ABSTRACT

Title of Thesis:	INVESTIGATING THE UTILITY OF ENVIRONMENTAL DNA ANALYSIS FOR THE MONITORING AND MANAGEMENT OF MID-ATLANTIC ALOSINE FISHES
	Chelsea Fowler, Master of Science, 2023
Thesis Directed By:	Louis Plough, Associate Professor, Horn Pt. Laboratory, University of Maryland Center for Environmental Science

Environmental DNA (eDNA) tools can address gaps in fish assessment data while reducing the cost and the impact of sampling on threatened anadromous alosine fishes in Chesapeake Bay. Here, I tested the ability of high-frequency eDNA sampling of river herring to predict fish abundances from sonar-based fish counts on the Choptank River and developed and validated novel species-specific eDNA assays for American and hickory shads. River herring eDNA concentrations from daily eDNA sampling were highly correlated to sonar-based fish counts (Spearman's Rho = 0.84). This relationship informed a model that could accurately predict fish count from eDNA and relevant covariates ($R^2 = 0.88$). The two new shad assays are highly specific and quantitative, and field testing validated detections in Delaware, Maryland, and North Carolina. This work provides a set of eDNA monitoring tools for the Mid-Atlantic alosines and highlights the capacity for eDNA data to generate quantitative metrics of fish abundance.

INVESTIGATING THE UTILITY OF ENVIRONMENTAL DNA ANALYSIS FOR THE MONITORING AND MANAGEMENT OF MID-ATLANTIC ALOSINE FISHES

by

Chelsea Michael Fowler

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2023

Advisory Committee:

Associate Professor Louis Plough, Chair

Associate Research Professor Geneviève Nesslage

Assistant Professor Rodney Richardson

Dr. Matthew Ogburn, Smithsonian Environmental Research Center

Dedication

This thesis is dedicated to Carolyn Fowler. Who sat through countless hours of a curious child asking a million and one questions about anything she could think of. You encouraged me to explore the world and find the answers on my own, leading me into an adventure and learning-filled life. If she were here, she would simply say, ‘Well, that’s what moms do!’. I aspire to bring your spirit into every aspect of my life.

Acknowledgements

This thesis is the product of the collaboration and support of numerous people and groups. I want to express thanks to the following:

My advisor, Dr. Louis Plough, who provided me with this life-changing opportunity and whose continued support has allowed me to explore my research interests and make progress on my professional goals. My committee, whose expertise and support along the way has been integral in my academic journey. The Smithsonian Environmental Research Center Fish Conservation Lab, who was a vital part of the high-frequency sampling work conducted on the Choptank River and who's contributions and feedback made this project a success. Thomas Fisher, Ericka Koontz, Benjamin Gregory and Catherine Fitzgerald for their contributions to the development and implementation of the eDNA field study conducted here.

I would also like to acknowledge and thank the Smithsonian Environmental Research Center, and Maryland Department of Natural Resources, for their valuable resources and contributions. Finally, Maryland SeaGrant, and University of Maryland Center for Environmental Science, for the funding that made this work possible.

Table of Contents

List of Tables

List of Equations

List of Figures

Chapter 1: Introduction to Environmental DNA Monitoring of Anadromous Fishes

Environmental DNA: Background, History, Ecology and Detection Techniques

DNA, eDNA, and Identifying the Source of Released Genetic Material

Deoxyribonucleic acids (DNA) are molecules found in cells that store and transmit hereditary information for all living organisms (Avery et al., 1943). DNA is composed of four bases; adenine, thymine, guanine, and cytosine (Watson & Crick, 1953), the sequence of which gives rise to genes that code for amino acids, the basic building blocks for proteins that underpin all of the biological processes an organism needs to function properly (e.g., Travers & Muskhelishvili, 2015). Since DNA is inherited from parent to offspring, the arrangement of DNA bases within genes is similar among related individuals. Regions of the genome can be similar among organisms of the same species or, more broadly, of the same taxonomic group. This makes DNA an effective identification tool, exemplified by its widespread use in forensic science to match an individual's unique DNA 'fingerprint' to that found in biological evidence (Budowle & Van Daal, 2009). While forensics typically relies on specimen-derived DNA, DNA can be extracted from other sources.

Environmental DNA (eDNA) is a general term for genetic material found in the environment (e.g., soil, ice, water, etc.; Lodge et al., 2012). While specimen-derived DNA is highly concentrated and typically from a single individual, eDNA is highly fragmented, from various

sources, and processed without isolating specific organisms (Mauvisseau et al., 2022). Sources of eDNA are highly diverse. From living macrofauna, eDNA can come from urine, feces, tissues, mucus and other secretions, and reproductive cells (Turner et al., 2014). eDNA can also come from dead individuals (Tillotson et al., 2018) or from secondary sources like the feces of a predator or scavenger (Barnes & Turner, 2016). Both biotic factors, like shedding rate and eDNA source, and abiotic environmental factors, like temperature and stream discharge, influence the detectability of eDNA in the environment.

The ecology of eDNA and its impact on detection

Estimates for eDNA shedding rates vary widely among taxa and are dependent on life stage, the size, and density of individuals in the sampling area, and the physiological status of individuals (Andruszkiewicz Allan et al., 2021; Maruyama et al., 2019; Jo et al., 2019; Tillotson et al., 2018; Yates et al., 2021). Once shed by the organism, eDNA is dispersed into the environment and begins to degrade. eDNA persistence and transport dynamics are highly dependent on local hydrological characteristics, and various lab- and field-based experiments have been aimed at better understanding the complexities of hydrodynamic dispersion of eDNA in riverine environments (Carraro et al., 2021; Curtis et al., 2020; Thalinger et al., 2019). eDNA is transported away from the source organism by water currents, after which it is subject to processes such as particle deposition, resuspension, and transport (Harrison et al., 2019); it can settle out of the water and be preserved in riverbed sediment, and it can become diluted by high flow events (Curtis et al., 2020).

eDNA degradation occurs at different rates depending on the state of eDNA (Mauvisseau et al., 2022) and environmental characteristics like pH, UV exposure, and temperature (Yates et al., 2021). Andruszkiewicz-Allan et al. (2021) evaluated how eDNA shedding and decay occurred across different taxonomic groups and temperatures. They found that the decay rates for mummichog eDNA increased with increasing temperature and were consistent with those found in other studies focused on fish. However, the slopes and intercepts of decay among fish species varied widely, suggesting that eDNA decay is mediated by variables aside from temperature.

Depending on how eDNA is sampled from the environment, decay can continue to occur after collection before it is filtered from the water. Progress has been made in reducing degradation between the time of collection and the time of filtration by placing water samples on ice immediately after sampling and filtering within 16-24 hours (Goldberg et al., 2016), adding preservatives to samples immediately after collection (Jo et al., 2019; Yamanaka et al., 2017), or conducting filtration in the field and preserving filter papers (Goldberg et al., 2016). Despite the complexities associated with eDNA persistence and detectability, eDNA signals can be detected in the environment and have been used in a variety of ways to gather information on natural environments and local macrofauna.

Techniques used to detect eDNA: single-species qPCR and metabarcoding

The utility of eDNA for species detection was first demonstrated by microbiology, where microbial communities were characterized using DNA extracted from marine sediments (e.g., Ogram et al., 1987; Venter et al., 2004). eDNA detection tools were first used for macro-

organisms by Ficetola et al. (2008) to detect the American bullfrog (*Rana catesbeiana*) from water samples taken in mesocosm and the field. These studies exemplify the two types of eDNA analysis most commonly used: metabarcoding community composition and species-specific detection.

Both metabarcoding and species-specific detection rely on similar sample collection, filtration, and extraction methods. Samples for eDNA analysis in the aquatic environment typically require an environmental water sample (0.5-2 L in volume, Tsuji et al., 2019), and eDNA is isolated from the water by filtration, using filters whose material and size are well-suited for attracting and retaining DNA while mitigating filter clogging. Turner et al. (2014) found that filters with pore sizes from 1-10 μm retained most of the target eDNA while reducing the membrane clogging that slows water filtration. Renshaw et al. (2015) evaluated four different filter types and found cellulose nitrate (CN) and polyethersulfone (PES) filters recovered the most eDNA from a mesocosm. eDNA sample extraction is commonly conducted using commercially available DNA extraction kits (Shu et al., 2020). Phenol-chloroform-isoamyl alcohol (PCI) extraction methods in conjunction with room temperature preservation buffers have gained traction for their high yield (leading to higher detection rates), the potential to reduce the cost per sample of extraction, and the high recovery of eukaryotic DNA (Deiner, 2015; Renshaw et al., 2015). After extraction, genetic markers are used to isolate and amplify eDNA depending on the purpose of the study.

eDNA studies typically use polymerase chain reaction (PCR) to amplify a genetic marker (region of DNA) for a desired target species or taxonomic group (e.g., Thomsen & Willerslev, 2015). The genetic targets and specificity of the primers used in PCR-based studies fall into two general categories: metabarcoding of multiple species or communities or single-species detection. Metabarcoding uses generic or 'universal' primers to amplify a taxonomic group of interest, followed by next-generation sequencing to assign species to all unique sequences found in a sample (Yao et al., 2022), while species-specific PCR approaches are designed to amplify a single species from a sample typically using quantitative PCR (qPCR; Tsuji et al., 2019). Single-species analysis has been the predominant eDNA tool of interest for studies focused on monitoring fish populations as their results are straightforward, simple to interpret in a management context, and easier to relate to fish abundance (Thomsen & Willerslev, 2015).

Single-species detection assays are designed and optimized to maximize target species detection and eliminate the amplification of closely related species, after which they are tested on eDNA samples from a mesocosm or the field. Quantitative PCR (qPCR) is often used for single-species detection as it quantifies the eDNA as a sample copy number or eDNA concentration by comparing it to a standard curve (Tsuji et al., 2019). Metrics of detection and quantification are established, and ecological and physical factors affecting the detectability of eDNA are investigated. After an eDNA tool is considered operational, it can detect the presence or absence of the target species with reasonable accuracy (Thalinger et al., 2021). Analyses of fishes have been a persistent interest in eDNA research (Tsuji et al., 2019) due to the potential of eDNA abundance data to revolutionize the management of aquatic resources (Lodge et al., 2012). The

detection capabilities of eDNA analysis have been established (Deiner et al., 2015), and eDNA tools are already at a point where they can be utilized for routine monitoring (Thalinger et al., 2021). The ability of eDNA to provide measures of abundance has been examined across various taxa and is a promising direction for single-species detection (Yao et al., 2022). However, the use of eDNA as a fish abundance monitoring tool still requires further refinement (Lacoursière-Roussel & Diener, 2019; Yates et al., 2019). Anadromous fish are an ideal species to evaluate the quantitative abilities of eDNA tools due to the large fluctuations in eDNA concentration they evoke in freshwater systems during spawning migrations (Rourke et al., 2022).

The Current State of eDNA Monitoring for Anadromous Fishes

Anadromous fish are well suited for eDNA monitoring due to their strong pattern of presence and absence in riverine systems during their spawning migration seasons. Anadromous fish also provide high-quality ecosystem services like acting as a forage base for predators (Hall et al., 2012), transferring nutrients between marine and riparian ecosystems (West et al., 2010), and several species (e.g., Pacific salmon; Criddle & Shimizu, 2014) support high-value fisheries. Maintaining these ecosystem services and sustainable harvesting requires careful monitoring of temporal trends in abundance. This has typically been achieved using traditional methods such as visual counts, electrofishing, netting, and trawling (Rourke et al., 2022), and there is interest in using eDNA as a proxy for conventional metrics of abundance (Levi et al., 2019; Searcy et al., 2022; Yao et al., 2022).

Detection of anadromous species

eDNA analysis can provide presence and abundance information for a species of interest that can extend over a larger area, both spatially and temporally, than is feasible with traditional methodologies. Pfleger et al., 2016, developed eDNA qPCR assays for endangered sturgeon species in the Moblie River Basin, Alabama, and detected the Alabama and Gulf sturgeon (*Scaphirhynchus suttkusi, Acipenser oxyrinchus desotoi*) in 17 and 43 of the 100 field samples collected, respectively. This was a massive improvement over traditional quantification methods for these species. For the Alabama sturgeon in particular, from 1997-2005, only five individuals were collected via traditional methods. eDNA monitoring has also been used to identify overwintering habitat in striped bass (*Morone saxatilis*; Harris et al., 2022) and to evaluate the distribution of the endangered chinook salmon (*Oncorhynchus tshawytscha*) in subbasins of the Columbia River (Laramie et al., 2015).

eDNA studies of anadromous fishes can also produce information on the sources and rates of eDNA generation, which can help with the downstream interpretation of eDNA results. Plough et al., 2021, developed a qPCR single-species detection assay for the endangered Atlantic sturgeon (*Acipenser oxyrhynchus oxyrhynchus*) that was tested and validated using a pond mesocosm. This study also evaluated important eDNA dynamics like the shedding rate and decay rate of sturgeon eDNA under ambient conditions. These metrics are vital to understanding how eDNA accumulates in the environment and determining if eDNA can be used quantitatively. The detection capabilities of eDNA provide a wealth of previously inaccessible information.

Using eDNA quantitatively combines the complexities of animal migration and individual behavior with hydrology, water chemistry, and the ecology of eDNA.

Quantification of anadromous species

Plough et al., 2018 developed a qPCR single-species detection assay for river herring (*Alosa spp.*) and used it for field testing in the Chesapeake Bay alongside ichthyoplankton sampling and visual surveys. eDNA data correlated with traditional abundance metrics and uncovered interesting species-specific spatial and temporal patterns. For example, eDNA and ichthyoplankton data showed that river herring were more prevalent on the eastern shore of the Chesapeake Bay compared to the western shore, and temporally, alewife (*Alosa pseudoharengus*) entered the system earlier than blueback herring (*Alosa aestivalis*). This assay was used in subsequent studies to evaluate patterns of habitat use among the two species (Ogburn et al., 2023) and to evaluate the effect of restoration of historic spawning habitat on river herring habitat use (Huang et al., 2023; Ogburn et al., 2023). The influence of environmental and physical factors on eDNA transport and decay have been evaluated to make eDNA data more quantitative. Tillotson et al. (2018) used a high-frequency sampling approach at different locations on a well-characterized riverine system to evaluate the ability of a single-species detection assay to quantify sockeye salmon (*Oncorhynchus nerka*) abundance and to test the utility of incorporating stream discharge, temperature, and spawning behavior into the relationship between eDNA and abundance as a means of addressing dynamics of eDNA persistence, decay, and transport. They found stream discharge had a dilution effect on eDNA detectability, and temperature had a minor effect on eDNA degradation. Levi et al. (2019)

conducted a similar study on both sockeye and coho salmon (*Oncorhynchus kisutch*) across different life stages at a weir in Alaska, where daily visual counts and daily eDNA samples were taken for out-migrating juveniles and returning spawners. This study developed a model that used a flow-corrected eDNA concentration to predict fish abundance to address the dilution effect of stream discharge on eDNA detectability. In summary, eDNA analysis of anadromous fishes can produce quantitative estimates of abundances, but further improvement is necessary. Refinement of eDNA tools could include incorporating environmental data or eDNA generation and fate (decay/shedding) information into eDNA concentration estimations. Salmonidae is one of the most studied families in quantitative eDNA research (Rourke et al., 2022), and the lessons learned can be applied to studies using these tools on other species with similar life histories.

The Mid-Atlantic Anadromous Alosine Fishes

The river herring: alewife and blueback herring

On the US east coast, Alewife and blueback herring, collectively referred to as river herring, are ecologically and commercially significant species that have sustained coastal communities for centuries (Hall et al., 2012). These important anadromous forage fishes provide a vital link between marine riparian and terrestrial ecosystems along the mid-Atlantic coastline (ASMFC, 2012). River herring spend most of their life at sea but move into coastal rivers and bays in the early spring as adults as they migrate to their natal freshwater habitats to reproduce (Hare et al., 2021). Larvae remain in freshwater and migrate toward the sea as they develop. As juveniles, they spend about a year in coastal brackish water before transitioning to the pelagic ocean as sub-adults, where they spend 1-4 years before reaching reproductive maturity. Reproductively mature

adults from the Mid-Atlantic stock will make the journey to spawn several times throughout their life, returning to the sea after spawning in freshwater (Hare et al., 2021). The predictability of this annual migration made river herring an easy-to-access food source, and as such, they were a crucial resource for coastal North American communities throughout history (Hall et al., 2012). This also made their populations highly susceptible to overexploitation.

Reports of local management efforts enacted to reduce the overexploitation of river herring populations date back to the beginning of the 18th century (Hare et al., 2021). Despite these efforts, river herring populations were depleted to historic lows in the late 20th century and have struggled to rebound due to the plethora of anthropogenic threats they continue to experience (Limburg & Waldman, 2009; Ogburn et al., 2017b). Management efforts have continued into the modern era. Most states on the eastern seaboard have prohibited commercial or recreational harvesting of river herring species (AMSFC, 2012). Current threats most impactful to river herring populations across their range are habitat fragmentation via dams and other barriers, declining water quality, and incidental catch (Hare et al., 2021; Hasselman et al., 2016). These ongoing threats, in combination with historic lows in abundance, have made the management of these species an increasing priority.

River herring populations are ideally managed by individual river systems. This is a challenging goal for management agencies with limited resources who must prioritize management actions for a wide variety of species under their jurisdiction (AMSFC, 2012). River herring populations require fishery-independent stock assessments due to the moratoria on river herring harvest.

Fishery-independent stock assessments can be costly, as they often require specialized equipment, personnel, and permitting. Additionally, studies on rare or low-abundance fish species can be difficult and invasive because traditional sampling approaches rely on the physical collection of specimens (Thomsen & Willerslev, 2015), and newer methodologies such as acoustic-based monitoring can be labor-intensive, straining resources available for stock assessments within an agency or management body. The variability in resource allocation to manage the river herring species contributes to gaps in data and inconsistent data quality in monitoring efforts from year to year. To advance conservation goals for the river herring species, environmental DNA (eDNA) monitoring has been investigated for its potential to provide accessible information on species distributions and abundances (see above: e.g., Plough et al., 2018; Ogburn et al., 2021). Other anadromous alosine fish are also present in these river systems, such as American shad (*Alosa sapidissima*) and hickory shad (*Alosa mediocris)*, and have experienced similar declines in abundance (AMSFC, 2020). Their occurrence in the same systems as river herring and similar management goals makes the development and application of eDNA monitoring potentially beneficial for the monitoring and management of these species.

American shad and hickory shad

American shad is North America's largest herring, with populations from St. John's River in Florida, USA, to the St. Lawrence River at the boundary between the USA and Canada (Zydlewski et al., 2021). Closely related to river herring, American shad show distinct differences in life history. Spawning run timing for the American shad varies latitudinally, and in the Mid-Atlantic, populations spawn from March to June. In the northern parts of their range,

they typically spawn later in the year, while populations further south spawn earlier in the year (Limburg et al., 2003). American shad show fidelity to their natal rivers during spawning like river herring, but their reproductive strategy also varies latitudinally, with northern populations exhibiting iteroparity and one southern population exhibiting semelparity (Zydlewski et al., 2021).

Hickory shad is intermediately sized between American shad and river herring and is the most data-limited compared to the other three species described. The range of hickory shad varies, with some publications citing populations as far north as the Bay of Fundy, Canada (Mansueti, 1962), Maine (Limburg & Waldman, 2009), and Connecticut (Murauskas & Rulifson, 2011), while others report northern spawning limits for Hickory shad in the northern regions of Maryland and Delaware (Rulifson et al., 2014; Hill et al., 2022). This species is anadromous, and iteroparity has been exhibited throughout its range (Rulifson et al., 2014). Due to its similarity to the other three species described and co-occurrence in coastal systems along the coast of North America, much of the information on hickory shad life history is based on river herring and American shad due to their larger abundances and historical importance as fisheries (Rulifson et al., 2014).

eDNA tools are well suited for data-poor, low-abundance fish stocks like the hickory shad, where catch or visual observation rates are too low to characterize depleted populations properly (Anderson et al., 2018; Pfleger et al., 2016). Generating more information on fish stocks using accessible and highly sensitive methods like eDNA can enable more rapid conservation and

restoration interventions for populations in peril (Lodge et al., 2012). Given the success of eDNA tool development and subsequent eDNA monitoring of local anadromous species like the Atlantic sturgeon and river herring, it is likely that eDNA analysis could be applied successfully to American and hickory shad as well. A panel of eDNA monitoring tools for alosines in the Mid-Atlantic would provide cost-effective and non-invasive methods to gather additional information on populations with limited data throughout these species' ranges.

Goals and Objectives

This thesis seeks to build on the findings of initial eDNA work conducted on river herring while providing new tools that can be used to implement eDNA sampling in other important anadromous forage fish species on the US East Coast. There were two overarching research foci: 1) high-frequency eDNA sampling for spawning run count estimation in river herring, and 2) the development and testing of tools for the American and hickory shads.

The goal of Chapter Two was to evaluate the capability of an eDNA detection assay for river herring to generate daily eDNA concentrations that could be associated with hourly sonar-based fish counts and weekly electrofishing species proportions collected at the same sampling location. This chapter improves the operationality of an existing eDNA molecular assay for river herring while providing a direct comparison to management-relevant metrics like daily fish count, species proportions, escapement, and spawning run count. The goal of Chapter 3 was to develop, optimize, and validate single-species eDNA detection assays for two species closely related to the river herring, the American shad and hickory shad. Standard statistical methods

were used to estimate critical performance parameters for the two new assays as well as the preexisting river herring assay. This work completed a panel of eDNA tools for the main Mid-Atlantic anadromous alosine fishes. These assays can provide a cost-effective method to study species distribution and habitat use and could be applied to estimate run counts similar to the river herring work from Chapter Two. Overall, the work in this thesis contributes new monitoring tools for shad species and advances the utility of eDNA methods for estimating fish abundances in the environment, which is critical to the potential future use of eDNA in management contexts.

Chapter 2: Evaluating the Capacity for High-Frequency Environmental DNA (eDNA) Sampling to Estimate Fish Count for River Herring Populations Utilizing the Choptank River

Abstract

Despite growing interest in eDNA monitoring as a low-cost, non-invasive tool for monitoring fish presence and relative abundance, quantitative assessments of fish biomass based on eDNA are still rare. Anadromous fishes, with their highly migratory spawning behavior that concentrates their populations in relatively small rivers over a short period of time, are ideal models for quantitative eDNA analysis. In this study, we assessed the eDNA abundance of river herring, an ecologically relevant anadromous duo, alongside sonar-generated river herring fish counts, at a well-characterized site on the Choptank River near Greensboro, Maryland, to examine the relationship between eDNA data and management-relevant abundance data. Over the 2021 spawning season, there was a strong correlation between eDNA concentration and the upstream fish count, with a Spearman's correlation coefficient of 0.86. A generalized additive model for location size and scale (gamlss) was used to estimate fish count as a function of flow-corrected eDNA concentration and relevant covariates. The model could generate predictions with a mean absolute error of 2949.56 and 0.88 R^2, and limited datasets with at least five weeks of data produced similar fit metrics. This work explores the potential utility of eDNA monitoring of river herring in the mid-Atlantic and provides a critical assessment of the quantitative capacity of eDNA for fish enumeration.

Introduction

Environmental DNA (eDNA) for detecting aquatic organisms

eDNA monitoring has gained traction in the past 15 years for its ability to detect organisms using the DNA they shed into their environment (Lodge et al., 2012; Takahara et al., 2012; Deiner et al., 2015; Goldberg et al., 2016). In aquatic environments, eDNA can be sampled by simply collecting and filtering water. Highly sensitive molecular tools can then be used to detect and quantify a species of interest by targeting a short diagnostic fragment of DNA, typically 80-200 bp in length (Tsuji et al., 2019). eDNA monitoring does not require interaction with the species of interest, making it an attractive option for monitoring rare or protected species (e.g., Pfleger et al., 2016). eDNA tools can also detect a target species' eDNA at highly dilute concentrations, making it useful in managing invasive species where quick management response is crucial to ecosystem recovery (Ficetola et al., 2008). eDNA monitoring can be highly cost-effective, reducing both the effort and cost of routine sampling compared to traditional methods (Evans et al., 2017). However, before eDNA monitoring can truly be integrated into standardized protocols for fisheries or wildlife management, more comparison studies with traditional metrics are needed, and some of the limitations of eDNA techniques must be addressed (e.g., Lacoursière-Roussel & Deiner, 2021).

Accurately quantifying eDNA requires understanding the underlying factors that influence the presence and persistence of eDNA over time and space (Yao et al., 2022). Several experiments have shown that fish eDNA persists in the water for varying lengths of time depending on both biotic and abiotic factors (Eichmiller et al., 2016; Harrison et al., 2019; Karlsson et al., 2022;

Sassoubre et al., 2016). For example, high stream discharge has been shown to have a dilution effect on eDNA in the environment, which can produce false negatives during high-flow events (Curtis et al., 2021). Environmental characteristics also affect the degradation of eDNA in the environment. The decay effect of increasing temperature has been shown in laboratory studies (Eichmiller et al., 2016; Jo et al., 2019; Andruszkiewicz Allan et al., 2020), but these are difficult to translate to the field where natural fluctuations in temperature may affect both eDNA production and degradation (Yates et al., 2021). The decay effect also complicates sample collection as eDNA could continue to decay after collection, potentially decoupling eDNA concentration from quantitative measures of abundance or causing false negatives due to the sample decaying below the limit of detection. There has been success in reducing eDNA decay by adding preservatives, freezing the sample, or filtering samples on-site (Takahara et al., 2015; Yamanaka et al., 2016, 2017). Understanding eDNA dynamics in the field or environment is vital to accurately interpreting eDNA- based data.

In order to establish the real-world application of eDNA monitoring as a fish enumeration tool, high-frequency sampling studies must be conducted at locations with well-characterized environmental data, ideally on species that increase in abundance predictably over a fixed period of time. The strong temporal pattern of presence and absence in riparian systems exhibited by anadromous species provides a straightforward scenario for assessing the capabilities of environmental DNA monitoring while simultaneously generating data to fill knowledge gaps on ecologically important species.

Advancing eDNA techniques using anadromous fishes

Anadromous species spend most of their life in the ocean but migrate into freshwater systems to reproduce. Spawning migrations occur on an annual basis and provide a scenario in which eDNA from a species of interest should be absent from freshwater systems before the spawning season, increase when the season begins as mature adults arrive and spawn, then decrease as adult fish occupancy declines and juveniles move downstream (Limburg et al., 2003; Hare et al., 2021). The absence of anadromous fishes in a system between spawning seasons provides a simple baseline from which increases and decreases in fish abundance, and presumably, eDNA can be examined and compared. Additionally, the conservation and management of anadromous species are of high priority as they provide valuable linkages between marine, riparian, and estuarine environments (Limburg & Waldman, 2009), and some species, like Pacific salmon, support active high-value fisheries (Levi et al., 2018). Therefore, sites typically utilized for traditional monitoring of these species are well-characterized and are ideal for integrating field-collected environmental variables into eDNA studies.

Studies applying eDNA surveillance to anadromous salmonid populations in the US Pacific Northwest have shown that eDNA concentration provides highly quantitative information on migrating fish populations when site-specific information is considered (Levi et al., 2019). Tillotson et al. (2018) used a site in Alaska where visual surveys (fish counts) for sockeye salmon (*Oncorhynchus nerka*) had been conducted for over 20 years. With this well-characterized site and a robust abundance metric (fish count), this study demonstrated how environmental variables like temperature and stream discharge altered the correlation between

eDNA concentration and visual fish count. A subsequent study on both sockeye salmon and coho salmon (*Oncorhynchus kisutch*) by Levi et al. (2019) found flow-corrected eDNA concentration was highly predictive of same-day daily adult counts for both species, emphasizing the need for accurate stream flow estimates and high-frequency eDNA sampling.

eDNA tools have also been developed to monitor anadromous species on the east coast of North America, though their application has been more focused on spatial distribution and habitat use rather than quantifying abundance (Harris et al., 2022; Huang et al., 2023; Ogburn et al., 2023; Plough et al., 2018, 2021). To equate the data collected by preexisting eDNA tools to management relevant metrics, high-frequency sampling must be conducted using both a traditional observation or catch method and an eDNA method at a well-characterized sampling location.

Study species: the mid-Atlantic river herring (*Alosa* spp.)

Plough et al. (2018) generated a dual-species qPCR eDNA detection assay for both alewife (*Alosa pseudoharengus*) and blueback herring (*Alosa aestivalis*) for monitoring and assessment of spatial distribution in the mid-Atlantic. Collectively referred to as river herring, these species occur from Newfoundland, Canada, to South Carolina, USA (alewife) and Nova Scotia, Canada, to Florida, USA (blueback herring; Limburg & Waldman, 2009). Similar to other anadromous fish, river herring are keystone species that transfer nutrients between marine, riparian, and estuarine environments through their highly migratory life history (Limburg & Waldman, 2009; West et al., 2010). River herring are born in rivers and streams and begin migrating toward the

sea within the first two months of life. Juvenile river herring stay in estuaries for up to 12 months before migrating to the sea, where they spend most of their adult lives (Hare et al., 2021). Reproductively mature adults return to their natal rivers to spawn once a year, and this strong pattern of natal homing gives rise to genetically distinct populations (Hasselman et al., 2016; Baetscher et al., 2017; Ogburn et al., 2017b).

Due to their highly migratory life history and genetically distinct sub-populations, the Atlantic Marine States Fisheries Commission manages these species on a river-by-river basis (AMSFC, 2012, 2017). The wide geographic distribution of spawning habitat and the difficulty of access to some sites makes monitoring these species a costly endeavor. Additionally, capture-based methods such as seining, gillnets, electrofishing, and trawl surveys can be laborious and expensive, limiting the number of streams that can be assessed in a single spawning season. Passive technologies are beginning to be utilized for river herring monitoring and have shown promise in generating metrics of abundance similar to capture-based methods (Magowan et al., 2012; Ogburn et al., 2017a), but these methods are reliant on expensive equipment and labor-intensive post-collection analysis. Any reduction in the cost or effort associated with monitoring can allow management activities to cover a wider area and gather more information.

eDNA monitoring has the potential to provide access to information at a broad spatial resolution with considerably less strain on resources compared to more traditional fish assessment techniques. We evaluated if high-frequency eDNA abundance data could estimate fish abundance (run count) on a single stream in the upper reaches of the Choptank River. We

evaluated the effectiveness of a preservative in reducing eDNA decay in sampling bottles and if a flow correction improved the correlation between eDNA and fish counts by addressing the dilution effect of stream discharge. Finally, generalized additive models were developed and tested using the entire dataset. The model was fit using reduced datasets to evaluate the amount of data needed to produce fish count estimates based on eDNA abundance.

Methods

Study site

A site on the Choptank River in Greensboro, Maryland (GPS 38.99491, -75.78702) was selected for both the eDNA and imaging sonar sampling (Fig. 2.1). The Choptank River is a coastal plain stream connected to the Chesapeake Bay on the Eastern Shore of Maryland and is a known spawning location for the river herring, where monitoring tools have been used previously (Ogburn et al., 2017; Plough et al., 2018). The sampling location was strategically located upstream of tidal influence and ~100 meters downstream of the U.S. Geological Survey (USGS) gauging station 01491000 (GPS 38.99719444, -75.7858056) to evaluate if stream discharge influenced eDNA abundance data (Curtis et al., 2021).

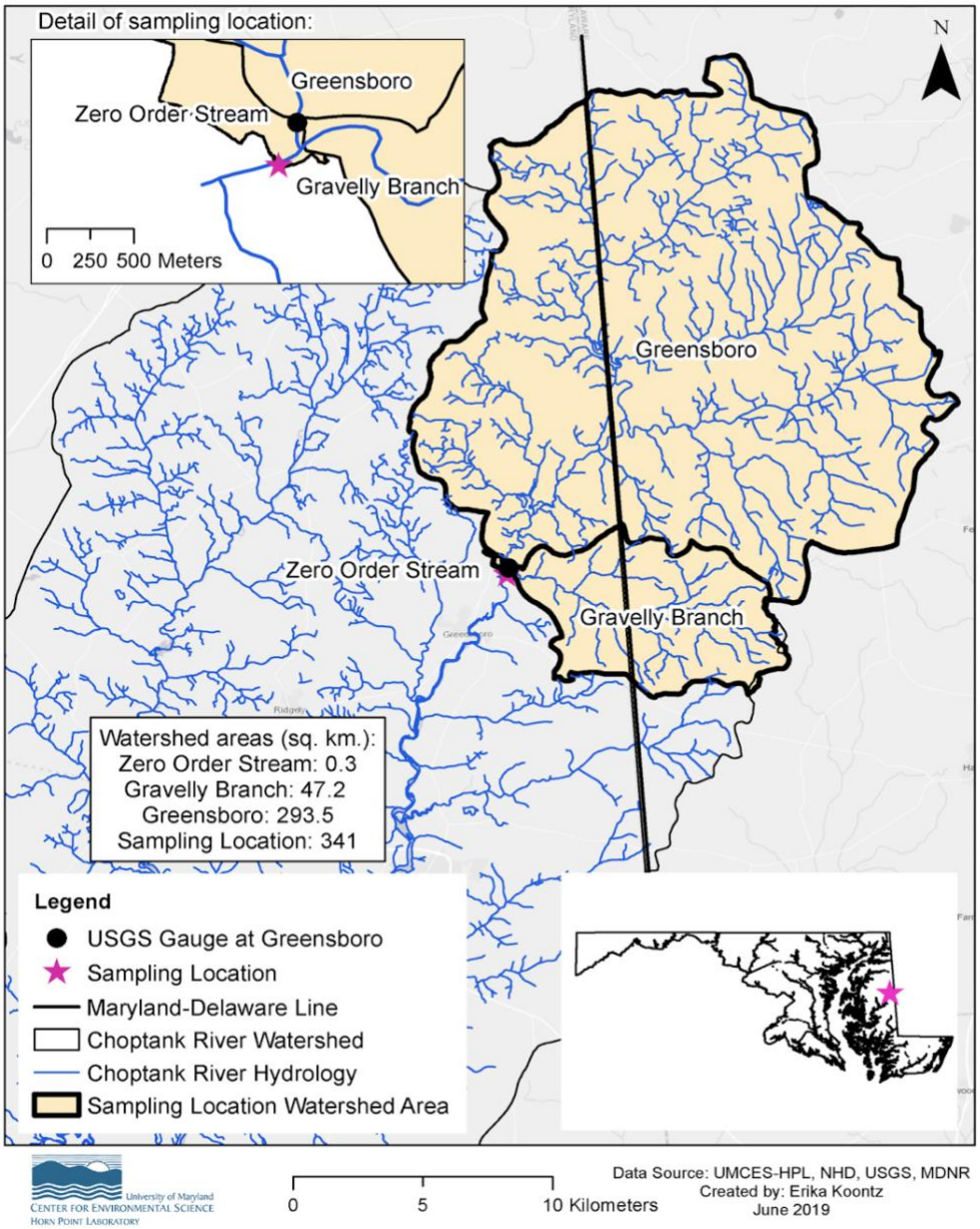

Figure 2.1. A map of the study site (pink star) on the Choptank River near Greensboro, Maryland. The Choptank River is a tributary of the Chesapeake Bay and is the spawning grounds of river herring populations and other anadromous species.

DIDSON sampling and fish enumeration

Imaging sonar monitoring can be used in shallow river systems and produces abundance estimates for species of interest based on their expected total length (Magowan et al., 2012). Dual frequency identification sonar (DIDSON) monitoring has been used on river herring in the Choptank River at the same site utilized for this study, where it proved to be an effective method for generating run counts for river herring species (Ogburn et al., 2017a). Counts of adult river herring were estimated during their spawning migration from March to May 2021 at the study site using a DIDSON (Fig. 2.2A, DIDSON V5.25.52, Sound Metrics Corp, Bellevue, WA). Sonar videos were collected at a resolution of 1.8 mHz within a 10 m field of view. Construction fencing was installed in the river to funnel fish into the DIDSON's field of view. The DIDSON device was placed in the river at the sampling site and, as such, was sensitive to disruptions from high-flow events. Some of these disruptions required the removal of the device from the sampling site. During the sampling period, a high-flow event required the removal of the DIDSON device from March 20-29. The device was placed back at the site on the 29th, and recordings resumed. Fish counts were not generated for nine days, March 20-29, 2021.

Sonar recordings were generated in 10-minute segments per hour with a randomized within-hour start time. This 10 min/hr sampling effort has been utilized for anadromous species in the past and results in accurate estimates of fish count when compared to visual counts (Xie & Martens, 2014). After collection, these videos were processed following established protocols (Ogburn et al., 2017a). Fish counts were generated by manually counting fish moving upstream in sonar recordings within the size range for adult river herring (200–350 mm TL) using DIDSON

V5.25.52 software (Fig. 2.2B). Fish were measured for the target size range digitally using a straight line. Counts were generated for a random 10-minute sample each hour and multiplied by a factor of six to generate fish count estimates per hour.

River herring species (alewife and blueback) are similar in size and cannot be differentiated using sonar recordings. To generate species composition data for DIDSON counts, weekly boat-based electrofishing was used to collect samples and estimate the proportions of the two species. Fish caught within the size range of adult river herring were visually identified to species (Ogburn et al., 2017a). Proportions of the river herring species, alewife, and blueback herring, were calculated from the electrofishing samples, and daily species proportions were generated using linear interpolation (Ogburn et al., 2017). Daily species proportions were applied to the hourly counts to generate species-specific counts per hour. Spearman's rank correlation coefficients were generated for eDNA and general count proportions, eDNA and electrofishing proportions, and eDNA and interpolated proportions. The correlations between the reduced datasets (eDNA and electrofishing proportions and eDNA and interpolated proportions) were analyzed using the boot package (ver. 1.3-28.1) to resample the data with replacement 1000 times and generate correlation coefficients for each repetition.

A. B.

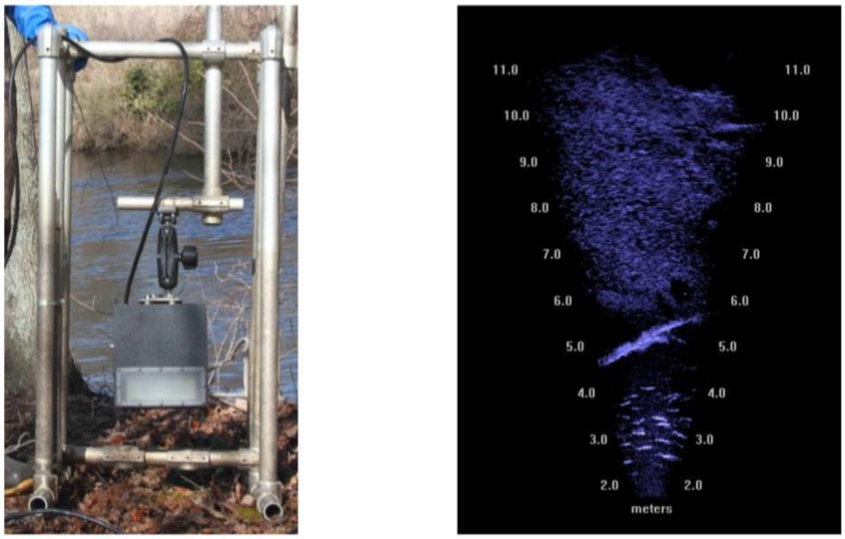

Figure 2.2: The Dual Frequency Identification Sonar Unit (DIDSON) for generation of fish count. Panel A is an image of the DIDSON device before its deployment in a shallow river system (a picture courtesy of SERC). Panel B is a screen capture of a DIDSON video where a river herring can be seen passing by (image credit: Sinclair, 2013).

Collection of environmental covariates

Environmental covariates were collected during the spawning migration by the United States Geological Survey (USGS) and the DIDSON device. The DIDSON device had a HOBO U20-001–2-Ti data logger (Onset Computer Corporation, Bourne, Massachusetts) in PVC pipe anchored underwater next to the DIDSON unit to collect hourly average water temperature (℃). Hourly water temperature estimates were averaged over the DIDSON fish count window to generate an average daily water temperature (℃). Stream discharge (m^3/s) rates were collected from the USGS gauge ~100 m upstream of the collection site (Choptank River, no. 01491000).

Milling activity was measured as a binary factor (present or absent) based on observer notes during sonar fish quantification.

Water sampling for eDNA

Daily eDNA samples were collected remotely using an ISCO model 6712 full-size portable sampler (Teledyne ISCO, Fig. 2.3A) from March to May 2021 to coincide with the known spawning migration for river herring. Weekly, 27 collection bottles were cleaned using a 10% bleach solution and then preloaded with 9 mL of 1% Benzalkonium Chloride (BAC) solution (final working concentration 0.01% BAC) to preserve eDNA between sample collection and sample filtration (Yamanaka et al., 2017). These bottles were deployed in the ISCO, and the instrument was programmed to draw three water samples of 891 mL (final volume 900 mL) per sampling event (Fig. 2.3B). Samples were taken daily at 6 pm to coincide with river herring migration activity (Ogburn et al., 2017a; Magowan et al., 2012).

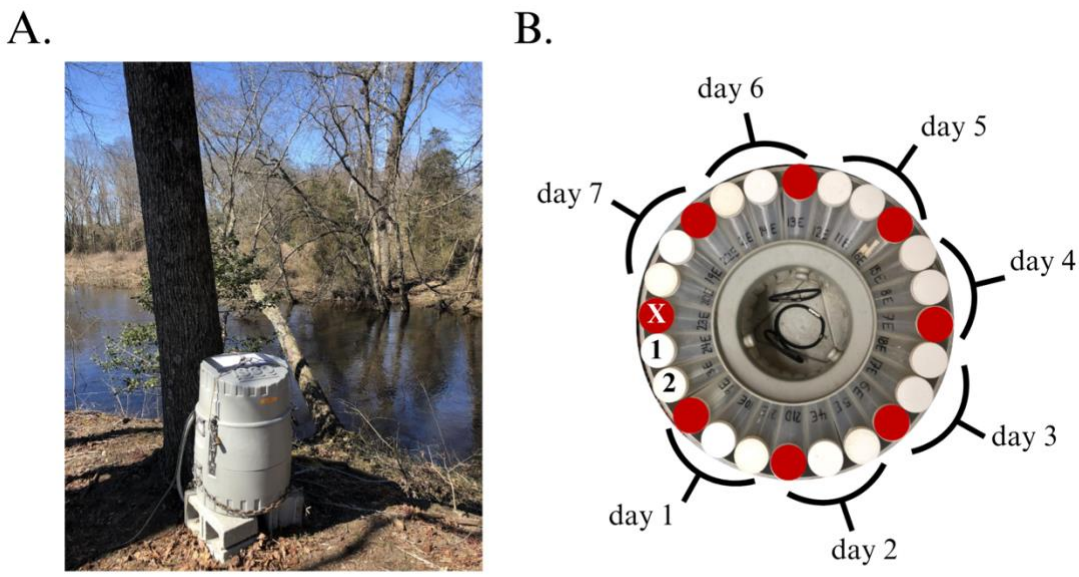

Figure 2.3 Automatic water sampling device for eDNA water sample collection. Panel A shows the ISCO device at the study site, situated on the edge of the Choptank River near Greensboro, MD. Panel B shows the interior of the device with 24 collection bottles. Bottles are grouped by collection day, and caps are colored to denote rinse bottles between sampling events (red caps) and duplicate samples used in downstream analysis (white caps). Samples were taken daily and collected from the device for filtration once a week.

At the beginning of each sampling event and between each draw within a sampling event, the device was programmed to rinse the sampling tubing three times to prevent contamination between water samples. As an additional contamination precaution, the first collection bottle was discarded. The following two bottles in every iteration of the sampling schema were collected once a week for filtration and downstream laboratory analysis (Fig. 2.3B). The full collection bottles were replaced with clean, sterilized bottles, and the program was reset each week. Days to filtration was calculated as the number of days between eDNA sample collection and filtration, ranging from 1-6 days, and was recorded for each sample.

Molecular analysis of eDNA: filtration, extraction, and qPCR

Samples were processed in a pre-PCR space at the Horn Point Laboratory in Cambridge, Maryland. Filtration and extraction methods were adapted from Plough et al. (2018). Briefly, duplicate sample bottles were each filtered (500-900 ml) with a vacuum pump and custom PVC manifold on 47 mm diameter 1.0-micron cellulose nitrate filter membranes (Cytiva-Whatman), which are optimal for aquatic eDNA detection of fishes (Turner et al., 2014). Filtering funnels were cleaned with a 10% bleach solution and rinsed with deionized water between samples. The date of collection, date of filtration, milliliters filtered, and total milliliters were recorded for each sample. Immediately after filtration, filter papers were placed in individual 15 mL centrifuge tubes and kept at -80°C until extraction.

eDNA was extracted using the CTAB method adapted from Renshaw et al. (2015) and previously used by Plough et al. (2018) and (2021). This method has a streamlined protocol, low-cost implementation, and high DNA recovery from environmental samples (Renshaw et al., 2015; Turner et al., 2014). Half the filter was used for eDNA extraction and quantification, while the other half was preserved. Extracted samples were stored in Eppendorf tubes and kept at -20°C until qPCR.

To quantify river herring eDNA in the field samples after extraction, quantitative real-time polymerase chain reactions (qPCR) were performed on a BioRad CFX 96 qPCR instrument

(Bio-Rad Laboratories Inc., Hercules, California, USA) using a molecular beacon assay targeting a 144 bp region of the CO1 gene in river herring developed by Plough et al. (2018). PCR reactions were run in triplicate in total volumes of 20 µL, consisting of 10 µL Bio-Rad 2x SSO advanced universal probe super mix (Bio-Rad Laboratories Inc., Hercules, California, USA), 0.7 µL of each primer or probe, 3.9 µL of PCR water, and 4 µL of extracted eDNA. A positive control spiked with a known copy number was run with each sample, and six negative controls (water blanks) were run with each 96-well plate of samples. PCR conditions for each quantification run were: 95°C for 45 seconds, 44 cycles of 95°C for 5 seconds, 59°C for 15 seconds, and 72°C for 10 seconds, with a fluorescence reading measured after each of the 44 cycles. A standard curve was run for each qPCR run as described by Plough et al. (2018). In brief, known quantities of synthesized DNA were included in duplicate to generate copy number estimates from Cq (cycle number) values for field samples with unknown quantities of DNA. Standards were created from a synthesized 164 bp fragment of Alewife DNA (Geneblock, Integrated DNA Technologies Inc., Coralville, IA) serially diluted from 300,000 copies/µL to 30 copies/µL. qPCR plates with standard curve regressions that had an R^2 of less than 0.98 were rerun.

Samples with positive detections (Cq <40) were sent out to MCLAB (San Francisco, USA) for Sanger sequencing to develop species proportions. Sequence files were downloaded, and FinchTV (ver 1.5.0, Geospiza Inc., 2004) was used to analyze chromatograms. A single diagnostic base was used to differentiate alewife and blueback herring signals (Plough et al., 2018). At this base site, relative peak heights were measured, and the proportional peak heights

were used as the species proportions. Sequences with high levels of interference were not used to generate species proportions. Those samples were reamplified, and Sanger sequencing was attempted again.

Field-based eDNA decay experiments

To estimate decay rates for river herring eDNA in bottles in the presence of 0.01% BAC under natural environmental conditions, five decay experiments were conducted over the course of the 2021 river herring spawning season. For each of the five experiments, one large container was cleaned using the same methods used on the individual sampling bottles. The bucket was preloaded with 35 mL of concentrated BAC (final concentration 0.01%). At the Choptank site, four draws were taken to achieve a final volume of 3,500 mL. This sample was transported back to Horn Point Laboratory, where it was divided into fourteen clean 250 mL Nalgene bottles. Two of these bottles were immediately filtered, while the other twelve were placed in a styrofoam box and left outside in the shade to mimic the conditions at the Choptank site. Temperature, both inside the box and inside the bottles, was continuously monitored throughout each experiment using HOBO UA-002-64 Pendant Temperature/Light 64K Data Loggers (Onset Computer Corporation, Bourne, Massachusetts). Two bottles were collected from the box each day over the course of one week. All laboratory procedures (filtering, extraction, qPCR) were identical to those conducted on the field samples (see above).

eDNA decay experiments in the presence of BAC preservative were run to evaluate if significant decay was occurring in the field between the time of sample collection and sample filtration. A

log-linear regression was used to model change in log-transformed eDNA concentration as a function of time in days for five independent experiments individually. These models were fitted and evaluated for model significance. A Bonferroni correction was used to account for multiple tests, and the alpha level of significance was set to 0.01.

Modeling DIDSON fish count as a function of eDNA concentration

All analyses were performed using R (ver 4.2.2, R Core Team, 2022). Daily eDNA concentrations were first evaluated for their correlation with a daily fish count response variable. Since the goal of this analysis was to generate fish count predictions, eDNA was used as the predictor, and fish count was the response variable. Using eDNA and fish count as the predictor and response, respectively, has backing in the literature (Takahara et al., 2012; Wilcox et al., 2016; Levi et al., 2019; Searcy et al., 2022), and while this makes the analysis more straightforward, it should be noted that this model does not reflect the actual biological mechanism through which eDNA is deposited in the environment by the fish. Hourly fish counts were summed across a variety of 24-hour windows based on the time of eDNA sampling to determine which summation should be used as the response variable in a predictive model. Daily eDNA concentrations were also flow-corrected, meaning daily eDNA concentration (copies/mL) was multiplied by stream discharge (m^3/s) to account for the dilution effect of stream flow (Tillotson et al., 2018; Levi et al., 2019). The correlation between fish counts, eDNA concentration, and flow-corrected eDNA concentration was evaluated using Spearman's Rank Correlation Coefficient. The predictor-response pair with ρ closest to one was used to inform a model that used eDNA to estimate fish count. The final model was built by 1) estimating the

distribution of the response variable, 2) optimizing the model structure for the main effect of eDNA on upstream fish count, and 3) adding covariates to maximize model fit and prediction accuracy.

Based on the fitDist function in the gamlss package (ver 5.4-12, (D. M. Stasinopoulos & Rigby, 2007) the distribution of the response variable was well characterized by the double Poisson, geometric, and negative binomial distributions and the zero adjusted and inflated versions of these distributions. The negative binomial distribution is a gamma mixture of a Poisson distribution and is often used with over-dispersed count data (Joe & Zhu, 2005). The negative binomial distribution models variance as a quadratic function of the mean and contains a separate parameter (sigma) for controlling the slope of this variance-mean relationship and, as a result, the severity of overdispersion (Stoklosa et al., 2022).

To evaluate the need for a complex and flexible model structure, like that provided by gamlss, a simple linear regression was used as a baseline model. The next level of complexity was a generalized linear model (stats ver 4.3.1, R Core Team, 2023), fitting a linear regression with a Poisson distribution, which is considered the simplest distribution used for modeling count data (Zeileis et al., 2008). The final (highest) level of complexity employed was the gamlss model, which provides a high degree of flexibility and includes a separate term for modeling variance (M. D. Stasinopoulos et al., 2017). Since eDNA concentration was selected as the independent variable, the model structures tested and used here violate the regression assumption of no error in the independent variable. This violation of regression assumptions is common in field science,

but noise in the independent variable can cause attenuation bias where model coefficients are biased toward zero (Spear et al., 2021).

During optimization, models were evaluated by their adherence to model assumptions and fit metrics. Model assumptions were tested through visual inspection of diagnostic residual plots: Residuals v Fitted plot, Normal Q-Q plot, Scale-Location Plot, and Residuals v Leverage Plot. Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to evaluate the fit of the model and its predictive accuracy. AIC tends to favor more complex models and is useful for variable selection and maximizing prediction accuracy, while BIC is useful for evaluating the functional form of the model and finding the best model for describing a general pattern (Aho et al., 2014). The coefficient of determination (R^2) was included as a measure of the model's predictive power. Other measures of predictive ability were also reported, specifically the root mean squared error (RMSE), which is the prediction error of the model, and the mean absolute error (MAE, Equation 1), which is an easy-to-interpret metric of error on the same scale as the predictions. These metrics were used throughout the analysis of model fit and covariate selection.

$$MAE = \frac{\sum_{n=1}^{w} |obs\_count_n - pred\_count_n|}{w}$$

Equation 1: Mean absolute error (MAE) for evaluating model fit. Obs_count is the observed count and pred_count is the prediction from the model. N is the observation in the dataset, and w is the total number of observations.

To arrive at the final model, the inclusion of the covariates was evaluated through forward and backward stepwise variable selection for all model terms using the function stepGAICAll.A() within gamlss (Stasinopoulos et al., 2017). stepGAICAll.A() uses the Akaike Information Criterion, a metric of prediction error, to test the effect of covariates on each parameter of a gamlss model as specified by the distribution family. From a null model, beginning with the mu parameter, single-term addition determines the covariate that causes the biggest decrease in AIC. That term is retained, and the remaining covariates are added individually again. If adding any of these terms decreases the AIC further, the term is retained, and the process repeats itself. This forward stepwise selection is carried out for the mu parameter until adding remaining covariates does not improve the model fit. The process is then carried out for the sigma parameter, given the full mu model. Once forward stepwise selection is complete for the sigma parameter, the backward stepwise selection is carried out for the mu parameter and then the sigma parameter. Models that verified model assumptions with low AIC values were used to generate predictions on the full dataset and species-specific datasets. Species-specific datasets were generated by multiplying fish count and eDNA concentrations by eDNA species proportions before being used to build the model and generate predictions.

Predictive capacity of the full model

To evaluate the predictive ability of the full model to produce fish count estimates using only model predictors, the dataset was divided into test and training sets iteratively. In Schema 1, observations were grouped by week and added to the ‘training set’ consecutively to evaluate how

much data was needed to make robust predictions later in the season, which mimics real-world sampling considerations. Although high-frequency sampling for both eDNA and DIDSON fish count were conducted for this study, high-frequency DIDSON analysis may not be feasible for regular or annual monitoring efforts. DIDSON device site preparation takes time to create weirs to funnel fish to the DIDSON's field of view and position the device correctly (Magowan et al., 2012). This means moving the DIDSON from site-to-site would be unrealistic over the course of a few days but could potentially be beneficial over the course of a few weeks if eDNA could be used to generate fish count estimates for the remainder of the season. The first model built in Schema 1 was trained on the first week of data collected in the season, March 9-14, the fitted model made predictions on the data that was withheld from the training set, the 'test set', data from March 15-May 25. Then fit metrics mean absolute error (MAE), R-squared, and Pearson's correlation coefficient were recorded, and the process was repeated. Each iteration of the process took one week of data from the test set, and moved it to the training set, until the training set included all observations (the full model).

In prediction Schema 2, observations were grouped by week and added to the training set at random. The first model built in Schema 2 was trained on one week of data, the remaining 10 weeks of data were used as the test set. This was repeated 10 times before moving to two weeks of data. Not all of the randomly selected training sets could reach convergence, and models that did not converge were omitted from the analysis. In total, for 2, 4, 5, 6, and 10 weeks of data, nine repetitions converged; for 9 weeks of data, eight repetitions converged; and for 3, 7, and 8 weeks of data, seven repetitions converged. The fit metrics from the repetitions that converged,

MAE, R-squared, and Pearson's correlation coefficient, were recorded and averaged. This process was repeated for 2-10 weeks of data, ten times each. Finally, the week group was dropped, and the model was tested on completely randomly selected data.

In Schema 3, a learning curve was used to evaluate how much data was needed to make robust predictions using the model. A learning curve generates model performance metrics for different sample sizes of the training set (e.g., Emmert-Streib & Dehmer, 2019). Observations were randomly selected ten times per sample size, ranging from 1-69. Predictions were generated using both the training and the test set, and the mean absolute error was calculated.

Results

Basic comparisons of count and eDNA

Daily eDNA concentrations and sonar fish counts varied significantly over the spawning period, with major pulses in run counts or eDNA and periods of low fish passage (low eDNA) at certain points in the season (Fig. 2.4A). Overall, raw, uncorrected eDNA concentration was highly and positively correlated with sonar-based run counts (Fig. 2.4A) with a Spearman's correlation coefficient $\rho = 0.841$ ($p<0.05$). When eDNA was corrected using stream discharge, this correlation increased slightly, $\rho = 0.860$ ($p<0.05$). Both eDNA concentration and sonar-based fish count showed species-specific patterns of presence over time (Fig. 2.4B). Alewife entered the system earlier in the season, with the first alewife detection and count on March 9th, while blueback herring entered the system later, with the first blueback herring detection occurring on

March 29th, and the first blueback herring count occurring on April 5th (Fig. 2.4B). Alewife and blueback eDNA were last detected in the system May 19th and 20th, respectively. The last fish counts for alewife and blueback were April 28th and May 25th, respectively.

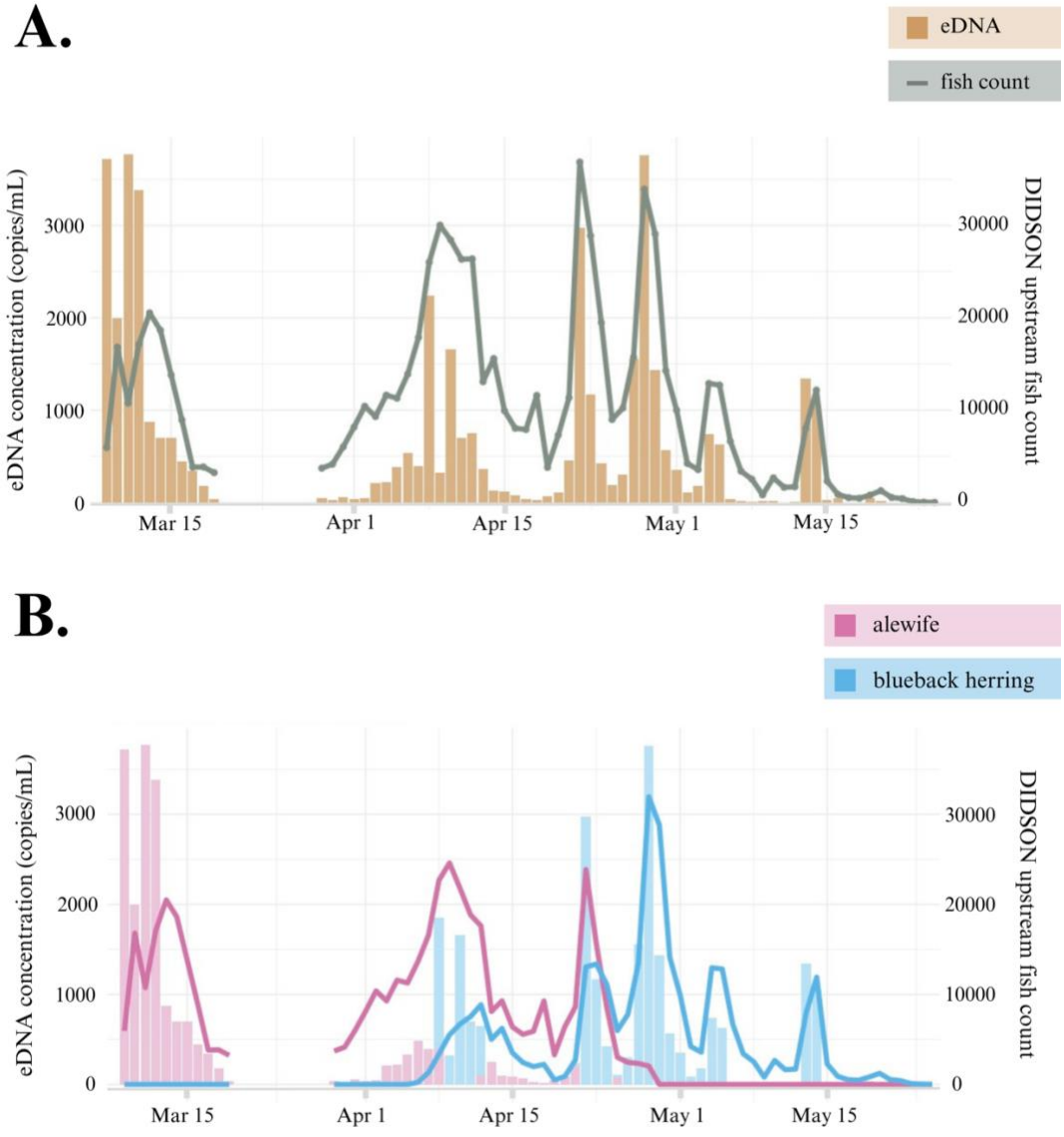

Figure 2.4: eDNA concentration and sonar fish counts for the 2021 river herring spawning run on the Choptank river near Greensboro, Maryland. Panel A shows raw, untransformed, daily eDNA concentrations in bars enumerated on the left axis. Sonar fish count, in lines enumerated on the right axis, is a summation of hourly fish counts 12 hours before and after eDNA sampling. Panel

B shows species-specific breakdowns for each method. eDNA concentration species proportions are based on Sanger Sequencing of daily eDNA samples. Sonar fish count species proportions are based on weekly electrofishing data. Linear interpolation was used to generate species proportion estimates between electrofishing events.

Species-specific analysis of eDNA and fish count datasets

Species proportions were generated for both eDNA concentrations and fish counts at different intervals. Sanger sequencing was used to generate species proportions for daily eDNA samples, weekly electrofishing was used to generate species proportions for fish counts, and proportions were linearly interpolated for days between weekly electrofishing surveys. Species proportions on a given day were highly correlated among methodologies with a Spearman's rank correlation coefficient of 0.811 ($n=69$; $p<0.05$). Since daily species proportions for fish counts were derived (interpolated) from weekly electrofishing data, an analysis of the absolute difference between eDNA and fish count data was conducted. This analysis showed that eDNA species proportions were more similar to electrofishing proportions (fish count species proportions collected on the days electrofishing was conducted) than interpolated proportions (fish count species proportions generated via linear interpolation between electrofishing events Fig. 2.5A). Since the sample sizes for these two categories were highly dissimilar (electrofishing n=10, interpolation n=59), a bootstrap analysis of correlation was conducted. This bootstrap analysis showed daily proportions generated by interpolation were significantly less correlated with eDNA species proportions ($\rho = 0.767$, $p<0.05$) than those proportions generated by electrofishing ($\rho = 0.883$, $p<0.05$; Fig. 2.5B).

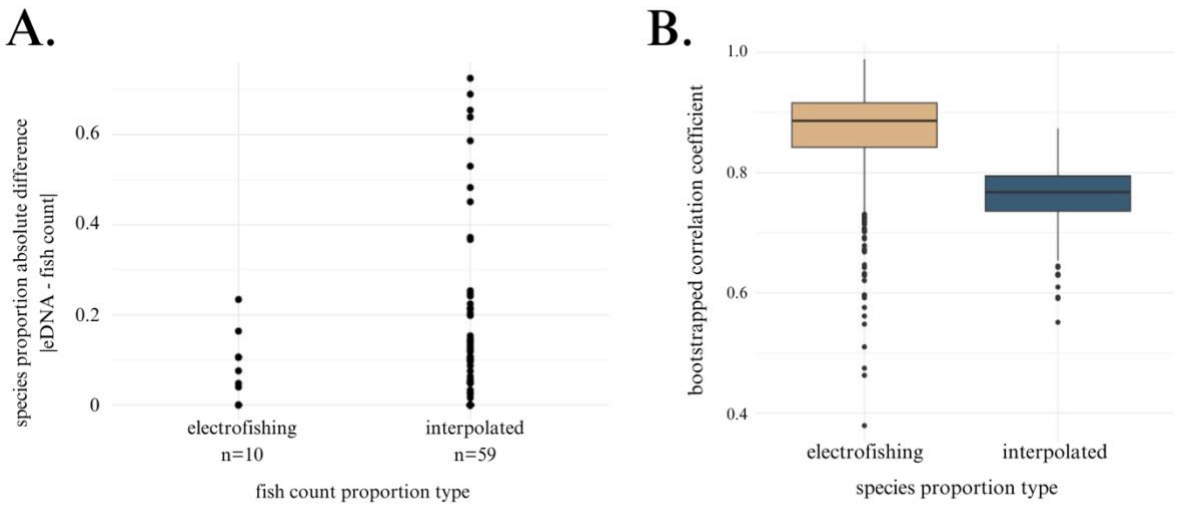

Figure 2.5: Analysis of eDNA and fish count species proportions using both actual and resampled data. Panel A shows the absolute difference between eDNA and fish count proportions. Species proportions for fish count data were grouped as "electrofishing" (n=10, days that both electrofishing was conducted and eDNA was sampled) or "interpolation" (n=59, days between electrofishing events when eDNA was sampled but fish count proportions were interpolated). Panel B shows the bootstrapped Spearman's rank correlation coefficient between eDNA and fish count species-proportion methods for simulated data. Species proportions for fish count data were grouped by electrofishing and interpolation and bootstrapped 1000 times. The upper limit of whisker for both categories within panel B is the maximum value, and the lower limit, below the interquartile range, is the minimum value. These maximum and minimums excluded outliers which are shown by the black points. The boxes are the interquartile ranges.

Field-based eDNA decay experiments

Five independent eDNA decay experiments in the presence of BAC were conducted over the span of the 2021 river herring spawning season and showed highly variable patterns of decay (Fig. 2.6). While eDNA appeared to decay slightly in some experiments, log-linear regression analysis indicated that decay was insignificant for four out of five of the experiments (Table 1). This result suggests that eDNA decay between the time of sampling and the time of filtration is

minimal in the presence of BAC preservative overall, and thus decay was not included in downstream analyses.

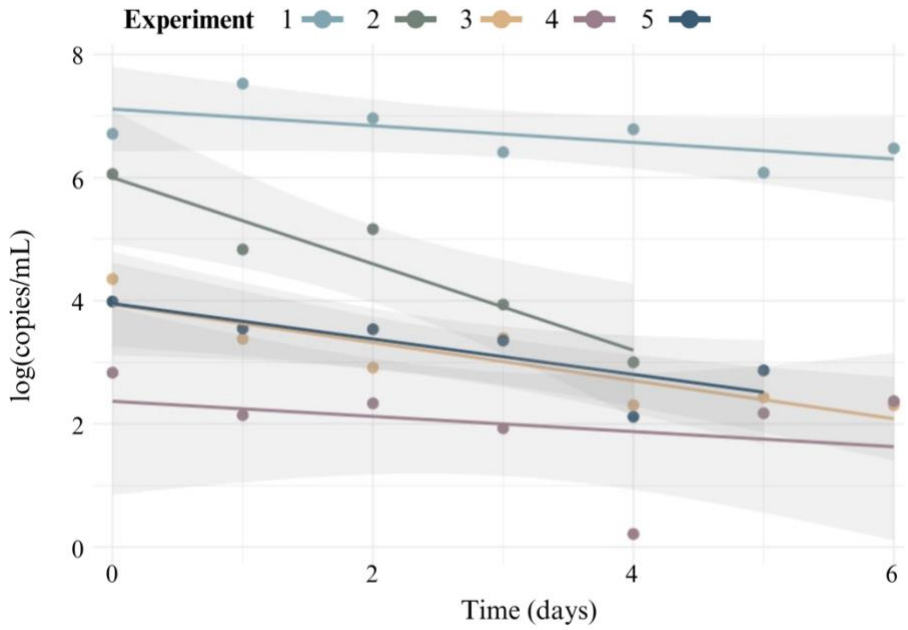

Figure 2.6: eDNA concentrations over the duration of five decay experiments in the presence of BAC preservative under field environmental conditions. The points represent the observed eDNA concentrations on a given day. The lines represent the log-linear regression for each experiment. The shaded region is the 95% confidence interval for the trendline.

	Experiment	**slope**	**intercept**	**multi r-squared**	**p-value**
llm_exp1	exp1	-0.135	7.113	0.400	0.128
llm_exp2	exp2	-0.309	3.940	0.781	**0.008****
llm_exp3	exp3	-0.123	2.368	0.101	0.487

llm_exp4	exp4	-0.288	3.956	0.673	0.046
llm_exp5	exp5	-0.701	6.001	0.894	0.015

Table 1: eDNA decay in the presence of BAC preservative under field conditions across five experiments. A log-linear regression (llm) predicted log(copies/mL) as a function of time in days for five independent week-long experiments. Alpha significance threshold was set to 0.01 to account for multiple testing, bolded p-values indicate experiments that are <0.01.

Generalized additive models: fish count as a function of eDNA concentration

The response variable, upstream fish count, was left-skewed and over-dispersed, with an average daily count of 12410.6 and a variance of 1.074×10^8. Ordinary least squares (OLS) residuals showed increasing variance with increasing river herring abundance indicating the need for a flexible mean-variance relationship. Based on the diagnostic plots, and residuals of the initial model containing a single predictor, eDNA, the negative binomial distribution was most appropriate for modeling this dataset (Table 2).

Predictors	**Response dist.**	**AIC**	**BIC**	**Rsq**	**RMSE**	**Diagnostics**	**MAE**
log(flow-corrected eDNA)	Normal	1411.41	1418.11	0.51	6407.32	poor	5035.85
log(flow-corrected eDNA)	Poisson	173453.24	173457.71	NA	6191.82	poor	3984.21
log(**μ**): log(flow-corrected eDNA) log(**σ**): log(flow-corrected eDNA)	Negative Binomial	1322.86	1331.79	0.54	6381.01	fair	1351.74
log(**μ**): log(flow corrected eDNA) +	Negative Binomial	1290.30	1312.64	0.88	4191.90	excellent	2949.56

milling + avg_watertemp + JD + JD:avg water temp + avg water temp:log(flow corrected eDNA) log(**σ**): avg water temp + milling							

Table 2: Model fit metrics and diagnostics for predicting upstream fish count from eDNA and covariates under various response distributions. Fit metrics generated include: Akaike information criterion (AIC), Bayesian information criterion (SBC), Coefficient of determination (R^2), root mean squared error (RMSE), model assumptions diagnostics rating, and mean absolute error (MAE). Model diagnostics is a qualitative rank based on the four residual diagnostic plots (Residuals v Fitted plot, Normal Q-Q plot, Scale-Location Plot, and Residuals v Leverage Plot) and their adherence to model assumptions.

The model structure was first optimized for the main effect of eDNA concentration on the upstream count. A generalized additive model was selected to describe upstream fish count as a function of log-transformed flow-corrected eDNA concentration. The log transformation allowed the predictor to better fit the data given the multiplicative log-link used in the negative binomial model. Through forward stepwise variable selection beginning with the mu term, environmental covariates/predictors known to impact eDNA or herring migration were assessed for their contributions to the model. Log-transformed flow-corrected eDNA concentration was the most important predictor, decreasing AIC by 92.5 points when compared to a null model. Average water temperature was added to the model, further decreasing AIC by 9.6 points. The conditional effect of water temperature increased the number of fish predicted by a given eDNA concentration at high temperatures (Fig. 2.7C). Adding an interaction between eDNA and average water temperature decreased the model AIC by 11.3 points. The inclusion of a milling factor decreased the model AIC by 5.4 points and predicted a higher fish count on days with

milling versus days without (Fig. 2.7D). Finally, day in season was added to address the residual variation attributed to the temporal dependence of the datatset (Fig. 2.7B). Forward stepwise selection was also performed on the sigma term. The addition of the milling factor decreased the AIC by 2.7 points. The addition of log-transformed flow-corrected eDNA decreased AIC by 0.9.

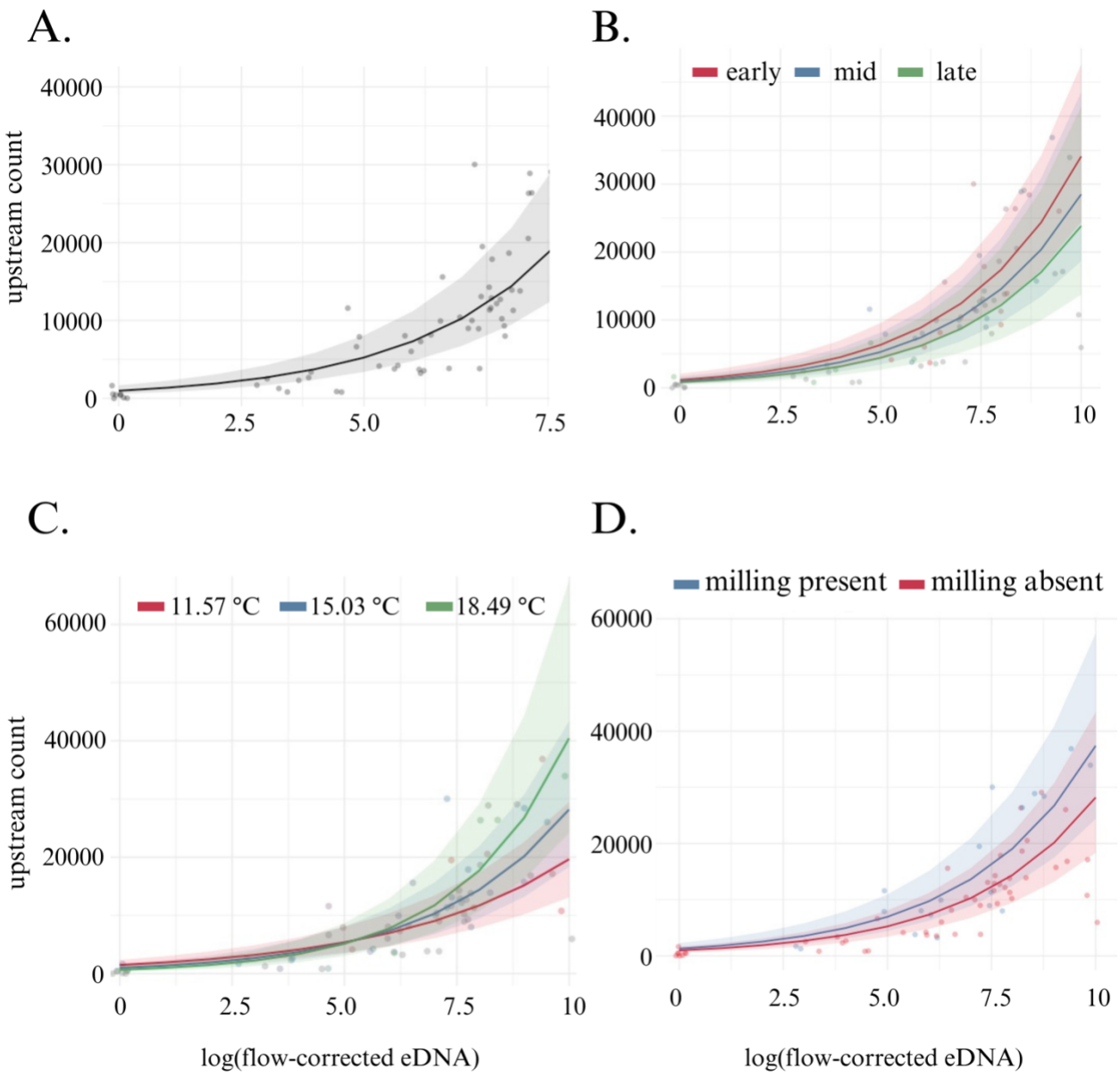

Figure 2.7: Full Model Predicted Fish Count for the Main Effect and Conditional Effects Across Covariates. Panel A shows full model predictions. Panel B shows the effect of the day of season on model predictions. Panel C shows the predicted counts broken down by average water temperatures and panel D shows the effect of the binary factor for milling presence/absence. 95% confidence intervals are included in all four panels.

The final model was exponential in shape and included log-transformed flow-corrected eDNA, milling, average water temperature (°C), an interaction term between average water temperature (°C) and log(flow-corrected eDNA), Julian Day, and an interaction between Julian Day and average water temperature in the mu term, and average water temperature and milling in the sigma term (Equation 2). This model explained 88% of the deviance (R-squared) and had a mean absolute error of 2949.56 (Table 2). The eDNA model produced predictions for the season, predicting a run count of 729,385 fish. The model underpredicted the fish count compared to sonar generated run count of 733,208 by a difference of 0.5%. The species-specific predictions generated slightly lower R^2 values (alewife 0.67; blueback 0.76) and lower MAE values (alewife 1841.68; blueback 2350.57) than the full dataset (Table 2; Fig. 2.8B).

$$upstream\ count = NBII(\mu, \sigma)$$

$$\log(\mu) = 5.54 - 0.0016 \log(FCeDNA) + 0.28\, m + 0.028\, JD + 0.15\, tmp + 0.022 \log(FCeDNA) : tmp - 0.0024\, JD : tmp$$

$$\log(\sigma) = 8.17 - 0.079\, tmp + 0.56\, m$$

Equation 2: Equation for the fitted gamlss model using all observations. Variables are as follows, *FCeDNA* is flow-corrected eDNA, *m* is a binary milling factor, *tmp* is average water temperature, *JD* is Julian day or day in season. The colon between two variables indicates an interaction term.

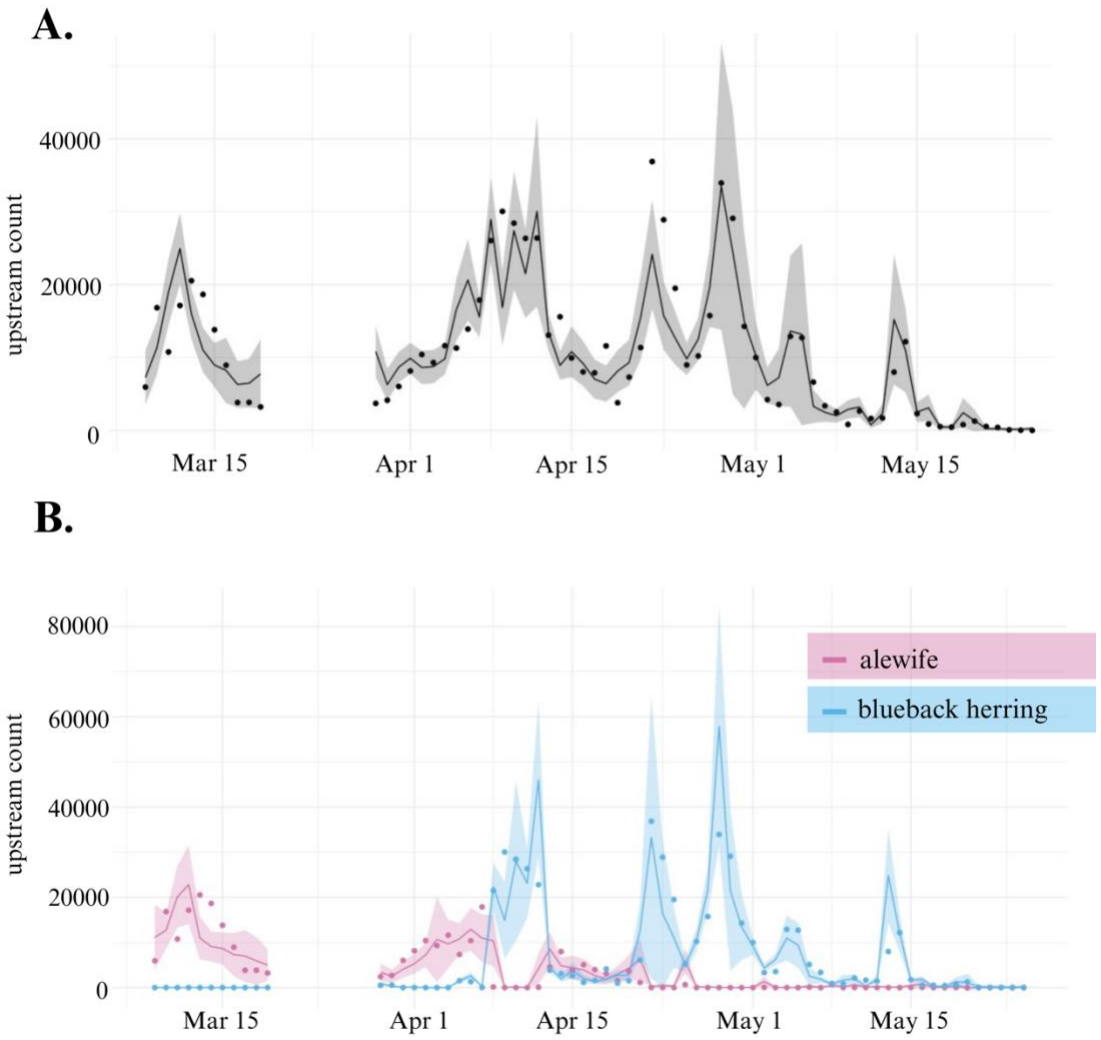

Figure 2.8 Predictions generated by the full GAMLSS Negative Binomial II model plotted with observed upstream fish counts. Panel A shows model predictions (line) and a 95% confidence interval. Observed upstream fish counts are plotted in points. Panel B takes model predictions and 95% confidence intervals and breaks them apart by species using eDNA species proportions. Observed species-specific fish counts are plotted in points and assigned to species based on eDNA species proportions. The gap in the data during March is due to the removal of one or both devices from March 19-29, 2021.

Predictive power of the model using subsets of the data

To evaluate the predictive power of the model, the complete dataset was divided into training and test sets, which were used repetitively to build the model and produce predictions of run counts. Sampling schema 1 consecutively added data grouped by week to the training set. The first two weeks of data did a poor job of making predictions on the remaining nine weeks of data (Fig. 2.10A). This result is not surprising as eDNA concentrations and sonar fish counts showed the poorest correlation in the first couple of weeks of the season ($\rho = 0.53$; Fig. 2.4). The correlation between schema 1's predictions and the observed fish counts increased with the addition of weeks 3 and 4 and reached an asymptote when the training set included 5-8 weeks of data when it reached a similar correlation coefficient as when using the entire dataset (Fig. 2.9A). At 9 and 10 weeks of data the correlations quickly decreased. The model trained on 10 weeks of data had only two values to make predictions for, fish counts of 0 and 22. The eDNA concentrations for these two points were both zeros but the predictions generated by the model were 256.99 and 390.87 fish, highlighting the difficulty this model has predicting zero values. The mean absolute error had a different pattern compared to the Pearson's correlation coefficient, showing a pattern of constant decrease with increasing training data (Fig. 2.9B).

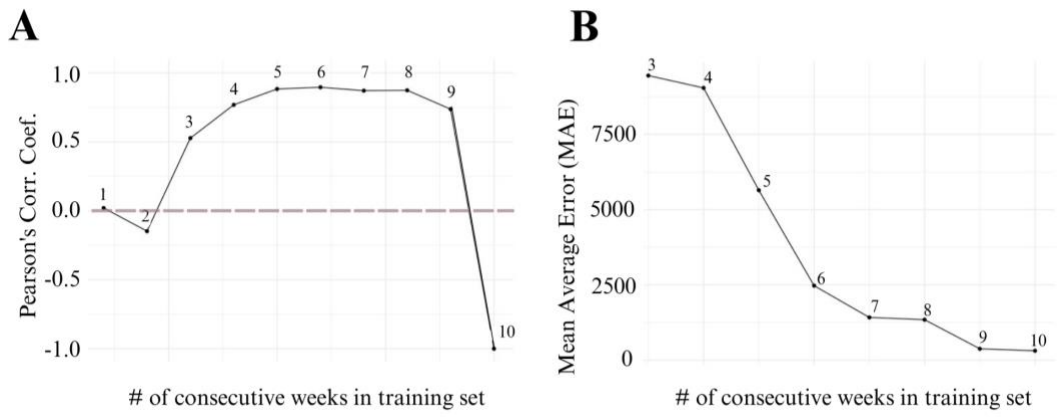

Figure 2.9: Fit metrics for schema 1, training the full model on consecutive sub-samples of data. Panel A shows the Pearson's Correlation Coefficient between the predictions made by the model trained on 1-10 weeks of data and the observed fish count values. Panel B shows the mean absolute error of predictions compared to the observed fish counts. Models generated on 1-2 weeks of data were omitted from the plot. Their mean absolute errors were 1.53×10^{13}, and 9.79×10^{45}, respectively.

Schema 2 randomized the addition of weekly data to examine if the model could cope with non-consecutive data. There was a similar pattern in Pearson's correlation coefficient to schema 1 (Fig. 2.9A & 2.10A), in which the correlation flattens out at 4 weeks of data but schema 2 did not have the same sharp decrease in correlation at the end of the season. Predictions generated by 2-6 weeks of data resulted in highly variable mean absolute errors (Figure 2.9B). For Schema 3, observations were randomly selected for sample sizes 1-69, and replicated 10 times. This analysis shows a steep reduction in average MAE when the sample size increases from 15 to 30 (Fig. 2.11). At a sample size of 40 the model predictions have a similar MAE to that of the full model suggesting the model needs at least five weeks of data to produce estimates similar to the observed values.

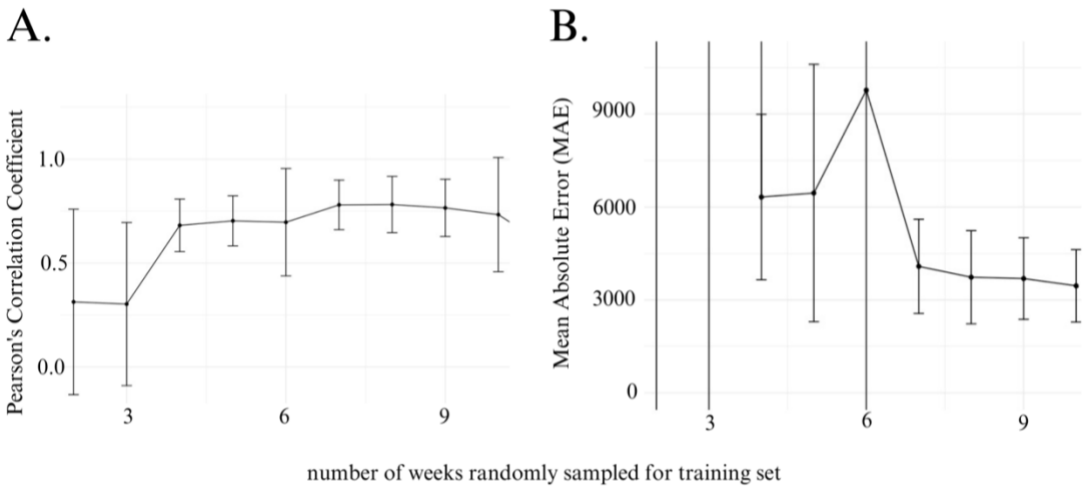

Figure 2.10: Schema 2: random sub-sampling by week. Observations were grouped by week and randomly selected to train the model; remaining observations were used to produce predictions. Panel A shows the Pearson's correlation coefficient between predicted and observed values in the test set over 10 repetitions and the standard deviation. Panel B shows the mean absolute error (MAE) for model predictions from subsets of data based on their expected values and the standard deviation.

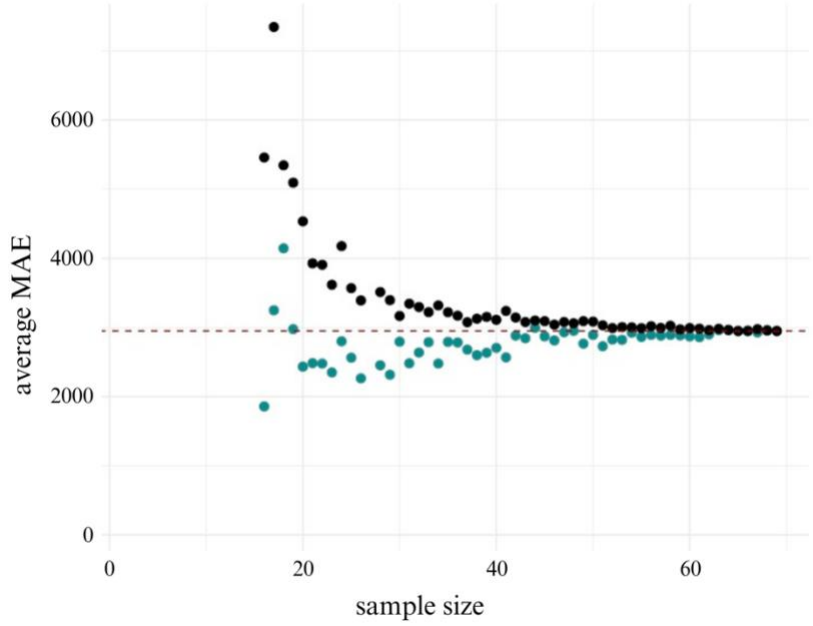

Figure 2.11. Learning curve for eDNA model predictions mean absolute error (MAE). Black points are from predictions made using the test set, teal points are from predictions made using the training set. The dashed red line is the full model's mean absolute error, 2949.56.

Discussion

The goals of this study were to improve the operationality of an existing eDNA qPCR assay for river herring by providing a direct comparison between eDNA abundance and fish count, a relevant management metric. We evaluated if high-frequency, remote eDNA sampling could be used to estimate fish count and species proportion, and assessed the necessity of including site-specific environmental variables that impact the detectability of eDNA in a predictive model.

Overall, our results suggest that daily sampling of eDNA concentrations can be used to quantify river herring abundance during their spawning runs. The inclusion of environmental covariates improved the predictions made by a model that used daily eDNA concentrations to generate fish count and indicates that information on migration triggers, site hydrology, and factors that influence eDNA persistence and transport are integral to making highly quantitative estimates of fish abundances.

eDNA concentrations effectively predict DIDSON-generated river herring counts

The capacity for eDNA monitoring to correlate to and expand upon information collected by other fish assessment methodologies has been extensively documented across ecosystems and species (Searcy et al., 2022; Levi et al., 2019, Thalinger et al., 2019; Tillotson et al., 2018; Thomsen & Willerslev, 2015, Takahara et al., 2012). These studies typically find strong correlations between eDNA concentrations and biomass, count or catch metric of choice, and associations are often improved by accounting for natural phenomena that influence eDNA detection probabilities (Yates et al., 2019). Our study supports the findings of these previous studies by demonstrating that daily eDNA concentrations were most highly correlated with 24-hour summations of fish count ($\rho = 0.84$), compared to a 72-hour count ($\rho = 0.72$). This highlights the short residency time of eDNA in the environment that makes it appropriate for generating estimates of local abundance at fine temporal scales.

As noted in the Results, the relationship between log-transformed flow-corrected eDNA and fish count was exponential in shape (Fig. 2.7) – i.e., as eDNA concentration rose, fish count

increased, but this increase slowed at higher values. This relationship is different than the simple assumption of linear increase of eDNA with increased fish count or biomass seen in some eDNA studies at low biomass levels (e.g., Erickson et al., 2016; Larson et al., 2017; Nevers et al., 2018). The nonlinear relationship we found has also been observed in eDNA studies working with large ranges of biomass (e.g., Levi et al., 2019; Coulter et al., 2019; Shelton et al., 2019) and could suggest the relationship between these two metrics saturates at some point. This saturation could be driven in part by the error and biases associated with either of the methodologies used. In regard to the fish count estimation, counts from imaging sonar may be less accurate at high densities (Holmes et al., 2006) or during certain fish behaviors (Magowan et al., 2012). During periods of high fish density, individual fish can be difficult to differentiate in sonar videos leading to undercounting. This could have caused the higher error around high fish count estimates (Fig. 2.7A). Fish milling is a behavior that can cause fish to swim back and forth in front of the imaging sonar, which can result in overcounting. Fish milling has been observed in river herring populations during their spawning migrations (Magowan et al., 2012; Rillahan et al., 2021) and was observed at the study site. A milling factor was added to the model, and higher fish counts were associated with eDNA when milling was present compared to when milling was absent (Fig. 2.7D); this could suggest milling activity was causing an overcounting of fish.

The exponential relationship between eDNA concentration and fish count at high values could also be due to error in estimating eDNA concentrations. This error is less straightforward to interpret due to the complex and competing relationships between eDNA and its environment

(Yates et al., 2021). Error in eDNA concentration measurement could be attributed to the different sources of eDNA associated with sampling during spawning runs (Tsuji et al., 2023), the dilution and dispersion effects of stream discharge (Levi et al., 2019; Searcy et al., 2023), and eDNA persistence in different sized biomasses (Jo et al., 2019). eDNA samples collected during spawning migrations can come from large communal releases of genetic material during spawning events that could impact the relationship between eDNA and fish count (Rourke et al., 2022; Tsuji & Shibata, 2023). Given the timing of eDNA samples taken in this study, it's likely the samples recovered contain spawning material (eggs and sperm). This could decouple eDNA and fish count at higher biomass or fish densities, causing eDNA to continue to increase even if the fish biomass isn't changing. eDNA decay dynamics could also cause the observed saturation with increasing biomass (Shelton et al., 2019). eDNA degradation occurs at different rates for differently sized biomass groups, and larger fish biomasses have been associated with increased rates of eDNA decay due to an increase in microbial activity, which is a major contributor to the degradation of eDNA in the environment (Jo et al., 2019). If eDNA degrades more quickly at high biomasses, changes in abundance at very high values may be harder to detect. Finally, stream discharge has a well-documented effect on eDNA detectability (Tillotson et al., 2018; Levi et al., 2019; Searcy et al., 2023). While multiplying eDNA by stream discharge has been cited to address this dilution effect (Levi et al., 2019) and was implemented in this study, this doesn't address other aspects of hydrodynamic dispersion. Stream discharge can also affect the patchiness of eDNA in the water column, with medium flow being the most efficient at dispersing eDNA evenly in the water column (Searcy et al., 2023). High or low flow events could disrupt the relationship between eDNA, and fish count due to dilution or lack of mixing,

respectively. Future studies could use hydrodynamic models that integrate the physical and chemical properties of eDNA with the hydrological parameters of a system (Harrison et al., 2019).

While the model produced using this dataset generated estimates of daily fish abundance that were close to the observed data, there are several caveats that should be addressed in future studies that apply similar modeling approaches. The selection of eDNA as the independent variable, while helpful for building a model that can make straightforward predictions on fish abundance, does not address the causality of eDNA in the environment. Obviously eDNA is produced (shed) by the fish, thus fish count is a more appropriate independent variable when addressing the causal relationship between these two variables. Future studies could utilize inverse prediction methods (LaMotte & Wells, 2016) to both capture the proper direction of the causal relationship and produce predictions of abundance. eDNA as the predictor also violates the regression assumption of no measurement error in the independent variable. eDNA concentration is highly variable in the natural environment, and measurement error could arise due to a number of factors associated with the ecology of eDNA or collection and analysis methods (Yao et al., 2022). Not accounting for this error can lead to attenuation bias in regression estimates. While this is a common violation in field studies (e.g., Spear et al., 2021), future studies could make use of error-in-variables models (e.g., Fuller, 1987) to address the error associated with eDNA as the independent variable.

Addressing the Ecology of eDNA at the Sampling Location

Overall, the strong relationship between fish abundance and flow-corrected eDNA concentration was sufficient to inform a model that could produce fish count estimates, but the predictive ability of this model was improved by the inclusion of average water temperature, day in season, and a binary milling factor. This highlights the importance of site-specific information when drawing conclusions about fish abundance in the field using eDNA data.

The utilization of flow-corrected eDNA in lotic systems has been applied in previous studies of anadromous fishes and other species as a way to mitigate the dilution effect increased stream discharge can have on the detectability of eDNA and has been shown to increase the correlation between eDNA and fish abundance datasets (Curtis et al., 2021; Tillotson et al., 2018; Levi et al., 2019). When fish pass by a fixed sampling point, they continue to contribute eDNA to the environment after that point, and that eDNA can be transported downstream towards the sampling site. The degree to which this transported eDNA contributes to an eDNA sample depends on how far the eDNA source is from the sampling site and the flow rate between shedding and collection (Nukazawa et al., 2018). The further upstream an individual travels, the less likely it is that their eDNA is contributing to eDNA samples at the sampling site. At our site, the inclusion of stream discharge as an interactive factor applied to eDNA concentration increased the Spearman's correlation coefficient between eDNA concentration and fish count (eDNA, = 0.84; flow-corrected eDNA = 0.86).

eDNA decay experiments in the presence of BAC preservative under ambient conditions did not show significant or consistent patterns of decay, which suggested minimal decay of the sample

with BAC and that the inclusion of decay in the eDNA model was not necessary. Simple preservatives that can be added to sampling bottles prior to sampling events reduce the need for cold storage between the time of collection and the time of filtration. Nevertheless, there are several other environmental factors and temporal dynamics that could have contributed to the absence of a strong decay effect observed in this study. The presence of humic compounds in the water could dampen the decay effect as DNA can form a strong bond to humic materials slowing down enzymatic activity (Eichmiller et al., 2016). In wetland areas, high levels of dissolved organic matter with increased humic content have been measured, which is a result of wetlands being at the nexus of land and water and the highly dynamic nature of biogeochemical processes in these environments (Hosen et al., 2018). The absence of a strong decay effect in eDNA decay experiments could also be due to the strategic timing and location of this study at a site with a strong river herring run during the height of the spawning migration from March to June. High concentrations of genetic material are released into the water during spawning activities and the ambient temperatures during the study period (avg air temp 14.3°C ± 5.0, Harris Market Station: KMDSALIS57) were moderate until the end of the season. Previous studies have shown that higher temperatures can increase eDNA decay (e.g., Tsuji et al., 2017; Jo et al., 2019; Andruszkiewicz Allan, et al., 2021), thus the limited decay we observed may be due to the relatively moderate temperatures during early spring in Maryland. However, eDNA decay occurring in the natural environment, between the time the DNA is released into the water and when it is collected, is still subject to environment-mediated decay rates and some decay likely occurred before we sampled. Water temperature was added to the full model because it appeared to increase the model fit (inclusion decreased the AIC of the full model). The model predicted

higher fish counts for samples taken at higher temperatures (see Fig. 2.7C), perhaps accounting for decay occurring in the environment prior to sampling.

The inclusion of a milling factor improved the fit of the model and model predictions were higher when milling was present compared to when milling was absent (Fig. 2.7D). This result indicates that high fish activity and a milling behavior where fish congregate and swim back and forth instead of swimming swiftly by the imaging sonar may reduce the accuracy of upstream fish counts. Milling behavior is often observed in river herring during fish counting and can artificially inflate run count estimates (Magowan et al., 2012). Milling may be driven by a feature of river topography or hydrology as certain small scale stretches of the system may induce more milling behavior than other areas. Magowan et al., (2012) hypothesized that the location of their imaging sonar device, in a deep and protected region of the Herring River, was the reason they observed high milling activity during the day, and they recommended the site not be used for future studies as it obscured the ability to generate accurate fish counts from DIDSON videos. Rillahan et al., (2021) evaluated milling behaviors in different species during sonar-based fish counting and found that alewife exhibited milling in a scour pool during high flow events. eDNA sampling may be less sensitive to milling behavior as there's no way to tell if the eDNA collected came from a unique individual or a milling individual. While we found interesting and significant effects of measured environmental variables that improved the modeling of eDNA and fish counts, there are many variables not considered here that may have an effect on the relationship between eDNA and fish abundance. Increasing the number of sampling locations to better understand how site-specific factors like stream discharge and river

topography influence the detectability of eDNA would improve site selection and could generalize the eDNA model proposed here and make it more useful across different river systems. Future studies utilizing the same sampling approach and location during a spawning run can add to these data and examine how well this model performs given interannual fluctuations in river herring species abundance (Bi et al., 2021).

eDNA produces comparable species proportions to electrofishing

Species proportions generated by eDNA (qPCR), and electrofishing were highly correlated (Fig. 2.6), suggesting that eDNA sampling (and Sanger validation) can produce species proportions comparable to electrofishing surveys. This result suggests that alewife and blueback may have similar eDNA shedding rates. If the species had significantly different shedding rates, the correlation between eDNA and electrofishing species proportions would likely be much weaker than estimated. For example, if alewife produced 2 or 3 times as much eDNA than blueback but were at a 1:1 ratio in the river (observed via electrofishing), eDNA vs electrofishing proportions would be less well correlated. It is perhaps somewhat expected that similarly sized, related alosine fishes would shed at similar rates, but these data are the first to explicitly indicate that shedding rates may be similar between the river herring in field-based sampling. This result is caveated by the error and bias in both methodologies. Electrofishing is highly size selective towards larger individuals, and the catchability of fish in lotic systems can be overwhelmed when high volumes of fish are present (Bohlin et al., 1989). Additionally, eDNA data may have size biases as different-sized animals are expected to have different eDNA shedding rates (Yates et al., 2019). Nevertheless, our results suggests that eDNA sampling (qPCR and Sanger

validation) on a spawning run is able generate species proportion data similar to that produced by electrofishing. Based on the relationship between the two methodologies in this instance and the limitations of electrofishing, eDNA has the potential to be a more cost effective and accurate method for determining species proportions compared to electrofishing (Evans et al., 2017). More broadly, eDNA could be used to mitigate data uncertainties in current monitoring efforts as river herring are commonly reported as a mixed stock and visual species identification can be difficult among closely related alosines (AMSFC, 2012).

Utility of eDNA in management of river herring and alosines

Effective monitoring of natural populations relies on obtaining accurate data at relevant ecological and management time scales (Thomsen & Willerslev, 2015). Traditional monitoring methods, such as visual or sonar counts, often fall short of these goals due to the effort-intensive and scale-constrained nature of these types of surveys (Yao et al., 2022). eDNA monitoring can reduce the cost and labor associated with fish monitoring (Evans et al., 2017) and could allow for monitoring on a larger scale, spatially or temporally, increasing the accessibility of information on natural populations that can be used to inform management decisions (Lodge et al., 2012). eDNA monitoring, as implemented in this study, produced a very similar run count of 729,385 based on eDNA, differing only by 0.5% percent from the sonar-based estimate. Given that run counts for river herring are only available for a small fraction of the water systems and rivers they use, the potential for eDNA data to produce lower-cost and quantitative run counts could greatly expand the number of systems that can be feasibly monitored for river herring and for which river herring run counts can be generated.

Our study indicates that relatively limited count-based sampling at this site (4-6 weeks) can inform a model that can generate fish count predictions based on eDNA sampling for the remainder of the season. Randomly subsampling the data over the spawning run and re-fitting the model to those subsets indicated the model needed at least 40-50 data points to produce MAE values similar to that of the full model (Fig. 2.11). Consecutive and randomized weekly sub-sampling of the season dataset highlighted how eDNA sampling could reduce the need for daily sonar sampling throughout the entire season. The model could produce estimates similar to those generated by the full model using at least five weeks of training data. Analyzing DIDSON fish count videos takes considerable time and effort and even a reduction of just 30 days can eliminate 15 workdays per season. Finally, it is worth briefly discussing the finding that a model trained on nine and ten weeks of data produced Pearson correlation coefficients lower than models trained on 5-8 weeks of data (See Fig. 2.8A). This may be a result of the fact that the estimation points in the nine- and ten-week test sets were from the end of the season, which contained more zero counts than the training dataset. This highlights the potential limited ability for the current model to produce accurate estimates when run counts are low or zero.

While this model performed well on this dataset, before quantitative measures of fish abundance based on eDNA data can be used more widely, studies like this one must be repeated on different temporal and spatial scales. Sampling before and after the spawning migration could provide a baseline should target eDNA be present in the system outside the spawning season, and it would also provide additional non-detection observations, which was a limitation for the model built in

this study. Sampling at the same location over multiple years could evaluate how a model developed and built in one year performed on eDNA sampling from a different year. Increasing the number of sampling locations on a river system would allow for a more complete understanding of the detection, persistence, and transport of eDNA for that system. A fine spatial scale study could build hydrological models for eDNA transport, aiding in a more accurate biological interpretation of eDNA results (Harrison et al., 2019). Finally, sampling over a wider geographic area would assess the site-specific dependence of the relationship between eDNA and fish count and test current recommendations on optimal sampling strategies in riverine networks (Carraro et al., 2021).

Conclusions

At our sampling location on the Choptank River in Greensboro, MD (Fig 2.1), we found that eDNA monitoring has the potential to be a valuable supplement to traditional fisheries assessment metrics. Using an automatic sampling device and an eDNA preservative strikes a balance between the efficacy of daily sampling and the resolution required to generate fish abundance estimates using eDNA. This study builds on preexisting river herring monitoring tools (Ogburn et al., 2017a; Plough et al., 2018, Ogburn et al., 2023), and establishes a model by which eDNA concentrations estimated daily can produce estimates of fish count, providing an empirical linkage between eDNA data and management-relevant measures of fish abundance. This work emphasizes the necessity for careful site selection, as the inclusion of important environmental covariates improved the fit of the model used to generate fish count predictions. Future studies could also increase the spatial breadth of this study within the same system or

across different systems with known river herring runs. High-frequency sampling over a fine spatial scale (i.e., more sampling sites along the Choptank river) would allow for a deeper understanding of the eDNA dynamics occurring within that system. Repeating this study design on different rivers with known river herring runs but different hydrological conditions would allow for the further exploration of site-specific factors on the relationship between eDNA and fish abundance. This study provides a critical connection between eDNA data and management-relevant metrics supporting the potential utility of eDNA monitoring in fish management, and future use and optimization of remote autosampler systems like the ISCO for eDNA collection may be key to providing highly quantitative eDNA abundance data for estimating run counts of anadromous species.

Chapter 3: Development and Testing of Molecular Beacon Assays for eDNA monitoring of Mid-Atlantic American and Hickory Shad

Abstract

Highly migratory anadromous fish populations are vital linkages between terrestrial and aquatic ecosystems. In the Mid-Atlantic, American shad, hickory shad, and river herring are anadromous species of conservation concern that can be difficult to monitor due to their broad distribution into hard-to-reach habitats during spawning migrations. Environmental DNA (eDNA) monitoring techniques have been used successfully to evaluate river herring populations throughout the Chesapeake Bay watershed. To contribute to the management and conservation of these shad species and complete a set of qPCR eDNA detection assays for the Mid-Atlantic alosines, we developed, optimized, and field validated species-specific qPCR assays for American shad and hickory shad. Limit of detection (LOD), qPCR efficiency, and limit of quantification (LOQ) were estimated for the two novel shad assays and the previously developed river herring assay showed high sensitivity (LOD < 13 copies). Field testing of the two novel shad assays showed validated detections of both species in Delaware and the Chesapeake Bay, and for American shad in North Carolina. This work completes a set of eDNA qPCR detection assays for the alosines in the mid-Atlantic, providing valuable new resources for the monitoring and management of these species.

Introduction

River herring, American and hickory shads are ecologically significant anadromous fishes that inhabit coastal waters off the east coast of North America and have seen substantial population declines over the last century (Limburg & Waldman, 2009; Hasselman & Limburg, 2012). Alewife and blueback herring, collectively known as river herring, range from Newfoundland, Canada, to South Carolina, USA and Nova Scotia, Canada, to Florida, USA, respectively. American shad (*Alosa sapidissima*) range from Florida, USA to Quebec, Canada, while hickory shad (*Alosa mediocris*) range from Florida, USA to Maine, USA (Limburg and Waldman, 2009). Their widespread distribution and highly migratory life history makes them an integral linkage between pelagic, littoral, benthic, and terrestrial food webs through-out their range (Bi et al., 20121; Limburg et al., 2003). At sea, mature adults are common forage fish and reduce predation pressure on endangered Atlantic salmon by acting as an alternative food source (Saunders et al., 2006). During spawning, mature adults transport marine-derived nutrients inland (Zydlewski et al., 2021). As juveniles, they feed on zooplankton, insects, and small aquatic invertebrates in both riparian and estuarine systems (Limburg et al., 2003). The predictability of annual spawning migrations has made anadromous fishes an important resource for North Americans for over 200 years (Cushman et al., 2012; Meehan, 1893).

At the turn of the 19th century, American shad was the top commercially fished species in the United States (Limburg et al., 2003). Due to overharvesting, habitat fragmentation, and declining water quality, commercial landings plummeted in the late 20th century, and the fishery was closed in Maryland in 1980 (AMSFC, 2020). A similar narrative exists for the river herring, who

are often subject to the same management regulations as the shads due to their highly similar life history and historic patterns of decline (Hare et al., 2021). Hickory shad is also managed with the previously mentioned species but is a severely understudied species (Murauskas & Rulifson, 2011). Despite management efforts dating back to the early 1900s, these species are at historic lows indicating the need for more accurate data to inform effective adaptive management strategies. Environmental DNA (eDNA) qPCR detection tools have tremendous potential when it comes to increasing the cost-effectiveness of bioassessments and providing timely information on natural populations.

eDNA techniques have been used to document species distributions, movements, and colonization by using highly fragmented genetic material, shed by an organism of interest into the environment, as an indicator of species presence (Yao et al., 2022). Species-specific population surveillance and monitoring is one of the major trajectories of macrobial eDNA research (Lodge et al., 2012) and has been revered for its higher detection limits and cost-effectiveness when compared to traditional assessment methodologies (Pfleger et al., 2016; Evans et al., 2017; Fediajevaite et al., 2021). Boothroyd et al., (2016) used an eDNA single species detection assay to examine habitat occupancy of the threatened spotted gar (*Lepisosteus oculatus*) in southern Ontario, Canada. This study uncovered positive eDNA detections in a region where the spotted gar was previously considered expatriated, highlighting the sensitivity of eDNA compared to capture-based methods. Evans et al., (2017) conducted a cost and effort analysis comparing eDNA single species detection with electrofishing for brook trout (*Salvelinus fontinalis*) in the upper Namekagon river watershed finding that eDNA single-species detection

was less labor intensive and 67% cheaper than multiple pass electrofishing. These studies exemplify the potential for eDNA monitoring to be used to address different questions on a species of interest while mitigating costs. There has been success utilizing eDNA monitoring tools for populations of river herring in the Mid-Atlantic, and eDNA monitoring tools already exist for the river herring species (Plough et al., 2018).

A dual-species molecular beacon qPCR assay for river herring has been used to successfully map the spatial distribution of each species (Plough et al., 2018; Ogburn et al., 2023), monitor the effect of restoration efforts on local populations (Huang et al., 2023), and predict run count estimates (Chapter 1). eDNA monitoring could potentially be used to examine habitat occupancy of the hickory shad, since little is known about the distribution of their populations (Murauskas & Rulifson, 2011). eDNA monitoring also has potential in evaluating habitat use, the impacts of fish stocking, and recolonization of reconnected historic spawning grounds for the American shad (Bi et al., 2021; Hasselman & Limburg, 2012; Zydlewski et al., 2021). Here we have developed new single species eDNA detection assays for Hickory and American shad, and have performed comprehensive analyses of assay specificity, assay efficiency, and assay performance (limits of detection and quantification). Novel eDNA assays for these species will provide valuable new resources for their monitoring and management.

Methods

Assay design

The assays were developed and tested using methods outlined by Plough et al., 2018. Sequences for mitochondrial genes commonly targeted by eDNA assays, cytochrome oxidase subunit I, and cytochrome b (Shu et al., 2020), were downloaded for the target species, the American and hickory shads, from the National Center for Biotechnology Information (NCBI) genetic sequence database, Genbank. Mitochondrial genes are commonly used for single-species detection as there are more copies of these genes per cell, they have a high mutation rate making them highly species-specific, and they have broad coverage in genetic databases (Tsuji et al., 2019).

Sequences for closely related non-target species that utilize the same habitats as the target shads, such as alewife (*Alosa pseudoharengus*) and blueback herring (*Alosa aestivalis*), Atlantic menhaden (*Brevoortia tyrannus*), American gizzard shad (*Dorosoma cepedianum*), and bay anchovy (*Anchoa mitchilli*), were also downloaded from NCBI to evaluate design specificity. Target and non-target sequences were aligned using AliView v.1.26 and visually inspected for areas of intra-species similarity and inter-species dissimilarity between target species and among closely related species. Molecular beacon assays for these regions were developed using Sigma-Aldrich's Oligo Architect (Merck KgaA, Darmstadt, Germany). Candidate assays primers and probes were between 18-50 bp in length, had melting temperatures ™ 50-60 °C, and at least one mismatch base with each non-target species. Assay designs that satisfied these initial constraints were ordered from IDT (Integrated DNA Technologies, Coralville, IA) for laboratory testing.

Laboratory testing and optimization

Assays ordered for laboratory testing were initially tested on DNA extracted from fin clip samples. qPCR reactions were run in 20 uL volumes with 10 uL of Biorad 2x SSO advanced universal probe supermix, 0.7 uL of each primer and the molecular beacon probe, 3.9 uL water, and 4 uL sample extract. Negative controls were run with every batch of samples to evaluate the potential for primer dimer fluorescence to rise above background fluorescence. Assays were initially run using the thermocycling parameters for the river herring assay (Plough et al., 2018). Briefly, PCR conditions for each quantification run were: 95°C for 45 seconds, 44 cycles of 95°C for 5 seconds, 59°C for 15 seconds, and 72°C for 10 seconds, with a plate read after each of the 44 cycles. Assay designs that had low fluorescence efficiency or high cross-amplification of non-target species were removed. Assay designs that had some cross-amplification of species were retained for testing as the modification of PCR conditions can improve assay performance and specificity (Lorenz, 2012). All qPCR analysis was conducted using a CFX 96 qPCR instrument (Biorad Hercules, CA).

Assays that had high efficiency and low non-target amplification were optimized by modifying the annealing temperature and modifying the qPCR reaction protocol. A temperature gradient from 50-65°C was run for annealing temperature, and the qPCR reaction protocol was modified by reducing the amount of 2x SSO supermix, primers, and probe. Optimized assays were used to amplify target species fin clip samples, and the resulting PCR product was run on a gel using gel electrophoresis to ensure the amplicon was of the proper size. qPCR efficiency is calculated using a standard curve, where known quantities of the target amplicon from a five-fold serial dilution are amplified. The known copy numbers from the serial dilution are $\log_{10}$-transformed

and regressed against qPCR cycle numbers (Cq). qPCR efficiency was defined as the fraction of target molecules copied in one PCR cycle (Svec et al., 2015) was calculated using Equation 3.

$$E = 10^{-\frac{1}{m}} - 1$$

Equation 3: qPCR efficiency calculated using the slope of the standard curve (m)

Limit of detection and limit of quantification

A limit of detection (LOD) and a limit of quantification (LOQ) analysis were generated for each of the molecular beacon assays designed here using a discrete threshold method (Forootan et al., 2017; Klymus et al., 2020). Using methods in Forootan et al., 2017, a serial dilution from 300,000 copies to 3 copies was prepared using a synthesized geneblock of the target species amplicon (Integrated DNA Technologies Inc., Coralville, IA). Three plates were prepared using the previously documented qPCR methods (previous section) with known quantities of synthetic DNA ranging from 300,000 copies to 3 copies. These samples were amplified in many replicates to determine the minimum amount of target eDNA needed to produce a positive detection at a defined level of confidence (LOD), typically at 95% (Klymus et al., 2020). The minimum amount of target eDNA at which the variation in replicates' Cq values is under 35% (LOQ) was also identified (Forootan et al., 2017; Klymus et al., 2020). DNA copy numbers 300,000, 30,000, 3,000, 300, 150, 100, 50, 30, 25, 15, 10, 5, 3, were run with 14-28 replicates each.

The fraction of positive detections was determined for each standard plotted against copy number. A sigmoidal curve was fit to these data to allow for the interpolation of the LOD

(Forootan et al., 2017). The LOD was established at a 90% confidence level due to the relatively low number of replicates used. In previous studies, LOD evaluated at 95% confidence used 64-128 replicates per standard (Forootan et al., 2017; Klymus et al., 2020). The LOQ was found using the coefficient of variation for log-normal data of each standard replicate (Equation 4). The LOQ was the lowest copy number standard with a coefficient of variation <35% (Klymus et al., 2020). LOD, LOQ, and efficiency was also measured for the river herring assay from Plough et al., (2018) as a comparison and to complete the assay's diagnostics.

$$CV_{ln} = \sqrt{(1+E)^{(SD(Cq))^2 \times \ln(1+E)} - 1}$$

Equation 4: Equation from Frootan et al., (2017), the coefficient of variation for log-normal distributed data. Used to estimate the limit of quantification, E is qPCR efficiency, SD(Cq) is the standard deviation of replicate Cq values.

Assay field testing

Once assays were optimized and working efficiently on lab samples, eDNA samples previously collected from across the Chesapeake Bay watershed, were selected for field testing (e.g., samples collected and processed in Plough et al., 2018, Ogburn et al., 2023). These samples were selected based on the probability of presence for the species of interest based on traditional methodology observations. Samples assessed were from 2015-2022 from the Choptank, Rappahannock, Patapsco, Northeast, Potomac, and Pamunkey Rivers (see Fig. 3.3 & 3.7). These samples were collected and extracted in Plough et al., (2018) and Ogburn et al., (2023) and thus followed those previously published methods. Briefly, samples were collected in 1-L volumes from sites across the Chesapeake Bay, filtered on 47 mm diameter cellulose nitrate filter papers

with a pore size of 1.0 micron (Cytiva-Whatman), and frozen at -80°C until extraction with the EZNA water kit (Omeg Biotek) or CTAB method (Renshaw et al., 2015).

Samples were quantified using the same methods used for laboratory testing of the assays. Due to low detection probabilities for American shad in the samples selected for field testing, additional samples were also collected from the Potomac River in April 2023 by Phong Trieu at Metropolitan Washington Council of Governments (MWCOG). These samples were filtered, extracted, and tested using the same methodology described above. Water samples from Cape Fear River in North Carolina were also collected by partners from the North Carolina Wildlife Resources Commission during electrofishing activities, providing a traditional metric (catch per unit effort, CPUE) to compare to eDNA detections. eDNA filtration and extraction methods for the North Carolina samples followed identical protocols to Plough et al., (2018).

Results

Two assays were designed for American and Hickory shad species, respectively. Based on the methods used to develop these tools, they can be categorized as level 4 (out of 5) assays by the eDNA assay validation scale generated by Thalinger et al., 2021. Level 4 assays have undergone assay development and optimization, field and lab testing, and have diagnostics like LOD, LOQ and qPCR efficiency. At level 4, the results generated by these assays are reliable. Positive detections indicate the species is very likely present, and negative detections indicate the species is likely absent. Environmental factors have a considerable effect on eDNA detection and have

been cited as a source of error in the development and testing of the river herring assay (Plough et al., 2018). Additional testing would clarify if the environment influences detection probabilities which can cause false negatives; this testing would elevate the assays to Level 5.

American shad eDNA qPCR molecular beacon assay

The American shad assay amplified an 88 bp region of the cytochrome oxidase subunit 1 gene in the mitochondrial genome (Table 3). The final assay had seven base mismatches with Hickory shad, five base mismatches with alewife, six base mismatches with blueback herring, eleven base mismatches with gizzard shad, and five base mismatches with menhaden (Fig. 3.1). Initial laboratory testing using fin clip samples from target and non-target species from the Chesapeake Bay showed some issues with PCR stringency (Fig. 3.2A). A range of annealing temperatures and $MgCl_2$ concentrations were tested to find the conditions with the most consistent amplification and lowest cross-species amplification. Full strength primer, probe, and SSO 2x supermix concentrations were most effective in maintaining high qPCR efficiency and an annealing temperature of 56°C eliminated cross-species amplification signals.

	5' ⟶ 3'
Forward Primer	**GTG AAC AGT CTA CCC ACC TTT G**
Reverse Primer	**ACC TGC TAG ATG AAG AGA GAA GA**
Molecular Beacon Probe	**6-FAM/CGC GAT CCC GGA GCA TCC GTC GAC CTA ACT AGA TCG CG/BHQ_1**

Table 3: American shad molecular beacon qPCR eDNA assay components.

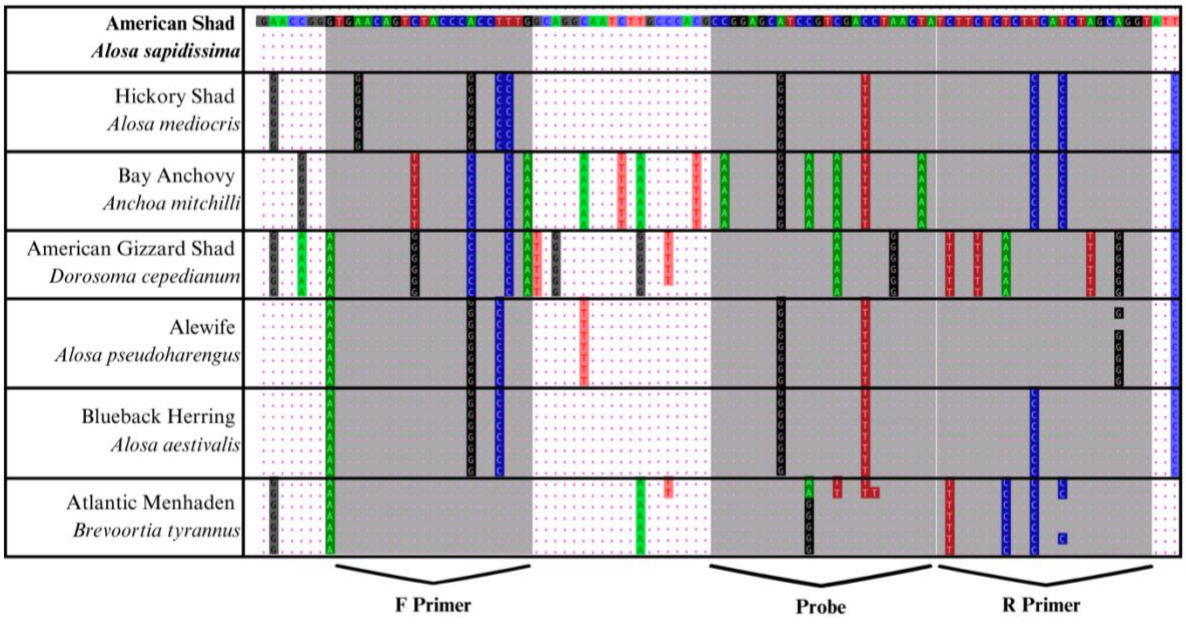

Figure 3.1: Alignment of the mitochondrial gene cytochrome oxidase 1 for American shad and closely related non-targets. Bases highlighted are non-consensus based on the target species sequence, bases denoted by a pink dot are the same as the base at that position in the top sequence. Gray highlighted regions are the primer and probe target regions.

Once the optimal conditions for qPCR quantification were established, the assay was tested on environmental samples taken from 2015-2023 around the Chesapeake Bay (Figure 3.3). In total 81 samples environmental samples were tested for the presence of American shad. Out of 81 samples, 32 (40%) had at least one qPCR replicate amplify (Fig. 3.3 and 3.4). PCR products from amplification using the American shad assay were sent out for Sanger sequencing validation (MClabs, San Francisco, CA).

A limit of detection (LOD) and limit of quantification (LOQ) were developed for the American shad assay (Fig. 3.5). The equation for the standard curve by which copy number is discerned from cycle number (Cq) was $Cq = -3.48\log_{10}(\text{eDNA concentration}) + 41.12$ with an R-squared

0.98 for all serial dilutions included in the LOD/LOQ analysis (Fig. 3.5A). The limit of detection, the quantity at which 90% of replicates amplify, for the American shad assay was 8.57 copies (Fig. 3.5B). The limit of quantification, the quantity at which replicate samples have a coefficient of variation of <35%, was 30 copies (Fig. 3.5C).

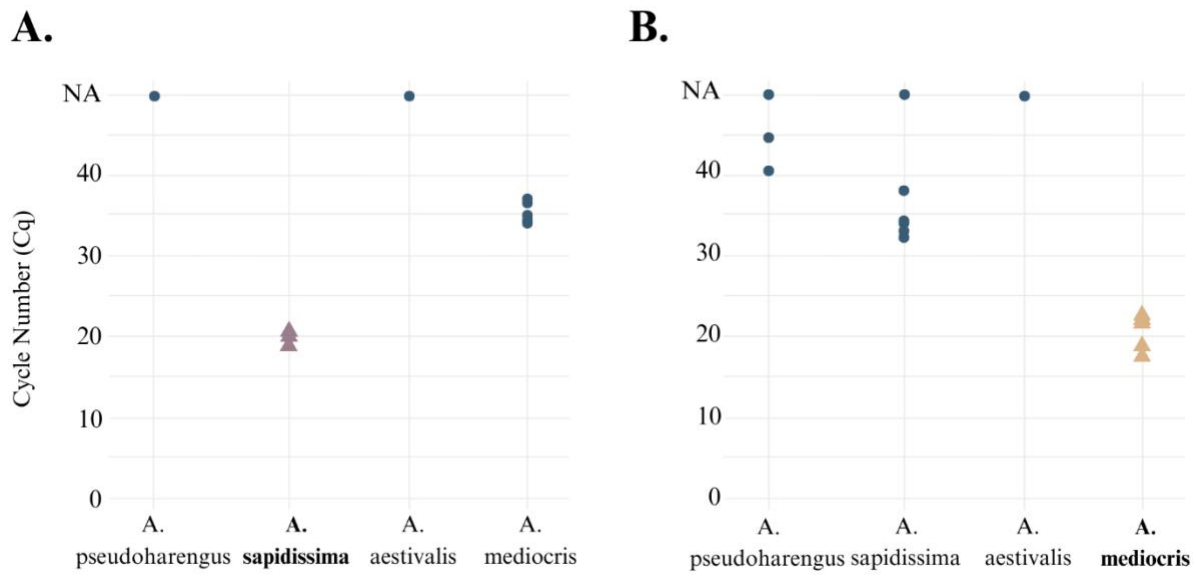

Figure 3.2: Results from laboratory testing of both assays using fin clip samples for target and non-target species. The y-axis is qPCR cycle number (Cq), those samples that did not amplify within 40 cycles are denoted by the 'NA' value at the top of the scale. Along the x-axis are the target and non-target species tested using the assay, the target species, for whom the assay was designed, is in bold. Panel A shows the amplification of target (American shad, *A. sapidissima*) and non-target species using the American shad assay. Panel B shows the amplification of target (Hickory shad, *A. mediocris*) and non-target species using the Hickory shad assay.

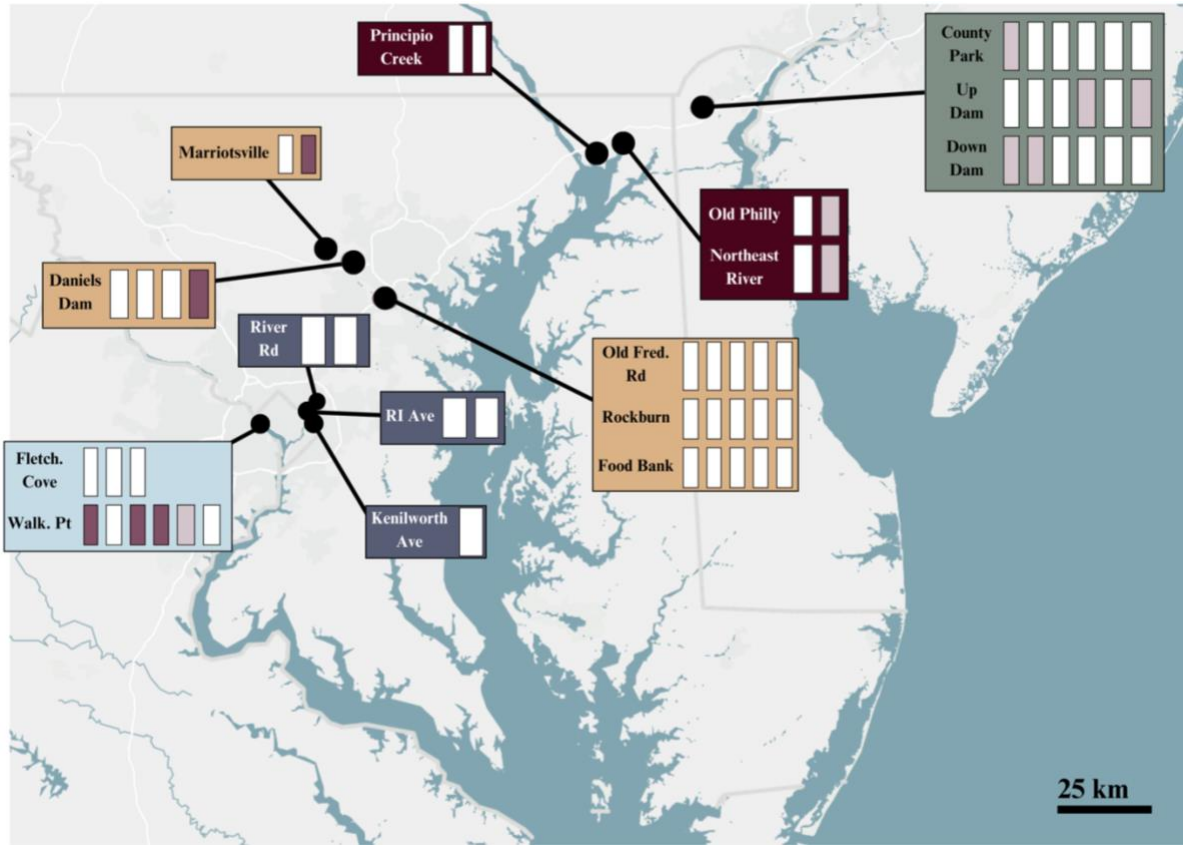

Figure 3.3 American Shad eDNA samples and detections in the Chesapeake Bay 2015-2023. Each rectangle represents one eDNA sample. The shade of the rectangle is the number of positive qPCR replicates (0-3). White is 0/3 replicates (non-detection), light pink 1/3 and dark pink 3/3. No samples had 2/3 positive replicates. The colored boxes indicate the river system: green is the Christiana River, maroon is the Northeast River, yellow is the Patapsco River, dark blue is the Anacostia River and light blue is the Potomac River.

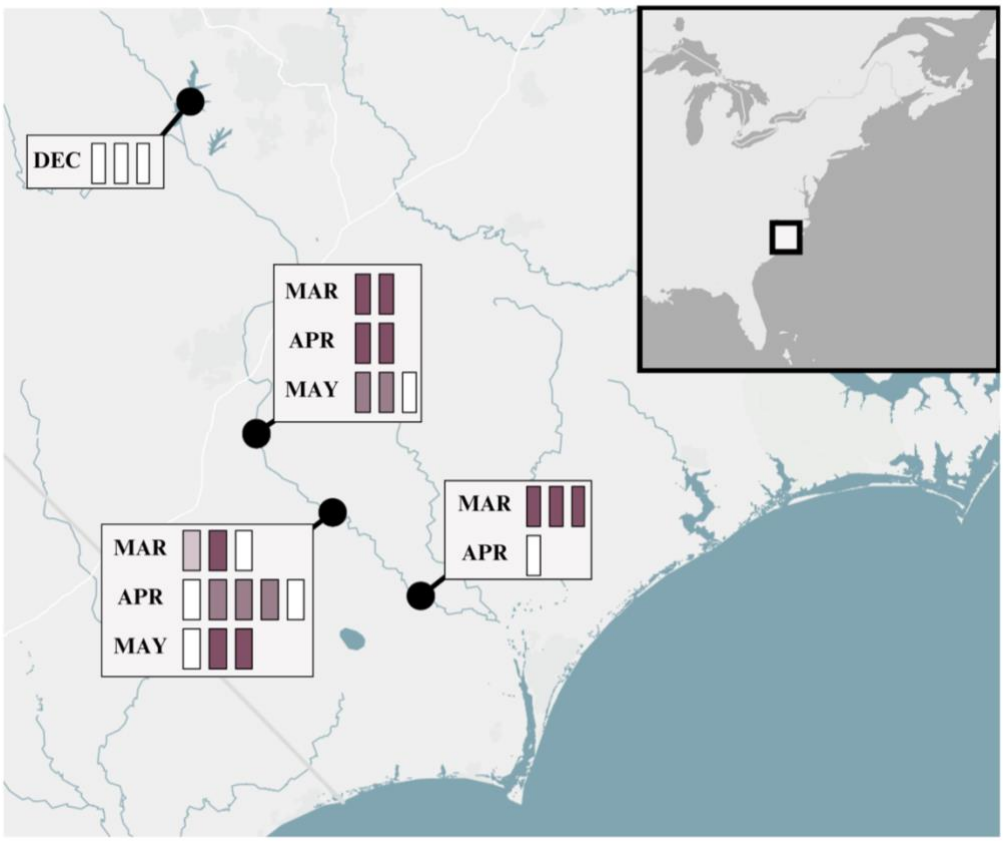

Figure 3.4. American Shad eDNA samples and detections in Cape Fear, North Carolina, 2021. Each rectangle represents one eDNA sample. The shade of the rectangle is the number of positive qPCR replicates (0-3). White is 0/3 (non-detection), the lightest shade of pink is 1/3, the intermediate shade is 2/3, and the dark pink is 3/3.

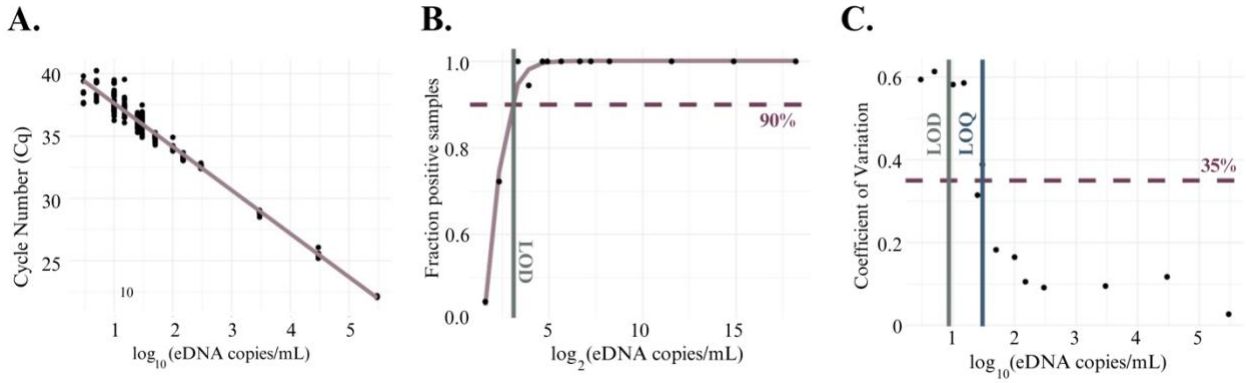

Figure 3.5: American Shad assay limit of detection and limit of quantification metrics. Panel A shows the quantification of all replicates (3-300000 copies) based on a standard curve. Panel B

shows the assay limit of detection at a 90% positive detections per standard threshold. Panel C shows the assay limit of quantification with the 35% coefficient of variation threshold in pink.

Hickory shad eDNA qPCR molecular beacon assay

A dual-labeled molecular beacon probe assay was also developed for Hickory shad amplifying a 108 bp region of the cytochrome oxidase subunit 1 gene in the mitochondrial genome (Table 4). The final assay had three base mismatches with American shad, seven base mismatches with alewife, eight base mismatches with blueback herring, nine base mismatches with gizzard shad, eleven mismatches with bay anchovy, and seven base mismatches with menhaden (Fig. 3.6). Similar to the American shad assay, initial laboratory testing showed some, late Cq value cross-amplification of non-target species (Fig. 3.2B). In this case, lowering the annealing temperature addressed cross-amplification.

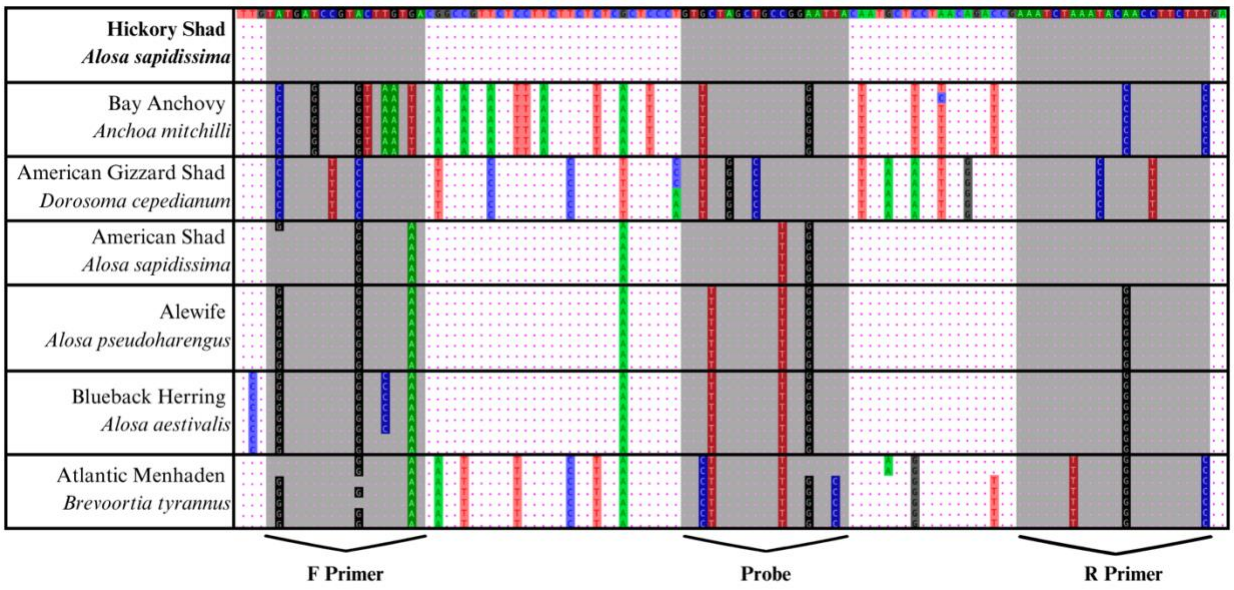

Figure 3.6: Alignment of the mitochondrial gene cytochrome oxidase 1 for Hickory shad and closely related non-targets. Bases highlighted are non-consensus bases based on the target species sequence, bases denoted by a pink dot are the same as the base at that position in the top

sequence. Gray highlighted regions are the primer and probe target regions. Alignment generated using AliView.

Once the optimal conditions for qPCR quantification were established, the assay was tested on environmental samples taken 2015-2023 from around the Chesapeake Bay. In total 57 environmental samples were tested for the presence of hickory shad. Out of 57 samples, 23 (40%) had at least one qPCR replicate amplify (Fig. 3.7). PCR product from amplification using the Hickory shad assay were sent out for Sanger sequencing validation (MClabs, San Francisco, CA).

Assay LOD and LOQ were estimated (Fig. 3.8). The equation for the standard curve by which copy number is discerned from cycle number (Cq) was $Cq = -3.36\log_{10}(\text{eDNA concentration}) + 39.14$ with an R-squared 0.98 for all serial dilutions included in the LOD/LOQ analysis (Fig. 3.8A). The LOD, for the Hickory shad assay was 4 copies per individual reaction (Fig. 3.8B) and the LOQ was 20 copies (Fig. 3.8C).

	5' ⟶ 3'
Forward Primer	**TAT GAT CCG TAC TTG TGA**
Reverse Primer	**AAA GAA GGT TGT ATT TAG ATT T**
Molecular Beacon Probe	**6-FAM/CGC GAT CTA ATT CCG GCA GCT AGC ACG ATC GCG/BHQ_1**

Table 4: Hickory shad molecular beacon assay description

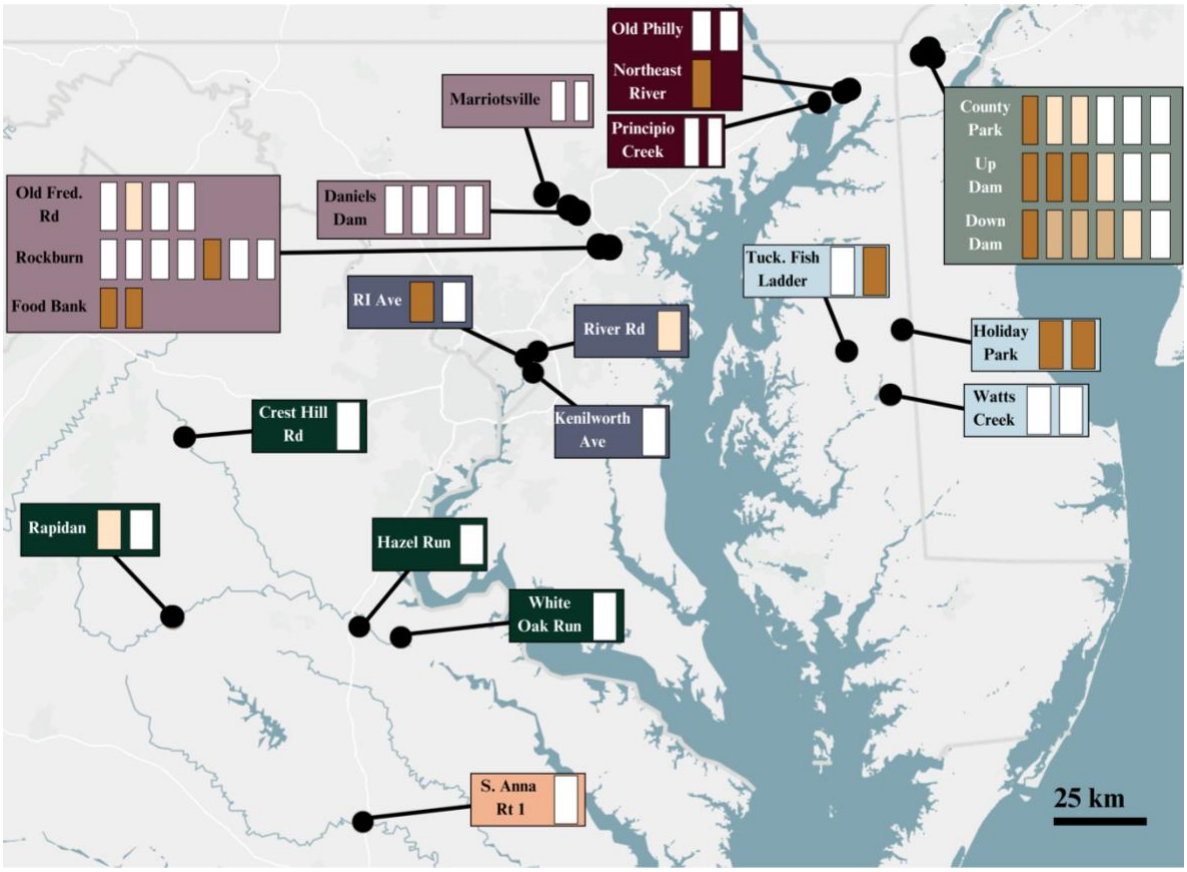

Figure 3.7: Hickory Shad eDNA samples and detections in the Chesapeake Bay 2015-2022. Each rectangle represents one eDNA sample. The shade of the rectangle is the number of positive qPCR replicates (0-3). White is 0/3 replicates (non-detection), light yellow 1/3, the intermediate shade is 2/3 and dark yellow is 3/3. The colored boxes indicate the river system: light green is the Christiana River, maroon is the Northeast River, pink is the Patapsco River, light blue is the Choptank River, dark blue is the Anacostia River and dark green is the Rappahannock River, and orange is the Pamunkey River.

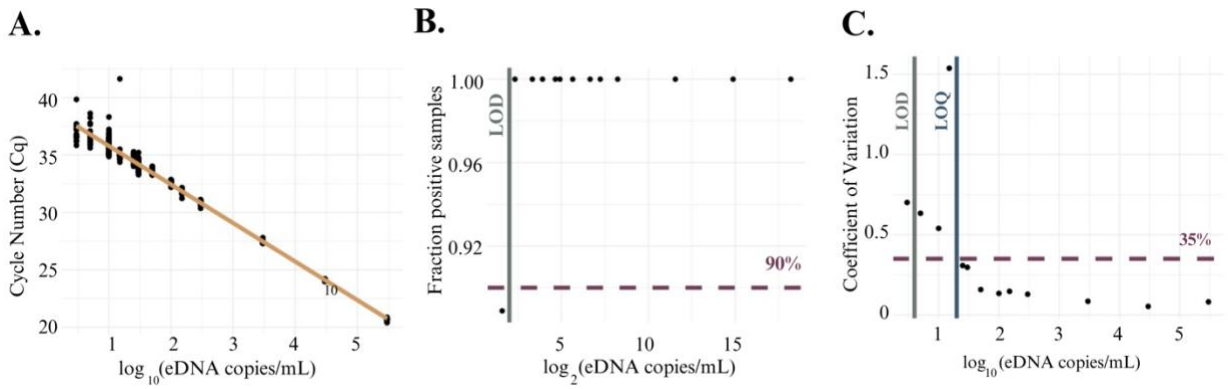

Figure 3.8: Hickory Shad assay limit of detection and limit of quantification metrics. Panel A shows the quantification of all replicates (3-300000 copies) based on a standard curve. Panel B shows the assay limit of detection at a 90% positive detections per standard threshold. The green 'LOD' line indicates the calculated LOD, 4 copies. Panel C shows the assay limit of quantification with the 35% coefficient of variation threshold in pink. The green line is the LOD, and the blue line is the LOQ, 20 copies.

Limits of detection and quantification for river herring molecular beacon assay

The equation of the standard curve for river herring cytochrome oxidase one assay based on all of the serial dilutions performed for LOD/LOQ determination was Cq = -3.55log10(eDNA concentration) + 41.22 with an R2 value of 0.86 (Fig. 3.9A). The limit of detection for the river herring assay was 12.55 copies for an individual qPCR reaction (Fig. 3.9B). The limit of quantification was 100 copies (Figure 3.9C).

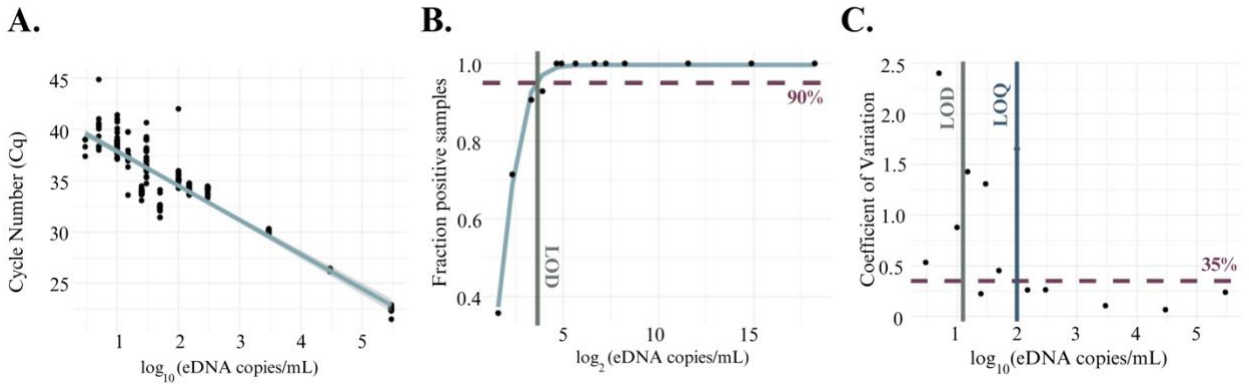

Figure 3.9: River Herring assay limit of detection and limit of quantification metrics. Panel A shows the quantification of all replicates (3-300000 copies) based on a standard curve. Panel B shows the assay limit of detection at a 90% positive detections per standard threshold. The green line is the calculated LOD, 12.55 copies. Panel C shows the assay limit of quantification with the 35% coefficient of variation threshold in pink. The green line is still the LOD, the blue line is the calculated LOQ, 100 copies.

Discussion

Shad and herring assay performance and LOD/LOQ

Both assays developed for the American and hickory shad had strong target amplification in the lab and in the field. Samples taken from fin clips showed some cross-species amplification at late cycles for both assays during initial laboratory testing. For the American shad assay, cross-amplification was observed with hickory shad fin clip DNA (Fig. 3.2A). While not ideal, this is not an entirely surprising result as these species are closely related, and an assessment of restriction site differences in the mitochondrial genomes of Alosa species showed American shad had a sequence divergence 6.1% with hickory shad (Bentzen et al., 1993). Within the target gene, cytochrome oxidase one, the sequences used to develop the assay had an even lower divergence percentage of 4.7%. The hickory shad assay amplified American shad and alewife samples (Fig.

3.2B). Alewife samples only amplified at >40 cycles or not within 45 cycles (NAs, see Fig. 3.2B), which are typically considered non-detections (e.g., for river herring: Plough et al., 2018). American shad and alewife had the lowest number of mismatched bases within the hickory shad amplicon which can lead to nontarget amplification. This cross-species amplification was addressed when qPCR conditions were optimized and was not a major concern with environmental samples as fin clip extracts contain concentrations of DNA far higher than that typically recovered in eDNA samples.

All three assays were highly efficient and had low limit of detections and limit of quantifications. Klymus et al., (2020) generated LOD and LOQ for 36 assays from seven independent laboratories and had discrete LODs that ranged from 4-192 copies per reaction, and discrete LOQs that ranged from 10-1920 copies per reaction. The LOD and LOQs generated for the assays reported here are on the low end of these ranges. Low LOD and LOQs are preferable as eDNA is highly fragmented and can be present at very low concentrations in the natural environment (Thomsen & Willerslev, 2015). Low LOD and LOQs indicate low concentrations of our target species DNA can be detected with accuracy and precision, respectively. These assays were also highly efficient and well within the recommended qPCR efficiency limit of ≥90% (Svec et al., 2015).

Field results for American and hickory shads

American and hickory shad assays both resulted in positive field detections (Fig. 3.3, 3.4, 3.7). The American shad had very limited detections in the Chesapeake Bay with the strongest

detections on the Potomac and Patapsco Rivers (Fig. 3.3). The Potomac River has a relatively large population of American shad and has been used to stock rivers in Maryland, Delaware, and Virginia since 1995 (AMSFC, 2020). Additionally, samples taken in the Potomac were collected during the spring when American shad were expected to be in the system. The only strong positive detections in the Patapsco River were collected in 2021 and are beyond Bloede Dam, a site that underwent a dam removal in 2018 and showed signs of river herring recolonization soon after removal (Huang et al., 2023). There were some sporadic single replicate detections in the Northwest and Christiana Rivers. Typically, at least 2/3 positive replicates are required to confirm a positive detection (e.g., for river herring; Plough et al. 2018) so it's possible these detections are false positives and additional testing would be required to make definitive conclusions on the validity of these detections. Due to the low number of American shad detections in Chesapeake Bay, additional samples were collected from the Cape Fear River in North Carolina, 19/25 of which had strong detections (76%, Fig. 3.4). Most of these samples were collected from March to May, the end of the American shad spawning season in this part of their range (Limburg et al., 2003). Interestingly, strong eDNA detections were found at the upstream Cape Fear River sampling locations into May (Fig. 3.4), when the species is supposed to be done spawning. This could be due to the persistence of juveniles in those parts of the river after the spawning migration since juveniles tend to remain in the lower parts of riverine and estuarine systems before migrating out to sea (Limburg et al., 2003). It's also worth noting that Cape Fear River is the northern boundary for the semelparous metapopulation (AMSFC, 2020), so perhaps the stronger detections than those found in the Chesapeake Bay could be due to the

differential shedding rates between alive and dead individuals as previously documented in Pacific salmon (Tillotson et al., 2018).

In the Mid-Atlantic, hickory shad are monitored using traditional methods opportunistically during American shad surveys (AMSFC, 2020) and have large gaps in information due to their low abundances. They've also been observed actively avoiding fyke and pound nets typically used for surveying by Maryland Department of Natural Resources which could artificially reduce estimates of abundance (Bi et al., 2021). eDNA detections for hickory shad were mostly found on Maryland's Eastern shore with a few strong positive detections in the Patapsco and Potomac Rivers on the western shore (Fig. 3.7). On the eastern shore, hickory shad were detected in the northern parts of Delaware (Fig. 3.7) on the Christiana and Northwest Rivers. These rivers aren't currently assessed for American or hickory shad and most of the Upper Chesapeake does not have reported observations of hickory shad, according to the most recent Atlantic Marine States Fisheries Commission's Benchmark Stock Assessment (2020). Samples from the Choptank River also had strong detections in 2015 (Fig. 3.7). Stocking activities for the hickory shad took place in the Choptank River from 2001-2014 and a report from 2014 stated the stocking efforts had established self-sufficient hickory shad populations (Stence et al., 2014). While the data collected here cannot be used to make any robust conclusions about the outcome of those stocking efforts, it does show how eDNA could be used to continue to monitor those populations in the Choptank. These recent (samples collected in 2022) and strong positive detections suggest that more investigation into American and hickory shad populations with eDNA or other

methods is warranted and could advance the knowledge on hickory shad distribution in Maryland and Delaware.

Conclusion

In this study, we developed, optimized, and validated eDNA assays for the American and hickory shad species. We developed metrics of assay performance for the new assays and the existing river herring assay, that showed all three assays were ready for implementation in eDNA field testing. American and hickory shad assays were used in field testing of samples from 2015-2023 that highlighted the capabilities of eDNA monitoring studies to advance conservation goals for these data poor species. This work completes a panel of eDNA monitoring tools for the Mid-Atlantic alosine fishes that can be used in future studies to generate more information on these species in a cost-efficient and non-invasive manner and make progress towards overarching conservation and management goals.

Works Cited

Aho, K., Derryberry, D., & Peterson, T. (2014). *Model selection for ecologists: The worldviews of AIC and BIC*. *95*(3).

Anderson, J. T., Schumer, G., Anders, P. J., Horvath, K., & Merz, J. E. (2018). Confirmed Observation: A North American Green Sturgeon *Acipenser medirostris* Recorded in the Stanislaus River, California. *Journal of Fish and Wildlife Management*, *9*(2), 624–630. https://doi.org/10.3996/012018-JFWM-006

Andruszkiewicz Allan, E., Zhang, W. G., Lavery, A., & Govindarajan, A. (2021). Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. *Environmental DNA*, *3*(2), 492–514. https://doi.org/10.1002/edn3.141

Atlantic States Marine Fisheries Commission. (2020). 2020 American Shad Benchmark Stock Assessment.

Atlantic States Marine Fisheries Commission. (2012). River Herring Benchmark Stock Assessment Volume I: Coastwide Summary.

Avery, O. T., MacLeod, C. M., McCarty, M. (1944). *Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types.* The Journal of experimental medicine, 79(2), 137–158. https://doi.org/10.1084/jem.79.2.137

Baetscher, D. S., Hasselman, D. J., Reid, K., Palkovacs, E. P., & Garza, J. C. (2017). Discovery and characterization of single nucleotide polymorphisms in two anadromous alosine fishes of conservation concern. *Ecology and Evolution*, *7*(17), 6638–6648. https://doi.org/10.1002/ece3.3215

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, *17*(1), 1–17. https://doi.org/10.1007/s10592-015-0775-4

Bi, R., Jiao, Y., Weaver, L. A., Greenlee, B., McClair, G., Kipp, J., Wilke, K., Haas, C., & Smith, E. (2021). Environmental and anthropogenic influences on spatiotemporal dynamics of *Alosa* in Chesapeake Bay tributaries. *Ecosphere*, *12*(6). https://doi.org/10.1002/ecs2.3544

Bohlin, T., Hamrin, S., Heggberget, T. G., Rasmussen, G., & Saltveit, S. J. (1989). Electrofishing—Theory and practice with special emphasis on salmonids. *Hydrobiologia*, *173*(1), 9–43. https://doi.org/10.1007/BF00008596

Budowle, B., & Van Daal, A. (2009). Extracting evidence from forensic DNA analyses: Future molecular biology directions. *BioTechniques*, *46*(5), 339–350. https://doi.org/10.2144/000113136

Canty, A., and Ripley, B. (2022). boot: Bootstrap R (S-Plus) Functions. R package version 1.3-28.1.

Carraro, L., Stauffer, J. B., & Altermatt, F. (2021). How to design optimal eDNA sampling strategies for biomonitoring in river networks. *Environmental DNA*, *3*(1), 157–172. https://doi.org/10.1002/edn3.137

Coulter, D. P., Wang, P., Coulter, A. A., Van Susteren, G. E., Eichmiller, J. J., Garvey, J. E., & Sorensen, P. W. (2019). Nonlinear relationship between Silver Carp density and their eDNA concentration in a large river. *PLOS ONE, 14*(6), e0218823. https://doi.org/10.1371/journal.pone.0218823

Criddle, K. R., & Shimizu, I. (2014). The Economic Importance of Wild Pacific Salmon. In P. T. K. Woo. & D. J. Noakes (Eds), *Salmon* (pp. 269-306). Nova Science Publishers, Inc.

Curtis, A. N., Tiemann, J. S., Douglass, S. A., Davis, M. A., & Larson, E. R. (2021). High stream flows dilute environmental DNA (eDNA) concentrations and reduce detectability. *Diversity and Distributions*, *27*(10), 1918–1931. https://doi.org/10.1111/ddi.13196

Cushman, E., Tarpey, C., Post, B., Ware, K., & Darden, T. (2012). Genetic Characterization of American Shad in the Edisto River, South Carolina, and Initial Evaluation of an Experimental Stocking Program. *Transactions of the American Fisheries Society*, *141*(5), 1338–1348. https://doi.org/10.1080/00028487.2012.692347

Davison, A. C. & Hinkley, D. V. (1997) Bootstrap Methods and Their Applications. Cambridge University Press, Cambridge. ISBN 0-521-57391-2

Deiner, K. (2015). Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biological Conservation*.

Eichmiller, J. J., Best, S. E., & Sorensen, P. W. (2016). Effects of Temperature and Trophic State on Degradation of Environmental DNA in Lake Water. *Environmental Science & Technology*, *50*(4), 1859–1867. https://doi.org/10.1021/acs.est.5b05672

Emmert-Streib, F., & Dehmer, M. (2019). Evaluation of Regression Models: Model Assessment, Model Selection and Generalization Error. *Machine Learning and Knowledge Extraction*, *1*(1), 521–551. https://doi.org/10.3390/make1010032

Erickson, R. A., Rees, C. B., Coulter, A. A., Merkes, C. M., McCalla, S. G., Touzinsky, K. F.,

Walleser, L., Goforth, R. R., & Amberg, J. J. (2016). Detecting the movement and spawning activity of bigheaded carps with environmental DNA. *Molecular Ecology Resources*, *16*(4), 957–965. https://doi.org/10.1111/1755-0998.12533

Evans, N. T., Shirey, P. D., Wieringa, J. G., Mahon, A. R., & Lamberti, G. A. (2017). Comparative Cost and Effort of Fish Distribution Detection via Environmental DNA Analysis and Electrofishing. *Fisheries*, *42*(2), 90–99. https://doi.org/10.1080/03632415.2017.1276329

Fediajevaite, J., Priestley, V., Arnold, R., & Savolainen, V. (2021). Meta-analysis shows that environmental DNA outperforms traditional surveys, but warrants better reporting standards. *Ecology and Evolution*, *11*(9), 4803–4815. https://doi.org/10.1002/ece3.7382

Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, *4*(4), 423–425. https://doi.org/10.1098/rsbl.2008.0118

Forootan, A., Sjöback, R., Björkman, J., Sjögreen, B., Linz, L., & Kubista, M. (2017). Methods to determine limit of detection and limit of quantification in quantitative real-time PCR (qPCR). *Biomolecular Detection and Quantification*, *12*, 1–6. https://doi.org/10.1016/j.bdq.2017.04.001

Fuller, W. A. (2009). *Measurement error models*. John Wiley & Sons.

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

Hall, C. J., Jordaan, A., & Frisk, M. G. (2012). Centuries of Anadromous Forage Fish Loss: Consequences for Ecosystem Connectivity and Productivity. *BioScience*, *62*(8), 723–731. https://doi.org/10.1525/bio.2012.62.8.5

Hare, J. A., Borggaard, D. L., Alexander, M. A., Bailey, M. M., Bowden, A. A., Damon-Randall, K., Didden, J. T., Hasselman, D. J., Kerns, T., McCrary, R., McDermott, S., Nye, J. A., Pierce, J., Schultz, E. T., Scott, J. D., Starks, C., Sullivan, K., & Beth Tooley, M. (2021). A Review of River Herring Science in Support of Species Conservation and Ecosystem Restoration. *Marine and Coastal Fisheries*, *13*(6), 627–664. https://doi.org/10.1002/mcf2.10174

Harris, M., Brodeur, N., LeBlanc, F., Douglas, S., Chamberland, P., Guyondet, T., Steeves, R., & Gagné, N. (2022). EDNA and Acoustic Tag Monitoring Reveal Congruent Overwintering Distributions of Striped Bass in a Hydrologically Complex Estuarine Environment. *Fishes*, *7*(4), 183. https://doi.org/10.3390/fishes7040183

Harrison, J. B., Sunday, J. M., & Rogers, S. M. (2019). Predicting the fate of eDNA in the environment and implications for studying biodiversity. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1915), 20191409. https://doi.org/10.1098/rspb.2019.1409

Hasselman, D. J., Anderson, E. C., Argo, E. E., Bethoney, N. D., Gephard, S. R., Post, D. M., Schondelmeier, B. P., Schultz, T. F., Willis, T. V., & Palkovacs, E. P. (2016). Genetic stock composition of marine bycatch reveals disproportional impacts on depleted river herring genetic stocks. *Canadian Journal of Fisheries and Aquatic Sciences*, *73*(6), 951–963. https://doi.org/10.1139/cjfas-2015-0402

Hill, C. R., Rulifson, R. A., & Halden, N. M. (2022). A quantitative approach to investigate natal philopatry in Hickory Shad (*Alosa mediocris*) using otolith chemistry. *Fisheries Research*, *249*, 106232. https://doi.org/10.1016/j.fishres.2022.106232

Holmes, J. A., Cronkite, G. M. W., Enzenhofer, H. J., & Mulligan, T. J. (2006). Accuracy and precision of fish-count data from a "dual-frequency identification sonar" (DIDSON) imaging system. *ICES Journal of Marine Science*, *63*(3), 543–555. https://doi.org/10.1016/j.icesjms.2005.08.015

Hosen, J. D., Armstrong, A. W., & Palmer, M. A. (2018). Dissolved organic matter variations in coastal plain wetland watersheds: The integrated role of hydrological connectivity, land use, and seasonality. *Hydrological Processes*, *32*(11), 1664–1681. https://doi.org/10.1002/hyp.11519

Huang, C. S., Legett, H. D., Plough, L. V., Aguilar, R., Fitzgerald, C., Gregory, B., Heggie, K., Lee, B., Richie, K. D., Harbold, W., & Ogburn, M. B. (2023). Early detection and recovery of river herring spawning habitat use in response to a mainstem dam removal. *PLOS ONE*, *18*(5), e0284561. https://doi.org/10.1371/journal.pone.0284561

Jo, T., Murakami, H., Yamamoto, S., Masuda, R., & Minamoto, T. (2019). Effect of water temperature and fish biomass on environmental DNA shedding, degradation, and size distribution. *Ecology and Evolution*, *9*(3), 1135–1146. https://doi.org/10.1002/ece3.4802

Joe, H., & Zhu, R. (2005). Generalized Poisson Distribution: The Property of Mixture of Poisson and Comparison with Negative Binomial Distribution. *Biometrical Journal*, *47*(2), 219–229. https://doi.org/10.1002/bimj.200410102

Karlsson, E., Ogonowski, M., Sundblad, G., Sundin, J., Svensson, O., Nousiainen, I., & Vasemägi, A. (2022). Strong positive relationships between eDNA concentrations and biomass in juvenile and adult pike (*Esox lucius*) under controlled conditions: Implications for monitoring. *Environmental DNA*, *4*(4), 881–893. https://doi.org/10.1002/edn3.298

Klymus, K. E., Merkes, C. M., Allison, M. J., Goldberg, C. S., Helbing, C. C., Hunter, M. E., Jackson, C. A., Lance, R. F., Mangan, A. M., Monroe, E. M., Piaggio, A. J., Stokdyk, J. P., Wilson, C. C., & Richter, C. A. (2020). Reporting the limits of detection and quantification for environmental DNA assays. *Environmental DNA*, *2*(3), 271–282. https://doi.org/10.1002/edn3.29

Lacoursière-Roussel, A., & Deiner, K. (2021). Environmental DNA is not the tool by itself. *Journal of Fish Biology*, *98*(2), 383–386. https://doi.org/10.1111/jfb.14177

LaMotte, L. R., & Wells, J. D. (2016). Inverse prediction for multivariate mixed models with standard software. *Statistical Papers*, *57*(4), 929–938. https://doi.org/10.1007/s00362-016-0815-2

Laramie, M. B., Pilliod, D. S., & Goldberg, C. S. (2015). Characterizing the distribution of an endangered salmonid using environmental DNA analysis. *Biological Conservation*, *183*, 29–37. https://doi.org/10.1016/j.biocon.2014.11.025

Larson, E. R., Renshaw, M. A., Gantz, C. A., Umek, J., Chandra, S., Lodge, D. M., & Egan, S. P. (2017). Environmental DNA (eDNA) detects the invasive crayfishes Orconectes rusticus and Pacifastacus leniusculus in large lakes of North America. *Hydrobiologia*, *800*(1), 173–185. https://doi.org/10.1007/s10750-017-3210-7

Levi, T., Allen, J. M., Bell, D., Joyce, J., Russell, J. R., Tallmon, D. A., Vulstek, S. C., Yang, C., & Yu, D. W. (2019). Environmental DNA for the enumeration and management of Pacific salmon. *Molecular Ecology Resources*, *19*(3), 597–608. https://doi.org/10.1111/1755-0998.12987

Limburg, K. E., Hattala, K. A., & Kahnle, A. (2003). *American Shad in Its Native Range*.

Limburg, K. E., & Waldman, J. R. (2009). Dramatic Declines in North Atlantic Diadromous Fishes. *BioScience*, *59*(11), 955–965. https://doi.org/10.1525/bio.2009.59.11.7

Lodge, D. M., Turner, C. R., Jerde, C. L., Barnes, M. A., Chadderton, L., Egan, S. P., Feder, J. L., Mahon, A. R., & Pfrender, M. E. (2012). Conservation in a cup of water: Estimating biodiversity and population abundance from environmental DNA. *Molecular Ecology*, *21*(11), 2555–2558. https://doi.org/10.1111/j.1365-294X.2012.05600.x

Lorenz, T. C. (2012). Polymerase Chain Reaction: Basic Protocol Plus Troubleshooting and Optimization Strategies. *Journal of Visualized Experiments*, *63*, 3998. https://doi.org/10.3791/3998

Magowan, K., Reitsma, J., & Murphy, D. (2012). Use of Dual-Frequency Identification Sonar to Monitor Adult River Herring in a Small Coastal Stream. *Marine and Coastal Fisheries*, *4*(1), 651–659. https://doi.org/10.1080/19425120.2012.730916

Mansueti, R. J. (1962). Eggs, Larvae, and Young of the Hickory Shad, *Alosa mediocris*, with Comments on Its Ecology in the Estuary. *Chesapeake Science*, *3*(3), 173. https://doi.org/10.2307/1350995

Maruyama, A., Nakamura, K., Yamanaka, H., Kondoh, M., & Minamoto, T. (2019). Correction: The Release Rate of Environmental DNA from Juvenile and Adult Fish. *PLOS ONE*, *14*(2), e0212145. https://doi.org/10.1371/journal.pone.0212145

Mauvisseau, Q., Harper, L. R., Sander, M., Hanner, R. H., Kleyer, H., & Deiner, K. (2022). The Multiple States of Environmental DNA and What Is Known about Their Persistence in Aquatic Environments. *Environmental Science & Technology*, *56*(9), 5322–5333. https://doi.org/10.1021/acs.est.1c07638

Meehan, W. E. (1893). *Fish, fishing and fisheries of Pennsylvania*. E. K. Meyers. https://doi.org/10.5962/bhl.title.54175

Murauskas, J. G., & Rulifson, R. A. (2011). Reproductive Development and Related Observations during the Spawning Migration of Hickory Shad. *Transactions of the American Fisheries Society*, *140*(4), 1035–1048. https://doi.org/10.1080/00028487.2011.607036

National Oceanic and Atmospheric Science Administration. (2019) Status Review Report: Alewife (*Alosa pseduoharengus*), Blueback Herring (*Alosa aestivalis*)

Nukazawa, K., Hamasuna, Y., & Suzuki, Y. (2018). Simulating the Advection and Degradation of the Environmental DNA of Common Carp along a River. *Environmental Science & Technology*, *52*(18), 10562–10570. https://doi.org/10.1021/acs.est.8b02293

Ogburn, M. B., Plough, L. V., Bangley, C. W., Fitzgerald, C. L., Hannam, M. P., Lee, B., Marafino, G., Richie, K. D., Williams, M. R., & Weller, D. E. (2023). Environmental DNA reveals anadromous river herring habitat use and recolonization after restoration of aquatic connectivity. *Environmental DNA*, *5*(1), 25–37. https://doi.org/10.1002/edn3.348

Ogburn, M. B., Spires, J., Aguilar, R., Goodison, M. R., Heggie, K., Kinnebrew, E., McBurney, W., Richie, K. D., Roberts, P. M., & Hines, A. H. (2017a). Assessment of River Herring

Spawning Runs in a Chesapeake Bay Coastal Plain Stream using Imaging Sonar. *Transactions of the American Fisheries Society*, *146*(1), 22–35. https://doi.org/10.1080/00028487.2016.1235612

Ogburn, M. B., Spires, J., Aguilar, R., Goodison, M. R., Heggie, K., Kinnebrew, E., McBurney, W., Richie, K. D., Roberts, P. M., & Hines, A. H. (2017b). Assessment of River Herring Spawning Runs in a Chesapeake Bay Coastal Plain Stream using Imaging Sonar. *Transactions of the American Fisheries Society*, *146*(1), 22–35. https://doi.org/10.1080/00028487.2016.1235612

Ogram, A., Sayler, G. S., & Barkay, T. (1987). The extraction and purification of microbial DNA from sediments. *Journal of Microbiological Methods*, *7*(2), 57–66. https://doi.org/10.1016/0167-7012(87)90025-X

Pfleger, M. O., Rider, S. J., Johnston, C. E., & Janosik, A. M. (2016). Saving the doomed: Using eDNA to aid in detection of rare sturgeon for conservation (Acipenseridae). *Global Ecology and Conservation*, *8*, 99–107. https://doi.org/10.1016/j.gecco.2016.08.008

Plough, L. V., Bunch, A. J., Lee, B. B., Fitzgerald, C. L., Stence, C. P., & Richardson, B. (2021). Development and testing of an environmental DNA (eDNA) assay for endangered Atlantic sturgeon to assess its potential as a monitoring and management tool. *Environmental DNA*, *3*(4), 800–814. https://doi.org/10.1002/edn3.186

Plough, L. V., Ogburn, M. B., Fitzgerald, C. L., Geranio, R., Marafino, G. A., & Richie, K. D. (2018). Environmental DNA analysis of river herring in Chesapeake Bay: A powerful tool for monitoring threatened keystone species. *PLoS ONE*, *13*(11), e0205578. https://doi.org/10.1371/journal.pone.0205578

R Core Team (2023). R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.

Renshaw, M. A., Olds, B. P., Jerde, C. L., McVeigh, M. M., & Lodge, D. M. (2015). The room temperature preservation of filtered environmental DNA samples and assimilation into a phenol–chloroform–isoamyl alcohol DNA extraction. *Molecular Ecology Resources*, *15*(1), 168–176. https://doi.org/10.1111/1755-0998.12281

Rigby R.A. and Stasinopoulos D.M. (2005). Generalized additive models for location, scale and shape, (with discussion), Appl. Statist., 54, part 3, pp 507-554.

Rillahan, C. B., Alcott, D., Castro-Santos, T., & He, P. (2021). Activity Patterns of Anadromous Fish below a Tide Gate: Observations from High-Resolution Imaging Sonar. *Marine and Coastal Fisheries*, *13*(3), 200–212. https://doi.org/10.1002/mcf2.10149

Rourke, M. L., Fowler, A. M., Hughes, J. M., Broadhurst, M. K., DiBattista, J. D., Fielder, S., Wilkes Walburn, J., & Furlan, E. M. (2022). Environmental DNA (eDNA) as a tool for assessing fish biomass: A review of approaches and future considerations for resource surveys. *Environmental DNA*, *4*(1), 9–33. https://doi.org/10.1002/edn3.185

Sassoubre, L. M., Yamahara, K. M., Gardner, L. D., Block, B. A., & Boehm, A. B. (2016). Quantification of Environmental DNA (eDNA) Shedding and Decay Rates for Three Marine Fish. *Environmental Science & Technology*, *50*(19), 10456–10464. https://doi.org/10.1021/acs.est.6b03114

Saunders, R., Hachey, M. A., & Fay, C. W. (2006). *Maine's Diadromous Fish Community: Past, Present, and Implications for Atlantic Salmon Recovery*. *31*(11).

Searcy, R. T., Boehm, A. B., Weinstock, C., Preston, C. M., Jensen, S., Roman, B., Birch, J. M., Scholin, C. A., Van Houtan, K. S., Kiernan, J. D., & Yamahara, K. M. (2022). High-frequency and long-term observations of eDNA from imperiled salmonids in a coastal stream: Temporal dynamics, relationships with environmental factors, and comparisons with conventional observations. *Environmental DNA*, *4*(4), 776–789. https://doi.org/10.1002/edn3.293

Shelton, A. O., Kelly, R. P., O'Donnell, J. L., Park, L., Schwenke, P., Greene, C., Henderson, R. A., & Beamer, E. M. (2019). Environmental DNA provides quantitative estimates of a threatened salmon species. *Biological Conservation*, *237*, 383–391. https://doi.org/10.1016/j.biocon.2019.07.003

Shu, L., Ludwig, A., & Peng, Z. (2020). Standards for Methods Utilizing Environmental DNA for Detection of Fish Species. *Genes*, *11*(3), 296. https://doi.org/10.3390/genes11030296

Stasinopoulos, D. M., & Rigby, R. A. (2007). Generalized Additive Models for Location Scale and Shape (GAMLSS) in *R*. *Journal of Statistical Software*, *23*(7). https://doi.org/10.18637/jss.v023.i07

Stasinopoulos, M. D., Rigby, R. A., Heller, G. Z., Voudouris, V., & Bastiani, F. D. (2017). *Flexible Regression and Smoothing: Using GAMLSS in R* (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/b21973

Stence, C. P., Baldwin, M. W., Bowermaster, M. T., Porta, M. J., & Barker, L. S. (2014). *Hickory shad restoration in three Maryland Rivers*.

Stoklosa, J., Blakey, R. V., & Hui, F. K. C. (2022). An Overview of Modern Applications of Negative Binomial Modelling in Ecology and Biodiversity. *Diversity*, *14*(5), 320. https://doi.org/10.3390/d14050320

Svec, D., Tichopad, A., Novosadova, V., Pfaffl, M. W., & Kubista, M. (2015). How good is a PCR efficiency estimate: Recommendations for precise and robust qPCR efficiency assessments. *Biomolecular Detection and Quantification*, *3*, 9–16. https://doi.org/10.1016/j.bdq.2015.01.005

Takahara, T., Minamoto, T., & Doi, H. (2015). Effects of sample processing on the detection rate of environmental DNA from the Common Carp (Cyprinus carpio). *Biological Conservation*, *183*, 64–69. https://doi.org/10.1016/j.biocon.2014.11.014

Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., & Kawabata, Z. (2012). Estimation of Fish Biomass Using Environmental DNA. *PLoS ONE*, *7*(4), e35868. https://doi.org/10.1371/journal.pone.0035868

Thalinger, B., Deiner, K., Harper, L. R., Rees, H. C., Blackman, R. C., Sint, D., Traugott, M., Goldberg, C. S., & Bruce, K. (2021). A validation scale to determine the readiness of environmental DNA assays for routine species monitoring. *Environmental DNA*, *3*(4), 823–836. https://doi.org/10.1002/edn3.189

Thalinger, B., Wolf, E., Traugott, M., & Wanzenböck, J. (2019). Monitoring spawning migrations of potamodromous fish species via eDNA. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-51398-0

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, *183*, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019

Tillotson, M. D., Kelly, R. P., Duda, J. J., Hoy, M., Kralj, J., & Quinn, T. P. (2018). Concentrations of environmental DNA (eDNA) reflect spawning salmon abundance at fine spatial and temporal scales. *Biological Conservation*, *220*, 1–11. https://doi.org/10.1016/j.biocon.2018.01.030

Travers, A., & Muskhelishvili, G. (2015). DNA structure and function. *FEBS Journal*, *282*(12), 2279–2295. https://doi.org/10.1111/febs.13307

Tsuji, S., Takahara, T., Doi, H., Shibata, N., & Yamanaka, H. (2019). The detection of aquatic macroorganisms using environmental DNA analysis—A review of methods for collection, extraction, and detection. *Environmental DNA*, *1*(2), 99–108. https://doi.org/10.1002/edn3.21

Tsuji, S., & Shibata, N. (2023). *Identification of fish spawning based on size-selected sampling and diurnal concentration changes of environmental DNA resulting from sperm-release* [Preprint]. Molecular Biology. https://doi.org/10.1101/2023.06.19.545639

Turner, C. R., Miller, D. J., Coyne, K. J., & Corush, J. (2014). Improved Methods for Capture, Extraction, and Quantitative Assay of Environmental DNA from Asian Bigheaded Carp (Hypophthalmichthys spp.). *PLoS ONE*, *9*(12), e114329. https://doi.org/10.1371/journal.pone.0114329

Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., Wu, D., Paulsen, I., Nelson, K. E., Nelson, W., Fouts, D. E., Levy, S., Knap, A. H., Lomas, M. W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., … Smith, H. O. (2004). Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science*, *304*(5667), 66–74. https://doi.org/10.1126/science.1093857

Watson, D., & Crick, F. H. C. (1953). Genetical Implications of the Structure of Deoxyribonucleic Acid. *Nature*, 171(4361), 964-967.

West, D. C., Walters, A. W., Gephard, S., & Post, D. M. (2010). Nutrient loading by anadromous alewife (*Alosa pseudoharengus*): Contemporary patterns and predictions for restoration efforts. *Canadian Journal of Fisheries and Aquatic Sciences*, *67*(8), 1211–1220. https://doi.org/10.1139/F10-059

Xie, Y., & Martens, F. J. (2014). An Empirical Approach for Estimating the Precision of Hydroacoustic Fish Counts by Systematic Hourly Sampling. *North American Journal of Fisheries Management*, *34*(3), 535–545. https://doi.org/10.1080/02755947.2014.892546

Yamanaka, H., Minamoto, T., Matsuura, J., Sakurai, S., Tsuji, S., Motozawa, H., Hongo, M., Sogo, Y., Kakimi, N., Teramura, I., Sugita, M., Baba, M., & Kondo, A. (2017). A simple method for preserving environmental DNA in water samples at ambient temperature by addition of cationic surfactant. *Limnology*, *18*(2), 233–241. https://doi.org/10.1007/s10201-016-0508-5

Yamanaka, H., Motozawa, H., Tsuji, S., Miyazawa, R. C., Takahara, T., & Minamoto, T. (2016). On-site filtration of water samples for environmental DNA analysis to avoid DNA degradation during transportation. *Ecological Research*, *31*(6), 963–967. https://doi.org/10.1007/s11284-016-1400-9

Yao, M., Zhang, S., Lu, Q., Chen, X., Zhang, S., Kong, Y., & Zhao, J. (2022). Fishing for fish environmental DNA: Ecological applications, methodological considerations, surveying designs, and ways forward. *Molecular Ecology*, *31*(20), 5132–5164. https://doi.org/10.1111/mec.16659

Yates, M. C., Cristescu, M. E., & Derry, A. M. (2021). Integrating physiology and environmental dynamics to operationalize environmental DNA (eDNA) as a means to monitor freshwater macro-organism abundance. *Molecular Ecology*, *30*(24), 6531–6550. https://doi.org/10.1111/mec.16202

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression Models for Count Data in *R. Journal of Statistical Software, 27*(8). https://doi.org/10.18637/jss.v027.i08

Zydlewski, J., Stich, D. S., Roy, S., Bailey, M., Sheehan, T., & Sprankle, K. (2021). What Have We Lost? Modeling Dam Impacts on American Shad Populations Through Their Native Range. *Frontiers in Marine Science*, *8*, 734213. https://doi.org/10.3389/fmars.2021.734213