portage Creek and Barnes Creek did not have stream gauges. Circles designate the day of sampling. For Padden Creek, the nearest 15 minute interval of flow was used. For Chuckanut and Squalicum Creeks, the correction factor from five years of historical data from Padden Creek was used (see methods section and Appendix S1: Figure S2 and Appendix S1: Figure S3).

Figure 3. Conceptual figure of different datasets and models used for analyses. * indicates that here, biological replicates are different dilutions of the synthetic gBlock. ** indicates that for most samples, only one technical replicate was sequenced but for one sample per sampling month, three technical replicates were sequenced to check for consistency across replicates. *** indicates that here, the three biological replicates indicate three different mock communities with varying species compositions, but all containing the four salmonids of interest. Created with BioRender.com.

Figure 4. Compositions of salmonid DNA in control creeks as determined by metabarcoding after correction for amplification bias. Grey shading denotes time points that were not sampled (Barnes and Chuckanut Creeks after March 2023 and Squalicum Creek in September 2021 which was dry). The empty bars in the Barnes upstream sites indicate that no salmonid DNA was found at those time points.

Figure 5. Compositions of salmonid DNA in Padden Creek as determined by metabarcoding after correction for amplification bias. The empty bar in the March 2023 in Up5 indicates that no salmonid DNA was found. The vertical dashed lines indicates the time period in which the culverts were replaced (SR-11 and I-5, respectively).]

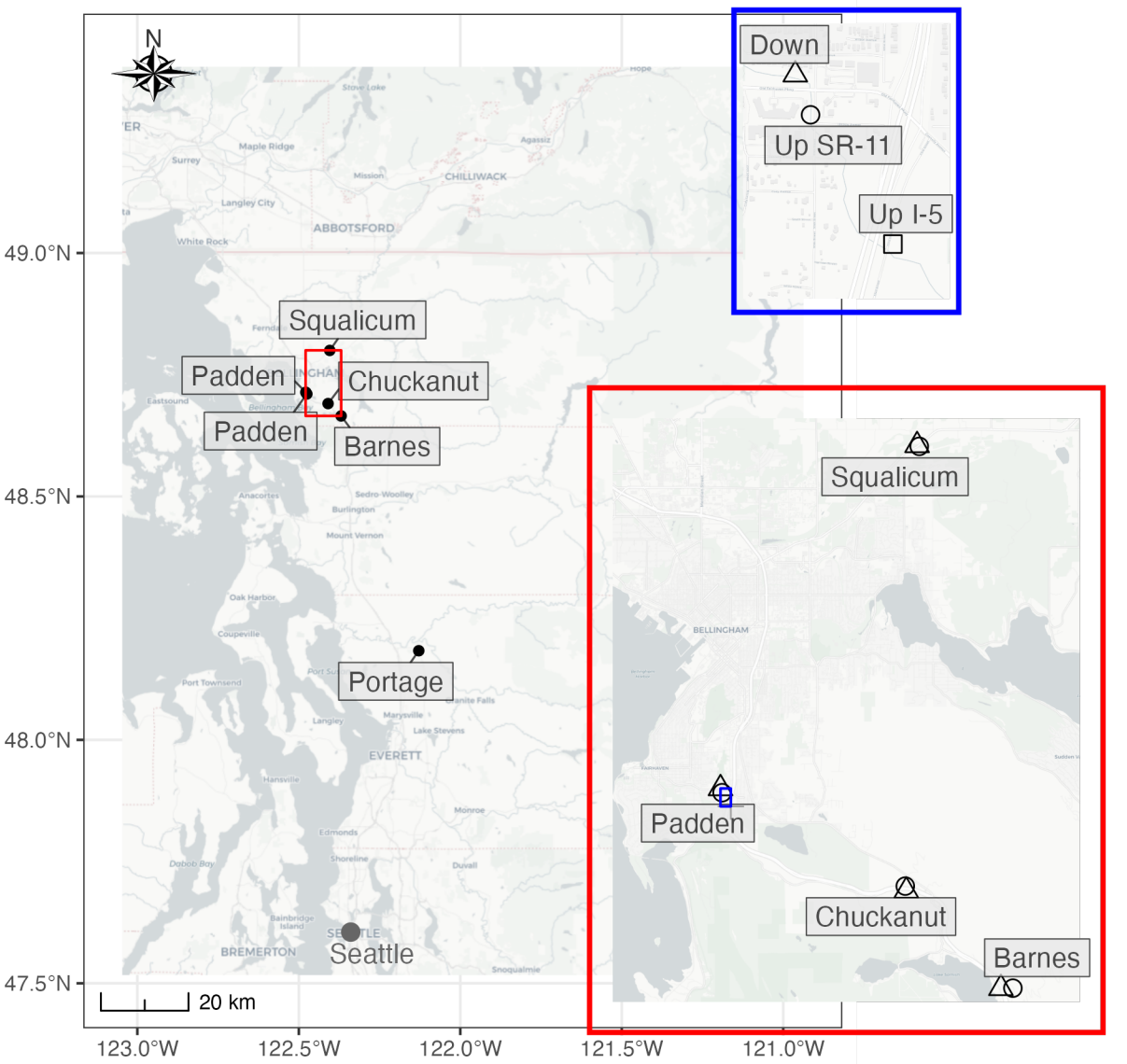
Figure 6. Absolute mass flow rate (log copies/s) of cutthroat trout (*O. clarkii*) as measured by qPCR after flow correction. Note that Barnes Creek and Chuckanut Creek were not sampled after February 2022. Red crosses show the limit of detection for each species and time point, which changes with flow rate and total volume filtered per sample.

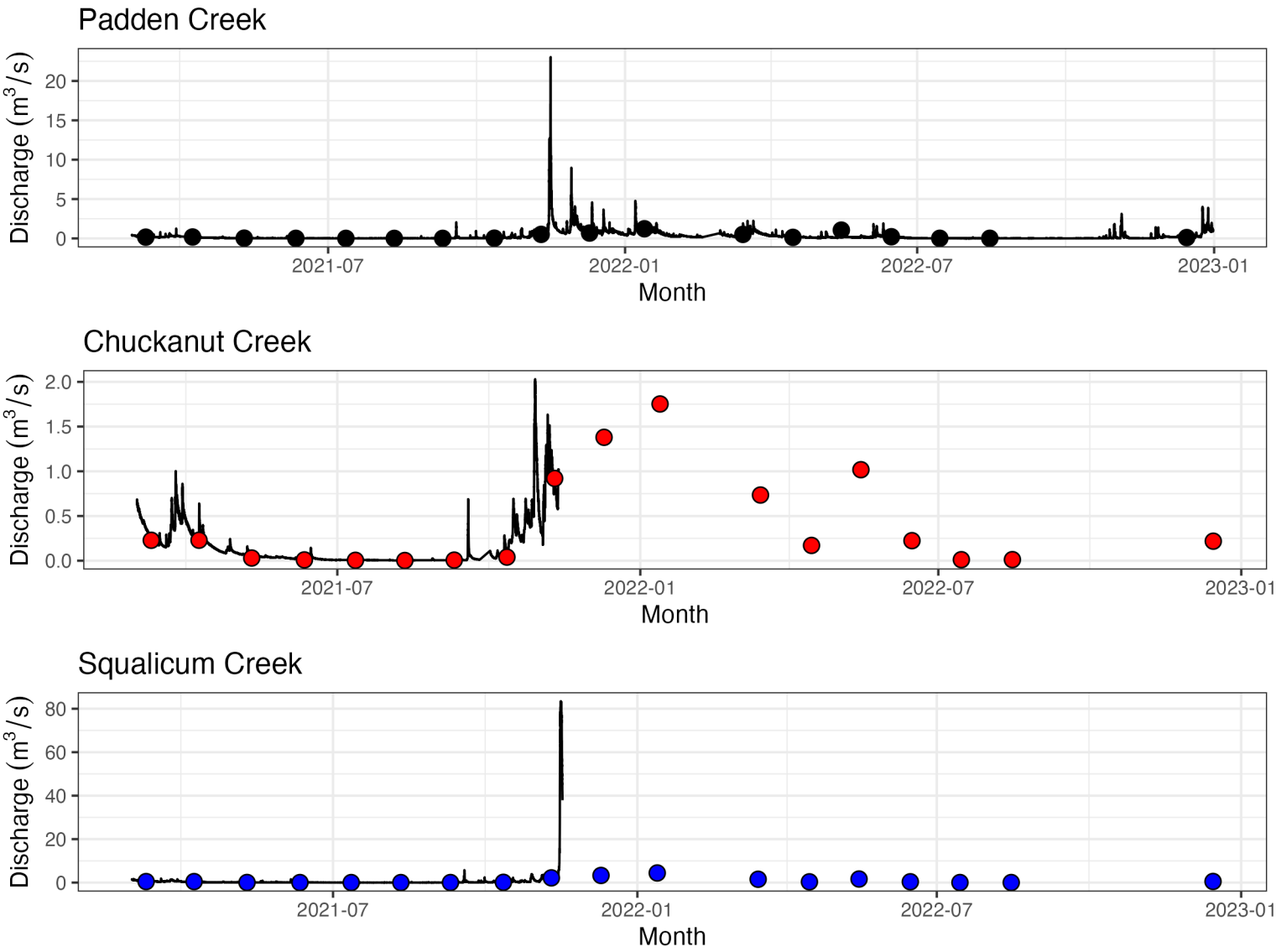
Figure 7. Trends in mass flow rate (log copies/s) for each of four salmonid species across creeks and across time as estimated by eDNA analysis. Points represent posterior means for the linear mixed effects model and error bars represent the 95% posterior confidence interval. Colors indicate station upstream (black) or downstream (grey) of the culvert. Padden has an additional sampling site upstream of the second culvert (I-5; blue). Yellow shading indicates the time period in which the culverts in the treatment creek (Padden Creek) were replaced. Grey shading indicates time points that were not sampled (Barnes and Chuckanut after February 2022). Time points with no data had no sequencing reads corresponding to that species or no quantifiable cutthroat DNA by qPCR.

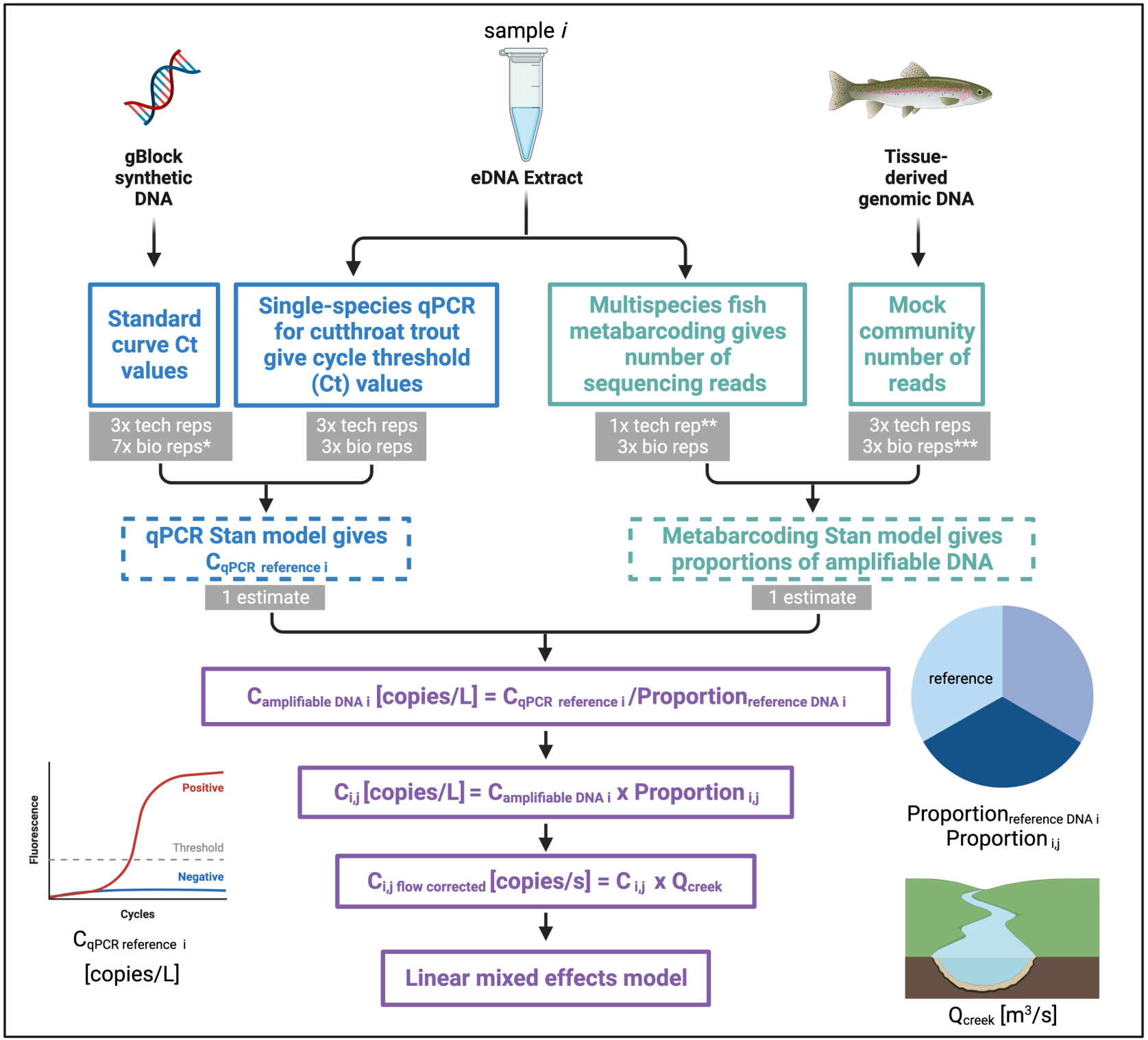
Figure 8. The effect of culvert on salmonid abundance summed across all species and creeks by time. The y-axis shows the log-fold change in eDNA mass flow rate (copies/s) between upstream and downstream, normalized by upstream mass flow rate. The box boundaries correspond to the 25th and 75th percentiles; the whiskers correspond to 1.5 times the interquartile range. Here, negative values imply that eDNA mass flow rates are higher downstream than upstream. Samples with very low eDNA mass flow rates (less than 150 copies/s) were removed before plotting to remove extreme proportional values due to large denominators. Grey stars indicate times when no samples were taken.

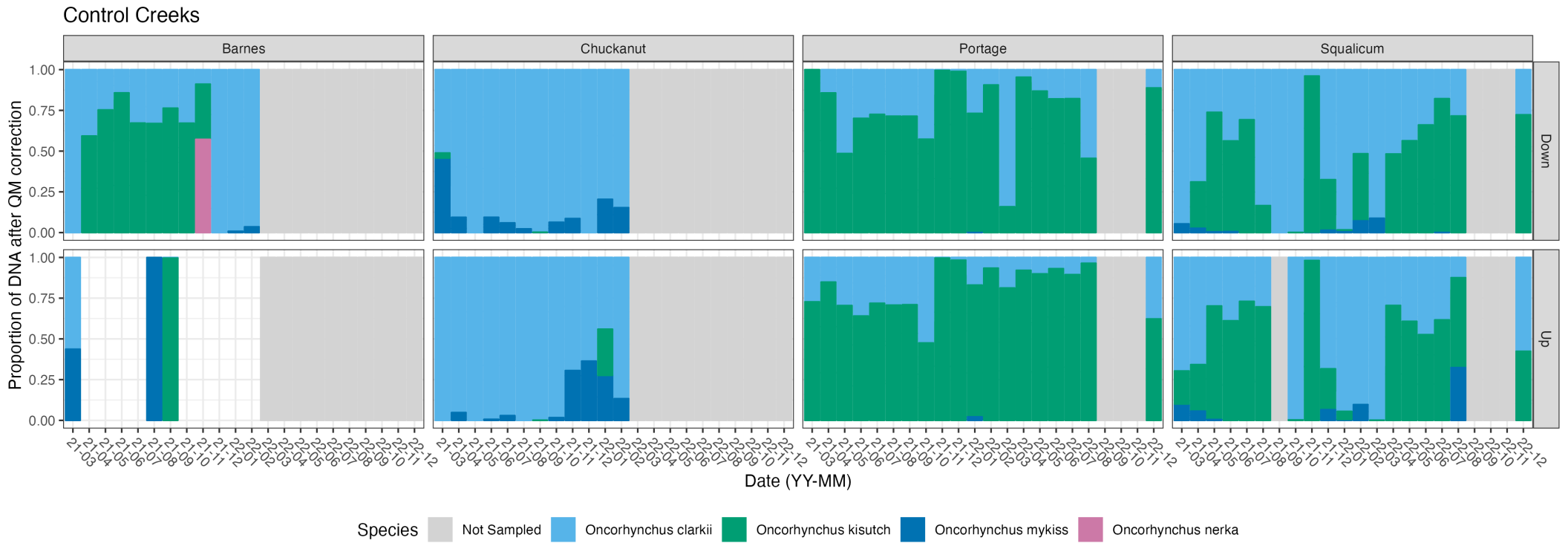
Figure 9. Log-fold change in eDNA mass flow rate between upstream and downstream over time. Size of circles corresponds to the discharge in each creek at each time point. Color of circles corresponds to each creek. Each creek and time point has up to four circles of the same color for the four salmonid species.

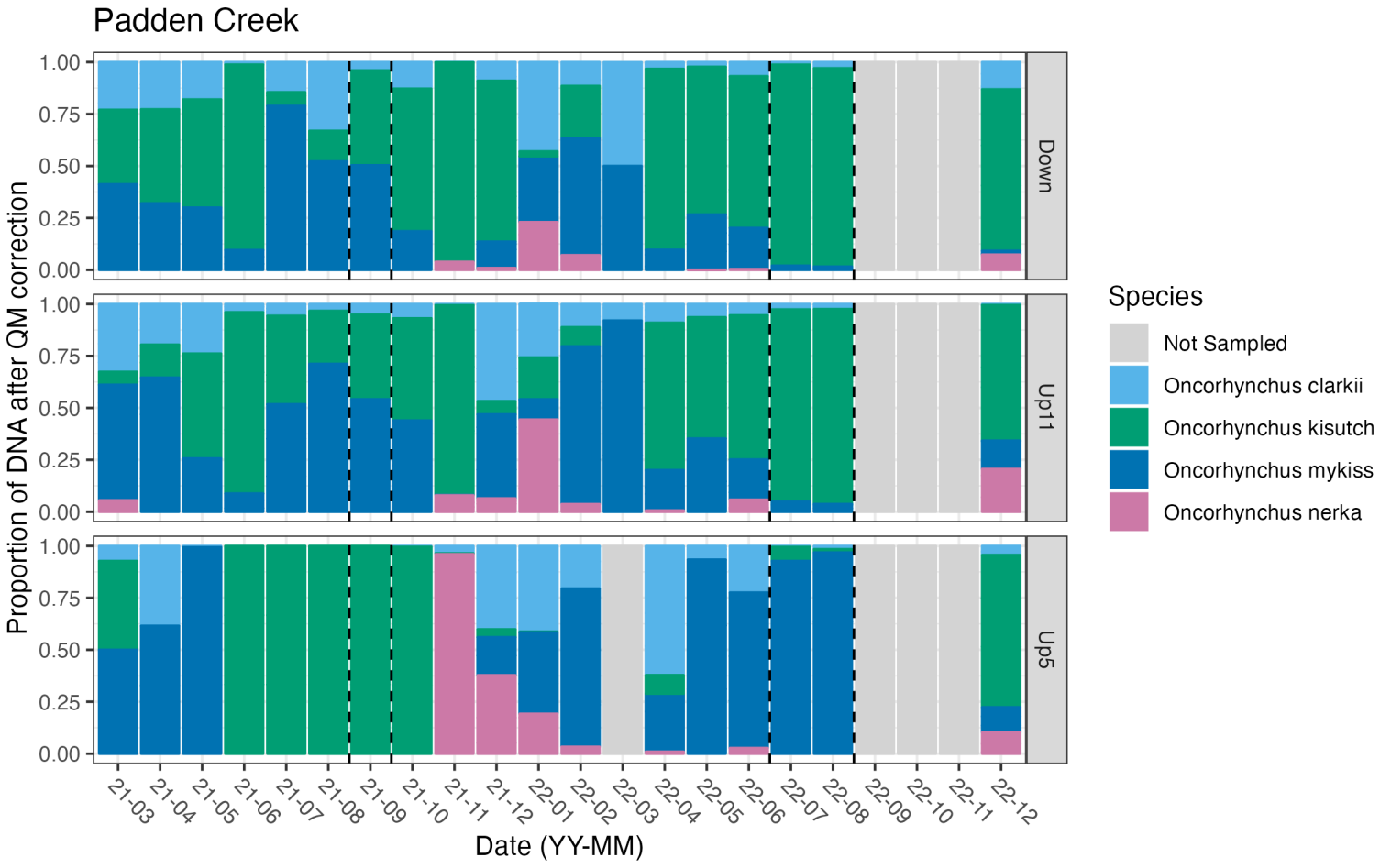
Figure 10. Effect of construction on log-fold change in eDNA mass flow rate upstream to downstream in Padden Creek. Yellow shading shows when construction started for each of the two culverts. Grey points show the corresponding log-fold changes in control creeks and black points show Padden Creek. Sockeye/kokanee salmon (*O. nerka*) was only found in Padden Creek so other creeks are not shown. Samples with very low eDNA mass flow rates (less than 150 copies/s) were removed before plotting to remove extreme proportional values due to large denominators.

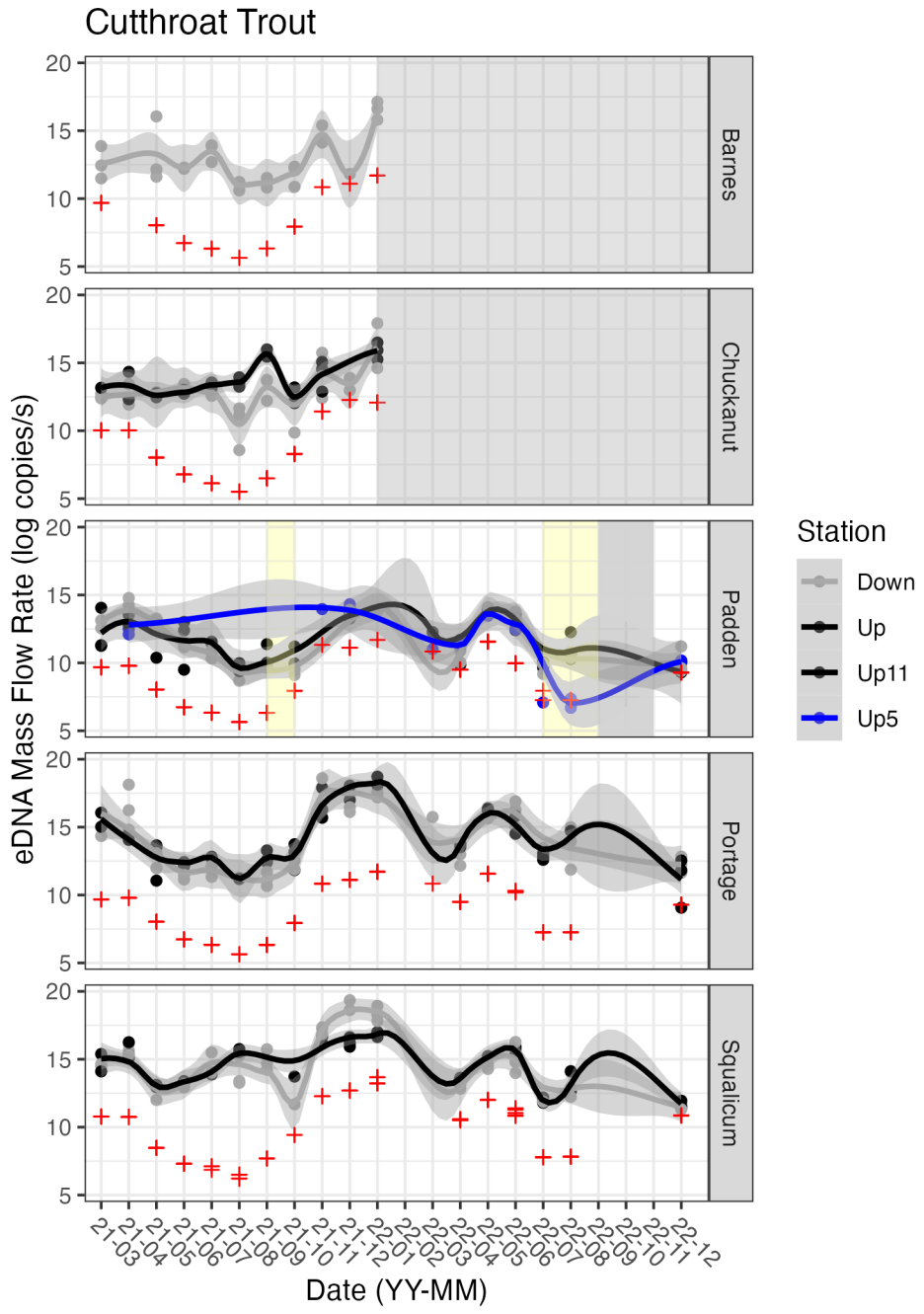

N
Down
Up SR-11
Up I-5
Squalicum
Padden
Chuckanut
Padden
Barnes
Portage
Seattle
Squalicum
Padden
Chuckanut
Barnes
20 km
49.0°N
48.5°N
48.0°N
47.5°N
123.0°W
122.5°W
122.0°W
121.5°W
121.0°W

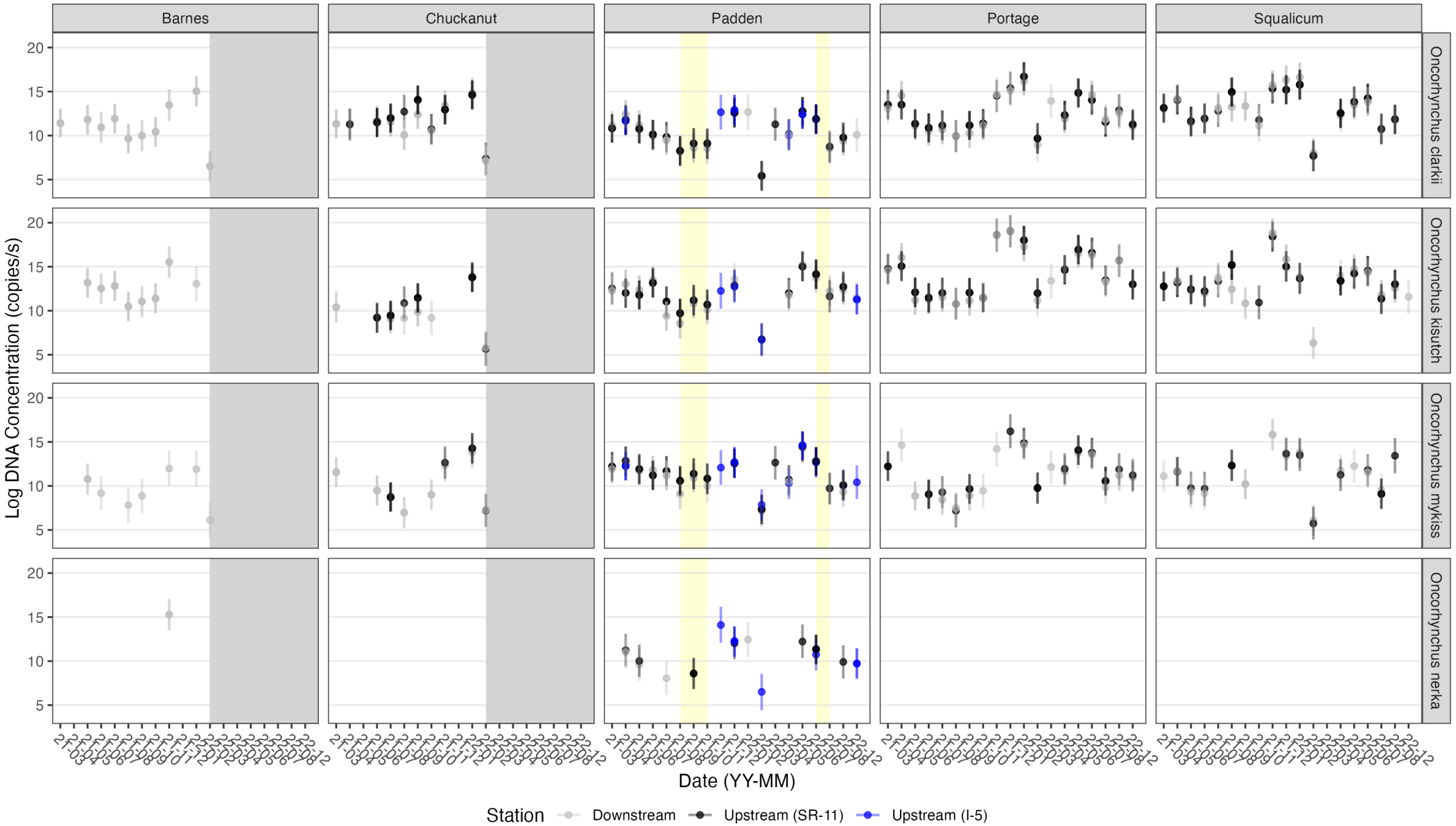

Padden Creek
Discharge (m3/s)
20
15
10
5
0
2021-07
2022-01
2022-07
2023-01
Month
Chuckanut Creek
Discharge (m3/s)
2.0
1.5
1.0
0.5
0.0
2021-07
2022-01
2022-07
2023-01
Month
Squalicum Creek
Discharge (m3/s)
80
60
40
20
0
2021-07
2022-01
2022-07
2023-01
Month

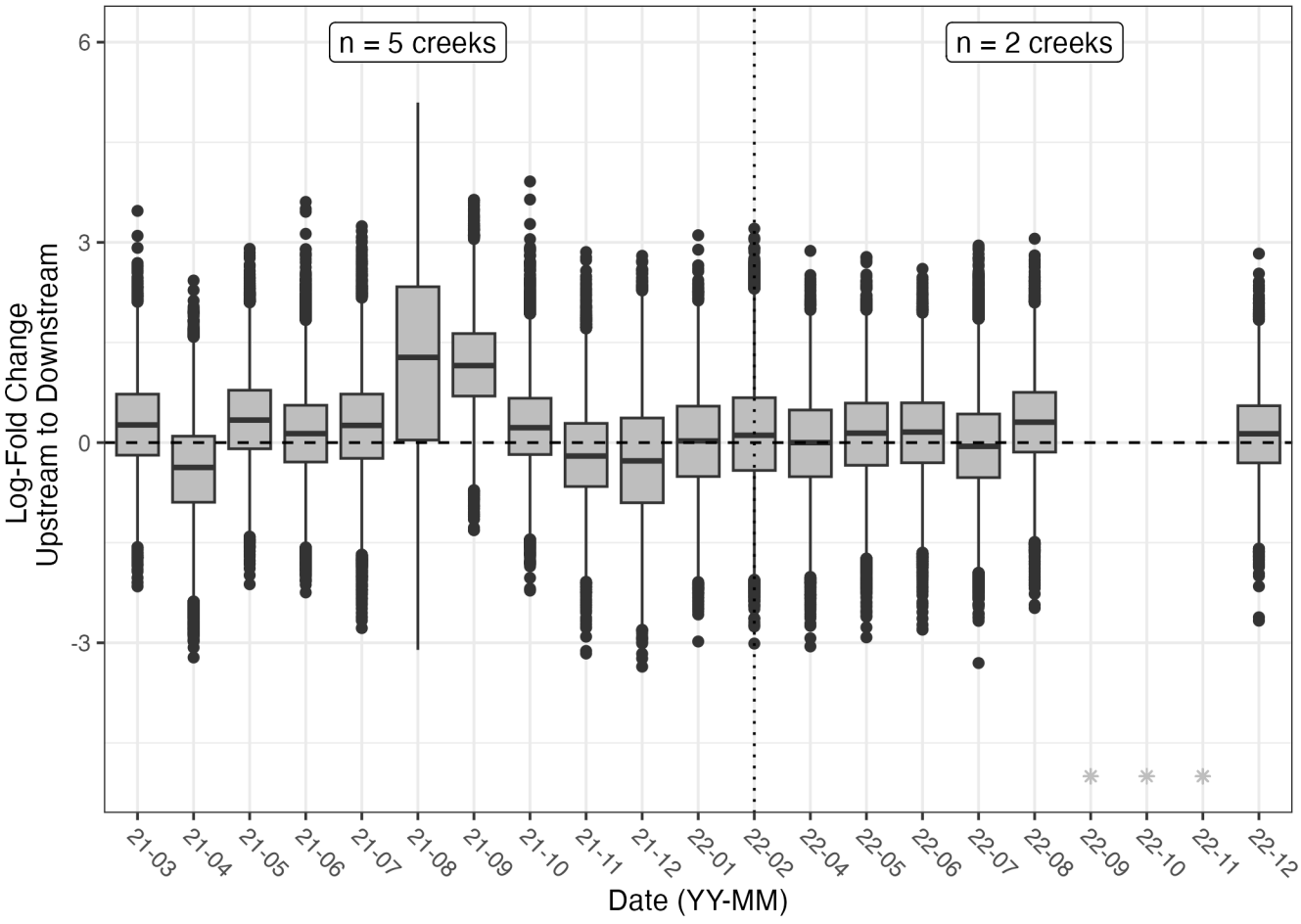

sample *i*
gBlock synthetic DNA
eDNA Extract
Tissue-derived genomic DNA
Standard curve Ct values
Single-species qPCR for cutthroat trout give cycle threshold (Ct) values
Multispecies fish metabarcoding gives number of sequencing reads
Mock community number of reads
3x tech reps 7x bio reps*
3x tech reps 3x bio reps
1x tech rep** 3x bio reps
3x tech reps 3x bio reps***
qPCR Stan model gives $C_{qPCR\ reference\ i}$
1 estimate
Metabarcoding Stan model gives proportions of amplifiable DNA
1 estimate
$C_{amplifiable\ DNA\ i}$ [copies/L] = $C_{qPCR\ reference\ i}$ /$Proportion_{reference\ DNA\ i}$
$C_{i,j}$ [copies/L] = $C_{amplifiable\ DNA\ i}$ x $Proportion_{i,j}$
$C_{i,j\ flow\ corrected}$ [copies/s] = $C_{i,j}$ x Q_{creek}
Linear mixed effects model
Positive
Threshold
Negative
Fluorescence
Cycles
$C_{qPCR\ reference\ i}$
[copies/L]
reference
$Proportion_{reference\ DNA\ i}$
$Proportion_{i,j}$
Q_{creek} [m^3/s]

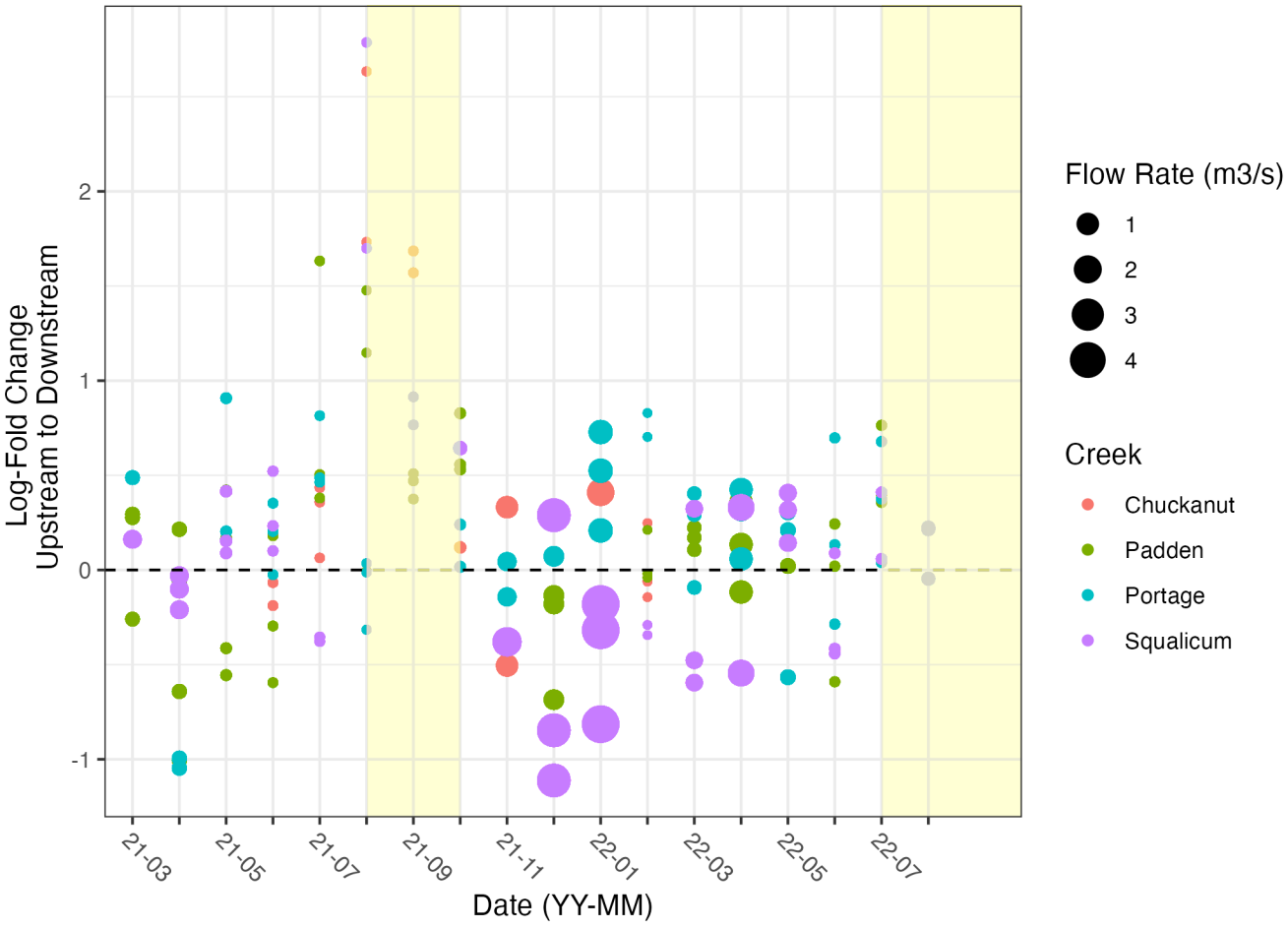

Control Creeks
Barnes
Chuckanut
Portage
Squalicum
Down
Up
Proportion of DNA after QM correction
1.00
0.75
0.50
0.25
0.00
Date (YY-MM)
Species
Not Sampled
Oncorhynchus clarkii
Oncorhynchus kisutch
Oncorhynchus mykiss
Oncorhynchus nerka

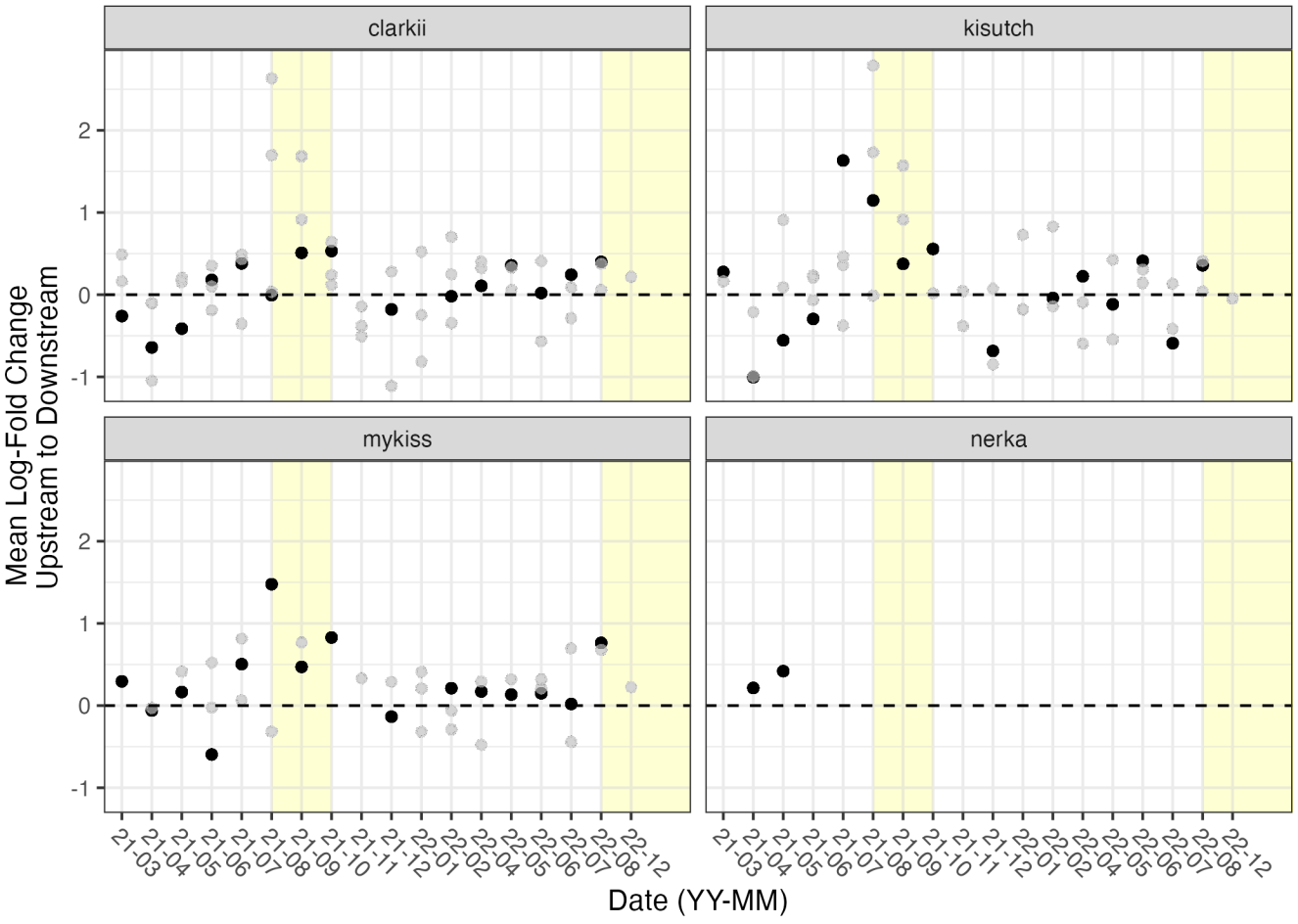

Padden Creek
Proportion of DNA after QM correction
1.00
0.75
0.50
0.25
0.00
Down
Up11
Up5
21-03
21-04
21-05
21-06
21-07
21-08
21-09
21-10
21-11
21-12
22-01
22-02
22-03
22-04
22-05
22-06
22-07
22-08
22-09
22-10
22-11
22-12
Date (YY-MM)
Species
Not Sampled
Oncorhynchus clarkii
Oncorhynchus kisutch
Oncorhynchus mykiss
Oncorhynchus nerka


Cutthroat Trout
eDNA Mass Flow Rate (log copies/s)
Barnes
Chuckanut
Padden
Portage
Squalicum
Station
Down
Up
Up11
Up5
Date (YY-MM)


Barnes
Chuckanut
Padden
Portage
Squalicum
Oncorhynchus clarkii
Oncorhynchus kisutch
Oncorhynchus mykiss
Oncorhynchus nerka
Log DNA Concentration (copies/s)
20
15
10
5
Date (YY-MM)
Station
Downstream
Upstream (SR-11)
Upstream (I-5)


n = 5 creeks
n = 2 creeks
6
3
0
-3
Log-Fold Change
Upstream to Downstream
21-03
21-04
21-05
21-06
21-07
21-08
21-09
21-10
21-11
21-12
22-01
22-02
22-04
22-05
22-06
22-07
22-08
22-09
22-10
22-11
22-12
Date (YY-MM)


Log-Fold Change
Upstream to Downstream
2
1
0
-1
21-03
21-05
21-07
21-09
21-11
22-01
22-03
22-05
22-07
Date (YY-MM)
Flow Rate (m3/s)
1
2
3
4
Creek
Chuckanut
Padden
Portage
Squalicum


clarkii
kisutch
mykiss
nerka
Mean Log-Fold Change
Upstream to Downstream
2
1
0
-1
21-03
21-04
21-05
21-06
21-07
21-08
21-09
21-10
21-11
21-12
22-01
22-02
22-04
22-05
22-06
22-07
22-08
22-12
Date (YY-MM)

Quantifying impacts of an environmental intervention using environmental DNA: Appendix S1

Elizabeth Andruszkiewicz Allan, Ryan P. Kelly, Erin D'Agnese, Maya Garber-Yonts, Megan Shaffer, Zachary Gold, Andrew O. Shelton

2023

Supplemental material for "Quantifying impacts of an environmental intervention using environmental DNA" in *Ecological Applications*

Field Sampling

Site selection and study design

There were two culverts in the treatment creek (Padden) that were suspected to be partially impassible and thus was removed and replaced during the course of the study; one of the control creeks had a bridge, which allowed fish passage (Portage), one control creek had a culvert classified as having limited fish passability (Squalicum), and two control creeks had culverts classified as preventing fish passage (Barnes and Chuckanut) (Washington Department of Fish and Wildlife 2019). These creeks were chosen due to their comparable size, flow, watersheds, and species presumed to be present to constrain as many ecological variables as possible.

Distance between sites and flow variability at sites

The average distance between upstream and downstream sampling within a creek was about 160 m; the largest distance between downstream and upstream sampling was at Barnes Creek, which was approximately 330 m, whereas the shortest distance between sampling was at Squalicum Creek at approximately 66 m.

Over the course of the year, flow within each creek varied. USGS flow gauges were located in three of the five creeks, relatively nearby to the sampling locations (Figure S1). The closest gauge to sampling locations was Padden Creek (~1.5 km); the gauge at Chuckanut Creek was ~5.5 km and the gauge at Squalicum Creek was ~7.9 km away (calculated using the Haversine distance in R).

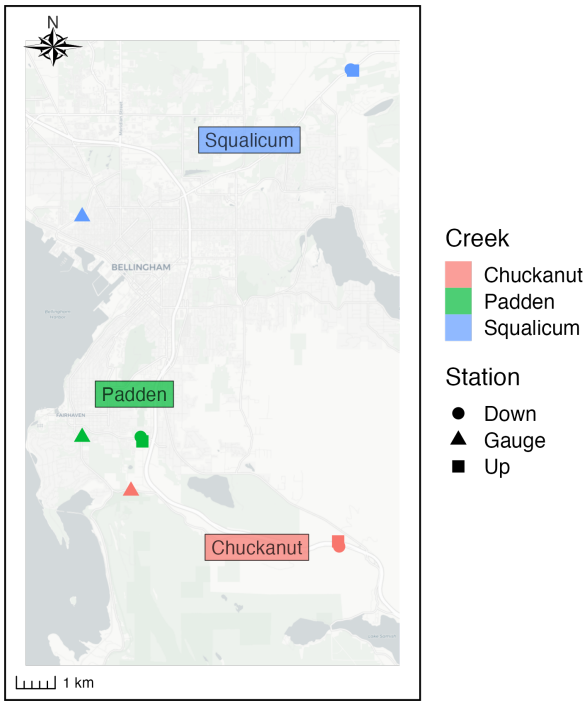


Figure S1. Location of flow gauges compared to sampling locations for Chuckanut, Padden, and Squalicum Creeks.

The flow meters at Squalicum Creek and Chuckanut Creek were offline from November 2021 for the remainder of the sampling period. The highest discharge seen during the course of the study from January to November 2021 occurred in November 2021 at Squalicum Creek. The mean discharge in each creek was: 0.42 m^3/s in Padden, 0.29 m^3/s in Chuckanut, and 1.14 m^3/s in Squalicum Creek. The lowest discharge registered by the flow meters is 0.0028 m^3/s, which occurred 8.5%, 1.6%, and 0.78% of the time in Padden, Chuckanut, and Squalicum, respectively.

Due to the lack of flow data in Squalicum and Chuckanut Creeks from November 2021 to February 2022, we used historical data from the three flow gauges to calculate the average discharge for each day of the year from about 2015-2021 (Figure S2). We then used the value for the day of the year that we sampled in either 2021 or 2022 when the gauges were offline. For consistency, we also did this at Padden Creek despite the gauge there being online for the entire sampling period.

We compared the different ways one could use flow data to correct the eDNA concentrations. We included (1) the value from the closest time point from the gauge to the time point sampling, (2) the average flow on the day of sampling from the gauge, (3) the monthly average for the month of sampling from the gauge, and (4) the correction factor approach. For (4), the values for Padden Creek represent the same as (1) for Padden Creek, the value from the closest time point in the gauge to the time point of sampling. The values

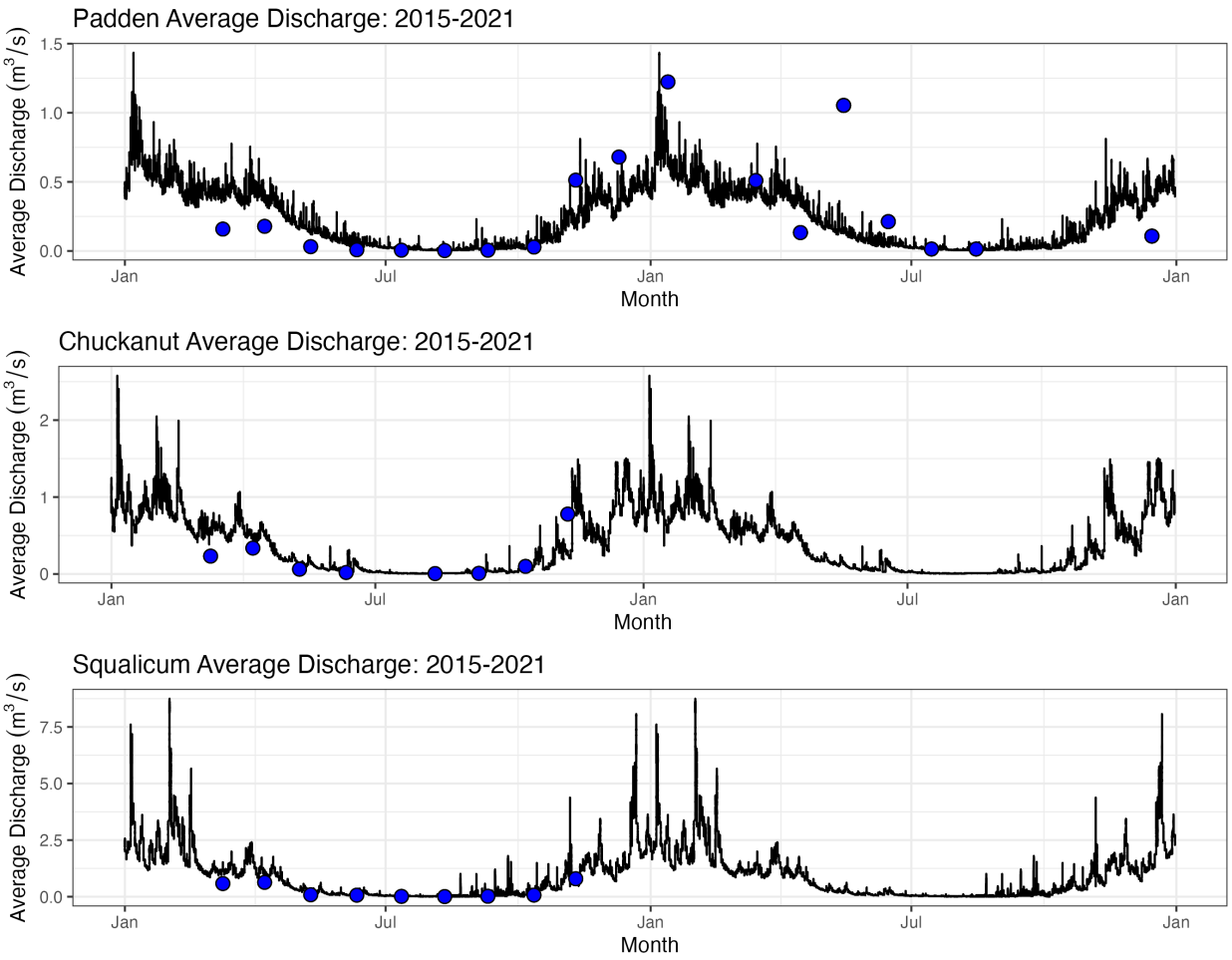


Figure S2. Daily average discharge from 2015-2020 in creeks from USGS gauges. Blue dots show the discharge time of sampling during the course of this study (2021-2022) for time points where gauges were online (note the missing data points after December through February in Chuckanut and Squalicum Creeks).

for Chuckanut and Squalicum Creek are based on the the correction factor from Padden Creek. First, five years of historical data (2015-2020) were used to find monthly averages for flow rates for each creek. Because the gauges in Squalicum and Chuckanut Creeks stopped metering in 2021, we solved for the ratio of the monthly average of each of those creeks to Padden Creek. Then, we used the closest values from Padden Creek (1) and multiplied by the monthly correction factor from the 5 years of historical data to find a value for Squalicum and Chuckanut Creeks to use for the year of sampling (2021-2022). For all three creeks, we demonstrate the relatively small changes in discharge depending on which way flow data were used (Figure S3). Though in the course of sampling, the discharge in Padden Creek ranged from no metered flow to 23 m^3/s, the discharge on the dates of sampling only reached a maximum of 1.3 m^3/s. For sites with no metered flow, half of the minimum verified discharge of the flow gauge was used (0.0014 m^3/s).

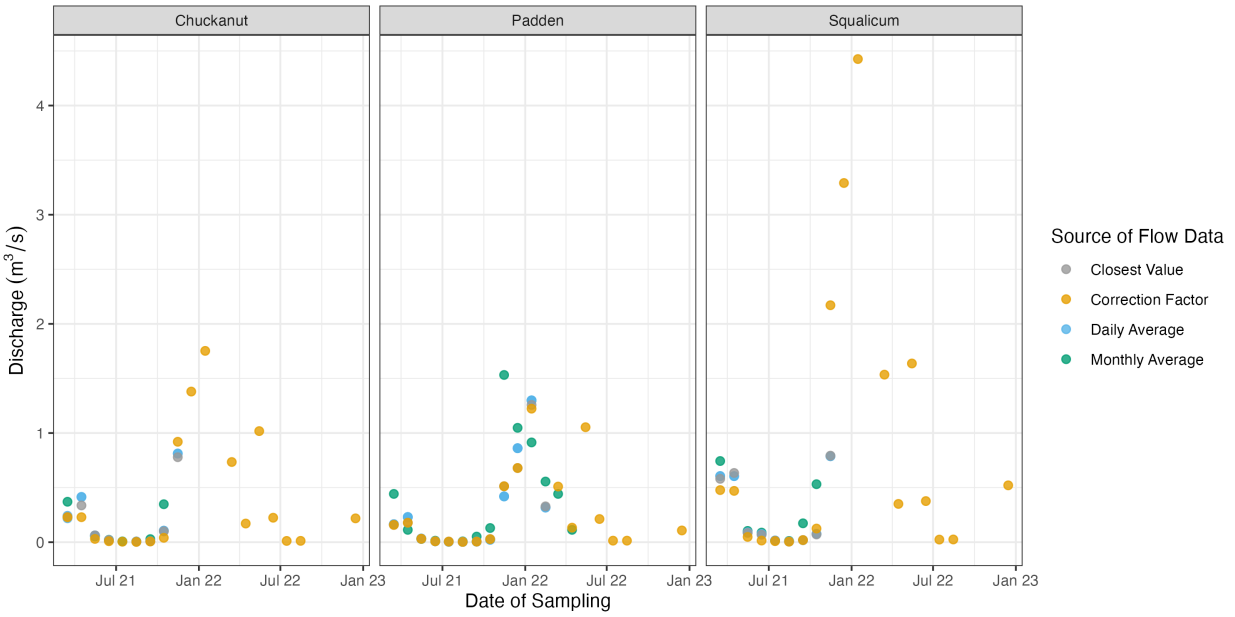


Figure S3. Comparison of different ways flow data can be used to correct eDNA concentrations. In the main text, the correction factor is used. Note that for Padden Creek, the "Correction Factor" method is the same as the "Closest Value" method. Also note that for Chuckanut and Squalicum Creeks, no data exist for cloest value, daily average, or monthly average after November 2021 when the gauges went offline.

Construction and Fish Exclusion

Fish were excluded from Padden Creek on August 30th, 2021 in preparation for the stream to be diverted on September 9th, 2021 and the diversion was removed October 7th, 2021. Water sampling occurred on September 10th, 2021, the day after the diversion, and on October 12, 2021, just 5 days after reconnecting the stream (Figure S4).

Stocking in Lake Padden

Padden Lake has historically been stocked with hatchery fish by the Washington Department of Fish and Wildlife (Figure S4). Rainbow trout (*O. mykiss*) and occasionally cutthroat trout (*O. clarkii*) and kokanee salmon (*O. nerka*) are stocked in Lake Padden, approximately 1.5 km upstream of the sampling sites. During the course of the study, rainbow trout were stocked in April 2021 and April 2022, kokanee salmon were stocked in May 2021 and October 2022, and cutthroat trout were stocked in November 2022 (Figure S4). However, despite the stocking of 30,000 kokanee salmon in May in Lake Padden, O. nerka was not detected by metabarcoding in May 2021 or at any point in 2021 until November (see main text Results). Importantly, this suggests that Lake Padden is far enough upstream that the eDNA signal at the sampling sites by the culvert is not a result of stocking the lake 1.5 km upstream (see main text Discussion for more information).

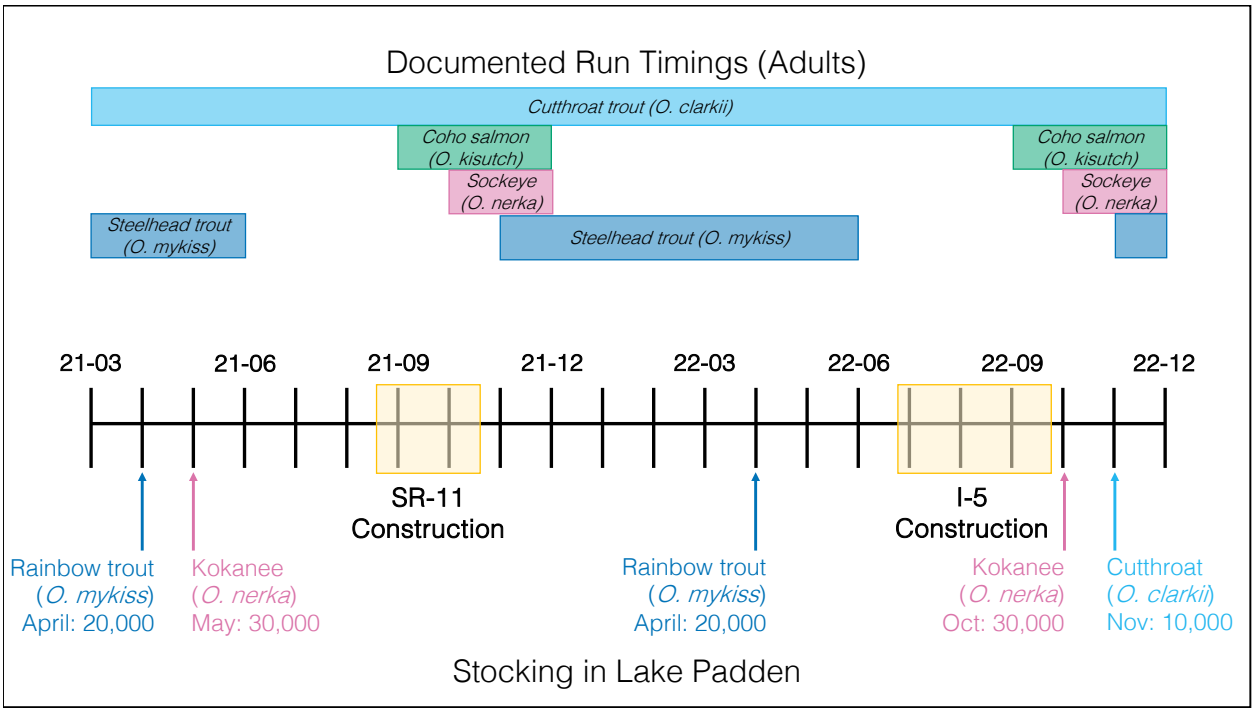


Figure S4. Timeline of runs for migrating species, stocking of Lake Padden, and construction at the intervention site (Padden Creek). Note, the dates of kokanee stocking do not include days, just the month and year. For plotting purposes, they are shown as the 15th.

Water Sampling

Water samples were collected using Smith Root's eDNA Backpack (Thomas et al. 2018), a portable pumping-and-filtering setup set to filter at 1 L/min at 82.7 kPa (12 psi). For most months, a trident sampler was used to collect all 3 biological replicates at the exact same time, for a total sampling time of about 5 minutes. Otherwise, the three replicates were collected consecutively, for a total sampling time of about 15 minutes.

Downstream sites were always sampled before upstream sites to ensure no potential DNA was introduced into the stream before sampling. In some samples, less than 2 L of water was filtered due to clogging (mean = 1.95 L).

Laboratory Processing

DNA Extraction, Amplification, Sequencing

Here, we used mock communities to determine the species-specific amplification efficiencies for each salmonid in the study. Briefly, we constructed five communities with known proportions of starting DNA from different species (total DNA as measured by Qubit). The communities ranged from having a total of 12 to 20 species, but six salmonid species were included in all five mock communities to have more information on the amplification efficiencies of salmonids (Supplemental Table 2). We sequenced these communities using the same metabarcoding primers and thermocycling conditions above and then determined the species-specific amplification rates given the discrepancy between the known starting proportion and the proportion of reads after sequencing. The mock community data were then used to correct the sequencing reads from the environmental samples to estimate the starting DNA proportions of each species in environmental samples, which is the metric of interest (Figure 3, green boxes). This is the first application of the model to correct eDNA data from water samples with mock community data as described in Shelton et al. (2022) (see Supplemental Text 2 for more information).

All molecular work prior to sequencing was performed at the University of Washington. Bench-tops were cleaned with 10% bleach for 10 minutes and then wiped with 70% ethanol. Molecular work was separated onto pre- and post-PCR benches; all DNA extractions and PCR preparation was conducted on a bench where no PCR product was handled.

DNA Extractions

We followed a protocol developed for extracting DNA off the self-preserving Smith Root filters (Thomas et al. 2019). Filters were removed from their housing with sterile tweezers and cut in half using sterile razor blades. One half was archived and the other half was used for extraction. DNA was extracted from half of each filter using a Qiashredder column (Qiagen, USA) and the DNeasy Blood and Tissue Kit (Qiagen, USA) with an overnight incubation (Thomas et al. (2019)), such that the effective filtering effort was 1 L/sample; the remaining half of each filter was archived at -20°C. Extracts were eluted in 100 μL of molecular grade water, quantified via Qubit (Invitrogen, USA) and stored at -20°C until PCR amplification within 2 months of extraction.

PCR Amplification

For the metabarcoding approach, we targeted a ~186 bp hypervariable region of the mitochondrial DNA 12S rRNA gene for PCR amplification (MiFish; Miya et al. 2015), but using modified primer sequences as given in Praebel and Wangensteen (unpublished; via personal communication) and including the Illumina Nextera overhang sequences for subsequent indexing. The primers used were as follows: F 5' *TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG*GCCGGTAAAACTCGTGCCAGC 3', R 5' *GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG*CATAGTGGGGTATCTAATCCCAGTTTG 3' (*italics* indicate Nextera overhang). PCR reactions included 10 μL of 5X Platinum ii Buffer, 0.4 μL of Platinum ii Taq, 1.25 μL of 8 mM dNTPS, 1.25 μL of 10 μM F primer, 1.25 μL of 10 μM R primer, 5 μL of template, and 30.85 μL of molecular grade water, for a total reaction volume of 50 μL. Cycling conditions were as follows: 95°C for 2 min, 35 cycles of 95°C for 30 sec, 60°C for 30 sec, 72°C for 30 sec, followed by a final extension of 72°C for 5 min.

Each month of samples was amplified on a single plate with the addition of a no template control (NTC; molecular grade water in lieu of template) and a positive control (genomic DNA from kangaroo). After PCR amplification, PCR products were visualized on a 1-2% gel. If no band was present for a given sample, a new amplification was attempted with extracts diluted 1:10 iteratively until a band was detected. PCR products were size-selected and cleaned using MagBind Beads (Omega Biotek, USA) at a sample:beads ratio of 1.2. Bead-cleaned PCR products were eluted in 30 μL of molecular grade water and quantified via Qubit (Invitrogen, USA).

An indexing PCR reaction added a unique index to each sample using Nextera indices (Illumina, USA) to allow pooling multiple samples onto the same sequencing run. For indexing, 10 ng of PCR product was used as template in a final volume of 11.25 μL. For samples with concentrations less than 0.88 ng/μL, 11.25 μL was added despite being less than 10 ng of amplicon. Each sample received a unique index; Nextera index sets A and B were used to avoid using the same index for more than one sample on a single sequencing run. The PCR reaction included the 11.25 μL of template, 12.5 μL of Kapa HiFi MMX (Roche, USA), and 1.25 μL of indexed primer. Cycling conditions were as follows: 95°C for 5 min, 8 cycles of 98°C for 20 sec, 56°C for 30 sec, 72°C for 3 min, and a final extension of 72°C for 5 min.

Indexed PCR products were also size-selected and purified using MagBind Beads (Omega Biotek, USA) at a sample:beads ratio of 0.8. Bead-cleaned PCR products were eluted in 30 μL of molecular grade water and quantified via Qubit. Indexed and bead-cleaned products were normalized before pooling into libraries, which were subsequently quantified via Qubit and visualized on a Bioanalyzer (Agilent, USA) before sequencing. Samples were randomized in 3-month blocks and each block split across 3 sequencing runs, for a total of

12 sequencing runs. The loading concentration of each library was 4-8 pM and 5-20% PhiX was included depending on the composition of the run. Sequencing was conducted using an Illumina Miseq with v3 2x300 chemistry at the NOAA Northwest Fisheries Science Center and the University of Washington's Northwest Genomics Center.

Species Specific qPCR

We quantified cutthroat trout (*O. clarkii*) DNA in each sample, targeting a 114 bp fragment of the cytochrome b gene with a qPCR assay (Duda et al. 2021). The primer/probe sequences were: F 5' CCGCTACAGTCCTTCACCTTCTA 3', R 5' GATCTTTGTATGAGAAGTAAGGATGGAA 3', P 5' 6FAM-TGAGACAGGATCCAAC-MGB-NFQ 3'. The qPCR assay was multiplexed with TaqMan Exogenous Internal Positive Control Reagents (EXO-IPC) (Applied Biosystems, USA) to check for the presence of PCR inhibitors (Duda et al. 2021). Each DNA sample was run in triplicate using Gene Expression Mastermix (ThermoFisher, USA), a final concentration of 0.375 μM F primer, 0.375 μM R primer, and 0.105 μM probe, as well as 1X EXO-IPC mix, 1X EXO-IPC DNA, 3.5 μL of template for a final reaction volume of 12 μL. The EXO-IPC mix includes the primers and probe for the EXI-IPC DNA, with the probe having a VIC reporter, allowing it to be multiplexed with the *O. clarkii* assay, which has a FAM reporter. All qPCRs were conducted on an Applied Biosystems StepOnePlus thermocycler.

Thermocycling was as follows: 50°C for 2 min, 95°C for 10 min, followed by 45 cycles of 95°C for 15 sec, 60°C for 1 min. The cycle threshold (Ct) value determined for the EXO-IPC assay from the NTC was compared to the Ct value for the EXO-IPC assay in each of the environmental samples. If the Ct value was >0.5 Ct values from the mean Ct for the NTCs, the sample was deemed inhibited and diluted 1:10 and re-assayed until the Ct value fell within the accepted range. After converting Ct values to DNA concentrations using the standard curve (see below), the concentration was multiplied by the dilution factor.

Each plate included a 8-point standard curve created using synthetic DNA (gBlocks) at the following concentrations: 100,000 copies/μL, 10,000 copies/μL, 1,000 copies/μL, 100 copies/μL, 10 copies/μL, 5 copies/μL, 3 copies/μL, 1 copy/μL Additionally, six no template controls (NTCs) were included on each plate: 3 with the IPC DNA mix and 3 with molecular grade water instead of template or IPC DNA mix. Plates were re-run if efficiency as determined by the standard curve was outside of the range of 90-110%.

To check for inhibition, the cycle threshold (Ct) value determined for the EXO-IPC assay from the NTC was compared to the Ct value for the EXO-IPC assay in each of the environmental samples. If the Ct value was >0.5 Ct values from the mean Ct for the NTCs, the sample was deemed inhibited and diluted 1:10 and re-assayed until the Ct value fell within the accepted range. The majority of environmental samples

(60%) were inhibited and accordingly diluted for analysis. In 80% of inhibited samples, a 1:10 dilution or less remedied the inhibition, but some samples (11%) required dilution by a factor of up to 1000.

Bioinformatics Processing

Primers were removed with cutadapt (Martin 2011) and then reads were de-noised, filtered, merged, and ASVs were generated using dada2 (Callahan et al. 2016). For each MiSeq run, the trimming lengths were determined by visually assessing the quality score plots. After ASVs were generated, taxonomy was assigned using the "classify" function in the insect package in R using the classifier published by the authors of the package (Wilkinson et al. 2018).

Quality Controls

Positive controls were included on each sequencing run to monitor for cross contamination that might have occurred in the laboratory or due to "tag jumping". With 13 MiSeq runs, we included one sample of kangaroo tissue on each run and then measured how many reads of kangaroo were found in environmental samples and how many reads of non-kangaroo were found in kangaroo samples (Figure S5).

We can also check to make sure that no reads assigning to kangaroo were in the environmental samples. We only found kangaroo in two environmental samples, both of which were a very small number (and proportion) of reads (2 and 136 reads found in samples with 40,425 and 28,725 reads respectively) (Figure S6).

Annotation

We first used a tree-based annotation method (insect package) and then followed up with a BLAST search for all ASVs that were not annotated to species level by insect. The percent of reads annotated did not correlate with sample read depth, creek, station, or month of sampling. Read depth across samples ranged from 1,011 to 311,879, with a mean of 79,709 and median of 75,967 reads (Figure S7). With a total of ~565 samples, 93% of samples had >20,000 reads and 65% of samples had over 50,000 reads. There did not seem to be a pattern with samples of low reads with creek or time. Additionally, of the low read depth samples (<20,000 reads, 40 samples), there was only one sample in which all three replicates were low (March 2022 Squalicum Upstream), meaning that it is very unlikely that low read depth samples would have lead to changing ecological results.

A total of 81 unique species were identified in the environmental samples by the MiFish primers, including 25 fish, 25 mammals, 23 birds , and 8 amphibians (Figure S8 and S9; see also Supplemental Table 1). Of the 81 species, 17 only were found in a single environmental sample. The three most commonly found species were coho salmon (*O. kisutch*), cutthroat trout (*O. clarkii*), and rainbow trout (*O. mykiss*).

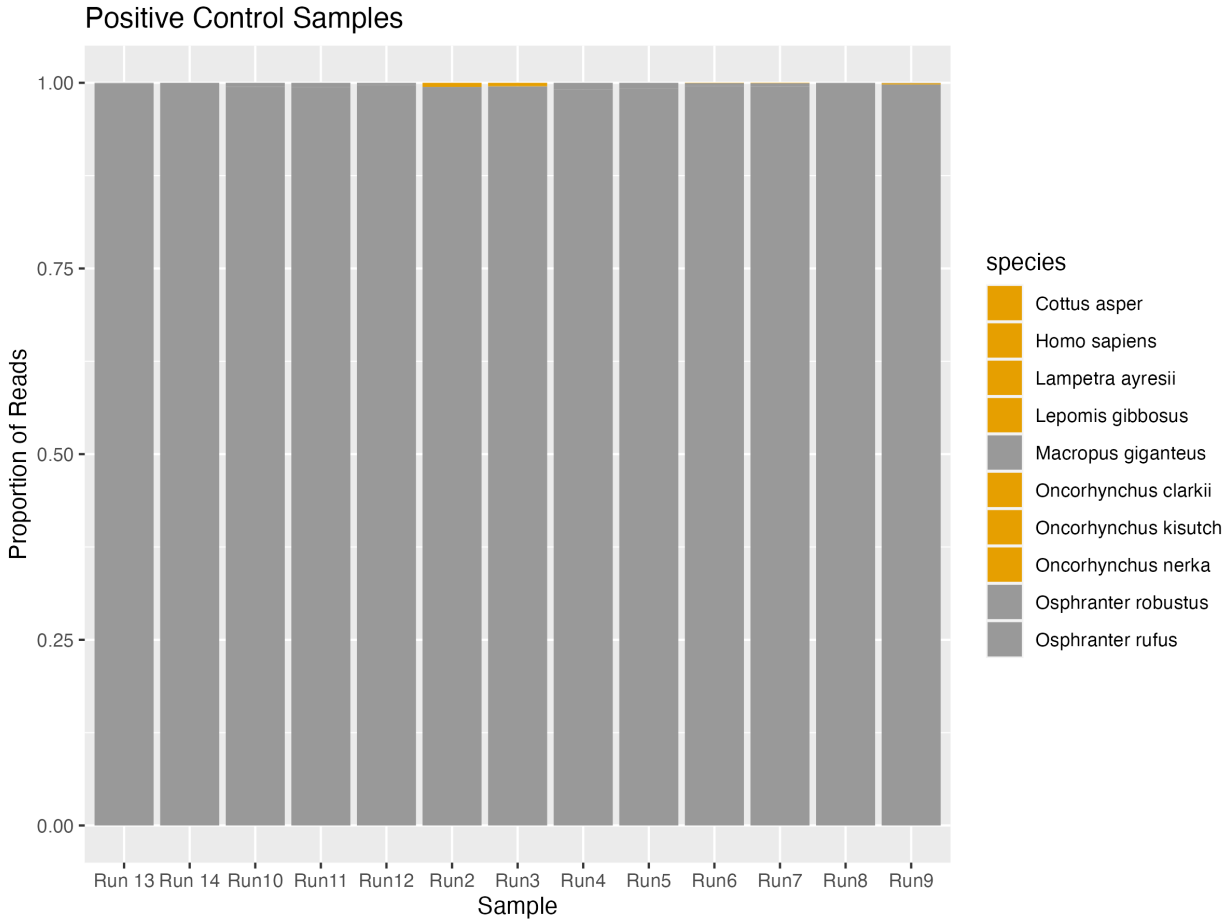


Figure S5. Proporiton of annotated reads found in positive controls. Grey colors are the three species of kangaroo used for positive controls and are what should be in each sample. Orange species should not be in the positive controls and indicate low level contamination from environmental samples.

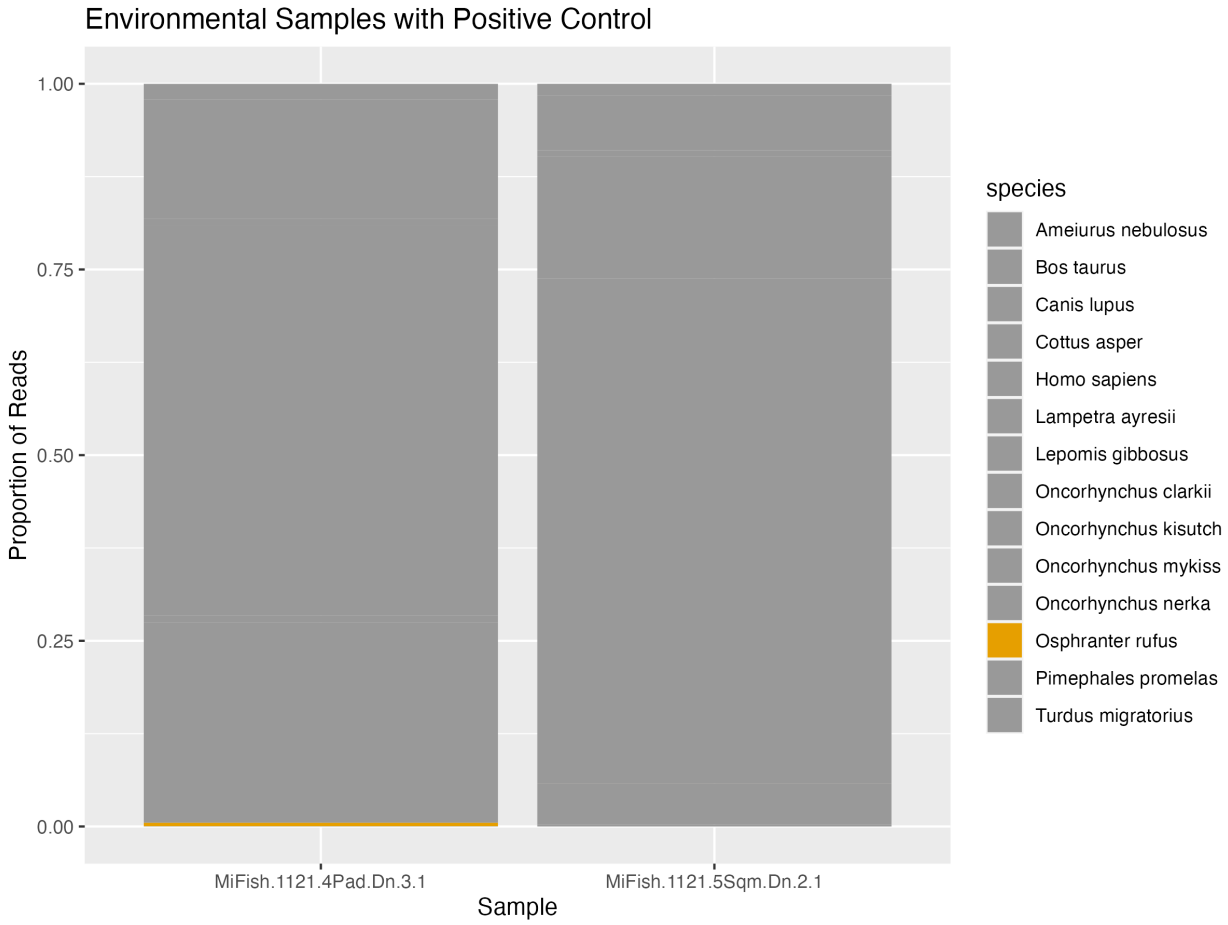


Figure S6. Proporiton of annotated reads found in environmental samples with positive control. Grey colors are non-kangaroo reads and therefore and are what should be in each sample. Orange species are kangaroo reads and therefore should not be in the environmental samples and indicate low level contamination from positive controls.

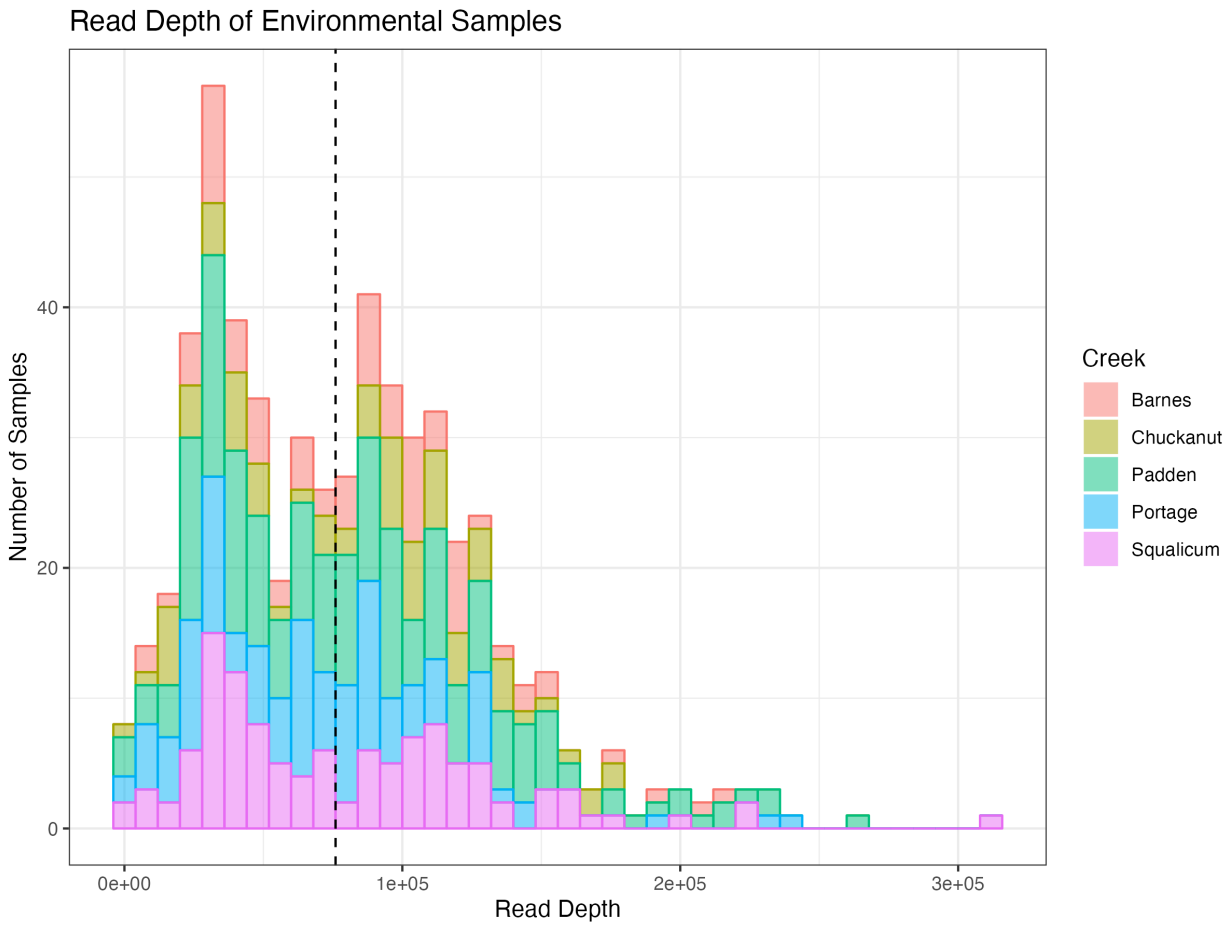


Figure S7. Read depth of samples colored by creek. Dashed line shows median read depth (87,698 reads).

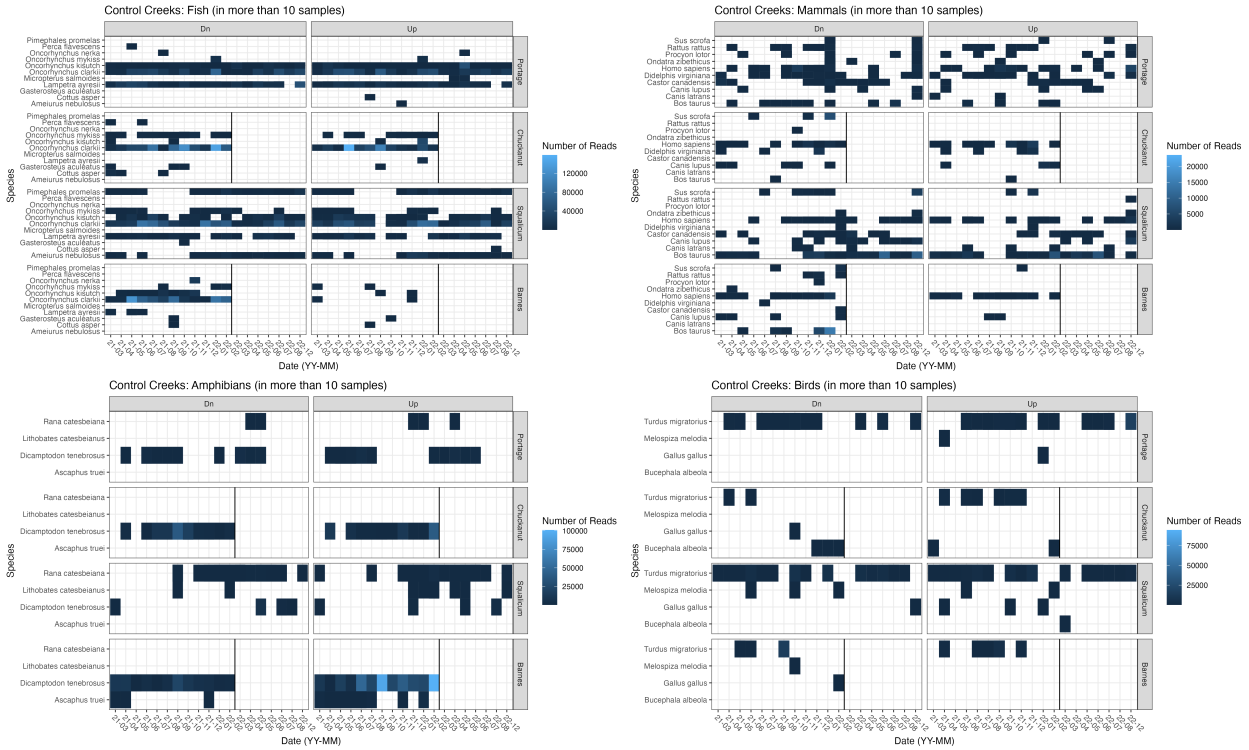


Figure S8. Heat map of all species found in control creeks in at least ten environmental samples.

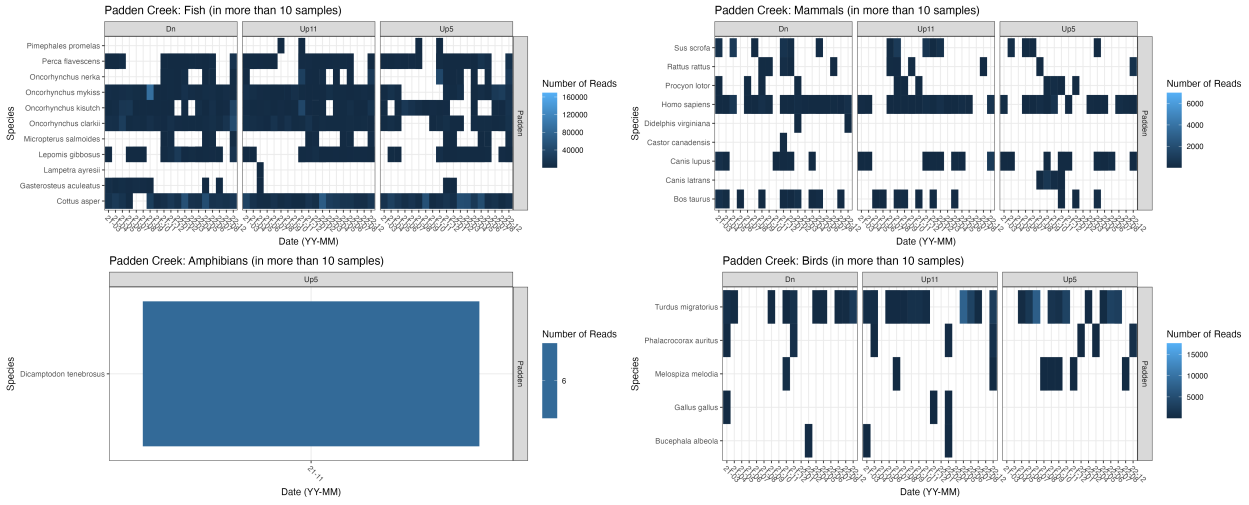


Figure S9. Heat map of all species found in Padden Creek in at least ten environmental samples.

Correcting metabarcoding data for amplification bias

Using our six mock communities (three different taxa compositions at two different proportions [even and skewed]), we can first check how well the quantitative metabarcoding model corrects for amplification bias. In one case, we consider the even mock communities as the mock community data and the skewed mock communities as unknown. We can then re-create what the model believes to be the original starting proportions of the skewed mock community given the proportions of reads found in the skewed mock communities and the proportion of DNA as compared to the proportion of reads found in the even mock communities. We can also do the same treating the skewed mock communities as known and even mock communities as unknown (Figure S10).

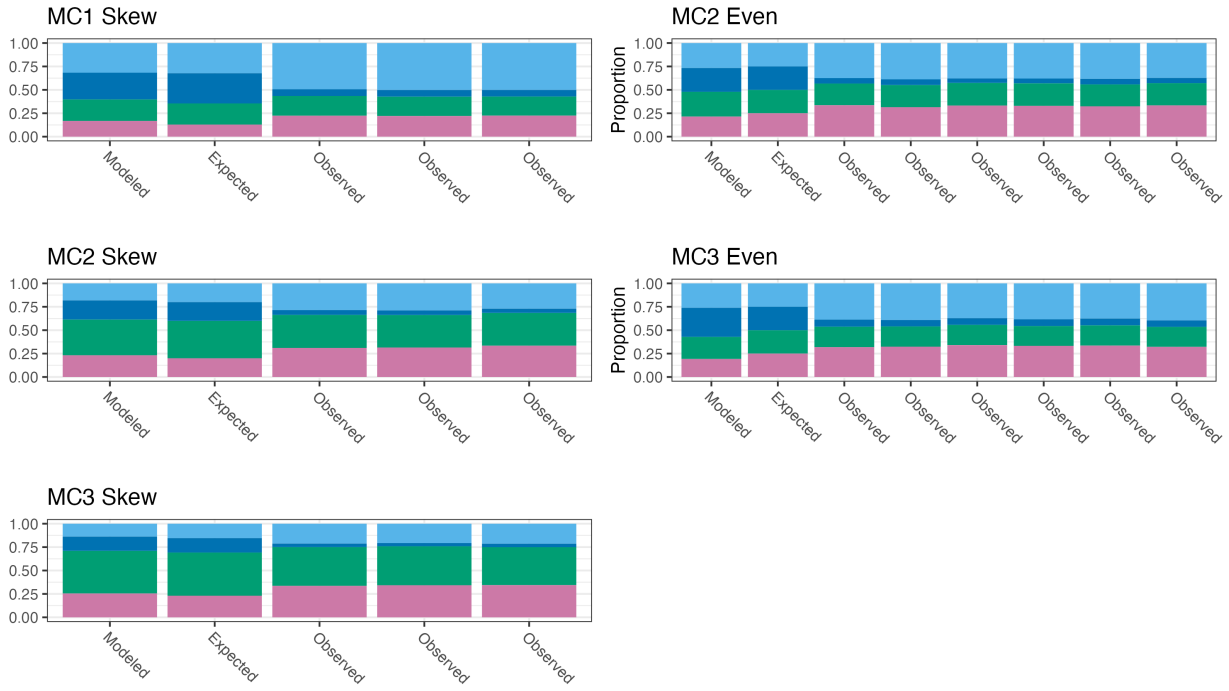


Figure S10. Intercalibration of mock communities used to correct environmental samples for amplification bias.

We can also check how well the calibration is working by comparing the alpha values by using different subsets of mock community data as true and unknown (Figure S11). We can then use the mock communities to correct the data from the MiSeq to account for the different alpha values. The corrected results are shown in the main text as Figure 4.

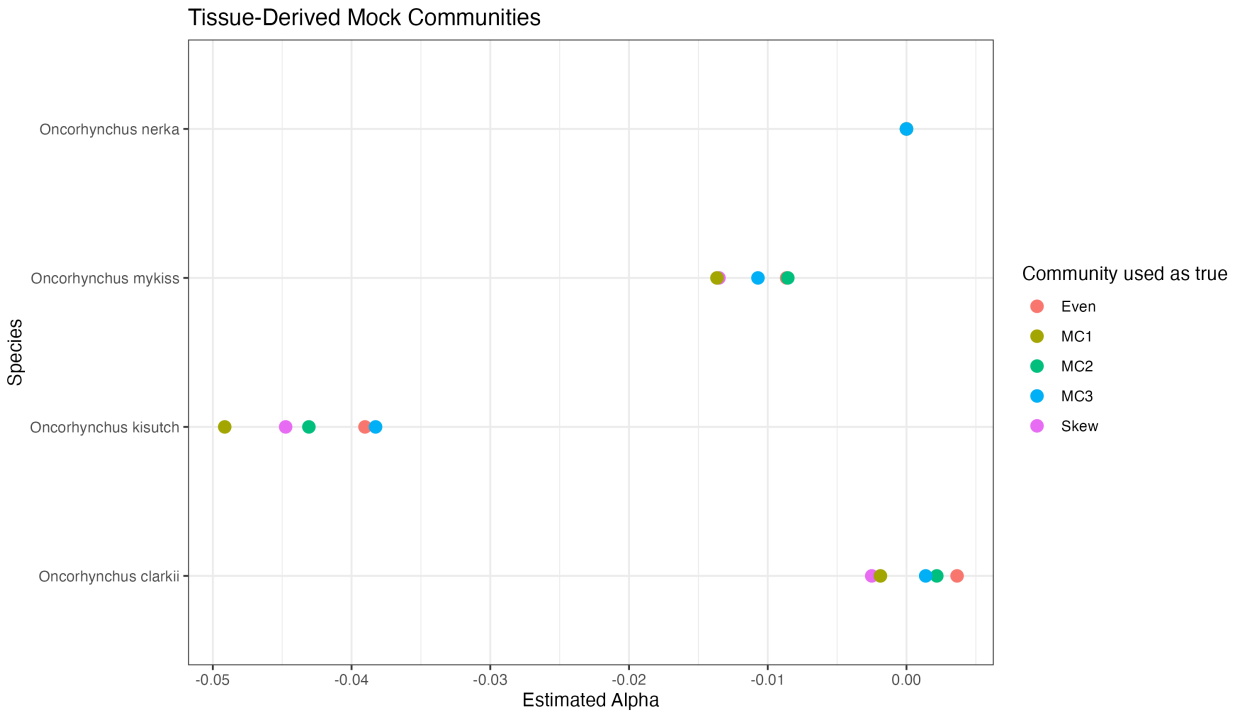


Figure S11. Estimated alpha values of salmonid species with different calibrations of the mock communities. Each color represents a different subset of mock community data treated as 'true' to calibrate the remainder of the mock community data.

Species-specific Effect of Culverts

In the main text, we show the effect of culverts averaged over creeks and species (Figure 8). Here, we show them separated by species and creek (Figure S12).

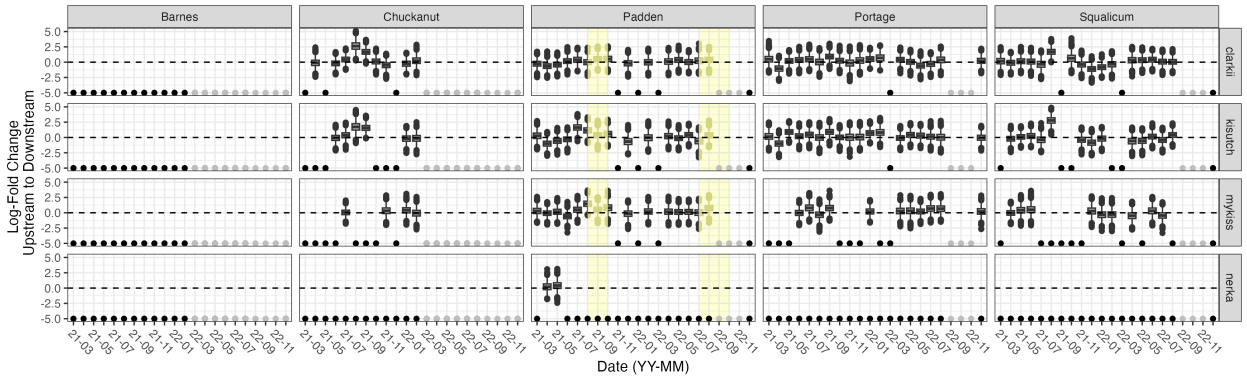


Figure S12. The effect of culvert on salmonid abundance separated by species and creeks across time. The y-axis shows the log-fold change in eDNA mass flow rate (copies/s) between upstream and downstream, normalized by upstream mass flow rate. The box boundaries correspond to the 25th and 75th percentiles; the whiskers correspond to 1.5 times the interquartile range. Here, negative values imply that eDNA mass flow rates are higher downstream than upstream. Samples with very low eDNA mass flow rates (< 150 copies/s) were removed before plotting to remove extreme proportional values due to large denominators. Grey points indicate times when no samples were taken. Black points indicate times when samples were taken, but no target DNA was found in either upstream or downstream samples and therefore the log-fold change can not be calculated.

References

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: High resolution sample inference from illumina amplicon data. Nature methods 13:581–583.

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 3:121–141.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10.

Shelton, A. O., Z. J. Gold, A. J. Jensen, E. D'Agnese, E. Andruszkiewicz Allan, A. Van Cise, R. Gallego, A. Ramón-Laca, M. Garber-Yonts, K. Parsons, and R. P. Kelly. 2022. Toward quantitative metabarcoding. Ecology n/a:e3906.

Thomas, A. C., J. Howard, P. L. Nguyen, T. A. Seimon, and C. S. Goldberg. 2018. ANDe ™: A fully integrated environmental DNA sampling system. Methods in Ecology and Evolution 9:13791385.

Thomas, A. C., P. L. Nguyen, J. Howard, and C. S. Goldberg. 2019. A self-preserving, partially biodegradable eDNA filter. Methods in Ecology and Evolution 10:1136–1141.

Washington Department of Fish and Wildlife. 2019. Fish passage inventory, assessment, and prioritization manual.

Wilkinson, S. P., S. K. Davy, M. Bunce, and M. Stat. 2018. Taxonomic identification of environmental DNA

with informatic sequence classification trees.

Title: Quantifying impacts of an environmental intervention using environmental DNA

Authors: Elizabeth Andruszkiewicz Allan
Ryan P. Kelly
Erin D'Agnese
Maya Garber-Yonts
Megan Shaffer
Zachary Gold
Andrew O. Shelton

Journal Name: Ecological Applications

Table S1. Complete list of species found over the course of the study in all environmental samples.

Table S2. Compositions of species in the mock communities used to correct metabarcoding data for relative proportions.

Table S1. Complete list of species found over the course of the study in all environmental sample

Amphibians	**Birds**	**Fish**	**Mammals**
Anaxyrus boreas	*Aix sponsa*	*Alburnus alburnus*	*Aplodontia rufa*
Ascaphus truei	*Ardea herodias*	*Ameiurus nebulosus*	*Bos taurus*
Dicamptodon tenebrosus	*Bombycilla cedrorum*	*Barbatula barbatula*	*Canis latrans*
Ensatina eschscholtzii	*Branta canadensis*	*Cottus asper*	*Canis lupus*
Lithobates catesbeianus	*Bucephala albeola*	*Cottus marginatus*	*Capra hircus*
Lithobates clamitans	*Catharus guttatus*	*Cottus pollux*	*Castor canadensis*
Rana catesbeiana	*Cinclus mexicanus*	*Entosphenus tridentatus*	*Cervus elaphus*
Taricha granulosa	*Colaptes auratus*	*Gasterosteus aculeatus*	*Didelphis virginiana*
	Corvus corax	*Gobio gobio*	*Equus caballus*
	Fulica americana	*Lampetra ayresii*	*Felis catus*
	Gallus gallus	*Lepomis gibbosus*	*Homo sapiens*
	Junco hyemalis	*Micropterus salmoides*	*Lynx rufus*
	Melospiza melodia	*Oncorhynchus clarkii*	*Myotis keenii*
	Phalacrocorax auritus	*Oncorhynchus gorbuscha*	*Myotis lucifugus alascensis*
	Pipilo maculatus	*Oncorhynchus keta*	*Neovison vison*
	Piranga ludoviciana	*Oncorhynchus kisutch*	*Neurotrichus gibbsii*
	Poecile atricapillus	*Oncorhynchus mykiss*	*Ondatra zibethicus*
	Regulus calendula	*Oncorhynchus nerka*	*Osphranter rufus*
	Regulus satrapa	*Oncorhynchus tshawytscha*	*Procyon lotor*
	Sturnus vulgaris	*Perca flavescens*	*Rattus norvegicus*
	Tachycineta bicolor	*Pimephales promelas*	*Rattus rattus*
	Tachycineta thalassina	*Prosopium williamsoni*	*Rattus sp. NH 2147*
	Turdus migratorius	*Salmo trutta*	*Scapanus orarius*
		Sardinops sagax	*Sus scrofa*
		Spirinchus thaleichthys	*Ursus americanus*

Table S2. Compositions of species in the mock communities used to correct metabarcoding data for relative proportions.

Species	Percent Composition in Each Community				
	MC1-Skew	MC2-Even	MC2-Skew	MC3-Even	MC3-Skew
Oncorhynchus clarkii	9.5%	5.6%	4.5%	5%	3.6%
Oncorhynchus kisutch	9.5%	5.6%	4.5%	5%	3.6%
Oncorhynchus tschawtscha	3.8%	5.6%	9.0%	5%	7.1%
Oncorhynchus nerka	3.8%	5.6%	4.5%	5%	5.3%
Oncorhynchus mykiss	6.7%	5.6%	9.0%	5%	10.6%
Oncorhynchus gorbuscha	19.1%	5.6%	22.4%	5%	17.7%
Salmo trutta	1.9%	0%	0%	0%	0.0%
Salvelinus malma	1.9%	0%	0%	5%	1.8%
Entosphenus tridentatus	7.6%	5.6%	13.5%	0%	0%
Lampetra ayresii	19.1%	0%	0%	5%	7.1%
Cottus asper	8.6%	5.6%	3.4%	0%	0%
Cottus marginatus	8.6%	0%	0%	5%	7.1%
Salvelinus confluentus	0%	5.6%	3.6%	5%	0.7%
Castor canadensis	0%	5.6%	6.7%	0%	0%
Turdus migratorius	0%	5.6%	1.8%	0%	0%
Ameiurus nebulosus	0%	5.6%	3.4%	0%	0%
Gasterosteus aculeatus	0%	5.6%	1.1%	0%	0%
Taricha granulosa	0%	5.6%	6.7%	0%	0%
Ondatra zibethicus	0%	5.6%	0.9%	0%	0%
Felis catus	0%	5.6%	0.9%	0%	0%
Rana pretiosa	0%	5.6%	3.4%	0%	0%
Ascaphus truei	0%	5.6%	0.9%	0%	0%
Novumbra hubbsi	0%	0%	0%	5%	2.7%
Procyon lotor	0%	0%	0%	5%	2.7%
Rana catesbeiana	0%	0%	0%	5%	8.9%
Micropterus salmoides	0%	0%	0%	5%	7.1%
Bos taurus	0%	0%	0%	5%	0.9%
Didelphis virginiana	0%	0%	0%	5%	0.9%
Neurotrichus gibbsii	0%	0%	0%	5%	0.9%
Neovison vison	0%	0%	0%	5%	0.9%
Anaxyrus boreas	0%	0%	0%	5%	3.6%
Ardea herodias	0%	0%	0%	5%	7.1%

Quantifying impacts of an environmental intervention using environmental DNA: Appendix S3

Elizabeth Andruszkiewicz Allan, Ryan P. Kelly, Erin D'Agnese, Maya Garber-Yonts, Megan Shaffer, Zachary Gold, Andrew O. Shelton

2023

Supplemental material for "Quantifying impacts of an environmental intervention using environmental DNA" in *Ecological Applications*

Our analysis depends upon a set of quantitative models, each linking our observations of metabarcoding reads or qPCR cycle-threshold values to an underlying concentration of target-species DNA in water samples.

In summary, we (1) use a mock community with a known composition to calibrate our environmental metabarcoding data as described in (Shelton et al. 2022). The result is a set of estimated proportions of DNA from each species in each sample. We then (2) relate qPCR cycle-threshold values for a reference species (here, cutthroat trout (*O. clarkii*)) from the same set of samples to a standard curve to yield quantitative estimates of the concentration of our reference species in each sample. We (3) use these absolute estimates of DNA concentration to expand the metabarcoding-derived proportion data into a complete set of quantitative estimates of DNA concentrations for each species in each sample. We account for the variable water-flow-rates of the sampled creeks by converting these concentrations from units of copies/L into units of copies/s, given an flow rate in L/s. Finally, we (4) construct a model describing changes in these species-specific concentrations over time. We give the statistical details of these steps below.

Calibration with a Mock Community

See Shelton et al. (2022); McLaren et al. (2019); Silverman et al. (2021) for similar analyses.

For ease of computation, we ran the metabarcoding-calibration model on data for each of our five creeks separately, using the same mock communities to calibrate each.

Model Diagnostics: 3 chains, 1500 iterations, for all parameters, $\hat{R} \leq 1.02$

qPCR Calibration

See Shelton et al. (2019); McCall et al. (2014) for similar analyses.

For all samples i, on qPCR plates j, we either observe ($z_{i,j} = 0$ or do not observe $z_{i,j} = 1$) amplification; we omit the subscripts i and j from the following description except where necessary for clarity. We assume an intercept of zero.

We model the probability of detection $P(z = 1))$ as a linear function of concentration and slope parameter ϕ, ($P(z = 1) = \theta = c\phi$), with a logit transform to constrain the inferred probability to between 0 and 1.

For those samples that amplify ($z = 1$), we model the observed Ct value (y) as a linear function of our parameter of interest, the log-concentration of target-species DNA under analysis (c). We treat y as drawn from a normal distribution $y \sim N(\mu_{i,j}, \sigma_{i,j}))$, where each triplicate sample on each qPCR plate has its own estimated mean and standard deviation. The means are estimated as a straightforward linear model, $\mu = \beta_{0,j} + \beta_{1,j}c$, but we allow the standard deviation to vary as a linear function of log-concentration so as to accurately capture decreasing precision with decreasing concentration: $\sigma = e^{\gamma_0 + \gamma_{1,j}c}$; we estimate these parameters as an exponent to constrain $\sigma > 0$.

Samples with known concentrations (i.e., standards) were fit jointly with unknown samples (i.e., environmental samples); because qPCR plate identity was shared among all environmental samples and standards within a plate, this has the effect of applying plate-specific slope and intercept values for the standard curve to each of the environmental samples on the plate (Figure S1).

We apply moderately informative priors that make use of background information in hand. For example, because qPCR standard curves of all kinds have slopes near -3, this slope becomes our background expectation as embodied in the prior on β_1, but the standard deviation of that prior leaves plenty of room for this background to be overwhelmed by the observed data. The same logic applies to the intercept of the standard curve, which in qPCR (for any given species) generally falls near 39 cycles, an expectation that we formalize by having β_0 drawn from a normal distribution with $\mu = 39$ and $\sigma = 3$.

Taken together with priors, the model is:

$$z_{i,j} \sim Bernoulli(\theta_{i,j})$$

$$\theta_{i,j} = logit^{-1}(\phi c_{i,j})$$

$$y_{i,j} \sim Normal(\mu_{i,j}, \sigma_{i,j}) \text{ if } z_{i,j} = 1$$

$$\mu_{i,j} = \beta_{0,j} + \beta_{1,j} c_{i,j}$$

$$\sigma_{i,j} = e^{\gamma_0 + \gamma_{1,j} c_{i,j}}$$

$$\beta_0 \sim normal(39, 3)$$

$$\beta_1 \sim normal(-3, 1)$$

$$\gamma_1 \sim normal(0, 5)$$

$$\gamma_0 \sim normal(-2, 1)$$

Model Diagnostics: 3 chains, 2500 iterations, for all parameters, $\hat{R} \leq 1.002$.

Expanding Proportions into Absolute Abundances

See Pont et al. (2022) and McLaren et al. (n.d.) (preprint) for examples of similar expansions.

As described in the main text, calibrated metabarcoding analysis yielded quantitative estimates of the proportions of species' DNA in environmental samples prior to PCR.

We then converted these proportions into absolute abundances by expansion, in light of the qPCR results for our reference species *O. clarkii*. We estimated the total amplifiable salmonid DNA in environmental sample i as $DNA_{salmonid_i} = \frac{[qPCR_{reference_i}]}{Proportion_{reference_i}}$, and then expanded species' proportions into absolute concentrations by multiplying these sample-specific total concentrations by individual species' proportions, such that for species j in sample i, $DNA_{i,j} = DNA_{salmonid_i} * Proportion_{i,j}$.

We transformed the resulting abundances to account for the creeks' flow-rates as described in the main text.

Ideally, we would have fit a joint model that simultaneously estimated species proportions (metabarcoding), absolute concentrations (qPCR), and developed the time-series trends for all species. As a practical computational matter, we had to create these models individually, which entailed some loss of information about parameter variability and cross correlation. For the mixed-effects model describing trends over time (described below), we used the product of posterior means from the metabarcoding and the concentrations

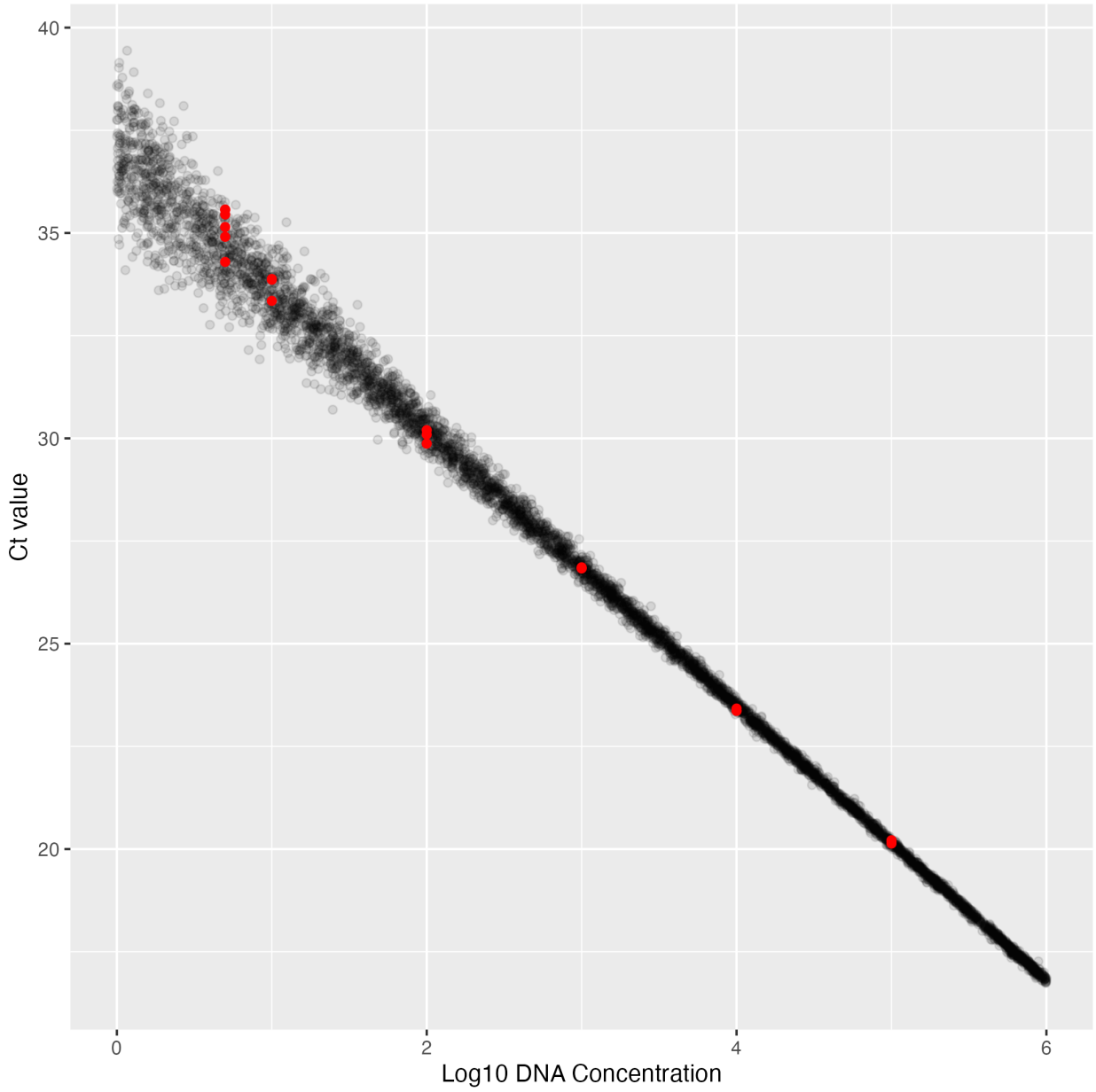


Figure S1. Example of 2500 samples from the joint posterior distribution of the model fit for a single representative qPCR plate. Red dots are standard-curve observations with known starting concentrations. The spread of black dots (posterior samples) indicates the shape of the calibration curve, with standard deviation increasing as concentration decreases.

of the qPCR model as observations, rather than being able to use the full posteriors for each input to the model. We deemed this acceptable because our metabarcoding proportions were quite precisely estimated: for example, in our focal Padden Creek, the coefficient of variation for estimated proportions of our reference species (*O. clarkii*) ranged from 0.008 to 0.25.

Modeling Changes in Concentration over Time

At a given station in a given creek, some DNA concentration exists for each species. For simplicity, we focus on a single species and a single station (downstream or upstream) for the moment.

Our observations of the (log) DNA concentration in creek i at time t are distributed as $Y_{i,t} \sim \mathcal{N}(\mu_{i,t}, \sigma^2)$. More complex versions of the model may let σ vary across creeks, time points, species, or with environmental covariates of interest.

We are interested in how the DNA concentration changes over time, so we model the expected value of DNA in a creek at time t, $\mu_{i,t}$.

We considered three ways of modeling the salmonid eDNA data, each in a Bayesian framework, but each treating non-independence among time points somewhat differently:

- A linear auto-regressive (AR(1)) model, written in `stan`. For each species in each creek, the expected concentration of eDNA of each month is a linear function of the expected value from the previous month. Within a species, the monthly autoregressive parameters are shared across creeks. For each species j – the subscript for which we omit here for clarity – we have a single overall model of the change in eDNA concentration among species, creeks (i), timepoints (t), and stations (d).

$$
\begin{aligned}
Y_{i,t,d} &\sim \mathcal{N}(\mu_{i,t,d},\, \sigma^2_{\text{obs}}) \\
\mu_{i,t,d} &= \alpha_{i,t} + \epsilon_{i,t,d} + \eta_{i,t,d} \\
\epsilon_{i,t,d} &\sim \mathcal{N}(\beta\mu_{i,t-1,d},\, \phi^2) \\
\alpha_{i,t} &\sim \mathcal{N}(\mu_\alpha, \sigma_\alpha) \\
\beta &\sim \mathcal{U}(-1, 1) \\
\sigma_{\text{obs}} &\sim gamma(1, 2) \\
\sigma_\alpha &\sim gamma(1, 2) \\
\eta &\sim \mathcal{N}(\mu_\eta, 1) \\
\phi &\sim gamma(1, 2) \\
\mu_\alpha &\sim \mathcal{N}(\mathbf{0}, \mathbf{10}) \\
\mu_\alpha &\sim \mathcal{N}(\mathbf{0}, \mathbf{5})
\end{aligned}
$$

- A generalized additive model (GAM), written in `brms` (which itself writes a `stan` model). For each species in each creek, an independent set of spline (weighting) parameters describes the temporal trends in expected eDNA concentration; the number of spline knots is shared across species and creeks. We follow (Pedersen et al. 2019) to create a hierarchical GAM in which the expected value for each species in each creek at each time point is a spline function of time, time-by-creek, and time-by-station, with random effects for creek and station. Here, time-by-creek and time-by-station splines are centered, requiring additional fixed-effect terms for station and creek. Because no information is shared across species in this model, we fit the model each species independently.

$$
\mu_{idt} = \beta_0 + s(t) + s_d(t) + s_i(t) + s(d) + s(i)
$$

In `R` code using `brms`, this model is coded as

```
brm(
  bf(
    log(observed) ~
              s(time_idx, bs="cc") +
              s(time_idx, by=station, m=1, bs="cc")+  #main effect, station
```

```
                s(station, bs="re") +  #random effect, station
                s(time_idx, by=creek, m=1, bs="cc")+  #main effect, creek
                s(creek, bs="re") + #random effect, creek
  )
)
```

- A linear mixed-effects (LME) model, written in `rstanarm`. For each species in each creek, time (i.e., sampling month) is treated as a random effect. Each species-creek-month effect is treated as an independent draw from a common distribution.

$$\mu_{ijdt} = \beta_0 + \beta_{1_{ij}} + Month * \beta_{2_{ij}} + \beta_{3_{ijdt}}$$

In `R` code using `rstanarm`, this model is coded as

```
stan_glmer(log(observed) ~ (1 + time_idx|creek:species) + (1|station:creek:species:time_idx)
```

Ultimately, the three models yielded very similar results (Figure S2). The LME model proved simplest and most flexible insofar as it could easily handle datasets with uneven sets of observations – for example, cases in which a species was detected downstream of a barrier, but not upstream. We accordingly used the LME as the model for the analysis given in the main manuscript.

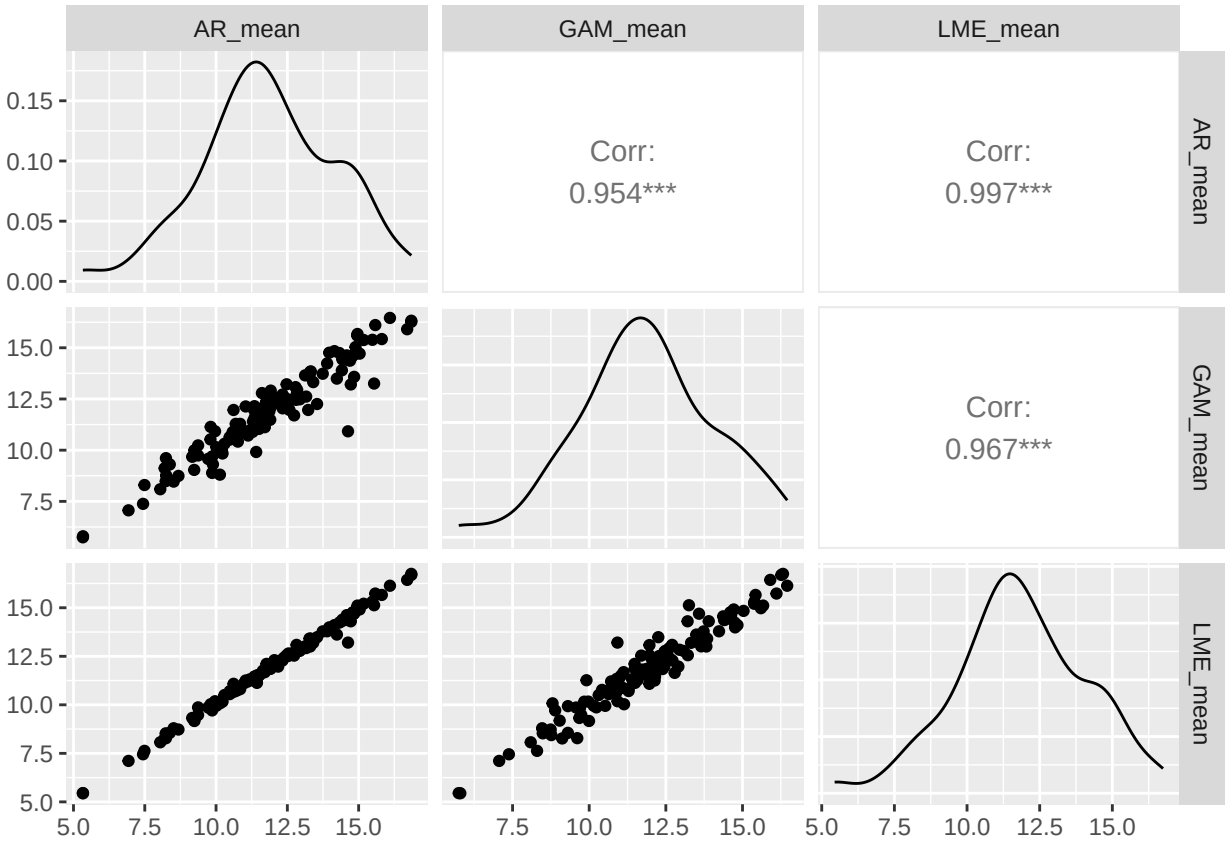


Figure S2. Comparison of three models (linear autoregressive, generalized additive model, and linear mixed effects model) shown for a subset of the data used in the main manuscript.

References

McCall, M. N., H. R. McMurray, H. Land, and A. Almudevar. 2014. On non-detects in qPCR data. Bioinformatics 30:2310–2316.

McLaren, M. R., J. T. Nearing, A. D. Willis, K. G. Lloyd, and B. J. Callahan. (n.d.). Implications of taxonomic bias for microbial differential-abundance analysis.

McLaren, M. R., A. D. Willis, and B. J. Callahan. 2019. Consistent and correctable bias in metagenomic sequencing experiments. eLife 8:e46923.

Pedersen, E. J., D. L. Miller, G. L. Simpson, and N. Ross. 2019. Hierarchical generalized additive models in ecology: an introduction with mgcv. PeerJ 7:e6876.

Pont, D., P. Meulenbroek, V. Bammer, T. Dejean, T. Erős, P. Jean, M. Lenhardt, C. Nagel, L. Pekarik, M. Schabuss, B. C. Stoeckle, E. Stoica, H. Zornig, A. Weigand, and A. Valentini. 2022. Quantitative monitoring of diverse fish communities on a large scale combining eDNA metabarcoding and qPCR. Molecular Ecology Resources n/a.

Shelton, A. O., Z. J. Gold, A. J. Jensen, E. D'Agnese, E. Andruszkiewicz Allan, A. Van Cise, R. Gallego, A. Ramón-Laca, M. Garber-Yonts, K. Parsons, and R. P. Kelly. 2022. Toward quantitative metabarcoding. Ecology n/a:e3906.

Shelton, A. O., R. P. Kelly, J. L. O'Donnell, L. Park, P. Schwenke, C. Greene, R. A. Henderson, and E. M. Beamer. 2019. Environmental DNA provides quantitative estimates of a threatened salmon species. Biological Conservation 237:383–391.

Silverman, J. D., R. J. Bloom, S. Jiang, H. K. Durand, E. Dallow, S. Mukherjee, and L. A. David. 2021. Measuring and mitigating PCR bias in microbiota datasets. PLoS Computational Biology 17:e1009113.