ENVIRONMENTAL DNA INNOVATIONS FOR CONSERVATION

EDITED BY: Matthew A. Barnes, Katy Klymus and Hiroki Yamanaka
PUBLISHED IN: Frontiers in Ecology and Evolution and
Frontiers in Environmental Science

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88971-975-4
DOI 10.3389/978-2-88971-975-4

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

ENVIRONMENTAL DNA INNOVATIONS FOR CONSERVATION

Topic Editors:
Matthew A. Barnes, Texas Tech University, United States
Katy Klymus, United States Geological Survey (USGS), United States
Hiroki Yamanaka, Ryukoku University, Japan

Citation: Barnes, M. A., Klymus, K., Yamanaka, H., eds. (2021). Environmental DNA Innovations for Conservation. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88971-975-4

Table of Contents

frontiers in Ecology and Evolution

EDITORIAL
published: 11 November 2021
doi: 10.3389/fevo.2021.785077

Editorial: Environmental DNA Innovations for Conservation

Matthew A. Barnes[1]*, *Katy Klymus*[2] *and Hiroki Yamanaka*[3]

[1] Department of Natural Resources Management, Texas Tech University, Lubbock, TX, United States, [2] U.S. Geological Survey, Columbia Environmental Research Center, Columbia, MO, United States, [3] Department of Environmental Solution Technology, Ryukoku University, Otsu, Japan

Keywords: ecology, conservation, innovation, eDNA, foundation, application, calibration, interpretation

Editorial on the Research Topic

Environmental DNA Innovations for Conservation

INTRODUCTION

Environmental DNA (eDNA) analysis refers to the collection of bulk environmental samples such as water, sediment, or air, and studying the genetic remnants that organisms have shed into their environment to gain information about species presence. The earliest efforts to learn about organisms from genetic analysis of environmental samples occurred in the field of microbiology, where many focal organisms cannot be cultured under laboratory conditions (e.g., Ogram et al., 1987). A major innovation occurred in the early 2000s when paleoecologists began to apply similar genetic methods to study ancient communities of extinct multicellular organisms (e.g., Willerslev et al., 2003). The contemporary application of eDNA analysis was realized when Ficetola et al. (2008) demonstrated that informative genetic material was not only retrievable from ancient sediments that had been preserved under dark, cold, and stabilizing conditions, but from contemporary environmental samples as well. Since then, eDNA analysis has burgeoned into a powerful tool for ecological research and management.

This Research Topic represents a collection of studies demonstrating the continued innovation of eDNA methods and applications, especially in the field of conservation. Several distinct trends emerge across this curated collection that make clear that innovation comes in many different forms, which we have envisioned as key "bases" of innovation (**Figure 1**). We believe that keeping these different bases- or possibilities- of innovation in mind will help researchers identify opportunities for continued innovation in the future, and we describe each base in more detail in the sections below as we introduce the contributions to our Research Topic, *Environmental DNA Innovations for Conservation*.

FOUNDATION

Occasionally overlooked during the excitement of the latest, innovative eDNA application is the fact that eDNA itself represents an ecological entity worthy of study in its own right. After an organism sheds genetic material, but before a researcher or manager collects it, eDNA interacts with its surrounding environment in myriad ways that influence its production and accumulation (e.g., Maruyama et al., 2014; Klymus et al., 2015), changes in form and state in the environment (e.g., Jo et al., 2019; Barnes et al., 2021), transport (e.g., Andruszkiewicz et al., 2019; Valentin et al., 2021), and ultimately its fate (e.g., Tsuji et al., 2017; Foucher et al., 2020). Collectively, Barnes and Turner (2016) referred to these dynamic processes and functions as "the ecology of eDNA," and they shape the inferences that researchers and managers can make based on detection (or non-detection)

OPEN ACCESS

Edited and reviewed by:
Orsolya Valkó,
Hungarian Academy of Science, Hungary

***Correspondence:**
Matthew A. Barnes
matthew.a.barnes@ttu.edu

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *28 September 2021*
Accepted: *20 October 2021*
Published: *11 November 2021*

Citation:
Barnes MA, Klymus K and Yamanaka H (2021) Editorial: Environmental DNA Innovations for Conservation. Front. Ecol. Evol. 9:785077. doi: 10.3389/fevo.2021.785077

FIGURE 1 | Environmental DNA innovations can be organized into several major "bases" or themes.

events. For example, if researchers seek to relate eDNA abundance to organismal abundance (e.g., Takahara et al., 2012), knowledge about the production and decay of eDNA as well as its movement in and out of the system is critical. Similarly, understanding the particle size distribution and other aspects of the state of eDNA in the environment is essential for the optimization of capture methods (Turner et al., 2014) and sensitivity.

We use the term "foundation innovations" to describe advancement of our understanding of the ecology of eDNA. Multiple studies provide foundation innovations within this Research Topic. Considering the origin of eDNA, Thalinger et al. examine eDNA shedding rates among several different fish species in an aquarium study and found that fish activity, energy use, and species-specific differences all influenced eDNA production. Multiple studies within this Research Topic address eDNA transport, including Mixumoti et al. and Wood et al., who both describe patterns in eDNA concentration related to local distribution and biomass of fish as well as the effects of dilution and downstream transport in lotic systems. Kasai et al. relate Japanese eel eDNA abundances and local environmental conditions to develop hypotheses about what factors promote recruitment, survival, and growth of the endangered species. Finally, with regards to eDNA degradation, in a meta-analysis of 28 previous eDNA degradation studies, Saito and Doi identify both water temperature and amplicon length as positive relationships with eDNA degradation rate.

APPLICATION

Perhaps the most consistent form of innovation in eDNA applications has been the excitement of confronting eDNA methods with new species targets in novel ecosystems. Environmental DNA methods have been applied to freshwater (Harper et al., 2018) and marine (Gilbey et al., 2021) aquatic samples, various terrestrial substrates (van der Heyde et al., 2020), and even air (Johnson et al., 2019), and similar methods have taken advantage of biotic collection "assistants" such as sponges in marine habitats (Mariani et al., 2019) and carrion flies or leeches for sampling terrestrial mammals (Schnell et al., 2012; Calvignac-Spencer et al., 2013). Moreover, many studies have advanced beyond the initial, single-species presence/absence determinations of early eDNA research toward the combination of eDNA and next-generation sequencing methods and whole-community characterization (Ruppert et al., 2019).

Together, we summarize the advancement of technological and methodological boundaries and the deployment of eDNA methods in novel habitats and with new targets as "application innovations." For example, although aquatic plant species have been relatively understudied compared to other taxonomic groups, in this Research Topic, Drummond et al. and Tsukamoto et al. apply eDNA analysis for the detection of algae, aquatic, and terrestrial plants across multiple study sites. Wilcox et al. perform eDNA analysis on drinking water sites, representing a novel substrate for eDNA analysis and a new addition to the toolbox for sampling rare jaguars and other elusive mammals. Kirse et al. demonstrate that collection and extraction methods as well as target sequence all affected the number of invertebrate species detected in soil samples. Sepulveda et al. demonstrate eDNA collection *via* robotic samplers associated with streamgages for the detection of a common fish species and a rare fish parasite.

CALIBRATION

Comparisons between eDNA and "traditional" survey methods have contributed to the quantification of sensitivity and accuracy of eDNA methods. Such experiments have been conducted in both aquatic (e.g., Evans et al., 2017) and terrestrial (e.g.,

Leempoel et al., 2020) habitats, and they represent a critical link between eDNA applications and existing toolboxes for ecological study and conservation. Indeed, Jerde (2021) identified the accumulation of comparisons and calibrations between eDNA and more traditional methods, as well as efforts to maximize quality assurance and quality control as key building blocks for continued growth in eDNA research and incorporation into management, conservation, and other ecological applications.

This Research Topic includes several studies aimed at locating eDNA within the broader toolbox of biodiversity surveillance methods, and we refer to these as "calibration innovations." For example, in an ambitious review, McElroy et al. synthesize 37 previous studies that compared metabarcoding and traditional methods for estimating fish richness. They conclude that eDNA metabarcoding tends to detect more species than traditional methods in low-diversity systems, whereas traditional methods and metabarcoding perform similarly in highly diverse systems. Govindarajan et al. compare eDNA metabarcoding and net-based sampling of the marine mesopelagic zone, and in this system, net collection recovered more animal taxa than metabarcoding, although this was attributable to the ability to sample relatively larger volumes of water with netting compared to eDNA analysis. Furthermore, eDNA analysis excelled for certain taxa, notably gelatinous animals that are known to be commonly missed or damaged by netting. Comparing methods in a well-studied terrestrial reserve, Meyer et al. demonstrate that eDNA analysis outperformed other tools for describing biodiversity. Finally, Sepulveda et al. provide a thoughtful review of contamination and associated remediation efforts that will be useful across all eDNA studies, regardless of target species, habitat, or future innovations.

INTERPRETATION

The types of innovation discussed so far all provide a foundation for research which pushes the bounds of what we can learn using eDNA analysis. Such research builds on a strong foundational understanding of eDNA; it requires innovative application using state of the art methods that push the tool to new targets and systems; and it must be grounded through comparison with existing technologies. We refer to innovations that advance the potential knowledge gained from eDNA analysis as "interpretation innovations," and several are represented within this Research Topic.

For example, one potential specter to interpretation of eDNA results has long been the possibility of "false positive" detections resulting from the detection of eDNA released by organisms that are no longer present in an ecosystem (e.g., due to migration or other movement or due to having died). Hirohara et al. apply propidium monoazide (PMA), commonly used in microbiology to distinguish live vs. dead single-celled organisms, during a laboratory-based eDNA survey of zebrafish. While this study also represents a methodological advancement and "application innovation," the finding that PMA helped differentiate intact and disrupted zebrafish cells and eDNA detection could help future eDNA studies determine whether eDNA detections are sourced from living or dead sources.

Other interpretation innovations have come from collection and analysis of eDNA samples over time. For example, Inui et al. analyze river water for detection of the amphidromous fish *Plecoglossus altivelis* over time to identify timing and location (including novel locations) of seasonal spawning events. Using repeated eDNA metabarcoding analyses in Puget Sound, Washington, USA, Jacobs-Palmer et al. monitor algal communities, noting particularly the growth of harmful algal bloom (HAB) forming species and the conditions that contribute to bloom events, which could benefit HAB monitoring and management. Johnson et al. provide a critical test of the utility of airborne eDNA monitoring for terrestrial plants and found that airborne eDNA abundances reflected both seasonal patterns in plant ecology as well as acute events on the landscape (i.e., a human restoration effort).

CONCLUSION

The "bases" of innovation (**Figure 1**) that have emerged in this Research Topic can help synthesize connections across the field of eDNA analysis and identify major questions and research priorities to promote continued growth and development of eDNA applications for conservation biology and beyond. No single form of innovation is more important than any other, nor are the forms of innovation exclusive of one another. Instead, like a DNA sequence, larger structure, advancements, and evolution of ideas will occur most rapidly when the bases work together. Indeed, the highest-impact studies will likely embody multiple types of innovation at once. Finally, as the field of eDNA analysis continues to mature, we hope that similar emerging fields such as the detection of RNA ("eRNA" e.g., Marshall et al., 2021) or proteins in the environment may find this framework useful for advancing their own fields, connecting with eDNA knowledge, and improving understanding and application overall.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Andruszkiewicz, E. A., Koseff, J. R., Fringer, O. B., Ouellette, N. T., Lowe, A. B., Edwards, C. A., et al. (2019). Modeling environmental DNA transport in the coastal ocean using Lagrangian particle tracking. *Front. Mar. Sci.* 6:477. doi: 10.3389/fmars.2019.00477

Barnes, M. A., Chadderton, W. L., Jerde, C. L., Mahon, A. R., Turner, C. R., and Lodge, D. M. (2021). Environmental conditions influence eDNA particle size distribution in aquatic systems. *Environ. DNA* 3, 643–653. doi: 10.1002/edn3.160

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

Calvignac-Spencer, S., Merkel, K., Kutzner, N., Kühl, H., Boesch, C., Kappeler, P. M., et al. (2013). Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity. *Mol. Ecol.* 22, 915–924. doi: 10.1111/mec.12183

Evans, N. T., Shirey, P. D., Wieringa, J. G., Mahon, A. R., and Lamberti, G. A. (2017). Comparative cost and effort of fish distribution detection *via* environmental DNA analysis and electrofishing. *Fisheries* 42, 90–99. doi: 10.1080/03632415.2017.1276329

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

Foucher, A., Evrard, O., Ficetola, G. F., Fielly, L., Poulain, J., Giguet-Covex, C., et al. (2020). Persistence of environmental DNA in cultivated soils: implication of this memory effect for reconstructing the dynamics of land use and cover changes. *Sci. Rep.* 10:10502. doi: 10.1038/s41598-020-67452-1

Gilbey, J., Carvalho, G., Castilho, R., Coscia, I., Coulson, M. W., Dahle, G., et al. (2021). Life in a drop: sampling environmental DNA for marine fishery management and ecosystem monitoring. *Mar. Policy* 124:104331. doi: 10.1016/j.marpol.2020.104331

Harper, L. R., Buxton, A. S., Rees, H. C., Bruce, K., Brys, R., Halfmaerten, D., et al. (2018). Prospects and challenges of environmental DNA (eDNA) monitoring in freshwater ponds. *Hydrobiologia* 826, 24–41. doi: 10.1007/s10750-018-3750-5

Jerde, C. L. (2021). Can we manage fisheries with the inherent uncertainty from eDNA? *J. Fish. Biol.* 98, 341–353. doi: 10.1111/jfb.14218

Jo, T., Arimoto, M., Murakami, H., Masuda, R., and Minamoto, T. (2019). Particle size distribution of environmental DNA from the nuclei of marine fish. *Environ. Sci. Technol.* 53, 9947–9956. doi: 10.1021/acs.est.9b02833

Johnson, M. D., Cox, R. D., and Barnes, M. A. (2019). Analyzing airborne environmental DNA: a comparison of extraction methods, primer type, and trap type on the ability to detect airborne eDNA from terrestrial plant communities. *Environ. DNA.* 1, 176–185. doi: 10.1002/edn3.19

Klymus, K. E., Richter, C. A., Chapman, D. C., and Paukert, K. (2015). Quantification of eDNA shedding rates from invasive bighead carp *Hypopthalmichthys nobilis* and silver carp *Hypophthalmichthys molitrix*. *Biol. Conserv.* 183, 77–84. doi: 10.1016/j.biocon.2014.11.020

Leempoel, K., Hebert, T., and Hadly, E. A. (2020). A comparison of eDNA to camera trapping for assessment of terrestrial mammal diversity. *Proc. R. Soc. B* 287:20192353. doi: 10.1098/rspb.2019.2353

Mariani, S., Baillie, C., Colosimo, G., and Riesgo, A. (2019). Sponges as natural environmental DNA samplers. *Curr. Biol.* 29, R401–R402. doi: 10.1016/j.cub.2019.04.031

Marshall, N. T., Vanderploeg, H. A., and Chaganti, S. R. (2021). Environmental (e)RNA advances the reliability of eDNA by predicting its age. *Sci. Rep.* 11:2769. doi: 10.1038/s41598-021-82205-4

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

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

Ruppert, K. M., Kline, R. J., and Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: a systematic review in methods, monitoring, and applications of global eDNA. *Glob. Ecol. Cons.* 17:e00547. doi: 10.1016/j.gecco.2019.e00547

Schnell, I. B., Thomsen, P. F., Wilkinson, N., Rasmussen, M., Jensen, L. R. D., Willerslev, E., et al. (2012). Screening mammal biodiversity using DNA from leeches. *Curr. Biol.* 22, R262–R263. doi: 10.1016/j.cub.2012.02.058

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

Tsuji, S., Ushio, M., Sakurai, S., Minamoto, T., and Yamanaka, H. (2017). Water temperature-dependent degradation of environmental DNA and its relation to bacterial abundance. *PLoS ONE* 12:e0176608. doi: 10.1371/journal.pone.0176608

Turner, C. R., Barnes, M. A., Xu, C. C. Y., Jones, S. E., Jerde, C. L., and Lodge, D. M. (2014). Particle size distribution and optimal capture of aqueous macrobial environmental DNA. *Methods Ecol. Evol.* 5, 676–684. doi: 10.1111/2041-210X.12206

Valentin, R. E., Kyle, K. E., Allen, M. C., Welbourne, D. J., and Lockwood, J. L. (2021). The state, transport, and fate of aboveground terrestrial arthropod eDNA. *Environ. DNA.* 1–12. doi: 10.1002/edn3.229

van der Heyde, M., Bunce, M., Wardell-Johnson, G., Fernandes, K., White, N. E., and Nevill, P. (2020). Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA metabarcoding. *Mol. Ecol.* 20, 732–745. doi: 10.1111/1755-0998.13148

Willerslev, E., Hansen, A. J., Binladen, J., Brand, T. B., Gilbert, M. T. P., Shapiro, B., et al. (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science* 300, 791–795. doi: 10.1126/science.1084114

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Ecology and Evolution

BRIEF RESEARCH REPORT
published: 28 August 2020
doi: 10.3389/fevo.2020.00276

Calibrating Environmental DNA Metabarcoding to Conventional Surveys for Measuring Fish Species Richness

*Mary E. McElroy[1], Terra L. Dressler[2], Georgia C. Titcomb[2], Emily A. Wilson[3], Kristy Deiner[4], Tom L. Dudley[5], Erika J. Eliason[2], Nathan T. Evans[6], Steven D. Gaines[7], Kevin D. Lafferty[5,8], Gary A. Lamberti[9], Yiyuan Li[10], David M. Lodge[11], Milton S. Love[5], Andrew R. Mahon[12], Michael E. Pfrender[9], Mark A. Renshaw[13], Kimberly A. Selkoe[14] and Christopher L. Jerde[5]**

[1] Interdepartmental Graduate Program in Marine Science, University of California, Santa Barbara, Santa Barbara, CA, United States, [2] Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, Santa Barbara, CA, United States, [3] Biology Department, Bakersfield College, Bakersfield, CA, United States, [4] Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, Zurich, Switzerland, [5] Marine Science Institute, University of California, Santa Barbara, Santa Barbara, CA, United States, [6] U.S. Fish and Wildlife Service, Carterville Fish and Wildlife Conservation Office, Wilmington, IL, United States, [7] Bren School of Environmental Science and Management, University of California, Santa Barbara, Santa Barbara, CA, United States, [8] Western Ecological Research Center, United States Geological Survey at Marine Science Institute, University of California, Santa Barbara, Santa Barbara, CA, United States, [9] Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, United States, [10] Department of Integrative Biology, The University of Texas at Austin, Austin, TX, United States, [11] Department of Ecology and Evolutionary Biology, Cornell Atkinson Center for a Sustainable Future, Cornell University, Ithaca, NY, United States, [12] Department of Biology, Central Michigan University, Mount Pleasant, MI, United States, [13] Oceanic Institute, Hawai'i Pacific University, Waimanalo, HI, United States, [14] National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara, Santa Barbara, CA, United States*

OPEN ACCESS

Edited by:
Hiroki Yamanaka, Ryukoku University, Japan

Reviewed by:
Elizabeth Andruszkiewicz Allan, Woods Hole Oceanographic Institution, United States
Satoshi Yamamoto, Kyoto University, Japan

***Correspondence:**
Christopher L. Jerde
cjerde@ucsb.edu

Specialty section:
This article was submitted to Conservation, a section of the journal Frontiers in Ecology and Evolution

Received: *06 May 2020*
Accepted: *03 August 2020*
Published: *28 August 2020*

Citation:
McElroy ME, Dressler TL, Titcomb GC, Wilson EA, Deiner K, Dudley TL, Eliason EJ, Evans NT, Gaines SD, Lafferty KD, Lamberti GA, Li Y, Lodge DM, Love MS, Mahon AR, Pfrender ME, Renshaw MA, Selkoe KA and Jerde CL (2020) Calibrating Environmental DNA Metabarcoding to Conventional Surveys for Measuring Fish Species Richness. Front. Ecol. Evol. 8:276. doi: 10.3389/fevo.2020.00276

The ability to properly identify species present in a landscape is foundational to ecology and essential for natural resource management and conservation. However, many species are often unaccounted for due to ineffective direct capture and visual surveys, especially in aquatic environments. Environmental DNA metabarcoding is an approach that overcomes low detection probabilities and should consequently enhance estimates of biodiversity and its proxy, species richness. Here, we synthesize 37 studies in natural aquatic systems to compare species richness estimates for bony fish between eDNA metabarcoding and conventional methods, such as nets, visual census, and electrofishing. In freshwater systems with fewer than 100 species, we found eDNA metabarcoding detected more species than conventional methods. Using multiple genetic markers further increased species richness estimates with eDNA metabarcoding. For more diverse freshwater systems and across marine systems, eDNA metabarcoding reported similar values of species richness to conventional methods; however, more studies are needed in these environments to better evaluate relative performance. In systems with greater biodiversity, eDNA metabarcoding

will require more populated reference databases, increased sampling effort, and multi-marker assays to ensure robust species richness estimates to further validate the approach. eDNA metabarcoding is reliable and provides a path for broader biodiversity assessments that can outperform conventional methods for estimating species richness.

Keywords: bland-altman analysis, Lin's concordance correlation coefficient, high-throughput sequencing, marine, freshwater, eDNA

INTRODUCTION

High-throughput sequencing (HTS) of macro-organismal DNA from environmental samples is an innovative conservation approach to detect and measure ecological communities (Thomsen and Willerslev, 2015; Deiner et al., 2017). This technique, hereafter referred to as environmental DNA (eDNA) metabarcoding, enhances conventional biodiversity monitoring because it targets a presumably more widespread particle (DNA) than the species itself (Lacoursière-Roussel and Deiner, 2019), which aids in the detection of rare and elusive species (Jerde, 2019). eDNA metabarcoding for macro-organism detection works in various substrates, including freshwater (Olds et al., 2016), seawater (Thomsen et al., 2012), soil (Epp et al., 2012), sediment (Guardiola et al., 2015), and even air and snow (Kraaijeveld et al., 2015; Kinoshita et al., 2019). Similarly, these methods have been used to characterize taxa from across the eukaryotic tree of life – including mammals (Foote et al., 2012; Ushio et al., 2017), amphibians (Lopes et al., 2017; Bálint et al., 2018), bony fishes (Yamamoto et al., 2017), elasmobranchs (Bakker et al., 2017; Boussarie et al., 2018), plants (Yoccoz et al., 2012), and macro-invertebrates (McGee and Eaton, 2015; Lacoursière-Roussel et al., 2018). This broad applicability across taxa and environments makes eDNA metabarcoding a potentially revolutionary biodiversity monitoring tool, but only if it provides reliable, accurate, and efficient assessments of communities on par with, or better than, conventional methods of species detection.

Numerous studies have now compared conventional species detection to that inferred from eDNA metabarcoding. This has provided valuable insight into the relative performance of eDNA metabarcoding covering limited spatial extents or taxonomic diversity (Deiner et al., 2017; Jerde et al., 2019). Agreement between how many and which species are detected has ranged from nearly identical (Olds et al., 2016) to very disparate (Cilleros et al., 2019). However, we currently lack a broad understanding of how eDNA metabarcoding calibrates to conventional surveys across diverse systems and taxa, particularly given differences in organisms' DNA shedding rates, degradation of DNA in variable environments, and fluctuation in eDNA transport (Barnes and Turner, 2016). Beyond appreciating the logistical and financial advantages over conventional methods (Evans et al., 2017b), we must also quantitatively evaluate how eDNA performs as a measure of species richness (Jerde et al., 2019). The need for enhanced biodiversity surveillance for conservation and management has never been more acute. On land and in the oceans, ecological communities are undergoing rapid compositional and geographic shifts (Pecl et al., 2017; Blowes et al., 2019) and are confronted with numerous threats (Tilman et al., 2017; Halpern et al., 2019; Reid et al., 2019), so it is critical that we better understand if eDNA metabarcoding could facilitate broad biodiversity assessment.

Because of the vast methodological differences and limited taxonomic coverage of published studies comparing eDNA metabarcoding to conventional surveys, a global meta-analysis remains difficult. However, bony fishes have been an early and popular focus of eDNA-based approaches (Jerde et al., 2011), and eDNA metabarcoding has been used to measure fish diversity across a broad range of environmental conditions and species richness values (Jerde et al., 2019). One requirement of eDNA metabarcoding is establishing comprehensive genetic reference libraries, consisting of genetic sequences sourced from reliably identified species, to compare with metabarcoding outputs (sequence reads). Several eDNA metabarcoding studies of bony fishes have used multiple gene markers (e.g., CO1, 12S, 16S) from different mitochondrial loci to increase taxonomic coverage (Olds et al., 2016; Evans et al., 2017a). This provides an opportunity to evaluate the impact of multi-marker methods, which should improve measures of species richness with eDNA metabarcoding.

Here, we synthesize peer-reviewed studies that used eDNA metabarcoding and conventional surveys to measure fish species richness and characterize fish community composition in natural aquatic systems. We test the agreement between the methods to evaluate how eDNA metabarcoding performs relative to conventional surveys as a measure of species richness. We also examine method complementarity by linking fish species identities to the method of detection (eDNA or conventional). Finally, we assess how complementarity in species detection differ across diverse aquatic systems and between single and multi-marker eDNA metabarcoding studies.

MATERIALS AND METHODS

We conducted a literature search following systematic review practices (Moher et al., 2015). Using *Google Scholar*[1] and *Web of Science*[2], we queried peer-reviewed articles published between 1 January 2008 and 1 April 2020 with the key terms "environmental DNA," "metabarcoding," and "fish." Records from the search results were screened and selected for analysis if the study

[1]https://scholar.google.com/

[2]https://webofknowledge.com

(1) sampled eDNA from water in natural aquatic systems, (2) used an eDNA metabarcoding approach, (3) measured fish species richness with eDNA and (4) compared eDNA-based species richness to species richness measured by conventional fish surveys in the same study area.

For each article retained, we extracted data on study context and methodology as reported or referenced in the main article and **Supplementary Material**. When necessary, authors were contacted to provide additional details. We identified comparative observations of fish species richness from eDNA metabarcoding and conventional surveys for independent study sites. Some articles contained multiple independent sites for species richness evaluation whereas others focused on a single study site. We used the same hydrological units defined in the primary studies to compare fish species richness observations between eDNA and conventional methods. However, to maintain consistency among lentic and lotic systems, we aggregated comparative observations from three studies to single observations of species richness at the river, canal, and lake level (Pont et al., 2018; McDevitt et al., 2019; Doble et al., 2020).

We used Lin's concordance correlation coefficient (CCC) (Lawrence and Lin, 1989) and Bland-Altman analyses to evaluate how well species richness measured by eDNA metabarcoding agreed with conventional surveys (Bland and Altman, 1986). Both approaches are widely used in medicine and engineering to validate new assays and instrumentation against an accepted method. We implemented them here to evaluate performance of eDNA metabarcoding for measuring species richness relative to conventional surveys. In general, CCC values provide a measure of agreement in the species richness estimates whereas Bland-Altman analyses show directionality in performance (i.e., is eDNA metabarcoding detecting more or fewer species than conventional methods?). We assessed overall agreement, as well as context-specific agreement to identify differences in relative performance based on system (freshwater and marine) and metabarcoding approach (single and multi-marker). CCC values were considered significant if 95% confidence intervals (CI) did not capture zero. Bland-Altman analyses allowed us to evaluate differences in species richness estimates as function of increasing site-level diversity. Differences were considered significant if the 95% CI around group means did not overlap with zero. Observations were considered outliers if they fell beyond the 95% CI around two standard deviations from the mean. We evaluated subsets of the data by comparing freshwater versus marine systems and single- versus multi-marker observations. Because freshwater systems were well represented, we also analyzed subsets of lentic and lotic systems (**Supplementary File S1**). All analyses were performed in R (R Core Team, 2020) using agRee (Feng, 2020) and blandr (Datta, 2018) packages.

To examine the extent of species-identity overlap between detection methods, we collated species lists for each study site and recorded the survey method (eDNA, conventional, or both) by which each species was detected. We note in the **Supplementary Material** when authors reported detections to the genus or family level, but we did not include these taxa in the analyses (**Supplementary File S2**). Although false-positive detections (misidentified DNA sequences to species presumably not present in the system or detections emerging from likely contamination) would incorrectly indicate eDNA performed better as a richness indicator, we assumed authors minimized such errors relative to total species richness at each site. We calculated total observed species richness at each site as the sum of fish species detected by (1) eDNA only, (2) conventional survey only, and (3) both methods, then determined the proportions of each mechanism of detection. Finally, we assessed the proportions of shared and method-specific detections between marine and freshwater systems and between single- and multi-marker eDNA metabarcoding.

For all analyses, we included species richness data from conventional surveys conducted alongside eDNA sampling, as well as data from many years of routine or historical conventional fish monitoring. When authors provided data from both historical monitoring and contemporary surveys for the same study site, we used species richness calculated from aggregated contemporary and historical data. Our intent was to capture the most complete picture of fish diversity possible through conventional methods to compare with eDNA metabarcoding. However, this could also provide an unfair measure of comparison due to substantially disproportionate effort between approaches and changes in community richness or species presence through time. As such, we provide additional analyses of observations when eDNA sampling and conventional surveys were conducted concurrently (**Supplementary File S3**). Although we collected data on the types of conventional methods and gene markers used in each study (**Supplementary File S4**), our analyses did not distinguish between multiple conventional survey types or between locus-specific detections when multiple genetic markers were used for eDNA metabarcoding because further partitioning of the data into subsets resulted in reduced power to detect differences.

RESULTS

Overview of Studies

Systematic review of the literature yielded 37 peer-reviewed studies meeting our synthesis criteria (**Supplementary Files S5**, **S6**). The earliest study was published in 2012, but all others were published between 2016 and 2020. Most were conducted in Europe (35%), Asia (24%), and North America (22%), and a few were conducted in Australia (11%), South America (5%), and Africa (3%) (**Figure 1**). Most studies (65%) occurred in freshwater systems, including lentic and lotic environments ranging from ditches, ponds, and small streams to large rivers and lakes in temperate and tropical locations. Studies of marine systems included temperate and tropical estuaries, bays, and coastal oceans. Within these studies, we identified 121 independent sites where authors compared fish species richness between eDNA metabarcoding and conventional surveys. The number of sites across studies was highly uneven – 56% of comparative observations originated from just three publications, all of which sampled freshwater systems (Valentini et al., 2016; Fujii et al., 2019; Li et al., 2019).

FIGURE 1 | Global distribution of studies included in synthesis (n = 37). Red circles indicate freshwater studies (n = 24), and blue circles indicate marine studies (n = 13).

Authors compared eDNA-based fish species richness to a variety of conventional survey methods both within and across studies. Nets were the most common conventional method (used in 22 studies), followed by traps, electrofishing, visual surveys, and angling. Impingement, acoustic, and toxicant-based surveys were also used. Most studies (57%) deployed multiple gears or derived observed species richness from a combination of methods to compare with eDNA metabarcoding (**Figures 2C,D**). Although using multiple gears for conventional surveys was common, using multiple gene markers for eDNA metabarcoding of fish diversity was not. Only 11 studies used a multi-marker approach targeting different mitochondrial gene regions for sequencing (**Figures 2A,B**). Overall, the most commonly used gene locus was 12S rRNA (used in 29 studies), followed by 16S rRNA (n = 9), cytochrome *b* (CytB, n = 9), and cytochrome-*c*-oxidase subunit I (COI, n = 4). Of the multi-marker studies, the most common combination was 12S and CytB (n = 4). For single-marker studies, the most frequently used locus was 12S (n = 18). Other components of eDNA metabarcoding workflows were inconsistently reported across studies (**Supplementary File S7**). When reported, water volumes filtered per site ranged widely (0.6 to 3540 L, n = 118), as did filter pore sizes (0.22 to 1.2 μm, n = 106), filter membrane materials, extraction methods, amplicon primers, and numbers of PCR replicates (2 to 12, n = 102).

Across study sites (n = 118), total observed fish species richness regardless of detection method ranged from 0 to 253 and averaged 30.44 ± 3.92 (mean ± *SEM*). On average, marine sites (89.00 ± 15.45, n = 17) were over four times more species-rich than freshwater sites (20.58 ± 2.80, n = 101). Conventional surveys detected more fish species than (i.e., outperformed) eDNA metabarcoding at 50 sites (41%), and eDNA outperformed conventional surveys at 54 sites (45%). Both methods detected the same number of fish species at 17 sites (14%).

Relative Performance of Methods as Species Richness Measures

Lin's CCC showed moderate agreement across all study sites (n = 121, CCC = 0.74, $CI_{95\%}$: 0.66, 0.80), suggesting similar performance of eDNA metabarcoding and conventional surveys as measures of fish species richness. However, notable disparities emerged when we assessed agreement by target system – freshwater systems showed good agreement (CCC = 0.86, $CI_{95\%}$: 0.81, 0.90; n = 104), but marine systems showed no agreement as Lin's CCC was not significantly different from zero (CCC = 0.35, $CI_{95\%}$: −0.04, 0.65; n = 17). Bland-Altman analyses reflected similar patterns (**Figure 3**). Here, performance of eDNA metabarcoding as a richness measure was not significantly different from conventional surveys in both freshwater and marine systems as means for both groups bounded zero (**Figures 3A,B**). Although the differences between methods in marine systems were not considered significant in the latter analysis, we note that wide confidence intervals around the mean and few observations for marine systems limit the insight of this particular result.

Lin's CCC showed agreement for multi-marker observations (CCC = 0.85, $CI_{95\%}$: 0.76, 0.91; n = 37) and for single-marker observations (CCC = 0.72, $CI_{95\%}$: 0.64, 0.79; n = 84). However, Bland-Altman analyses indicated multi-marker eDNA metabarcoding outperformed conventional surveys while single-marker eDNA metabarcoding did not (**Figures 3C,D**). These results highlight the importance of considering not only the agreement between methods with CCC values, but also the direction of performance as demonstrated by Bland-Altman analysis. For freshwater sites, we conducted a *post hoc* evaluation of the Bland-Altman formatted data (**Figure 3A**) using a bent cable model (grid size 30) to identify thresholds implemented by applying the SiZer package in R (Sonderegger et al., 2009). The peak difference in relative performance for eDNA metabarcoding compared to conventional surveys, occurred at a species richness of 30, and the species richness threshold where the bent cable model intersected relative performance parity (eDNA species richness estimate – conventional species richness estimate = 0) was at approximately 100 species. There was insufficient sample size and species richness coverage to implement a similar analysis for marine systems.

Across all subsets of data except for marine systems, CCC values showed agreement between eDNA metabarcoding and conventional surveys (**Supplementary File S8**). Additionally, Bland-Altman analyses indicated eDNA metabarcoding performs as well as conventional surveys in both lentic and lotic environments (**Supplementary File S1**). When conventional surveys are conducted concurrently with eDNA sampling, eDNA metabarcoding performed as well as conventional methods. In contrast, non-concurrent sampling resulted in eDNA metabarcoding slightly underperforming conventional surveys (**Supplementary File S3**).

Method Complementarity Using Species Identities

Conventional surveys and eDNA metabarcoding exhibited a high degree of overlap when accounting for shared species detections within study sites (**Figure 4**). For most sites (65%), the proportion of shared detections was 0.50 or greater (specifically, average overlap was 0.56 ± 0.02, n = 115). At 25 different sites, the degree of overlap was 0.75 or greater, including 7 sites with complete

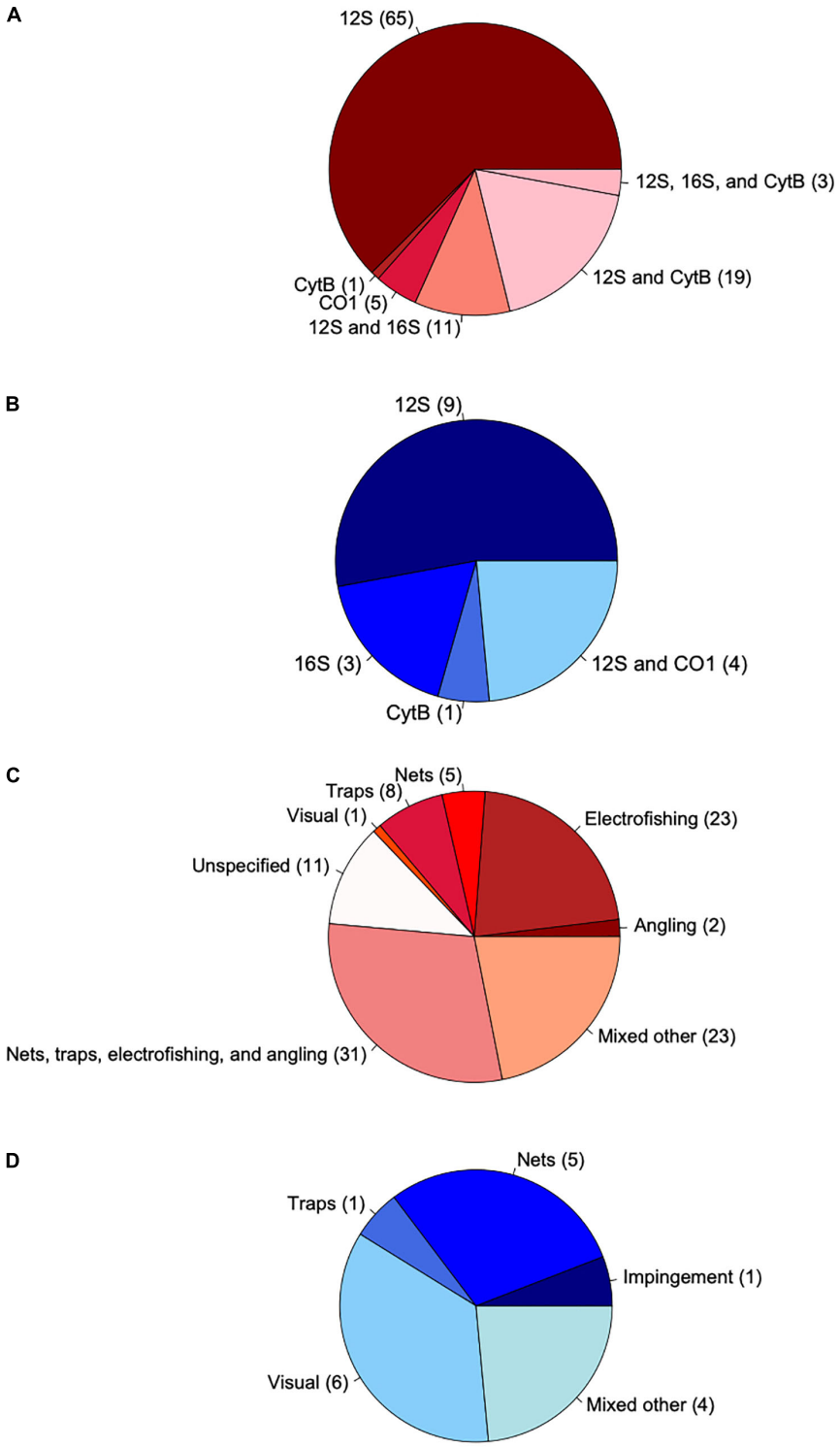


FIGURE 2 | Pie charts characterizing the genetic loci sequenced in eDNA metabarcoding studies and conventional survey types for freshwater (red) and marine systems (blue). Numbers in parentheses show the number of observations (n = 121). **(A)** Loci used in freshwater systems. Dark and light shading indicate single- and multi-marker observations, respectively. **(B)** Loci used in marine systems. Dark and light shading indicate single- and multi-marker observations, respectively. **(C)** Conventional surveys used in freshwater systems. Dark and light shading indicate where single and multiple survey methods were used, respectively. White indicates observations for which primary study authors did not report the type of conventional methods used to create species lists. **(D)** Conventional surveys used in marine systems. Dark and light shading indicate where single and multiple survey methods were used, respectively.

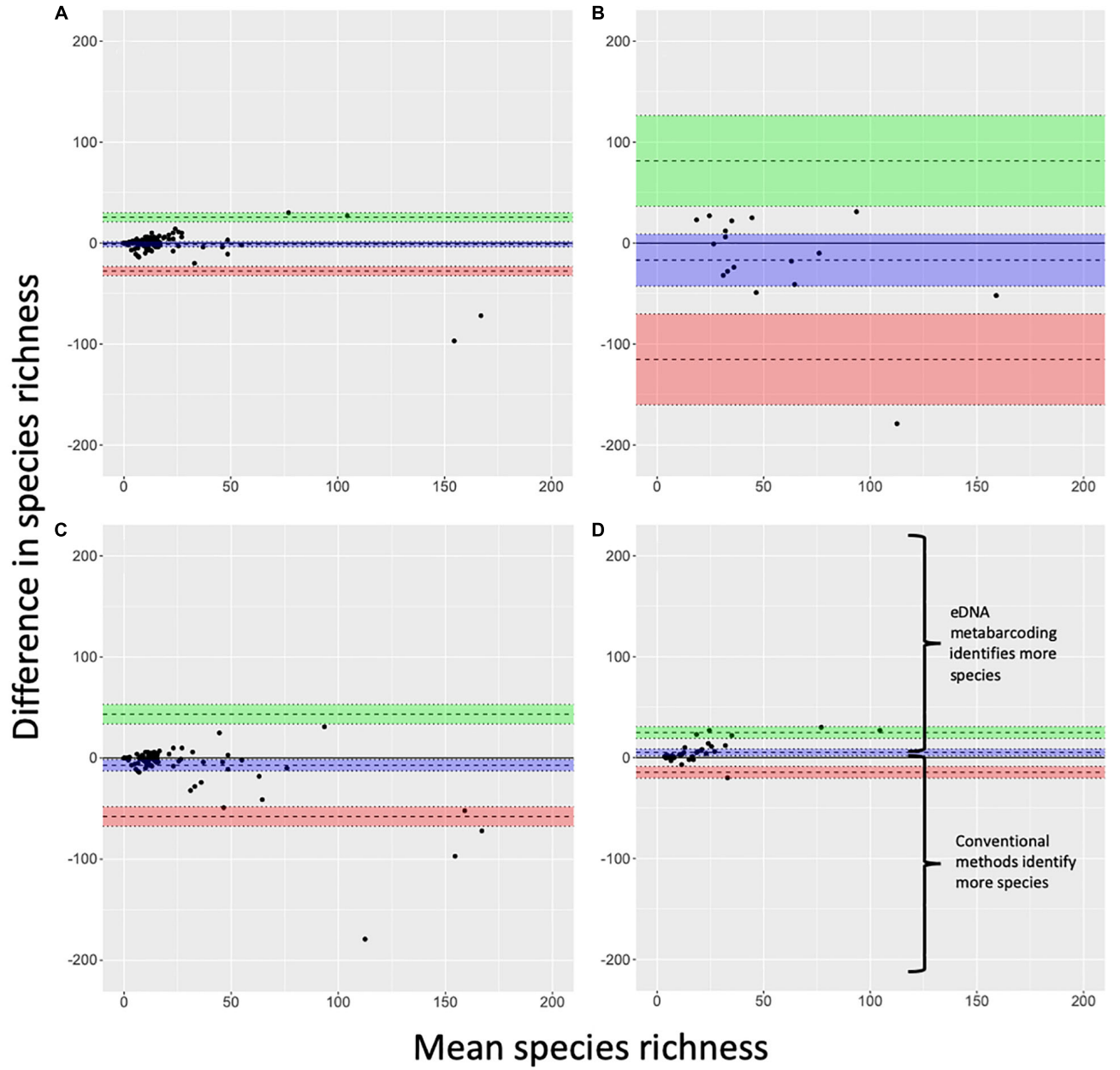


FIGURE 3 | Bland-Altman plots for **(A)** freshwater and **(B)** marine systems, **(C)** single-marker and **(D)** multi-marker observations for freshwater and marine systems combined. Blue band indicates a mean (wide dashed line) with 95% confidence intervals. Green and red bands indicate two standard deviations (wide dashed lines) from the mean with 95% confidence intervals. Outliers are observations that lie above and below the green and red bands, respectively. Regions of the figure where one method differentially detects more or fewer species than the other method are detailed in panel **(D)** and are consistent across all panels.

overlap (1.00). These seven sites originated from two studies (Valentini et al., 2016; Li et al., 2019) and were all low-diversity freshwater lentic systems (<7 species). Additionally, there were four sites with no shared detections where eDNA metabarcoding failed to detect 7 to 14 fish species identified in conventional surveys. These sites occurred in a single study of 31 oxbow and backwater lakes in Japan (Fujii et al., 2019).

In freshwater systems, the proportion of shared detections (0.61 ± 0.02, n = 98) was more than twice as in marine systems (0.26 ± 0.04, n = 17). The proportion of shared detections in multi-marker studies (0.63 ± 0.04, n = 37) was also higher than in single-marker studies (0.53 ± 0.03, n = 78). Interestingly, the proportion of eDNA-only detections was higher in marine systems (0.31 ± 0.06) than in freshwater systems (0.19 ± 0.02), and it was higher in multi-marker studies (**Figure 4A**; 0.27 ± 0.03) than in single-marker studies (**Figure 4B**; 0.18 ± 0.02).

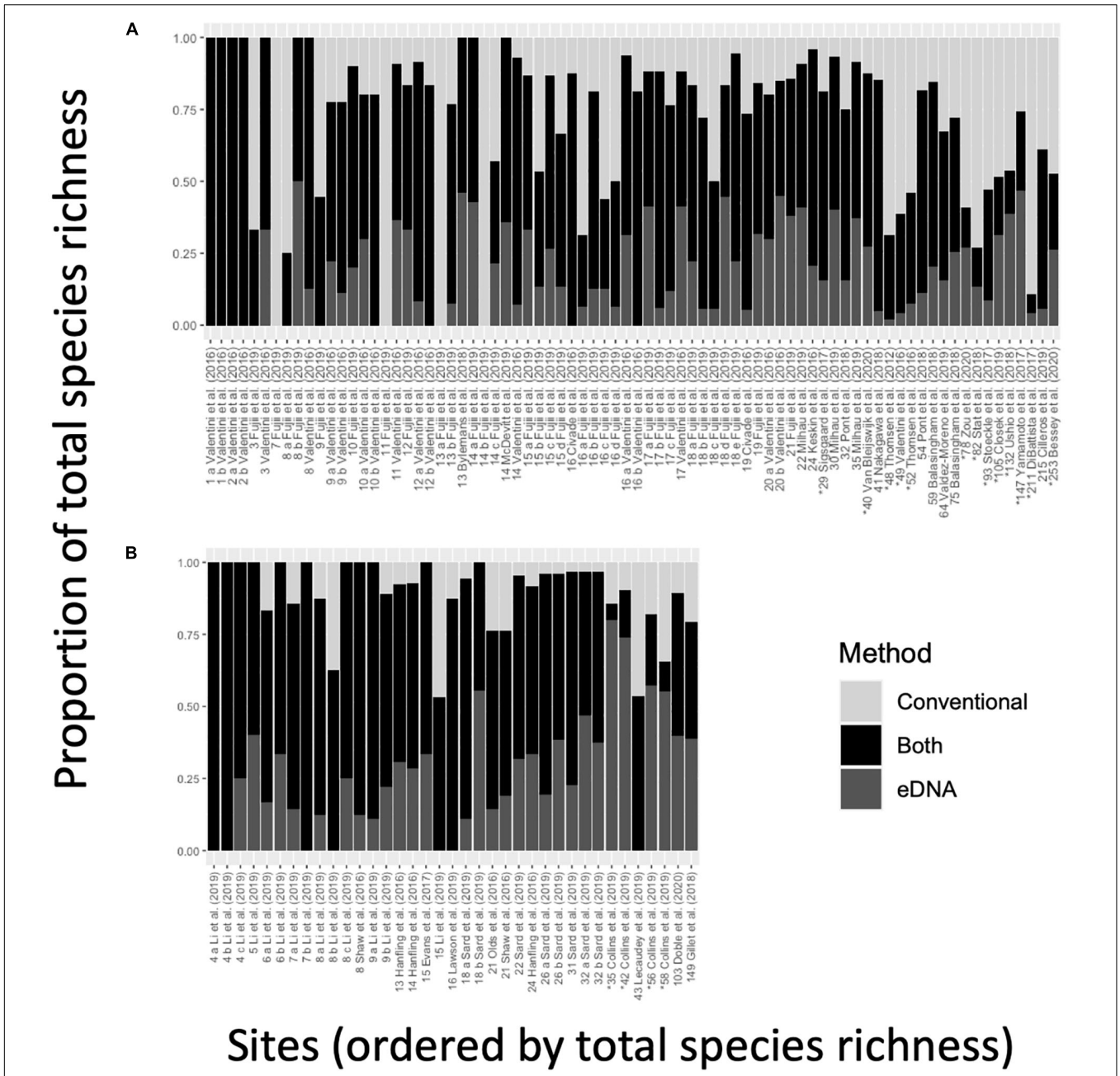


FIGURE 4 | The proportion of species detected by only eDNA (bottom, dark gray), eDNA and conventional surveys (middle, black), and only conventional surveys (top, light gray), ordered by total observed species richness for **(A)** single-marker and **(B)** multi-marker observations. Labels on the *x*-axis indicate the study from which the data were sourced – the alphabetical label is a unique observation within a study, and the preceding number indicates total species richness associated with an observation. Asterisks (*) indicate marine and estuarine observations. A cross-referenced table of values can be found in **Supplementary File S7**.

The high degree of shared species detections was matched by a similarly high degree of method-specific detections. At 77% of sites, eDNA metabarcoding revealed at least one additional fish species beyond those also detected by conventional surveys – at times adding up to 69 species (Yamamoto et al., 2017). At 78% of sites, conventional surveys also identified at least one additional fish species beyond those also detected by eDNA metabarcoding – in one case, adding up to 188 species (DiBattista et al., 2017).

DISCUSSION

As a measure of fish species richness, eDNA metabarcoding calibrates well to conventional surveys in low to moderately diverse freshwater systems (<100 species). In marine and more diverse systems, the relative performance of eDNA is unclear given few observations from these sites. It is possible there are proportionally many undetected taxa by eDNA metabarcoding

in high-diversity systems due in part to false negatives stemming from single-marker assays, incomplete reference libraries, PCR inhibition, and insufficient sampling effort or sequencing depth. Nevertheless, eDNA metabarcoding performs as well as conventional methods in many freshwater systems. This result should impart confidence in eDNA metabarcoding applications for ecological study and natural resource management (Kelly et al., 2014; Jerde, 2019; Sepulveda et al., 2020), particularly if further refinement of the approach improves performance in more diverse and marine systems.

Critically, we also show that both eDNA metabarcoding and conventional surveys detect unique fish species in all aquatic systems despite a high degree of overlapping detections. Thus, when species identities matter, such as for assessments of community composition, they are currently complementary methods. Although this is the case now, eDNA capacity may improve as we optimize metabarcoding approaches to better detect rare species and distinguish closely related species. Furthermore, eDNA offers the possibility of reanalyzing archived samples as we develop more powerful assays and instruments (Singer et al., 2019), which may reveal species that initially went undetected. Reanalysis of archival samples represents a significant advantage over conventional surveys, which are unlikely to experience a similar degree of technological advancement moving forward and for which retroactive species detection is impossible without some form of specimen capture and preservation.

One way of optimizing eDNA metabarcoding for biodiversity assessment involves using multiple genetic markers. Our results highlight that multi-marker assays improve species detection, albeit with added sequencing cost. For example, Doble et al. (2020) sampled eDNA concurrently with visual surveys at 21 sites in Lake Tanganyika to characterize the lake's highly diverse, endemic fish communities. The authors used four primer sets – two previously published and two newly developed for their study, including a cichlid-specific marker. With multiple markers targeting different loci, good genetic reference database coverage (83% of 431 known species), and deeper sequencing, eDNA metabarcoding identified 30 more fish species than aggregated detections from 27 concurrent snorkel surveys. Although deeper sequencing and multi-marker methods involve increased costs, these costs are unlikely to be greater than those associated with more intensive conventional field surveys, especially for remote and sensitive habitats.

However, multi-marker approaches do not guarantee better performance if reference databases are insufficiently populated or if primers impart amplification bias. The impact of these two factors on biodiversity inferences from eDNA metabarcoding has been explored (Hajibabaei et al., 2019; Kelly et al., 2019) but is inconsistently reported in field studies. Lecaudey et al. (2019) detected only 23 of 43 fish species (53%) known to occur in their study area despite using three gene markers (cyt*b*, 12S, 16S) in their analysis of Volga River eDNA samples. In a follow-up study reanalyzing the metabarcoding data, Schenekar et al. (2020) revealed that an incomplete reference database led to several false negatives and mis-assigned species. They also highlighted significant differences in primer efficiencies between markers and among species and the associated potential for false negatives. Addressing these pitfalls could involve *a priori* analyses of primer amplification bias using tissue from target species or *in silico* tests of primer specificity to inform appropriate genetic marker selection (Collins et al., 2019). In another multi-marker study assessing marine fish diversity in a public aquarium, Morey et al. (2020) only detected ∼50% of 107 known tank species with eDNA metabarcoding. While using three markers (12S, 16S, COI) improved taxonomic recovery over using just one or two, the approach was limited in part by poor quality of available reference sequences. Until we develop regional databases and molecular markers can ensure consistent species detection, eDNA metabarcoding will remain a complementary tool for aquatic biodiversity surveys in many systems rather than a stand-alone monitoring approach (McGee et al., 2019). Along with genetic reference development and primer optimization, two additional areas are ripe for improvement. We identified four instances of eDNA failing to detect any present fish species at sites where PCR amplification failed due to chemical inhibition (Fujii et al., 2019), which can occur when soil debris and humic substances are extracted along with DNA from environmental samples. Overcoming false negatives from PCR inhibition will enable robust species richness estimation with eDNA metabarcoding data. This is a well-acknowledged pitfall in the general eDNA approach (Goldberg et al., 2016) that can be mitigated with simple protocol adjustments to remove inhibitors from samples (McKee et al., 2015). Less acknowledged is the pitfall of insufficient sampling or spatial coverage to make eDNA metabarcoding inferences comparable to intensive conventional sampling. Justification of sampling effort and configuration for eDNA metabarcoding is a knowledge gap requiring more attention moving forward (Dickie et al., 2018), but see Evans et al. (2017a).

In the context of our synthesis, it is difficult to discern the impact of variable effort on the relative performance of eDNA metabarcoding and conventional surveys in part because effort can be characterized and quantified in myriad ways for both approaches. We made a limited attempt to explore the effect of differential effort in time using observations from concurrent and non-concurrent surveys. When water samples were collected concurrently with conventional surveys, eDNA metabarcoding performed as well as conventional methods, but with non-concurrent surveys, conventional methods slightly outperformed eDNA (**Supplementary File S3**). Although this may suggest some bias due to mismatched effort favoring conventional methods, it was not possible to disentangle this effect from other sources of variable effort like using multiple conventional gear types. Furthermore, the studies we analyzed varied considerably in their metabarcoding workflows – from the volume and number of samples collected, spatial coverage, filtration and extraction methods, to the selected markers and primers, sequencing platforms, and bioinformatics pipelines – all of which have documented impact on biodiversity inferences from eDNA (Djurhuus et al., 2017; Evans et al., 2017a; Alberdi et al., 2018; Grey et al., 2018; Zhang et al., 2018).

Despite substantial methodological variation across the synthesized studies (**Figure 2** and **Supplementary File S7**), we

found eDNA metabarcoding performed well compared to, in some cases, many years of conventional surveys (Olds et al., 2016; Yamamoto et al., 2017). A few explanations are possible. First, species richness inferences from eDNA metabarcoding may be robust against methodological variation. Second, conventional methods may significantly and systematically underestimate species richness. Third, eDNA metabarcoding may overestimate species richness in detecting DNA transported from locally absent species (Shogren et al., 2017; Andruszkiewicz et al., 2019), or false positives emerging from sources such as contamination or reference database errors in species identification (Jerde, 2019). Untangling biases in species richness (under- or overestimation) from eDNA metabarcoding will require some reference to species present in the target system, which could come from calibration experiments using complex mesocosms with known species composition. Although there have been recent calls for standardized approaches in eDNA metabarcoding (Shu et al., 2020), it is unclear if standardized protocols are needed for purposes of measuring community composition or if protocols should be optimized for each system. Still, best practices for eDNA metabarcoding are useful for minimizing contamination during sample and sequence processing and for maximizing yield with DNA capture, extraction, and amplification protocols (Goldberg et al., 2016).

There has also been a call for increased applications of eDNA metabarcoding in marine systems (Ausubel et al., 2018), which were poorly represented in our synthesis due primarily to a lack of species-level comparative studies. Indeed, ambiguous eDNA metabarcoding performance compared to conventional methods reflects an insufficient number of observations across a wide range of richness values (n = 17, range: 32 to 253 species). Nonetheless, a higher proportion of eDNA-only detections at marine sites demonstrate added value in eDNA-based marine monitoring even if overall relative performance is still unclear. In many cases, eDNA detected cryptic, nocturnal, rare, or elusive species missed by conventional surveys (Thomsen et al., 2016; Closek et al., 2019; Bessey et al., 2020). Further, more recent studies highlight improved species detection in marine systems with better populated reference libraries (Stoeckle et al., 2020) and multiple markers (Lafferty et al., 2020). Marine systems are especially difficult to sample comprehensively with conventional methods, and eDNA metabarcoding could expand the scale and resolution of monitoring at lower relative cost and effort. However, data from both approaches are needed to robustly assess the degree of agreement between them. To this end, we encourage collaboration between eDNA-samplers and the divers, seiners, snorkelers, trappers, and trawlers who together can provide such critical data.

One of the most game-changing promises of eDNA metabarcoding is the ability to detect biodiversity across the tree of life from simple environmental samples (Stat et al., 2017; Sawaya et al., 2019). We focused exclusively on water samples and the target group of bony fish here because there was a relatively large number of comparative observations across aquatic systems. Our findings should motivate similar data-gathering efforts and analyses for a wider range of organisms and habitats. Importantly, further investigations of the robustness of eDNA metabarcoding may reveal critical insights for taxa that are notoriously challenging to observe using conventional methods. Environmental DNA metabarcoding offers the promise of a unified approach to whole-ecosystem assessments, which would reduce monitoring costs, facilitate conservation and management, and enhance studies of ecological responses to growing global impacts (Trisos et al., 2020). The present analysis supports continued development and expansion of eDNA metabarcoding as an integral component of biodiversity monitoring in a world where innovative approaches are needed to track the effects of fast-paced and far-reaching ecological change.

DATA AVAILABILITY STATEMENT

All data are contained in the **Supplementary Material** or can be found within the article and **Supplementary Material** of the original sources of this synthesis.

AUTHOR CONTRIBUTIONS

All authors contributed to content and editing of the manuscript. CJ and MM wrote the manuscript, conceived the ideas, and conducted the analyses. MM, TeD, and EW compiled the database.

FUNDING

This research was partially supported by NASA (NNX14AR62A), BOEM (MC15AC00006), NOAA's support of the Santa Barbara Channel Marine Biodiversity Observation Network, USAID (AID-OAA-A-00057), DOD-SERDP [W912HG-12-C-0073 (RC-2240) and W912HQ19C0064 (RC19-1004)], California Regional Water Quality Control Board (16-023-140), NFWF (57446 and 55495), NSF GRF (1650114) and NGS Early Career (EC-333R-18), the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program (852621), and the U.S. Geological Survey Ecosystem Mission Area.

ACKNOWLEDGMENTS

We thank Adam Sepulveda, the Wonders of the Mekong workgroup, and participants of the *How to Achieve a True Consensus for Best Environmental DNA Practices* symposium held at UC Davis (January 29, 2020) for providing feedback and guidance on previous iterations of this manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.00276/full#supplementary-material

FILE S1 | Bland-Altman analyses for lentic and lotic freshwater systems.

FILE S2 | Species detection data for each study used in the synthesis ordered by first author. Each tab consists of one or more species lists and detection category (eDNA or conventional) to identify the method of detection for each species at independent sites. For each site, we calculate the sum of species detected by each method and include the data sourcing information referencing the primary research article.

FILE S3 | Bland-Altman analyses comparing species richness estimates between eDNA sampling with concurrent and non-concurrent conventional surveys.

FILE S4 | Summary of sites for which fish species richness was measured by both eDNA metabarcoding and conventional surveys. Asterisks (*) indicate richness observations generated from multiyear conventional survey data or when studies referenced 'all previous records' without specifying a survey method. Bold font indicates richness observations generated from surveys using multiple conventional gear types.

FILE S5 | Database search records and stepwise selection process. Includes list of records found, rejected, rationale for rejection, records retained, manual additions, and publication information.

FILE S6 | Flow diagram of study selection process used in this synthesis following PRISMA-P systematic review practices outlined in Moher et al. (2015).

FILE S7 | Data used to perform summary statistics, agreement and overlap analyses. Includes additional details of eDNA metabarcoding workflows.

FILE S8 | Lin's concordance correlation analyses for all sites and subsets of data.

REFERENCES

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., and Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods Ecol. Evol.* 9, 134–147. doi: 10.1111/2041-210x.12849

Andruszkiewicz, E. A., Koseff, J. R., Fringer, O. B., Ouellette, N. T., Lowe, A. B., Edwards, C. A., et al. (2019). Modeling environmental DNA transport in the coastal ocean using Lagrangian particle tracking. *Front. Mar. Sci.* 6:477. doi: 10.3389/fmars.2019.00477

Ausubel, J. E., Stoeckle, M. Y., and Gaffney, P. (2018). *Final Report: 1st US National Conference on Marine Environmental DNA (eDNA)*. New York, NY: The Marine Science & Policy Series.

Bakker, J., Wangensteen, O. S., Chapman, D. D., Boussarie, G., Buddo, D., Guttridge, T. L., et al. (2017). Environmental DNA reveals tropical shark diversity in contrasting levels of anthropogenic impact. *Sci. Rep.* 7:16886.

Balasingham, K. D., Walter, R. P., Mandrak, N. E., and Heath, D. D. (2018). Environmental DNA detection of rare and invasive fish species in two Great Lakes tributaries. *Mol. Ecol.* 27, 112–127. doi: 10.1111/mec.14395

Bálint, M., Nowak, C., Márton, O., Pauls, S. U., Wittwer, C., Aramayo, J. L., et al. (2018). Accuracy, limitations and cost efficiency of eDNA-based community survey in tropical frogs. *Mol. Ecol. Resour.* 18, 1415–1426. doi: 10.1111/1755-0998.12934

Barnes, M. A., and Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conserv. Genet.* 17:17.

Bessey, C., Jarman, S. N., Berry, O., Olsen, Y. S., Bunce, M., Simpson, T., et al. (2020). Maximizing fish detection with eDNA metabarcoding. *Environ. DNA* doi: 10.1002/edn3.74 [Epub ahead of print].

Bland, J. M., and Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 327, 307–310. doi: 10.1016/s0140-6736(86)90837-8

Blowes, S. A., Supp, S. R., Antão, L. H., Bates, A., Bruelheide, H., Chase, J. M., et al. (2019). The geography of biodiversity change in marine and terrestrial assemblages. *Science* 366, 339–345.

Boussarie, G., Bakker, J., Wangensteen, O. S., Mariani, S., Bonnin, L., Juhel, J. B., et al. (2018). Environmental DNA illuminates the dark diversity of sharks. *Sci. Adv.* 4:eaa9661.

Cantera, I., Cilleros, K., Valentini, A., Cerdan, A., Dejean, T., Iribar, A., et al. (2019). Optimizing environmental DNA sampling effort for fish inventories in tropical streams and rivers. *Sci. Rep.* 9, 1–11.

Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., et al. (2019). Unlocking biodiversity and conservation studies in high-diversity environments using environmental DNA (eDNA): a test with Guianese freshwater fishes. *Mol. Ecol. Resour.* 19, 27–46. doi: 10.1111/1755-0998.12900

Civade, R., Dejean, T., Valentini, A., Roset, N., Raymond, J. C., Bonin, A., et al. (2016). Spatial representativeness of environmental DNA metabarcoding signal for fish biodiversity assessment in a natural freshwater system. *PLoS One* 11:e0157366. doi: 10.1371/journal.pone.0157366

Closek, C. J., Santora, J. A., Starks, H. A., Schroeder, I. D., Andruszkiewicz, E. A., Sakuma, K. M., et al. (2019). Marine vertebrate biodiversity and distribution within the central California current using environmental DNA (eDNA) metabarcoding and ecosystem surveys. *Front. Mar. Sci.* 6:732. doi: 10.3389/fmars.2019.00732

Collins, R. A., Bakker, J., Wangensteen, O. S., Soto, A. Z., Corrigan, L., Sims, D. W., et al. (2019). Non-specific amplification compromises environmental DNA metabarcoding with COI. *Methods Ecol. Evol.* 10, 1985–2001. doi: 10.1111/2041-210X.13276

Datta, D. (2018). *blandr: Bland-Altman Method Comparison. R package version 0.5.1*.

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., et al. (2017). Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Mol. Ecol.* 26, 5872–5895. doi: 10.1111/mec.14350

DiBattista, J. D., Coker, D. J., Sinclair-Taylor, T. H., Stat, M., Berumen, M. L., and Bunce, M. (2017). Assessing the utility of eDNA as a tool to survey reef-fish communities in the Red Sea. *Coral Reefs* 36, 1245–1252. doi: 10.1007/s00338-017-1618-1

Dickie, I. A., Boyer, S., Buckley, H. L., Duncan, R. P., Gardner, P. P., Hogg, I. D., et al. (2018). Towards robust and repeatable sampling methods in eDNA-based studies. *Mol. Ecol. Resour.* 18, 940–952. doi: 10.1111/1755-0998.12907

Djurhuus, A., Port, J., Closek, C. J., Yamahara, K. M., Romero-Maraccini, O., Walz, K. R., et al. (2017). Evaluation of filtration and DNA extraction methods for environmental DNA biodiversity assessments across multiple trophic levels. *Front. Mar. Sci.* 4:314. doi: 10.3389/fmars.2017.00314

Doble, C. J., Hipperson, H., Salzburger, W., Horsburgh, G. J., Mwita, C., Murrell, D. J., et al. (2020). Testing the performance of environmental DNA metabarcoding for surveying highly diverse tropical fish communities: a case study from Lake Tanganyika. *Environ. DNA* 2, 24–41. doi: 10.1002/edn3.43

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., et al. (2012). New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Mol. Ecol.* 21, 1821–1833. doi: 10.1111/j.1365-294x.2012.05537.x

Evans, N. T., Li, Y., Renshaw, M. A., Olds, B. P., Deiner, K., Turner, C. R., et al. (2017a). Fish community assessment with eDNA metabarcoding: effects of sampling design and bioinformatic filtering. *Can. J. Fish. Aqu. Sci.* 74, 1362–1374. doi: 10.1139/cjfas-2016-0306

Evans, N. T., Shirey, P. D., Wieringa, J. G., Mahon, A. R., and Lamberti, G. A. (2017b). Comparative cost and effort of fish distribution detection via environmental DNA analysis and electrofishing. *Fisheries* 42, 90–99. doi: 10.1080/03632415.2017.1276329

Feng, D. (2020). *agRee: Various Methods for Measuring Agreement. R package version* 0.5.3.

Foote, A. D., Thomsen, P. F., Sveegaard, S., Wahlberg, M., Kielgast, J., Kyhn, L. A., et al. (2012). Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. *PLoS One* 7:e41781. doi: 10.1371/journal.pone.0041781

Fujii, K., Doi, H., Matsuoka, S., Nagano, M., Sato, H., and Yamanaka, H. (2019). Environmental DNA metabarcoding for fish community analysis in backwater lakes: a comparison of capture methods. *PLoS One* 14:e0210357. doi: 10.1371/journal.pone.0210357

Gillet, B., Cottet, M., Destanque, T., Kue, K., Descloux, S., Chanudet, V., et al. (2018). Direct fishing and eDNA metabarcoding for biomonitoring during a 3-year survey significantly improves number of fish detected around a South East Asian reservoir. *PLoS One* 13:e0208592. doi: 10.1371/journal.pone.0208592

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., et al. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods Ecol. Evol.* 7, 1299–1307. doi: 10.1111/2041-210x.12595

Grey, E. K., Bernatchez, L., Cassey, P., Deiner, K., Deveney, M., Howland, K. L., et al. (2018). Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. *Sci. Rep.* 8:8843.

Guardiola, M., Uriz, M. J., Taberlet, P., Coissac, E., Wangensteen, O. S., and Turon, X. (2015). Deep-sea, deep-sequencing: metabarcoding extracellular DNA from sediments of marine canyons. *PLoS One* 10:e0139633. doi: 10.1371/journal.pone.0139633

Hajibabaei, M., Porter, T. M., Wright, M., and Rudar, J. (2019). COI metabarcoding primer choice affects richness and recovery of indicator taxa in freshwater systems. *PLoS One* 14:e0220953. doi: 10.1371/journal.pone.0220953

Halpern, B. S., Frazier, M., Afflerbach, J., Lowndes, J. S., Micheli, F., O'Hara, C., et al. (2019). Recent pace of change in human impact on the world's ocean. *Sci. Rep.* 9:11609.

Hänfling, B., Lawson Handley, L., Read, D. S., Hahn, C., Li, J., Nichols, P., et al. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Mol. Ecol.* 25, 3101–3119. doi: 10.1111/mec.13660

Jerde, C. L. (2019). Can we manage fisheries with the inherent uncertainty from eDNA?. *J. Fish Biol.* doi: 10.1111/jfb.14218 [Epub ahead of print].

Jerde, C. L., Mahon, A. R., Chadderton, W. L., and Lodge, D. M. (2011). "Sight-unseen" detection of rare aquatic species using environmental DNA. *Conserv. Lett.* 4, 150–157. doi: 10.1111/j.1755-263x.2010.00158.x

Jerde, C. L., Wilson, E. A., and Dressler, T. L. (2019). Measuring global fish species richness with eDNA metabarcoding. *Mol. Ecol. Resour.* 19, 19–22. doi: 10.1111/1755-0998.12929

Kelly, R. P., Port, J. A., Yamahara, K. M., Martone, R. G., Lowell, N., Thomsen, P. F., et al. (2014). Harnessing DNA to improve environmental management. *Science* 344, 1455–1456. doi: 10.1126/science.1251156

Kelly, R. P., Shelton, A. O., and Gallego, R. (2019). Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. *Sci. Rep.* 9, 1–14.

Keskin, E., Unal, E. M., and Atar, H. H. (2016). Detection of rare and invasive freshwater fish species using eDNA pyrosequencing: Lake Iznik ichthyofauna revised. *Biochem. Syst. Ecol.* 67, 29–36. doi: 10.1016/j.bse.2016.05.020

Kinoshita, G., Yonezawa, S., Murakami, S., and Isagi, Y. (2019). Environmental DNA collected from snow tracks is useful for identification of mammalian species. *Zool. Sci.* 36, 198–207.

Kraaijeveld, K., De Weger, L. A., Ventayol García, M., Buermans, H., Frank, J., Hiemstra, P. S., et al. (2015). Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Mol. Ecol. Resour.* 15, 8–16. doi: 10.1111/1755-0998.12288

Lacoursière-Roussel, A., and Deiner, K. (2019). Environmental DNA is not the tool by itself. *J. Fish Biol.* doi: 10.1111/jfb.14177 [Epub ahead of print].

Lacoursière-Roussel, A., Howland, K., Normandeau, E., Grey, E. K., Archambault, P., Deiner, K., et al. (2018). eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. *Ecol. Evol.* 8, 7763–7777. doi: 10.1002/ece3.4213

Lafferty, K. D., Garcia-Vedrenne, A. E., McLaughlin, J. P., Childress, J. N., Morse, M. F., and Jerde, C. L. (2020). At Palmyra Atoll, the fish-community environmental DNA signal changes across habitats but not with tides. *J. Fish Biol.* doi: 10.1111/jfb.14403 [Epub ahead of print].

Lawrence, I., and Lin, K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45, 255–268.

Lawson Handley, L., Read, D. S., Winfield, I. J., Kimbell, H., Johnson, H., Li, J., et al. (2019). Temporal and spatial variation in distribution of fish environmental DNA in England's largest lake. *Environ. DNA* 1, 26–39. doi: 10.1002/edn3.5

Lecaudey, L. A., Schletterer, M., Kuzovlev, V. V., Hahn, C., and Weiss, S. J. (2019). Fish diversity assessment in the headwaters of the Volga River using environmental DNA metabarcoding. *Aqu. Conserv. Mar. Freshw. Ecosyst.* 29, 1785–1800. doi: 10.1002/aqc.3163

Li, J., Hatton-Ellis, T. W., Lawson Handley, L. J., Kimbell, H. S., Benucci, M., Peirson, G., et al. (2019). Ground-truthing of a fish-based environmental DNA metabarcoding method for assessing the quality of lakes. *J. Appl. Ecol.* 56, 1232–1244. doi: 10.1111/1365-2664.13352

Lim, N. K., Tay, Y. C., Srivathsan, A., Tan, J. W., Kwik, J. T., Baloğlu, B., et al. (2016). Next-generation freshwater bioassessment: eDNA metabarcoding with a conserved metazoan primer reveals species-rich and reservoir-specific communities. *R. Soc. Open Sci.* 3:160635. doi: 10.1098/rsos.160635

Lopes, C. M., Sasso, T., Valentini, A., Dejean, T., Martins, M., Zamudio, K. R., et al. (2017). eDNA metabarcoding: a promising method for anuran surveys in highly diverse tropical forests. *Mol. Ecol. Resour.* 17, 904–914. doi: 10.1111/1755-0998.12643

McDevitt, A. D., Sales, N. G., Browett, S. S., Sparnenn, A. O., Mariani, S., Wangensteen, O. S., et al. (2019). Environmental DNA metabarcoding as an effective and rapid tool for fish monitoring in canals. *J. Fish Biol.* 95, 679–682. doi: 10.1111/jfb.14053

McGee, K. M., and Eaton, W. D. (2015). A comparison of the wet and dry season DNA-based soil invertebrate community characteristics in large patches of the bromeliad *Bromelia pinguin* in a primary forest in Costa Rica. *Appl. Soil Ecol.* 87, 99–107. doi: 10.1016/j.apsoil.2014.11.009

McGee, K. M., Robinson, C., and Hajibabaei, M. (2019). Gaps in DNA-based biomonitoring across the globe. *Front. Ecol. Evol.* 7:337. doi: 10.3389/fevo.2019.00337

McKee, A. M., Spear, S. F., and Pierson, T. W. (2015). The effect of dilution and the use of a post-extraction nucleic acid purification column on the accuracy, precision, and inhibition of environmental DNA samples. *Biol. Conserv.* 183, 70–76. doi: 10.1016/j.biocon.2014.11.031

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., et al. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* 4:1.

Morey, K. C., Bartley, T. J., and Hanner, R. H. (2020). Validating environmental DNA metabarcoding for marine fishes in diverse ecosystems using a public aquarium. *Environ. DNA* 2, 330–342. doi: 10.1002/edn3.76

Nakagawa, H., Yamamoto, S., Sato, Y., Sado, T., Minamoto, T., and Miya, M. (2018). Comparing local-and regional-scale estimations of the diversity of stream fish using eDNA metabarcoding and conventional observation methods. *Freshw. Biol.* 63, 569–580. doi: 10.1111/fwb.13094

Olds, B. P., Jerde, C. L., Renshaw, M. A., Li, Y., Evans, N. T., Turner, C. R., et al. (2016). Estimating species richness using environmental DNA. *Ecol. Evol.* 6, 4214–4226.

Pecl, G. T., Araújo, M. B., Bell, J. D., Blanchard, J., Bonebrake, T. C., Chen, I. C., et al. (2017). Biodiversity redistribution under climate change: impacts on ecosystems and human well-being. *Science* 355:eaai9214.

Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., et al. (2018). Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Sci. Rep.* 8:10361.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T., et al. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biol. Rev.* 94, 849–873.

Sard, N. M., Herbst, S. J., Nathan, L., Uhrig, G., Kanefsky, J., Robinson, J. D., et al. (2019). Comparison of fish detections, community diversity, and relative abundance using environmental DNA metabarcoding and traditional gears. *Environ. DNA* 1, 368–384. doi: 10.1002/edn3.38

Sawaya, N. A., Djurhuus, A., Closek, C. J., Hepner, M., Olesin, E., Visser, L., et al. (2019). Assessing eukaryotic biodiversity in the Florida Keys National Marine Sanctuary through environmental DNA metabarcoding. *Ecol. Evol.* 9, 1029–1040. doi: 10.1002/ece3.4742

Schenekar, T., Schletterer, M., Lecaudey, L. A., and Weiss, S. J. (2020). Reference databases, primer choice, and assay sensitivity for environmental metabarcoding: lessons learnt from a re-evaluation of an eDNA fish assessment in the Volga headwaters. *River Res. Appl.* doi: 10.1002/rra.3610 [Epub ahead of print].

Sepulveda, A. J., Nelson, N. M., Jerde, C. L., and Luikart, G. (2020). Are environmental DNA methods ready for aquatic invasive species management? *Trends Ecol. Evol.* 35, 668–678. doi: 10.1016/j.tree.2020.03.011

Shaw, J. L., Clarke, L. J., Wedderburn, S. D., Barnes, T. C., Weyrich, L. S., and Cooper, A. (2016). Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biol Conserv.* 197, 131–138. doi: 10.1016/j.biocon.2016.03.010

Shogren, A. J., Tank, J. L., Andruszkiewicz, E., Olds, B., Mahon, A. R., Jerde, C. L., et al. (2017). Controls on eDNA movement in streams: transport, retention, and resuspension. *Sci. Rep.* 7:5065.

Shu, L., Ludwig, A., and Peng, Z. (2020). Standards for methods utilizing environmental DNA for detection of fish species. *Genes* 11:296. doi: 10.3390/genes11030296

Sigsgaard, E. E., Nielsen, I. B., Carl, H., Krag, M. A., Knudsen, S. W., Xing, Y., et al. (2017). Seawater environmental DNA reflects seasonality of a coastal fish community. *Mar. Biol.* 164:128.

Singer, G. A. C., Fahner, N. A., Barnes, J. G., McCarthy, A., and Hajibabaei, M. (2019). Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater. *Sci. Rep.* 9:5991.

Sonderegger, D. L., Wang, H., Clements, W. H., and Noon, B. R. (2009). Using SiZer to detect thresholds in ecological data. *Front. Ecol. Environ.* 7:190–195. doi: 10.1890/070179

Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., et al. (2017). Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Sci. Rep.* 7:12240.

Stat, M., John, J., DiBattista, J. D., Newman, S. J., Bunce, M., and Harvey, E. S. (2018). Combined use of eDNA metabarcoding and video surveillance for the assessment of fish biodiversity. *Conserv. Biol.* 33, 196–205. doi: 10.1111/cobi.13183

Stoeckle, M. Y., Das Mishu, M., and Charlop-Powers, Z. (2020). Improved environmental dna reference library detects overlooked marine fishes in New Jersey, United States. *Front. Mar. Sci.* 7:226. doi: 10.3389/fmars.2020.00226

Stoeckle, M. Y., Soboleva, L., and Charlop-Powers, Z. (2017). Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. *PLoS One* 12:e0175186. doi: 10.1371/journal.pone.0175186

Thomsen, P. F., Kielgast, J., Iversen, L. L., Møller, P. R., Rasmussen, M., and Willerslev, E. (2012). Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS One* 7:e41732. doi: 10.1371/journal.pone.0041732

Thomsen, P. F., Møller, P. R., Sigsgaard, E. E., Knudsen, S. W., Jørgensen, O. A., and Willerslev, E. (2016). Environmental DNA from seawater samples correlate with trawl catches of subarctic, deepwater fishes. *PLoS One* 11:e0165252. doi: 10.1371/journal.pone.0165252

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

Tilman, D., Clark, M., Williams, D. R., Kimmel, K., Polasky, S., and Packer, C. (2017). Future threats to biodiversity and pathways to their prevention. *Nature* 546, 73–81. doi: 10.1038/nature22900

Trisos, C. H., Merow, C., and Pigot, A. L. (2020). The projected timing of abrupt ecological disruption from climate change. *Nature* 580, 496–501. doi: 10.1038/s41586-020-2189-9

Ushio, M., Fukuda, H., Inoue, T., Makoto, K., Kishida, O., Sato, K., et al. (2017). Environmental DNA enables detection of terrestrial mammals from forest pond water. *Mol. Ecol. Resour.* 17, e63–e75. doi: 10.1111/1755-0998.12690

Ushio, M., Murakami, H., Masuda, R., Sado, T., Miya, M., Sakurai, S., et al. (2018). Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. *Metabarcod. Metagenomics* 2:e23297. doi: 10.1101/113472

Valdez-Moreno, M., Ivanova, N. V., Elias-Gutierrez, M., Pedersen, S. L., Bessonov, K., and Hebert, P. D. (2019). Using eDNA to biomonitor the fish community in a tropical oligotrophic lake. *PLoS One* 14:e0215505. doi: 10.1371/journal.pone.0215505

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., et al. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Mol. Ecol.* 25, 929–942. doi: 10.1111/mec.13428

Yamamoto, S., Masuda, R., Sato, Y., Sado, T., Araki, H., Kondoh, M., et al. (2017). Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. *Sci. Rep.* 7:40368. doi: 10.1038/srep40368

Yoccoz, N. G., Bråthen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., et al. (2012). DNA from soil mirrors plant taxonomic and growth form diversity. *Mol. Ecol.* 21, 3647–3655. doi: 10.1111/j.1365-294X.2012.05545.x

Zhang, G. K., Chain, F. J., Abbott, C. L., and Cristescu, M. E. (2018). Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities. *Evol. Appl.* 11, 1901–1914. doi: 10.1111/eva.12694

Zou, K., Chen, J., Ruan, H., Li, Z., Guo, W., Li, M., et al. (2020). eDNA metabarcoding as a promising conservation tool for monitoring fish diversity in a coastal wetland of the Pearl River Estuary compared to bottom trawling. *Sci. Total Environ.* 702:134704. doi: 10.1016/j.scitotenv.2019.134704

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 06 November 2020
doi: 10.3389/fevo.2020.569425

An Environmental DNA Survey on Distribution of an Endangered Salmonid Species, *Parahucho perryi*, in Hokkaido, Japan

*Hiroki Mizumoto[1], Takashi Mitsuzuka[2] and Hitoshi Araki[1]**

[1] Research Faculty of Agriculture, Hokkaido University, Sapporo Japan, [2] Pacific Consultants Co., Ltd., Tokyo, Japan

For protecting endangered species, precise understanding of their distribution is crucial. However, it is often very difficult to estimate at a large scale with conventional methods (e.g., casting nets or electrofishing for aquatic species) because of their low densities in the wild. Sakhalin taimen (*Parahucho perryi*) is one of the largest and most critically endangered freshwater salmonid fishes in the world. In this study, we applied an environmental DNA (eDNA) detection system for this species to 120 rivers in Hokkaido, the second largest main island of Japan. We successfully detected eDNA from Sakhalin taimen in seven rivers (5.8%). Although these rivers were widely distributed across the island, > 95% of the total amounts of eDNA were detected from region-A and -I, indicating that local populations in the other regions of Hokkaido are very small and on the brink of extinction. In addition, principal component analyses based on the eDNA-based estimation of Sakhalin taimen distribution and GIS revealed their distribution determinants including limited topographic relief of watershed as well as presence of wetlands and lagoons. Our results suggest that eDNA-based detection systems are an efficient means of monitoring the population status of endangered freshwater species at large scales.

Keywords: environmental DNA, endangered species, Sakhalin taimen, distribution, habitat degradation

OPEN ACCESS

Edited by:
Matthew A. Barnes,
Texas Tech University, United States

Reviewed by:
Peter Rand,
Prince William Sound Science Center, United States
Andrew R. Mahon,
Central Michigan University, United States

****Correspondence:***
Hitoshi Araki
arakih@res.agr.hokudai.ac.jp

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *04 June 2020*
Accepted: *09 October 2020*
Published: *06 November 2020*

Citation:
Mizumoto H, Mitsuzuka T and Araki H (2020) An Environmental DNA Survey on Distribution of an Endangered Salmonid Species, Parahucho perryi, in Hokkaido, Japan. Front. Ecol. Evol. 8:569425. doi: 10.3389/fevo.2020.569425

INTRODUCTION

Precise information about the current distribution and biomass of threatened species is essential for their conservation and appropriate management. Over 30,000 species are currently listed as threatened species in the Red List by the International Union for Conservation of Nature (IUCN), and the number is increasing (International Union for Conservation of Nature and Natural Resources, 2020). While many kinds of threats are considered as factors for extinction (e.g., loss or degradation of habitat, illegal trade, invasive species, or human activity), 85% of the endangered species are facing threats of habitat loss or degradation (Wilcove et al., 1998). These threats potentially accelerate further reduction of population sizes through a process known as an 'extinction vortex" (Gilpin and Soulé, 1986). However, vast sampling efforts are often required for field surveys with conventional methods (e.g., visual observation or physical capture). In addition, specialized skills are typically needed for species identification, especially in the case of

juveniles. Due to these requirements, studies on current distribution of endangered species are often incomplete and may be limited to small parts of a species' distribution.

The Sakhalin taimen (*Parahucho perryi*) has been listed as a critically endangered species (CR) in the IUCN Red List since 2006 (Rand, 2006). They are a typical example of an endangered species threatened by habitat loss or degradation due to human activities (Fukushima, 2006; Rand, 2006; Fukushima et al., 2007; Zolotukhin et al., 2013). Historically, they were distributed in the Russian Far East and northern Japan (Rand, 2006; Zolotukhin et al., 2013). Hokkaido is the second largest main island of Japan (83,456 km^2), and currently the only island hosting Sakhalin taimen in the country (Fukushima et al., 2011). According to a previous research (Fukushima et al., 2011), Sakhalin taimen populations require various habitats including upstream habitats, estuarine and/or coastal habitats, and lagoons and wetlands through their life history. However, even the best available scientific report on the distribution of this species above suffers from spotty and sometimes anecdotal records across large time spans, bringing some conclusions about the species' ecology and distribution into doubt.

Environmental DNA (eDNA) techniques have been developed as a new assessment tool for aquatic organisms since 2008 (Ficetola et al., 2008; Fukumoto et al., 2015; Pfleger et al., 2016). These techniques have many advantages for detecting target species in the field, such as objectivity, high detectability, and reduced sampling effort. They have been shown to be especially useful for monitoring endangered species that are otherwise difficult to capture or observe (Fukumoto et al., 2015; Laramie et al., 2015; Pfleger et al., 2016; Carlsson et al., 2017; Maruyama et al., 2018; Atkinson et al., 2019; Iwai et al., 2019). However, the previous studies tended to apply the eDNA techniques to specific freshwater/saltwater systems or to relatively small geographical regions. For comprehensive understanding of distribution of endangered species and its determinants, an applicability of this technique to an investigation in large geographical regions with a variety of environments is crucial.

In this study, we aimed to estimate the distribution and population status of Sakhalin taimen in 120 rivers covering the entirety of Hokkaido, Japan, using an eDNA detection system established during a previous study (Mizumoto et al., 2018). Furthermore, we sought to understand key environmental determinants for the presence/absence of Sakhalin taimen using Geographic Information System (GIS), so that the current distribution could be estimated together with limiting factors for this endangered species. Given the size of Hokkaido (83,450 km^2) and heterogeneity in river environments in it, this study provides a textbook example of an eDNA application to understand a wide-range distribution, population status and their environmental determinants of endangered species.

MATERIALS AND METHODS

Sample Collections

Field sampling was conducted at 120 rivers in Hokkaido, Japan, from 2015 to 2018 (**Figure 1** and **Supplementary Table A1**). Sampling sites were selected around estuaries on each river for the sake of convenience and uniformity of the sampling effort. Because of the Sakhalin taimen's conservation status and risk of increased fishing pressure on them due to this study, eDNA detection results were projected on a map of the 14 subprefectures (A–N) of Hokkaido.

Filtering procedures are described in Mizumoto et al. (2018). In brief, collected water samples were filtered at field stations or visitor centers near sampling sites as soon as possible. When we estimated that it had taken > 6 h from water collection to filtration, we added benzalkonium chloride (BAC) for preserving DNA at a final concentration of 0.01% (Yamanaka et al., 2017; **Supplementary Table A1**). At most sampling sites, 2 L of environment water were collected and filtered separately to provide 1 L duplicates, but in some samples, only 500 mL could be filtered due to clogging by fine particles in turbid water. All water samples were filtered through glass-membrane filters (Whatman GF/F, GE Healthcare Japan, Tokyo, Japan) using an aspirator. After filtration, 10 mL of 70% molecular grade ethanol was filtered for preserving DNA in the sample. Negative control samples were collected by filtering 500 mL of sterilized distilled water before filtering each river sample. All filter samples were stored around −20°C at field stations and at −25°C in the lab until further analyses. To prevent cross-contamination, all filtration equipment was bleached using a 6% sodium hypochlorite solution and carefully rinsed with sterilized distilled water after each filtration.

DNA Extraction and qPCR Procedures

DNA extractions were conducted using DNeasy Blood and Tissue kits (Qiagen, Hilden, Germany) following Mizumoto et al. (2018). The final volume of the extracted DNA was 110 μL/sample. qPCRs were carried out in 20 μL volumes with 400 nM of Sakhalin taimen specific primer and probe (Forward-primer: 5′-GCAATAGGCCTCCTATCAACAA-3′, Reverse-primer: 5′-CGAATTGTAAAATTAGTACTATCCATCCA-3′, Probe: 5′-FAM-AGCATACTCTTCAATCGCCCACCT-TAMRA-3′) designed by Mizumoto et al. (2018) in a Brilliant III Ultra-Fast qPCR Master Mix with Low ROX (Agilent Technologies, Inc.) and 2 μL of the extracted DNA as a template sample (Mizumoto et al., 2018). Triplicates of 2 μL were applied to each qPCR reaction on a Stratagene Mx3000P (Agilent Technologies, Inc.) with the same thermal-cycling regime as in the previous study (Mizumoto et al., 2018). For the quantification of environmental samples, we applied synthetic linear DNA of the target region containing 2×10^1, 2×10^2, 2×10^3, 2×10^4 and 2×10^5 copies per tube as standards in all qPCR assays ($n = 21$, average $R^2 = 0.992$, $SD = 0.007$, and average PCR efficiency = 0.894, $SD = 0.126$). In addition, sterilized distilled water samples were applied in all qPCR assays as PCR negative control samples.

Estimation of Distribution and GIS analyses

We accepted sites as positive detections when we had at least one detection in six PCR replicates per site (triplicate PCR reactions x two filters per site). The lowest detection limit was set at "0.01

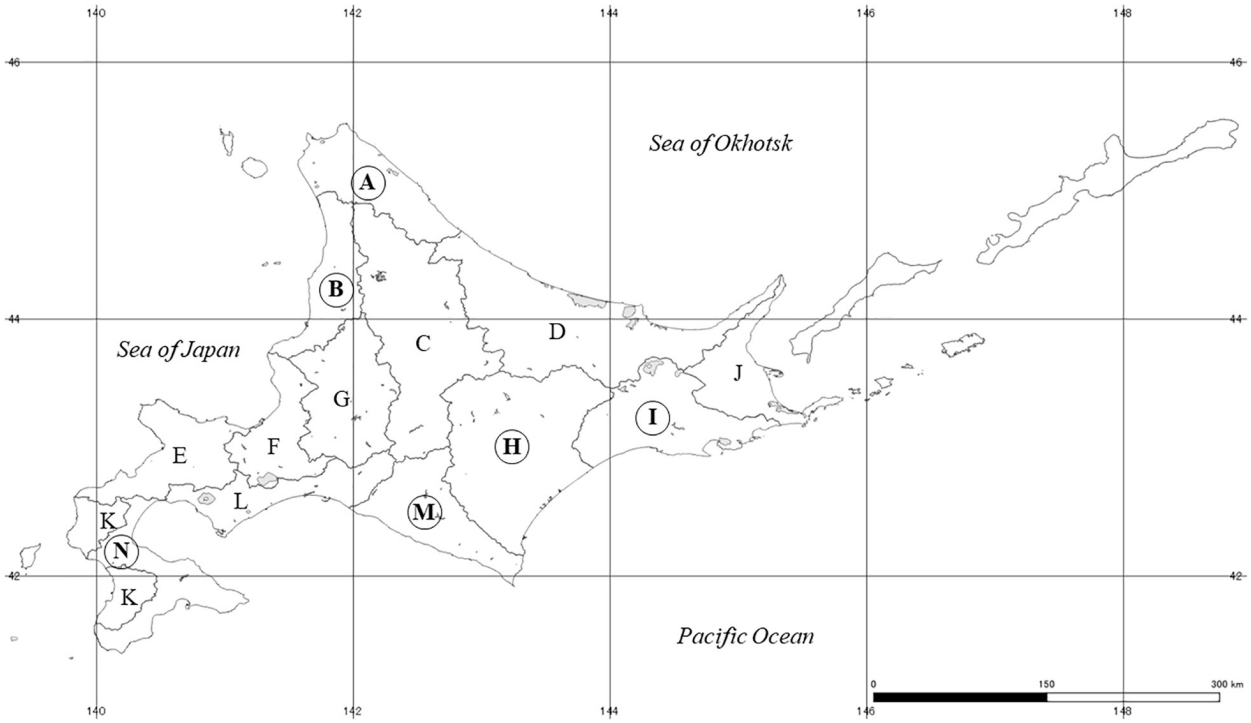


FIGURE 1 | The map of our survey areas. Sampling was conducted from 2015 to 2018 at 120 rivers in total. All sampling sites were located around estuaries, however, detailed geographical information was not presented for the protection and conservation of Sakhalin taimen. Instead a larger scale review was included for the 14 subprefectures (A-N) of Hokkaido. The number of rivers located each region were 16 (region A), 8 (region B), 12 (region D), 12 (region E), 3 (region F), 6 (region H), 16 (region I), 18 (region J), 7 (region L), 8 (region M) and 14 (region N). Circled alphabets represent areas where Sakhalin taimen DNA was detected. The river and coastline data were provided by the National Land Information Division, Ministry of Land, Infrastructure, Transport and Tourism of Japan, under the CC BY 4.0 license.

copy/2 μL*site." To adjust for the variation in water volumes we sampled (500 mL or 1,000 mL, see above), estimated DNA concentrations were represented as/1,000 mL of environmental water after averaging over the six replicates. In the following analyses, qPCR outputs (copies/1,000 mL) that were obtained from the same river system's samples were averaged across the river system (**Supplementary Table A1**). Ultimately, 116 river systems were used for GIS analyses.

To investigate factors determining the settlement of Sakhalin taimen, we examined 23 environmental variables (18 for natural environment and 5 for human activity) in 116 rivers (**Table 1**). For this study, we looked into the same variables that were tested in Fukushima et al. (2011) to confirm the current population status of Sakhalin taimen. Among the environmental variables, information on the distribution of marshes were collected from the 5th National Survey on Natural Environment (The Ministry of the Environment, 1995), and the other variable data were provided by the Geospatial Information Authority of Japan (Ministry of Land, Infrastructure, Transport and Tourism, 2020). In general, an aggregation of variables is often used by a mesh as a division unit. However, in order to grasp the habitat of Sakhalin taimen, it is effective to use a regional classification based on basin. The basins of division unit were merged from the basin mesh data (30 × 45 s for latitude and longitude, about 1 km per side) based on information published by Geospatial Information Authority of Japan. For each unit basin created, the variable mesh was calculated as the sum (population, number of lagoons, number of wetlands), average (temperature, precipitation, altitude, slope), or ratio (low land area, land use). The degree of meandering was obtained by dividing the flow path extension from the start point to the end of the longest river in the relevant water system by the linear distance. We used ArcGIS 10.7 (ESRI) to aggregate these environment variables.

To summarize the aggregated environmental variables into characteristic environments, we used principal component analysis (PCA). From the obtained PCA scores, we performed cluster analyses using the k-means method and categorized the rivers. The number of clusters was determined using the largest score according to the Calinski-Harabasz score. Statistical analyses were performed using R3.6.1 (R Core Team, 2019), cluster analyses were performed using Package "cclust," and the Calinski-Harabasz score was calculated using Package "vegan."

RESULTS

We detected eDNA of Sakhalin taimen in 7 of 120 rivers in Hokkaido (5.8%). No detection was observed in any negative control samples. Among the 14 subprefectures of Hokkaido (A-N), two sites were positive in region-A, and one site was positive

in region-B, -H, -I, -M, and -N (**Table 2**). Three of the seven river systems with positive eDNA detection had no previous reports of Sakhalin taimen, whereas Sakhalin taimen in the other four river systems were previously reported (Fukushima et al., 2011). Among the eDNA detections in the seven river systems, the estimated copy numbers varied largely (**Table 3**). The highest number of DNA copies were obtained in river-A-11 and -I-2, being consistent with healthy populations of Sakhalin taimen reported previously in these river systems (Fukushima et al., 2011).

TABLE 1 | Summary of the 23 environmental variables (18 for natural environment and 5 for human activity) used for evaluation of 116 rivers.

Variable type	Variable	Unit	Number of variables
Natural variables	Drainage area	ha	1
	Elevation	m	1
	Area below 10, 20, and 30 m	%	3
	Lagoon presence, number, and area	ha	3
	Marsh presence, number, and area	ha	3
	Forest	%	1
	Slope	Degree	1
	Annual air temperature (max, min, and average)	℃	3
	Precipitation	mm	1
	Degree of meandering		1
Human activity	Dam number and density		2
	Human population and density	/ha	2
	Farmland	%	1

Marsh distribution was collected from the 5th National Survey on Natural Environment (The Ministry of the Environment, 1995), and the other variables were prepared by the Geospatial Information Authority of Japan (Ministry of Land, Infrastructure, Transport and Tourism, 2020).

TABLE 2 | Summary of the results of qPCR from 120 river water samples.

Region	Detection	Copies	Ratio
A	2/16	351.69	56.36%
B	1/8	11.73	1.88%
D	0/12	–	–
E	0/12	–	–
F	0/3	–	–
H	1/6	0.67	0.11%
I	1/16	246.40	39.49%
J	0/18	–	–
L	0/7	–	–
M	1/8	8.89	1.42%
N	1/14	4.62	0.74%
Total	7/120	624.01	100%

Copies represent the sum of the estimated DNA copy numbers per 1,000 mL from each site. Three of the seven river systems with positive eDNA detections from region-B, -M, and -N had no previous reports of Sakhalin taimen.

TABLE 3 | Results of qPCR from 7 rivers where Sakhalin taimen DNA was positively detected.

River	Detection	eDNA concentration (SD)	Ratio
A-8	4/6	103.92 (105.18)	16.65%
A-11	4/6	247.78 (190.10)	39.71%
B-5	1/6	11.73 (26.24)	1.88%
H-4	1/6	0.67 (1.49)	0.11%
I-2	4/6	246.40 (289.71)	39.49%
M-6	1/6	8.89 (19.88)	1.42%
N-14	1/6	4.62 (10.34)	0.74%

River names were reassigned with alphanumeric codes (A-8, A-11, B-5, H-4, I-2, M-6, and N-14) based on the region where these rivers were located. Detection represents the number of positive detections from six qPCR replicates at each site and eDNA concentration represents the estimated DNA copy number per 1,000 mL from each site. SD represents standard deviation. Ratio represents contribution rate for the total estimated eDNA concentration.

Among the results of GIS-based PCA, percentages of the top three explained variances were 29.3, 18.2, and 10.4% (illustrated as Dimension 1, 2, and 3, respectively, in **Figure 2**). Regarding the three explained variances, ten parameters (slope, below 30 m, forest, below 20, elevation, wetland presence, below 10 m, lagoon presence, precipitation, and farmland) that affected geographical flatness had high variance contributions in dimension 1, six parameters (drainage area, dam number, human population, wetland number, minimum air temperature and average air temperature) affecting river scales had high variance contributions in dimension 2, and seven parameters (average air temperature, minimum air temperature, maximum air temperature, dam density, dam number, wetland number, and human population) affecting climates had high variance contributions in dimension 3 (**Figure 2**). Forest appeared to have a negative effect of Sakhalin taimen's presence although it was estimated to have a positive effect in the previous study (Fukushima et al., 2011). In contrast, the number of dams and farmland appeared to have positive effects of Sakhalin taimen's presence, whereas these factors were estimated to have negative effects in the previous study (Fukushima et al., 2011).

DISCUSSION

In this study, we applied an eDNA detection system to estimate the distribution of a critically endangered salmonid species, Sakhalin taimen (*P. perryi*), across Hokkaido, Japan. As a result, we detected Sakhalin taimen eDNA in seven out of 120 rivers in Hokkaido. Of the seven rivers where Sakhalin taimen's eDNA were successfully detected, three rivers had no previous reports (river-B-5, -M-6, and -N-14, **Tables 2**, **3**).

Tank experiments in our previous study (Mizumoto et al., 2018) revealed a high sensitivity of the Sakhalin taimen eDNA detection system and its potential for biomass estimation in flowing waters. In general, however, the detectability of eDNA from aquatic organisms is affected by both biotic and abiotic factors such as the range from the target species, sizes and density of them, water temperature, pH, water current, flow volume and so on (e.g., Barnes et al., 2014; Song et al., 2017;

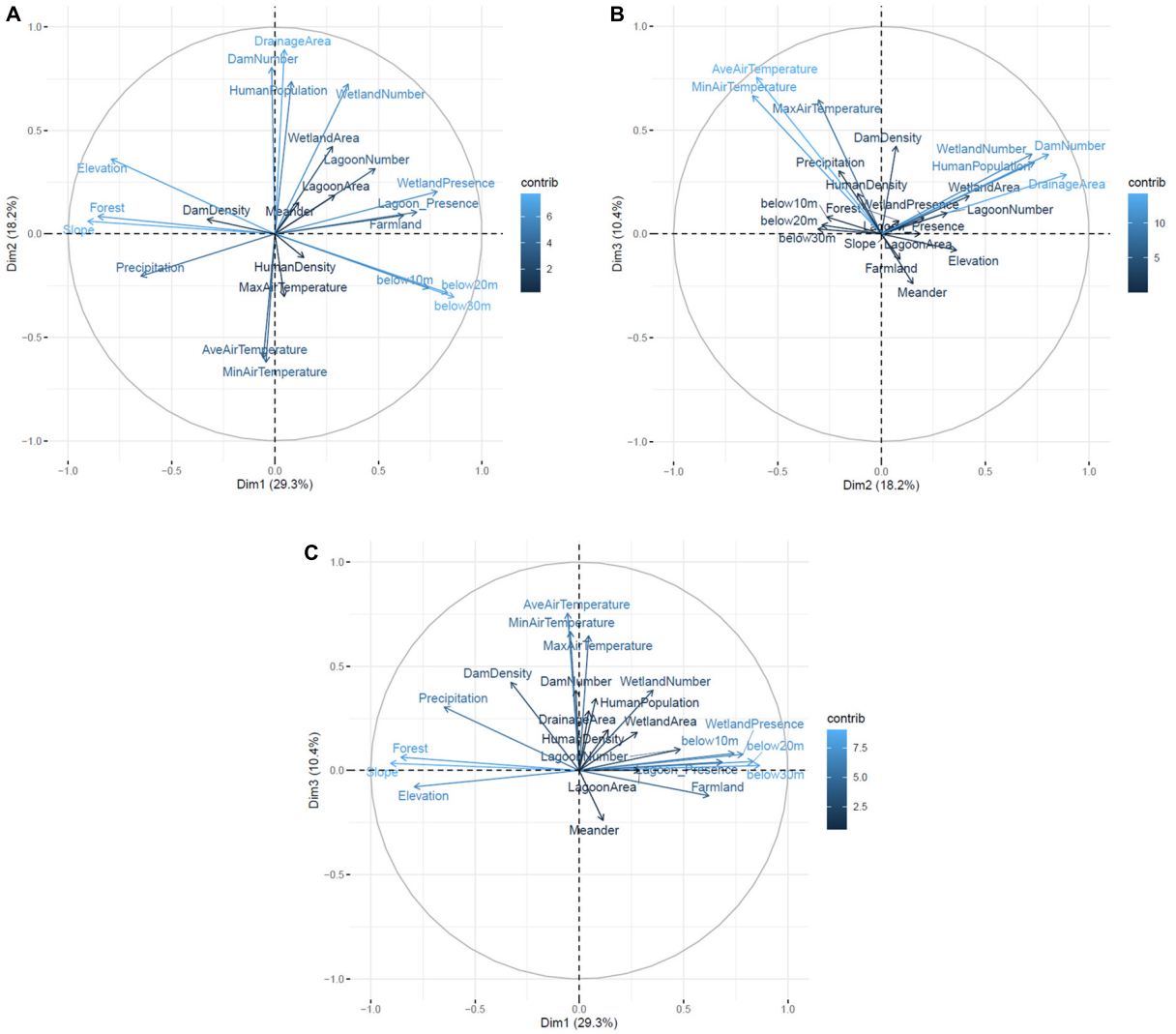


FIGURE 2 | Results of the GIS-based PCA. The top three explained variances (Dimensions 1–3) were described bidimensionally in each combination [**(A)** Dimension 1 vs. 2, **(B)** Dimension 2 vs. 3, **(C)** Dimension 1 vs. 3]. Blue colored gradation and arrow length represent the contribution percentages and arrow direction represents directions of effects onto each explained variance.

Maruyama et al., 2018; Kasai et al., 2020). Therefore, actual detectability of Sakhalin taimen's eDNA in the natural river systems would rather be limited, and we are almost certain that we missed eDNA detections from some river systems, especially from large rivers or areas with low density of Sakhalin taimen.

For the biomass estimation, eDNA concentration is considered as a relatively good proxy of biomass according to previous studies (Takahara et al., 2012; Doi et al., 2017). If we assumed that the eDNA was evenly distributed within each river and that the estimated eDNA concentrations reflected the relative biomass of this species among rivers, the distributions of Sakhalin taimen in this island should be considered to have a serious bias (> 95% in region-A and -I, **Table 3**). Interestingly, these two regions were previously suggested to support the most stable Sakhalin taimen populations in Japan (Fukushima et al., 2011). For better understanding of fine-scale distribution within/among river systems, however, further studies with more eDNA samplings and/or with conventional sampling methods would be required.

The GIS-based PCA revealed that top three dimensions (dimensions 1–3) had effects on geographical flatness, river scales, and climates (**Figure 2**). Sakhalin taimen is known as a long-living freshwater fish, and river connectivity between upstream and estuarine habitats is essential for them to complete their life history (for spawning, growing foraging and wintering. (Fukushima, 1994; Edo et al., 2000; Esteve et al., 2009; Honda et al., 2012, 2014, 2017). According to Fukushima et al. (2011), Sakhalin taimen are more likely to persist in rivers with low and flat areas like wetlands or lagoons, and average air temperature combined with agricultural development were cited as seriously contributing to their risk of extinction. The results of our 2015–2018 eDNA survey were generally consistent with

previous reports. Namely, they provided a further support for the importance of wetlands and lagoons for the protection of endangered species.

However, some parameters showed non-consistent results between the two studies (e.g., forest, the number of dams, farmland). There are two potential explanations on them; one is the temporal changes of the restrictions for their presence from the previous study (Fukushima et al., 2011) to the present one. For example, Fukushima et al. (2019) reported that there was a Sakhalin taimen population that was landlocked by a huge reservoir in northern Hokkaido, and this population potentially helped sustain the metapopulation dynamics of them at the watershed scale. In addition, these kinds of artificial barriers can be the restrictions of invasive species (Sharov, 2004; Vélez-Espino et al., 2011). In Hokkaido, Sakhalin taimen is considered to be threatened by the impacts of introduced rainbow trout (*Oncorhynchus mykiss*) because rainbow trout disturbed Sakhalin taimen's spawning beds (Nomoto et al., 2010), therefore these kinds of barriers might restrict invasion of rainbow trout. Considering these potential effects of barriers on Sakhalin taimen, the effect of dams on the presence of Sakhalin taimen that were estimated in the previous study may change with the times. However, the effects of artificial barriers should be carefully considered in conservation programs because artificial barriers can also be restrictions of Sakhalin taimen's migration (Fukushima et al., 2019). Another possible explanation is differently estimated distributions of the species between the two studies. Three rivers (river-B-5, -M-6, and -N-14) were newly suggested to hold Sakhalin taimen populations, and one river was reassessed to have avoided extinction (river-H-4). The difference might influence the determinant evaluation because of the added regional variations, for example, river-N-14 was located in southern Hokkaido where had relatively warmer air temperature (Max, Min and Ave) than the other regions in this study. In addition, river-H-4 had the largest values of drainage area and human population, and the second largest number of dams in this study. Regardless, more sensitive filtration system and/or more frequently spatio-temporal sampling are still needed to help estimating their distribution determinants more accurately.

CONCLUSION

In conclusion, we estimated the current distribution of Sakhalin taimen in Hokkaido and illustrated some of the population's main vulnerabilities. While conventional capture surveys or visual observations can provide data on individuals of a target species (e.g., body size, sex, and age), our study shows that eDNA surveys have strong advantages for collecting population-level data on endangered species (e.g., distribution, population status, and threats of extinction) both non-invasively and objectively compared to traditional methods. Assuming that estimated DNA concentrations are reflective of a target species' relative biomass, we believe that the Sakhalin taimen's biomass distribution in Japan was heavily skewed into a few regions during our sampling. In drawing these conclusions, we believe that eDNA technologies might not only help establish species' distribution, but may also show the spatial distribution of biomass and therefore infer population health over large scales. These findings should help to optimize conservation efforts for wildlife populations that are facing critical threats.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

ETHICS STATEMENT

Ethical review and approval was not required for the animal study because we did not find/touch any animals.

AUTHOR CONTRIBUTIONS

HM and HA conceived ideas and collected samples. HM wrote the manuscript. HM and TM collected data and conceived models. TM conducted data analyses and visualized analyzed data. HA acquired funding and gave final approval for publication. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the Asahi Glass Foundation, Environment Research and Technology Development Fund (4-1602 and JPMEERF20204004) of the Environmental Restoration and Conservation Agency, Japan, and JSPS KAKENHI grant no. 17H03623 (HA).

ACKNOWLEDGMENTS

We thank K. Osanai, M. Kawahara, M. Fukushima, and the other members of the Sarufutsu Itou Conservation Council for supporting our study. We also thank Dr. M. Nakaoka and the other members of the Akkeshi Marine Station of the Field Science Center for Northern Biosphere Hokkaido University, Dr. K. Miyashita and the other members of the Field Science Center for Northern Biosphere Aquatic Research Station Division of Marine Bioresource and Environmental Science, and K. Takahashi and the other members of the Council of Biodiversity in Shiribeshi Region for supporting our sampling in Hokkaido. For English corrections and comments, we thank T. E. Squires. We declare that the experiments complied strictly with the current laws in Japan.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.569425/full#supplementary-material

REFERENCES

Atkinson, S., Carlsson, J. E., Ball, B., Kelly-Quinn, M., and Carlsson, J. (2019). Field application of an eDNA assay for the threatened white-clawed crayfish *Austropotamobius pallipes*. *Freshw. Sci.* 38, 503–509. doi: 10.1086/704712

Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., and Lodge, D. M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environ. Sci. Technol.* 48, 1819–1827. doi: 10.1021/es404734p

Carlsson, J. E., Egan, D., Collins, P. C., Farrell, E. D., Igoe, F., and Carlsson, J. (2017). A qPCR MGB probe based eDNA assay for European freshwater pearl mussel (*Margaritifera margaritifera* L.). *Aquat. Conserv. Mar. Freshw. Ecosyst.* 27, 1341–1344. doi: 10.1002/aqc.2788

Doi, H., Inui, R., Akamatsu, Y., Kanno, K., Yamanaka, H., Takahara, T., et al. (2017). Environmental DNA analysis for estimating the abundance and biomass of stream fish. *Freshw. Biol.* 62, 30–39. doi: 10.1111/fwb.12846

Edo, K., Kawamula, H., and Higashi, S. (2000). The structure and dimensions of redds and egg pockets of the endangered salmonid, Sakhalin taimen. *J. Fish. Biol.* 56, 890–904. doi: 10.1111/j.1095-8649.2000.tb00879.x

Esteve, M., Gilroy, D., and McLennan, D. A. (2009). Spawning behaviour of taimen (*Hucho taimen*) from the Uur River, Northern Mongolia. *Environ. Biol. Fish.* 84, 185–189. doi: 10.1007/s10641-008-9407-x

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

Fukumoto, S., Ushimaru, A., and Minamoto, T. (2015). A basin-scale application of environmental DNA assessment for rare endemic species and closely related exotic species in rivers: a case study of giant salamanders in Japan. *J. Appl. Ecol.* 52, 358–365. doi: 10.1111/1365-2664.12392

Fukushima, M. (1994). Spawning migration and redd construction of Sakhalin taimen, *Hucho perryi* (Salmonidae) on northern Hokkaido Island, Japan. *J. Fish. Biol.* 44, 877–888. doi: 10.1111/j.1095-8649.1994.tb01261.x

Fukushima, M. (2006). The effects of damming on masu salmon and the Sakhalin taimen and the assessment of their conservation areas based on predictive habitat models. *Ecol. Civil Eng.* 8, 233–244.

Fukushima, M., Harada, C., Yamakawa, A., and Iizuka, T. (2019). Anadromy sustained in the artificially land-locked population of Sakhalin taimen in northern Japan. *Environ. Biol. Fish.* 102, 1219–1230. doi: 10.1007/s10641-019-00904-4

Fukushima, M., Kameyama, S., Kaneko, M., Nakao, K., and Ashley, S. E. (2007). Modelling the effects of dams on freshwater fish distributions in Hokkaido, Japan. *Freshw. Biol.* 52, 1511–1524. doi: 10.1111/j.1365-2427.2007.01783.x

Fukushima, M., Shimazaki, H., Rand, P. S., and Kaeriyama, M. (2011). Reconstructing Sakhalin taimen *Parahucho perryi* historical distribution and identifying causes for local extinctions. *Trans. Am. Fish. Soc.* 140, 1–13.

Gilpin, M. E., and Soulé, M. E. (1986). "Minimal viable populations: processes of species extinction," in *Conservation Biology: The Science of Scarcity and Diversity*, ed. M. E. Soulé (Sunderland, MA: Sinauer Associates Inc), 19–34.

Honda, K., Kagiwada, H., Takahashi, N., and Miyashita, K. (2012). Seasonal stream habitat of adult Sakhalin taimen, *Parahucho perryi*, in the Bekanbeushi River system, eastern Hokkaido, Japan. *Ecol. Freshw. Fish.* 21, 640–657. doi: 10.1111/j.1600-0633.2012.00585.x

Honda, K., Kagiwada, H., Takahashi, N., and Miyashita, K. (2014). Movement patterns of adult Sakhalin taimen, *Parahucho perryi*, between stream habitats of the Bekanbeushi River system, eastern Hokkaido, Japan. *Ichthyol. Res.* 61, 142–151. doi: 10.1007/s10228-013-0387-2

Honda, K., Takahashi, N., Yamamoto, K., Kagiwada, H., Tsuda, Y., Mitani, Y., et al. (2017). First documentation of detailed behaviors of endangered adult Sakhalin taimen *Parahucho perryi* in the Bekanbeushi River system, eastern Hokkaido, Japan, using bio-logging and acoustic telemetry concurrently. *Ichthyol. Res.* 64, 357–364. doi: 10.1007/s10228-016-0570-3

International Union for Conservation of Nature and Natural Resources (2020). *The IUCN Red List of Threatened Species classified in CR, EN and VU*. Available online at: http://www.iucnredlist.org/search (accessed February 20, 2020).

Iwai, N., Yasumiba, K., and Takahara, T. (2019). Efficacy of environmental DNA to detect and quantify stream tadpoles of *Odorrana splendida*. *R. Soc. Open Sci.* 6:181798. doi: 10.1098/rsos.181798

Kasai, A., Takada, S., Yamazaki, A., Masuda, R., and Yamanaka, H. (2020). The effect of temperature on environmental DNA degradation of Japanese eel. *Fish. Sci.* 86, 465–471. doi: 10.1007/s12562-020-01409-1

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

Maruyama, A., Sugatani, K., Watanabe, K., Yamanaka, H., and Imamura, A. (2018). Environmental DNA analysis as a non-invasive quantitative tool for reproductive migration of a threatened endemic fish in rivers. *Ecol. Evol.* 8, 11964–11974. doi: 10.1002/ece3.4653

Ministry of Land, Infrastructure, Transport and Tourism (2020). *National Regional Planning Bureau: Digital National Land Information*. Available online at: http://nlftp.mlit.go.jp/ksj-e/gml/gml_datalist.html (accessed February 20, 2020).

Mizumoto, H., Urabe, H., Kanbe, T., Fukushima, M., and Araki, H. (2018). Establishing an environmental DNA method to detect and estimate the biomass of Sakhalin taimen, a critically endangered Asian salmonid. *Limnology* 19, 219–227. doi: 10.1007/s10201-017-0535-x

Nomoto, K., Omiya, H., Sugimoto, T., Akiba, K., Edo, K., and Higashi, S. (2010). Potential negative impacts of introduced rainbow trout on endangered Sakhalin taimen through redd disturbance in an agricultural stream, eastern Hokkaido. *Ecol. Freshw. Fish.* 19, 116–126. doi: 10.1111/j.1600-0633.2009.00396.x

Pfleger, M. O., Rider, S. J., Johnston, C. E., and Janosik, A. M. (2016). Saving the doomed: using eDNA to aid in detection of rare sturgeon for conservation (Acipenseridae). *Glob. Ecol. Conserv.* 8, 99–107. doi: 10.1016/j.gecco.2016.08.008

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rand, P. S. (2006). *Hucho perryi: IUCN (International Union for the Conservation of Nature) 2020 Red List of Threatened Species, Version 2020-1. IUCN*. Available online at: https://www.iucnredlist.org/species/61333/12462795 (accessed February 20, 2020).

Sharov, A. A. (2004). Bioeconomics of managing the spread of exotic pest species with barrier zones. *Risk Anal.* 24, 879–892. doi: 10.1111/j.0272-4332.2004.00486.x

Song, J. W., Small, M. J., and Casman, E. A. (2017). Making sense of the noise: the effect of hydrology on silver carp eDNA detection in the Chicago area waterway system. *Sci. Total Environ.* 605, 713–720. doi: 10.1016/j.scitotenv.2017.06.255

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

The Ministry of the Environment (1995). *Report of Wetland Survey, the 5th National Survey On The Natural Environment*. Available online at: http://www.biodic.go.jp/reports3/5th/5_wetland/5_wetland.pdf in Japanese (accessed February 20, 2020).

Vélez-Espino, L. A., McLaughlin, R. L., Jones, M. L., and Pratt, T. C. (2011). Demographic analysis of trade-offs with deliberate fragmentation of streams: control of invasive species versus protection of native species. *Biol. Conserv.* 144, 1068–1080. doi: 10.1016/j.biocon.2010.12.026

Wilcove, D. S., Rothstein, D., Dubow, J., Phillips, A., and Losos, E. (1998). Quantifying threats to imperiled species in the United States. *Bioscience* 48, 607–615. doi: 10.2307/1313420

Yamanaka, H., Minamoto, T., Matsuura, J., Sakurai, S., Tsuji, S., Motozawa, H., et al. (2017). A simple method for preserving environmental DNA in water samples at ambient temperature by addition of cationic surfactant. *Limnology* 18, 233–241. doi: 10.1007/s10201-016-0508-5

Zolotukhin, S., Makeev, S., and Semenchenko, A. (2013). Current status of the Sakhalin taimen, *Parahucho perryi* (Brevoort), on the mainland coast of the Sea of Japan and the Okhotsk Sea. *Arch. Pol. Fish.* 21, 205–210.

Conflict of Interest: TM was employed by the company Pacific Consultants Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 10 November 2020
doi: 10.3389/fevo.2020.588430

Molecular Ecological Network Analyses: An Effective Conservation Tool for the Assessment of Biodiversity, Trophic Interactions, and Community Structure

Jordana M. Meyer[1], Kevin Leempoel[1], Gianalberto Losapio[1] and Elizabeth A. Hadly[1,2,3]*

[1] Department of Biology, Stanford University, Stanford, CA, United States, [2] Jasper Ridge Biological Preserve, Stanford University, Stanford, CA, United States, [3] Stanford Woods Institute for the Environment, Stanford University, Stanford, CA, United States

OPEN ACCESS

Edited by:
Hiroki Yamanaka, Ryukoku University, Japan

Reviewed by:
Haruko Ando, National Institute for Environmental Studies (NIES), Japan
Hirokazu Toju, Kyoto University, Japan

***Correspondence:**
Jordana M. Meyer
jordana7@stanford.edu

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *28 July 2020*
Accepted: *28 September 2020*
Published: *10 November 2020*

Citation:
Meyer JM, Leempoel K, Losapio G and Hadly EA (2020) Molecular Ecological Network Analyses: An Effective Conservation Tool for the Assessment of Biodiversity, Trophic Interactions, and Community Structure. Front. Ecol. Evol. 8:588430. doi: 10.3389/fevo.2020.588430

Global biodiversity is threatened by the anthropogenic restructuring of animal communities, rewiring species interaction networks in real-time as individuals are extirpated or introduced. Conservation science and adaptive ecosystem management demand more rapid, quantitative, and non-invasive technologies for robustly capturing changing biodiversity and quantifying species interactions. Here we develop molecular ecological network analyses (MENA) as an ecosystem assessment tool to address these needs. To construct the ecological network, we used environmental DNA from feces to identify the plant and mammal diet of two carnivores: puma (*Puma concolor*) and bobcat (*Lynx rufus*); two omnivores: coyote (*Canis latrans*) and gray fox (*Urocyon cinereoargenteus*); and two herbivores: black-tailed deer (*Odocoileus hemionus*) and black-tailed jackrabbit (*Lepus californicus)* in a well-studied Californian reserve. To evaluate MENA as a comprehensive biodiversity tool, we applied our framework to identify the structure of the network, patterns of trophic interactions, key species, and to assess its utility in capturing the biodiversity of the area. The high dietary taxonomic resolution enabled the assessment of species diversity, niche breadth and overlap. The network analysis revealed a dense ecological network with a high diversity of weakly connected species and a community that is highly modular and non-nested. The significant prevalence of tri-trophic chain and exploitative competition patterns indicates (i) the removal or reintroduction of a top predator would trigger a trophic cascade within this community, directly affecting their prey and indirectly the plant communities, and (ii) the potential impact of indirect effects between two predators that consume the same prey. These results suggest that the recent resurgence of puma in the study area may impact the herbaceous and woody vegetation and the population size of other predators. This effect of fluctuating predator populations and plant communities could be predicted through MENA's fine-scale assessment of the diet selection and the identified keystone species. Although just using a subset of species, MENA more rapidly,

accurately, and effectively captured the broader biodiversity of the area in comparison to other methodologies. MENA reconstructed and unveiled the hidden complexity in trophic structure and interaction networks within the community, providing a promising toolkit for biodiversity and ecosystem management.

Keywords: food web, network analysis, diet ecology, eDNA, DNA metabarcoding, conservation

INTRODUCTION

Anthropogenic impacts threaten global biodiversity, as species and their networks of interactions (ecological networks) are forced to adapt or be lost (Gonzalez et al., 2011; Dirzo et al., 2014; Munguía et al., 2016; Start et al., 2018). As a consequence of species loss and invasion, ecological networks, such as food webs and pollinator networks, are restructuring and species interactions are rewiring in real-time, fundamentally impacting whole ecosystems and their functions (Bartley et al., 2019; Daam et al., 2019). To comprehend these altered ecological networks, conservation biologists are increasingly focusing on extinction cascades caused by weakened or lost species interactions (Baiser et al., 2012; Valiente-Banuet et al., 2015; Losapio and Schöb, 2017) and are motivated to take a multitrophic perspective to address biodiversity and ecosystem functioning (Eisenhauer et al., 2019). However, we are in need of more rapid, non-invasive, and quantitative technologies for biodiversity assessments and monitoring for the implementation of adaptive management strategies.

Counting species (alpha diversity) rather than their interactions is often more tractable, however, it is not nearly as informative for understanding how ecosystems function (Jordan and Scheuring, 2004; McCann, 2007). Ecosystems are composed of thousands of species interactions (billions if including microbes) that directly or indirectly impact biological populations (Kéfi et al., 2016) and contribute to network robustness (Losapio and Schöb, 2017). Recent advances in network analysis enable the interpretation of large numbers of species interactions and assessment of the predictability of dynamic ecological systems (Dale and Fortin, 2010; Delmas et al., 2019). Yet, building robust ecological networks, particularly in response to perturbations and across communities that are comprised of different species, is challenging given the difficulty of observing trophic events and quantifying the strength of interactions between species. The advent of high-throughput sequencing (HTS) makes this kind of research remarkably more tractable (Roslin et al., 2019).

Innovations in HTS, and, more specifically, DNA metabarcoding, enable accurate and cost-effective biodiversity assessments at a level of taxonomic coverage and precision previously unavailable (Ji et al., 2013; Clare, 2014; Deiner et al., 2017; Pawlowski et al., 2018; Bush et al., 2019; Makiola et al., 2020). Recent studies have used metabarcoding techniques to investigate the feeding ecology of carnivores (Shehzad et al., 2012; Torre et al., 2013; Walsh, 2015; Xiong et al., 2017), herbivores (Soininen et al., 2009; Czernik et al., 2013; Kartzinel et al., 2015; Coverdale et al., 2016; Iwanowicz et al., 2016; Erickson et al., 2017; Pansu et al., 2019), and omnivores (De Barba et al., 2014; Robeson et al., 2018; for a review see Sousa et al., 2019). For example, Kartzinel et al. (2015) measured dietary niche partitioning among large herbivores and was able to address how generalist consumer species coexist on a limited range of resource types because metabarcoding allowed for unprecedented plant identification. While many of these studies targeted the diet of a single species or a trophic guild, few compared across mammal species and multiple trophic levels, a shortcoming that hinders our full knowledge of complex ecological processes (Eisenhauer et al., 2019). Multitrophic perspectives provide a better understanding of ecological and evolutionary processes than do typical pairwise interactions between trophic levels (Abdala-Roberts et al., 2019). Despite this, our knowledge of ecosystem function in multitrophic communities is limited to date (Eisenhauer et al., 2019).

Constructing multitrophic ecological networks using traditional methods is difficult and labor intensive, especially in poorly studied and highly diverse systems (Derocles et al., 2018). Non-invasively collected DNA from the environment (eDNA) [i.e., a complex mixture of genomic DNA from different organisms found in soil, water, or feces (Taberlet et al., 2018)] may help to construct complex ecological networks by rapidly detecting trophic interactions that would otherwise be impossible to observe, and to do so more quickly than traditional methods (Ruppert et al., 2019). Only recently has molecular network analysis been applied to the assessment of ecological systems. Metabarcoding studies in this context have focused on varying scales: the diet of European hake (Riccioni et al., 2018); coral reef fish (Casey et al., 2019); freshwater ecosystems (Compson et al., 2019); aquatic macroinvertebrates (Bush et al., 2019); and bat (Clare et al., 2019), forest (Evans et al., 2016), and herbivore (Kartzinel et al., 2015) communities. However, to the best of our knowledge (see Sousa et al., 2019 for a review), to date a metabarcoding study has yet to use fecal eDNA to reconstruct ecological networks and unravel trophic interactions among carnivores, omnivores, and herbivores within a terrestrial mammal community to inform biodiversity assessments.

Here we combine DNA metabarcoding and network-based approaches in a non-invasive molecular ecological network analyses (MENA) (Deng et al., 2012) and demonstrate its utility to assess biodiversity, trophic interactions, and structure within a mammalian community. We construct a food web by identifying the diets of mammals using HTS of eDNA from feces across three feeding guilds – two carnivores: puma (*Puma concolor*) and bobcat (*Lynx rufus*); two omnivores: coyote (*Canis latrans*) and gray fox (*Urocyon cinereoargenteus*); and two herbivores: black-tailed deer (*Odocoileus hemionus*) and black-tailed jackrabbit (*Lepus californicus)*. To evaluate MENA

as a comprehensive biodiversity assessment tool, we apply our framework to answer the following questions: (i) What is the structure of the network and who are key species? (ii) Are specific patterns of trophic interactions (i.e., trophic cascade, omnivory, and exploitative competition) occurring and prevailing within the community?; (iii) Does MENA provide an accurate portrait of biodiversity? We validate findings with long-term studies in the study area and discuss how these molecular and network metrics provide increasingly valuable information and a rapid bioassessment tool for conservation decision making.

MATERIALS AND METHODS

Study Site and Feces Collection

Jasper Ridge Biological Preserve (JRBP) is a small (5 km^2) protected area on the urban fringe of Silicon Valley, a highly urbanized region in California, United States. It is located in the eastern foothills of the Santa Cruz Mountains and surrounded by various land use types. JRBP has a Mediterranean-type climate, consisting of warm, dry summers and cool, wet winters, and averaging temperatures of 5–22°C across the year. The mean annual precipitation is 605 mm (Zavaleta and Kettley, 2006). JRBP has been an active site of research for over a century, protected as a preserve since 1973, and the fauna and flora of the preserve have been well-documented. Dominant vegetation types include grassland, woodland, chaparral, coastal scrub, marshes, and forests, with the most common vegetation type being oak (Perea et al., 2017).

A long-term camera-trapping effort (2009–2018) at JRBP elucidated dynamic shifts within the predator and herbivore populations, triggered when an apex predator, the puma, became resident after at least a decade of very low abundance (Leempoel et al., 2019). The (re)establishment of the puma coincides with a change in coyote behavior and abundance, which in turn, opened a niche for the smaller, omnivorous gray fox. JRBP thus provides an ideal opportunity to test the MENA tool in identifying biodiversity and patterns of trophic interactions, because the flora, fauna and dynamic shifts within this system are known.

Fecal samples (scats) from puma, coyote, bobcat, gray fox, black-tailed deer, and black-tailed jackrabbit were collected across all six habitat types in JRBP within 2-week survey windows [following (Murphy et al., 2007; Ruell et al., 2009)] in the dry season (October 26 – November 12, 2017) and wet season (April 4 – 17, 2018). Before the start of the survey, all scat was removed from the trails to ensure the age of the collected samples was all less than 2 days old. This kept freshness of the sample consistent and minimized fecal DNA degradation. Sampling transects were non-random and predator scat was collected along 34 paths (trails 17 km; roads 7 km) throughout JRBP (**Supplementary Figure S1**). Camera trapping demonstrates that predators use these pathways at a high frequency (Kohn et al., 1999; Leempoel et al., 2019). During collection, one-half of the scat sample was left *in situ* to ensure that the relevant inter- and intra-species scent cues were not disturbed. Scat samples were collected into a sterile bag, using gloves to avoid contamination, and GPS coordinates and metadata were recorded at the collection site. All samples were stored at −20°C until DNA extraction was performed. In total, + 175 km of trails were traversed through all dominate habitats over 12 collection days in each season (finding scat on 19 of the 24 traversing days).

A combination of morphological identification (i.e., size, shape, composition) and DNA sequencing (e.g., the mammal primer confirmed all predator identifications through the most abundant predator reads per sample) enabled us to determine the identity of species that deposited the scat. Multiple jackrabbit fecal samples were genotyped to confirm the identification of the collected samples, which are visually much larger, but similar in shape, to scats of the smaller brush rabbit (*Sylvilagus bachmani*).

Fecal DNA Extraction and Metabarcoding

We conducted all DNA extractions at Stanford University in a sterilized laminar airflow hood to avoid contamination. Before DNA extractions, fecal samples were thawed, homogenized, and processed (∼0.2 g) utilizing the Zymo Quick-DNA Fecal/Soil Microbe Miniprep Kit (Kartzinel et al., 2015). Samples were processed in small batches (∼14) with an extraction blank to monitor for potential cross-contamination in the laboratory. The eluted DNA was quantified using a Nanodrop 2000 (Thermo Fisher Scientific Inc.) and stored at −20°C until used for PCR.

We used DNA metabarcoding on the extracted DNA to characterize carnivore, omnivore, and herbivore diets, by quantifying large numbers of short, unique regions of DNA extracted from individual fecal samples (Pompanon et al., 2012). Two different primer pairs were used to amplify mammal and plant DNA. All primers were modified with the Illumina adaptor preceding the target primers and separated by 6-N spacers as designed by Ushio et al. (2017).

Mammals

We use the MiMammal-U metabarcoding primers for the 12S mtDNA gene targeting 210 bp amplicons because it can amplify and distinguish DNA from a diverse and wide range of mammals that are well represented in public databases (MiMammal-UF: *GGGTTGGTAAATTTCGTGCCAGC* and MiMammal-UR: *CATAGTGGGGTATCTAATCCCAGTTTG*) (Ushio et al., 2017). We compared our primer reference database to the list of known JRBP mammal species (46 species) and found that these primers identify 96% of the mammals present in JRBP to the genus and species level and 4% to the family level (Leempoel et al., 2020). The PCR comprised 20 μL: 10 μl of GoTaq® Colorless Master Mix, 1 μL of each primer (5 mM), 4 μL of DNA template, and 4 μL of water. Cycling conditions used initial denaturing at 95°C for 10 min, followed by 35 cycles of denaturing at 95°C for 30 s, annealing at 60°C for 30 s and extension at 72°C for 10 s.

Plants

The most widely applied DNA barcode for herbivore diet analysis is the P6 loop of the chloroplast trnL (UAA) gene (Taberlet et al., 2007). However, it has a lower taxonomic

coverage than a recently designed second internal transcribed spacer of nuclear ribosomal DNA (ITS2) primer (Moorhouse-Gann et al., 2018). For this study, we targeted 200–387 bp amplicons of the ITS2 region for the amplification of plants in the diet using primers UniPlantF: *TGTGAATTGCARRATYCMG* and UniPlantR: *CCCGHYTGAYYTGRGGTCDC* (Moorhouse-Gann et al., 2018). We compared our primer reference database to the list of known JRBP plant species (762 species) and found that these primers identify 88% of the plants present at JRPB to the species and genus levels, 7% only to the family level, and 5% were not found in the database (GenBank). The first PCR comprised of 20 μL reactions using GoTaq® Colorless Master Mix, 0.6 μL (10 mM) of each primer, and 6 μL of DNA template and 2 μL of water. Thermocycling followed a program of initial denaturing at 95°C for 10 min, followed by 35 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s, with a 2-min final extension at 72°C.

The integrity of DNA extracted by each primer was assessed by gel electrophoresis. Specifically, 3 μL of each DNA extract was analyzed in a 2% agarose gel and was visualized by UV illumination. If positive, the PCR products were cleaned with the Qiagen PCR Purification Kit (Valencia, CA). For the two-step PCR, appropriate Illumina barcodes were ligated to each sample as an index tag for each unique sample (Ushio et al., 2017). The index PCR was performed as a 20 μl reaction: 10 μl of Amplitaq Gold360 Mastermix reactions (with 2.5 mM MgCl2,200 lM each dNTP, 0.1 mg/mL BSA, 4% DMSO), 1 μl (of each primer), 3 μl of purified DNA amplicons and 6 μl of H_20. Cycling conditions used initial denaturing at 95°C for 10 min, followed by 12 cycles of denaturing at 95°C for 30 s, annealing at 60°C for 30 s and extension at 72°C for 10 s.

The indexed secondary PCR products were quantified using a Fragment Analyzer™ (Automated CE System from Advanced Analytical Technologies), normalized to equimolar concentrations and pooled together before purification using QIAquick PCR Purification Kit (Qiagen). Sequencing was performed on a MiSeq platform using the Reagent Kit Nano v3 for 2 × 300 bp (plant primers) and 2 × 150 bp (mammal primers) paired-end reads (Illumina, San Diego, CA, United States) and run at the Stanford University PAN Facility. PhiX DNA spike-in control (18–30%) was added to improve the data quality.

Sequence Filtering and Taxonomic Assignment

We used a series of conservative filtering steps developed by Leempoel et al. (2020), to retain as much of the "true" DNA diversity present in the scat samples as possible. We used software packages in Obitools (Boyer et al., 2016) and R version 3.6.3 (R Core Team, 2020) for demultiplexing and quality control. Each sequence was assigned to its sample of origin based on exact matches to both index identifier. Sequences were paired with Obitools *illuminapairedend* and aligned sequences with a score of <40 were discarded (Kartzinel et al., 2015). Quality scores of paired sequences were checked using FastQC, prior to adapter trimming (with a mismatch tolerance with primers of 10%) in *Cutadapt* (Martin, 2011), and low-quality sequences <Q30 were removed. After assignment of sequences to their corresponding samples, we used *obiuniq* to dereplicate reads into unique sequences, eliminated potential PCR and sequencing errors with *obiclean*, and kept only sequences occurring at least 10 times. See Boyer et al. (2016) for a more detailed explanation. *Obiclean* was applied sample by sample, with a maximum of one difference between two-variant sequences and a threshold ratio between counts of one, meaning that all less-abundant sequences were considered as variants. Only sequences with "head" (true sequences or chimera product and can have multiple variants) or "singleton" (either true sequences or chimeras but are not related to any other sequences) status in at least one sample were kept. Further, sequences whose status in the global dataset were more commonly "internal" (amplification/sequencing errors) than "head" or "singleton" were discarded (Giguet-Covex et al., 2014).

The remaining sequences were matched in Obitools against the reference databases built using *EcoPCR* to identify the molecular operational taxonomic units (mOTUs) (Blaxter et al., 2005; Ficetola et al., 2010). First, we downloaded all standard sequences for vertebrates and plants from EMBL ([1]release 141) and converted the recovered file to *EcoPCR* format. *EcoPCR* was then used to simulate an *in silico* PCR, using the two primer pairs and maximum three mismatches, and a minimum and maximum length identical to the length of each metabarcode. The taxonomic assignment of mOTUs was performed using *Ecotag*, keeping only sequences with an identity ≥ 95%, and further inspecting and revising taxonomic assignments to ensure validity.

We removed sequences with relative read abundance (RRA) < 0.001% within samples to reduce the likelihood of including data that might be the result of cross-contamination (Kartzinel et al., 2015). Obvious contaminants (i.e., human DNA) and counter-marking species mOTUs were removed – e.g., fox and bobcat frequently urinate or defecate on each other's scats. Non-target mOTUs were removed because the MiMammal primer is known to identify bird sequences in the database (Ushio et al., 2017) and we identified 7 different bird mOTUs (**Supplementary Table S4**) (resulting removal of one coyote and bobcat scat) and the ITS2 primer amplified fungi and bacteria (Moorhouse-Gann et al., 2018). *Canis lupus familiaris* (domestic dog) was present in many fox samples yet the sequence was a 95% match to *Vulpes vulpes* (red fox) in the database, therefore we believe reads were misidentified and we removed this mOTU from diet analyses. Finally, all sequences with > 1% of the total reads in the negative controls were removed from our samples, a threshold that is 10 times more conservative than other recent studies (e.g., Siegenthaler et al., 2019).

Diet Composition and Biodiversity Analysis

We converted the filtered sequence read data into read abundance and occurrence data to examine the utility of both data types for food web construction. Three methods were used

[1]http://ftp.ebi.ac.uk/pub/databases/embl/release/std/

to quantify diet composition, as suggested by Deagle et al. (2019), including Frequency of Occurrence (FOO), which is calculated as the number of occurrences divided by the total number of samples; Percent of Occurrence (POO), rescaled to account for all food items; and the Relative Read Abundance (RRA) defined as the proportion of unique Illumina sequence reads in a sample divided by the final (i.e., after quality control) number of sequence reads in that sample. Fecal samples were sequenced separately for the plant and mammal DNA; therefore, we analyzed each independently for RRA and combined the FOO results in the case of omnivore diets.

Diet items were categorized into functional groups. Plants were identified as grasses, herbs, or woody vegetation (trees and shrubs), while small mammals were defined as species with an average adult body mass < 1.5 kg and large animals as those with a body mass > 1.5 kg. The mean percent of RRA was used to compare diet type (e.g., grass vs. trees vs. small mammals) among the focal species (Pansu et al., 2019).

To determine if there is an empirical relationship in the predator–prey mass ratio in JRBP, we compared predator body size (in grams) to prey body size (both natural log transformed) at the individual level as suggested by Nakazawa (2017) using Pearson's correlation coefficient. Predators can take down prey of a certain size or smaller and scavenging carcasses does not involving killing of prey. We have evidence from camera trap imagery that deer carcasses are being scavenged by bobcats and gray fox within JRBP, therefore these correlations were performed with and without the deer.

To capture important baseline information about the food web dynamics, we compared different diversity, niche breadth and diet overlap metrics. We determined diet diversity (raw richness, Shannon diversity), niche breadth (individual, Levins'), niche overlap (Pianka's niche overlap), and compared consumed prey species diversity (at mOTU level) following Razgour et al. (2011). The Shannon's diversity index (H) accounts for both abundance and evenness of the species present in the diet and was calculated for each species diet using the *vegan* package (Oksanen et al., 2019) in R version 3.6.3 (R Core Team, 2020). A species' niche breadth describes the suite of resources that it can use (Gaston et al., 1997). We determined how uniformly resources (mOTUs) are being utilized by each species using the standardized Levins' measure of niche breadth index (B_A) (Razgour et al., 2011; Lyngdoh et al., 2014):

$$B = \left[\frac{1}{\sum p_i^2}\right]$$

Standardize as: $B_A = \frac{B-1}{n-1}$

where p_i is the proportion of fecal samples in which the mOTU i was found and n is the number of possible mOTUs in the diet.

Pianka's adaptation of the niche overlap (O_{jk}) metric was used to determine dietary overlap among all pairs of target species (Pianka, 1973; Woodward and Hildrew, 2002; Brown et al., 2014; Pansu et al., 2019):

$$O_{jk} = \left[\frac{\sum_i^n p_{ij} p_{ik}}{\sqrt{\sum_i^n p_{ij}^2 \sum_i^n p_{ik}^2}}\right]$$

where p_{ij} is the proportion of prey species i in one carnivore species j diet, p_{ik} is the proportion prey species i in another carnivore species k diet, n = total number of available prey species. Ojk = 0 represents no overlap, whereas a value of O_{jk} = 1 represents complete overlap in prey species between all carnivore species.

To determine whether the number of collected fecal samples captured enough of the diet richness for this study, we used the function *specaccum* in the R package *vegan* to determine the mOTUs accumulation curve (without replacement) (Gotelli and Colwell, 2001). We used EcoSim null models (version 1[2]) to test whether the extent of niche overlap is greater than expected by chance (Gotelli et al., 2015). We generated 1,000 simulated matrices of randomized mOTU diet composition, using the randomization algorithm 3, where the niche utilization values are reshuffled within each row of the matrix to detect non-random niche overlap patters (Winemiller and Pianka, 1990). The observed niche overlap (O_{jk}) is then compared to the simulated niche overlap values (Gotelli et al., 2015). For all of these pairwise comparisons, we deemed niche overlap to be significant when the observed value was greater than at least 99% of the simulated values.

We performed a non-metric multidimensional scaling (nMDS) ordination based on a Bray–Curtis dissimilarity matrix using the occurrences and the FOO of mOTUS to examine the patterns of niche space across different species (permutations = 999, trymax = 500, k = 2) (Kartzinel et al., 2015; Casey et al., 2019). We also performed principle component analysis (PCA) on the FOO of mammal and plant diet taxa to detect differences in the diet composition among species and to investigate whether certain taxa were more commonly found in certain diets.

Molecular Ecological Network Analyses

Molecular ecological network analyses consisted of three main steps: (i) mapping species interactions, (ii) quantifying network structure, and (iii) quantifying specific interaction patterns.

First, we built a JRBP food web as a unipartite network, which consists of one set of nodes where two species can be connected through trophic interactions (Delmas et al., 2019). Because network metrics can be impacted by mOTU identification protocols (Clare et al., 2019), we were conservative with the filtering process. Thus, unique sequence reference number mOTUs were grouped at species, genus, or family level, as this approach circumvents the potential taxonomic misclassification (Lupatini et al., 2014). The network was directed from predator to prey and interactions were weighted using the FOO values. To calculate network complexity and visualize the JRBP food web, we used the software Gephi ([3]Bastian and Heymann, 2009). The Yifan Hu layout algorithm was used for the construction of the directed network and belongs to the category of force-directed algorithms (Force Atlas and Fruchterman Reingold algorithms) (Hu, 2006).

[2]http://grayentsminger.com/ecosim.htm

[3]http://gephi.org

Second, we quantified overall network complexity and structure calculating the following network metrics: *Connectance* (proportion of realized interactions among all possible ones) to quantify the level of arrangement of a network; *Average degree* (average number of links per species) to understand the average level of specialization of the network; *Average clustering coefficient* (degree to which the neighbors of a node are connected) to estimate groupings in closely connected subsets; and *Path length* (average distance between any pair of nodes) to determine shortest distance between species (Delmas et al., 2017; Landi et al., 2018).

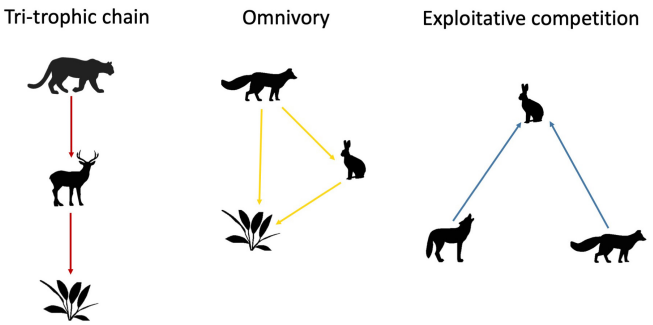
To measure the relative importance of each taxon within the network we calculated: (i) *Weighted in-degree centrality*, which also indicates the degree to which losing a central species from the network will trigger secondary extinctions and impact the community (Elhesha et al., 2017) and (ii) the *Eigenvector centrality*, which measures the node's importance within the network while accounting for the importance of its neighbors, representing potential keystone species (Allesina and Pascual, 2009). Keystone species play a structuring role within the food web, strongly influencing the abundances of other species (Power et al., 1996). The higher the centrality values within the network, the greater the functional role within the network (Allesina and Pascual, 2009).

To better understand the structure of the JRBP food web, we measured modularity and nestedness. Modularity is defined as the degree to which networks are organized in discrete groups, i.e., modules. A modular network is characterized by species belonging to the same module having high connection among themselves and few connections to species belonging to other modules (Moore et al., 2017; Landi et al., 2018; Ma et al., 2020). Modularity and module partitions were computed by simulated annealing (Doulcier et al., 2016) using *rnetcarto*, which allowed us also to determine which species were most connected among modules and to calculate *z*-values for within-module connectivity. Nestedness is defined as the degree to which more specialist species interact only with subsets of species that interact with more generalist species (Jonsson, 2001; Bastolla et al., 2009). The role of nestedness for biodiversity depends on the type of species interactions. In mutualistic networks, nestedness enhances the number of coexisting species, but in food webs nestedness increases niche overlap among consumers and thus may prevent their coexistence (Kondoh et al., 2010). Nestedness (η) was calculated (Bastolla et al., 2009; Losapio et al., 2019) for unipartite networks as:

$$\eta = \sum_{i<j} \frac{n_{ij}^{(P)}}{\min(n_i^{(P)}, n_j^{(P)})}$$

where n_{ij} is the number of prey species between two predator species i and j and min (n_i, n_j) is the smaller of the two values. Values of nestedness η range between 0 (perfectly non-nested, full trophic complementarity) and 100 (perfectly nested).

Thirdly, we identified specific interaction patterns using network motif analysis (Milo et al., 2002). Specifically, we focused on patterns of: (i) tri-trophic chains (e.g., predator that consumes prey who feed on plants), (ii) omnivory (species at the top of a food chain that feed on both plants and animals), and (iii) exploitative competition (two predators competing for the same shared prey) (Giling et al., 2019). These three motif patterns are represented as:



We used figure the RANDESU motif finder algorithm of *igraph* (Csardi and Nepusz, 2006) to find and count these subnetworks. Then, we compared the JRBP food web motifs with random expectation. Notice that network links were directed from predators to prey, highlighting the structure of "who eats whom" rather the "matter and energy flow." For this reason, the motif of exploitative competition involving two predators and one prey can resemble apparent competition in theory. However, apparent competition involves two prey and one predator (Wootton, 1994). Thus, we opted for a biological definition of network motifs involving indirect interactions between two predators as exploitative competition.

To assess the significance of nestedness, modularity, trophic cascade, omnivory and exploitative competition, we used randomizations of the empirical JRBP food web. Our null model randomized the food web probabilistically maintaining total number of interactions and species-specific frequencies (Jonsson, 2001; Oksanen et al., 2019). To assess statistical significance for each of these metrics we assessed whether (i) the empirical food web metric was higher or lower than the 95% CI expected by chance, and (ii) the probability, *p*, that a randomization was equal or higher than observed empirically. All analyses were conducted in R version 3.6.3 (R Core Team, 2020).

Bioassessment Comparison

Mammal diversity recovered from fecal DNA was compared to an expert-curated list of known mammals in the region, and both camera trap soil eDNA surveys at JRBP. The fauna and flora of JRBP are well documented, including 762 plant species (Oakmead Herbarium, 2019) and 46 mammal species (JRBP, 2019). Camera trap and soil eDNA data sets and their analysis are described in Leempoel et al. (2020) and Leempoel et al. (2019). Briefly, during the time of this study, 18 wireless camera traps were continuously recording wildlife along trails. We used images recorded from October 15, 2017 – April 20, 2018 to overlap both fecal collection sessions for a total of 2,754 capture days (153 days × 18 cameras). To determine if some scats were over- or under-represented, we tested for a correlation between the number of capture days per species and the number of scats. The soil eDNA study was conducted in October 2017, contemporaneously with

some scat collection. It relied on 12 soil samples sampled in front of 6 of the camera traps, and used the same MiMammal-U 12 s metabarcode. We compared mammal species identified in this soil eDNA study with those identified in the scat samples.

RESULTS

We collected 158 fecal samples over the two sampling periods, consisting of 15 puma, 12 coyote, 31 bobcat, 71 gray fox, 14 deer, and 15 jackrabbit samples. The majority of the samples (87%) were collected in the dry season (dry season = 137 and wet season = 12), which may be a result of seasonal usage of the JRBP by animals, seasonal differences in trail use, and/or seasonal differences in scat preservation. We combined all samples for further analysis to represent the temporal averaging of the focal species' diets and plotted sample-based species-accumulation curves, all of which are approaching asymptotes (**Supplementary Figures S2A,D**), indicating adequate sampling of guild diets for this study.

Mammal and Plant DNA Metabarcoding

We successfully sequenced either mammal DNA, plant DNA or both from 113 fecal samples post filtering (**Table 1** and **Supplementary Table S1** for sample size filtering). A total of 75% of gray fox samples contained plant DNA ($n = 36$) and 63% contained mammal DNA ($n = 30$). Two of the coyote samples had both plant and mammal DNA, with one containing plant DNA only. We retained all mOTUs as representatives of the number of species that were both in the environment (biodiversity assessment) and being consumed (network assessment) at the time of collection.

From a combined 24-day sampling period, we identified a total of 212 known mOTUs with unique reference sequence numbers from 31 carnivore, 57 omnivore, and 25 herbivore fecal samples (**Supplementary Table S2**). These mOTUs represent 55 families and 119 genera. We combined multiple mOTUs identified to the same genus or family, making up 167 individual mOTUs (identified to each taxonomic level: 85 species/sub-species, 67 genera, 10 tribes/sub-tribes, and 5 families/sub-families) (**Supplementary Table S3**). We classified 91% of the mOTUs to the genus or species/sub-species level.

The mtDNA 12S gene was sequenced to assess the mammal diet of four predator species. From the 102 samples sequenced, a total of 20,373,994 raw reads were generated with 4,459,979 reads remaining after removing host and contaminant DNA. Host DNA made up the majority of the reads for puma 63 ± 0.12%, bobcat 89 ± 0.04%, coyote 81 ± 0.09%, and gray fox 93 ± 0.02%, and therefore able to identify any unidentified scat. After taxonomic assignment, a total of three mOTUs (genera *Equus, Rattus*, and *Felis*) were removed because their abundance in the extraction-negative control was > 1% of the total reads for those mOTUs, suggesting upstream contamination. We removed samples that consisted only of host reads (i.e., samples showing no prey species other than host DNA) or samples deemed contaminated, leaving 69 samples for downstream analysis (**Supplementary Table S1**).

We identified a total of 16 unique mammal mOTUs from the scat samples of the predator species for downstream diet analysis (**Table 2**). Predator diet diversity comprised 11 families and 13 genera. If an mOTU was identified to the genus-level and that genus was monotypic in the San Francisco Bay Area, then cf. (*conferre*) was indicated and the species level identification was included (i.e., *Neotoma* cf. *fuscipes*). Therefore, 75% (12/16) of the mOTUs were identified to species.

The ITS2 marker was sequenced to assess the plant diet of the two herbivore and two omnivore species (Moorhouse-Gann et al., 2018; **Table 1**). From the 64 samples sequenced, a total of 1,970,916 post filtering and a final 1,813,698 reads were used for downstream analysis. Plant DNA was identified as 196 unique mOTUs to the family level and below, marked by unique reference sequence numbers (**Supplementary Table S5**). The mOTUs represented 44 families and 106 different genera and 77 species of plants. Of the 196 unique reference sequence numbers, there were 151 individual mOTUs after combining the same family and same genus together (identified to each taxonomic level: 77 species, 61 genera, 10 tribes, and 3 families) (**Supplementary Table S6**). The ITS2 primers were specific in identifying plants to the species level (51% of mOTUs) and genus level (40% of mOTUs).

Diet Composition and Richness

When evaluating diet by functional groups, we found that puma consumed mainly large mammals (**Figure 1** and **Table 3**), primarily black-tailed deer, and frequently augmented with small mammals and single occurrences of coyote, raccoon (*Procyon lotor*), and feral pig (*Sus scrofa*). The next largest predator, the coyote, consumed an even mix of large mammals,

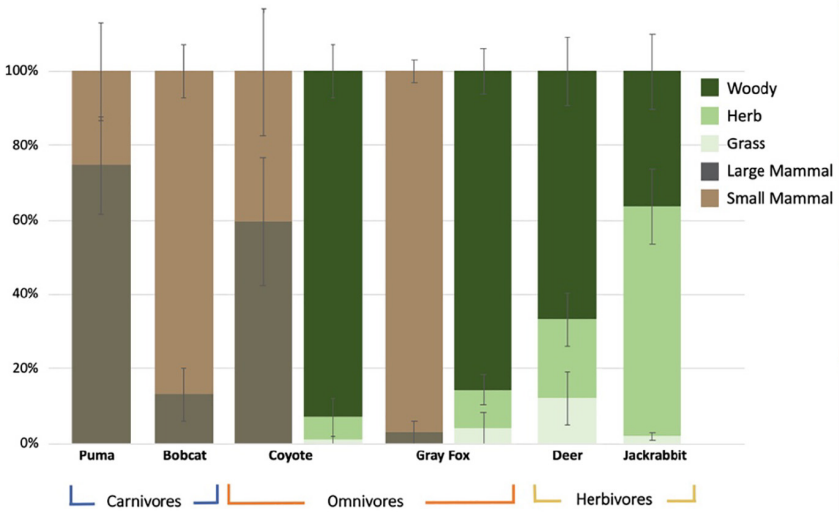
TABLE 1 | *A priori* guild assignment according to the plant and mammal DNA, represented as the percent of samples that contained only mammal, only plants and both.

Focal Species	Common name	*A priori* guild assignment	Mammal	Mammal & Plants	Plant
Puma concolor	Puma	Carnivore	100% (12)	–	–
Lynx rufus	Bobcat	Carnivore	100% (19)	–	–
Canis latrans	Coyote	Omnivore	67% (6)	22% (2)	11% (1)
Urocyon cinereoargenteus	Gray fox	Omnivore	25% (12)	38% (18)	38% (18)
Odocoileus hemionus	Black-tailed deer	Herbivore	–	–	100% (12)
Lepus californicus	Black-tailed jackrabbit	Herbivore	–	–	100% (13)

The final sample size of sequenced samples is represented in brackets ().

TABLE 2 | Mammal diet mOTUs found within scats from four predator species collected from JRBP, expressed by the Count (number of scats containing prey species); FOO (frequency of occurrence = number of occurrences/total scat sample size); POO (percentage of occurrence = number of occurrences/total number of prey species across all samples); and RRA (mean relative read abundance = number of reads per species/total number of reads per scat sample).

Functional Group	mOTUs	Common Name	Family	RefSeq	Puma (*n* = 12)				Bobcat (*n* = 19)				Coyote (*n* = 8)				Gray Fox (*n* = 30)			
					Count	FOO	POO	RRA	Count	FOO	POO	RRA	Count	FOO	POO	RRA	Count	FOO	POO	RRA
	Canis latrans	Coyote	Canidae	DQ480509	1	0.08	5.9%	0.08 ± 0.08	0	0.00	0.0%	0 ± 0	0	0.00	0.0%	0 ± 0	0	0.00	0.0%	0 ± 0
	Odocoileus hemionus	Black-tailed Deer	Cervidae	AF091708	7	0.58	41.2%	0.58 ± 0.14	2	0.11	6.3%	0.02 ± 0.02	1	0.13	2.6%	0.09 ± 0.09	1	0.03	0.3%	0.03 ± 0.03
Large Mammals	*Ondatra zibethicus*	Muskrat	Cricetidae	KU177045	1	0.08	5.9%	0.0 ± 0.0	0	0.00	0.0%	0 ± 0	0	0.00	0.0%	0 ± 0	0	0.00	0.0%	0 ± 0
	Lepus cf. *californicus*	Black-tailed Jackrabbit	Leporidae	U58924	0	0.00	0.0%	0 ± 0	2	0.11	6.3%	0.1 ± 0.06	4	0.50	10.3%	0.37 ± 0.15	0	0.00	0.0%	0 ± 0
	Procyon lotor	Racoon	Procyonidae	AB462203	1	0.08	5.9%	0.08 ± 0.08	0	0.00	0.0%	0 ± 0	0	0.00	0.0%	0 ± 0	0	0.00	0.0%	0 ± 0
	Sus scrofa	Wild Boar	Suidae	AB292606	1	0.08	5.9%	0.0 ± 0.0	0	0.00	0.0%	0 ± 0	0	0.00	0.0%	0 ± 0	0	0.00	0.0%	0 ± 0
	Arvicolinae	Vole	Cricetidae	AJ972918	1	0.08	5.9%	0.08 ± 0.08	8	0.42	25.0%	0.26 ± 0.09	1	0.13	2.6%	0.1 ± 0.1	8	0.27	2.8%	0.24 ± 0.07
	Neotoma cf. *fuscipes*	Dusky-Footed Woodrat	Cricetidae	KU745736	0	0.00	0.0%	0 ± 0	5	0.26	15.6%	0.16 ± 0.08	0	0.00	0.0%	0 ± 0	6	0.20	2.1%	0.19 ± 0.07
	Peromyscus spp.	Deer Mouse	Cricetidae	KY707306	0	0 0.00	0.0%	0 ± 0	0	0.00	0.0%	0 ± 0	0	0.00	0.0%	0 ± 0	5	0.17	1.7%	0.08 ± 0.04
	Thomomys cf. *bottae*	Botta's Pocket gophers	Geomyidae	AM407912	1	0.08	5.9%	0.0 ± 0.0	4	0.21	12.5%	0.15 ± 0.08	0	0.00	0.0%	0 ± 0	1	0.03	0.3%	0.0 ± 0.0
Small Mammals	*Sylvilagus bachmani*	Brush Rabbit	Leporidae	KU057239	1	0.08	5.9%	0.08 ± 0.08	4	0.21	12.5%	0.17 ± 0.08	0	0.00	0.0%	0 ± 0	4	0.13	1.4%	0.07 ± 0.04
	Mus musculus	House Mouse	Muridae	AB042432	1	0.08	5.9%	0.0 ± 0.0	2	0.11	6.3%	0.02 ± 0.02	0	0.00	0.0%	0 ± 0	2	0.07	0.7%	0.03 ± 0.03
	Neotominae	Woodrat/ Deer Mouse	Neotominae	KY707303	1	0.08	5.9%	0.0 ± 0.0	2	0.11	6.3%	0.05 ± 0.05	0	0.00	0.0%	0 ± 0	8	0.27	2.8%	0.18 ± 0.06
	Sciurus spp.	Tree Squirrels	Sciuridae	U59174	0	0.00	0.0%	0 ± 0	1	0.05	3.1%	0 ± 0	0	0.00	0.0%	0 ± 0	0	0.00	0.0%	0 ± 0
	Sciurus niger	Fox Squirrel	Sciuridae	U67289	1	0.08	5.9%	0.08 ± 0.08	2	0.11	6.3%	0.02 ± 0.02	2	0.25	5.1%	0.05 ± 0.05	1	0.03	0.3%	0.01 ± 0.01
	Scapanus latimanus	Broad-footed mole	Talpidae	KX754499	0	0.00	0.0%	0 ± 0	0	0.00	0.0%	0 ± 0	3	0.38	7.7%	0.16 ± 0.1	4	0.13	1.4%	0.13 ± 0.06



Functional Group	Plant Genera	Most abundant plant genera consumed within each functinal group in JRBP
Woody	Prunus	49.52%
	Ficus	14.26%
	Baccharis	14.19%
	Quercus	8.97%
	Rosa	3.55%
	Rubus	2.90%
	Diospyros	1.81%
	Corylus	1.35%
Herbaceous	Hemizonia	23.96%
	Plantago	12.58%
	Verbena	11.85%
	Centaurea	10.85%
	Carduus	10.31%
	Trifolium	9.04%
	Medicago	3.59%
	Lotus	3.56%
	Crepis	3.03%
	Castilleja	2.82%
	Convolvulus	2.76%
Grasses	Lolium	36.43%
	Bromus	13.75%
	Sorghum	11.47%
	Phalaris	8.79%
	Leymus	7.74%
	Gastridium	7.63%
	Festuca	4.59%
	Amelichloa	3.33%
	Brachypodiu	2.01%
	Nassella	1.23%
	Poa	1.22%

FIGURE 1 | Mean percentage ± SE of the relative read abundance per species per diet functional groups. The two omnivores (coyote and gray fox) have two columns representing both their plant and animal diets. The panel represents the percent of total plant genera reads identified in the diets of herbivore and omnivores within JRBP, showing only genera that comprised 1% or more of the diet.

TABLE 3 | Diet composition of six focal species, described at the functional groups of plants and mammals consumed and expressed by the Count, FOO, POO, and RRA.

Functional Groups	Jackrabbit (*n* = 13)				Deer (*n* = 12)				Gray Fox (*n* = 36)				Coyote (*n* = 3)			
	Count	FOO	POO	RRA	Count	FOO	POO	RRA	Count	FOO	POO	RRA	Count	FOO	POO	RRA
Grass	11	0.85	31%	0.02 ± 0.01	12	1.00	34%	0.12 ± 0.07	15	0.42	20%	0.04 ± 0.04	1	0.33	17%	0.01 ± 0.01
Herb	12	0.92	33%	0.61 ± 0.1	12	1.00	34%	0.21 ± 0.07	25	0.69	33%	0.1 ± 0.04	2	0.67	33%	0.06 ± 0.05
Woody	13	1.00	36%	0.36 ± 0.1	11	0.92	31%	0.66 ± 0.09	35	0.97	47%	0.84 ± 0.06	3	1.00	50%	0.92 ± 0.07
	Puma (*n* = 12)				**Bobcat (*n* = 19)**				**Gray Fox (*n* = 30)**				**Coyote (*n* = 8)**			
Large Mammals	9	0.75	64%	0.74 ± 0.13	4	0.21	18%	0.13 ± 0.07	1	0.03	3%	0.03 ± 0.03	5	0.63	50%	0.59 ± 0.17
Small Mammals	5	0.42	36%	0.25 ± 0.13	18	0.95	82%	0.86 ± 0.07	29	0.97	97%	0.96 ± 0.03	5	0.63	50%	0.40 ± 0.17

Gray fox and coyote had a combination of plant and mammal species in their diets, but RRA were not be calculated together because of differences in the primers used.

mainly black-tailed jackrabbit (*Lepus* cf. *californicus)*, and small mammals, mainly broad-footed mole (*Scapanus latimanus*). The smaller mesopredators, the bobcat and gray fox, regularly consumed small mammals (95 and 97% of samples, respectively) including dusky-footed woodrats (*Neotoma* cf. *fuscipes*), voles (Arvicolinae), and brush rabbits (*Sylvilagus bachmani macrorhinus*) (**Figure 1** and **Tables 2**, **3**). All mesopredators consumed black-tailed deer, but at a very low frequency of 1–2 samples for all species.

Based on PCA, it is evident that the body size of both predator and prey drive the differences in the diets of predators (**Figure 2A**). The two principal components explained 82.6% of the variation, with small mammals contributing 70.6% of variance to PC1 and large mammals 60.6% in the PC2. Predator-prey body mass were positively correlated when scavenged deer was included ($n = 98$, $r = 0.55$, $p < 0.0001$), but the correlation is stronger when deer was excluded ($n = 95$, $r = 0.62$, $p < 0.0001$) (**Supplementary Figure S3**).

Dietary prey richness was the same for the small predators [gray fox and bobcat (10 mOTUs)], 11 for the puma, but only five for the coyote, although the number of samples for the latter was low (**Table 2**). The number of mammal taxa per scat ranged between 1 and 6 (individual niche breadth, mean mOTUs ± SE per species) (**Supplementary Table S7**). Individual niche breadth was similar among the four predator species ($F_{3,65} = 0.69$, $p = 0.56$).

The mean plant taxa per scat sample was much larger for herbivores, with 18.63 ± 1.72 mOTUs/fecal sample, as opposed to 7.05 ± 1.06 for omnivores. The dietary dissimilarity within species was greatest for the gray fox and least for deer, while fox, deer, and jackrabbit diets were almost completely overlapping (**Supplementary Figure S4**). Deer are mixed feeders (i.e., grazers and browsers), and here their diet comprised of 66% woody and 33% grass and herbaceous vegetation, while jackrabbits are grazers, with 63% of their diet containing grasses and/or herbaceous plants (**Figure 1** and **Table 3**). The plants with the highest overall FOO for the herbivores included: the trees *Quercus* spp. and *Cedrus deodara,* (FOO = 0.68, 0.40, respectively); the forbs *Plantago* spp. and *Medicago polymorpha* (0.48, 0.36); and the grasses *Amelichloa* spp. and *Nassella* spp. (0.40) (**Supplementary Table S5**). Note that grass was found in all deer fecal samples, however, it only represented 12% of the RRA of the diet. For the gray fox, herbs comprised just 1% of RRA, despite being frequent in the diet (FOO) (**Table 3**). Clear differences in the diets of plant consumers were evident in the PCA (**Figure 2B**). The two principal components explained 77.7% of the variation, with woody vegetation explaining 43.5% of the variance in PC1 and herbaceous vegetation explaining 53.4% of the variance in PC2.

Gray fox were frequent plant consumers (75% of samples contained plant DNA), selecting a wide diversity of plants (representing 59 genera). The fox and coyote both frequently consumed woody vegetation (85–92% of the diet) which produce fruit. The top plants consumed by the two omnivores were the woody plants Catalina cherry (*Prunus ilicifolia*), coyote brush (*Baccharis pilularis*), and oaks (Fagaceae, *Quercus* spp.) (FOO = 0.36, 0.28), and herbaceous English ivy (*Hedera helix*) (0.31) (**Supplementary Table S5**). Although differing in average daily intake, the gray fox diet richness was similar to that of the herbivores (**Supplementary Table S7**).

Dietary Diversity, Niche Breadth and Overlap

We identified diet diversity and niche overlap among carnivores, and across all feeding guilds. Between the two carnivores, the bobcat had a relatively wider niche breadth than the puma (Levins' measure: B_A = 0.41 and 0.26, respectively) but approximately the same dietary diversity (Shannon diversity index: H = 2.13 and 2.03, respectively) (**Supplementary Table S7**). The coyote had the narrowest niche breadth and dietary diversity of all four predators (B_A = 0.19, H = 1.47). The Shannon's Index shows that the herbivores and omnivores have a significantly more diverse diet than do the carnivores ($F_{2,110}$ = 29.65, $p < 0.0001$) (**Figure 2C**), along with a higher individual niche breadth ($F_{2,110}$ = 58.99, $p < 0.0001$) and wider niche breadth using Levins' Standardized Index (**Supplementary Table S7**).

Pianka's index was used to calculate the niche overlap of predator species, first considering only mammal species in the diet (**Supplementary Table S8A**). The diets of the coyote and the puma did not significantly overlap [O_{jk} = 0.23, P(Obs > null) = 37%], as the diet of the coyote consisted mostly of small mammals and jackrabbits while the diet of the puma consisted mostly of deer. Gray fox and bobcat niches overlap significantly with each other [O_{jk} = 0.77, P(Obs > null) = 99%], and to a lesser extent with the coyote [O_{jk} = 0.26, 0.33, respectively; P(Obs > null) < 95%]. Gray fox and bobcat differ from the coyote mainly in the breadth of small mammals they consume. As expected, the bobcat does not significantly overlap with the puma [O_{jk} = 0.39, P(Obs > null) = 65%]. However, the nMDS for Bray-Curtis dissimilarity of diet shows that when plant and mammal diet are considered, the carnivores group together, the herbivores group together, and the omnivores group in the first nMDS axis but split in the second nMDS axis (**Figure 2D**).

We expanded the dietary niche overlap analysis to include plant mOTUs for omnivores and herbivores (**Supplementary Table S8B**). The coyote shares the greatest niche overlap with the gray fox [Ojk = 0.32, P(Obs > null) = 99%] while sharing the least with the puma (O_{jk} = 0.16). As expected, deer and jackrabbit diets have the greatest overlap [O_{jk} = 0.50, P(Obs > null) = 99%]. Although there is a 50% niche overlap in the herbivores diets, the PCA visually represents how the woody and herbaceous plant taxa are driving a significant differences between the two species (**Figure 2B**). The gray fox diet significantly overlaps with diet of both herbivores [Ojk = 0.46, P(Obs > null) = 99%]. The jackrabbit and black-tailed deer have the widest niche breadth (B_A = 0.32), and a relatively diverse and even diet (H = 4.45 and 4.35) (**Supplementary Table S7**).

Molecular Ecological Network Analyses

The diet data were used to construct an empirical food web composed of 151 plant (**Supplementary Table S6**) and 19 mammal mOTUs (**Table 2**), totaling 170 nodes and 310 weighted edges/links (predation interactions) by FOO (**Figure 3A**). The network was directed with links going from the predator node to prey/plant node. The analysis of the topological properties showed that the JRBP network has an *average clustering coefficient* of 0.104 among the feeding guilds with an *average path length* of 1.608, *average degree* of 1.82, and a low *connectance* of 0.011.

The metrics to determine key species within the network identified similar species. The most central nodes were identified through the *weighted in-degree centrality* metric and the top plant mOTUS were: *Quercus, Cedrus deodara, Plantago, Baccharis pilularis* and mammals were: deer, vole (Arvicolinae), squirrel (*Sciurus niger*) and jackrabbit. Within the network, each node was given an *eigenvector centrality* score, the closer to 1 the greater the level of influence within the network. The most important plant nodes were: *Quercus, Plantago, Cedrus deodara, Baccharis pilularis, Centaurea solstitialis, Brassica, Hedera helix*; and the mammal nodes were: deer, vole, squirrel, and jackrabbit (for full list see **Supplementary Table S9**).

The JRBP community network was significantly more modular than expected by chance (Q = 0.36, 95% CI = 0.26–0.30, $p < 0.05$). The JRBP food web was comprised of four modules (**Figure 3B**). The most significantly connected (*Z-score* > 1.96) species (hub connectors) within this community are deer (7.95), jackrabbit (7.08), gray fox (5.66), coyote (3.96) and bobcat (2.74)

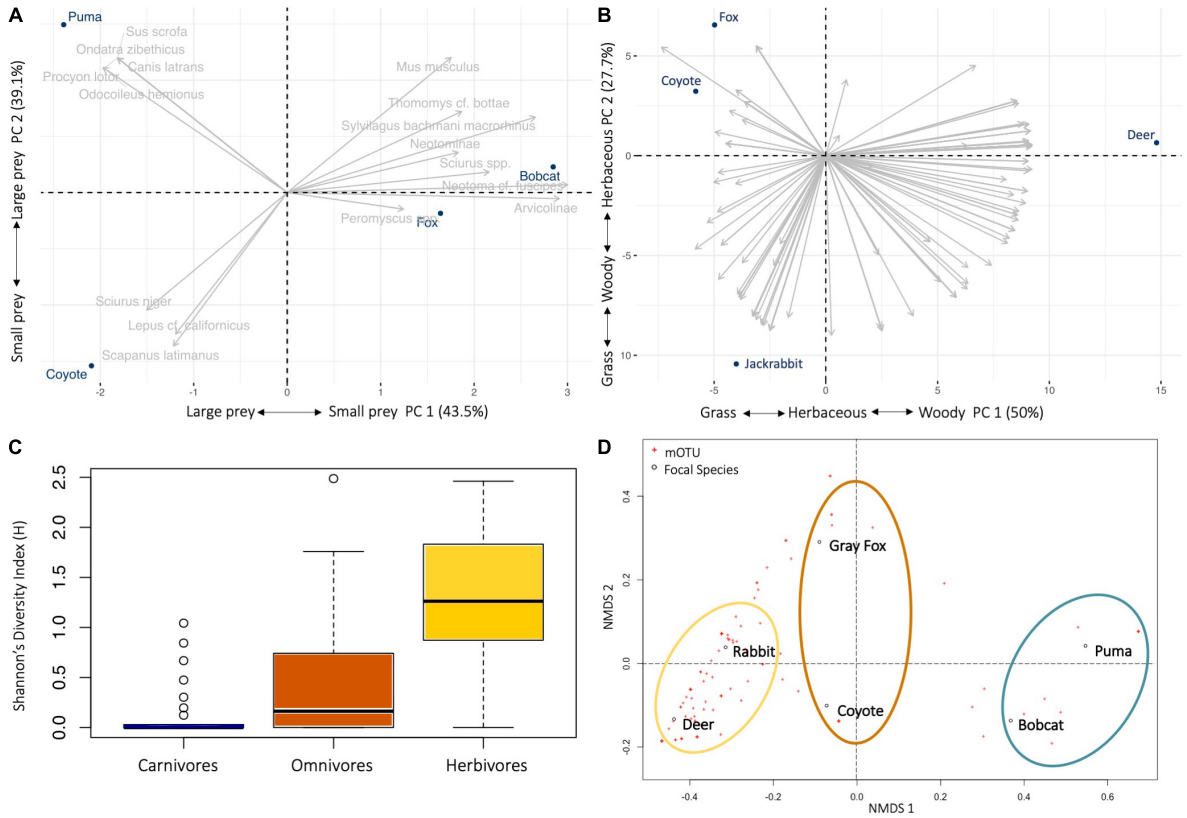


FIGURE 2 | Diet ecology of six mammal species on Jasper Ridge Biological Preserve. Principal component analysis (PCA) biplot of the mammal (n = 16) and plant (n = 196) taxa frequency of occurrence within fecal samples of the **(A)** predators (puma, coyote, bobcat, gray fox), showing the position of consumer in relation to the different prey items; and **(B)** herbivore (deer, jackrabbit) and omnivore (coyote, gray fox) (taxa names removed for clarity). Explained variances of the two principal components are shown in brackets. **(C)** Box-plots alpha diversity indices of Shannon diversity ($F_{2,110}$ = 29.65, $p < 0.0001$) in diet across fecal samples from the carnivore, omnivore and herbivore trophic level in JRBP. Lines in boxes are medians, box ends are quartiles, whiskers show ranges, and o indicates outliers. **(D)** Patterns of niche space across the six different focal species using non-metric multidimensional scaling (nMDS) for Bray-Curtis dissimilarity of frequency of occurrence data showing dissimilarity of carnivore (blue), omnivore (orange), and herbivore (yellow) diets. The distance between points represents the level of difference. The closer the species in the graph, the higher their similarity.

(for a full list see **Supplementary Table S10**). The JRBP network is not significantly nested (falls within the 95% CI expected from random networks;η = 0.40, 95% CI = 0.39–0.49, p = 0.09).

The tri-trophic chain motif was significantly over-represented, meaning this pattern occurred within the network more times than expected by chance (JRBP Network = 464; 95% CI = 43–240). Omnivory, however, was not more prevalent in this network than expected by chance (JRBP Network = 74; 95% CI = 42–184). Finally, the exploitative competition motif was also significantly more represented than expected by chance (JRBP Network = 128, 95% CI = 17–123).

Bioassessment Comparison

We compared our fecal eDNA data to camera trap and soil eDNA surveys from the same time period of scat collection (October 2017 – April 2018), and to an expert-curated mammal list of mammals in the study area (see Leempoel et al., 2020 for camera trap and soil eDNA methodologies). The scat sample size of predators was significantly correlated with the number of camera trap days for predator occurrences – in fact the rank order abundance in each data set was the same (Spearman's r_s = 1, p(2-tailed) = 0): in order, gray fox was most abundant (scat n = 71, camera n = 186), followed by bobcat (n = 31, 136), puma (n = 15, 84) and finally coyote (n = 12, 22).

This study detected 73% of known (listed) mammals in the study area (excluding bats), equal to the percent of mammals detected from soil eDNA, but greater than detected of the camera trap array (59%). Through our diet analysis, we found five species undetected by the camera traps that were small mammals (*Thomomys sp.*, *Scapanus latimanus*, *Peromyscus spp.*, and Arvicolinae), two large mammals undetected by the soil eDNA survey (*Ondatra zibethicus* and *Procyon lotor*), and an additional two species previously not known from JRBP (*Mus musculus* and *Sus scrofa*) (**Figure 4**). These two species, the house mouse (*Mus musculus)* and feral pig (*Sus scrofa)*, are not known to inhabit the preserve, suggesting that consumption of these prey may have taken place outside the preserve.

The camera trap and soil eDNA survey both captured two species that were not found in the diets of the four predators: opossum (*Didelphis virginiana*) and the striped skunk (*Mephitis*

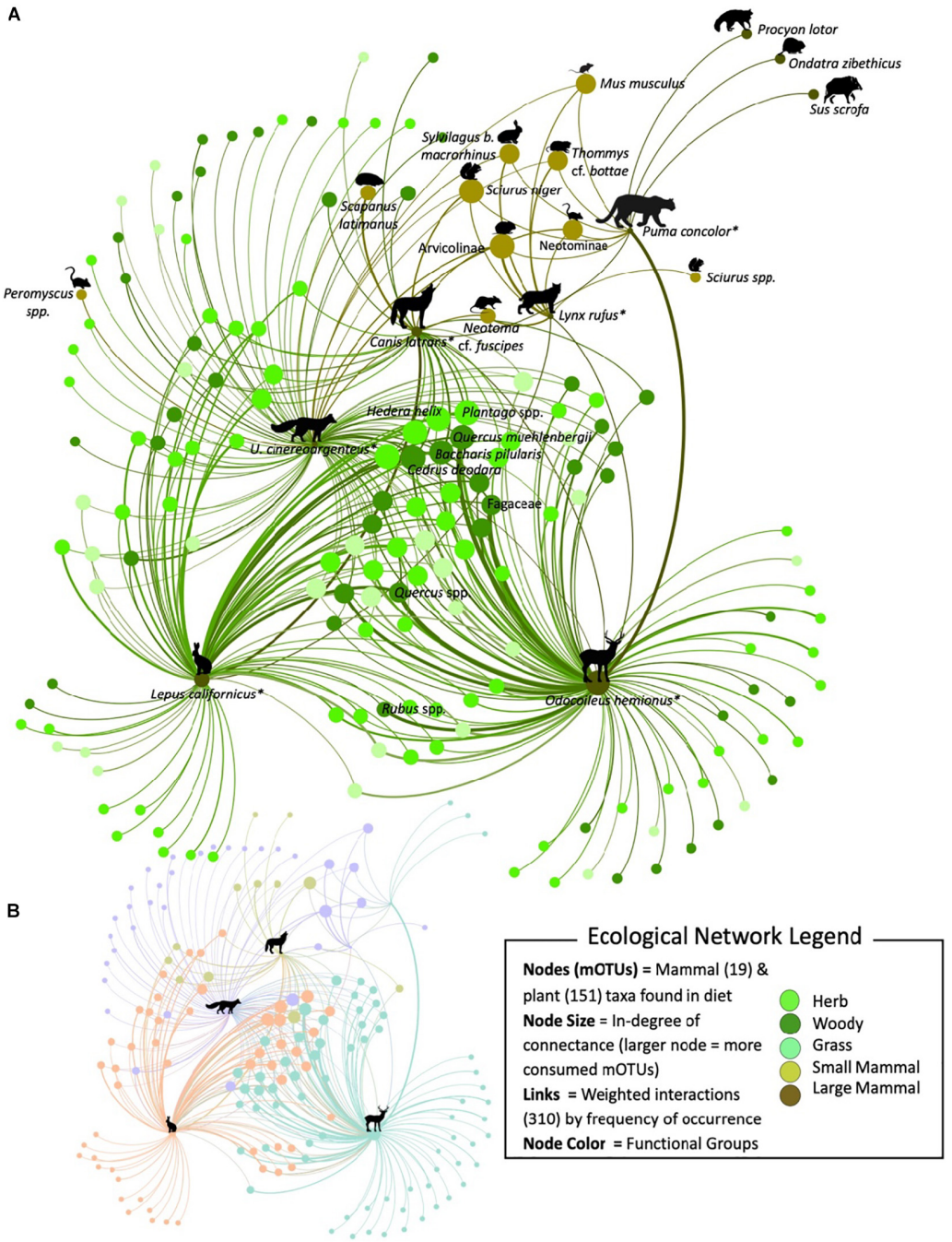


FIGURE 3 | A subset of the molecularly constructed food web of JRBP targeting the diets of black-tailed jackrabbit, black-tailed deer, gray fox, coyote, bobcat, and puma (*) represented by 170 nodes and 310 interactions. **(A)** Nodes are representing the predator – prey mOTUs and are color coded according to their functional group. The size of the node is proportional to the number of links connected (in-degree), and the width of the link is proportional to the frequency of occurrence the prey item was found in the diet across scat samples. The curve of the edge is always in a clockwise direction representing a weighted link from the consumer to the food item. All of the mammal mOTUs have been labeled, along with the most consumed plant mOTUs. **(B)** The colors in the scaled network represent the four different modular communities detected within the JRBP community and the icon represents the key species within each module.

mephitis) (**Figure 4**; Leempoel et al., 2019, 2020). However, one misidentified skunk scat sample was sequenced, which we included in the biodiversity assessment. Ground squirrels (*Otospermophilus beecheyi*) are abundant in JRBP and captured by the cameras as prey items, however, were undetected in the fecal or soil eDNA. We found that neither this species nor genus were in our 12S database relying on GenBank accessions, and thus could not be detected using our methods. The soil

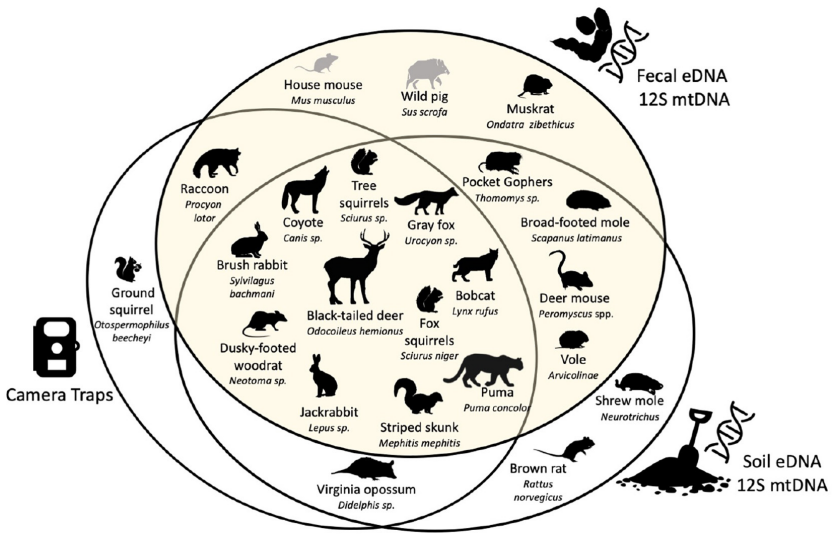


FIGURE 4 | Venn diagram between species recorded by camera traps, and species detected with soil eDNA and fecal eDNA both using the mtDNA 12S metabarcode. Scientific names are given at the maximum rank reached with each metabarcode. Species known to be present in the study area are in black. Species absent from the study area but detected with fecal eDNA are in gray [figure modified from Leempoel et al. (2020)].

eDNA study documented two small mammal species not found in this study: the native shrew mole (*Neurotrichus gibbsii*) and non-native brown rat (*Rattus norvegicus)*.

When we include both plant genera and mammal mOTUs, we found that 32% of our mOTUs have not been recorded at JRBP previously, comprising identified plant species (41/77), genera (45/61), family/tribe (13/13) and mammal taxa (17/19) in the diet. These previously unrecord species could be new species to the reserve, or our study species may be going beyond the reserve to forage. It is also possible that detection of these new taxa is caused by poor resolution or errors in the GenBank database, or unknown sequencing errors. To be more conservative with the mOTU identification, we collapsed each plant species to the genus-level, which revealed that 71% (84/119) were on the JRBP plant list.

DISCUSSION

Studies of multitrophic interactions and ecological network analysis are increasingly being used to monitor biodiversity and trophic interactions (Compson et al., 2019), but the fundamental assessment and quantification of these interactions is challenging. We assembled a real-world, high-resolution multitrophic food web (**Figure 3**), demonstrating that molecular ecological network analyses (MENA) from readily available fecal environmental DNA (eDNA) provides a powerful assessment of an ecosystem. MENA identified the plant and mammal diet composition for six large mammals across three feeding guilds, accurately capturing the biodiversity of the area and characterizing the trophic structure and interaction patterns within the community at an exceptional taxonomic resolution, thereby improving traditional food web analysis and biomonitoring. Although previous studies have validated the use of DNA metabarcoding (Roslin et al., 2019), and created bipartite networks (Kartzinel et al., 2015; Pansu et al., 2019) or even a network of networks (Clare et al., 2019), few studies have yet characterized multitrophic networks with high taxonomic resolution, to characterize the extent of indirect links in a terrestrial system.

Assessing Community Structure and Keystone Species

Food web structure is one mechanism to examine ecosystem organization (Allesina et al., 2009; Thompson et al., 2012; Saint-béat et al., 2015; Monteiro and Faria, 2016) and is key in forecasting the effects of trophic degradation on ecosystem function (Duffy et al., 2007; Estes et al., 2011; Naiman et al., 2012), however, is often underutilized in bioassessments. The structure of the multitrophic network of Jasper Ridge Biological Preserve (JRBP) was found to be significantly more modular and non-nested. This highlights the presence of discrete groups of predators and preys that are strongly connected among each other and less tightly with other groups. The highly modular community indicates that disturbances (e.g., defaunation and local extinctions) would spread more slowly throughout the multiple smaller communities (Tylianakis et al., 2010). The most connected species within each hub or module were, in order of connectance, deer, jackrabbit, gray fox, and coyote. The non-nestedness of this food web indicates low niche overlap among consumers and thus a more stable coexistence (Kondoh et al., 2010).

The centrality metrics we chose identified a list of candidate keystone species that play a functional role in the food web dynamics (Power et al., 1996; Delmas et al., 2019). Studies have shown that the removal of such keystone species or nodes with a high centrality value collapse the network (Allesina

et al., 2009), thus should be prioritized for monitoring to avoid cascading extinctions within an ecosystem (Dunne et al., 2002). Within Jasper Ridge, we identified species contributing most to supporting higher trophic levels, such as the oaks, plantains, and fleshy fruit bearing shrubs/trees being fundamental to the herbivore and omnivore community and the deer, jackrabbit, vole, and squirrel are primarily supporting the predator community (**Supplementary Table S9**). Accurately identifying the diets of these key prey within the system will contribute to the mapping of ecological feedback loops, allowing for quantitative vegetation projections (Bowman et al., 2015). For example, the two focal herbivores of this study were the most frequently consumed prey species, and had the widest niche breadth and most diverse diets. If we focused exclusively on the frequency of functional group occurrence in the diet, we might infer that the three plant groups were consumed equally among both herbivores, however, when considering the mean abundance of reads within the diet, we see a much different conclusion (**Figure 1** and **Table 3**). We find that both species are feeding on a broad variety of plant taxa, the deer are consuming mostly woody vegetation (trees and shrubs) and the jackrabbit more herbaceous plants. Consequently, despite the high overlap of mOTUs consumed by the herbivores, the impact on the vegetation from the removal or introduction of a prey-specific predator (e.g., puma consume mostly deer, whereas gray fox consumes more rabbits) is expected to be markedly different (**Figure 2B**). With our fine-scale understanding of the diet selection by these herbivores and identified keystone species, we can now better predict the impact fluctuating predator populations may have on this ecosystem.

Patterns of Trophic Interactions

Our MENA framework provides a real-time assessment of empirical species interactions and a technique to quantify and predict the patterns of trophic interactions. Trophic interactions in food webs are, however, often modeled through species traits (McGill et al., 2006; Gravel et al., 2013), for example body mass, where predator and prey body masses scale with each other in natural food webs (Brose et al., 2019). This predator-prey mass ratio (PPMR) is often a driver of the patterns and strength of trophic interactions within a food web, in turn contributing to food web stability (Emmerson and Raffaelli, 2004; Brose, 2010). To be confident in the identified species interactions and trophic patterns, we confirmed that the predator and prey body masses scale with each other in JRBP, adding validity to this method and allowing for stronger predictive modeling in the future (Brose et al., 2019). As expected according to recent results (Allen, 2014), we found the largest predator in the region, the puma, predominantly consumed deer and other large mammals (including coyote), supplemented with small mammals. The smaller mesopredators, bobcat and gray fox, had a similar diet richness to the puma, but primarily consumed small mammals, similar to diets observed nearby in the Santa Cruz Mountains (**Table 2** and **Figure 2A**; Smith et al., 2018).

We captured the same hierarchical patterns through network motifs as the long-term camera trap study, where the natural reoccurrence of puma to JRBP triggered a top-down effect on mesopredators and herbivores (Leempoel et al., 2019). Network motifs are the basic building blocks of communities and can be identified as overrepresented patterns throughout the network. Here we focused on the three most informative patterns: tri-trophic chains, exploitative competition, and omnivory (Milo et al., 2002; Stouffer and Bascompte, 2010; Paulau et al., 2015; Delmas et al., 2019). Within the JRBP food web, tri-trophic chains and exploitative competition patterns were significantly overrepresented compared to chance expectation. These motifs may emerge through the functional importance of these direct and indirect interactions within the assembly of this ecological network. In fact, the prevalence of tri-trophic chain patterns indicate that the removal or reintroduction of top predators would trigger a trophic cascade within this community, directly affecting not only their prey but also, indirectly, the plant communities on which the prey feed. Projecting the long-term consequences of these trophic cascade motifs, we predict that the recent resurgence of puma population may positively affect the woody vegetation, favoring forest regrowth and benefiting the overall ecosystem function.

Exploitative competition predicts indirect links between predators consuming the same resources and may lead to resource exclusion by some of these competitors, results that are concordant with the decrease in coyotes following an increase in pumas in the study area (Leempoel et al., 2019). These predicted network responses are also supported by the significant niche overlap of the omnivorous coyote and gray fox diets and the subsequent predator release of gray fox that occurred at JRBP due to the exclusion of coyote by puma (Leempoel et al., 2019). This suggest the small overlap between the puma and coyote diets is driven by interspecific competition and resource partitioning, while the larger overlap of the bobcat and gray fox diet points to limited competition and a sharing of resources (Gotelli and Graves, 1996). Although omnivory patterns were not significant within this network, among the six mammals we studied, the gray fox was responsible for 58% of omnivore motifs within the food web. Since omnivory is significantly reduced without fox, it is possible that a future change in fox population would shift the prevalence and importance of omnivory within the community. Taken together, our results suggest that the JRBP food web is favoring the persistence of apex predators, despite the consequences of exploitative competition, and is enhancing the regulation of herbivores and the survival of identified key woody and herbaceous vegetation. We can then generate expectations about the impacts of shifting community dynamics, which would provide a framework for management decisions.

Capturing Biodiversity With MENA

This study showcases the value of non-invasive fecal eDNA surveys as a biodiversity assessment tool, highlighted by the depth of biodiversity captured and the greater insight into the ecology of species and their interactions. We identified a rich diversity of plant and mammal mOTUs, classifying 91% to the genus or species level (**Supplementary Table S2**), allowing for high-resolution assessments of species dietary richness, niche breadth and overlap, and diversity of the community. To the best of our knowledge, this is the first published molecular analysis of gray

fox, coyote and black-tail jackrabbit plant diets. Our approach revealed a much wider diet breadth of plants than previously recorded for gray fox (107 mOTUs), jackrabbit (97) and, to a lesser extent, coyote (24). Micro histological identification techniques used to determine fox diet found only 4–20 different plant species (Wilson and Thomas, 1999; Cunningham et al., 2006), whereas we identified 61 different genera, tripling previous estimates. We identified 65 plant genera within the black-tailed jackrabbit diet at this site, while previous morphological studies only identified 14–32 plant species across numerous sites and habitats (Fagerstone et al., 1980; Wansi et al., 1992).

We compared MENA to the plant and mammal species list of Jasper Ridge, and two concurrently run surveys, soil eDNA and camera traps (Leempoel et al., 2020), to determine its effectiveness. Although this is just a subset of the larger biodiversity network in this protected area (e.g., other mammals, microorganisms, birds, pollinators), fecal eDNA detected a greater number of mammal species over a shorter period of time than the dense array of camera traps (4 cameras/km^2) and a greater number of species than the soil eDNA (**Figure 4**). Two species, the possum and skunk, were not found in the fecal eDNA sampled diets, although we did collect a skunk scat (initially misidentified as gray fox) and so included skunk in the biodiversity assessment. Puma have previously been identified consuming both skunk and possum on camera traps, and the populations of both have declined since puma returned (Leempoel et al., 2019). The low abundances of these species could explain their absence from scat samples over the scat collection period. Fecal eDNA identified a significant proportion of the known small mammal community, which were not detected by the camera trap survey. Camera traps are set up to survey either medium to large mammals or small mammal communities, but seldom both. Accurate identification from images is also challenging (De Bondi et al., 2010; Meek and Pittet, 2012; Meek et al., 2013), and consequently, small mammals are often left out of camera trap biodiversity studies, or are poorly distinguished, yet they are important indicators of ecosystem health (Rowe and Terry, 2014). The soil eDNA detected the same small mammals as fecal eDNA, however, did not detect some of the larger mammals that occur at low occupancy in JRBP, including raccoon and muskrat. Also notable, the positive correlation between scats and camera-trap images suggests that scat collection may also be used as an indicator of relative abundance for these species, similar to previous studies on Iberian lynx (Garrote et al., 2014) and jaguar abundance (Sollmann et al., 2013).

Fecal eDNA provides a perspective of a system through a species lens, rather than that of a predefined human boundary – often used to delimit conventional surveys or assessments. For example, the diet of plants is representative throughout a herbivores range, rather than just that recorded in JRBP (29% of mOTUs genera were previously unrecorded in JRBP). MENA can also be deployed for the early detection and monitoring of non-native species that are occurring within the animal's range (e.g., two prey items never before recoded within JRBP). For example, the presence of feral pigs (*Sus scrofa*) in the puma diet is an early sign that this highly invasive and destructive species [$1.5 billion in economic damage to United States agriculture and environment annually (Finzel and Baldwin, 2015]) is moving into the region, as the nearest known records are 15 km + from the study area.

Current Limitations of MENA

Molecular biodiversity assessments can only be as good as the reference database on which they rely for species identification. A good example of this is ground squirrels (*Otospermophilus beecheyi*), an abundant species in JRBP that we know is frequently preyed upon from photographic evidence. Ground squirrels were not detected in the diet of our four predator species and the concurrent eDNA soil study because it was not in our 12S mtDNA database. This highlights the importance of collaborative initiatives, such as The Bar Code of Life data system (Ratnasingham and Hebert, 2007), to increase reference library coverage and database capacities. Only with a comprehensive database can we thoroughly assess the biodiversity of an area using a metabarcoding approach, especially in regions that have been historically understudied (Bush et al., 2019).

Due to the sensitivity of eDNA, it is important to be aware of confounding factors. For example, we identified bobcat and gray fox DNA in many of each other's scat samples. We have evidence from camera traps that these species countermark one another (urinate and defecate on scats from the other species), and it is highly unlikely that they are consuming each other. Therefore, it is important to know the behavioral ecology of a study species, to rule out possible cross-contaminations, while at the same time exploring the possibility of novel and poorly recognized interactions. A second example is wind-dispersed plant spores that could seasonally contaminate scat samples. We considered possible contamination from Deodar Cedar (*Cedrus deodar*) pollen as it is released in the autumn (September – October) (Sharma and Khanduri, 2012). However, Deodar occurred in less than half of the samples from species consuming plants. It is possible they were consuming the seed-producing cones that were readily available. A third consideration is that the prey's diet (via the gut contents) could be represented in a predator's molecular diet analysis. It is not yet feasible to detect this, however, we did not have positive PCR results for plant DNA in the majority of the predator scats, possibly due to the larger fragment size (200–387 bp) of the ITS2 amplicon region that we used (Moorhouse-Gann et al., 2018). Regardless, as a bioassessment of the region, the identified plants were either consumed by the omnivores directly or indirectly and thus still part of the functional system. Studies reviewing these and other limitations of these techniques should be considered prior to bioassessments (Clare, 2014; Alberdi et al., 2018; Pawlowski et al., 2018; Delmas et al., 2019; McGee et al., 2019; Ruppert et al., 2019; Zinger et al., 2019).

CONCLUSION

We have demonstrated that MENA is a promising tool for monitoring biodiversity, unveiling and understanding multitrophic interactions and community structure, and

identifying key and vulnerable species within a terrestrial system. Only when we know what species are present and understand their function in the ecosystem, will we be equipped to protect and holistically manage these systems. The consequences of ecological network rewiring are drastic (Olivier et al., 2019), and the current rapidity of changes to ecological networks makes it difficult to detect and respond. By repeating MENA, temporal changes can be quickly identified and tracked. For example, the impact of wildlife reintroductions or changes in wildlife abundance and occupancy can be tracked and assessed through MENA and fed back into adaptive management plans to monitor for predicted impacts (Pires, 2017). Here we have explored MENA using a few mammal species scats on three trophic levels, but future assessments could drastically broaden the picture by including more vertebrates, invertebrates, and interaction types that also include pathogens and parasites. The non-invasive and quick turn-around of MENA, along with the decreasing costs of HTS, will accelerate its implementation at local and global scales (Ruppert et al., 2019).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JM designed the study, performed the field and lab work, and wrote the first draft. JM and KL performed the bioinformatics related to DNA metabarcoding. GL and JM performed the network analysis. EH designed and supervised the research. GL, KL, and EH contributed to writing and editing the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

JM acknowledges support from the Philippe S. Cohen Graduate Fellowship and JRBP Mellon grant. KL acknowledges support from the Swiss National Science Foundation Early *Post doc* Mobility grant (P2ELP3_175075), JRBP Mellon grant. GL acknowledges support from the Swiss National Science Foundation (P2ZHP3_187938). The work was made possible by funding from the HHMI Professorship (EH).

ACKNOWLEDGMENTS

The manuscript was greatly improved after reviews by Simon Morgan, Allison Stegner, and Katie Solari. We thank Rodolfo Dirzo for initiating the long-term camera trap array and inspiring the continuous research, and for invaluable advising and edits. We thank Nona Chiariello, Trevor Herbert, and the Docents of Jasper Ridge Biological Preserve for their contributions to the tagging of camera trap images, scat collection, plant identifications, and historical information.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.588430/full#supplementary-material

REFERENCES

Abdala-Roberts, L., Puentes, A., Finke, D. L., Marquis, R. J., Montserrat, M., Poelman, E. H., et al. (2019). Tri-trophic interactions: bridging species, communities and ecosystems. *Ecol. Lett.* 22, 2151–2167. doi: 10.1111/ele.13392

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., and Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods Ecol. Evol.* 9, 134–147. doi: 10.1111/2041-210X.12849

Allen, M. L. (2014). *The Ecology and Behaviour of Pumas (Puma Concolor) in Northern California, U.S.A.* Ph.D. Diss, Victoria University of Wellington, Wellington

Allesina, S., Bodini, A., and Pascual, M. (2009). Functional links and robustness in food webs. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1701–1709. doi: 10.1098/rstb.2008.0214

Allesina, S., and Pascual, M. (2009). Googling food webs: can an eigenvector measure species' importance for coextinctions? *PLoS Comput. Biol.* 5:e1000494. doi: 10.1371/journal.pcbi.1000494

Baiser, B., Buckley, H. L., Gotelli, N. J., and Ellison, A. M. (2012). Predicting food-web structure with metacommunity models. *Oikos* 122, 001–015. doi: 10.2307/41937696

Bartley, T. J., McCann, K. S., Bieg, C., Cazelles, K., Granados, M., Guzzo, M. M., et al. (2019). Food web rewiring in a changing world. *Nat. Ecol. Evol.* 3, 345–354. doi: 10.1038/s41559-018-0772-3

Bastian, M., and Heymann, S. (2009). "Gephi: an open source software for exploring and manipulating networks," in *Proceedings of the Third International Conference on Weblogs and Social Media, ICWSM 2009*, San Jose, CA, 4–6. doi: 10.13140/2.1.1341.1520

Bastolla, U., Fortuna, M. A., Pascual-García, A., Ferrera, A., Luque, B., and Bascompte, J. (2009). The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature* 458, 1018–1020. doi: 10.1038/nature07950

Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R., et al. (2005). Defining operational taxonomic units using DNA barcode data. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 1935–1943. doi: 10.1098/rstb.2005.1725

Bowman, D. M. J. S., Perry, G. L. W., and Marston, J. B. (2015). Feedbacks and landscape-level vegetation dynamics. *Trends Ecol. Evol.* 30, 255–260. doi: 10.1016/j.tree.2015.03.005

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., and Coissac, E. (2016). obitools: a unix-inspired software package for DNA metabarcoding. *Mol. Ecol. Resour.* 16, 176–182. doi: 10.1111/1755-0998.12428

Brose, U. (2010). Body-mass constraints on foraging behaviour determine population and food-web dynamics. *Funct. Ecol.* 24, 28–34. doi: 10.1111/j.1365-2435.2009.01618.x

Brose, U., Archambault, P., Barnes, A. D., Bersier, L. F., Boy, T., Canning-Clode, J., et al. (2019). Predator traits determine food-web architecture across ecosystems. *Nat. Ecol. Evol.* 3, 919–927. doi: 10.1038/s41559-019-0899-x

Brown, D. S., Burger, R., Cole, N., Vencatasamy, D., Clare, E. L., Montazam, A., et al. (2014). Dietary competition between the alien Asian Musk Shrew (Suncus murinus) and a re-introduced population of Telfair's Skink (*Leiolopisma telfairii*). *Mol. Ecol.* 23, 3695–3705. doi: 10.1111/mec.12445

Bush, A., Compson, Z. G., Monk, W. A., Porter, T. M., Steeves, R., Emilson, E., et al. (2019). Studying ecosystems With DNA metabarcoding: lessons from biomonitoring of aquatic macroinvertebrates. *Front. Ecol. Evol.* 7:434. doi: 10.3389/fevo.2019.00434

Casey, J. M., Meyer, C. P., Morat, F., Brandl, S. J., Planes, S., and Parravicini, V. (2019). Reconstructing hyperdiverse food webs: gut content metabarcoding as a tool to disentangle trophic interactions on coral reefs. *Methods Ecol. Evol.* 10, 1157–1170. doi: 10.1111/2041-210X.13206

Clare, E. L. (2014). Molecular detection of trophic interactions: emerging trends, distinct advantages, significant considerations and conservation applications. *Evol. Appl.* 7, 1144–1157. doi: 10.1111/eva.12225

Clare, E. L., Fazekas, A. J., Ivanova, N. V., Floyd, R. M., Hebert, P. D. N., Adams, A. M., et al. (2019). Approaches to integrating genetic data into ecological networks. *Mol. Ecol.* 28, 503–519. doi: 10.1111/mec.14941

Compson, Z. G., Monk, W. A., Hayden, B., Bush, A., O'Malley, Z., Hajibabaei, M., et al. (2019). Network-based biomonitoring: exploring freshwater food webs with stable isotope analysis and DNA metabarcoding. *Front. Ecol. Evol.* 7:395. doi: 10.3389/FEVO.2019.00395

Coverdale, T. C., Kartzinel, T. R., Grabowski, K. L., Shriver, R. K., Hassan, A. A., Goheen, J. R., et al. (2016). Elephants in the understory: opposing direct and indirect effects of consumption and ecosystem engineering by megaherbivores. *Ecology* 97, 3219–3230. doi: 10.1002/ecy.1557

Csardi, G., and Nepusz, T. (2006). The igraph software package for complex network research. *Inter J. Complex Sys.* 1695, 1–9.

Cunningham, S. C., Kirkendall, L., and Ballard, W. (2006). gray fox and coyote abundance and diet responses after a wildfire in central arizona. *West. North Am. Nat.* 66, 169–180. doi: 10.3398/1527-0904(2006)66[169:gfacaa]2.0.co;2

Czernik, M., Taberlet, P., Świsłocka, M., Czajkowska, M., Duda, N., and Ratkiewicz, M. (2013). Fast and efficient DNA-based method for winter diet analysis from stools of three cervids: moose, red deer, and roe deer. *Acta Theriol.* 58, 379–386. doi: 10.1007/s13364-013-0146-9

Daam, M. A., Teixeira, H., Lillebø, A. I., and Nogueira, A. J. A. (2019). Establishing causal links between aquatic biodiversity and ecosystem functioning: status and research needs. *Sci. Total Environ.* 656, 1145–1156. doi: 10.1016/j.scitotenv.2018.11.413

Dale, M. R. T., and Fortin, M.-J. (2010). From graphs to spatial graphs. *Annu. Rev. Ecol. Evol. Syst.* 41, 21–38. doi: 10.1146/annurev-ecolsys-102209-144718

Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., et al. (2019). Counting with DNA in metabarcoding studies: how should we convert sequence reads to dietary data? *Mol. Ecol.* 28, 391–406. doi: 10.1111/mec.14734

De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., et al. (2014). DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. *Mol. Ecol. Resour.* 14, 306–323. doi: 10.1111/1755-0998.12188

De Bondi, N., White, J. G., Stevens, M., and Cooke, R. (2010). A comparison of the effectiveness of camera trapping and live trapping for sampling terrestrial small-mammal communities. *Wildl. Res.* 37, 456–465. doi: 10.1071/WR10046

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., et al. (2017). Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Mol. Ecol.* 26, 5872–5895. doi: 10.1111/mec.14350

Delmas, E., Besson, M., Brice, M.-H., and Burkle, L. A. (2017). Analyzing ecological networks of species in- teractions. *bioRxiv* [Preprint]. doi: 10.1101/112540

Delmas, E., Besson, M., Brice, M. H., Burkle, L. A., Dalla Riva, G. V., Fortin, M. J., et al. (2019). Analysing ecological networks of species interactions. *Biol. Rev.* 94, 16–36. doi: 10.1111/brv.12433

Deng, Y., Jiang, Y. H., Yang, Y., He, Z., Luo, F., and Zhou, J. (2012). Molecular ecological network analyses. *BMC Bioinformatics* 13:113. doi: 10.1186/1471-2105-13-113

Derocles, S. A. P., Bohan, D. A., Dumbrell, A. J., Kitson, J. J. N., Massol, F., Pauvert, C., et al. (2018). Biomonitoring for the 21st Century: integrating next-generation sequencing into ecological network analysis. *Adv. Ecol. Res.* 58, 1–62. doi: 10.1016/bs.aecr.2017.12.001

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., and Collen, B. (2014). Defaunation in the antrhopocene. *Science* 401, 401–406. doi: 10.1126/science.1251817

Doulcier, G., Guimera, R., and Stouffer, D. (2016). *Fast Network Modularity and Roles Computation by Simulated. R Packag. Version 0.2.4*. doi: 10.1038/nature03288

Duffy, J. E., Cardinale, B. J., France, K. E., McIntyre, P. B., Thébault, E., and Loreau, M. (2007). The functional role of biodiversity in ecosystems: incorporating trophic complexity. *Ecol. Lett.* 10, 522–538. doi: 10.1111/j.1461-0248.2007.01037.x

Dunne, J. A., Williams, R. J., and Martinez, N. D. (2002). Network structure and biodiversity loss in food webs: robustness increase with connectance. *Ecol. Lett.* 5, 558–567. doi: 10.1046/j.1461-0248.2002.00354.x

Eisenhauer, N., Schielzeth, H., Barnes, A. D., Barry, K., Brose, U., Bruelheide, H., et al. (2019). A multitrophic perspective on biodiversity–ecosystem functioning research. *Adv. Ecol. Res.* 61, 1–48. doi: 10.1016/bs.aecr.2019.06.001

Elhesha, R., Kahveci, T., and Baiser, B. (2017). Motif centrality in food web networks. *J. Complex Netw.* 5, 641–664. doi: 10.1093/comnet/cnw032

Emmerson, M. C., and Raffaelli, D. (2004). Predator-prey body size, interaction strength and the stability of a real food web. *J. Anim. Ecol.* 73, 399–409. doi: 10.1111/j.0021-8790.2004.00818.x

Erickson, D. L., Reed, E., Ramachandran, P., Bourg, N. A., McShea, W. J., and Ottesen, A. (2017). Reconstructing a herbivore's diet using a novel rbcL DNA mini-barcode for plants. *AoB Plants* 9:lx015. doi: 10.1093/aobpla/plx015

Estes, J., Terborgh, J., and Brashares, J. (2011). Trophic downgrading of planet Earth. *Science* 333, 301–306. doi: 10.1126/science.1205106

Evans, D. M., Kitson, J. J. N., Lunt, D. H., Straw, N. A., and Pocock, M. J. O. (2016). Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. *Funct. Ecol.* 30, 1904–1916. doi: 10.1111/1365-2435.12659

Fagerstone, K. A., Lavoie, G. K., and Griffith, R. E. (1980). Black-tailed jackrabbit diet and density on rangeland and near agricultural crops. *J. Range Manag.* 33, 229–233. doi: 10.2307/3898292

Ficetola, G. F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., et al. (2010). An In silico approach for the evaluation of DNA barcodes. *BMC Genomics.* 11:434. doi: 10.1186/1471-2164-11-434

Finzel, J. A., and Baldwin, R. A. (2015). *Pest Notes: Wild Pigs*. Davis, CA: UC ANR Publ, 74170.

Garrote, G., de Ayala, R. P., and Tellería, J. L. (2014). A comparison of scat counts and camera-trapping as means of assessing Iberian lynx abundance. *Eur. J. Wildl. Res.* 60, 885–889. doi: 10.1007/s10344-014-0855-7

Gaston, K. J., Blackburn, T. M., and Lawton, J. H. (1997). Interspecific abundance-range size relationships: an appraisal of mechanisms. *J. Anim. Ecol.* 66, 579–601. doi: 10.2307/5951

Giguet-Covex, C., Pansu, J., Arnaud, F., Rey, P. J., Griggo, C., Gielly, L., et al. (2014). Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nat. Commun.* 5:3211. doi: 10.1038/ncomms4211

Giling, D. P., Ebeling, A., Eisenhauer, N., Meyer, S. T., Roscher, C., Rzanny, M., et al. (2019). Plant diversity alters the representation of motifs in food webs. *Nat. Commun.* 10, 1–7. doi: 10.1038/s41467-019-08856-0

Gonzalez, A., Rayfield, B., and Lindo, Z. (2011). The disentangled bank: how loss of habitat fragments and disassembles ecological networks. *Am. J. Bot.* 98, 503–516. doi: 10.3732/ajb.1000424

Gotelli, N., Hart, E., Ellison, A., and Maintainer. (2015). *PFackage "EcoSimR": Null Model Analysis for Ecological Data. R Packag.*

Gotelli, N. J., and Colwell, R. K. (2001). Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol. Lett.* 4, 379–391. doi: 10.1046/j.1461-0248.2001.00230.x

Gotelli, N. J., and Graves, G. R. (1996). *Null Models in Ecology*. Washington, D.C: Smithsonian Institution Press, doi: 10.1086/284619

Gravel, D., Poisot, T., Albouy, C., Velez, L., and Mouillot, D. (2013). Inferring food web structure from predator-prey body size relationships. *Methods Ecol. Evol.* 4, 1083–1090. doi: 10.1111/2041-210X.12103

Hu, Y. (2006). Efficient and high quality force-directed graph drawing. *Math. J.* 10, 37–71.

Iwanowicz, D. D., Vandergast, A. G., Cornman, R. S., Adams, C. R., Kohn, J. R., Fisher, R. N., et al. (2016). Metabarcoding of fecal samples to determine herbivore diets: a case study of the endangered Pacific pocket mouse. *PLoS One* 11:e0165366. doi: 10.1371/journal.pone.0165366

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., et al. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecol. Lett.* 16, 1245–1257. doi: 10.1111/ele.12162

Jonsson, B. G. (2001). A null model for randomization tests of nestedness in species assemblages. *Oecologia* 127, 309–313. doi: 10.1007/s004420000601

Jordan, F., and Scheuring, I. (2004). Network ecology: topological constraint on ecosystem dynamics. *Phys. Life Rev.* 1, 139–172. doi: 10.1016/j.plrev.2004.08.001

JRBP (2019). *Jasper Ridge Biological Preserve Mammal Species List*. Available online at: https://jrbp.stanford.edu/search/site/mammal list (accessed February 10, 2019).

Kartzinel, T. R., Chen, P. A., Coverdale, T. C., Erickson, D. L., Kress, W. J., Kuzmina, M. L., et al. (2015). DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8019–8024. doi: 10.1073/pnas.1503283112

Kéfi, S., Miele, V., Wieters, E. A., Navarrete, S. A., and Berlow, E. L. (2016). How structured is the entangled bank? The surprisingly simple organization of multiplex ecological networks leads to increased persistence and resilience. *PLoS Biol.* 14:e1002527. doi: 10.1371/journal.pbio.1002527

Kohn, M. H., York, E. C., Kamradt, D. A., Haught, G., Sauvajot, R. M., and Wayne, R. K. (1999). Estimating population size by genotyping faeces. *Proc. R. Soc. B Biol. Sci.* 266, 657–663. doi: 10.1098/rspb.1999.0686

Kondoh, M., Kato, S., and Sakato, Y. (2010). Food webs are built up with nested subwebs. *Ecology* 91, 3123–3130. doi: 10.2307/20788145

Landi, P., Minoarivelo, H. O., Brännström, Å, Hui, C., and Dieckmann, U. (2018). Complexity and stability of ecological networks: a review of the theory. *Popul. Ecol.* 60, 319–345. doi: 10.1007/s10144-018-0628-3

Leempoel, K., Hebert, T., and Hadly, E. A. (2020). A comparison of eDNA to camera trapping for assessment of terrestrial mammal diversity. *Proc. R. Soc. B Biol. Sci.* 287:20192353. doi: 10.1098/rspb.2019.2353

Leempoel, K., Meyer, J., Hebert, T., Nova, N., and Hadly, E. A. (2019). Return of an apex predator to a suburban preserve triggers a rapid trophic cascade. *bioRxiv* [Preprint]. doi: 10.1101/564294

Losapio, G., Fortuna, M. A., Bascompte, J., Schmid, B., Michalet, R., Neumeyer, R., et al. (2019). Plant interactions shape pollination networks via nonadditive effects. *Ecology* 100, 1–9. doi: 10.1002/ecy.2619

Losapio, G., and Schöb, C. (2017). Resistance of plant–plant networks to biodiversity loss and secondary extinctions following simulated environmental changes. *Funct. Ecol.* 31, 1145–1152. doi: 10.1111/1365-2435.12839

Lupatini, M., Suleiman, A. K. A., Jacques, R. J. S., Antoniolli, Z. I., de Siqueira Ferreira, A., Kuramae, E. E., et al. (2014). Network topology reveals high connectance levels and few key microbial genera within soils. *Front. Environ. Sci.* 2:10. doi: 10.3389/fenvs.2014.00010

Lyngdoh, S., Shrotriya, S., Goyal, S. P., Clements, H., Hayward, M. W., and Habib, B. (2014). Prey preferences of the snow leopard (Panthera uncia): regional diet specificity holds global significance for conservation. *PLoS One* 9:e0088349. doi: 10.1371/journal.pone.0088349

Ma, J., Lu, Y., Chen, F., Li, X., Xiao, D., and Wang, H. (2020). Molecular ecological network complexity drives stand resilience of soil bacteria to mining disturbances among typical damaged ecosystems in China. *Microorganisms* 8:433. doi: 10.3390/microorganisms8030433

Makiola, A., Compson, Z. G., Baird, D. J., Barnes, M. A., Boerlijst, S. P., Bouchez, A., et al. (2020). Key questions for next-generation biomonitoring. *Front. Environ. Sci.* 7:197. doi: 10.3389/fenvs.2019.00197

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. doi: 10.14806/ej.17.1.200

McCann, K. (2007). Protecting biostructure. *Nature* 446:29. doi: 10.1038/446029a

McGee, K. M., Robinson, C. V., and Hajibabaei, M. (2019). Gaps in DNA-based biomonitoring across the globe. *Front. Ecol. Evol.* 7:337. doi: 10.3389/fevo.2019.00337

McGill, B. J., Enquist, B. J., Weiher, E., and Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends Ecol. Evol.* 21, 178–185. doi: 10.1016/j.tree.2006.02.002

Meek, P. D., and Pittet, A. (2012). User-based design specifications for the ultimate camera trap for wildlife research. *Wildl. Res.* 39, 649–660. doi: 10.1071/WR12138

Meek, P. D., Vernes, K., and Falzon, G. (2013). On the reliability of expert identification of small-medium sized mammals from camera trap photos. *Wildl. Biol. Pract.* 9, 1–19. doi: 10.2461/wbp.2013.9.4

Milo, R., Shen-Orr, S., Itzkovitz, S., and Kashtan, N. (2002). Network motif: simple building blocks of complex networks. *Science* 298, 824–827. doi: 10.1126/science.298.5594.824

Monteiro, A. B., and Faria, L. D. B. (2016). The interplay between population stability and food-web topology predicts the occurrence of motifs in complex food-webs. *J. Theor. Biol.* 409, 165–171. doi: 10.1016/j.jtbi.2016.09.006

Moore, J. C., de Ruiter, P. C., McCann, K. S., and Wolters, V. (2017). in *Adaptive Food Webs: STABILITY and Transitions of Real and Model Ecosystems*, eds J. C. Moore, P. C. de Ruiter, K. S. McCann, and V. Wolters (Cambridge: Cambridge University Press), doi: 10.1017/9781316871867

Moorhouse-Gann, R. J., Dunn, J. C., De Vere, N., Goder, M., Cole, N., Hipperson, H., et al. (2018). New universal ITS2 primers for high-resolution herbivory analyses using DNA metabarcoding in both tropical and temperate zones. *Sci. Rep.* 8, 1–15. doi: 10.1038/s41598-018-26648-2

Munguía, M., Trejo, I., González-Salazar, C., and Pérez-Maqueo, O. (2016). Human impact gradient on mammalian biodiversity. *Glob. Ecol. Conserv.* 6, 79–92. doi: 10.1016/j.gecco.2016.01.004

Murphy, M. A., Kendall, K. C., Robinson, A., and Waits, L. P. (2007). The impact of time and field conditions on brown bear (*Ursus arctos*) faecal DNA amplification. *Conserv. Genet.* 8, 1219–1224. doi: 10.1007/s10592-006-9264-0

Naiman, R. J., Alldredge, J. R., Beauchamp, D. A., Bisson, P. A., Congleton, J., Henny, C. J., et al. (2012). Developing a broader scientific foundation for river restoration: columbia river food webs. *Proc. Natl. Acad. Sci. U.S.A.* 109, 21201–21207. doi: 10.1073/pnas.1213408109

Nakazawa, T. (2017). Individual interaction data are required in community ecology: a conceptual review of the predator-prey mass ratio and more. *Ecol. Res.* 32, 5–12. doi: 10.1007/s11284-016-1408-1

Oakmead Herbarium (2019). *Introdcution to the Vascular Palnt List, Oakmead Herbarium*. Available online at: https://jrbp.stanford.edu/content/introduction-vascular-plant-list-oakmead-herbarium (accessed February 10, 2019).

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., et al. (2019). *Package "vegan" Title Community Ecology Package. Community Ecol. Packag. 2, 1–297*. Available online at: https://cran.r-project.org/web/packages/vegan/vegan.pdf (accessed August 15, 2019).

Olivier, P., Frelat, R., Bonsdorff, E., Kortsch, S., Kröncke, I., Möllmann, C., et al. (2019). Exploring the temporal variability of a food web using long-term biomonitoring data. *Ecography* 42, 2107–2121. doi: 10.1111/ecog.04461

Pansu, J., Guyton, J. A., Potter, A. B., Atkins, J. L., Daskin, J. H., Wursten, B., et al. (2019). Trophic ecology of large herbivores in a reassembling African ecosystem. *J. Ecol.* 107, 1355–1376. doi: 10.1111/1365-2745.13113

Paulau, P. V., Feenders, C., and Blasius, B. (2015). Motif analysis in directed ordered networks and applications to food webs. *Sci. Rep.* 5, 12–15. doi: 10.1038/srep11926

Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéloz-Perret-Gentil, L., Beja, P., Boggero, A., et al. (2018). The future of biotic indices in the ecogenomic era: integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Sci. Total Environ.* 637–638, 1295–1310. doi: 10.1016/j.scitotenv.2018.05.002

Perea, R., López-Sánchez, A., and Dirzo, R. (2017). Differential tree recruitment in California oak savannas: are evergreen oaks replacing deciduous oaks? *For. Ecol. Manag.* 399, 1–8. doi: 10.1016/j.foreco.2017.05.018

Pianka, E. R. (1973). The structure of lizard communities. *Annu. Rev. Ecol. Syst.* 4, 53–74. doi: 10.1146/annurev.es.04.110173.000413

Pires, M. M. (2017). Rewilding ecological communities and rewiring ecological networks. *Perspect. Ecol. Conserv.* 15, 257–265. doi: 10.1016/j.pecon.2017.09.003

Pompanon, F., Deagle, B. E., Symondson, W. O. C., Brown, D. S., Jarman, S. N., and Taberlet, P. (2012). Who is eating what: diet assessment using next generation sequencing. *Mol. Ecol.* 21, 1931–1950. doi: 10.1111/j.1365-294X.2011.05403.x

Power, M. E., Tilman, D., Estes, J. A., Menge, B. A., Bond, W. J., Mills, L. S., et al. (1996). Challenges in the Quest for Keystones. *Bioscience* 46, 609–620. doi: 10.2307/1312990

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundations for Statistical Computing. Available online at: https://www.r-project.org/ (accessed May 26, 2020).

Ratnasingham, S., and Hebert, P. D. N. (2007). The Barcode of Life Data System. *Mol. Ecol. Notes* 7, 355–364. doi: 10.1111/j.1471-8286.2006.01678.x

Razgour, O., Clare, E. L., Zeale, M. R. K., Hanmer, J., Schnell, I. B., Rasmussen, M., et al. (2011). High-throughput sequencing offers insight into mechanisms of resource partitioning in cryptic bat species. *Ecol. Evol.* 1, 556–570. doi: 10.1002/ece3.49

Riccioni, G., Stagioni, M., Piccinetti, C., and Libralato, S. (2018). A metabarcoding approach for the feeding habits of European hake in the Adriatic Sea. *Ecol. Evol.* 8, 10435–10447. doi: 10.1002/ece3.4500

Robeson, M. S., Khanipov, K., Golovko, G., Wisely, S. M., White, M. D., Bodenchuck, M., et al. (2018). Assessing the utility of metabarcoding for diet analyses of the omnivorous wild pig (*Sus scrofa*). *Ecol. Evol.* 8, 185–196. doi: 10.1002/ece3.3638

Roslin, T., Traugott, M., Jonsson, M., Stone, G. N., Creer, S., and Symondson, W. O. C. (2019). Introduction: special issue on species interactions, ecological networks and community dynamics – Untangling the entangled bank using molecular techniques. *Mol. Ecol.* 28, 157–164. doi: 10.1111/mec.14974

Rowe, R. J., and Terry, R. C. (2014). Small mammal responses to environmental change: integrating past and present dynamics. *J. Mammal.* 95, 1157–1174. doi: 10.1644/13-MAMM-S-079

Ruell, E. W., Riley, S. P. D., Douglas, M. R., Pollinger, J. P., and Crooks, K. R. (2009). Estimating bobcat population sizes and densities in a fragmented urban landscape using noninvasive capture–recapture sampling. *J. Mammal.* 90, 129–135. doi: 10.1644/07-MAMM-A-249.1

Ruppert, K. M., Kline, R. J., and Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: a systematic review in methods, monitoring, and applications of global eDNA. *Glob. Ecol. Conserv.* 17:e00547. doi: 10.1016/j.gecco.2019.e00547

Saint-béat, B., Baird, D., Saint-béat, B., Baird, D., Asmus, H., Asmus, R., et al. (2015). Trophic networks: how do theories link ecosystem structure and functioning to stability properties? A review Trophic networks: how do theories link ecosystem structure and functioning to stability properties? A review. *Ecol. Indic.* 52, 458–471. doi: 10.1016/j.ecolind.2014.12.017

Sharma, C. M., and Khanduri, V. P. (2012). Pollen cone characteristics, pollen yield and pollen-mediated gene flow in Cedrus deodara. *Curr. Sci.* 102, 394–397.

Shehzad, W., Riaz, T., Nawaz, M. A., Miquel, C., Poillot, C., Shah, S. A., et al. (2012). Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. *Mol. Ecol.* 21, 1951–1965. doi: 10.1111/j.1365-294X.2011.05424.x

Siegenthaler, A., Wangensteen, O. S., Benvenuto, C., Campos, J., and Mariani, S. (2019). DNA metabarcoding unveils multiscale trophic variation in a widespread coastal opportunist. *Mol. Ecol.* 28, 232–249. doi: 10.1111/mec.14886

Smith, J. A., Thomas, A. C., Levi, T., Wang, Y., and Wilmers, C. C. (2018). Human activity reduces niche partitioning among three widespread mesocarnivores. *Oikos* 127, 890–901. doi: 10.1111/oik.04592

Soininen, E. M., Valentini, A., Coissac, E., Miquel, C., Gielly, L., Brochmann, C., et al. (2009). Analysing diet of small herbivores: the efficiency of DNA barcoding coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures. *Front. Zool.* 6:16. doi: 10.1186/1742-9994-6-16

Sollmann, R., Tôrres, N. M., Furtado, M. M., De Almeida Jácomo, A. T., Palomares, F., Roques, S., et al. (2013). Combining camera-trapping and noninvasive genetic data in a spatial capture-recapture framework improves density estimates for the jaguar. *Biol. Conserv.* 167, 242–247. doi: 10.1016/j.biocon.2013.08.003

Sousa, L. L., Silva, S. M., and Xavier, R. (2019). DNA metabarcoding in diet studies: unveiling ecological aspects in aquatic and terrestrial ecosystems. *Environ. DNA* 1, 199–214. doi: 10.1002/edn3.27

Start, D., Bonner, C., Weis, A. E., and Gilbert, B. (2018). Consumer-resource interactions along urbanization gradients drive natural selection. *Evolution* 72, 1863–1873. doi: 10.1111/evo.13544

Stouffer, D. B., and Bascompte, J. (2010). Understanding food-web persistence from local to global scales. *Ecol. Lett.* 13, 154–161. doi: 10.1111/j.1461-0248.2009.01407.x

Taberlet, P., Bonin, A., Zinger, L., and Coissac, E. (2018). *Environmental DNA – For Biodiversity Research and Monitoring - Oxford Scholarship Online Find in Library Find in Worldcat Environmental DNA: For Biodiversity Research and Monitoring*. Oxford: Oxford University Press, doi: 10.1093/OSO/9780198767220.001.0001

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., et al. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Res.* 35:e14. doi: 10.1093/nar/gkl938

Thompson, R. M., Brose, U., Dunne, J. A., Hall, R. O., Hladyz, S., Kitching, R. L., et al. (2012). Food webs: reconciling the structure and function of biodiversity. *Trends Ecol. Evol.* 27, 689–697. doi: 10.1016/j.tree.2012.08.005

Torre, I., Arrizabalaga, A., Freixas, L., Ribas, A., Flaquer, C., and Díaz, M. (2013). Using scats of a generalist carnivore as a tool to monitor small mammal communities in Mediterranean habitats. *Basic Appl. Ecol.* 14, 155–164. doi: 10.1016/j.baae.2013.01.005

Tylianakis, J. M., Laliberté, E., Nielsen, A., and Bascompte, J. (2010). Conservation of species interaction networks. *Biol. Conserv.* 143, 2270–2279. doi: 10.1016/j.biocon.2009.12.004

Ushio, M., Fukuda, H., Inoue, T., Makoto, K., Kishida, O., Sato, K., et al. (2017). Environmental DNA enables detection of terrestrial mammals from forest pond water. *Mol. Ecol. Resour.* 17, e63–e75. doi: 10.1111/1755-0998.12690

Valiente-Banuet, A., Aizen, M. A., Alcántara, J. M., Arroyo, J., Cocucci, A., Galetti, M., et al. (2015). Beyond species loss: the extinction of ecological interactions in a changing world. *Funct. Ecol.* 29, 299–307. doi: 10.1111/1365-2435.12356

Walsh, A. J. (2015). *Carnivore Diet Identification Through Scat and Genetic Analysis in Namibia, Africa*. Durham, NH: University of New Hampshire.

Wansi, T., Pieper, R., Beck, R., and Murray, L. (1992). Botanical content of black-tailed jackrabbit diets on semidesert rangeland. *Gt. Basin Nat.* 52, 300–308.

Wilson, J., and Thomas, B. (1999). Diet and Seed dispersal efficiency of the gray fox (*Urocyon cinereoargenteus*) in chaparral. *Bull. South. Calif. Acad. Sci.* 98, 119–126.

Winemiller, K. O., and Pianka, E. R. (1990). Organization in natural assemblages of desert lizards and tropical fishes. *Ecol.Monogr.* 60, 27–55. doi: 10.2307/1943025

Woodward, G., and Hildrew, A. G. (2002). Food web structure in riverine landscapes. *Freshw. Biol.* 47, 777–798. doi: 10.1046/j.1365-2427.2002.00908.x

Wootton, J. T. (1994). The nature and consequences of indirect effects in ecological communities. *Annu. Rev. Ecol. Syst.* 25, 443–466. doi: 10.1146/annurev.es.25.110194.002303

Xiong, M., Wang, D., Bu, H., Shao, X., Zhang, D., Li, S., et al. (2017). Molecular dietary analysis of two sympatric felids in the Mountains of Southwest China biodiversity hotspot and conservation implications. *Sci. Rep.* 7, 1–12. doi: 10.1038/srep41909

Zavaleta, E. S., and Kettley, L. S. (2006). Ecosystem change along a woody invasion chronosequence in a California grassland. *J. Arid Environ.* 66, 290–306. doi: 10.1016/j.jaridenv.2005.11.008

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., et al. (2019). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Mol. Ecol.* 28, 1857–1862. doi: 10.1111/mec.15060

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

SYSTEMATIC REVIEW
published: 07 December 2020
doi: 10.3389/fevo.2020.609973

The Elephant in the Lab (and Field): Contamination in Aquatic Environmental DNA Studies

Adam J. Sepulveda, Patrick R. Hutchins, Meghan Forstchen†, Madeline N. Mckeefry† and Anna M. Swigris†*

U.S. Geological Survey, Northern Rocky Mountain Science Center, Bozeman, MT, United States

OPEN ACCESS

Edited by:
Katy Klymus, United States Geological Survey (USGS), United States

Reviewed by:
Anna McKee, USGS South Atlantic Water Science Center, United States
Allan D. McDevitt, University of Salford, United Kingdom

***Correspondence:**
Adam J. Sepulveda
asepulveda@usgs.gov

†Present address:
Meghan Forstchen, University of Notre Dame, Notre Dame, IN, United States
Madeline N. Mckeefry, National Park Service, Yellowstone National Park, Mammoth, WY, United States
Anna M. Swigris, South Florida Water Management District, West Palm Beach, FL, United States

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *24 September 2020*
Accepted: *16 November 2020*
Published: *07 December 2020*

Citation:
Sepulveda AJ, Hutchins PR, Forstchen M, Mckeefry MN and Swigris AM (2020) The Elephant in the Lab (and Field): Contamination in Aquatic Environmental DNA Studies. Front. Ecol. Evol. 8:609973. doi: 10.3389/fevo.2020.609973

The rapid evolution of environmental (e)DNA methods has resulted in knowledge gaps in smaller, yet critical details like proper use of negative controls to detect contamination. Detecting contamination is vital for confident use of eDNA results in decision-making. We conducted two literature reviews to summarize (a) the types of quality assurance measures taken to detect contamination of eDNA samples from aquatic environments, (b) the occurrence, frequency and attribution (i.e., putative sources) of unexpected amplification in these quality assurance samples, and (c) how results were interpreted when contamination occurred. In the first literature review, we reviewed 156 papers and found that 91% of targeted and 73% of metabarcoding eDNA studies reported inclusion of negative controls within their workflows. However, a large percentage of targeted (49%) and metabarcoding (80%) studies only reported negative controls for laboratory procedures, so results were potentially blind to field contamination. Many of the 156 studies did not provide critical methodological information and amplification results of negative controls. In our second literature review, we reviewed 695 papers and found that 30 targeted and 32 metabarcoding eDNA studies reported amplification of negative controls. This amplification occurred at similar proportions for field and lab workflow steps in targeted and metabarcoding studies. These studies most frequently used amplified negative controls to delimit a detection threshold above which is considered significant or provided rationale for why the unexpected amplifications did not affect results. In summary, we found that there has been minimal convergence over time on negative control implementation, methods, and interpretation, which suggests that increased rigor in these smaller, yet critical details remains an outstanding need. We conclude our review by highlighting several studies that have developed especially effective quality assurance, control and mitigation methods.

Keywords: aquatic, false positive, metabarcoding, negative control (NC), PCR, targeted, review

INTRODUCTION

Environmental (e)DNA refers to sampling and detection techniques for deoxyribonucleic acid (DNA) released by organisms into the environment (e.g., water, soil, or air). The DNA can be queried for specific taxa in targeted techniques (Ficetola et al., 2008) or can be surveyed for many taxonomic groups with metabarcoding approaches (Thomsen et al., 2012). Because

most eDNA methods utilize polymerase chain reaction (PCR) of relatively short DNA fragments (generally < 200 nucleotides), these methods are sensitive enough to detect DNA at extremely low concentrations. This sensitivity is a key advantage, as it allows users to make inferences about taxa presence even when they are at abundances too low to be detected by traditional, non-molecular techniques. However, this extreme sensitivity presents a challenge, as it heightens susceptibility to contaminating DNA.

The rapid evolution of eDNA methods over the past decade has resulted in knowledge gaps in smaller, yet critical details. Here, we argue that contamination detection is a critical detail that has been overlooked, but is deserving of attention, especially as eDNA methods transition from research to application. Detecting contamination of eDNA samples is vital for confident use of eDNA results in natural resource management, as positive eDNA results can initiate a costly chain of control and containment actions. Costly actions based on false positives can cause decision-makers to question the use of eDNA as a monitoring tool (Jerde, 2019; Sepulveda et al., 2020). Effective means for detecting contamination are needed to not only inform potentially costly management decisions, but to also identify and strengthen weak points in current workflow protocols.

Those using eDNA sampling have been combating contamination since the inception of these techniques, and many guidelines and procedures have been developed to prevent, detect, and quantify false positives resulting from contamination at every stage of the work flow. Most eDNA research and monitoring programs have instituted general molecular best practices to minimize contamination potential in the field (e.g., single-use supplies, bleach sterilization) and in the lab (e.g., separation of low-template vs. high-template DNA work spaces), as described in Goldberg et al. (2016). However, these best practices are imperfect—multiple published examples report unexpected amplification of negative control samples despite adherence to best practices to minimize contamination (e.g., Maruyama et al., 2014; Serrao et al., 2018; Sepulveda et al., 2019b). For example, Serrao et al. (2018) analyzed 258 negative controls samples for redside dace (*Clinostomus elongatus*) DNA and found that 30% of samples amplified, though 98.4% of these samples had less than 1 copy reaction^{-1}. Similarly, Sepulveda et al. (2019b) analyzed 619 samples for dreissenid mussel (*Dreissena* spp.) DNA and two negative field control samples amplified despite the nearest known dreissenid population being > 1,000 km away. Additionally, some researchers with quality assurance results indicative of contamination likely opt to not publish studies, thus slowing the progress of the field.

Outside of the general guidance of analyzing field and laboratory negative controls, specific guidance for contamination detection does not exist. Consequently, a broad range of approaches are currently used to detect contamination in the field and laboratory. Some programs ensure a minimum of 10% of samples collected are field blanks comprised of target DNA-free water handled in the field (e.g., Woldt et al., 2019), while others assess field contamination only by analyzing field samples where the target species is presumed absent (e.g., Carim et al., 2019).

Moreover, there is no clear guidance on how to proceed when negative control samples amplify. Some studies discarded associated field samples (Sepulveda et al., 2019b), others attributed unexpected amplification to random noise and ignored the amplified negative controls (Maruyama et al., 2014), while others established a "limit of blank" that delimited detection thresholds above which was considered significant (Serrao et al., 2018). Additional examples in the peer-reviewed literature can be found that follow each of these paths, producing confusion and doubt for researchers and decision-makers alike. The need for clarification on how to proceed is elevated when associated field samples also amplify for the target DNA, as these detections could be true positives. A better understanding of the known or potential rate of error and standards for the control of the technique's operation are required for eDNA results to be considered reliable scientific evidence (Sepulveda et al., 2020).

We conducted a literature review to summarize (a) the types of quality assurance measures taken to detect contamination of eDNA samples, (b) the occurrence, frequency, and attribution (i.e., putative sources) of unexpected amplification in these quality assurance samples, and (c) how results were interpreted when contamination occurred. We also assess how these response variables have changed since 2008, when eDNA approaches were initially used to detect aquatic macroorganisms (Ficetola et al., 2008). Convergence in quality assurance measures and a decrease in the frequency and occurrence of contamination would suggest general agreement in best practices and that these best practices are effective, while divergence in quality assurance measures and an increase in contamination would suggest that contamination detection is still a critical detail deserving of attention.

MATERIALS AND METHODS

We conducted two literature reviews to synthesize contamination detection methods, contamination occurrence, and result interpretation related to the eDNA detection of aquatic organisms in peer-reviewed studies published between January 2008 and April 2020. Both reviews were inclusive of targeted and metabarcoding approaches across freshwater and marine environments. The objectives of the first review were to document the quality assurance measures used for contamination detection and to estimate how frequently amplification of negative controls has been reported. The objectives of the second review were to identify which stages of the eDNA workflow have been most susceptible to contamination and to document how evidence of contamination influenced result interpretations.

We used Web of Science to conduct the first literature review. The following topical terms had to appear in an article's title, abstract, or keywords: "environmental DNA" AND "aquatic*" OR "water*" OR "marine" OR "ocean*" OR "estuary." We reviewed the abstract of each article to ensure that it was applicable and included primary data (i.e., not a review paper). This resulted in 876 entries (**Supplementary Figure 1**). We randomly sampled 25% of the articles that were published each year from these 876 entries (**Supplementary Table 1**). We then reviewed each article and recorded the information listed in

TABLE 1 | The occurrence of factors within each category was recorded for all studies in Literature Review #1 or #2.

Review	Category		Factors
1, 2	Publication year		Year study was published
1, 2	Geographic location		Continent or major ocean
1, 2	Environment		Freshwater, marine
1, 2	Habitat		Deep ocean, coastal ocean, surface ocean, pond/wetland, river, lake, other
1, 2	Study type		Field, mesocosm, lab
1, 2	eDNA approach		Targeted or metabarcoding
1, 2	Principle analytical tool		Conventional PCR, digital droplet PCR, quantitative PCR, sequencing, other
1, 2	Negative control type		Field sampling, DNA capture, DNA extraction, PCR, other
1, 2	For each negative control type:	Negative control material	Deionized water, distilled water, double distilled water, molecular grade water, tap water, other water, negative environmental site, extraction kit reagent, filter
1, 2		Ratio of negative controls to samples	The number of negative controls relative to samples
1, 2		Did negative controls amplify	0, 1
2		How did amplification of negative controls influence results	No mention, rationalized then ignored, used to inform limit of blank, removed associated field samples, other

Table 1. Any article that was deemed not applicable when read in full was replaced by a randomly selected article published the same year.

Web of Science was not a useful tool for our second literature review because amplification of negative controls was seldom mentioned in a study's title, abstract or keywords. Consequently, we used Google Scholar to conduct our second literature review since this tool searches within the text as well as the title, abstract and keywords. Comparable search terms in Google Scholar as used in the previous Web of Science literature review returned over 8,000 papers. Thus, we used studies that were already filtered by Tsuji et al. (2019) in a recent review of eDNA detection methods for aquatic macroorganisms. These authors used a Google Scholar search for studies published between 2008 and 2018 that including the keywords: "eDNA" and "environmental DNA". The search results were filtered by hand to 388 papers based on the following criteria: (1) detection of macroorganisms (not micoorganisms, virus, or bacteria); (2) published in international journals (except preprint servers); and (c) peer-reviewed. To update papers published between 2019 and April 2020, we repeated these methods which resulted in an additional 307 papers (**Supplementary Table 2**). We then used Google Scholar to search within these papers for the following terms: "false positive*" OR "contaminat*." This resulted in 193 articles, but after further review of these articles, we found that only 64 articles reported negative control samples that amplified (**Supplementary Figure 2**). We then recorded the information listed in **Table 1** for these 64 articles and for any articles in our first literature review that did not appear in our Google Scholar search, yet had amplification of negative controls.

We summarized the number of studies per publication year, geographic location, environment, habitat, study type, and eDNA approach to place the reviewed studies into appropriate context. We then grouped the remaining data fields by eDNA approach (targeted or metabarcoding), since these approaches are used to address different types of study objectives and the potential for amplification of negative controls is much greater in metabarcoding approaches. For each eDNA approach group, we calculated the frequency of occurrence of factors within each data category (**Table 1**). We binned eDNA workflow steps into the following categories: field sampling; DNA capture, defined as the concentration of DNA material using filtration or centrifugation; DNA extraction; PCR and; other. We report results using descriptive statistics (e.g., frequency of occurrence, percentages) because sample sizes were often too small for inferential statistics, thus our results are not generalizable to all eDNA studies.

RESULTS

Literature Review 1

We reviewed 155 papers that met our inclusion criteria (**Supplementary Figure 1**). The number of studies using eDNA methods nearly doubled each year since 2012, consequently 80% of the studies that we reviewed were published between 2016 and April 2020 (**Figure 1**). The most commonly reviewed studies used targeted eDNA methods and took place in North American and European freshwater, lotic ecosystems (**Table 2**). Quantitative PCR analysis was the dominant analytical platform used for targeted eDNA studies (**Table 2**).

One-hundred (91%) of the 110 targeted eDNA studies and 33 (73%) of the 45 metabarcoding studies reported collection of at least one negative control (**Figure 1**). PCR controls were reported in 71%, field controls were reported in 51%, extraction controls were reported in 36%, and DNA capture controls were reported in 25% of the 100 targeted eDNA studies that collected negative controls (**Figure 2**). PCR controls were reported in 71%, extraction controls were reported in 44%, and field controls and DNA capture controls were each reported in 20% of the 33 metabarcoding studies that collected negative controls (**Figure 2**). The reporting of other negative control categories (e.g., travel controls) was less common (1–12%). We documented high annual fluctuations in the percent of targeted and metabarcoding studies that included controls from the most common categories (e.g., PCR and field; **Figure 2**). These fluctuations did not dampen over time. Temporal trends for metabarcoding studies were especially vague since most years had few studies.

Targeted and metabarcoding eDNA studies used a wide variety of water sources for field negative controls (**Supplementary Figure 3**). Deionized (28% of studies with field negative controls), environmental (27%) and distilled water (16%) were most common in targeted studies across years.

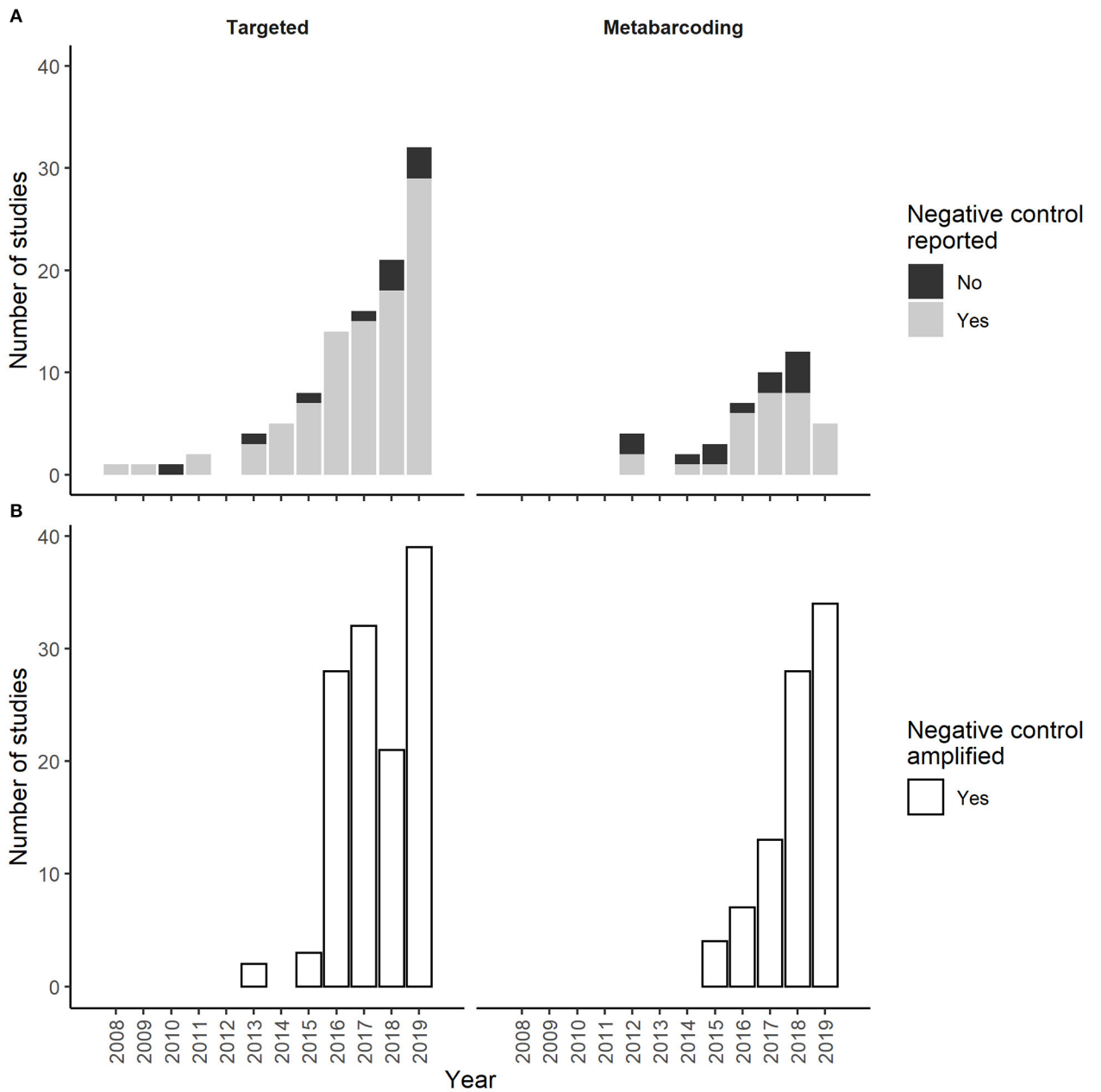


FIGURE 1 | The number of targeted and metabarcoding eDNA studies that **(A)** reported (Yes) or did not report (No) negative controls in the first literature review and that **(B)** reported (Yes) amplification of negative controls in the second literature review.

Distilled (22%) and tap water (22%) were most common in metabarcoding studies across years, though a similar proportion of studies (22%) did not report the water source.

We first documented reporting of negative control amplification in targeted and metabarcoding eDNA studies in 2016 (**Figure 3**). Thereafter, negative control amplification was reported in ∼6% of targeted studies and 25% of metabarcoding studies that included negative controls each year. Many targeted and metabarcoding studies that reported use of negative controls failed to report negative control results (**Figure 3**). Amplification was reported at similar low proportions across all negative control categories for targeted and metabarcoding studies (**Figure 4**).

Most studies failed to provide explicit data on the ratio of negative controls samples to field samples. For example, studies reported that extraction controls were collected per batch of extractions but failed to report the number of extraction batches. Raw data were not always publicly accessible, and

TABLE 2 | The proportion of studies in Literature Review #1 ($N = 155$) and #2 ($N = 62$) by geographic location, ecosystem and habitat, and eDNA approach.

Location			Ecosystem and habitat				eDNA approach			
Review	#1	#2			#1	#2			#1	#2
Africa	0.03	0.03	Freshwater		0.72	0.70	Targeted		0.71	0.48
Asia	0.18	0.05		Lentic	0.36	0.48		cPCR	0.13	0.15
Europe	0.27	0.20		Lotic	0.49	0.38		qPCR	0.67	0.70
North America	0.36	0.48		Lab	0.07	0.06		ddPCR	0.07	0.12
Oceania	0.10	0.13		Other	0.08	0.08		other	0.13	0.03
South America	0.01	0.03	Marine		0.25	0.30	Metabarcoding		0.29	0.52
Antarctic Ocean	0.00	0.02		Ballast	0.05	0.00				
Atlantic Ocean	0.01	0.00		Coastal	0.48	0.68				
Indian Ocean	0.01	0.02		Lab	0.00	0.05				
Pacific Ocean	0.01	0.02		Pelagic, deep	0.23	0.05				
Global	0.02	0.02		Pelagic, surface	0.08	0.16				
				Other	0.05	0.05				
			Brackish		0.03	0.00				

when these data were available, negative control results were infrequently reported.

Literature Review 2

We reviewed 62 studies that met our inclusion criteria of unexpected negative control amplification (**Supplementary Figure 2**). Thirty of these studies used targeted eDNA methods and 32 used metabarcoding eDNA methods (**Figure 1**). Twenty-two of the 30 targeted eDNA studies and 19 of the 32 metabarcoding eDNA studies provided enough description to attribute amplification to a specific negative control category (e.g., field or PCR). The other studies only reported the general occurrence of unexpected negative control amplification. The characteristics of studies reviewed were similar to those reviewed in the first literature review (**Table 1**), with the exception of the proportional representation of targeted vs. metabarcoding studies. In this second review, metabarcoding studies were as common as targeted studies.

For targeted studies, amplification was reported most frequently in field negative controls (**Figure 4**). For metabarcoding studies, contamination was reported most frequently in PCR negative controls (**Figure 4**). Amplification was reported in a variety of field negative control water sources, but sample sizes were too small to assess if specific water sources amplified more frequently than others. Similar to the first literature review, most studies failed to provide explicit data on the ratio of negative controls samples to field samples so it was not possible to characterize negative control effort.

We documented a variety or study responses to negative control amplification. Most targeted eDNA studies provided rationale for why the unexpected amplifications did not affect results; whereas, metabarcoding studies used amplified negative controls to delimit a detection threshold above which is considered significant (**Figure 5**). Fewer studies removed samples that were associated with negative controls that amplified or failed to provide rationale for why these results could be ignored.

DISCUSSION

A substantial number of eDNA studies from across the globe have been published in the past 12 years, which underscores the rapid technological advancements in this field and the applicability of eDNA methods to a broad range of taxa and habitats. Yet the inherent sensitivity of eDNA methods to contamination is a principal reason for why managers have been reluctant to use eDNA results for decision-making. Consequently, managers and eDNA practitioners have called for increased rigor in quality assurance and control measures to prevent and detect contamination (Loeza-Quintana et al., 2020; Sepulveda et al., 2020). Our review of eDNA studies published over the past 12 years suggests that this call for increased rigor remains an outstanding need.

We reviewed ∼25% of eDNA studies published each year, 2008–2020, and found 100 of 110 targeted studies and 33 of 45 metabarcoding studies reported inclusion of negative controls within their workflow (**Figure 1**). It is encouraging that most eDNA practitioners have included quality assurance methods in their workflows, but it is disconcerting that many studies failed to report method specifics, such as the negative control water source and the number of negative control samples analyzed. This result is in line with Dickie et al. (2018), which found that 95% of reviewed eDNA metabarcoding studies failed to provide critical methodological information required for reproducibility by independent researchers. These kinds of omissions set up the potential for a replication crisis that has hampered the advancement of other disciplines. Moreover, these omissions make it difficult to discern general best practices and to identify workflow steps that are consistently susceptible to contamination and therefore require improved quality assurance.

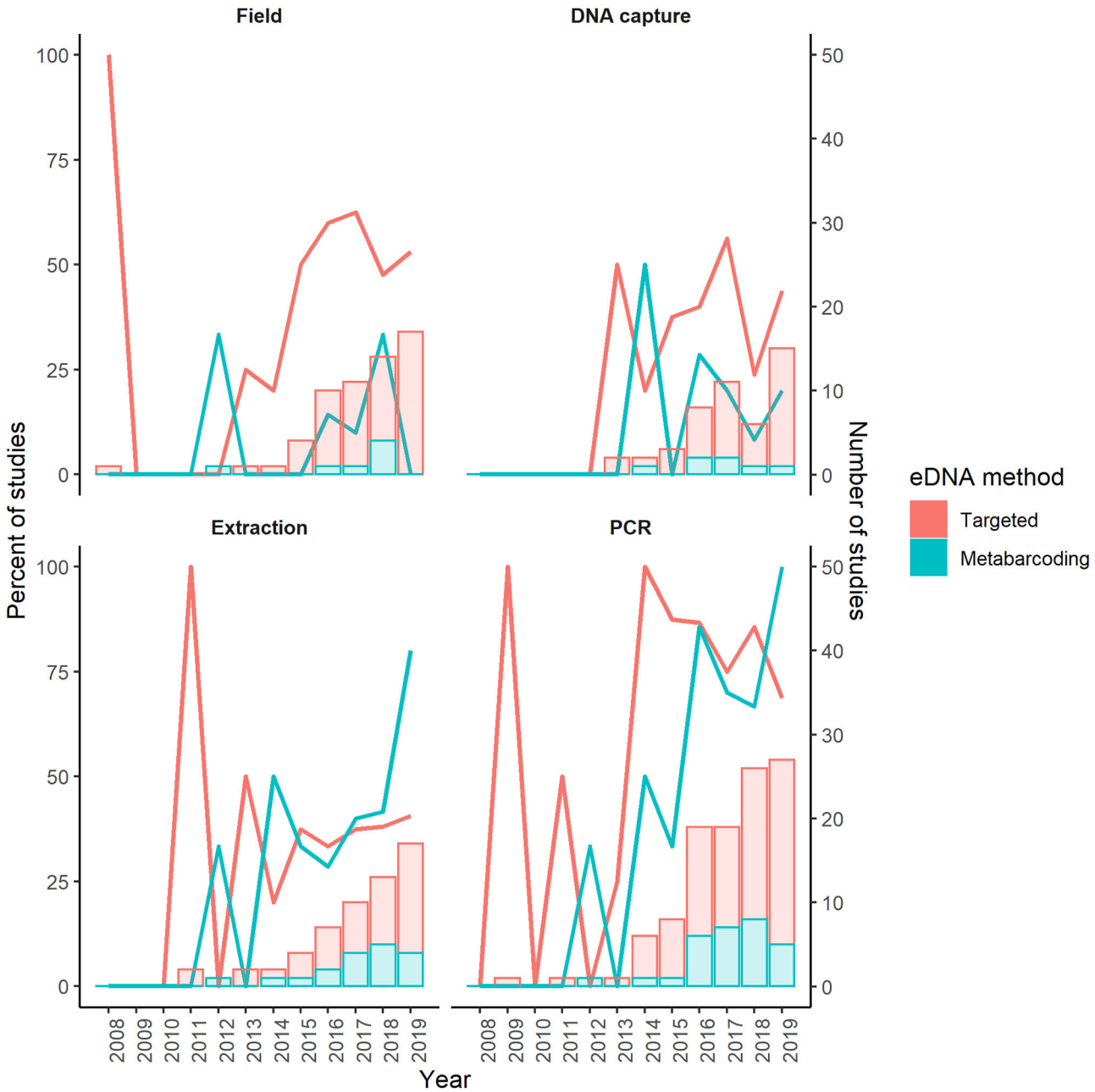


FIGURE 2 | Temporal trends in the percent of studies with reported negative controls associated with field, DNA capture, extraction, and PCR workflow steps in targeted and metabarcoding eDNA studies in the first literature review. Lines indicate the percent of studies and columns indicate the number of studies for a given year.

The negative control methods that were described with enough detail to be reproduced varied greatly among studies, and this variability has not decreased over time despite calls for standardization (**Figure 2**; Loeza-Quintana et al., 2020; Minamoto et al., 2020). For example, field and PCR negative controls used a wide variety of water types and collection schemes that are likely influenced by study objectives (**Supplementary Figure 3**). Examples included collecting field negative controls once per site vs. once per day, as the first vs. last sample collected at a site, or laboratory (e.g., deionoized water) vs. environmental (e.g., presumed negative field site) water sources. Different negative control methods may provide similar results when contamination is systemic (e.g., contamination of laboratory reagents), but it is unknown how these methods vary in their ability to reliably detect cross-contamination among samples collected from multiple

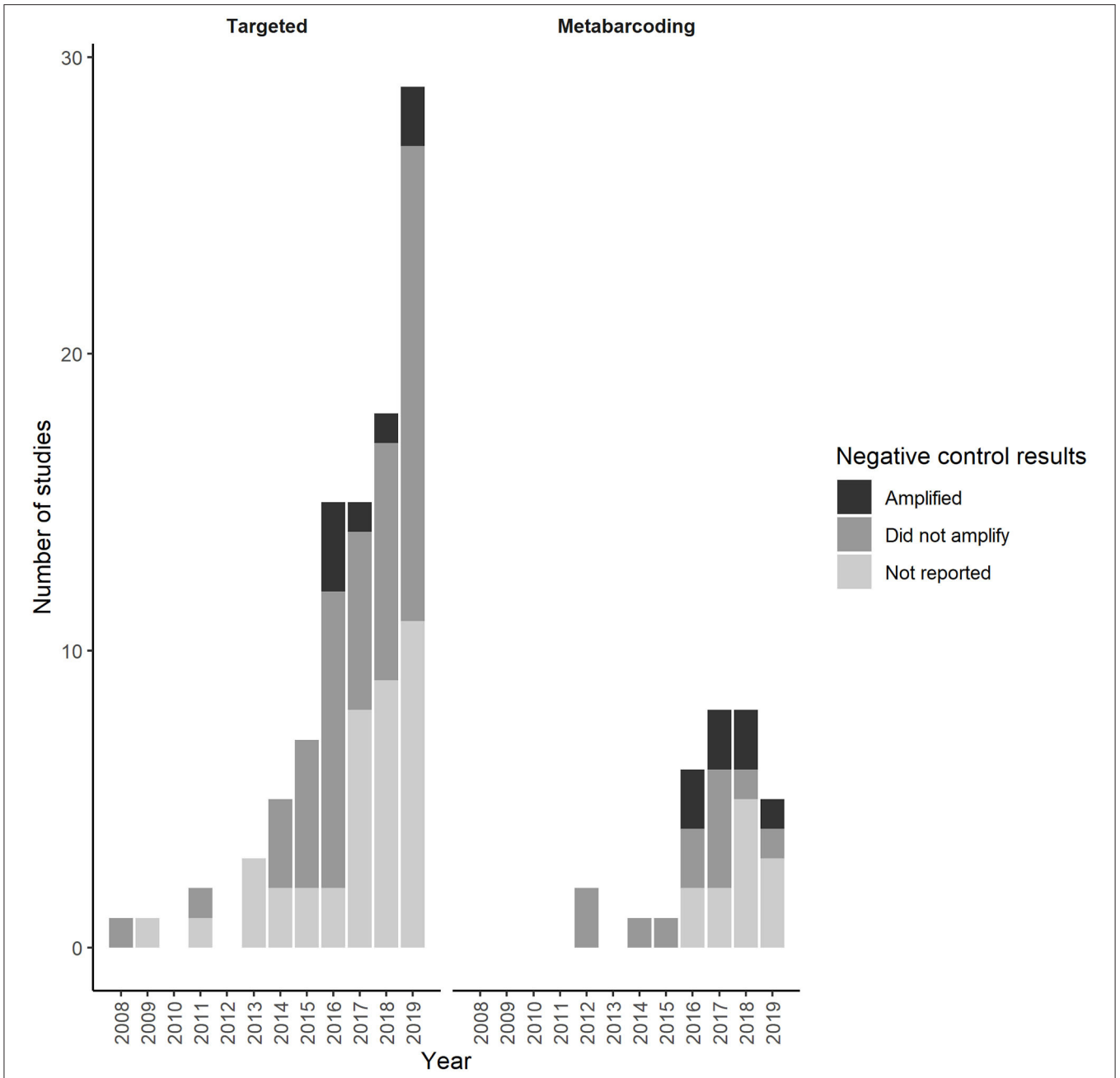


FIGURE 3 | Annual trends in the number of studies in the first literature review that reported amplification of negative controls (Amplified), no amplification of negative controls (Did not amplify), or collected negative controls but did not report amplification results (Not reported).

sites on the same day and random contamination (i.e., that which does not affect all samples in a batch equitably). Our review indicates that systemic contamination is rare, as amplification was only reported in a small subset of negative controls per study. For example, Guillera-Arroita et al. (2017) filtered and extracted 50 negative controls in the lab and found that 2–8% of these samples amplified for the DNA of four target amphibian species. Indeed, it is likely that studies with systemic contamination never make it to publication.

Overall, the number of studies reporting amplification of negative controls was low; ∼6% of targeted studies and 25% of metabarcoding studies that included negative controls each year (**Figure 3**). While these low percentages seem reassuring, we suspect that they are underestimates for at least two reasons. First, a large percentage of targeted (49%) and metabarcoding (80%) studies limited negative controls to laboratory procedures (**Figure 4**). These studies were blind to any contamination that may have occurred during field collection, transport to the lab and DNA capture (e.g., filtration). This is a surprising

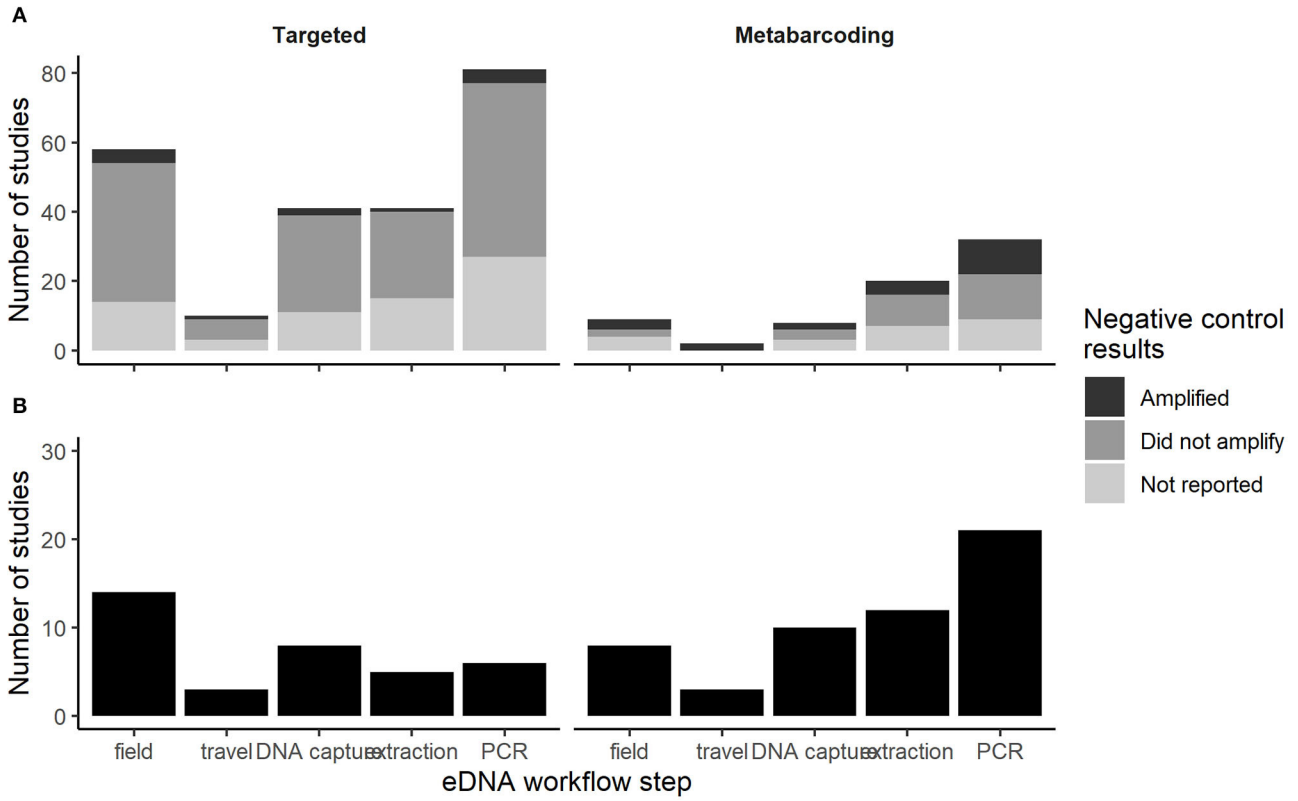


FIGURE 4 | The number of studies in the first **(A)** and second **(B)** literature reviews with eDNA workflow steps associated with negative controls that amplified, did not amplify, or amplification results were not reported.

omission given the attention to developing field protocols (e.g., single-use supplies; Spens et al., 2017) that reduce risk of cross-contamination. Indeed, multiple papers over the past decade have indicated that the inclusion of controls throughout the entire eDNA workflow is required for strong inference about species presence (Darling and Mahon, 2011; Goldberg et al., 2016; Jerde, 2019; Sepulveda et al., 2019b). Our results from the second literature review support this recommendation since amplification of pre-lab workflow negative controls occurred as much or more frequently than lab workflow negative controls (**Figure 4**). However, amplification of pre-lab workflow negative controls does not unequivocally indicate contamination occurred prior to the lab since these samples are also susceptible to lab contamination. Second, metabarcoding studies especially may have higher rates of contamination than we documented because the discipline of DNA metabarcoding (inclusive of eDNA and DNA samples) has only recently become aware of multiple contamination issues that can cause incorrect assignment of sequences to samples, including false tag combinations in the sequencing output (Schnell et al., 2015) and amplicon contamination (Schnell et al., 2015; Ballenghien et al., 2017). We found considerably more agreement among reviewed studies on how to proceed when negative control samples had unexpected amplification. The majority of targeted eDNA studies attributed the amplification to low-level noise and ignored unexpected amplification (**Figure 5**), whereas most metabarcoding studies used the quantitative information provided by analyses to delimit a detection threshold above which is considered significant (**Figure 5**). Consensus was stronger in metabarcoding studies, which had much higher occurrence of negative control amplification (and/or sequence reads) and a general acceptance that low-level contamination is unavoidable.

These *ad hoc* approaches for dealing with unexpected amplification have been criticized, given that they are subjective and can lead to underestimation of species occurrences (Ficetola et al., 2016; Lahoz-Monfort et al., 2016). Site occupancy-detection models (SODMs) provide a more objective means of accounting for detection errors caused by false positives. Contamination rates derived from amplified negative controls (Ficetola et al., 2016), calibration experiments that explicitly assess contamination rates at different steps of the eDNA work flow (e.g., Guillera-Arroita et al., 2017), or unambiguous eDNA data collected from sites with known absences (Lahoz-Monfort et al., 2016; Smith and Goldberg, 2020) can be used to parameterize SODMs. Model output informs the probability that an eDNA detection is a true presence given the number of detections. However, SODMs that account for false positives have infrequently been used in eDNA studies because they are relatively new, computationally intensive advancements. It is unclear how managers and other eDNA result end-users would integrate false positive probabilities into decision making.

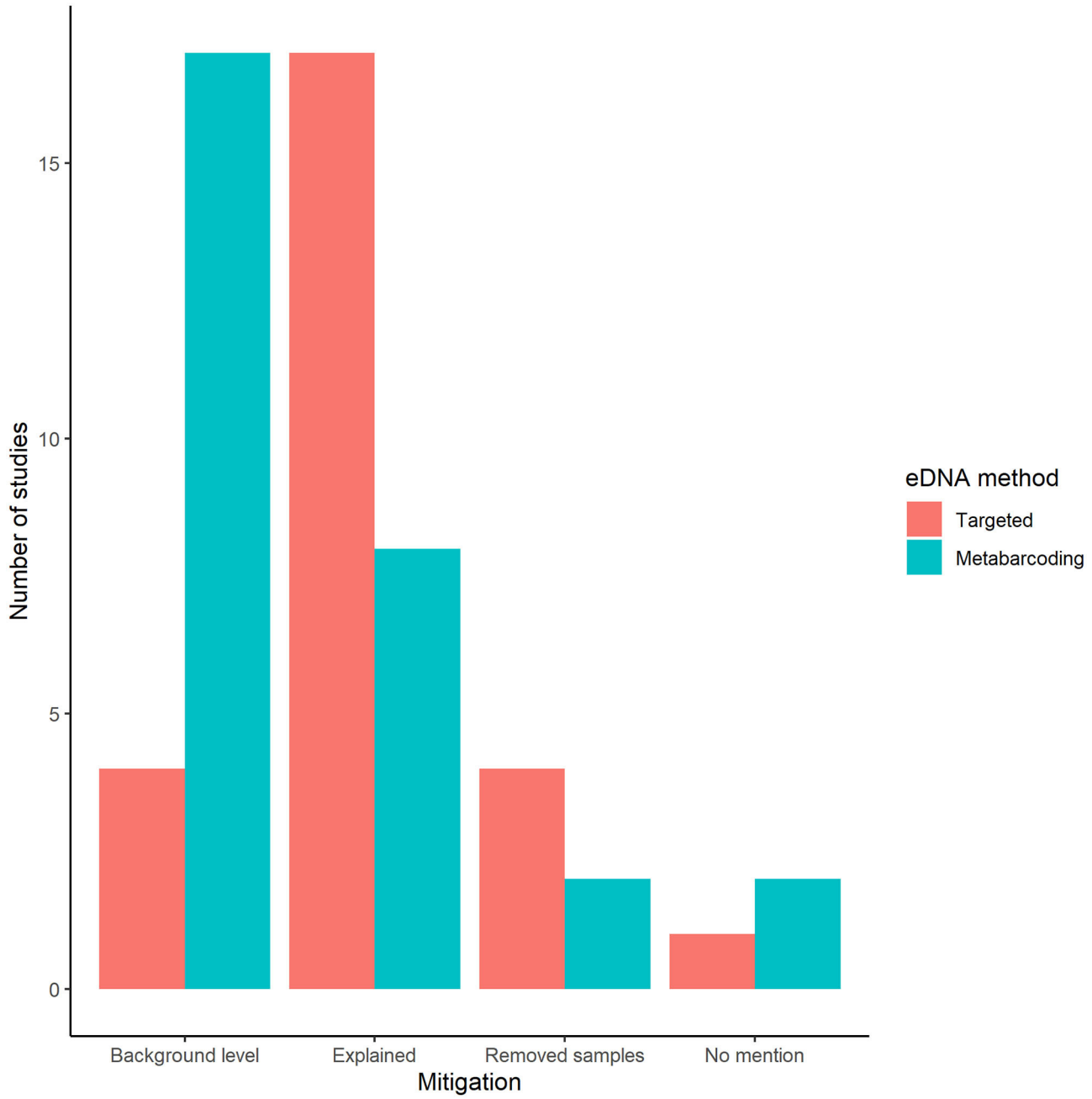


FIGURE 5 | The number of studies in the second literature review that used amplified negative controls to delimit a detection threshold above which is considered significant (Set as background level), provided rationale for why the unexpected amplifications did not affect results (Explained away), removed field samples that were associated with negative controls that amplified (Removed samples), or reported amplification but did not provide rationale for why these results could be ignored (No mention).

Even with implementation of appropriate negative controls, *ad hoc* approaches for dealing with negative control amplification, and advanced statistical tools, it is critical to follow strict procedures at each step of the workflow to limit the potential for false positives. Many of these procedures have been described elsewhere, especially in Goldberg et al. (2016). Here we draw attention to several procedures that have reduced false positive sample rates associated with contamination: development of assays that target multiple genomic locations, cleverly designed positive PCR controls, and single-use field sampling gear.

The use of multiple assays that each target different genomic locations provides independent tests of detection. The probability that an amplification in a sample is due to contamination is

the product of the contamination rates of the multiple assays (i.e., the product rule). If the contamination probabilities are low, as is the case we documented in this review, then the probability that an amplification is due to contamination is multiplicatively lower with each additional assay. This strategy increases certainty that observed amplifications are the result of the target organism's DNA in the original sample, as opposed to contaminating DNA or other false-positive signals that act independently on each assay. Calibration studies, such as those by Guillera-Arroita et al. (2017), are needed to quantify false positive sample probabilities. Multiple independent tests of detection, via the statistical product rule should also help to reduce uncertainty caused by cross-contamination and base-rate bias (i.e., when the prevalence of the target is extremely low, the test results in a significant proportion of false positives). Consequently, several eDNA programs that monitor for controversial species or in controversial locations use this type of approach. In the Asian Carp eDNA Monitoring Quality Assurance Project Plan, a sample must be positive for a genus-specific COI assay and for a species-specific ND2 or ND6 assay (Woldt et al., 2019). Similarly, for dreissenid mussels, a sample must be positive for a genus-specific 16S assay and for species-specific COI or Cytb assays (Sepulveda et al., 2019b). Moreover, water samples collected during proceeding field surveys (i.e., resampling verification) must also amplify for the suite of assays in order for the initial samples to be scored as positive.

Metabarcoding eDNA approaches have also begun to use multiple primer sets to minimize false positive taxa assignments. For example, a few studies only retained sequences that are shared by PCR replicates (Giguet-Covex et al., 2014; Alberdi et al., 2018). Metabarcoding studies also commonly remove singletons or doubletons from sequence reads to account for potential low levels of contamination (Evans et al., 2017). While these conservative approaches decrease the potential for false positives, they do increase false-negatives rates and may lead to incorrect inference about target species presence or diversity (Alberdi et al., 2018; Zinger et al., 2019). These approaches also inflate the costs of analyses. The tradeoffs among cost, decreasing false-positive rates and increasing false-negative rates should be carefully considered when designing an eDNA monitoring program.

Even the most cautious laboratories have the potential for sample contamination because high-template positive control material are handled adjacent to analytical samples and negative controls. Standard curves that include positive control DNA at orders of magnitude greater than that found in field samples are a common practice; standard curves are a quality assurance check that the assay is performing as expected and are a means to quantify the amount of target DNA in a sample. Multiple studies have suggested that DNA can aerosolize (e.g., Hebsgaard et al., 2005; Newton et al., 2015; Sepulveda et al., 2019a) and act as a contamination source. Wilson et al. (2016) proposed use of synthetic oligonucleotides with the addition of a readily detectable insert sequence for use as positive PCR controls. A simpler approach may be to scramble the non-priming regions of synthetic DNA. If negative controls amplify, both options permit sequencing of the amplicon to distinguish between real target detections in field samples and positive control-derived contamination. However, sequence inserts/modifications could affect tertiary structures of the DNA molecules (e.g., hairpin loops) and alter the melting temperature of and polymerase binding affinity to the template DNA (Fan et al., 2019). Great care must be taken when designing these synthetic genes to validate them *in silico*. *In vitro* comparisons between native sequences and modified sequences can also be performed to determine any changes in efficacy.

Sample collection and DNA capture methods (e.g., filtration and precipitation) are also important for limiting contamination in eDNA surveys. Collection and DNA capture methods have evolved over time to employ single-use supplies (e.g., gloves and sample collection containers) in order to minimize the potential of cross-contamination. However, there is still high variance among studies in specific methods since each study faces different challenges when attempting to optimize the tradeoffs between contamination risk, sample volume, collection time, and cost. For example, single-use enclosed filters have minimal contamination risk because they are pre-loaded and require no handling of the filter membrane since DNA extraction takes place within the filter capsule (Spens et al., 2017). Relative to open filters that have a higher contamination risk since they do require handling of the filter membrane both pre- and post-sampling, enclosed filters are more expensive, require more time to process, and the volume of water that can be processed may be limited in sites with turbid water since the filters clog easily (Uthicke et al., 2018; Tingley et al., 2019; Tsuji et al., 2019). More recently, Thomas et al. (2019) introduced a self-preserving eDNA filter housing that can process larger volumes of water yet limits handling of the filter membrane to the lab, where it is removed from the housing for DNA extraction.

NEXT STEPS

The potential for contamination-caused uncertainty in eDNA sampling and analysis has eroded confidence in the method because making decisions on incorrect inference can be socio-economically, politically and ecological costly (Jerde, 2019; Sepulveda et al., 2020). We reviewed the eDNA literature over the past twelve years and found contamination did occur, though at a very low rate relative to the hundreds of published studies. Over this period of time, much progress has been made on developing and applying quality control and assurance measures for preventing, alerting to, and source-tracing contamination. More recently, statistical methods have been developed to guide result interpretation in light of false positives. Though these efforts have strengthened the eDNA science and application, there is still ample room for improvement. Specifically, inclusion of critical methodological information is required to quantitatively identify best practices for negative control implementation. There is also an outstanding need for calibration experiments to quantify contamination rates under ideal and realistic field and laboratory conditions. Ultimately, eDNA researchers and end-users must acknowledge that contamination is more commonly observed when using an extremely sensitive molecular tool to search for rare taxa, and that this is an inversely proportional trade-off

between false positive and false negative inferences. Awareness of this tradeoff and due diligence to prevent, identify, and correct for contamination should bolster the use of eDNA results in confident decision-making and management applications.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

AJS and PH: study design and data analyses. AJS, PH, MF, MM, and AMS: data collection and writing. All authors contributed to the article and approved the submitted version.

FUNDING

The USGS Ecosystems Mission Area Invasive Species Program providing funding to support this work.

ACKNOWLEDGMENTS

We thank Cameron Turner for motivating us to conduct this review. We also thank Yale Passamaneck with the US Bureau of Reclamation for helpful comments on earlier versions of this manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.609973/full#supplementary-material

REFERENCES

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., and Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods Ecol. Evolut.* 9, 134–147. doi: 10.1111/2041-210X.12849

Ballenghien, M., Faivre, N., and Galtier, N. (2017). Patterns of cross-contamination in a multispecies population genomic project: detection, quantification, impact, and solutions. *BMC Biol.* 15, 1–16. doi: 10.1186/s12915-017-0366-6

Carim, K. J., Caleb Dysthe, J., McLellan, H., Young, M. K., McKelvey, K. S., and Schwartz, M. K. (2019). Using environmental DNA sampling to monitor the invasion of nonnative *Esox lucius* (northern pike) in the Columbia River basin, USA. *Environ. DNA* 1, 215–226. doi: 10.1002/edn3.22

Darling, J. A., and Mahon, A. R. (2011). From molecules to management: adopting DNA-based methods for monitoring biological invasions in aquatic environments. *Environ. Res.* 111, 978–988. doi: 10.1016/j.envres.2011.02.001

Dickie, I. A., Boyer, S., Buckley, H. L., Duncan, R. P., Gardner, P. P., Hogg, I. D., et al. (2018). Towards robust and repeatable sampling methods in eDNA-based studies. *Mol. Ecol. Resourc.* 18, 940–952. doi: 10.1111/1755-0998.12907

Evans, N. T., Li, Y., Renshaw, M. A., Olds, B. P., Deiner, K., Turner, C. R., et al. (2017). Fish community assessment with eDNA metabarcoding: effects of sampling design and bioinformatic filtering. *Can. J. Fish. Aquat. Sci.* 74, 1362–1374. doi: 10.1139/cjfas-2016-0306

Fan, H., Wang, J., Komiyama, M., and Liang, X. (2019). Effects of secondary structures of DNA templates on the quantification of qPCR. *J. Biomol. Struct. Dyn.* 37, 2867–2874. doi: 10.1080/07391102.2018.1498804

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

Ficetola, G. F., Taberlet, P., and Coissac, E. (2016). How to limit false positives in environmental DNA and metabarcoding? *Mol. Ecol. Resources* 16, 604–607. doi: 10.1111/1755-0998.12508

Giguet-Covex, C., Pansu, J., Arnaud, F., Rey, P.-J., Griggo, C., Gielly, L., et al. (2014). Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nat. Commun.* 5:4211. doi: 10.1038/ncomms4211

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., et al. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods Ecol. Evolut.* 7, 1299–1307. doi: 10.1111/2041-210X.12595

Guillera-Arroita, G., Lahoz-Monfort, J. J., van Rooyen, A. R., Weeks, A. R., and Tingley, R. (2017). Dealing with false-positive and false-negative errors about species occurrence at multiple levels. *Methods Ecol. Evolut.* 8, 1081–1091. doi: 10.1111/2041-210X.12743

Hebsgaard, M. B., Phillips, M. J., and Willerslev, E. (2005). Geologically ancient DNA: fact or artefact? *Trends Microbiol.* 13, 212–220. doi: 10.1016/j.tim.2005.03.010

Jerde, C. L. (2019). Can we manage fisheries with the inherent uncertainty from eDNA? *J. Fish Biol*. doi: 10.1111/jfb.14218

Lahoz-Monfort, J. J., Guillera-Arroita, G., and Tingley, R. (2016). Statistical approaches to account for false-positive errors in environmental DNA samples. *Mol. Ecol. Resources* 16, 673–685. doi: 10.1111/1755-0998.12486

Loeza-Quintana, T., Abbott, C. L., Heath, D. D., Bernatchez, L., and Hanner, R. H. (2020). Pathway to increase standards and competency of eDNA surveys (PISCeS)—advancing collaboration and standardization efforts in the field of eDNA. *Environ. DNA* 2, 255–260. doi: 10.1002/edn3.112

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

Minamoto, T., Miya, M., Sado, T., Seino, S., Doi, H., Kondoh, M., et al. (2020). An illustrated manual for environmental DNA research: water sampling guidelines and experimental protocols. *Environ DNA*. doi: 10.1002/edn3.121

Newton, J., Sepulveda, A., Sylvester, K., and Thum, R. (2015). Potential utility of environmental DNA for early detection of Eurasian watermilfoil (*Myriophyllum spicatum*). *J. Aquat. Plant Manage.* 54, 46–49.

Schnell, I. B., Bohmann, K., and Gilbert, M. T. P. (2015). Tag jumps illuminated–reducing sequence-to-sample misidentifications in metabarcoding studies. *Mol. Ecol. Resour.* 15, 1289–1303. doi: 10.1111/1755-0998.12402

Sepulveda, A. J., Nelson, N. M., Jerde, C. L., and Luikart, G. (2020). Are Environmental DNA methods ready for aquatic invasive species management? *Trends Ecol. Evolut*. 35, 668–678. doi: 10.1016/j.tree.2020.03.011

Sepulveda, A. J., Schabacker, J., Smith, S., Al-Chokhachy, R., Luikart, G., and Amish, S. J. (2019a). Improved detection of rare, endangered and invasive trout in using a new large-volume sampling method for eDNA capture. *Environ. DNA*. 2, 13–23. doi: 10.1002/edn3.23

Sepulveda, A. J., Schmidt, C., Amberg, J., Hutchins, P., Stratton, C., Mebane, C., et al. (2019b). Adding invasive species biosurveillance to the US Geological Survey streamgage network. *Ecosphere* 10:e02843. doi: 10.1002/ecs2.2843

Serrao, N. R., Reid, S. M., and Wilson, C. C. (2018). Establishing detection thresholds for environmental DNA using receiver operator characteristic (ROC) curves. *Conservat. Genet. Resour.* 10, 555–562. doi: 10.1007/s12686-017-0817-y

Smith, M. M., and Goldberg, C. S. (2020). Occupancy in dynamic systems: accounting for multiple scales and false positives using environmental DNA to inform monitoring. *Ecography* 43, 376–386. doi: 10.1111/ecog.04743

Spens, J., Evans, A. R., Halfmaerten, D., Knudsen, S. W., Sengupta, M. E., Mak, S. S., et al. (2017). Comparison of capture and storage methods for aqueous

macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. *Methods Ecol. Evolut.* 8, 635–645. doi: 10.1111/2041-210X.12683

Thomas, A. C., Nguyen, P. L., Howard, J., and Goldberg, C. S. (2019). A self-preserving, partially biodegradable eDNA filter. *Methods Ecol. Evolut.* 10, 1136–1141. doi: 10.1111/2041-210X.13212

Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., et al. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Mol. Ecol.* 21, 2565–2573. doi: 10.1111/j.1365-294X.2011.05418.x

Tingley, R., Greenlees, M., Oertel, S., van Rooyen, A., and Weeks, A. (2019). Environmental DNA sampling as a surveillance tool for cane toad *Rhinella marina* introductions on offshore islands. *Biol. Invasions* 21, 1–6. doi: 10.1007/s10530-018-1810-4

Tsuji, S., Takahara, T., Doi, H., Shibata, N., and Yamanaka, H. (2019). The detection of aquatic macroorganisms using environmental DNA analysis—A review of methods for collection, extraction, and detection. *Environ. DNA* 2, 3–12. doi: 10.1002/edn3.21

Uthicke, S., Lamare, M., and Doyle, J. R. (2018). eDNA detection of corallivorous seastar (*Acanthaster cf. solaris)* outbreaks on the Great Barrier Reef using digital droplet PCR. *Coral Reefs* 37, 1229–1239. doi: 10.1007/s00338-018-1734-6

Wilson, C. C., Wozney, K. M., and Smith, C. M. (2016). Recognizing false positives: synthetic oligonucleotide controls for environmental DNA surveillance. *Methods Ecol. Evolut.* 7, 23–29. doi: 10.1111/2041-210X.12452

Woldt, A., Baerwaldt, K., Monroe, E., Tuttle-Lau, M., Grueneis, N., Holey, M., et al. (2019). *Quality assurance project plan: eDNA monitoring of Bighead and Silver carps*. Bloomington, MN.

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., et al. (2019). DNA metabarcoding—need for robust experimental designs to draw sound ecological conclusions. *Mol. Ecol.* 28, 1857–1862. doi: 10.1111/mec.15060

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 25 January 2021
doi: 10.3389/fenvs.2020.563431

Airborne eDNA Reflects Human Activity and Seasonal Changes on a Landscape Scale

Mark D. Johnson[1], Robert D. Cox[1], Blake A. Grisham[1], Duane Lucia[2] and Matthew A. Barnes[1]*

[1]Department of Natural Resources Management, Texas Tech University, Lubbock, TX, United States, [2]United States Fish and Wildlife Service, Lubbock, TX, United States

Recent research on environmental DNA (eDNA), genetic material shed by organisms into their environment that can be used for sensitive and species-specific detection, has focused on the ability to collect airborne eDNA released by plants and carried by the wind for use in terrestrial plant populations, including detection of invasive and endangered species. Another possible application of airborne eDNA is to detect changes in plant communities in response to activity or changes on a landscape-scale. Therefore, the goal of this study was to demonstrate how honey mesquite, blue grama, and general plant airborne eDNA changes in response to human activity on a landscape-scale. We monitored airborne eDNA before, during, and after a rangeland restoration effort that included honey mesquite removal. As expected, restoration activity resulted in a massive increase in airborne honey mesquite eDNA. However, we also observed changes in abundance of airborne eDNA from the grass genus *Bouteloua*, which was not directly associated with the restoration project, and we attribute these changes to both human activity and seasonal trends. Overall, we demonstrate for the first time that activity and changes on a landscape-scale can be tracked using airborne eDNA collection, and we suggest that airborne eDNA has the potential to help monitor and assess ecological restoration projects, track changes due to global warming, or investigate community changes in response to encroachment by invasive species or extirpation of threatened and endangered species.

Keywords: environmental DNA, airborne, human impact, restoration, qPCR

OPEN ACCESS

Edited by:
Hideyuki Doi,
University of Hyogo, Japan

Reviewed by:
Kevin Leempoel,
Royal Botanic Gardens, Kew, United Kingdom
Lynsey Rebecca Harper,
Liverpool John Moores University, United Kingdom

****Correspondence:***
Mark D. Johnson
mark.johnson@ttu.edu

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Environmental Science

Received: *18 May 2020*
Accepted: *07 December 2020*
Published: *25 January 2021*

Citation:
Johnson MD, Cox RD, Grisham BA, Lucia D and Barnes MA (2021) Airborne eDNA Reflects Human Activity and Seasonal Changes on a Landscape Scale. Front. Environ. Sci. 8:563431. doi: 10.3389/fenvs.2020.563431

INTRODUCTION

Monitoring is a critical component of successful restoration before, during, and after management actions (Walters 1986; Lake 2001; Galatowitsch 2012). For example, conventional field-based plant community monitoring is generally accomplished via methods such as the line-point intercept, belt transect, and gap intercept, using quadrats, transects, or points (Herrick et al., 2005a; Elzinga et al., 1998). Monitoring methods can provide detailed information about a study site, but they can also be time consuming and withdraw logistical and financial resources from low budget projects. In addition, conventional monitoring activity typically incites elevated disturbance to target areas, which may be counterproductive to restoration, and the results can vary based on the intensity of the sampling (Herrick et al., 2005a; Herrick et al., 2005b; Garrard et al., 2008).

A novel method that could address the limitations of conventional monitoring relies upon collection and analysis of environment DNA (eDNA), the genetic material shed by an organism into its environment and collected by researchers from environmental samples such as soil, water, or air (Thomsen et al., 2012; Barnes and Turner 2016). A primary benefit of eDNA monitoring is the ability to detect organisms without the need of a captured target species or direct tissue sample. Indeed, recent reviews by Ruppert et al. (2019) and Makiola et al. (2020) have predicted an increasingly large role for eDNA analysis and metabarcoding in community surveillance and monitoring landscape changes in response to activities such as ecological restoration.

Research into eDNA has focused primarily on aquatic and sediment samples, including assessments on invasive species detection, endangered species monitoring, and the ecology of eDNA (Willerslev et al., 2007; Goldberg et al., 2011; Lodge et al., 2012; Taberlet et al., 2012; Barnes et al., 2014; Barnes and Turner 2016). Recently, Johnson et al. (2019a) demonstrated that airborne eDNA samples, genetic material collected from air, could detect both anemophilous and non-anemophilous plant species. These findings portend the utility of airborne eDNA analysis for broad conservation, management, and research applications such as whole plant community monitoring. However, the ecology of airborne eDNA, including spatial and temporal dynamics of airborne eDNA abundance relative to seasonal changes and other activity on the landscape, remain unexplored but represent critical gaps in current understanding that should be addressed before the potential management and conservation impact of airborne eDNA monitoring can be maximized.

Like eDNA research in general, the existing research addressing how eDNA responds to biotic and abiotic landscape changes primarily comes from aquatic systems. For example, Bista et al. (2017) found that macroinvertebrate eDNA levels displayed community level shifts in diversity throughout the year due to changes in species ecology, biotic, and abiotic factors. Additionally, Buxton et al. (2018) found, when examining eDNA of the great crested newt (*Triturus cristatus*) in aquatic and sediment samples, that detection varied throughout the year based on habitat suitability and species ecology. Work has also been done in ocean systems with studies ranging from examining shark diversity response to anthropogenic disturbance to understanding how anthropogenic activities such as oil spills and development impact micro and macro coastal communities (Bakker et al., 2017; Xie et al., 2018; DiBattista et al., 2020). Sun et al. (2019) used eDNA metabarcoding to understand Dipteria and other organism populations in human caused roadside stormwater ponds. Additionally, Klymus et al. (2017) examined how anthropogenic uranium containment ponds impacted the biodiversity of vertebrate species. Early applications also include monitoring of the re-introduction of Rhine sculpin (*Cottus rhenanus*) into its native range (Hempel et al., 2020) and at-risk, pre-restoration coral community monitoring (Nichols and Marko 2019).

Despite the apparent focus on aquatic systems to date, similar eDNA applications will likely also benefit conservation and management in terrestrial habitats. For example, airborne eDNA analysis could be used to track the changes in plant community composition throughout a restoration project and afterward to provide information for robust adaptive restoration. Airborne eDNA analysis could also assess the impacts of climate change, assist in identifying the spread and location of rangeland invasive species, track endangered species, and detect disease within a restored site or before a restoration (Dejean et al., 2012; Huver et al., 2015; Scriver et al., 2015; Barnes and Turner 2016). With the use of airborne eDNA, tedious surveys that require large amounts of time could be replaced by a network of airborne eDNA traps. However, a critical first step in the evaluation of such methods is to examine the extent to which airborne eDNA reflects landscape changes. Human activity on a landscape-scale encompasses a large variety of activities that could influence airborne eDNA dynamics, from building roads, farming, construction, habitat fragmentation, and invasive species introduction to name a few. We believe that airborne eDNA (both species specific and global) can be used reflect landscape-scale changes from human activity.

In our study, we used ecological restoration as one example of human activity on the landscape. Our goal was to demonstrate the use of airborne eDNA as a surveillance tool during removal of honey mesquite (*Prosopis glandulosa*) on the Texas Tech University native short-grass prairie. Specifically, we examined whether airborne eDNA changed in response to activity on the landscape by: 1) tracking the removal of honey mesquite during a rangeland restoration project; and 2) monitoring changes within the plant community using the *Bouteloua* genus and global plant eDNA as surrogates. The results of this work will help quantify the feasibility of using airborne eDNA to monitor human activities such as restoration and other landscape management.

METHODS

Study Site

The 130-acre Texas Tech University Native Rangeland (33.60327 N, −101.9003 W) acts as a natural area for teaching and research within the Department of Natural Resources Management (**Figure 1**). The site consists of short-grass prairie, with a large variety of bunchgrasses, forbs, and cacti, and a large population of honey mesquite due to the suppression of fire and grazing (Ansley et al., 2001). This site, despite being fragmented and isolated within an urban matrix, represents a native short-grass prairie in post-climax seral stage, and has not been disked, plowed, or reseeded.

Restoration Project

A rangeland restoration project was performed by the Texas Tech Student Chapter of the Wildlife Society, students from the NRM 4309: Range Wildlife Habitat course, and Texas A&M Forest Service in November and December 2017. During the restoration, honey mesquite, a thorny shrub/tree that can re-sprout and form multi-stem thickets (Ansley et al., 1997), was targeted for removal due to its negative impact on forage production, grazing, and local grass biodiversity (Mohamed et al., 2011). The project was

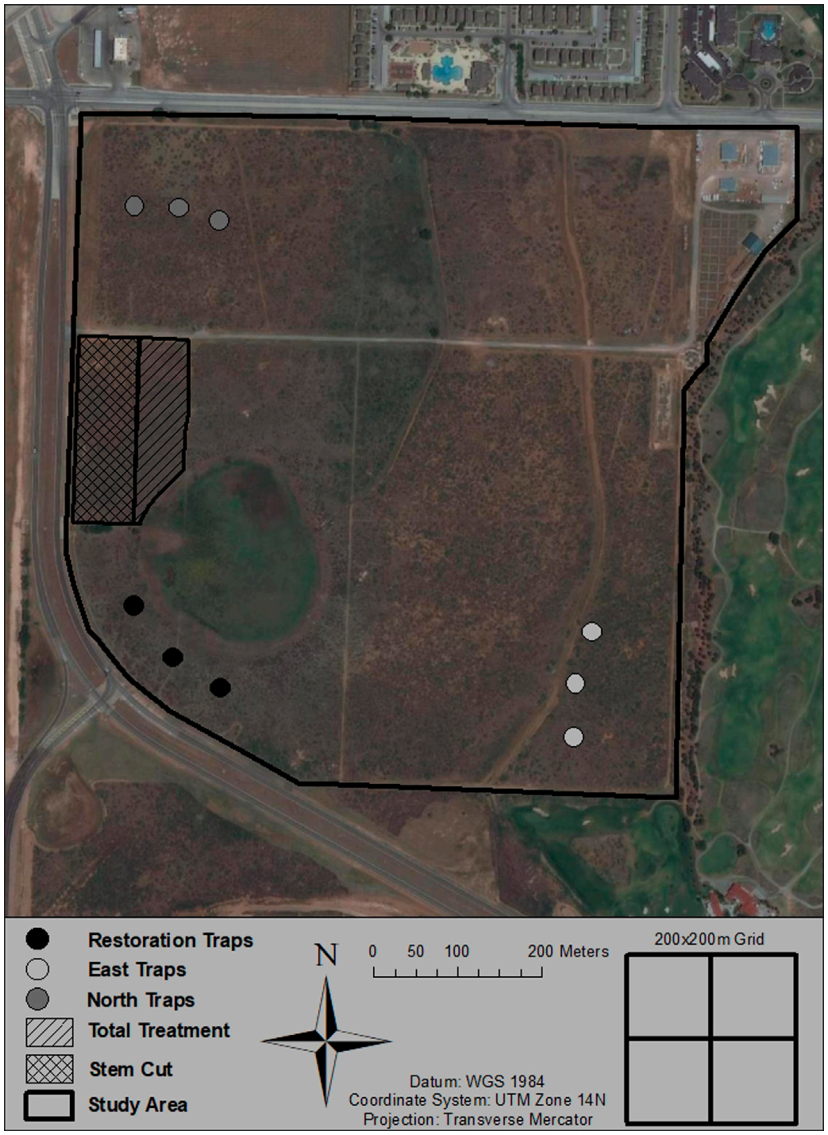


FIGURE 1 | The Texas Tech University Native Rangeland (Lubbock County, TX, United States) study site where the restoration and airborne eDNA sampling took place. The hatched lines show the two different restoration methods used and the clusters of points represent the three groupings of airborne eDNA Big Spring Number Eight Dust Traps.

completed in two treatments, each lasting 3 days. The first, 1.2-hectare treatment began on November 18, 2017. This treatment consisted of cutting off the entirety of the honey mesquite aboveground biomass (i.e., main stem and adjacent minor stems), and then chipping the cut material on site and treating the stump with 25% Triclopyr and 75% diesel with a 1% blue marker dye. This first effort was completed on the eastern side of our study site (**Figure 1**). The second, 1.4-hectare treatment began on December 2, 2017 and was changed to a reduced treatment due to logistical constraints, where only the larger mesquite trees (i.e., main and adjacent minor stems) were targeted and removed. As with the previous treatment, the cut material was chipped, and the stump was treated with the same herbicide mix. This reduced treatment occurred directly west and adjacent to the first total treatment (**Figure 1**). Both treatments attempted to kill the sprout buds of the underground main stem with both cutting and herbicide (Fisher, 1950).

eDNA Collection, Extraction, and Amplification

To examine how airborne eDNA responds to human activity on a landscape-scale, we collected airborne eDNA before, during, and after the restoration project. To collect airborne eDNA, we set up three eDNA sampling locations: one south of the restoration taking place ("restoration traps"), one plot along the northwest edge of the study site ("north traps") and one plot to the east of the restoration treatments ("east traps"; **Figure 1**). Within each

sampling location, three Big Spring Number Eight (BSNE) passive dust traps were deployed. Johnson et al. (2019b) demonstrated that BSNE traps performed well compared to other passive dust traps. The BSNE traps (**Figure 2**) consisted of two triangular traps 0.914 and 0.406 m above the ground. Each trap has an opening at the tip of the triangular piece of metal and a metal mesh vent on top, allowing air to enter the through the opening and flow out through the metal mesh on top, depositing carried material into a collection tray below (Zobeck 2006). Each triangular trap is attached to a metal sheet that acts as sail to consistently orient them into the wind to maximize the amount of material collected.

Traps were sampled four times between November 3 and December 8, 2017. Sampling Events I (November 10th) and II (November 17th) took place 8 and 1 day before the first total treatment restoration, respectively. Event III (November 27th) occurred 9 days after the first total treatment restoration, and Event IV (December 8th) occurred 6 days after the reduced treatment restoration. The collection of repeated samples through time allowed us to examine the response in airborne eDNA to activity associated with the restoration effort. At each sampling event, each trap was rinsed with approximately 1 L deionized water, and the water was collected into individual, sterile 1 L bottles. Since the BSNE arrays each have two collection traps, each tray at a single trap was washed and combined into a single water sample to avoid pseudoreplication (Hurlbert 1984). Rinse water samples were then transported to the laboratory within a cooler and vacuum filtered with 1 µm Isopore membrane filters. Filters were stored at −20 °C for approximately 1 month. Next, DNA extractions were completed using a DNeasy PowerPlant Pro DNA Isolation Kit, which has demonstrated high efficiency in previous airborne eDNA analyses (Johnson et al., 2019b). We followed the manufacturer's protocol, except we added an extra grinding step with a sterile plastic pestle and frequent vortex agitation to ensure homogenization (Johnson et al., 2019b). Extracted genomic DNA was stored at −20 °C until analysis (approximately 6 months later). To confirm that there was no contamination throughout this process, extraction blanks (i.e., sterile samples extracted alongside experimental samples as negative controls with just buffer and no filter) were processed with every extraction event. Additionally, we used sterile containers, bleached all laboratory surfaces, and used sterile gloves. Due to the nature of airborne eDNA, we could not develop confident field or filtration controls (i.e., we have not developed a method in which control samples are not exposed to the air representing our sample). As a result, we have only included extraction and PCR negative controls ("blanks").

To broadly characterize changes in airborne eDNA in response to the restoration activity, we used three different quantitative polymerase chain reaction (qPCR) assays: a honey mesquite species-specific primer, a *Bouteloua* genus assay, and a global plant primer. Species- and genus-specificity for honey mesquite and *Bouteloua* assays, respectively, were confirmed in silico using NCBI Primer-BLAST (Ye et al., 2012) as well as in-lab PCR confirmation using tissues of the nine most common plants found within our study site. We quantified both the limits of detection and quantification as the lowest amount of DNA our primers detect any of the technical replicates and where all the technical replicates were detected, respectively. Operationally, we followed the recommendations of low-copy qPCR analysis of Ellison et al. (2006) and assigned all non-detections a value of zero rather than omitting them from the analysis or missing out on information provided by samples in which fewer than all technical replicates amplified. We previously observed high rates of PCR inhibition in airborne eDNA samples (Johnson et al., 2019a; Johnson et al., 2019b), which led us to dilute samples in this study by 1:10 with pure water. All qPCR reactions were completed on a QuantStudio 3 Real-Time qPCR machine (ThermoFisher Scientific, Foster City, California). The honey mesquite assay targeted the focal species of the restoration effort using 25 µl reactions with 1x PowerSYBR Green qPCR Master Mix, 1 µM forward and reverse primer (Johnson et al., 2019a; **Table 1**) 2 µl diluted genomic DNA template. The thermocycling program for the honey mesquite assay began with an initial 95 °C step for 10 min followed by 40 cycles of 15 s at 95 °C and 1 min at 70.1 °C, and a final melt curve analysis. The *Bouteloua* genus assay targeted grasses that represents the most common wind pollinated species on our study site. Each 25 µl qPCR reaction contained 1x PowerSYBR Green qPCR Master Mix, 1 µM forward and reverse primer (Johnson et al., 2019a; **Table 1**), and 2 µl diluted genomic DNA template. The

FIGURE 2 | The Big Spring Number Eight dust traps that were used to collect airborne eDNA in this experiment.

TABLE 1 | The three different primer sets (forward and reverse) used over the course of this experiment. The honey mesquite and *Bouteloua* genus assays were produced by Johnson et al., 2019a. The *trnL* global chloroplast was taken from Taberlet et al. (2007).

Target	Forward (5–3′)	Reverse (5–3′)	Size (bp)
rbcL honey mesquite	CTGAAGAAGCAGGTGCTGCG	TTGAGTTTCTTCTCCAGGAACAGG	140
rbcL Bouteloua genus	ACCCGTTCCTGGAGAAGATAGT	CAGGAGGAATTCGTAGATCCTCCA	164
trnL global chloroplast	CGAAATCGGTAGACGCTACG	CCATTGAGTCTCTGCACCTATC	~200

thermocycling program used an initial 95 °C step for 10 min followed by 40 cycles of 15 s at 95 °C and 1 min at 66 °C, and a final melt curve analysis. Reactions in the honey mesquite and *Bouteloua* genus assays were quantified using five-point standard curves based on a 1:10 serial dilution of tissue-derived DNA from honey mesquite and *Bouteloua gracilis*, respectively. Non-detections were treated as zeros when averaged with other technical replicates (Ellison et al., 2006). For both assays, samples were run in triplicate and non-template controls were included to ensure no contamination occurred.

Finally, as an indicator of changes in the overall plant community beyond the two focal groups, honey mesquite and *Bouteloua* spp., we amplified all plant eDNA in our samples with a global plant assay targeting the chloroplast *trnL* gene (Taberlet et al., 2007; **Table 1**). For this assay, each 25-µl qPCR reaction contained 1x *Power*SYBR Green qPCR Master Mix, 1 µM forward and reverse primer concentrations, and 2 µl diluted genomic DNA template. The thermocycling program began with an initial 95 °C step for 10 min followed by 32 cycles of 2 min at 94 °C, 1 min at 55 °C, and 30 s at 72 °C, and a final extension at 72 °C for 2 min (Craine et al., 2016). Since amplification with the global plant assay could be the result of a variable mix of plant eDNA sources, we could not create a standard curve for quantification. Therefore, rather than absolute quantification of eDNA concentration in each reaction, we relied on comparison of the average number of cycles needed for the samples to display enough fluorescence to be considered positive (cycle threshold, determined using the default settings of the QuantStudio three Real-Time qPCR machine and abbreviated C_T; Heid et al., 1996).

To analyze whether airborne eDNA changed in response to restoration activity, we completed three separate repeated measures analyses of variance (rmANOVAs) with IBM SPSS statistics (IBM Corp. 2017), separately analyzing honey mesquite, *Bouteloua* genus, and global plant eDNA results. In each analysis, sampling units consisted of nine total BSNE traps, and three replicate traps at each plot were combined into experimental units to make comparisons between locations. Sampling event represented the repeated measure in our experimental design, we interpreted Wilks' Lambda as our test statistic, and we assumed $\alpha = 0.05$ for determination of statistical significance. Following significant rmANOVAs, we used Tukey-Kramer post-hoc tests to examine how the different trap locations varied from one another for each sampling event.

When collecting airborne eDNA, it is common to see very small leaf fragments in the samples; however, occasionally large plant leaves can be collected which in turn results in extremely high amounts of target DNA in the traps. While these detections are real, the extremely large amounts of airborne eDNA may obfuscate underlying trends or patterns. Therefore, we removed two outliers from consideration: one from the global eDNA assay during Event III in the east traps (>100x more eDNA than the other two traps at the same site), and the other from the honey mesquite assay from Event III, also in the east traps (600x more eDNA than replicate traps).

RESULTS

Honey Mesquite eDNA

We found during Events I and II that the average honey mesquite eDNA quantities for all three trap locations were consistently low (**Table 2**). After the total restoration treatment, the average quantity of honey mesquite eDNA spiked in the restoration traps but remained low in the east and undetected in north traps (**Table 2**). After Event IV, we observed honey mesquite eDNA in high concentrations in the restoration traps with low concentrations for the east traps and no detections for the north traps (**Table 2**). Overall, when comparing the amounts of airborne eDNA for each of our three sampling locations, eDNA significantly differed between trap locations (rmANOVA $F_{3,19} = 7.36$, $p = 0.0018$; **Figure 3**). Pairwise comparisons revealed that for Events I and II, the east traps were significantly different from both the restoration ($p = 0.0052$ and $p < 0.0001$, respectively) and north trap locations ($p = 0.0052$ and $p < 0.0001$), but the north and restoration traps were not significantly different from each other for either Event (**Table 3**). However, after the first restoration treatment (Event III), the restoration traps were significantly different from both the north traps ($p < 0.0001$) and the east traps ($p < 0.0001$), and the north and east traps were not significantly different ($p = 0.98$; **Table 3**). Lastly, during Event IV, the restoration traps were again significantly different from both the north ($p < 0.0001$) and east ($p < 0.0001$) trap locations while the north and east traps did not significantly differ from one another ($p = 0.996$; **Table 3**). Amplification percentage (i.e., number of samples displaying any amplification, regardless of eDNA quantification) for Events III and IV showed a large spike for the restoration traps, rising from 33% to 100% for both Events III and IV (**Figure 4A**). This spike was mirrored slightly in the east traps but not in the north traps.

Across all honey mesquite qPCR plates, negative controls failed to amplify as expected, and qPCR efficiencies were on average 67% with an average R-squared of 0.99. The limit of detection and limit of quantification for the honey mesquite primer were both 3.8×10^{-4} ng/µL.

TABLE 2 | The amounts of honey mesquite, *Bouteloua* genus, and global plant airborne eDNA collected for each trap location and all four sampling events. The honey mesquite and *Bouteloua* genus units are mean concentration ± standard deviation. The global plant eDNA shows the mean cycle threshold ± standard deviation. The complete raw results can be found in the **Supplementary Table**.

		Amount of eDNA captured			
	Trap location	**Event I**	**Event II**	**Event III**	**Event IV**
Honey mesquite eDNA (pg/µl)	Restoration	0	0.08 ± 0.24	4.5 ± 4.5	2.8 ± 1.5
	East	0.70 ± 1.3	0.37 ± 0.46	0.16 ± 0.30	0.08 ± 0.11
	North	0	0.11 ± 0.34	0	0
Bouteloua genus eDNA (pg/µl)	Restoration	0.08 ± 0.08	0.04 ± 0.07	0.04 ± 0.03	0.01 ± 0.02
	East	0.64 ± 0.67	0.15 ± 0.21	0.21 ± 0.13	0
	North	0.15 ± 0.06	0.02 ± 0.04	0.08 ± 0.08	0
Global plant eDNA (cycles)	Restoration	30.1 ± 0.94	29.5 ± 0.79	26.9 ± 0.48	29.5 ± 0.31
	East	29.1 ± 1.0	30.1 ± 0.55	29.4 ± 0.58	30.9 ± 0.40
	North	29.2 ± 0.86	30.9 ± 0.25	28.9 ± 0.49	31.1 ± 0.53

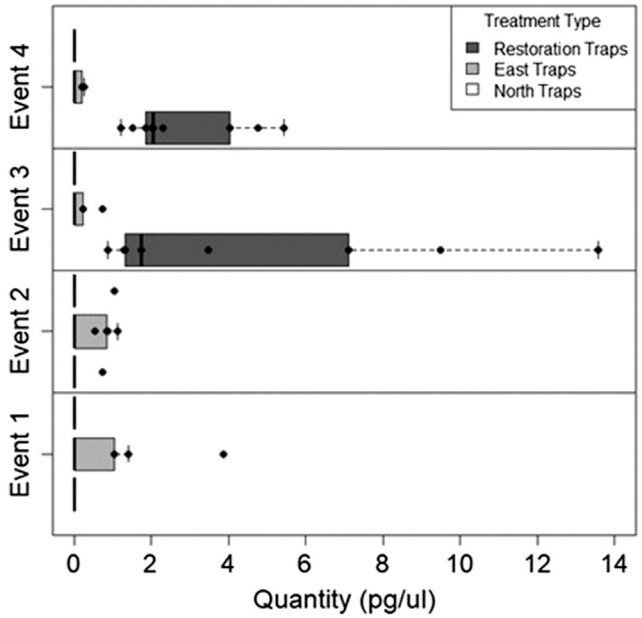


FIGURE 3 | The quantity of honey mesquite eDNA for all three trap locations over four sampling Events. Honey mesquite eDNA increased following restoration activity, especially within the restoration site.

Bouteloua Genus eDNA

The quantity of *Bouteloua* airborne eDNA changed over time as well. At Event I the quantity of *Bouteloua* varied greatly between trap locations, with the east traps having a larger quantity than the restoration and north traps. Between Events I and II the amount of *Bouteloua* eDNA declined (**Table 2**). After the first full treatment restoration event, increases in the quantity of *Bouteloua* DNA were detected in the east and northern traps while the restoration traps stayed the same. After the reduced treatment, there were no *Bouteloua* detections other than by the restoration traps (**Table 2**). Across all four Events, we observed significant differences in the amount of *Bouteloua* eDNA between each trap location (rmANOVA $F_{3,22} = 329.85$, $p < 0.0001$; **Figure 5**). At Event I the east traps were significantly different from both the restoration ($p = 0.0069$) and north ($p = 0.0198$) traps, whereas the restoration and north traps did not significantly differ ($p = 0.8934$; **Table 3**). At Event II, none of the trap locations significantly differed from one another (**Table 3**). At Event III, the east traps were significantly different from both restoration ($p < 0.0001$) and north ($p < 0.0001$) traps, while the north and restoration traps did not differ from one another (**Table 3**). Lastly, at Event IV, the restoration traps were significantly different from both the east ($p = 0.0319$) and north ($p = 0.0319$) traps, while the east and north traps were not significantly different (**Table 3**). *Bouteloua* amplification percentages were consistently 66% or higher in Events I and II, and Event III notably had 100% amplification for all groups (**Figure 4B**). At Event IV, we only observed *Bouteloua* amplification in restoration traps (**Figure 4B**).

Across all *Bouteloua* qPCR plates, negative controls failed to amplify as expected, and qPCR efficiencies were an average of 83% with an average R-squared of 0.99. The *Bouteloua* genus primer's limit of detection and limit of quantification were both 6.7×10^{-6} ng/µl.

Global Plant eDNA

For the first two sampling events, the average cycle threshold (C_T; note that decreasing C_T values indicate increasing eDNA concentrations) for all three trap locations were consistent and averaged between approximately 29.1 and 30.1 cycles (**Table 2**). After the total treatment restoration, all three traps detected more global plant eDNA with a large spike in the amount of restoration trap airborne eDNA. Lastly, the average C_T values for the samples taken after the reduced treatments were lower compared to those in Event III (**Table 2**). We found a significant effect of sampling Event on the amount of global eDNA (rmANOVA $F_{3,19} = 8{,}158.88$, $p < 0.0001$; **Figure 6**). Pairwise analyses revealed that at Event I, none of the sampling events were significantly different from each other (**Table 3**). At Event II, the restoration traps were significantly different from both the east ($p = 0.0142$) and north ($p < 0.0001$), while the north and east traps also differed significantly ($p = 0.0251$; **Table 3**). At Event III, following the first restoration, the restoration traps were significantly different from both the north ($p < 0.0001$) and east ($p < 0.0001$) traps, while the north and east traps did not differ from one another (**Table 3**). Finally, at Event IV, the

TABLE 3 | The *p* values for the Tukey-Kramer post-hoc comparison tests for each event examining the honey mesquite, *Bouteloua* genus and global plant eDNA.

		Event 1		Event 2		Event 3		Event 4	
		Restoration	East	Restoration	East	Restoration	East	Restoration	East
Honey mesquite	East	0.0052		<0.0001		0.0001		<0.0001	
	North	1	0.0052	0.8656	<0.001	<0.001	0.98	<0.0001	0.996
Bouteloua genus	East	0.0069		0.1872		<0.0001		0.0319	
	North	0.8934	0.0198	0.8881	0.0786	0.1059	<0.0001	0.0319	1
Global plant eDNA	East	0.0595		0.0142		<0.0001		<0.0001	
	North	0.1412	0.7879	<0.0001	0.0251	<0.0001	0.1111	<0.0001	0.7993

restoration traps were again significantly different from both the north ($p < 0.0001$) and east ($p < 0.0001$) traps, and the north and east traps remained non-significantly different (**Table 3**). All three trap locations amplified global plant airborne eDNA 100% of the time across all four events (**Figure 4C**). Across all global plant qPCR plates, negative controls failed to amplify, as expected.

DISCUSSION

Using three different qPCR assays targeting honey mesquite, *Bouteloua* genus, and a global plant chloroplast gene, we demonstrated that airborne eDNA is affected by human activity during a restoration event, and we argue that these changes track intuitively through time with different stages of the restoration. Notably, we found that airborne eDNA from species that were not even the target of the restoration changed over time. Collectively, our observations demonstrate that airborne eDNA reflects human activity and phenological changes on a landscape-scale and point to an expanding role that airborne eDNA surveillance may play in terrestrial conservation.

Since the restoration focused on the removal of honey mesquite, we first quantified how the amount of honey mesquite changed over the course of the restoration. We found that there was a significant difference between the restoration eDNA traps and the east and north traps. Before the total treatment restoration took place, mesquite eDNA quantity and detection was low. It is useful to consider the ecology of honey mesquite to put these results into context. Honey mesquite is insect-pollinated and flowers in the spring before losing its leaves and going dormant for the winter months (Lopez-Portillo et al., 1993; Golubov et al., 1999). At the time of our restoration effort, honey mesquite was inactive with most of their leaves gone and no flowering or pollination occurring. Accordingly, a "typical" paucity of airborne honey mesquite eDNA was illustrated by the low concentrations and detection percentages for this species in our first two sampling Events (**Figure 3** and **Figure 4A**). However, after the first restoration event, Event III demonstrated a large increase in the concentration of honey mesquite airborne eDNA for the restoration traps. A spike in airborne eDNA occurred again at Event IV but was not as large, likely because the second restoration was farther away and lower intensity. Together, these results demonstrate the potential for airborne eDNA analysis to distinguish between different types or intensities of activity on a landscape-scale. Additionally, airborne eDNA analysis may reveal spatial information. The amplification percentage of honey mesquite detected for each restoration group trap jumped to 100% for both Events III and IV (**Figure 4A**). We simultaneously observed amplification percentage in east traps (i.e., away from the site of the restoration activity), though not in as high concentrations as shown in the restoration traps, increase after the restorations, suggesting that honey mesquite eDNA traveled downwind to the east trap grouping.

On the other hand, the concentration of *Bouteloua* airborne eDNA, which was not the target of restoration activities, also changed over time. In general, *Bouteloua* eDNA decreased throughout our study. Of the four *Bouteloua* species on our study site - blue grama (*Bouteloua gracilis*), buffalo grass (*Bouteloua dactyloides*), sideoats grama (*Bouteloua curtipendula*), and sixweeks grama (*Bouteloua barbata*) - blue grama is by far the most common species. The steady decreasing trend that was observed corresponds to the ecology of blue grama, which is summer-active, releases pollen in early fall, and then goes dormant for the winter months (Riegal 1941; Anderson 2003). Johnson et al. (2019a) monitored the changes in *Bouteloua* genus eDNA for the early fall and showed there was a spike in *Bouteloua* airborne eDNA associated with their early fall pollination event. However, the results in the present study track the opposite direction and appear to document that *Bouteloua* airborne eDNA concentrations decline in alignment with approaching winter dormancy.

Therefore, if not impacted, *Bouteloua* airborne eDNA would likely have decreased uniformly throughout the study period. This trend was observed for the first two sampling events, especially in the eastern traps where large amounts of blue grama grow. However, after the first total treatment restoration we observed a modest increase in *Bouteloua* airborne eDNA concentrations during Event III. To give further evidence that *Bouteloua* airborne eDNA was impacted, we can examine the concentrations in conjunction with the percentage of traps that significantly detected *Bouteloua* DNA (**Figure 4B** and **Figure 5**). In addition to the increase in *Bouteloua* eDNA concentration after the total treatment, we saw each trap location detect *Bouteloua* eDNA 100% of the time, which should be unlikely to occur naturally since blue grama is becoming dormant for the year at the time of this study. For Event IV, the more limited restoration activity did not promote *Bouteloua* detection in either north or east trap locations, but we observed 67% amplification at the restoration trap location. Again, this

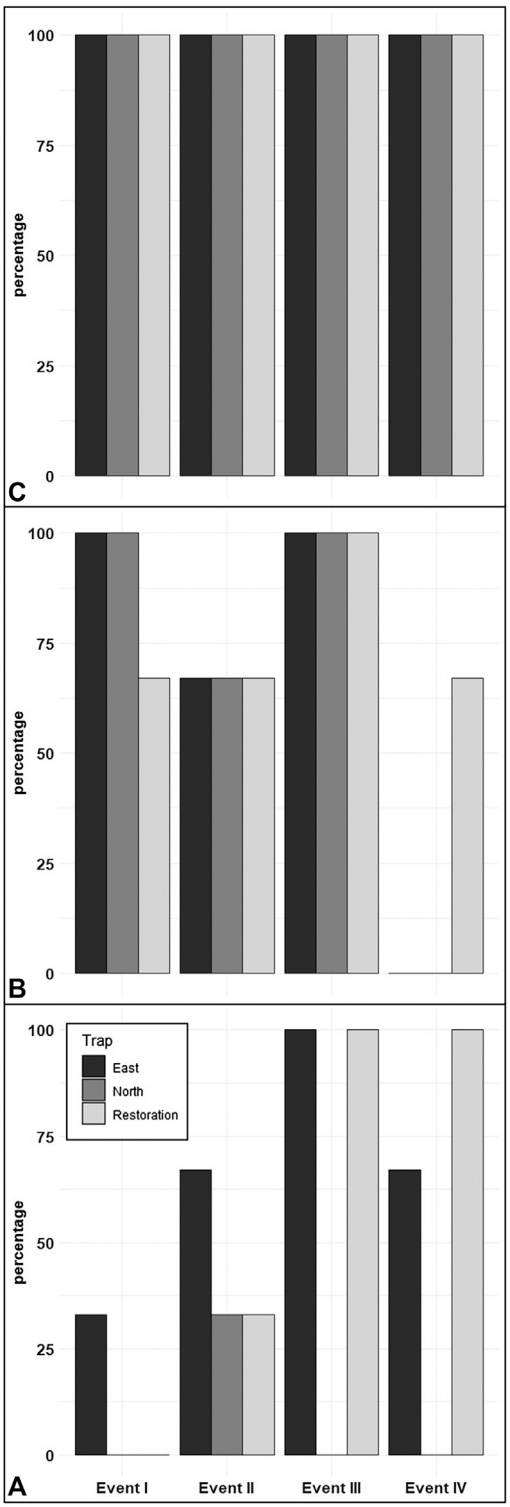


FIGURE 4 | The percentage of significant amplification/detections for the traps from all three sampling locations (East, North, and Restoration). The primers focused on honey mesquite airborne eDNA **(A)**, *Bouteloua* airborne eDNA **(B)**, and global plant airborne eDNA **(C)**.

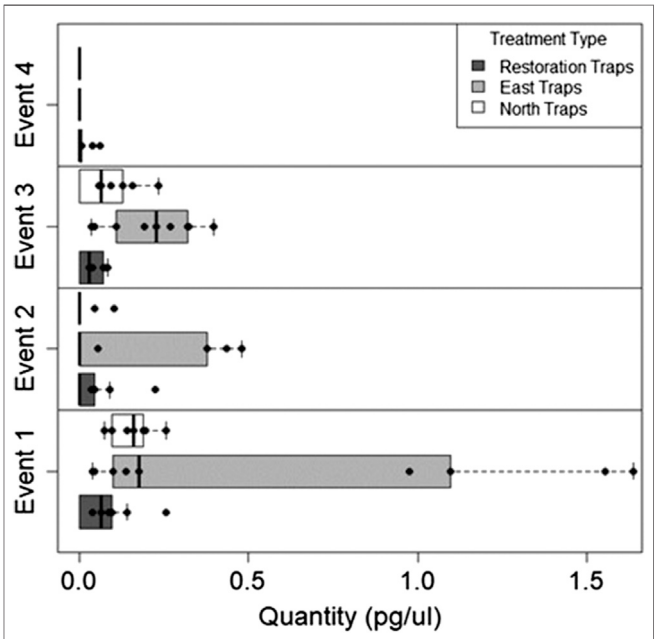


FIGURE 5 | The quantity of *Bouteloua* genus eDNA collected for all three trap locations over four sampling Events. *Bouteloua* genus eDNA concentrations generally decreased over time with notable spikes after the restoration treatments.

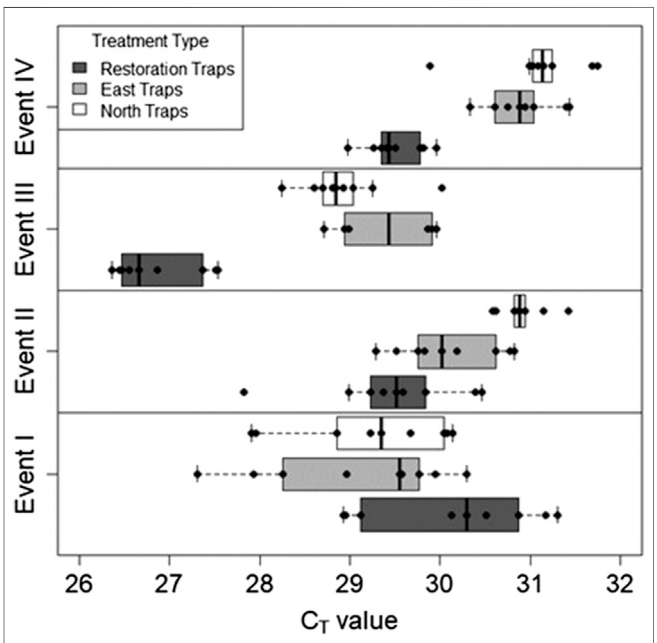


FIGURE 6 | The global plant eDNA collected for all three trap locations for all four sampling Events. The global plant eDNA is measured as cycle thresholds (C_T) where decreasing values indicate more eDNA. Global plant eDNA increased in response to restoration activity.

pattern points to a spatial interpretation of airborne eDNA results, with restoration traps closest to the restoration activity demonstrating impacts from the reduced treatment and farther east and north traps remaining unaffected.

We observed that *Bouteloua* eDNA appeared to impact all three trap locations whereas honey mesquite eDNA, the subject of the restoration, impacted only the restoration traps and moderately the eastern traps. This could be a result of the eDNA content being released into the air. The *Bouteloua* DNA is assumed to come primarily from blue grama which is at the tail end of its pollination season. Any disruption would result in a plume of leftover pollen being released. *Bouteloua* pollen is designed to travel on the wind so it would be logical for it to travel greater distances and effect the eastern and northern trap groupings. The honey mesquite on the other hand, is insect pollinated and had sparse leaves available when the restoration took place. As a result, the DNA being released into the air was most likely living wood fragments, cells, and a small amount of leaf material. This type of DNA is not designed to travel on the wind, so only seeing a spike at the closest trap site and slight increases in detection percentage elsewhere would make sense.

Finally, we observed a significant change in the global plant airborne eDNA in our traps closest to the restoration treatments over time. In this study, the global plant airborne eDNA acts as a qualitative surrogate for the plant community. Species specific primers are still rare for most plant species so without a metabarcoding approach, global plant eDNA allows us to monitor general patterns across all plants in addition to focusing on single species (Wallinger et al., 2012). For the first two Events, the global assay showed consistent levels of airborne eDNA across all three sites. After the total treatment restoration occurred, there was a large spike in restoration trap site eDNA and smaller increases in both the north and east traps. After the limited restoration activity, the restoration trap site showed a higher amount of airborne eDNA compared to the other two trap locations, which were unaffected. As shown in the *Bouteloua* analysis, the amount of *Bouteloua* eDNA (the most common genus on the landscape) was steadily dropping over the study period which is reflected by the east and north traps showing less global eDNA amounts. This again points to spatial information contained within airborne eDNA analyses as well as the fact that airborne eDNA analysis can distinguish between different activity intensities.

Overall, we have shown that airborne eDNA can assist with the evaluation of current species on a landscape-scale, and that airborne eDNA reflects human activity and seasonal changes on a landscape-scale. In a conservation or management setting, we believe airborne eDNA analysis can aid site selection and monitoring, which will prove especially valuable if it can supplement or even provide an alternative to time-consuming and potentially disruptive conventional plant community surveys. To maximize the conservation potential of airborne eDNA analysis, further examination of the ability of airborne eDNA analysis to detect rare species on a landscape-scale is warranted. Furthermore, combination of airborne eDNA analysis and metabarcoding approaches could allow airborne eDNA to act as a plant community monitoring tool, further increasing its utility for conservation, management, and research.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

MJ designed the research, collected and analyzed the data, and wrote the manuscript. RC collected the data and commented on the manuscript. BG performed the restoration, analyzed the data, and commented on the manuscript. DL performed the restoration and commented on the manuscript. MB designed the research, collected and analyzed the data, commented on the manuscript, and provided funding.

FUNDING

This work was funded by Texas Tech University as startup funding to MB.

ACKNOWLEDGMENTS

Peter Dotray, Karin and Ardon-Dryer provided guidance as members of MJ's doctoral dissertation committee. Robert Scott Van Pelt provided the BSNE dust collector traps from the United States Department of Agriculture. Elizabeth Roesler, Sadie Roth, Hayden Hays, Kristin Kabat, and two anonymous reviewers provided comments on an earlier draft of this manuscript. Use of product names within this manuscript does not constitute endorsement.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenvs.2020.563431/full#supplementary-material.

REFERENCES

Anderson, M. (2003). Bouteloua gracilis. *Fire effects information system*. Washington, DC: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Science Laboratory.

Ansley, R. J., Huddle, J. A., and Kramp, B. A. (1997). *Mesquite ecology*. Vernon, Tx: Texas A&M Agrilife Extension, Texas Agricultural Experiment Station.

Ansley, R. J., Wu, B., and Kramp, B. A. (2001). Observation: long-term increases in mesquite canopy cover in a North Texas savanna. *J. Range Manag.* 54, 171–176. doi:10.2307/4003179

Bakker, J., Wangensteen, O. S., Chapman, D. D., Boussarie, G., Buddo, D., Guttridge, T. L., et al. (2017). Environmental DNA reveals tropical shark diversity in contrasting levels of anthropogenic impact. *Sci. Rep.* 7, 16886. doi:10.1038/s41598-017-17150-2

Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., and Lodge, D. M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environ. Sci. Technol.* 48, 1819–1827. doi:10.1021/es404734p

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

Bista, I., Carvalho, G. R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., et al. (2017). Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. *Nat. Commun.* 8, 14087. doi:10.1038/ncomms14087

Block, W. M., Franklin, A. B., Ward, J. P., Ganey, J. L., and White, G. C. (2001). Design and implementation of monitoring studies to evaluate the success of ecological restoration on wildlife. *Restor. Ecol.* 9, 293–303. doi:10.1046/j.1526-100x.2001.009003293.x

Buxton, A., Groombridge, J. J., and Griffiths, R. A. (2018). Seasonal variation in environmental DNA detection in sediment and water samples. *PloS One* 13, e0191737. doi:10.1371/journal.pone.0191737

Clewell, A. F., and Aronson, J. (2013). *Ecological restoration. The science and practice of ecological restoration*. Washington, DC: Island Press.

Craine, J. M., Barberan, A., Lynch, R. C., Menninger, H. L., Dunn, R. R., and Fierer, N. (2016). Molecular analysis of environmental plant DNA in house dust across the United States. *Aerobiologia* 33, 71–86. doi:10.1007/s10453-016-9451-5

Dejean, T., Valentini, A., Miquel, C., Taberlet, P., Bellemain, E., and Miaud, C. (2012). Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog Lithobates catesbeianus. *J. Appl. Ecol.* 49, 953–959. doi:10.1111/j.1365-2664.2012.02171.x

DiBattista, J. D., Reimer, J. D., Stat, M., Masucci, G. D., Biondi, P., De Brauwer, M., et al. (2020). Environmental DNA can act as a biodiversity barometer of anthropogenic pressures in coastal ecosystems. *Sci. Rep.* 10, 8365. doi:10.1038/s41598-020-64858-9

Ellison, S. L., English, C. A., Burns, M., and Keer, J. T. (2006). Routes to improving the reliability of low level DNA analysis using real-time PCR. *BMC Biotechnol.* 6, 33. doi:10.1186/1472-6750-6-33

Elzinga, C. L., Salzer, D. W., and Willoughby, J. W. (1998). "Measuring and monitoring plant populations," in USDI-BLM Technical Reference 1730-1. (Denver, CO: USDI-BLM).

Fisher, C. E. (1950). The mesquite problem in the Southwest. *J. Range Manag.* 3, 60–70.

Galatowitsch, S. M. (2012). *Ecological restoration. Sunderland, mass.* Sunderland, MA: Sinauer Associates.

Garrard, G. E., Bekessy, S. A., McCarthy, M. A., and Wintle, B. A. (2008). When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. *Austral. Ecol.* 33, 986–998. doi:10.1111/j.1442-9993.2008.01869.x

Goldberg, C. S., Pilliod, D. S., Arkle, R. S., and Waits, L. P. (2011). Molecular detection of vertebrates in stream water: a demonstration using rocky mountain tailed frogs and Idaho giant salamanders. *PLoS One* 6, e22746. doi:10.1371/journal.pone.0022746

Golubov, J., Mandujano, M. D. C., Franco, M., Montana, C., Eguiarte, L. E., and Lopez-Portillo, J. (1999). Demography of the invasive woody perennial *Prosopis glandulosa* (honey mesquite). *J. Ecol.* 87, 955–962. doi:10.1046/j.1365-2745.1999.00420.x

Goudie, A. S. (1986). *The human impact on the natural environment.* Cambridge, MA: MIT Press.

Heid, C. A., Stevens, J., Livak, K. J., and Williams, P. M. (1996). Real time quantitative PCR. *Genome Res.* 6, 986–994. doi:10.1101/gr.6.10.986

Hempel, C. A., Peinert, B., Beermann, A. J., Elbrecht, V., Macher, J.-N., Macher, T.-H., et al. (2020). Using environmental DNA to monitor the reintroduction success of the Rhine sculpin (*Cottus rhenanus*) in a restored stream. *Front. Ecol. Evol.* 8, 81. doi:10.3389/fevo.2020.0008

Herrick, J. E., Van Zee, J. W., Havstad, K. M., Burkett, L. M., and Whitford, W. G. (2005b). *Monitoring manual for grassland, shrubland and savanna ecosystems volume II.* Tucson, AZ: USDA-ARS Jornada Experimental Range University of Arizona Press.

Herrick, J. E., Van Zee, J. W., McCord, S. E., Courtright, E. M., Karl, J. W., and Burkett, L. M. (2005a). *Monitoring manual for grassland, shrubland and savanna ecosystems volume I.* Tucson, AZ: USDA-ARS Jornada Experimental Range University of Arizona Press.

Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecol. Monogr.* 54, 187–211. doi:10.2307/1942661

Huver, J. R., Koprivnikar, J., Johnson, P. T., and Whyard, S. (2015). Development and application of an eDNA method to detect and quantify a pathogenic parasite in aquatic ecosystems. *Ecol. Appl.* 25, 991–1002. doi:10.1890/14-1530.1

IBM Corp. (2017). *IBM SPSS statistics for windows, Version 25.0.* Armonk, NY: IBM Corp.

Johnson, M. D., Cox, R. D., and Barnes, M. A. (2019a). The detection of a non-anemophilous plant species using airborne eDNA. *PLoS One* 14, e0225262. doi:10.1371/journal.pone.0225262

Johnson, M. D., Cox, R. C., and Barnes, M. A. (2019b). Analyzing airborne environmental DNA: a comparison of extraction methods, primer type, and trap type on the ability to detect airborne eDNA from terrestrial plant communities. *Environmental DNA* 1, 176–185. doi:10.1002/edn3.19

Klymus, K. E., Richter, C. A., Thompson, N., and Hinck, J. E. (2017). Metabarcoding of environmental DNA samples to explore the use of uranium mine containment ponds as a water source for wildlife. *Diversity* 9, 54. doi:10.3390/d9040054

Lake, P. S. (2001). On the maturing of restoration: linking ecological research and restoration. *Ecol. Manag. Restor.* 2, 110–115. doi:10.1046/j.1442-8903.2001.00074.x

Lodge, D. M., Turner, C. R., Jerde, C. L., Barnes, M. A., Chadderton, L., Egan, S. P., et al. (2012). Conservation in a cup of water: estimating biodiversity and population abundance from environmental DNA. *Mol. Ecol.* 21, 2555–2558. doi:10.1111/j.1365-294X.2012.05600.x

Lodge, D. M., Simonin, P. W., Burgiel, S. W., Keller, R. P., Bossenbroek, J. M., Jerde, C. L., et al. (2016). Risk analysis and bioeconomics of invasive species to inform policy and management. *Annu. Rev. Environ. Resour.* 41, 453–488. doi:10.1146/annurev-environ-110615-085532

Lopez-Portillo, J., Eguiarte, L., and Montana, C. (1993). Nectarless honey mesquites. *Funct. Ecol.* 7, 452–461. doi:10.2307/2390032

Makiola, A., Compson, Z. G., Baird, D. J., Barnes, M. A., Boerlijst, S. P., Bouchez, A., et al. (2020). Key questions for next-generation biomonitoring. *Frontiers in Environmental Science* 7, 197. doi:10.3389/fenvs.2019.00197

Mohamed, A. H., Holecheck, J. L., Bailey, D. W., Campbell, C. L., and DeMers, M. N. (2011). Mesquite encroachment impact on southern New Mexico rangelands: remote sensing and geographic information systems approach. *J. Appl. Remote Sens.* 5. 3514. doi:10.1117/1.3571040

Nichols, P. K., and Marko, P. B. (2019). Rapid assessment of coral cover from environmental DNA in Hawai'i. *Environmental DNA* 1, 40–53. doi:10.1002/edn3.8

Pimental, D., Zuniga, R., and Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecol. Econ.* 52, 273–288. doi:10.1016/j.ecolecon.2004.10.002

Riegal, A. (1941). Life history and habits of blue grama. *Trans. Kans. Acad. Sci.* 44, 76–85.

Ruppert, K. M., Kline, R. J., and Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: a systematic review in methods, monitoring, and applications of global eDNA. *Glob. Ecol. Conserv.* 17, e00547. doi:10.1016/j.gecco.2019.e00547

Scriver, M., Marinich, A., Wilson, C., and Freeland, J. (2015). Development of species-specific environmental DNA (eDNA) markers for invasive aquatic plants. *Aquat. Biol.* 122, 27–31. doi:10.1016/j.aquabot.2015.01.003

Sun, Z., Majaneva, M., Sokolova, E., Rauch, S., Meland, S., and Ekrem, T. (2019). DNA metabarcoding adds valuable information for management of biodiversity in roadside stormwater ponds. *Ecol. Evol.* 9, 9712–9722. doi:10.1002/ece3.5503

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., et al. 2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Res.* 35, e14. doi:10.1093/nar/gkl938

Taberlet, P., Prud'Homme, S., Campione, E., Roy, J., Miquel, C., Shehzad, W., et al. (2012). Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. *Mol. Ecol.* 21, 1816–1820. doi:10.1111/j.1365-294X.2011.05317.x

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., et al. (2004). Extinction risk from climate change. *Nature* 427, 145–148. doi:10.1038/nature02121

Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T., et al. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Mol. Ecol.* 21, 2565–2573. doi:10.1111/j.1365-294X.2011.05418.x

Wallinger, C., Juen, A., Staudacher, K., Schallhart, N., Mitterrutzner, E., Steiner, E. M., et al. 2012). Rapid plant identification using species- and group-specific primers targeting chloroplast DNA. *PLoS One* 7, e29473. doi:10.1371/journal.pone.0029473

Walters, C. (1986). *Adaptive management of renewable resources.* New York, NY: Macmillan.

Willerslev, E., Cappellini, E., Boomsma, W., Nielsen, R., Hebsgaard, M., Brand, T. B., et al. (2007). Ancient biomolecules from deep ice cores reveal a forested Southern Greenland. *Science* 317, 111–114. doi:10.1126/science.1141758

Xie, Y., Zhang, X., Yang, J., Kim, S., Hong, S., Giesy, J. P., et al. (2018). eDNA-based bioassessment of coastal sediments impacted by an oil spill. *Environ. Pollut.* 238, 739–748. doi:10.1016/j.envpol.2018.02.081

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinf.* 13, 134. doi:10.1186/1471-2105-13-134

Zavaleta, E. S., Hobbs, R. J., and Mooney, H. A. (2001). Viewing invasive species removal in a whole-ecosystem context. *Trends Ecol. Evol.* 16, 454–459. doi:10.1016/S0169-5347(01)02194-2

Zobeck, T. M. (2006). Erosion by wind: field measurement. *Encyclopedia of Soil Science* 1, 607–612.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 28 January 2021
doi: 10.3389/fevo.2020.622291

Detection of Endangered Aquatic Plants in Rapid Streams Using Environmental DNA

Yuta Tsukamoto[1], *Satoru Yonezawa*[1], *Natsu Katayama*[2] *and Yuji Isagi*[1]*

[1] Laboratory of Forest Biology, Division of Forest & Biomaterials Science, Graduate School of Agriculture, Kyoto University, Kyoto, Japan, [2] Department of Biology, Faculty of Science, Chiba University, Chiba, Japan

Podostemaceae are a eudicot family of plants that grow on rapid streams and waterfalls. Two genera and six species of this family are distributed in Japan, all of which are threatened with extinction. It is difficult to find these species from the river side and it takes much effort to investigate their distribution. In this study, we attempted to determine the presence and absence of the Podostemaceae species by environmental DNA (eDNA) metabarcoding. Four species of Podostemaceae were detected near four known habitats, and the detected species were in perfect agreement with the results of a past survey that was based on visual observation. The marker used in this study had sufficient resolution to distinguish all six Podostemaceae species distributed in Japan and detected multiple species growing in a site. These results show that eDNA metabarcoding can quickly detect rare aquatic plants that are difficult to find by visual observation and can provide important information regarding their conservation.

Keywords: DNA metabarcoding, Podostemaceae, endangered species, habitat screening, *trn*L

OPEN ACCESS

Edited by:
Katy Klymus,
United States Geological Survey (USGS), United States

Reviewed by:
Hideyuki Doi,
University of Hyogo, Japan
Naiara Guimaraes Sales,
University of Salford, United Kingdom

***Correspondence:**
Yuji Isagi
isagiy@gmail.com

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *28 October 2020*
Accepted: *28 December 2020*
Published: *28 January 2021*

Citation:
Tsukamoto Y, Yonezawa S, Katayama N and Isagi Y (2021) Detection of Endangered Aquatic Plants in Rapid Streams Using Environmental DNA. Front. Ecol. Evol. 8:622291. doi: 10.3389/fevo.2020.622291

INTRODUCTION

Podostemaceae are aquatic angiosperms that live in rapid streams and waterfalls (Cook and Rutishauser, 2007; Koi et al., 2012). They are distributed in tropical and subtropical regions in America, Africa, Madagascar, Asia, and Australia and consist of ∼50 genera and 300 species (Koi et al., 2012). Two genera and six species of Podostemaceae (*Cladopus doianus*, *C. fukienensis*, *Hydrobryum floribundum*, *H. japonicum*, *H. koribanum*, and *H. puncticulatum*) are distributed in the Kyushu, southwestern region of Japan (Kato, 2008, 2013). Most of these species and populations are registered as natural monuments by national or local governments, and all species are listed in the Red List of Japan (Kato, 2013). It is difficult to find them because they look like bryophytes and are flat (**Figure 1**). Therefore, unknown populations still could exist. The improvement of detection methods for hard-to-find taxa is needed to accurately assess local biodiversity. Recently, the environmental DNA (eDNA) survey was developed as an efficient and sensitive approach to determine the distribution of rare aquatic species (Thomsen and Willerslev, 2015). Several studies applied the eDNA survey to aquatic plants and successfully detected the target species among collected samples in aquariums, rivers, and ponds (Scriver et al., 2015; Fujiwara et al., 2016; Matsuhashi et al., 2016; Gantz et al., 2018; Kuzmina et al., 2018; Anglès d'Auriac et al., 2019; Chase et al., 2020; Coghlan et al., 2020; Doi et al., 2020; Kuehne et al., 2020; Miyazono et al., 2020). Some of these works report the detection of some populations that had not been recorded using conventional observation (Matsuhashi et al., 2016; Kuzmina et al., 2018; Miyazono et al., 2020). Most of the previous eDNA studies of aquatic plants focused on detecting invasive alien species

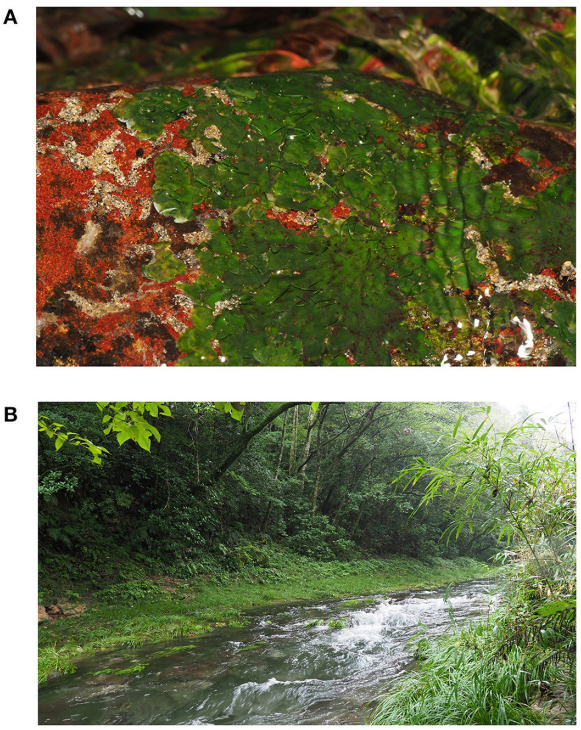


FIGURE 1 | *Hydrobryum japonicum* **(A)** and its habitat **(B)**. This corresponds to "TM1" in **Figure 2**.

(Scriver et al., 2015; Fujiwara et al., 2016; Gantz et al., 2018; Anglès d'Auriac et al., 2019; Chase et al., 2020; Doi et al., 2020; Kuehne et al., 2020; Miyazono et al., 2020), and no eDNA studies have attempted to search for rare aquatic plant species. In the case of Podostemaceae, an eDNA survey could be especially useful because of their moss-like morphology, hard-to-explore habitats, low-frequency occurrence, and possibly small population size. While quantitative PCR is primarily used to detect species- or taxon-specific plant eDNA (Scriver et al., 2015; Fujiwara et al., 2016; Matsuhashi et al., 2016; Gantz et al., 2018; Anglès d'Auriac et al., 2019; Chase et al., 2020; Doi et al., 2020; Kuehne et al., 2020; Miyazono et al., 2020), the DNA metabarcoding method has some advantages: it does not require the development of a species-specific primer set so it can easily be applied to various ecosystems (Yonezawa et al., 2020), and it can detect multiple target species using a single marker. In this study, we attempted to detect Podostemaceae from water samples using DNA metabarcoding and show that the method is useful for investigating the distribution of rare and difficult-to-find aquatic plants.

MATERIALS AND METHODS

Collection of eDNA Samples and DNA Extraction

We collected 16 water samples from 16 sites in the Miyazaki and Kagoshima prefectures, southwestern Japan, in September of 2017 (seven samples) and May of 2019 (nine samples) (**Figure 2** and **Table 1**). Among the sampling sites, *H. koribanum* grows in the Iwase River (IW2-IW4) in Miyazaki Prefecture, *C. doianus* and *H. floribundum* grow in the Anraku River (AR1) in Kagoshima Prefecture, *H. japonicum* grows in the Tomano River (TM1), and *C. doianus* grows in the Nakatsu River (NK2, NK3). In the Red List of Japan (https://ikilog.biodic.go.jp/Rdb/), *H. koribanum* is listed as "critically endangered" (CR), *C. doianus* is listed as "endangered" (EN), and *H. floribundum* and *H. japonicum* are listed as "vulnerable" (VU). Other samples were collected downstream of the known habitats (NK4, TM2–TM4), upstream of the known habitats (AM1, NK1, and IW1), or in other rivers (US1 and KN1). Immediately after collection, the water samples were filtered using 50 mL sterile syringes (Thermo Co., Tokyo, Japan) and 0.45 μm Sterivex filter cartridges (Millipore, MA, USA) until the Sterivex filters were clogged (250–1500 mL, see **Table 1**). One field negative control was prepared at the first sampling points of each sampling day in 2019 (May 18: IW1 and May 19 US1) to monitor the contamination from the sampling instrument. A measure of 1,000 mL of pure water was filtered using sterile syringes and Sterivex filter cartridges before conducting the field sampling. The Sterivex filter cartridges were kept at 4°C during transportation to the laboratory, and then kept at −20°C until DNA extraction. In the 2019 sampling, we added RNAlater (Qiagen, Hilden, Germany) to Sterivex after filtration to preserve the DNA better. We extracted eDNA from the Sterivex filter cartridges using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) and the method reported by Miya et al. (2016), with modification of the incubation time to 1 h. The extraction step was conducted in the same room as the generating reference database below, and one negative control was prepared to monitor contamination from the laboratory equipment and cross-contamination between samples. An extra lysis buffer was processed following the extraction protocol and treated in the same way as the other samples in subsequent steps.

Development of the *trn*L Reference Database for Podostemaceae

We used the universal primer *trn*L g/h (Taberlet et al., 2007) for DNA metabarcoding. This marker is short and widely used for analyzing degraded DNA in seed plants (Taberlet et al., 2018). To develop a reference database of Podostemaceae in Japan, we obtained their *trn*L P6 loop sequences according to Ando et al. (2013). Briefly, the whole chloroplast *trn*L introns of six Podostemaceae species from Japan (*C. doianus*, *C. fukienensis*, *H. floribundum*, *H. japonicum*, *H. koribanum*, and *H. puncticulatum*) were amplified using the *trn*L c/d primer (Taberlet et al., 2007). Cycle sequencing was performed using a Big Dye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) according to the standard protocol. The cycle-sequencing products were visualized on an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems). Subsequently, P6 loop sequences (Taberlet et al., 2007) were extracted from the entire *trn*L intron sequences. The sequences were aligned using MEGA X (Kumar et al., 2018), which confirmed the presence of at least one differences among them (**Table 2**). These sequences

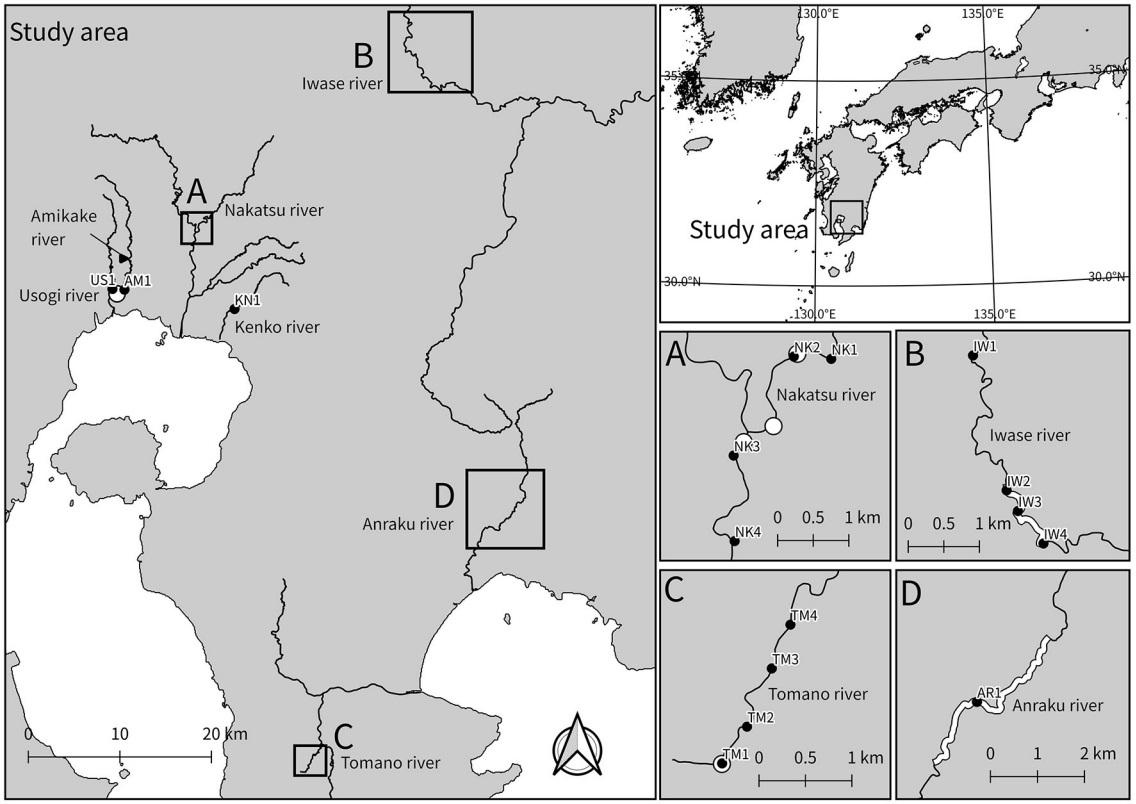


FIGURE 2 | Sampling locations for eDNA analysis. In Japan, all species of Podostemaceae are distributed in Kyushu, southwestern region. The closed circles denote the sampling points, and the white circles and lines denote the known habitats. This map was constructed using the data provided National Land Information Division, National Spatial Planning and Regional Policy Bureau, Ministry of Land, Infrastructure, Transport and Tourism of Japan.

were also searched using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi), which confirmed that they are not similar to any other plant species in the NCBI database. Finally, a custom database for BLAST+ 2.6.0 (Camacho et al., 2009) was generated using the "makeblastdb" command.

PCR, Library Preparation, and Illumina iSeq Sequencing

A two-step PCR protocol was used for library preparation. All instruments and tubes were autoclaved in advance, and separate rooms were used for procedures with and without PCR products. Negative controls were prepared for the first-round PCR and second-round PCR using MilliQ instead of sample eDNA solution to monitor cross-contamination during the procedure. In the first-round PCR, we used *trn*L g/h primer (Taberlet et al., 2007) concatenated with sequencing primer binding region and six random bases (Forward: 5′-ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNGGGCAATCCTGAGCCAA-3′, Reverse: 5′-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNCCATTGAGTCTCTGCACCTATC-3′, the "N" indicates random base). The first-round PCR was conducted in a volume of 11 μL containing 1 μL of the extracted eDNA, 0.2 μL of KOD FX Neo (Toyobo, Osaka, Japan), 6 μL of 2× PCR buffer, 2.4 μL of 2 mM dNTP, 0.7 μL of each forward/reverse primer (5 μM), and 2 μL of MilliQ water. The thermal cycling profile was as follows: 94°C for 2 min, followed by 35 cycles of 98°C for 10 s, 57°C for 30 s, 68°C for 30 min, and a final incubation at 68°C for 5 min. We performed triplicate first-round PCRs, and those replicates were pooled. The first-round of PCR products were purified using Exonuclease I (TaKaRa Bio, Otsu, Japan) and TSAP (Promega, Madison, WI, USA). In the second-round PCR, we used the primers that consisted of Illumina adapters, indices, and sequencing primer binding region (Forward: 5′-AATGATACGGCGACCACCGAGATCTACACXXXXXXXXACACTCTTTCCCTACACGACGCTCTTCCGATCT-3′, Reverse: 5′-CAAGCAGAAGACGGCATACGAGATXXXXXXXXGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3′, the "X" indicates index). The second-round PCR was carried out in a volume of 24 μL containing 2 μL of the first-round PCR products, 12 μL of KAPA HiFi HotStart ReadyMix (2×) (KAPA Biosystems, Wilmington, WA, USA), 1.4 μL of each forward/reverse primer (5 μM), and 7.2 μL of MilliQ water. The thermal cycling profile was as follows: 95°C for 3 min, followed by 12 cycles of 98°C for 20 s, 72°C for 15 s, and a final incubation at 72°C for 5 min. The 21 second-round PCR products (16 field samples, 2 field negative controls, and 3 laboratory negative controls) were mixed in equal amount and the mixed library was purified using Agencourt AMPure XP (Beckman Coulter,

TABLE 1 | Sampling sites, the inhabiting species of Podostemaceae at the sites, and their classification in the Red List of Japan.

Sites	River	Locality	Date	Inhabiting species (Red List Category)	Amount of water (mL)
US1	Usogi River	Kajikicho, Aira, Kagoshima	2019/5/19	Not reported	400
AM1	Amikake River	Kajikicho, Aira, Kagoshima	2019/5/19	Not reported	250
NK1	Nakatsu River	Makizonocho, Kirishima, Kagoshima	2019/5/19	Not reported	400
NK2	Nakatsu River	Makizonocho, Kirishima, Kagoshima	2019/5/19	*C. doianus* (endangered)	1,000
NK3	Nakatsu River	Hayatocho, Kirishima, Kagoshima	2019/5/19	*C. doianus* (endangered)	800
NK4	Nakatsu River	Hayatocho, Kirishima, Kagoshima	2019/5/19	Not reported	750
KN1	Kenko River	Kokubudaimyoji, Kirishima, Kagoshima	2019/5/19	Not reported	950
IW1	Iwase River	Higashikata, Kobayashi, Miyazaki	2019/5/18	Not reported	250
IW2	Iwase River	Nojiricho, Kobayashi, Miyazaki	2019/5/18	*H. koribanum* (critically endangered)	350
IW3	Iwase River	Nojiricho, Kobayashi, Miyazaki	2017/9/20	*H. koribanum* (critically endangered)	300
IW4	Iwase River	Nojiricho, Kobayashi, Miyazaki	2017/9/20	*H. koribanum* (critically endangered)	400
TM1	Tomano River	Airacho, Kanoya, Kagoshima	2017/9/21	*H. japonicum* (vulnerable)	1,500
TM2	Tomano River	Airacho, Kanoya, Kagoshima	2017/9/21	Not reported	1,500
TM3	Tomano River	Airacho, Kanoya, Kagoshima	2017/9/21	Not reported	700
TM4	Tomano River	Airacho, Kanoya, Kagoshima	2017/9/21	Not reported	1,500
AR1	Anraku River	Shibushicho, Shibushi, Kagoshima	2017/9/20	*C. doianus* (endangered) *H. floribundum* (vulnerable)	1,000

TABLE 2 | Variable sites of six Japanese Podostemaceae species in *trn*L P6 loop sequences (66–71 bp).

	Site									
Species	**8**	**9**	**20**	**21**	**22**	**34**	**39**	**57**	**66**	**70**
C. doianus	G	T	C	A	A	T	A	A	C	A
C. fukienensis	G	T	A	A	A	T	A	A	C	A
H. floribundum	–	–	C	A	A	T	C	G	A	G
H. japonicum	–	–	A	A	A	T	C	G	A	G
H. koribanum	–	–	–	–	–	G	C	G	A	A
H. puncticulatum	–	–	–	–	–	T	C	G	A	A

All sequences differed from each other in at least one base.

High, Wycombe, UK). The concentration of the DNA library was measured using a Qubit dsDNA HS Assay Kit and a Qubit 3.0 Fluorometer (Life Technologies, Paisley, UK). We confirmed that the size of the PCR products was within the expected range (200–400 bp; those of Podostemaceae were around 250 bp) using MultiNA (Shimadzu, Kyoto, Japan). Subsequently, the concentration of the library was adjusted to 35 pM with resuspension buffer (RSB) (Illumina, San Diego, CA, USA), and a 25% PhiX spike-in was added. Paired-end sequencing (150 bp × 2) was performed on an iSeq 100 platform (Illumina, San Diego, CA, USA) with iSeq 100 i1 Reagent (300 cycles) (Illumina, San Diego, CA, USA).

Data Analysis

Raw reads from iSeq 100 were processed following DADA2 ITS Pipeline Workflow (1.8) (https://benjjneb.github.io/dada2/ITS_workflow.html) (parameters as default, but minLen = 10 at filtering and trimming step) using a DADA2 package ver. 1.16.0 (https://github.com/benjjneb/dada2) on R 3.6.3 (R Core Team, 2020) and cutadapt 2.4 (Martin, 2011). The DADA2 algorithm (Callahan et al., 2016) corrects amplicon errors with single-nucleotide resolution and this pipeline is used to analyze highly variable regions, which are suitable for the *trn*L P6 loop sequences (10–220 bp, Taberlet et al., 2018). This pipeline consists of trimming, filtering, a learning error model from the sequence data, dereplication, correcting errors based on the error model, merging paired-end reads, removing chimera, and a generating amplicon sequence variants (ASV) table. The ASV can be analyzed in a similar way to the traditional operational taxonomic unit (van der Heyde et al., 2020). We processed the ASV table in a specific manner to reduce false positives of target species: for some Illumina sequencers (e.g., Hiseq X and iSeq 100) with patterned flow cells and ExAmp chemistry, free-floating indexing primers that remain in the library can cause a read misassignment called index hopping (Costello et al., 2018), which can cause false positives of the target taxa in DNA metabarcoding. Although the rate of index hopping in iSeq 100 is unknown, in Hiseq X, which uses a patterned flow cell and ExAmp chemistry, as in iSeq 100, the rate is estimated to be 0.470% using a single-indexed library (van der Valk et al., 2019). Assuming that hopping occurs independently at both ends of the library, the percentage of reads with at least one wrong index in this study was estimated to be $1 - (1 - 0.00470)^2 = 0.00937 = 1\%$. To ensure that index hopping would not cause false positives for Podostemaceae, when the number of reads of an ASV detected in each sample were <1% of the number of reads of the ASV detected in all samples, the reads of the ASV in the samples were removed. The ASVs were assigned using the *trn*L reference database for Podostemaceae and BLAST+ 2.6.0 (Camacho et al., 2009). Referring to similar studies using *trn*L P6 loop sequences (Nakahama et al., 2020), the top hits with at least 98% matching and e-values lower than 1.0×10^{-25} were used for species assignments. The ASVs which were not assigned to Podostemaceae were also identified using "clidentseq" and "classingtax" commands of Claident 0.2.2019.05.19 (Tanabe and

Toju, 2013) and its "semiall_family" database. This tool can identify the query sequence in the appropriate taxon level using the query-centric auto-k-nearest-neighbor (QCauto) method (Tanabe and Toju, 2013) based on the lowest common ancestor algorithm (Huson et al., 2007). The "semiall_family" database is based on NCBI Nucleotide database, and the *trn*L sequence of Podostemaceae in Japan was not registered at the time of analysis.

RESULTS

Sequencing Results

iSeq sequencing generated 2,773,218 reads, and sequence processing resulted in 400 ASVs of 2,443,502 reads from 21 samples, including five negative controls (**Supplementary Material 2**). All sequences can be found at DDBJ with accession numbers: DRA011228. Five negative controls (two field negative controls and three experimental negative controls) yielded 238,492 processed reads, 153,002 of which were attributed to plants at least at the family level. However, none of the negative controls included the sequence of Podostemaceae and no significant cross-contamination was considered to have occurred.

Detection of Podostemaceae

The *trn*L P6 loop sequences of Podostemaceae in Japan differed from each other in at least one base (**Table 2**) and no plants other than Podostemaceae were hit in the BLAST analysis. Those of *C. doianus* and *H. floribundum*, which grow sympatrically in the Anraku River, differed in four loci. As a result of BLAST analysis using the *trn*L reference database, Six ASVs were assigned to four Podostemaceae species in Japan. They were detected near or downstream of the known habitats (**Table 3**), with the exception of NK4, which was collected downstream, about 2 km away from known habitats (**Figure 1A**). The detected species agreed with previous records. In the Nakatsu River (NK2 and NK3), where *C. doianus* grows, and the Iwase River (IW2, IW3, and IW4), where *H. koribanum* grows, each species was detected. In the Tomano River, *H. japonicum* was detected from the vicinity of the known habitat (TM1) up to 2 km downstream (TM2–TM4). In the Anraku River (AR1), where *C. doianus* and *H. floribundum* grow sympatrically (Terada and Ohya, 2009), both species were distinguished and detected, which indicates the usefulness of this method for detailed species identification. Podostemaceae were not detected upstream of known habitats (IW1, AM1, and NK1) or in water systems outside of known habitats (US1 and KN1).

Taxon Assignment of Other Seed Plants

As a result of the Claident analysis to identify plants other than Podostemaceae, 180 of 400 ASVs were identified in 71 families, 57 of which were identified in 45 genera other than Podostemaceae (**Supplementary Material 1**). As the *trn*L P6 loop sequences are short and sometimes they do not have enough resolutions at the species level, only three non-target species (*Phragmites australis*, *Potamogeton crispus*, and *Ginkgo biloba*) were identified to species-level. As most of the rivers with sampling points run through a forest, many reads were assigned to trees that are typical in this region, such as Fagaceae, Theaceae, and Lauraceae. From the samples collected downstream of paddy fields, plants of the Oryza genus, which includes the rice crop species (*O. sativa*), were detected. Some aquatic taxa were also detected, such as *Phragmites australis*, *Potamogeton*, and *Acorus*.

DISCUSSION

Distribution of Podostemaceae Around the Study Area

eDNA derived from Podostemaceae, which are difficult to find from the river side, was all successfully detected in water samples collected near the known habitats. Especially, in the Anraku River (AR1), where *C. doianus* and *H. floribundum* are known to grow sympatrically (Terada and Ohya, 2009), both species were successfully detected and distinguished, indicating the effectiveness of the method for multi-species distribution. A large number of *H. japonicum* sequences were obtained from TM2–TM4, which were collected at 500 m to 2 km downstream of a known locality of *H. japonicum* (TM1). This could be because the eDNA from the known population at TM1 traveled downstream, or because there are unknown populations near TM2–TM4. It is difficult to draw a clear conclusion about their distribution because the transport and detection distances of eDNA in lotic ecosystems can be influenced by various biological and non-biological factors (Beng and Corlett, 2020). Presence in such an area can be confirmed by a detailed field survey. Alternatively, the information from eDNA surveys can be used to select an area that requires more extensive visual surveying. For taxa that are difficult to find by visual observation from the river side, this eDNA approach of screening sites where unknown populations are likely to exist was effective.

Usefulness of eDNA Survey for Determining the Distribution of Podostemaceae

In Japan, all species of Podostemaceae are endangered and their habitats are very limited. Some populations have declined or disappeared because of development and deterioration of water quality (Noro et al., 1993; Kato, 2013), while others have been newly found far from known localities (Seno and Hattori, 2013). Continuous monitoring and searching for unknown populations are essential for the conservation of this family of plants. Although the *trn*L P6 loop sequences of Podostemaceae in Japan constitute a short region of only about 70 bp, all six species could be distinguished using this region (**Table 2**). This method can be applied to detect other species of Podostemaceae in Japan. Although it was raining at the time of the 2019 survey, which can dilute eDNA concentrations and lead to false negatives for target species (Curtis et al., 2020; Sales et al., 2021), Podostemaceae eDNA was detectable in all samples near known habitats. Therefore, the method employed in this study seems to be effective under various weather conditions. However, the effectiveness of the detection of a new population of rare species using this approach would increase if the survey was conducted during the season of low rainfall. The investigation of the distribution of Podostemaceae has been performed using

TABLE 3 | Number of reads assigned to Podostemaceae in the samples collected near the known habitats.

Detected species	Sampling sites and known inhabiting species									
	NK2	NK3	IW2	IW3	IW4	AR1	TM1	TM2	TM3	TM4
	C. doianus		*H. koribanum*			*C. doianus* and *H. floribundum*	*H. japonicum*	Not reported*		
C. doianus	2,888	6,033	0	0	0	5,924	0	0	0	0
H. koribanum	0	0	802	9,395	4,109	0	0	0	0	0
H. floribundum	0	0	0	0	0	30,600	0	0	0	0
H. japonicum	0	0	0	0	0	0	41,377	226,419	60,771	127,248
Processed Reads	186,111	134,491	14,576	66,405	106,324	201,556	133,787	228,010	198,177	169,597

The name of Podostemaceae species under sampling site indicate known inhabiting species.
**No Podostemaceae species have been reported at TM2–TM4.*

visual observation, which is time- and labor-consuming and requires specific knowledge. If eDNA can be used to survey the distribution of Podostemaceae in rivers without investigating the inside of rivers, it will be possible to identify their presence and range more efficiently.

Application of eDNA Survey to Other Aquatic Plants

As eDNA metabarcoding system does not require the development of species-specific markers, our assay can be applied easily to other rare aquatic plants if these species can be distinguished from others using the *trn*L P6 loop sequence. In fact, *Potamogeton crispus*, a common submerged plant, was detected at three sites, although it could not be found visually during the field sampling. Distribution survey is essential for the conservation of rare aquatic plants, however, it is difficult to find them because of their submerged nature and aquatic habitats are often hard to approach (Kuzmina et al., 2018; Coghlan et al., 2020). eDNA metabarcoding using a short marker will quickly inform on the distribution of rare aquatic plants and provide useful information for their conservation.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: DDBJ; DRA011228; LC600213-LC600218.

AUTHOR CONTRIBUTIONS

YI conceived the study and revised the manuscript critically for important intellectual content. YT, SY, and YI conducted the field sampling. NK obtained the DNA samples for all Podostemaceae species in Japan. NK, YT, and SY performed the experiments. YT and SY analyzed the data and wrote the manuscript. All authors read and approved the final manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.622291/full#supplementary-material

REFERENCES

Ando, H., Setsuko, S., Horikoshi, K., Suzuki, H., Umehara, S., Inoue-Murayama, M., et al. (2013). Diet analysis by next-generation sequencing indicates the frequent consumption of introduced plants by the critically endangered red-headed wood pigeon (*Columba janthina* nitens) in oceanic island habitats. *Ecol. Evol.* 3, 4057–4069. doi: 10.1002/ece3.773

Anglès d'Auriac, M. B., Strand, D. A., Mjelde, M., Demars, B. O. L., and Thaulow, J. (2019). Detection of an invasive aquatic plant in natural water bodies using environmental DNA. *PLoS ONE* 14:e0219700. doi: 10.1371/journal.pone.0219700

Beng, K. C., and Corlett, R. T. (2020). Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. *Biodivers. Conserv.* 29, 2089–2121. doi: 10.1007/s10531-020-01980-0

Callahan, B. J., Mcmurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2 : high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421

Chase, D. M., Kuehne, L. M., Olden, J. D., and Ostberg, C. O. (2020). Development of a quantitative PCR assay for detecting *Egeria densa* in environmental DNA samples. *Conserv. Genet. Resour.* 12, 545–548. doi: 10.1007/s12686-020-01152-w

Coghlan, S. A., Shafer, A. B. A., and Freeland, J. R. (2020). Development of an environmental DNA metabarcoding assay for aquatic vascular plant communities. *Environ. DNA*. doi: 10.1002/edn3.120

Cook, C. D. K., and Rutishauser, R. (2007). "Podostemaceae," in *Flowering Plants Eudicots*, ed K. Kubitzki (Berlin; Heidelberg: Springer), 304–344. doi: 10.1007/978-3-540-32219-1_40

Costello, M., Fleharty, M., Abreu, J., Farjoun, Y., Ferriera, S., Holmes, L., et al. (2018). Characterization and remediation of sample index swaps by non-redundant dual indexing on massively parallel sequencing platforms. *BMC Genomics* 19:332. doi: 10.1186/s12864-018-4703-0

Curtis, A. N., Tiemann, J. S., Davis, M. A., Larson, E. R., and Douglass, S. A. (2020). High stream flows dilute environmental DNA (eDNA) concentrations and reduce detectability. *Divers. Distrib.* 1–14. doi: 10.1111/ddi.13196

Doi, H., Akamatsu, Y., Goto, M., Inui, R., and Nagano, M. (2020). Broad-scale detection of environmental DNA for an invasive macrophyte and the relationship between DNA concentration and coverage in rivers. *Biol. Invasions* 9, 1–14. doi: 10.1007/s10530-020-02380-9

Fujiwara, A., Matsuhashi, S., Doi, H., Yamamoto, S., and Minamoto, T. (2016). Use of environmental DNA to survey the distribution of an invasive submerged plant in ponds. *Freshw. Sci.* 35, 748–754. doi: 10.1086/685882

Gantz, C. A., Renshaw, M. A., Erickson, D., Lodge, D. M., and Egan, S. P. (2018). Environmental DNA detection of aquatic invasive plants in lab mesocosm and natural field conditions. *Biol. Invasions* 20, 2535–2552. doi: 10.1007/s10530-018-1718-z

Huson, D. H., Auch, A. F., Qi, J., and Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Res.* 17, 377–386. doi: 10.1101/gr.5969107

Kato, M. (2013). *The Illustrated Book of Plant Systematics in Color: Podostemaceae of the World*. Tokyo: Hokuryukan.

Kato, M. (2008). A taxonomic study of podostemaceae of Japan. *Bull. Natl. Museum Nat. Sci. Ser. B Bot.* 34, 63–73.

Koi, S., Kita, Y., Hirayama, Y., Rutishauser, R., Huber, K. A., and Kato, M. (2012). Molecular phylogenetic analysis of Podostemaceae: implications for taxonomy of major groups. *Bot. J. Linn. Soc.* 169, 461–492. doi: 10.1111/j.1095-8339.2012.01258.x

Kuehne, L. M., Ostberg, C. O., Chase, D. M., Duda, J. J., and Olden, J. D. (2020). Use of environmental DNA to detect the invasive aquatic plants *Myriophyllum spicatum* and *Egeria densa* in lakes. *Freshw. Sci.* 39, 521–533. doi: 10.1086/710106

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Kuzmina, M. L., Braukmann, T. W. A., and Zakharov, E. V. (2018). Finding the pond through the weeds: eDNA reveals underestimated diversity of pondweeds. *Appl. Plant Sci.* 6:e01155. doi: 10.1002/aps3.1155

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Matsuhashi, S., Doi, H., Fujiwara, A., Watanabe, S., and Minamoto, T. (2016). Evaluation of the environmental DNA method for estimating distribution and biomass of submerged aquatic plants. *PLoS ONE* 11:e156217. doi: 10.1371/journal.pone.0156217

Miya, M., Minamoto, T., Yamanaka, H., Oka, S., Sato, K., Yamamoto, S., et al. (2016). Use of a filter cartridge for filtration of water samples and extraction of environmental DNA. *J. Vis. Exp.* 54741. doi: 10.3791/54741

Miyazono, S., Kodama, T., Akamatsu, Y., Nakao, R., and Saito, M. (2020). Application of environmental DNA methods for the detection and abundance estimation of invasive aquatic plant *Egeria densa* in lotic habitats. *Limnology*. doi: 10.1007/s10201-020-00636-w

Nakahama, N., Furuta, T., Ando, H., Setsuko, S., Takayanagi, A., and Isagi, Y. (2020). DNA meta-barcoding revealed that sika deer foraging strategies vary with season in a forest with degraded understory vegetation. *For. Ecol. Manage* 2020:118637. doi: 10.1016/j.foreco.2020.118637

Noro, T., Hiroshi, S., Fujita, T., and Toshihide, K. (1993). Distribution of Podostemaceae in Southern Japan. *J. Jpn. Bot.* 68, 253–260.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Available online at: https://www.r-project.org/

Sales, N. G., Wangensteen, O. S., Carvalho, D. C., Deiner, K., Præbel, K., Coscia, I., et al. (2021). Space-time dynamics in monitoring neotropical fish communities using eDNA metabarcoding. *Sci. Total Environ.* 754:142096. doi: 10.1016/j.scitotenv.2020.142096

Scriver, M., Marinich, A., Wilson, C., and Freeland, J. (2015). Development of species-specific environmental DNA (eDNA) markers for invasive aquatic plants. *Aquat. Bot.* 122, 27–31. doi: 10.1016/j.aquabot.2015.01.003

Seno, J., and Hattori, M. (2013). The northernmost locality of *Cladopus japonicus* imamura (Podostemaceae). *J. Jpn. Bot.* 88, 61–63.

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., et al. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Res.* 35:e14. doi: 10.1093/nar/gkl938

Taberlet, P., Bonin, A., Coissac, E., and Zinger, L. (2018). *Environmental DNA: For Biodiversity Research and Monitoring*. Oxford: Oxford University Press.

Tanabe, A., and Toju, H. (2013). Two new computational methods for universal DNA barcoding : a benchmark using barcode sequences of bacteria, archaea, animals, fungi, and land plants. *PLoS ONE* 8:e0076910. doi: 10.1371/journal.pone.0076910

Terada, J., and Ohya, S. (2009). The podostemaceae plants: the vegetation in anrakugawa and maegawa rivers in Sibushi City, Kagoshima. *Kagoshima Prefect. Museum* 28, 1–28.

Thomsen, P. F., and Willerslev, E. (2015). Environmental DNA–an emerging tool in conservation for monitoring past and present biodiversity. *Biol. Conserv.* 183, 4–18. doi: 10.1016/j.biocon.2014.11.019

van der Heyde, M., Bunce, M., Wardell-Johnson, G., Fernandes, K., White, N. E., and Nevill, P. (2020). Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA metabarcoding. *Mol. Ecol. Resour.* 20, 732–745. doi: 10.1111/1755-0998.13148

van der Valk, T., Vezzi, F., Ormestad, M., Dalén, L., and Guschanski, K. (2019). Index hopping on the Illumina HiseqX platform and its consequences for ancient DNA studies. *Mol. Ecol. Resour.* 20, 1171–1181. doi: 10.1111/1755-0998.13009

Yonezawa, S., Ushio, M., Yamanaka, H., Miya, M., Takayanagi, A., and Isagi, Y. (2020). Environmental DNA metabarcoding reveals the presence of a small, quick-moving, nocturnal water shrew in a forest stream. *Conserv. Genet.* 21, 1079–1084. doi: 10.1007/s10592-020-01310-5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 25 February 2021
doi: 10.3389/fevo.2021.621461

Distribution of Japanese Eel *Anguilla japonica* Revealed by Environmental DNA

Akihide Kasai[1], Aya Yamazaki[1], Hyojin Ahn[1], Hiroki Yamanaka[2], Satoshi Kameyama[3], Reiji Masuda[4], Nobuyuki Azuma[5], Shingo Kimura[6], Tatsuro Karaki[1,7], Yuko Kurokawa[8] and Yoh Yamashita[4]*

[1] Faculty of Fisheries Sciences, Hokkaido University, Hakodate, Japan, [2] Faculty of Science and Technology, Ryukoku University, Otsu, Japan, [3] Center for Environmental Biology and Ecosystem Studies, National Institute for Environmental Studies, Tsukuba, Japan, [4] Field Science Education and Research Center, Kyoto University, Kyoto, Japan, [5] Faculty of Agriculture and Life Science, Hirosaki University, Aomori, Japan, [6] Graduate School of Frontier Sciences, Atmosphere and Ocean Research Institute, The University of Tokyo, Kashiwa, Japan, [7] Center for the Promotion of Interdisciplinary Education and Research, Kyoto University, Kyoto, Japan, [8] Division of Food Science and Nutrition, Tohoku Seikatsu Bunka Junior College, Sendai, Japan

OPEN ACCESS

Edited by:
David Jack Coates, Conservation and Attractions (DBCA), Australia

Reviewed by:
Ryan P. Kelly, University of Washington, United States
Miguel Alfonso Ortega-Huerta, National Autonomous University of Mexico, Mexico

***Correspondence:**
Akihide Kasai
akihide@fish.hokudai.ac.jp

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *26 October 2020*
Accepted: *02 February 2021*
Published: *25 February 2021*

Citation:
Kasai A, Yamazaki A, Ahn H, Yamanaka H, Kameyama S, Masuda R, Azuma N, Kimura S, Karaki T, Kurokawa Y and Yamashita Y (2021) Distribution of Japanese Eel Anguilla japonica Revealed by Environmental DNA. Front. Ecol. Evol. 9:621461. doi: 10.3389/fevo.2021.621461

The abundance of Japanese eel *Anguilla japonica* has rapidly decreased in recent decades. Following a re-evaluation of the possibility of extinction, the Japanese Ministry of the Environment and the International Union for Conservation of Nature listed the Japanese eel as an endangered species in 2013 and 2014, respectively. However, their abundance and precise distribution have never been clarified owing to their nocturnality and difficulty in their capture. In this study, the distribution of Japanese eels was investigated by monitoring for environmental DNA (eDNA), a non-invasive and efficient detection method. A total of 365 water samples were collected from 265 rivers located throughout Japan. High concentrations of eDNA of Japanese eels were detected in rivers on the Pacific side, but were low in the Sea of Japan side. In particular, very little eDNA amplification was confirmed from Hokkaido and the north of the Sea of Japan. The eDNA distribution in Japanese rivers coincides with the transport of the larvae in the ocean, as estimated by numerical simulations. Generalized linear mixed models were developed to explain the distribution of eDNA concentrations. The total nitrogen concentration emerged as an important factor in the best model. These results indicate that the distribution of Japanese eel is mostly determined by the maritime larval transport, and their survival and growth depend on the abundance of food in the river. The findings of the present study are useful for the management of populations and in the conservation of Japanese eels.

Keywords: distribution, endangered species, environmental DNA, Japanese eel, *Anguilla japonica*, numerical simulation

INTRODUCTION

The Japanese eel (*Anguilla japonica*) is a catadromous species distributed in the western Pacific Ocean. It is a commercially important fishery species in East Asia and is very popular as traditional seafood in Japan. However, the catch of naturally occurring eels in Japanese rivers has rapidly decreased in recent decades, indicating a remarkable decline in its biomass. In response to this

situation, the Japanese eel was classified as an endangered species in Japan by the Ministry of the Environment (2013). The International Union for the Conservation of Nature has also classified it as Endangered in the Red List of Threatened Species (Jacoby and Gollock, 2014). This decline could be related to factors such as overfishing, loss of habitat (Tatsukawa, 2003), and changes in the ocean environment affecting larval survival and recruitment (e.g., Kimura et al., 2001). However, the specific cause remains unclear.

Japanese eels spawn in the west of the Mariana Ridge (e.g., Tsukamoto, 2006). After hatching, the larvae (leptocephali) are transported westward by the North Equatorial Current for more than thousands of kilometers from the spawning site. The North Equatorial Current bifurcates into the northward Kuroshio Current and the southward Mindanao Current at its westernmost boundary off the east coast of the Philippines. Larvae enter the Kuroshio Current in the bifurcation zone, where they metamorphose into glass eels, and are transported to the coast (**Figure 1**). They then swim toward nearby estuaries and rivers for further growth (Zenimoto et al., 2009; Han et al., 2012; Hsiung et al., 2018). Although it has been suggested that some eels spend their entire lives in seawater without experiencing freshwater environments (Tsukamoto and Arai, 2001), freshwater and brackish waters are generally the main habitats for this species. A recent study on otolith strontium isotope ratios of adults collected from the spawning area has shown that they inhabited Japanese rivers and/or estuaries in the juvenile stage (Otake et al., 2019). This result strongly supports the idea that Japanese rivers and/or estuaries act as nurseries for Japanese eels and are hence important for their reproduction. It is essential to note that the lower reaches of rivers are the main nurseries for eels because they spend the longest part of their life in these habitats, and humans can manage and preserve these environments.

However, much remains unknown about the ecology of the Japanese eel. Despite its acknowledged existence, the life history of eels in the sea is still largely unknown. Research into the mechanism of spawning migration is not yet complete, although some progress has been made. In addition, Japanese eels migrate to East Asian countries, but little research has been done on their detailed distribution within those countries. To promote effective conservation measures within the constraints of human resource, time, and money, it is essential to determine their distribution within specific political boundaries and to identify areas where conservation measures should be prioritized.

One of the major difficulties in determining or confirming the distribution of an aquatic species such as the eel is the low detection rate when using conventional methods. Backpack electrofishers are often used for monitoring eels in freshwater areas (Reid, 2011), but they require advanced skills to collect eels from turbid waters. Eels tend to hide in refuges such as holes and crevices or burrows in the mud during the day (Aoyama et al., 2005), which further contributes to the difficulty in their capture. Furthermore, electrofishers are ineffective in deep areas and are disabled in saline water. Consequently, accurate data on the spatial and temporal distribution of Japanese eels are lacking. Alternative methods are required to monitor their precise distribution and abundance.

Animals and plants release their cellular material including DNA into aquatic or terrestrial environments through excrement, body mucus, blood, and sloughed tissues. This DNA suspended in environmental samples, such as water and sediment is called environmental DNA (eDNA, Miya et al., 2020). By filtering a certain amount of natural water, eDNA is concentrated and captured on a filter, from which it is extracted and subjected to various molecular biology experiments for detection of organisms (Deiner et al., 2015; Turner et al., 2015; Miya et al., 2016). Recent studies have demonstrated that eDNA detection can be a reliable method for determining the distribution of various species of fish in freshwater ecosystems (e.g., Minamoto et al., 2012; Thomsen et al., 2012; Yamanaka and Minamoto, 2016). eDNA detection methods are more sensitive than traditional sampling methods, such as electrofishing or visual surveys, particularly when determining the presence of species that are found at low densities (Jerde et al., 2011). Takeuchi et al. (2019) successfully detected eDNA in the southern west Mariana Ridge, which is one of the most plausible spawning areas of Japanese eels, but spawning eels have rarely been captured there (Kurogi et al., 2011). Itakura et al. (2019) conducted surveys of Japanese eels in several small rivers using both electrofishing and eDNA analysis. In their study, the eDNA of Japanese eels was detected at 56 (91.8%) of the 61 study sites from which individuals were collected by electrofishing, and at an additional 35 sites where individuals were not directly collected. This indicates that eDNA analysis was more sensitive for detecting the presence of eels than electrofishing. They also indicated the possibility of estimating eel abundance and biomass from eDNA concentrations, by showing a weak but significant positive relationship between the eDNA concentration and the abundance and biomass of eels.

It is also difficult to survey many places in a short period using conventional methods. Since eels are an important fishery resource, habitat surveys have been conducted in various rivers by many Prefectures (e.g., Research Institute of Environment, Agriculture and Fisheries, Osaka Prefecture, 2014). However, each survey is limited to less than several rivers in each target region, and there are no studies that cover the entire country. On the other hand, the Ministry of Land, Infrastructure, Transport, and Tourism conducts the National Census on River Environments to monitor the environment and living organisms of 109 first-class rivers across the county. The census provides valuable information on the distribution of organisms, including eels. However, it takes 5 years to survey all the rivers. Nakamura et al. (2015) mentioned that the number of survey points in each river is small, and it is necessary to increase the sampling effort for obtaining detailed data on the environment and biodiversity of the rivers, and based on this data, appropriate river improvements should be undertaken. However, it is difficult to increase the scope and frequency of surveys because of the costs (for labor and to undertake the surveys) involved (Fujita, 2018). Recently, the applicability of the eDNA metabarcoding assay for fish surveys was investigated using case studies of the National Census on River Environments (Kitagawa et al., 2020). The checklist of fish compiled using eDNA metabarcoding and traditional sampling methods were compared, and the results

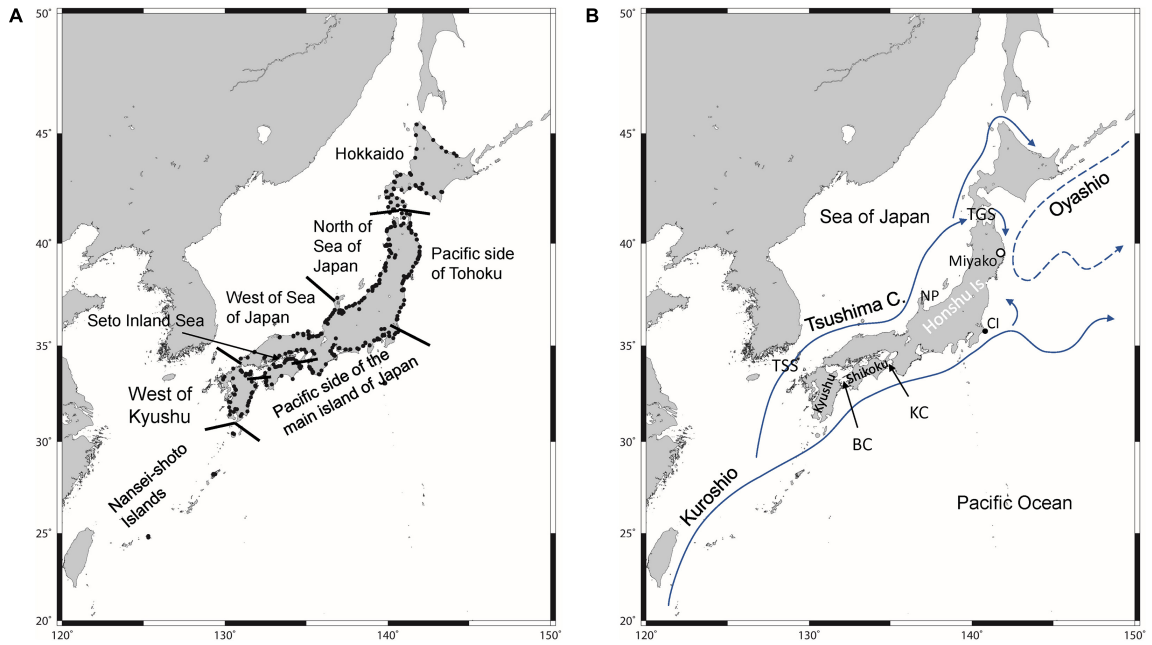


FIGURE 1 | Study area: **(A)** locations from where samples were collected to test for environmental DNA of Japanese eel. Black dots indicate sampling points. Solid arrows in panel **(B)** indicate warm currents and a dashed arrow indicates cold current. TSS, Tsushima Strait; BC, Bungo Channel; KC, Kii Channel; CI, Cape Inubozaki; TGS, Tsugaru Strait; NP, Noto Peninsula.

showed good agreement, indicating the effectiveness of the eDNA method for fish monitoring. These recent studies suggest that eDNA could be a powerful tool for monitoring the spatial distribution of eels across large landscapes.

Therefore, in this study, we used the eDNA method to detect the distribution of *A. japonica* in Japan. We collected water samples from rivers across the country and analyzed them for eDNA. As highlighted above, Japanese eels lay their eggs in the west of the Mariana Ridge, and larvae are transported to East Asia. The transport process and the location of the juveniles transported by the currents would be important in determining the distribution of eels in Japan. Therefore, numerical simulations were used to investigate the transport process of eel larvae to the Japanese coast. Survival after reaching the coast could be dependent on the environment of each river. Therefore, the data on various environmental factors in rivers were compiled using a Geographic Information System (GIS). The results of the eDNA analysis were compared with the results from the simulation, and modeled with the environmental factors to identify the factors determining the distribution of the Japanese eel.

MATERIALS AND METHODS

Study Areas

To analyze the eDNA of Japanese eels, samples were collected from the lower reaches of rivers located throughout Japan (**Figure 1**). Based on climatological and oceanographic insights, the studied area was divided into eight regions: the Nansei-Shoto Islands (NSI), west of Kyushu (WK), Pacific side of the main island of Japan (PMJ), the Seto Inland Sea (SIS), the western part of the Sea of Japan (WSJ), northern part of the Sea of Japan (NSJ), Pacific side of Tohoku (PT), and Hokkaido (HD) (**Figure 1A**). NSI is located in the most southern area, and the Kuroshio flows northeastward along these islands (**Figure 1B**). This area is the first location where eel larvae reach. However, the rivers in NSI are very short and small because of the narrow area and flat terrain. After passing through the NSI, the Kuroshio mainstream flows eastward in the south of Japan along the Kyushu, Shikoku, and western part of the Honshu area, and moves away from Japan at Cape Inubozaki. In WK, the Tsushima Current diverges from the Kuroshio and flows northward. The Tsushima Current enters the Sea of Japan *via* the shallow Tsushima Strait (∼120 m depth) and flows eastward along the western Honshu. In NSJ, the Tsushima Current shows unstable flows away from the coast. A weak branch of the Tsushima Current passes from the Sea of Japan to the Pacific Ocean through the Tsugaru Strait and turns south along the coast of the Pacific side of northern Honshu. However, PT is mainly influenced by the cold Oyashio Current. A warm water mass (warm core ring) sometimes separates from the Kuroshio and approaches PT. HD is in the northernmost part of the study area so that the Kuroshio and Tsushima Currents have little effect. It is located in the subarctic zone and water temperatures in the HD rivers are low and severe for eels, especially in winter. SIS is the largest estuary in Japan and connects the Pacific Ocean *via* the narrow Bungo Channel and Kii Channel. The seawater in SIS is originally from the Pacific Ocean, so that water temperature is moderate, but strongly affected by the terrestrial water.

It was expected that the eDNA of eels would be detected in high concentrations in the PMJ, which is directly under the effect of the Kuroshio Current, and where eel larvae are easily transported. After reaching the estuaries, water quality influences the survival and growth of eels, which are represented by the eDNA concentration in each river.

eDNA Field Sampling

A total of 365 study sites were selected from the downstream reaches of 265 rivers spatially spread across Japan (**Figure 1A**). In rivers with dams or weirs downstream, water was sampled from both the upper and lower sides of the dams or weirs. Most of the sites were in the freshwater area, but some were influenced by the tide. The sampling was conducted in the warm season when eels are active.

The researchers who collected the water samples wore new disposable rubber gloves at each sampling site. Samples for the eDNA analysis were directly collected by suctioning the surface water 2–10 times (50 mL at a time), for which a disposable plastic pump (TERUMO syringe pump, Terumo Corp., Tokyo, Japan) was used. A plastic bucket fastened to a rope was used for water collection at some sites that could not be accessed. The bucket and tip of the rope tied to the bucket were decontaminated by spraying with sodium hypochlorite. Then, the bucket and rope tip were washed twice with environmental water. Samples measuring 50–600 mL of water were collected depending on the turbidity and immediately filtered at the sampling site using a syringe and an enclosed filter cartridge with a nominal pore size of 0.45 μm (Sterivex, Merck Millipore, Burlington, MA). When we could not filter immediately after sampling, benzalkonium chloride was added to the sampled water to a final concentration of 0.1% to suppress the degradation of DNA (Yamanaka et al., 2017). The water samples were transported and filtered in the laboratory within 24 h. A 500 mL sample of pure water was filtered to monitor contamination at the filtration and subsequent extraction steps. Filters were preserved at −30°C until DNA extraction.

Water temperature and salinity were measured (Horiba Kyoto, Japan, or YSI Inc., Ohio, United States) before the sample were collected, and the area surrounding the sampling points were surveyed for information on the environment including the presence/absence of a weir on the upper or lower side of the sampling point, bottom material, vegetation, and revetment.

eDNA Extractions and qPCR Analyses

The eDNA extractions and qPCR analyses followed the Environmental DNA Sampling and Experiment Manual (Minamoto et al., 2020). Before the commencement of the experiments, all instruments and equipment were cleaned with DNA-Off (Takara Bio Inc., Shiga, Japan) to avoid DNA cross-contamination. All extractions and quantitative polymerase chain reaction (qPCR) setups were performed in a designated eDNA laboratory in a room free of PCR products (i.e., no PCR machines and no prior PCR amplification occurring in the room). eDNA was extracted from the filters using the modified Qiagen DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany) extraction method (Miya et al., 2016). Briefly, the RNAlater solution was first removed by passing the Milli-Q water through the filter membrane. Then, 200 mL lysis buffer (AL) and 20 μL proteinase K were added to the filter, and the filter was incubated with rotation (20 rpm) (Roller 6 digital, IKA, Staufen im Breisgau, Germany) at 56°C for 20 min. Subsequent extraction steps followed the standard Qiagen DNeasy extraction protocol. DNA was diluted using 100 μL elution buffer (AE), and all extracted DNA samples were preserved at −20°C before qPCR experiments. Buffer AL, AE, and proteinase K were provided by the extraction kit.

Real-time quantitative PCR (qPCR) was conducted with a total reaction volume of 15 μL, which included 900 nM of each primer, 125 nM of a probe, 0.075 μL of AmpErase^{TM} Uracil N-Glycosylase (Thermo Fisher Scientific, Waltham, MA, United States), and 2 μL of DNA template in 1x Environmental Master Mix 2.0 (Life Technologies, CA, United States) in a StepOne^{TM} real-time PCR system (Life Technologies). Species-specific primers and a probe designed by Kasai et al. (2020) were used as follows: forward primer Aja-Dlp-F2 5′-TACATTTAATGGAAAACAAGCATAAGCC-3′, reverse primer Aja-Dlp-R3 5′-CGTTAACATTACTCTGTCAACTTACCTG-3′, and a probe Aja-Dlp-Pr 5′-FAM-ACCCATAAACTGATAAATAG-NFQMGB-3′. The primer set was specifically designed to discriminate *A. japonica* among the *Anguilla* species potentially inhabiting the inland and coastal areas of Japan because some foreign species such as the European eel (*Anguilla anguilla*) was imported and cultured in Japan (Kishikawa, 1997) and have been found in natural waters as they escaped from cultured ponds (Zhang et al., 1999). In Kasai et al. (2020), all of the other species and subspecies in the genus *Anguilla* were taken into account to design the assay, and *in vitro* test were also conducted to verify the specificity using mitochondrial DNA samples of *A. anguilla, Anguilla bicolor bicolor, Anguilla marmorata*, and *Anguilla rostrata*. Thermal cycle settings for PCR were as follows: 2 min at 50°C, 10 min at 95°C, and 55 cycles of 15 s at 95°C, and 1 min at 60°C. PCR was performed in triplicate for each eDNA sample. The limit for quantification was defined as three copies per reaction, and any one of the detections out of the three replicates was considered a true detection (Takeuchi et al., 2019; Kasai et al., 2020). For the PCR negative control, ultrapure water was used instead of the DNA template. Each PCR plate included a fourfold dilution series of commercially synthesized cloned DNA, including the target sequence of *A. japonica* (30,000 copies/reaction to 30 copies/reaction) and PCR negative control including 2 μL of ultrapure water in place of the template DNA, in triplicate.

Numerical Model of Larval Transport

Larvae of *A. japonica* are transported to Japanese coastal waters by the Kuroshio Current (**Figure 1B**, Shinoda et al., 2011). Therefore, the distribution of larvae in terrestrial waters would depend on how they are transported by ocean currents and reach each river estuary. The swimming speed of larvae is slow compared to the velocity of the surrounding currents; thus, they are assumed to be particles floating in the water. If the details of the currents are recognized

around Japan, it is possible to simulate how many larvae will reach each river.

Recently, the Japan Meteorological Agency developed a fine-scale ocean model (JPN model) covering the coastal seas of East Asia including Japan (117°E–160°E, 20°N–50°N, Sakamoto et al., 2019). The model can realistically reproduce the distributions and seasonal variations of water temperature, salinity, and sea level. The horizontal resolution of the JPN model is 1/33° in the zonal direction and 1/50° in the meridional direction, corresponding to approximately 2 km, which is sufficient for the reproduction of currents in the coastal regions. Small-scale features in the coastal seas, such as fronts and mesoscale eddies, were also well simulated with high resolution. In this study, therefore, larval transport was simulated using the daily average flow reproduced by the JPN model to investigate the recruitment success of Japanese eel.

As an initial condition, 80,000 particles that were regarded as larvae were released in the east of Taiwan because all the larvae transported to the waters around Japan originate from the Kuroshio Current (Hsiung and Kimura, 2019, **Figure 1B**). It is known that *A. japonica* larvae show diel vertical migration, remaining in the upper surface waters at night and diving to deeper waters during the day (Otake et al., 1998). Therefore, the particles were set to have diurnal vertical movement from the surface to a depth of 100 m (deeper in the daytime and shallower at nighttime) in the model. The particles were transported by the currents reproduced by the JPN model, and their positions were calculated by the Runge–Kutta fourth-order method. The simulation was conducted from 1 December to 30 April, based on the information of larval catch in the Japanese rivers (Yoneta et al., 2019). The recruitment success was evaluated by the cumulative number of particles reaching the shore. Particles were considered to have reached the shore if they were in the grids within a 2 km (horizontal resolution of the JPN model) distance from the shore. The simulation was repeated nine times using the velocity fields from 2008 to 2017. The results of the simulation were compared with those from the eDNA detection.

Collection of Geographical, Water Quality, and Demographic Data

To investigate the relationship between the river environment and eDNA concentration, which is supposed to represent the fish biomass, we used water quality data from the Water Quality Survey of Public Water Areas dataset of the Ministry of the Environment of Japan. The time-averaged data in 2016 were collected from water quality monitoring sites throughout Japan for dissolved oxygen (DO), chemical oxygen demand (COD), suspended solids (SS), total nitrogen (TN), minimum pH, and maximum pH. Basin mesh data (ver. 2.1) were collected from the National Land Numerical Information database of the Ministry of Land, Infrastructure, Transport, and Tourism of Japan. Water quality data within 1 km of the eDNA sampling site were selected and used for statistical processing. The function "it falls inside" in ArcMap 10.5.1 (ESRI, Redlands, CA, United States) was used to link 123 points.

The distance of each eDNA sampling point from the river mouth and the width of the river at the sampling point were measured using Google Earth Pro (ver. 7.3.2.5776, Map data: Google, Digital Globe). Data on the number of farmed eel that were released were taken from the Census of Fisheries (2003, 2008, and 2013).

Statistical Analyses

The Tukey HSD test was used to compare the differences in eDNA concentrations among the regions.

Generalized linear models (GLMs) with a Gaussian distribution were developed to identify surrounding environmental factors influencing eDNA concentration. Sites with positive DNA concentrations were employed in the model to assess the effect of the environment on the survival and growth of larvae in each river after they were transported. Previous comparisons between visual observation of fish and eDNA concentration in rivers suggest that eDNA quantification can detect the order of magnitude variations in fish abundance (Doi et al., 2017; Maruyama et al., 2018). Therefore, the logarithmically converted eDNA concentration (>0) was used as the dependent variable. Prior to modeling, the collinearity of all explanatory variables (distance from the river mouth, river width, water temperature, TN, DO, COD, SS, minimum pH, and maximum pH) was evaluated using correlation values and variance inflation factors (VIFs). Since the explanatory variables used in the model need to be independent of each other, VIFs were used to examine the multicollinearity among the explanatory variables and to remove variables that are highly correlated. As all values of the VIFs were below 5, non-collinearity was assumed in this study (Zuur et al., 2009; **Table 1**).

The best model was selected using Akaike's information criteria (AIC). The best model for any candidate set applied to a given dataset was that with the lowest AIC value, and models with $\Delta AIC < 2$ were assumed to be reasonable alternatives to the best model, and thus were retained (Burnham and Anderson, 2010). All statistical analyses, GLM analyses, and graphics employed R, version 4.0.2 (R Core Team, 2020), and the lme4 and MuMIn packages (Bates et al., 2015; Bartoń, 2019).

RESULTS

Environmental DNA Concentration

Successful amplification of eDNA was quantified at 181 sampling sites, accounting for 49.6% of the total sites that were sampled

TABLE 1 | Summary of the variance inflation factor (VIF) analyses.

Distance from river mouth	River width	Water temperature	DO	TN	COD	SS	Min pH	Max pH
1.175	1.268	1.372	2.229	2.239	2.487	1.451	1.388	1.986

DO, dissolved oxygen; COD, chemical oxygen demand; SS, suspended solids; TN, total nitrogen; Min Ph, minimum pH; Max pH, maximum pH.

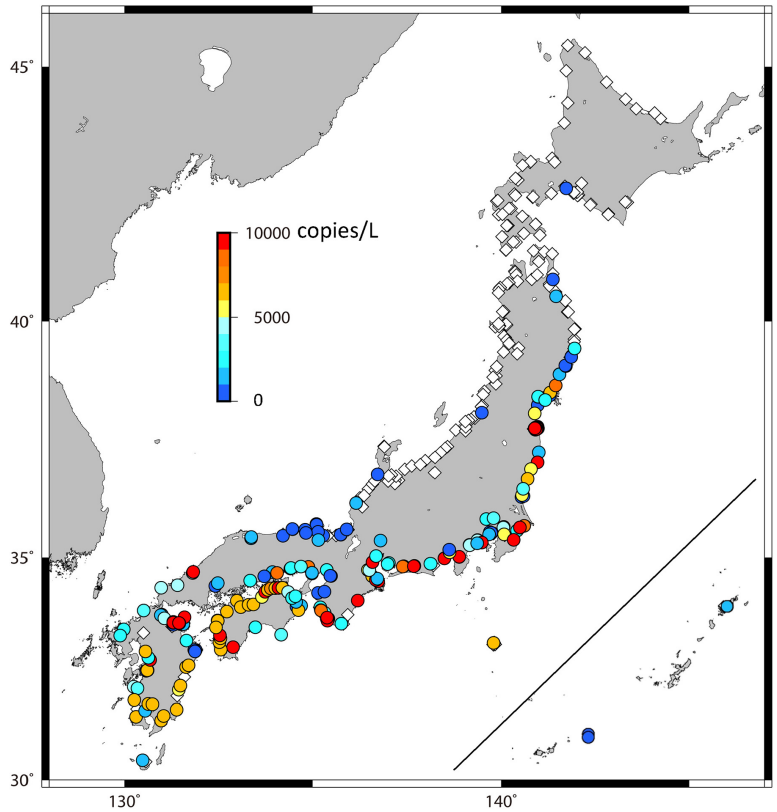


FIGURE 2 | Concentration of environmental DNA at each observation point. Warmer and colder colors represent larger and smaller values, respectively, and dark red points represent more than 10,000 copies/L. Open diamonds indicate zero.

(**Figure 2**). In the PCR analysis, no positives were detected in the negative controls. The eDNA of eels was detected in some sites in the NSI and WSJ, although the concentrations were not high (**Figure 3**). In WK, PMJ, and SIS, the ratio of detected sites was over 80%, and the eDNA concentrations were high, although the variations were large with no detection in some rivers. The three largest values (44,356.6 copies/L from the Asaragi River, 41,177.0 copies/L, and 38,407.1 copies/L from the Aku River) recorded across all sampling sites were detected in the PMJ. The southern part of PT showed large values, while the detected values were small or zero in the northern part. The eDNA was rarely detected in NSJ and HD. In these regions, the proportion of sites where eDNA was detected was small, and even when detected, their concentrations were very low.

The average eDNA concentration was significantly different among regions ($p < 0.001$, **Table 2**). The concentrations in HD and NSJ were significantly lower, while those in SIS and PMJ were significantly higher than in the other areas.

Numerical Simulation of Larval Transport

Figure 4 shows the results of the simulation of larval transport; the number of particles that reached each coastal area by April 30. The simulation results are generally well consistent with the eDNA results (**Figure 2**). A large number of particles reached the coasts in NSI, PMJ, WK, and WSJ every year, while low in SIS. Particles reached the coastal areas of PT in half of the years we examined. Even in those cases, they only reached the south of PT and were rare in the north. Further, particles were rarely transported beyond the Noto Peninsula, and they never reached HD. Only a few particles reached the east of the Noto

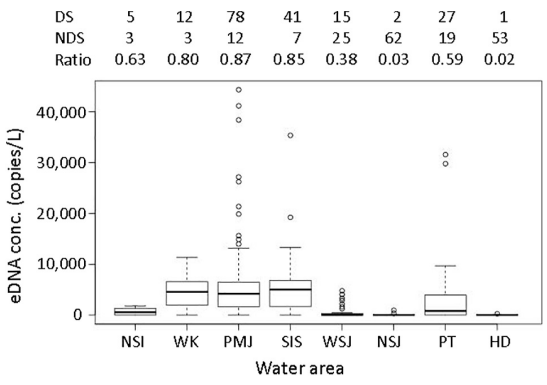


FIGURE 3 | Concentrations of environmental DNA in each region. In the box plots, the midline represents the median, the upper and lower limits of the box represent the third and first quartiles, respectively, the whiskers represent 1.5 times the interquartile range from the top (bottom) of the box to the furthest datum within that distance, and the circles represent the individual data beyond that distance. The numbers given above the graph are the number of sites where eDNA was detected (DS), was not detected (NDS), and the ratio of detected sites to the total number of sites that were sampled in each region.

TABLE 2 | Results of the Tukey HSD test comparing the average eDNA concentrations between the regions.

Region	NSI	WK	PMJ	SIS	WSJ	NSJ	PT
WK	0.786						
PMJ	0.111	0.882					
SIS	0.314	0.996	0.991				
WSJ	1.000	0.260	**0.000**	**0.001**			
NSJ	1.000	0.080	**0.000**	**0.000**	1.000		
PT	0.851	1.000	0.170	0.780	0.156	**0.016**	
HD	1.000	0.091	**0.000**	**0.000**	1.000	1.000	**0.023**

The regions include Nansei-Shoto Islands (NSI), west of Kyushu (WK), Pacific side of the main island of Japan (PMJ), the Seto Inland Sea (SIS), the western part of the Sea of Japan (WSJ), northern part of the Sea of Japan (NSJ), Pacific side of Tohoku (PT), and Hokkaido (HD). Significant differences (<0.05) are given in bold font style.

Peninsula in the year 2011–2012 and the southern tip of HD in the year 2015–2016.

The number of particles reaching an area was compared with whether eDNA was detected. GLM analysis was performed using a binomial distribution with eDNA detection as the dependent variable and the average number of particles reaching the area as the explanatory variable (**Figure 5**). The binomial test showed that eDNA was detected in the rivers where the number of particles reaching the area was high ($p < 0.01$).

Important Environmental Factor for Eel Survival and Growth

To determine the river environments that are most suitable for eels, we modeled the detected eDNA concentrations using environmental factors as explanatory variables. Among the GLMs used to examine eDNA concentration, seven models with $\Delta AIC < 2$ were retained (**Table 3**). Four common explanatory variables (distance from the river mouth, TN, DO, and the minimum pH) were selected in these models. The relationship of the eDNA concentration to the distance from the river mouth was negative, while that with TN, DO, and the minimum pH was positive. Only TN concentration was significantly correlated with eDNA concentration.

Effect of Released Eels on eDNA Detection

Young eels are released by prefectural Inland Fisheries Cooperatives throughout the country, other than in HD, to allow their growth and to increase their population sizes in the natural freshwater conditions. This indicates that individuals that were grown in eel farms and artificially released and those that had migrated to the natural environment as glass eels were mixed in the natural environment. Because eels form a single spawning population throughout their distribution area, it is impossible to genetically discriminate between naturally recruited and released individuals.

Figure 6 shows the number of farmed eel individuals released in each prefecture. Large numbers were not released in any of the regions, but were released evenly throughout the country except for HD. There was no correlation between the number of released eels and eDNA (**Figure 2**, $p > 0.05$). Thus, although some of the results of the eDNA distribution in this study may

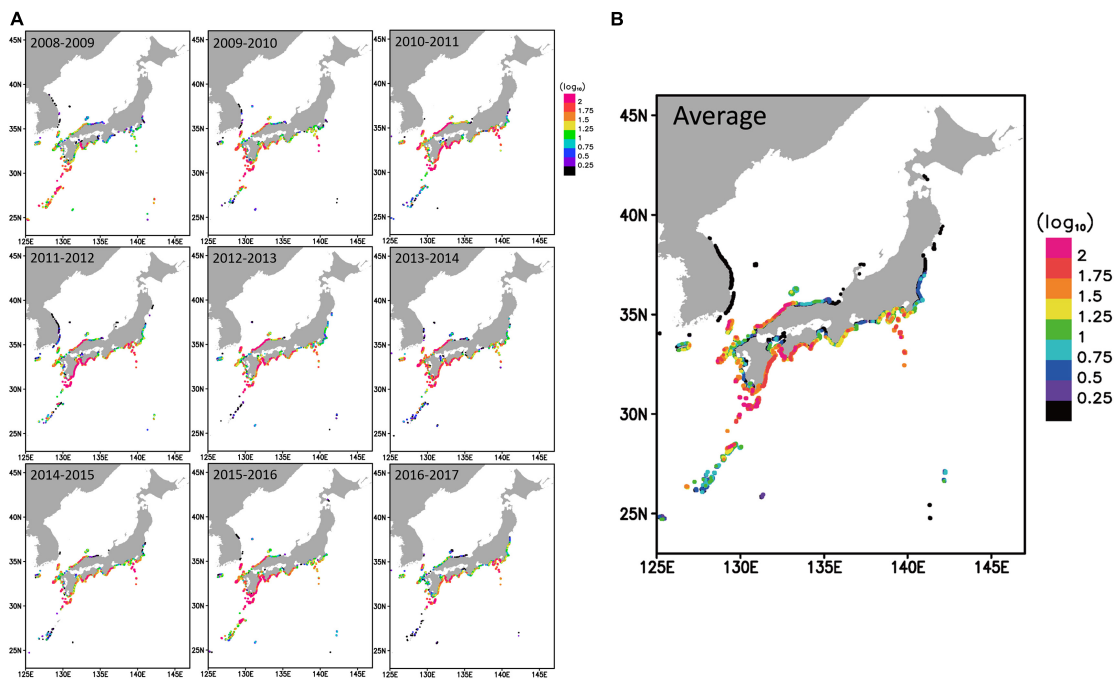


FIGURE 4 | The number of particles that reached each coastal area by April 30 in **(A)** each year and **(B)** 9 years average for the period 2008–2017. The number of particles is expressed in color on a log scale. Warm colors indicate more, and cooler colors indicate fewer particles.

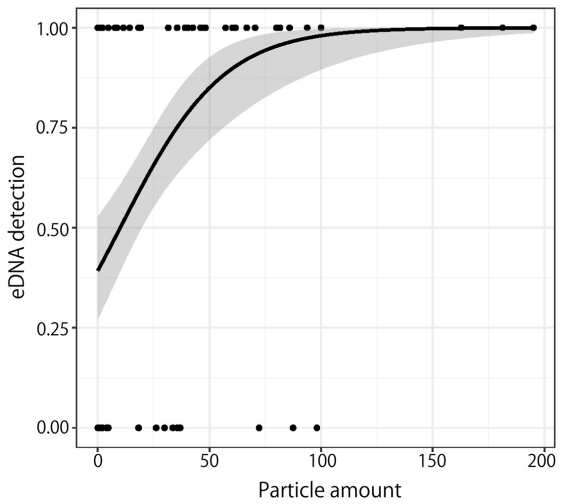


FIGURE 5 | Relationship between the number of particles that reached near the rivers and eDNA detection. The solid line and gray area indicate the best regression curve and the 95% confidence interval, respectively.

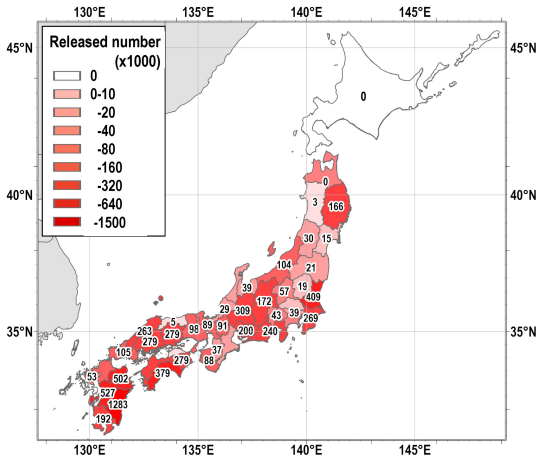


FIGURE 6 | The total number of farmed eels released in 2003, 2008, and 2013 in each prefecture by prefectural Inland Fisheries Cooperatives (Data from the Fishery Census 2003, 2008, and 2013).

TABLE 3 | Summary of the GLMs with $\Delta AIC < 2$.

(a)

Rank	Distance from river mouth	TN	DO	Min pH	AIC	ΔAIC
1		0.110	0.086		90.4	0.00
2		0.082			90.5	0.17
3	−0.013	0.112	0.096		90.9	0.55
4	−0.016	0.115	0.106	0.293	91.3	0.92
5		0.111	0.092	0.226	91.4	1.01
6	−0.010	0.081			91.6	1.23
7		0.081		0.188	91.9	1.50

(b)

	Estimate	Standard error	*t* value	*p* value
(Intercept)	0.255	1.904	0.134	0.894
Distance from river mouth	−0.016	0.011	−1.398	0.167
TN	0.115	0.043	2.675	0.009**
DO	0.106	0.060	1.759	0.084
Minimum PH	0.293	0.237	1.235	0.222

*(a) Selected parameters and AIC values in GLM, and (b) coefficient values and associated probability of the best model. ** < 0.01. TN, total nitrogen; DO, dissolved oxygen; Min pH, Minimum pH; AIC, Akaike's Information Criteria.*

reflect farmed eels, it would broadly reflect the distribution of natural individuals.

DISCUSSION

This is the first study to investigate the distribution of Japanese eels in detail on a national scale, and this could not be done without using the eDNA method. The eDNA of Japanese eels was detected in an unexpectedly wide area (**Figure 2**), even though *A. japonica* is listed as an endangered species by both the Ministry of the Environment and the IUCN. The detection rates were as high as over 80% in PMJ, WK, and SIS (**Figure 3**). This result coincides with the distribution of larval reaching rates estimated by the numerical simulation (**Figures 4**, **5**, $p < 0.01$). This indicates that the major distribution of *A. japonica* in Japanese rivers is determined by the maritime transport of larvae. Although a quantitative relationship between eDNA concentrations and biomass has not been assured, it can be assumed that the eDNA concentrations are reflective of the biomass (Maruyama et al., 2018; Itakura et al., 2019). Generally, our results revealed that eels in the river are distributed south of Noto Peninsula, in the Sea of Japan, and south of Miyako on the Pacific side. This is roughly consistent with the results compiled and reviewed by Yoneta et al. (2019), although our results show a much more detailed distribution. These areas are presumed to be continuously populated by naturally recruited individuals based on the results of the larval transport simulation. This is also consistent with recent reports on the harvests of glass eels, the collection of Japanese eels in river systems where eels were not released, and otolith stable isotope ratio-based origin determinations (Yoneta et al., 2019). The small amounts of eDNA detected north of PT (**Figure 2**) possibly reflect the artificially released individuals, because only one of the 19 individuals was a natural individual from a lake in the north of PT (Yoneta et al., 2019).

The distribution of eDNA showed a clear difference between the Sea of Japan side and the Pacific side of the Tohoku region (**Figure 2**). eDNA was detected in the southern region of PT rivers, while only local and very little eDNA was detected in the NSJ. This difference reflects the transport process of *A. japonica* to Japanese coastal waters from the south. Although some previous studies suggest the existence of natural individuals in PT (Kume et al., 2020), the transport process of larvae to PT rivers is unknown. In other words, it has been unclear whether eel larvae are transported to the south of PT by the Tsushima Current from

the Sea of Japan through the Tsugaru Strait or migrate northward from the Kuroshio and/or Kuroshio Extension (**Figure 1B**). To address this problem, our eDNA distribution (**Figure 2**) and numerical experiments (**Figure 4**) indicate that the latter process is plausible. Generally, the Kuroshio flows eastward off Cape Inubozaki, but a part of it often moves northward as a branch along the Pacific coast of Japan (**Figure 1B**). It is also well known that the Kuroshio meanders east of the Japanese Islands. As it meanders and is unstable, a part of the Kuroshio sometimes cuts off the mainstream and moves northward as a mesoscale eddy (Yasuda et al., 1992). Eel larvae would use the branch or warm core rings to migrate northward and reach the south of the PT coast.

From the perspective of the larval transport in the ocean, the occurrence of these eels should have been more in the NSI, but this was not reflected in our eDNA survey. There are two possible reasons for this discrepancy. The first reason is the shortness of rivers in the NSI. There are no large rivers, and the river discharges are small in small islands. Glass eels approach estuaries and rivers by smelling the odor originating from rivers (Fukuda et al., 2019). Therefore, around NSI, where there are no large rivers, eel larvae cannot smell the river and cannot approach the mouth of the river. The second reason is habitat segregation. Kano et al. (2017) surveyed the occurrence of *A. japonica* and the Indo-Pacific eel (*A. marmorata*) in many streams in Kyushu and NSI. They found a considerably higher abundance of *A. marmorata* than *A. japonica* in NSI, while no *A. Marmorata* was found in Kyushu. *A. marmorata* is widely distributed in tropical and subtropical terrestrial waters (Watanabe et al., 2008). *A. japonica* may be outcompeted by *A. marmorata*, which is larger than *A. japonica*, and their abundance is hence likely to be low in NSI. For these reasons, Japanese eels may not be well adapted to living on small islands where a tropical anguillid eel resides.

Based on our simulation, only a few particles reached the SIS; however, this is not the true picture as shown by the eDNA detected. This divergence is likely due to the characteristics of the model. As shown in **Figure 1B**, the two entrances of SIS, the Bungo Channel and the Kii Channel are very narrow with complicated topographies, making the friction and viscosity large in the model. In addition, it is difficult to reproduce the sufficient water exchange through these channels, because the minimum scale of the phenomena that can be reproduced by the model is about 5–10 times the grid size of the model (Lévy et al., 2012). Therefore, it is difficult for particles to pass through the channels and enter the SIS. Another improvement is the preference for low salinity areas. Eel larvae detect odor and migrate to estuaries. Including the effect of low salinity preference into the model may help to reproduce a more realistic distribution of eDNA.

The distance from the river mouth, TN, DO, and the minimum pH were selected in the GLMs. Considering the fact that the larvae migrate from the sea to the river, it is a natural consequence of the negative correlation between the distance from the river mouth and the eDNA concentration. Since oxygen is necessary for all animals to sustain life, it is not surprising that there was a positive correlation between DO and eDNA concentrations. Rather, it is worth noting that TN was a significantly important factor for determining the concentration of eDNA (**Table 3**). This indicates that the survival and growth of eels are better in nutritious rivers in Japan. Japanese rivers were highly eutrophicated in the 1970s and the 1980s, when there was a large growth in the Japanese economy. Red tides and consequent hypoxia often occurred, and the biodiversity in the rivers was reduced during this high economic growth period. Since the 1990s, however, sewage treatment technology has been developed, and the inputs of nutrients and organic matter from households and industries have been reduced to protect the environments of rivers, estuaries, and coastal areas. This policy has worked, and the river water condition has drastically recovered (Ye and Kameyama, 2020), and many species are returning to the Japanese rivers. The average TN value in this study was 1.2 mg/L. Considering that TN concentrations in formerly eutrophic rivers exceeded 5 mg/L in the 1980s (Ye and Kameyama, 2020), this value is very low. Therefore, the current high TN value does not necessarily mean that the river is highly eutrophicated. The high abundance of eels in TN-rich rivers is likely to be related to the abundance of food.

Unlike conventional methods that capture fish, the eDNA method requires less labor and a shorter time and reduces on-site survey costs. It enables surveys to be conducted over a wider area at multiple locations. In addition, unlike conventional surveys, the eDNA survey does not use fishing gears, and no special techniques are required to collect samples and only a small amount of water needs to be collected. Thus, these surveys can be conducted easily. This makes it possible to conduct unprecedented multisite surveys, similar to that undertaken in this study. Such a multisite survey in a short period would not be possible using conventional methods. Furthermore, this method does not kill or injure any organisms by capturing the target species and requires little access to their habitats. It is especially beneficial for endangered and rare species such as the Japanese eels.

One drawback of the eDNA method is that the demography of the population cannot be determined. The length and weight of the individuals of the target species also cannot be estimated. In the case of eels, we also cannot determine whether the individuals originated from farms or from the wild. Further, the eDNA also flows down from the upstream side of the river, and hence may not reflect the biomass at the site, but the sum of the transported and on-site DNA (Jerde et al., 2011). Therefore, if necessary, a combination of eDNA with conventional capture surveys might be helpful to obtain more detailed information that is needed for conservation planning and to take prompt and effective measures.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

Ethical review and approval was not required for the animal study because this research does not treat any vertebrate directly.

AUTHOR CONTRIBUTIONS

AK delineated the study. AK, AY, HA, SaK, RM, NA, ShK, YK, and YY collected the samples. AY, HY, and HA analyzed the DNA. TK conducted the numerical simulation. AK, AY, and SaK analyzed the data. AK and AY wrote the manuscript. All authors checked the manuscript and agreed to be accountable for the content of the work.

FUNDING

This work was supported by JSPS KAKENHI, Grant-in-Aid for Scientific Research (A), Grant Number17H01412.

ACKNOWLEDGMENTS

The data on currents, estimated by the JPN model, were kindly provided by the Meteorological Research Institute, Japan. We would like to thank Dr. Manabu Kume, Mr. Yuki Terashima, Mr. Yosuke Sadayuki, Mr. Shingo Takada, and Mr. Hideto Nakanishi for their supporting sampling. We also would like to thank Editage (www.editage.com) for English language editing.

REFERENCES

Aoyama, J., Shinoda, A., Sasai, S., Miller, M. J., and Tsukamoto, K. (2005). First observations of the burrows of *Anguilla japonica*. *J. Fish Biol.* 67, 1534–1543. doi: 10.1111/j.1095-8649.2005.00860.x

Bartoń, K. (2019). *MuMIn:Multi-Model Inference. R package version 1.43.6*. Available online at: https://CRAN.R-project.org/package=MuMIn (accessed 30 September, 2020).

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Burnham, K. P., and Anderson, D. A. (2010). *Model Selection and Multimodel Inference -A Practical Information-Theoretic Approach*, 2nd Edn. New York, NY: Springer.

Deiner, K., Walser, J. C., Mächler, E., and Altermatt, F. (2015). Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biol. Conserv.* 183, 53–63. doi: 10.1016/j.biocon.2014.11.018

Doi, H., Inui, R., Akamatsu, Y., Kanno, K., Yamanaka, H., Takahara, T., et al. (2017). Environmental DNA analysis for estimating the abundanceand biomass of stream fish. *Freshw. Biol.* 62, 30–39. doi: 10.1111/fwb.12846

Fujita, T. (2018). "The national census on river environments," in *The Encyclopedia of Ichthyology*, ed. The ichthyological Society of Japan (Tokyo: Maruzen), 544–545.

Fukuda, N., Yokouchi, K., Yamamoto, T., Kurogi, H., and Yada, T. (2019). Salinity and odor preferences of Japanese eel during the first year of post-recruitment growth in saline water. *J. Ethol.* 37, 93–99. doi: 10.1007/s10164-018-0572-x

Han, Y. S., Zhang, H., Tseng, Y. H., and Shen, M. L. (2012). Larval Japanese eel (*Anguilla japonica*) as sub surface current bio tracers on the east Asia continental shelf. *Fish. Oceanogr.* 21, 281–290. doi: 10.1111/j.1365-2419.2012.00624.x

Hsiung, K. M., and Kimura, S. (2019). Impacts of global warming on larval and juvenile transport of Japanese eels (*Anguilla japonica*). *Deep Sea Res. Part II* 169–170:104685. doi: 10.1016/j.dsr2.2019.104685

Hsiung, K. M., Kimura, S., Han, Y. S., Takeshige, A., and Iizuka, Y. (2018). Effect of ENSO events on larval and juvenile duration and transport of Japanese eel (*Anguilla japonica*). *PLoS One* 13:e0195544. doi: 10.1371/journal.pone.0195544

Itakura, H., Wakiya, R., Yamamoto, S., Kaifu, K., Sato, T., and Minamoto, T. (2019). Environmental DNA analysis reveals the spatial distribution, abundance, and biomass of Japanese eels at the river-basin scale. *Aquatic Conserv. Mar. Freshw. Ecosyst.* 29, 361–373. doi: 10.1002/aqc.3058

Jacoby, D., and Gollock, M. (2014). *Anguilla Japonica*. The IUCN Red List of Threatened Species 2014: e.T166184A1117791.

Jerde, C. L., Mahon, A., Chadderton, W. L., and Lodge, D. M. (2011). Sight-unseen" detection of rare aquatic species usingenvironmental DNA. *Conserv. Lett.* 4, 150–157. doi: 10.1111/j.1755-263X.2010.00158.x

Kano, Y., Yamashita, T., Tanaka, W., Koyama, A., and Kanno, K. (2017). Distribution of *Anguilla japonica* and *A. marmorata* in the Nansei Islands, Japan,and their habitat segregation based on local names. Japan. *J. Ichthyol.* 64, 43–53.

Kasai, A., Takada, S., Yamazaki, A., Masuda, R., and Yamanaka, H. (2020). The effect of temperature on environmental DNA degradation of Japanese eel. *Fish. Sci.* 86, 465–471. doi: 10.1007/s12562-020-01409-1

Kimura, S., Inoue, T., and Sugimoto, T. (2001). Fluctuation in the distribution of low salinity water in the north equatorial current and its effect on the larval transport of the Japanese eel. *Fish. Oceanogr.* 10, 51–60. doi: 10.1046/j.1365-2419.2001.00159.x

Kishikawa, T. (1997). The poor catch of glass eels and survival strategy of the eel farm. *Yoshoku* 34, 52–68.

Kitagawa, T., Muraoka, K., Yamada, T., and Nakamura, K. (2020). Analysis for trial of environmental DNA metabarcoding to fish survey in the national census on river environments. *Adv.River Eng.* 26, 319–324.

Kume, M., Terashima, Y., Kawai, F., Kutzer, A., Wada, T., and Yamashita, Y. (2020). Size-dependent changes in habitat use of Japanese eel *Anguilla japonica* during the river life stage. *Environ. Biol. Fish.* 103, 269–281. doi: 10.1007/s10641-020-00957-w

Kurogi, H., Okazaki, M., Mochioka, N., Jinbo, T., Hashimoto, H., Takahashi, M., et al. (2011). First capture of post-spawning female of the Japanese eel *Anguilla japonica* at the southern West Mariana Ridge. *Fish. Sci.* 77, 199–205. doi: 10.1007/s12562-010-0318-3

Lévy, M., Resplandy, L., Kleinc, P., Capet, X., Iovino, D., and Ethéa, C. (2012). Grid degradation of submesoscale resolving ocean models: benefits for offline passive tracer transport. *Ocean Model.* 48, 1–9. doi: 10.1016/j.ocemod.2012.02.004

Maruyama, A., Sugatani, K., Watanabe, K., Yamanaka, H., and Imamura, A. (2018). Environmental DNA analysis as a non-invasive quantitative tool for reproductive migration of a threatened endemic fish in rivers. *Ecol. Evol.* 8, 11964–11974. doi: 10.1002/ece3.4653

Minamoto, T., Miya, M., Sado, T., Seino, S., Doi, H., Kondoh, M., et al. (2020). An illustrated manual for environmental DNA research: water sampling guidelines and experimental protocols. *Environ. DNA* 3, 8–13. doi: 10.1002/edn3.121

Minamoto, T., Yamanaka, H., Takahara, T., Honjo, M., and Kawabata, Z. (2012). Surveillance of fish species composition using environmental DNA. *Limnology* 13, 193–197. doi: 10.1007/s10201-011-0362-4

Ministry of the Environment (2013). Available online at: https://www.env.go.jp/press/15619.html (accessed September 16, 2020).

Miya, M., Gotoh, R., and Sado, T. (2020). MiFish metabarcoding: a high-throughput approach for simultaneous detection of multiple fish species from environmental DNA and other samples. *Fish. Sci.* 86, 939–970. doi: 10.1007/s12562-020-01461-x

Miya, M., Minamoto, T., Yamanaka, H., Oka, S.i., Sato, K., Yamamoto, S., et al. (2016). Use of a filter cartridge for filtration of water samples and extraction of environmental DNA. *J. Vis. Exp.* 25:54741. doi: 10.3791/54741

Nakamura, K., Hattori, A., Fukuhama, M., and Kayaba, Y. (2015). Issues and directions for promoting environmental management of rivers. *Adv. River Eng.* 21, 31–36.

Otake, T., Amano, Y., Shirai, K., Mochioka, N., Takahashi, T., Chow, S., et al. (2019). Evaluation from otolith Sr stable isotope ratios of possible juvenile growth areas of Japanese eels collected from the West Mariana Ridge spawning area. *Fish. Sci.* 85, 483–493. doi: 10.1007/s12562-019-01304-4

Otake, T., Inagaki, T., Hasumoto, H., Mochioka, N., and Tsukamoto, K. (1998). Diel vertical distribution of *Anguilla japonica* leptocephali. *Ichthyol. Res.* 45, 208–211. doi: 10.1007/bf02678565

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Reid, S. M. (2011). Comparison of point and transect-based electrofishing to sampleAmerican eel (*Anguilla rostrata*) in wadeable riverine habitats. *Aquat. Living Resour.* 24, 79–83. doi: 10.1051/alr/2011106

Research Institute of Environment, Agriculture and Fisheries, Osaka Prefecture (2014). *Yodogawa Survey Results Report 2014*. Available online at: http://www.kannousuiken-osaka.or.jp/kankyo/gijutsu/seika/yod/7-2014.html (accessed September 16, 2020).

Sakamoto, K., Tsujino, H., Nakano, H., Urakawa, S., Toyoda, T., Hirose, N., et al. (2019). Development of a 2-km resolution ocean model coveringthe coastal seas around Japan for operational application. *Ocean Dyn.* 69, 1181–1202. doi: 10.1007/s10236-019-01291-1

Shinoda, A., Aoyama, J., Miller, M. J., Otake, T., Mochioka, N., Watanabe, S., et al. (2011). Evaluation of thelarval distribution and migration of the Japanese eel in the western north Pacific. *Rev. Fish Biol. Fish.* 21, 591–611. doi: 10.1007/s11160-010-9195-1

Takeuchi, A., Watanabe, S., Yamamoto, S., Miller, M. J., Fukuba, T., Miwa, T., et al. (2019). First use of oceanic environmental DNA to study the spawning ecology of the Japanese eel *Anguilla japonica*. *Mar. Ecol. Prog. Ser.* 609, 187–196. doi: 10.3354/meps12828

Tatsukawa, K. (2003). "Eel resources in East Asia," in *Eel Biology*, eds K. Aida, K. Tsukamoto, and K. Yamauchi (Tokyo: Springer), 293–298. doi: 10.1007/978-4-431-65907-5_20

Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., et al. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Mol. Ecol.* 21, 2565–2573. doi: 10.1111/j.1365-294x.2011.05418.x

Tsukamoto, K. (2006). Oceanic biology: spawning of eels near a seamount. *Nature* 439:929. doi: 10.1038/439929a

Tsukamoto, K., and Arai, T. (2001). Facultative catadromy of the eel *Anguilla japonica* between freshwater and seawater habitats. *Mar. Ecol. Prog. Ser.* 220, 265–276. doi: 10.3354/meps220265

Turner, C. R., Uy, K. L., and Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biol. Conserv.* 183, 93–102. doi: 10.1016/j.biocon.2014.11.017

Watanabe, S., Aoyama, J., Miller, M. J., Ishikawa, S., Feunteun, E., and Tsukamoto, K. (2008). Evidence ofpopulation structure in the giant mottled eel, *Anguilla marmorata*, using total number of vertebrae. *Copeia* 3, 680–688. doi: 10.1643/cg-07-080

Yamanaka, H., and Minamoto, T. (2016). The use of environmental DNA of fishes as an efficient method of determining habitat connectivity. *Ecol. Indic.* 62, 147–153. doi: 10.1016/j.ecolind.2015.11.022

Yamanaka, H., Minamoto, T., Matsuura, J., Sakurai, S., Tsuji, S., Motozawa, H., et al. (2017). A simple method for preserving environmentalDNA in water samples at ambient temperature by addition of cationicsurfactant. *Limnology* 18, 233–241. doi: 10.1007/s10201-016-0508-5

Yasuda, I., Okuda, K., and Hirai, M. (1992). Evolution of a Kuroshio warm-core ring -variability of the hydrographic structure. *Deep Sea Res. Part A* 39, S131–S161.

Ye, F., and Kameyama, S. (2020). Long-term spatiotemporal changes of 15 water-quality parameters in Japan: an exploratory analysis of countrywide data during1982-2016. *Chemosphere* 242:125245. doi: 10.1016/j.chemosphere.2019.125245

Yoneta, A., Itakura, H., Arai, K., Kaifu, K., Yoshinaga, T., Miyake, Y., et al. (2019). Distribution of naturally recruited wild Japanese eels in Japan revealed by otolith stable isotopic ratios and document investigation. *Nippon Suisan Gakkaishi* 85, 150–161. doi: 10.2331/suisan.18-00038

Zenimoto, K., Kitagawa, T., Miyazaki, S., Sasai, Y., Sasaki, H., and Kimura, S. (2009). The effects of seasonal and interannual variability of oceanic structure in the western Pacific North Equatorial Current on larval transport of the Japanese eel (*Anguilla japonica*). *J. Fish Biol.* 74, 1878–1890. doi: 10.1111/j.1095-8649.2009.02295.x

Zhang, H., Mikawa, N., Yamada, Y., Horie, N., Okamura, A., Utoh, T., et al. (1999). Foreign eel species in the natural waters of Japan detected by polymerase chain reaction of mitochondrial cytochrome b region. *Fish. Sci.* 65, 684–686. doi: 10.2331/fishsci.65.684

Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A., and Smith, G. M. (2009). *Mixed Effects Models and Extensions in Ecology With R*. New York, NY: Springer.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 26 February 2021
doi: 10.3389/fevo.2021.623718

The Effect of Activity, Energy Use, and Species Identity on Environmental DNA Shedding of Freshwater Fish

Bettina Thalinger[1,2,3], Andreas Rieder[1], Anna Teuffenbach[1], Yannick Pütz[1], Thorsten Schwerte[1], Josef Wanzenböck[4] and Michael Traugott[1,5]*

[1] Department of Zoology, University of Innsbruck, Innsbruck, Austria, [2] Centre for Biodiversity Genomics, University of Guelph, Guelph, ON, Canada, [3] Department of Integrative Biology, College of Biological Sciences, University of Guelph, Guelph, ON, Canada, [4] Research Department for Limnology, Mondsee, University of Innsbruck, Mondsee, Austria, [5] Sinsoma GmbH, Voels, Austria

The quantitative measurement of environmental DNA (eDNA) from field-collected water samples is gaining importance for the monitoring of fish communities and populations. The interpretation of these signal strengths depends, among other factors, on the amount of target eDNA shed into the water. However, shedding rates are presumably associated with species-specific traits such as physiology and behavior. Although such differences between juvenile and adult fish have been previously detected, the general impact of movement and energy use in a resting state on eDNA release into the surrounding water remains hardly addressed. In an aquarium experiment, we compared eDNA shedding between seven fish species occurring in European freshwaters. The investigated salmonids, cyprinids, and sculpin exhibit distinct adaptions to microhabitats, diets, and either solitary or schooling behavior. The fish were housed in aquaria with constant water flow and their activity was measured by snapshots taken every 30 s. Water samples for eDNA analysis were taken every 3 h and energy use was determined in an intermittent flow respirometer. After controlling for the effect of fish mass, our results demonstrate a positive correlation between target eDNA quantities as measured with digital PCR, fish activity, and energy use, as well as species-specific differences. For cyprinids, the model based on data from individual fish was only partly transferable to groups, which showed lower activity and higher energy use. Our findings highlight the importance of fish physiology and behavior for the comparative interpretation of taxon-specific eDNA quantities. Species traits should therefore be incorporated into eDNA-based monitoring and conservation efforts.

Keywords: digital PCR, video-analysis, respirometry, aquarium experiment, environmental DNA, fish tank

OPEN ACCESS

Edited by:
Hiroki Yamanaka, Ryukoku University, Japan

Reviewed by:
Matthew Yates, Université du Québec à Montréal, Canada
Meredith B. Nevers, United States Geological Survey (USGS), United States

****Correspondence:***
Bettina Thalinger
Bettina.Thalinger@gmail.com

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *30 October 2020*
Accepted: *29 January 2021*
Published: *26 February 2021*

Citation:
Thalinger B, Rieder A, Teuffenbach A, Pütz Y, Schwerte T, Wanzenböck J and Traugott M (2021) The Effect of Activity, Energy Use, and Species Identity on Environmental DNA Shedding of Freshwater Fish. Front. Ecol. Evol. 9:623718. doi: 10.3389/fevo.2021.623718

INTRODUCTION

The sensitivity, non-invasiveness, and cost-efficiency of environmental DNA (eDNA) based methods have been proven for diverse habitats and species making them powerful new tools for conservation biology and biodiversity assessments (Barnes and Turner, 2016; Deiner et al., 2017; Huerlimann et al., 2020). Regarding the detection of fish species, eDNA-based monitoring

outperforms traditional methods such as electrofishing: for example, for the detection of the endangered European weather loach, *Misgurnus fossilis* (Sigsgaard et al., 2015), the assessment of fish communities in Australian streams (McColl-Gausden et al., 2020), and the distribution of brook trout, *Salvelinus fontinalis* in a US watershed (Evans et al., 2017). The manifold successes of eDNA-based species detection lead to a call for more standardization and better reporting practices (Goldberg et al., 2016; Minamoto et al., 2020; Thalinger et al., 2021a) and to an international effort for implementing the technology into routine species monitoring (Leese et al., 2016; Pilliod et al., 2019). Although reporting the presence/absence of particular species is the starting point of these endeavors, a more quantitative interpretation of field-derived eDNA data is key for the general application of this technology.

Different processes influence the distribution of eDNA in space and time and the detection probabilities of species from environmental samples, namely the origin, degradation, suspension, resuspension, and transport of eDNA (Barnes and Turner, 2016; Harrison et al., 2019). The latter processes are directly linked to local hydrology [e.g., flow and substrate type (Shogren et al., 2017; Pont et al., 2018; Thalinger et al., 2021b)] and environmental conditions [e.g., water temperature, pH, UV-radiation (Strickler et al., 2015; Lacoursière-Roussel et al., 2016; Tsuji et al., 2017)]. The amount of eDNA in the water column is directly linked to fish biomass and originally, this was confirmed for common carp (*Cyprinus carpio*) in an aquarium trial and in experimental ponds (Takahara et al., 2012). In subsequent experiments, the positive relationship was confirmed for a range of freshwater and marine fish species (Evans et al., 2016; Lacoursière-Roussel et al., 2016; Sassoubre et al., 2016; Horiuchi et al., 2019; Jo et al., 2020). However, these results were primarily obtained for individuals at the same life stage.

Environmental DNA is released into the environment in the form of mucus, feces, scales, and gametes (Merkes et al., 2014; Barnes and Turner, 2016; Sassoubre et al., 2016; Bylemans et al., 2017). Under natural conditions, differences in fish physiology, diet, and behavior are likely to affect this process and confound the interpretation of eDNA-based results from a water body (Klymus et al., 2015). For perch and eel, Maruyama et al. (2014) and Takeuchi et al. (2019), respectively, found lower mass-specific eDNA shedding rates for adults in comparison to juveniles, which is likely caused by the scaling in metabolic rates, excretion rates, and surface area with body mass (discussed in Yates et al., 2020). However, these findings could not be confirmed in another experiment with a salmonid species (Mizumoto et al., 2018). In general, the metabolic rate and activity differ between fish species due to distinct physiology and behavior with pelagic species being more active and displaying higher resting metabolic rates than benthic species (Johnston et al., 1988; Killen et al., 2010). A stress response characterized by elevated metabolism and activity is frequently hypothesized as underlying cause for spiking eDNA levels at the beginning of aquarium experiments. Furthermore, metabolism and activity could generally explain mismatching quantitative results in studies comparing eDNA levels between species in the same water body (Takahara et al., 2012; Maruyama et al., 2014; Evans et al., 2016).

Here, we investigate the effect of fish activity (i.e., movement), energy use (i.e., oxygen use × oxycaloric factor), and species identity in an aquarium experiment with seven fish species commonly occurring in European rivers and streams (**Figure 1**). We hypothesized that higher activity leads to higher eDNA concentrations as there is more shearing between the fish surface and the surrounding water, and higher volumes are pumped through the gills due to the elevated oxygen demand. Independent of activity, fish species with higher energy use in a resting state potentially also emit more eDNA. Additionally, the species-specific composition of the constantly renewed cutaneous mucus layer (Ángeles Esteban, 2012) might lead to differences between individual taxa.

MATERIALS AND METHODS

Study Species

The examined species comprised four salmonids (*Salmo trutta, S. fontinalis, Oncorhynchus mykiss,* and *Thymallus thymallus*), two cyprinids (*Phoxinus phoxinus* and *Squalius cephalus*), and one sculpin (*Cottus gobio;* **Figure 1**). *S. trutta* is a rhithral species, territorial especially in later life stages, and primarily feeds on benthic organisms and insect drift on the surface. *S. fontinalis* and *O. mykiss* were anthropogenically introduced into European freshwaters and are less territorial than *S. trutta* (Freyhof and Kottelat, 2007). If possible, these three species choose areas with reduced current close to the main riverbed as preferential microhabitat. *T. thymallus* is also a rhithral species, but its scales are larger and adults primarily use the main riverbed (Spindler, 1997; Freyhof and Kottelat, 2007). *P. phoxinus* is a schooling, small fish species in the rhithral. It feeds on a mixture of plant debris, algae, and small invertebrates. The juveniles prefer vegetation-rich microhabitats without current, while adults switch to gravel substrate with low to intermediate flow. *S. cephalus* is eurytopic (rhithral to potamal) and can occur in habitats with strong to low current. Its juveniles are schooling and omnivorous with adults predominantly preying on fish. *C. gobio* is a rheophilic and benthic species primarily feeding on small bottom invertebrates. It has no swim bladder and mostly resides in interstices between large boulders or on coarse gravel characterized by low current (Muus and Dahlström, 1968; Spindler, 1997; Freyhof and Kottelat, 2007).

Experimental Setup

The aquarium experiment was carried out between March 2, 2017 and July 17, 2017 at the Research Department for Limnology, Mondsee of the University of Innsbruck, Austria. The juvenile salmonid individuals were purchased from commercial hatcheries, *P. phoxinus* and *S. cephalus* were caught with permission in Lake Mondsee and *C. gobio* were caught with permission in rivers in Tyrol (Austria). Fish individual sizes were chosen as similar as possible within and between species. As *P. phoxinus* and *C. gobio* are smaller in comparison to the other species (**Figure 1**), these individuals were supposedly closer to reproductive maturity. Until the start of the experiment, the fish species were kept separately in aquaria fed with lake water.

	Salmo trutta	*Salvelinus fontinalis*	*Oncorhynchus mykiss*	*Thymallus thymallus*	*Phoxinus phoxinus*	*Squalius cephalus*	*Cottus gobio*
family	Salmonidae	Salmonidae	Salmonidae	Salmonidae	Cyprinidae	Cyprinidae	Cottidae
length	resident stream individuals up to 300 mm SL	in European freshwaters up to 500 mm SL	up to 1000 mm SL	up to 600 mm TL	up to 140 mm TL	up to 600 mm SL	up to 180 mm TL
status and association with river zonation	native; rhithral zone	introduced	introduced; frequently stocked	native; rhithral zone	native; rhithral zone	native; rhithral to potamal zone	native
distinct morphological features	small scales	small scales	small scales	large scales	small scales	large scales	no swim bladder and no true scales
feeding	aquatic invertebrates, insect drift, and small fish	aquatic invertebrates, insect drift, and small fish	aquatic invertebrates, insect drift, and small fish	predominantly drifting insects, small worms and crustaceans	plant debris, algae, and small invertebrates	juveniles: omnivorous; adults: primarily piscivorous	small bottom invertebrates
preferred microhabitat	locations with reduced current close to the main riverbed	locations with reduced current close to the main riverbed	locations with reduced current close to the main riverbed	main river bed with potentially strong current	juveniles: vegetation-rich locations without current; adults: gravel substrate with low to intermediate current	eurytopic; can occur in habitats with strong or low current	benthic: resides in interstices between large boulders or on coarse gravel characterized by low current
social behavior	territorial (already as juveniles)	less territorial and at equal size dominated by *S. trutta*	less territorial than *S. trutta*	schooling	schooling	juveniles: schooling; adults: solitary or in small groups	solitary and territorial

FIGURE 1 | A summary of the morphological and ecological traits of the fish species used in the aquarium experiment. The provided information describes the situation in Central European freshwaters and is not necessarily transferable to other geographic regions. Depending on the source, different maximum fish length measurements were available with "total length (TL)" measured from the tip of the snout to the longest tip of the caudal fin and "standard length (SL)" measured from the tip of the snout to the base of the caudal fin (Muus and Dahlström, 1968; Spindler, 1997; Freyhof and Kottelat, 2007) (https://upload.wikimedia.org/wikipedia/commons/b/bb/CottusGobioSpreadingFins.JPG separated from background; Piet Spaans, CC BY-SA 2.5; https://creativecommons.org/licenses/by-sa/2.5, via Wikimedia Commons; https://upload.wikimedia.org/wikipedia/commons/9/99/Thymallus_thymallus2.jpg separated from background; Gilles San Martin, CC BY-SA 2.0 https://creativecommons.org/licenses/by-sa/2.0, via Wikimedia Commons).

In accordance with the regulations of the Austrian Animal Experiment Act (December 28, 2012) (Tierversuchsrechtsänderungsgesetz, part 1, section 1, §1, and point 2), and with the Directive 2010/63/EU of the European Parliament and of the Council of the European Union (September 22, 2010) on the protection of animals used for scientific purposes (chapter 1, article 1, and point 5a), all fish were reared according to regular agriculture (aquaculture) practice, including the provision of appropriate tank size, sufficient rate of waterflow, natural photoperiod, *ad libitum* food supply, and temperatures within the species' thermal tolerance range. This ensured that no pain, suffering, distress or lasting harm was inflicted on the animals, confirmed by the fact that mortality rates were low and equal between rearing groups. Based on the legislative provisions above, no ethics approval and no IACUC protocol was required for the experiments performed. In particular, the respirometry experiments were discussed with the legislative authorities (Austrian Federal Ministry of Education, Science and Research and the University of Veterinary Medicine, Vienna) and the conclusion was that the assessment of basic metabolism under these conditions (small fish sizes in relatively large chambers) does not incur pain, suffering or distress to the fish and no formal animal experimentation protocol was required.

Five aquaria (60 l) and corresponding plastic lids were used in the experiment, which were thoroughly cleaned with sodium hypochlorite (5%) and then rinsed with tap water (fish-DNA-free) prior to each experimental run (i.e., changing the fish under investigation). The flow-through rate for the tap-water fed aquaria was set to 5.45 l/min to mimic natural conditions and keep eDNA concentrations in the fish tanks constant based on the results of previous test runs (**Supplementary Material 1**). The water temperature in the aquaria was stabilized at 15°C by centrally heating the inflowing water to this temperature. Each tank was further equipped with an air-stone to ensure water mixing. At the start of each experimental run, a water sample (negative control) was taken from one of the aquaria and processed as described below. Then, five fish individuals per species were selected aiming at similar size. Each fish was placed individually in an aquarium using DNA-free fishnets (**Figure 2**). For *P. phoxinus* and *S. cephalus*, the experiment was carried out twice: once with individual fish, and once with groups of three fish per aquarium. The day before the experiment and for its duration, the respective fish were not fed to avoid contamination by fish feed and minimize the effects of defecation. Each run started with 1 day of familiarization.

Water Sampling, Filtration, and pH Measurements

All equipment used for this process was cleaned with sodium hypochlorite (5%) and rinsed with tap water prior to each use;

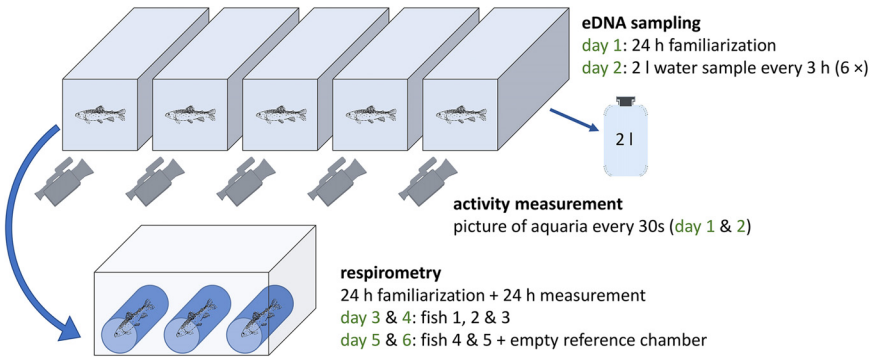


FIGURE 2 | The setup of the aquarium experiment carried out with seven fish species: five individual fish were put in fish tanks for water sampling (eDNA) and activity recordings (days 1 and 2) followed by respirometer measurements (three individuals on days 3 and 4; two individuals plus empty control chamber on days 5 and 6). For *Phoxinus phoxinus* and *Squalius cephalus* the experiment was repeated using groups of three individuals per tank and respirometer chamber.

DNA-free gloves were always worn. On the second day, 2 l water samples were taken every 3 h from 9:00 AM to 12:00 AM (six samples) at the back end of each aquarium (opposite to the inflow) using flexible tubes and 2 l wide neck bottles (**Figure 2**). Due to the high flow rates (entire water volume replaced every 11 min), the water level in each aquarium self-adjusted automatically after every sampling. The water samples were immediately filtered in an adjacent laboratory using glass microfiber filters (1.2 µm pore width, 47 mm diameter, Whatman GF/C) and one negative control (2 l MilliQ-water) was included per sampling event. Thereafter, the filters were individually placed in 2 ml reaction tubes and stored at −28°C until further processing in a special diagnostic molecular laboratory at the Department of Zoology, University of Innsbruck (Austria). After each sampling, pH was measured in three arbitrarily selected aquaria using a Hach HQ40 device.

Activity Measurement

During the familiarization time (day 1) and between water samplings, fish swimming activity was quantified using a custom-made activity monitoring system consisting of one high-definition USB camera (Ziggi HD Plus, IPEVO.COM) per aquarium. The cameras were placed at the front of each tank and the focus was set toward the back end (**Figure 2**). To enable recordings during the night, aquaria were lighted throughout the two recording days. Additionally, white polystyrene plates were used to cover the bottom and the sides to exclude influences from neighboring aquaria and standardize reflections. The signals from the cameras were acquired with a frame rate of 2 frames per minute (fpm) with a macro using the image analysis software FIJI[1] (a distribution of ImageJ) for MacOS (Schindelin et al., 2012; Rueden et al., 2017). For each aquarium, a region of interest (ROI) excluding the inflow, air-stone and sidewalls was set manually (**Supplementary Material 2**). Subsequent frames were arithmetically subtracted and the average gray-scale within the region of interest, as a quantification of fish activity, was extracted from the difference-images. The dataset was manually checked to exclude artifacts produced by changes in illumination (light/dark illumination of the fish), water sampling, measurement of abiotic factors, fogged-up aquarium front and few camera movements sometimes leading to a changing ROI in the recordings (**Supplementary Material 2**).

Respirometry

A custom-made intermittent-flow respirometer was used (Forstner, 1983; Svendsen et al., 2016) including three measurement chambers placed in a larger tank (**Figure 2**). The device was cleaned prior to each fish change using a mixture of 3% hydrogen peroxide and 3 l of tap water. The volume of each chamber was determined prior to the experiment and oxygen saturation (100%) and temperature (8–9°C) were kept constant in the tank via an airstone and heating/cooling device (Lauda DLK 10 and Alpha 1, Lauda Germany). The three chambers of the respirometer were connected to the respirometers' water circuit, constantly pumping O_2-saturated water from the large tank through the three chambers. For measurements of oxygen consumption, a chamber was cut off from this circuit and a closed-loop was established. Dissolved oxygen was measured in this chamber every 30 s for a period of 15 min using a YSI ProODO probe (YSI Inc.) and logged to a computer before the system switched to the next chamber for a 15 min measuring period. On the third day of an experimental run, three of the five fish were placed individually into the chambers avoiding air bubbles and kept there for 24 h for familiarization. On the fourth day, respirometer measurements were carried out for 24 h. Thereafter, the remaining two fish individuals were placed in two measurement chambers for 1 day of familiarization followed by 1 day of measurements (days 5 and 6; **Figure 2**). The third chamber was left empty, but measured as well, to evaluate potential microorganism-induced oxygen decrease. After the respirometer measurement day, the mass [g] and total length [mm] of each fish were determined before placing them together in a fish tank. For respirometer measurements of fish groups, the three individuals previously sharing an aquarium were put together in a respirometer chamber.

[1]http://fiji.sc/

Filter Processing and Molecular Analysis

After defrosting, each filter was soaked with 200 µl of lysis buffer consisting of TES-buffer (0.1 M TRIS, 10 mM EDTA, 2% sodium dodecyl sulfate; pH 8) and proteinase K (20 mg/ml) in a ratio of 19:1 and incubated at 56°C over night in a rocking platform. On the next day, filters were transferred with DNA-free forceps to a perforated inset which was repositioned in the top half of the original 2 ml reaction tube and centrifuged for 10 min at 20,000 *g*. Afterward, filters were discarded and the lysate at the bottom of the reaction tube (300–800 µl) was used for DNA extraction. Insets were cleaned in sodium hypochlorite (2.5%) for at least 30 min, thoroughly washed with MilliQ-water (10 wash steps) and reused.

DNA extraction was carried out with the Biosprint 96 instrument (Qiagen) using the Biosprint 96 DNA blood Kit (Qiagen) and the Biosprint 96 tissue extraction protocol following the manufacturer's instructions except for using 100 µl of TE-buffer instead of AE-buffer for DNA elution. Extractions were carried out in 96-well plates and four negative controls (containing TES-buffer instead of lysate) were included per plate. To process the whole lysate volume, a custom DNA-uptake program was set up: three uptake plates were used and 300 µl of lysate, 300 µl AL-buffer and 300 µl isopropanol were mixed per well in each plate. Missing lysate volumes (i.e., if only a total of 400 µl were available after centrifugation) were replaced by TES-buffer. Additionally, 30 µl MagAttract was added per well in the first plate. Using custom "binding" steps of the robotic platform, the DNA contained in the first plate was transferred to the second one. Next, a binding step was carried out in the second plate before transferring and releasing the entire collected DNA into the third plate, which was then used for the Biosprint 96 tissue extraction protocol. After extraction, each eluate was transferred to a 1.5 µl reaction tube for subsequent PCR.

All used primers (**Table 1**) have been previously published after extensive specificity and sensitivity testing (Thalinger et al., 2016, 2021b) and additional specificity tests were carried out on the digital PCR (dPCR) system (see below) confirming the specificity of the molecular assays under the following conditions: each 22 µl dPCR master mix for droplet generation on the QX200 AutoDG (Biorad) consisted of one-time EvaGreen Supermix (Biorad), 0.25 µM forward and reverse primer (**Table 1**) and up to 10.5 µl DNA extract. Depending on the results of initial tests with capillary electrophoresis PCR (i.e., the Relative Fluorescence Units (RFU) of the resulting band; see **Supplementary Material 3**), extracts were diluted with molecular grade water for dPCR as follows: RFU < 0.2: undiluted; 0.2 ≤ RFU < 1.3: 1:1 dilution; 1.3 ≤ RFU < 2: 1:3 dilution; 2 ≤ RFU: 1:7 dilution. Optimized thermo-cycling conditions were 5 min at 95°C, 40 cycles of 30 s at 95°C, 1 min at 58°C (*O. mykiss*, *P. phoxinus,* and *S. cephalus*), or 60°C (*C. gobio*, *S. fontinalis*, *S. trutta*, and *T. thymallus*), 1 min at 72°C, followed by one step of 5 min at 4°C and 5 min at 90°C. dPCR results were analyzed on the QX200 Droplet Reader with the corresponding QuantaSoftTM Analysis Pro Software (Version 1.7; Biorad). As target signal amplitude varied with the length of the amplified fragment, amplitude thresholds were set individually per primer pair (**Table 1**) prior to determining target copy numbers per µl for each DNA extract. Each sample was subjected to dPCR once, based on previous studies indicating a high precision of dPCR for low target DNA concentrations (e.g., Doi et al., 2015). Per primer pair, a positive control (DNA extract from target species tissue) and a negative control (molecular grade water) were included in dPCR, all of which resulted positive and negative, respectively. All filtration and extraction controls resulted negative as well.

Statistical Analysis

All calculations and visualizations were carried out in R Version 4.0.2 (R Core Team, 2020) using the packages "ggplot2" (Wickham, 2016), "gridExtra" (Auguie, 2017), "ggpubr" (Kassambara, 2019), "lme4" (Bates et al., 2015), "AICcmodavg" (Mazerolle, 2020), "MuMIn" (Barton, 2019), "rsq" (Zhang, 2020), and "sjPlot" (Lüdecke, 2020). As pH was not measured in all aquaria after each water sampling, missing values were estimated by averaging measurements taken at the respective fish tank before and after the skipped time step. If measurements at the first or last water sampling were missing, the values of the following or previous time step, respectively, were carried over.

The cleared activity dataset was visually inspected and summarized for each time step: for example, data obtained during the preceding day were associated with the first eDNA sampling event at 9:00 AM and measurements between 9:00 AM and 12:00 PM were considered relevant for the second water sampling at 12:00 PM. Mean activity was calculated per time interval. No cleared activity data was available for one *S. trutta* and *S. fontinalis* individual, respectively, and for one *P. phoxinus* and *T. thymallus* individual at a single time step each.

The total respirometry dataset was cleared of all 15 min measurement series showing an increase in dissolved O_2. As this value is expected to decrease linearly over the course of a measurement (Svendsen et al., 2016), a linear regression for the oxygen decrease in a measurement chamber over time was calculated for each measurement series. All intervals for which the obtained values showed an insufficient fit to a linear decrease ($R^2 < 0.8$) were also excluded from further analyses. For each of the remaining measurement intervals, oxygen consumption (OC) in mg / h was calculated as $OC = -s \times 60 \times vol$ where "s" denotes the slope of the linear regression and "vol" the volume of the respective measurement chamber minus the mass of the fish. Per fish species, the obtained value was corrected for the mean oxygen consumption in the empty chamber before calculating total energy use (oxygen consumption × 13.6 J/mg [oxycaloric factor (Brett and Groves, 1979)] per fish. Finally, energy use [J/h] was averaged across the values obtained from individual measurement intervals for each fish and fish group. Due to data clearing, this was not possible for one individual and one group of *C. gobio* and *S. cephalus*, two individuals of *S. fontinalis* and *S. trutta* and three individuals of *T. thymallus*. For these fish, energy use was estimated as the mean of the available values.

Concerning the fish eDNA copy numbers obtained from dPCR, 21 filtered water samples did not lead to an amplification. They were removed from the dataset, as other fish individuals of comparable size and other samplings reliably produced positive results and/or eDNA was detected in celPCR. Hence, processing errors during sampling and in the laboratory were deemed

TABLE 1 | Digital PCR assays used to amplify fish eDNA.

Target taxon	Primer name	Primer sequence (5′ – 3′)	Primer conc. in dPCR (μM)	Target gene	Fragment length (bp)	Amplitude threshold (dPCR)	Source
Cottus gobio	Cot-gob-S632	GAATAAAGGACTAAACCAAGTGGG	0.25	16S	118	13,500	Thalinger et al., 2016
	Cot-gob-A641	GCTGTAGCTCTCAGTTGTAGGAAAA	0.25				
Salmo trutta	Sal-tru-S1002	TCTCTTGATTCGGGCAGAACTC	0.25	COI	89	8,400	Thalinger et al., 2021b
	Sal-tru-A1002	CGAAGGCATGGGCTGTAACA	0.25				
Oncorhynchus mykiss	Onc-myk-S655	TCTCCCTTCATTTAGCTGGAATC	0.25	COI	82	12,500	Thalinger et al., 2016
	Onc-myk-S655	GCTGGAGGTTTTATGTTAATAATGGTC	0.25				
Salvelinus spp.	Sal-vel-S651	ATAGTCGGCACCGCCCTT	0.25	COI	112	14,000	Thalinger et al., 2016
	Sal-vel-A651	TAACGAAGGCATGGGCTGTT	0.25				
Thymallus thymallus	Thy-thy-S653	ATCAAATTTATAATGTGATCGTCACG	0.25	COI	179	14,000	Thalinger et al., 2016
	Thy-thy-A653	AAGAAAGGACGGGGGAAGC	0.25				
Phoxinus phoxinus	Pho-pho-S639	CGTGCAGAAGCGGATATAAATAC	0.25	16S	128	15,750	Thalinger et al., 2016
	Pho-pho-A648	CCAACCGAAGGTAAAGTCTTATTG	0.25				
Squalius cephalus	Squ-cep-S669	CAGTATACCCACCGCTTGCG	0.25	COI	130	14,250	Thalinger et al., 2016
	Squ-cep-A669	TTAATAATTGTGGTAATGAAGTTGACC	0.25				

Columns denote the target taxon of each primer combination, primer names, sequences, their respective concentration in dPCR, target gene, amplicon sizes, and threshold values for positive droplets in dPCR. Additionally, the source column shows the original publication. Please note that the Salvelinus spp. primer pair was designed to amplify both S. fontinalis and Salvelinus umbla.

the most likely cause for failing amplification. One group of *P. phoxinus* had to be excluded from further analyses, as two of three individuals were identified as *S. cephalus* when removed from the aquarium after the experiment. To determine whether the pH measurements, mean activity and eDNA copy numbers were significantly influenced by sampling (i.e., time of the day), a one-way repeated measurements ANOVA with rank transformation was calculated for each variable using a combination of fish species and aquarium as random factor. A significant trend could not be detected (**Table 2**). Despite efforts to standardize the mass of the chosen fish individuals within and between species, fish mass was identified as confounding variable (**Supplementary Material 4**). Hence, eDNA copies, mean activity, and energy use were normalized by the mass of the respective fish individual prior to all further analyses.

Generalized Linear Mixed-Effects Models (GLMM) for a Gamma-distributed dependent variable (i.e., eDNA copies) were set up with a log-link function to investigate the effects of mean activity, energy use, fish species, and pH (Faraway, 2016). Data obtained from fish groups were excluded from the comparison of model performance. Fish individuals were used as random intercept to account for repeated measurements and models were fit using Gauss-Hermite quadrature (nAGQ = 20) and the BOBYQA algorithm (Bolker et al., 2009; Powell, 2009). The variable "fish species" was entered via dummy coding into the models using *C. gobio* as base category. Corresponding with the focus of this study to investigate the effect of species identity, energy use, and activity on eDNA shedding, a set of six candidate models was chosen (**Table 3**). AICc, ΔAICc, and AICc weights (ω) were used to evaluate the strength of the six models for describing the data including marginal and conditional pseudo-R^2 values (Burnham and Anderson, 2002; Nakagawa and Schielzeth, 2013). Simulated, scaled residuals were calculated based on the best-performing candidate model, [package: DHARMa (Hartig, 2020); function: "simulateResiduals"; n = 1,000]. The best performing model passed the consecutive check for outliers and overdispersion; the 95%-CI for its fixed effects were derived via bootstrapping (200 simulations).

To test the differences between single and grouped fish in the different stages of the experiment, a data subset containing only values obtained from single and grouped *P. phoxinus* and *S. cephalus* was analyzed. Target eDNA copies, energy use, and mean activity (all normalized by fish mass) of the four distinct fish categories were tested for normality and homogeneity of variance with Shapiro-Wilk and Bartlett tests. Then, differences between groups were examined via Kruskal-Wallis tests followed by Wilcoxon Rank Sum tests with Benjamini-Hochberg-corrected p-values. In a final step, target eDNA copies for groups of *P. phoxinus* and *S. cephalus* were predicted using the model

TABLE 2 | The results of one-way repeated measurements ANOVA with rank transformation examining a potential effect of sampling on pH, mean activity, and target eDNA copy numbers.

		F-value	*p*-value
pH	**Intercept**	**34.52**	**<0.001**
	Sampling	2.22	0.053
Mean activity	**Intercept**	**78.78**	**<0.001**
	Sampling	0.76	0.57
Target copies per μl	**Intercept**	**39.00**	**<0.001**
	Sampling	0.38	0.86

Significant differences ($p < 0.05$) are in bold.

TABLE 3 | Covariate structures of the candidate Gamma GLMM with a log link function.

Model #	Covariate structure (fixed effects)
1	Mean activity + energy use + fish species + pH + sampling
2	Mean activity + energy use + fish species + sampling
3	Mean activity + energy use + fish species
4	Mean activity + energy use
5	Mean activity + fish species
6	Fish species

The models were compared for their potential to explain the target eDNA copy numbers per gram fish and µl extract obtained from single-fish aquaria with the following parameters: fish species identity (seven species), mean activity (per gram fish), energy use (per gram fish), and pH. In all models, individual fish were included as random effect (random intercept) to account for repeated measurements and the sampling event was included primarily to show its insignificance.

previously established for single fish (only possible when not incorporating the random effect of fish individual). Pairwise Wilcoxon tests were used to verify whether there was a significant difference between predicted and measured target eDNA copy numbers for both species separately and combined.

RESULTS

The mean mass of individually housed fish was 3.06 g ± 1.56 g (SD) and *C. gobio* individuals had the highest mass [5 g ± 2.1 g (SD); **Table 4**]. Water samples from *P. phoxinus* and *T. thymallus* aquaria had the highest eDNA copy numbers per µl extract and gram fish mass [31.13 ± 53.23 (SD) and 47.68 ± 41.13 (SD), respectively; **Figure 3** and **Table 4**]. The normalized mean activity was highest for *S. fontinalis* [1.08 ± 0.33 (SD)] and lowest for *C. gobio* [0.34 ± 0.10 (SD); **Figure 3** and **Table 4**]. The energy use per gram fish mass was highest for *O. mykiss* [1.81 J/h ± 0.91 J/h (SD)], while *S. fontinalis* and *S. trutta* aquaria had the lowest pH.

The ΔAICc-based comparison of model weight (single fish only) resulted in model #3 outperforming five other candidate models (**Tables 3**, **5**). Therein, mean activity, energy use, and fish species were contained as explanatory variables (conditional pseudo-R^2 = 0.59). Increased activity had a significantly positive effect on eDNA copy numbers ($p < 0.05$) and *P. phoxinus*, *S. cephalus*, and *T. thymallus* displayed significantly higher copy numbers compared to *C. gobio* (base group) after controlling for the effect of fish mass. The relationship between energy use and copy numbers was also positive, but not significant ($p = 0.08$; **Table 6** and **Figure 4**).

For single and grouped individuals of *P. phoxinus* and *S. cephalus*, target eDNA copies per gram fish were significantly higher for grouped fish in general [28.96 ± 35.44 (SD) compared to 22.44 ± 38.64 (SD); $Chi^2 = 5.96$; $p < 0.05$]. Specifically, they were significantly higher for grouped *P. phoxinus* [42.61 ± 48.04 (SD)] compared to single *P. phoxinus* and single and grouped *S. cephalus* and characterized by few outliers with particularly high eDNA concentration (**Figures 3**, **5**). Significant differences were also detected between the four groups regarding mean activity ($Chi^2 = 80.95$; $p < 0.001$) and energy use ($Chi^2 = 36.77$; $p < 0.001$): mean activity was significantly higher when fish were kept solitary compared to having them in groups for both species ($p < 0.001$). Contrastingly, energy use was significantly higher for grouped individuals of *P. phoxinus* and *S. cephalus* ($p < 0.01$).

To test the suitability of model #3 for describing eDNA shedding also for grouped fish, model #3-predicted eDNA copies were compared to the measured copy numbers in the group treatments. For the two species combined, there was no significant difference between predicted and measured copy numbers ($W = 1612$, $p = 0.35$). For *P. phoxinus* alone, no such difference was detected either ($W = 274$; $p = 0.78$; **Figure 6**), while predicted and measured copy numbers of *S. cephalus* showed a significant difference ($W = 609$; $p < 0.05$; **Figure 6**).

DISCUSSION

This experiment confirms the hypothesized positive relationship between eDNA shedding and fish activity. The species identity and thereby associated physiological differences were found to influence the amount of released eDNA, and the positive relationship between energy use and eDNA signals was not significant. Furthermore, our data show that models of eDNA shedding cannot always be generalized from experiments with individual fish to fish groups. For a conclusive habitat-scale

TABLE 4 | The means and standard deviations of mass, activity, and energy use for each fish species in the experiment.

	Mass [g] ± SD [g]	Activity ± SD	Energy use [J/h] ± SD [J/h]	eDNA copies ± SD
Cottus gobio	5.00 ± 2.10	0.34 ± 0.10	0.68 ± 0.17	2.48 ± 2.94
Oncorhynchus mykiss	3.30 ± 0.31	0.51 ± 0.10	1.81 ± 0.91	9.60 ± 5.16
Phoxinus phoxinus	3.48 ± 0.55	0.55 ± 0.09	0.67 ± 0.16	31.13 ± 53.23
Salvelinus fontinalis	1.40 ± 0.42	1.08 ± 0.33	1.06 ± 0.00	11.46 ± 8.72
Salmo trutta	2.22 ± 0.77	0.68 ± 0.29	0.92 ± 0.00	5.98 ± 8.08
Squalius cephalus	2.72 ± 0.66	0.69 ± 0.09	1.06 ± 0.44	14.14 ± 11.56
Thymallus thymallus	3.53 ± 1.42	0.56 ± 0.30	0.43 ± 0.01	47.68 ± 41.13
Phoxinus phoxinus grouped	12.06 ± 1.86	0.26 ± 0.09	1.00 ± 0.33	42.61 ± 48.04
Squalius cephalus grouped	7.75 ± 1.30	0.32 ± 0.11	1.41 ± 0.50	18.04 ± 13.69

The eDNA copies (per µl extract), activity, and energy use are provided per gram fish mass; for grouped fish, the mass is displayed per aquarium (i.e., sum of three fish individuals).

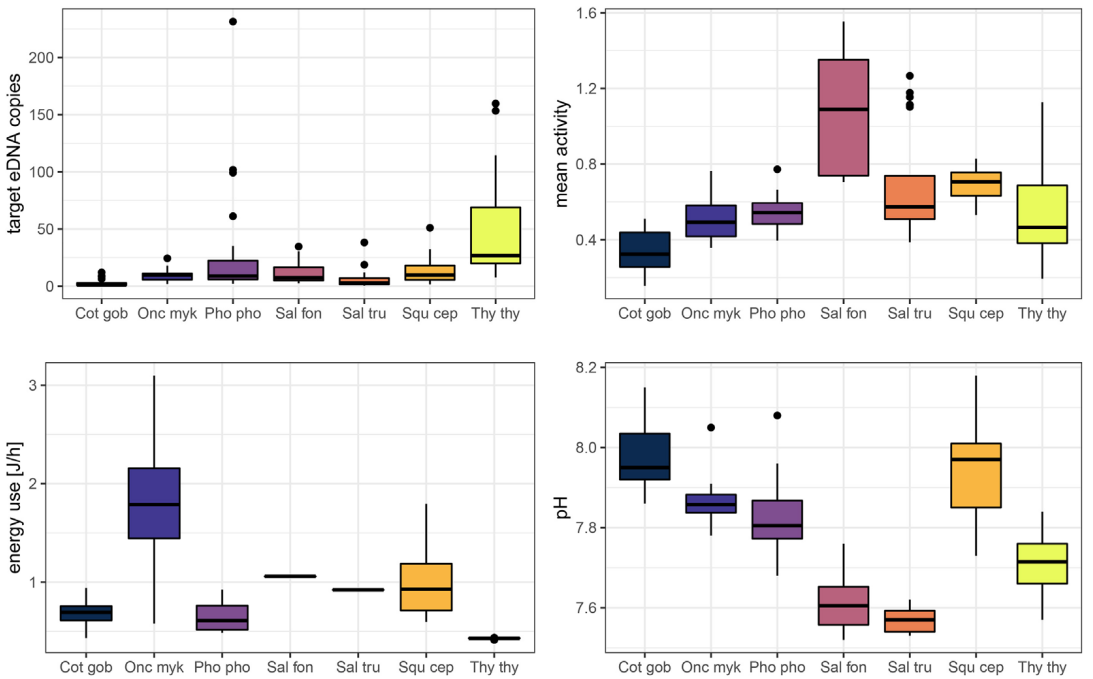


FIGURE 3 | Key parameters obtained during the experiment for single fish. Boxplots display target eDNA copies per µl extract, energy use [J/h], mean activity, and pH per fish species. Fish species are abbreviated: "Cot gob," *Cottus gobio*; "Onc myk," *Oncorhynchus mykiss*; "Pho pho," *Phoxinus phoxinus*; "Sal fon," *Salvelinus fontinalis*; "Sal tru," *Salmo trutta*; "Squ cep," *Squalius cephalus*; "Thy thy," *Thymallus thymallus*. The variables target eDNA copies, mean activity, and energy use were normalized by fish mass to control for the effect of this confounding variable.

estimation of fish communities with eDNA-based methods it is therefore necessary to incorporate species physiology and behavior into the analysis.

In early aquarium experiments, the strongest eDNA signals were found right after the introduction of fish into tanks without water circulation and often explained by elevated stress levels through handling and adaption to the new environment (Takahara et al., 2012; Maruyama et al., 2014; Klymus et al., 2015). Hence, many recent studies allow for one or several days of accommodation prior to eDNA sampling (Lacoursière-Roussel et al., 2016; Jo et al., 2019; Takeuchi et al., 2019). We can confirm the positive relationship between fish activity (i.e., movement) and eDNA shedding independent of the introductory phase of an experiment. However, it was not possible to determine the actual reason for the elevated eDNA levels associated with higher activity, as both higher metabolic rates during movement and higher water volumes shearing against the fish body could be responsible for this effect. For eDNA-based field studies, this result indicates that signals emitted by highly active fish (e.g., during spawning or predatory behavior) potentially mimic higher levels of fish biomass.

Energy use in a resting state as measured with an intermittent-flow respirometer, was also positively correlated with eDNA production, albeit not significant. In case this trend is confirmed in the future, it could be attributed to the higher metabolic rate and larger gill size of active species in combination with higher water volumes pumped through them (Wegner et al., 2009). However, the elevated eDNA signals could also stem from other physiological processes (e.g., defecation), which are known to positively influence eDNA production rates (Klymus et al., 2015). As fish were not fed during the entire experiment, the latter factor is potentially negligible, albeit it might substantially influence eDNA levels under natural conditions. Except for *T. thymallus*, the energy use of the species preferring microhabitats with strong currents and preying on fish as adults (primarily *O. mykiss* and *S. cephalus*) was higher than for *C. gobio* and *P. phoxinus*. This is in concordance with general differences in resting metabolic rates between these ecological guilds (Roberts, 1975; Johnston et al., 1988; Killen et al., 2010). In this experiment, the smaller sized *S. fontinalis* and *S. trutta* individuals, were more difficult to measure with the chosen respirometer setup (i.e., fewer measurements passed our quality filtering), which could be the cause for the weak relationship between energy use and eDNA copy numbers. In the future, more emphasis should be placed on a ratio of 20–50 between the volume of the measurement chamber in the respirometer and the fish individual to facilitate respirometer measurements (Svendsen et al., 2016).

There were distinct differences in eDNA shedding between the species, with *T. thymallus*, *P. phoxinus*, and *S. cephalus* emitting the most eDNA. The adaptation to habitats with stronger currents (Freyhof and Kottelat, 2007), namely an increased mucus production in combination with comparably large scales, might explain this result for *T. thymallus* and *S. cephalus*. The underlying taxonomy could also contribute to this pattern if cyprinids (*P. phoxinus* and *S. cephalus*)

TABLE 5 | Results of the ordinal ranking based on ΔAICc for the GLMM (**Table 3**).

Model #	K	AICc	Δ AICc	ω	Marginal pseudo-R^2	Conditional pseudo-R^2
3	11	159.57	0	0.52	0.56	0.59
5	10	160.28	0.71	0.36	0.56	0.60
6	9	162.74	3.17	0.11	0.53	0.60
2	16	167.10	7.53	0.01	0.57	0.60
1	17	169.55	9.98	0.00	0.57	0.60
4	5	194.18	34.6	0.00	0.04	0.52

Models are sorted from high to low weight and K denotes for the number of estimable parameters, AICc for the second-order variant of Akaike's Information Criterion, ΔAICc for AICc difference, ω for Akaike weight, and marginal/conditional pseudo-R^2 represent the variance explained by the fixed effects only and by the entire model, respectively.

TABLE 6 | The highest weight (ω = 0.52) GLMM (model #3) describing the measured eDNA copy numbers via (A) the fixed effects: mean activity, energy use and fish species identity, and (B) the random effect fish individual (31 groups, σ = 0.88).

(A)	Parameter estimate	Standard error	Lower 95% CI	Upper 95% CI	*t*-value	*p*-value
Intercept	0.19	0.31	−0.55	0.85	0.63	0.53
Mean activity	**1.00**	**0.42**	**0.19**	**1.95**	**2.39**	**0.02**
Energy use [J/h]	0.41	0.23	−0.07	0.94	1.77	0.08
Oncorhynchus mykiss	0.79	0.43	−0.21	1.57	1.85	0.06
Phoxinus phoxinus	**2.40**	**0.35**	**1.71**	**3.17**	**6.78**	**< 0.001**
Salvelinus fontinalis	0.65	0.47	−0.45	1.67	1.38	0.17
Salmo trutta	0.32	0.38	−0.58	1.28	0.83	0.40
Squalius cephalus	**1.25**	**0.38**	**0.41**	**2.01**	**3.28**	**< 0.01**
Thymallus thymallus	**2.93**	**0.33**	**2.30**	**3.58**	**8.75**	**< 0.001**
(B)	**Variance**	**Standard deviation**				
Fish individual (intercept)	0.06	0.24				

Significant p-values of fish species in the model refer to a significant difference between Cottus gobio (used as base category for dummy coding) and the respective fish species. Significant differences ($p < 0.05$) are in bold.

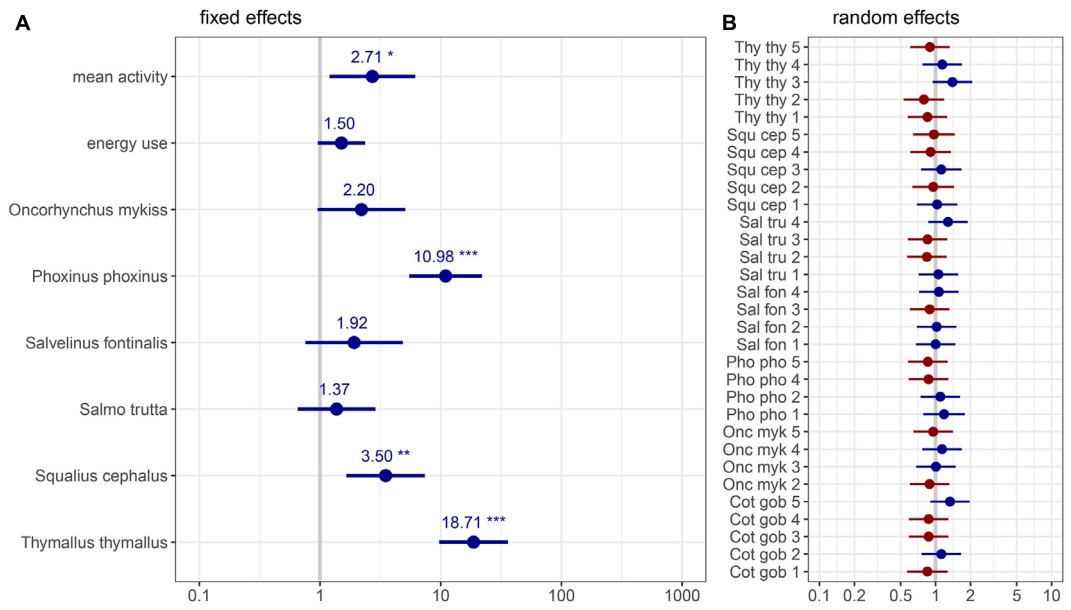


FIGURE 4 | Graphic representation of the GLMM estimates (model #3) best describing the obtained target eDNA copy numbers: **(A)** fixed effects and **(B)** random effect of individual fish. Coefficients are exponentiated, significance codes of denoted fish species indicate differences in comparison to the base category *Cottus gobio*, whiskers display the 95%-CI. Fish species are abbreviated: "Cot gob," *Cottus gobio*; "Onc myk," *Oncorhynchus mykiss*; "Pho pho," *Phoxinus phoxinus*; "Sal fon," *Salvelinus fontinalis*; "Sal tru," *Salmo trutta*; "Squ cep," *Squalius cephalus*; "Thy thy," *Thymallus thymallus* in addition to individual numbers from 1 to 5. Asterisks denote *p*-values smaller than: 0.05 (*), 0.01 (**), and 0.001 (***).

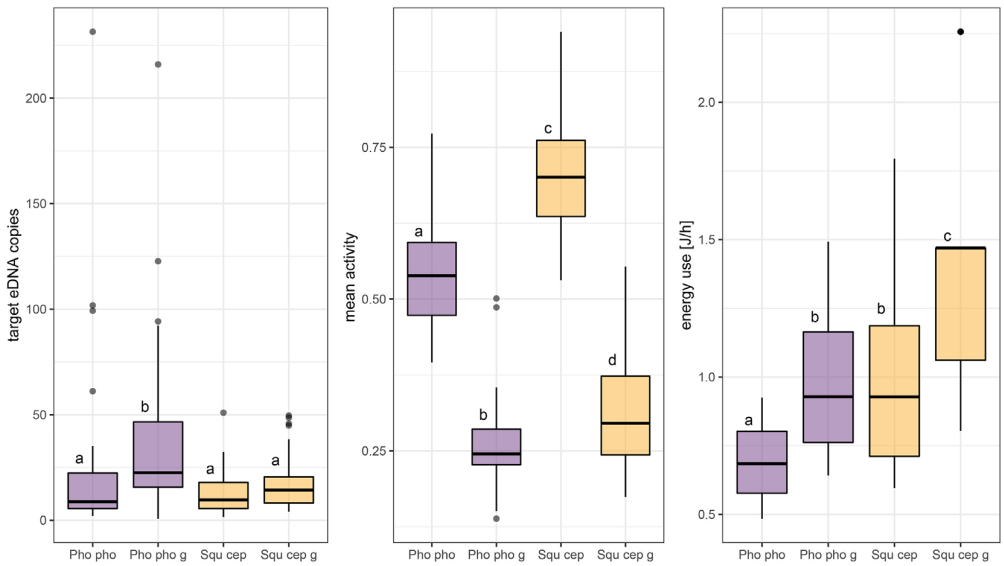


FIGURE 5 | Comparison of target eDNA copies, mean activity, and energy use (normalized by fish mass) in aquaria obtained from single and grouped individuals of *Phoxinus phoxinus* and *Squalius cephalus*. Different lower case letters above boxplots code for significant differences ($p < 0.05$) between categories, which are abbreviated as: "Pho pho," *Phoxinus phoxinus* (single fish); "Pho pho g," *Phoxinus phoxinus* grouped fish; "Squ cep," *Squalius cephalus* (single fish); "Squ cep g," *Squalius cephalus* grouped fish.

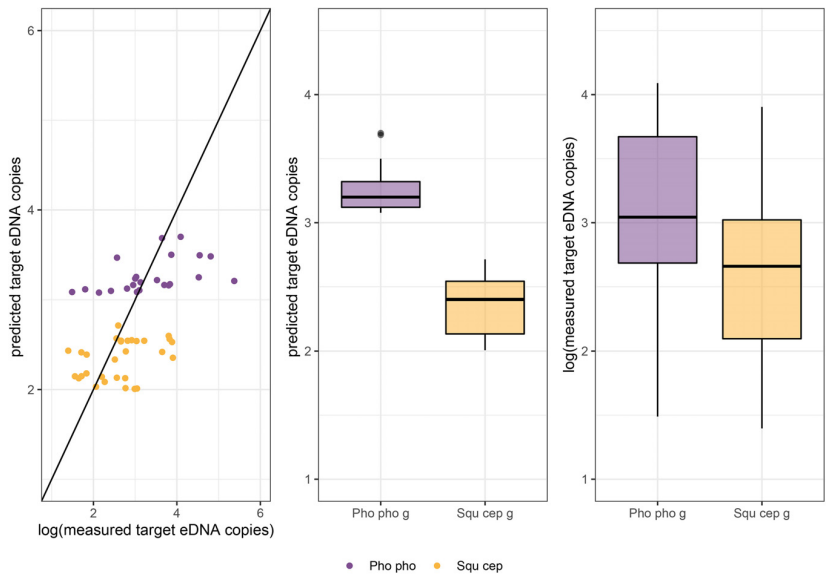


FIGURE 6 | For groups of *Phoxinus phoxinus* (Pho pho g) and *Squalius cephalus* (Squ cep g) measured and predicted copy numbers are plotted: left, against each other; middle, predicted copy numbers are compared between species; right, comparison of measured copy numbers between the two species. The measured copy numbers were log-transformed to enable a direct comparison with the values predicted by the Gamma GLMM with log-link function; the random effect of individual fish could not be taken into account for this prediction. For *S. cephalus* a significant difference between measured and predicted copy numbers was detected ($W = 609$; $p < 0.05$).

generally release more DNA into the surrounding water via their gills, feces or mucus. Another explanation for the high eDNA shedding of cyprinids in this experiment could be the stress induced by solitary housing. The model estimating eDNA concentrations for individual fish could not fully explain the findings obtained for grouped fish: the activity of both *P. phoxinus* and *S. cephalus* was significantly lower when fish were held in groups, while their energy use was significantly higher. A change in measurement precision regarding activity and energy use (respirometer more precise, activity measurement less precise for fish groups) could explain these contradictory results. Nevertheless, eDNA copies of grouped *P. phoxinus*

individuals did not differ from values predicted with a model based on single fish.

Generally, the measured eDNA concentrations per µl DNA extract were right-skewed and a few exceptionally high values showed a considerable influence on the size of standard deviations. These results were independent of fish handling and stress during the introduction phase as eDNA sampling started only after 24 h and the aquaria had constant flow with the entire volume being renewed every 11 min. Such "outliers" were also detected in other aquarium experiments (Klymus et al., 2015; Wilcox et al., 2016) and cell-conglomerates released into the surrounding water were previously deemed responsible for this pattern (Wilcox et al., 2016). Additionally, the size distribution of eDNA particles (starting from <0.2 µm and exceeding >180 µm) and commonly detected fragment sizes suggest intact cells or organelles as the primary source of eDNA in the water column (reviewed by Harrison et al., 2019). Our data support the hypothesis of constant eDNA shedding rates at constant environmental conditions indicating the potential to determine this variable for a broad range of species and add to the interpretation of field sampling results. We could not observe any effects of sampling time, possibly due to the constant illumination of the aquaria. Hence, this aspect is not necessarily transferable to natural environments where fish are known to exhibit distinct diurnal movement patterns (Helfman, 1986).

The influence of fish mass on eDNA concentrations was not in the focus of this experiment and fish individuals were as similar in size/mass as possible. However, adult fish of *P. phoxinus* and *C. gobio* are considerably smaller in comparison to the other species (Freyhof and Kottelat, 2007) and the respective juveniles were thus closer to sexual maturity. Hence, the allometric change of metabolic processes (Brown et al., 2004) could be an alternative explanation for the comparably low energy use of these two species. For studies investigating eDNA shedding directly from live animals, biomass will always be an influential and potentially confounding variable and should thus be considered carefully already during experimental design. Recently, the allometrically scaled mass was found to be the best index variable for describing eDNA concentrations in lakes (Yates et al., 2020); since excretion rate, metabolic rate and surface area all scale allometrically too (Brown et al., 2004; O'Shea et al., 2006; Vanni and McIntyre, 2016), future experiments could greatly benefit from the incorporation of this concept. For activity measurements via videotaping, fish length had to be used as an index variable. In this context, considerations of body shapes and fins, which differ a lot between taxa (Freyhof and Kottelat, 2007), are also advisable. Finally, individual differences are well documented for fish behavior and metabolic rates (Metcalfe et al., 2016). The number of study animals in future experiments should thus be increased to better control for such effects within a species.

Our results demonstrate that for the successful application of eDNA-based methods on a habitat scale it is necessary to incorporate fish physiology and behavior not only in the study design and sampling process [e.g., by sampling at different depths and in different micro-habitats (Littlefair et al., 2020)], but also during data analysis (Barnes and Turner, 2016; Thalinger et al., 2021a). Seasonal patterns could have a much stronger effect on eDNA concentrations in the water column as previously assumed: for instance, many cyprinids in European freshwaters seek calm areas without current during the winter. Their eDNA is less likely to spread through the water column and additionally, their decreased activity lowers the detection probability even further. In the future, the eDNA shedding of diverse fish species and families in relation to their biomass, activity, and energy use should be investigated to deepen our understanding of taxon-specific effects. Until then, estimations of fish biomass from eDNA quantities in field-collected samples should at least take distinct physiology and behavior into account, especially for comparative analyses between species or seasons.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: All data on eDNA signals, fish activity, energy use, fish mass, and pH have been uploaded to Figshare and are available at https://doi.org/10.6084/m9.figshare.13151180.v1.

ETHICS STATEMENT

Ethical review and approval was not required for the animal study because of Austrian legislative provisions. For a detailed statement describing the regulations and recommendations given by the legislative authorities (Austrian Federal Ministry of Education, Science and Research and University of Veterinary Medicine, Vienna) see section "Materials and Methods."

AUTHOR CONTRIBUTIONS

MT and JW conceived the study. BT, MT, TS, and JW designed the experiments. AR and AT carried out the experiments under the supervision of BT and JW. TS was responsible for the processing of activity data. AR, AT, and YP were responsible for laboratory processing of the eDNA samples under the supervision of BT who also carried out statistical analysis and wrote the first draft of the manuscript which was revised by all co-authors. All authors contributed to the article and approved the submitted version.

FUNDING

This research was conducted within the eDNA-Alpfish project funded by the Austrian Research Promotion Agency (FFG); project number 853219, and this open access publication was co-funded by the University of Innsbruck.

ACKNOWLEDGMENTS

We thank R. Vogt for his support during the experiment, M. Böcker for assistance with the background literature, J. Harvie for input on the statistical analysis, and C. Moritz and D. Kirschner for their help in obtaining the *C. gobio* individuals. We also thank two reviewers for their extensive and constructive feedback on the original manuscript. This manuscript has been uploaded as preprint to bioRxiv https://doi.org/10.1101/2020.10.28.359653 (Thalinger et al., 2020).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.623718/full#supplementary-material

REFERENCES

Ángeles Esteban, M. (2012). An overview of the immunological defenses in fish skin. *Isrn Immunol*. 2012:853470. doi: 10.5402/2012/853470

Auguie, B. (2017). *gridExtra: Miscellaneous Functions for "Grid" Graphics*. Available online at: https://cran.r-project.org/package=gridExtra (accessed September 15, 2020).

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

Barton, K. (2019). *MuMIn: Multi-Model Inference*. Available online at: https://cran.r-project.org/package=MuMIn doi: 10.5402/2012/853470 (accessed September 15, 2020).

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48.doi: 10.18637/jss.v067.i01

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., et al. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol.* 24, 127–135. doi: 10.1016/j.tree.2008.10.008

Brett, J. R., and Groves, T. D. D. (1979). "Physiological energetics," in *Fish Physiology, Bioenergetics and Growt*, Vol. 8, eds W. S. Hoar, D. J. Randall, and J. R. Brett (New York, NY: Academic Press).

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., and West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology* 85, 1771–1789.

Burnham, K. P., and Anderson, D. R. (2002). *Model Selection and Multimodel Inference*, 2nd Edn. (New York, NY: Springer). doi: 10.1007/b97636

Bylemans, J., Furlan, E. M., Hardy, C. M., McGuffie, P., Lintermans, M., and Gleeson, D. M. (2017). An environmental DNA-based method for monitoring spawning activity: a case study, using the endangered Macquarie perch (*Macquaria australasica*). *Methods Ecol. Evol.* 8, 646–655. doi: 10.1111/2041-210X.12709

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., et al. (2017). Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Mol. Ecol.* 26, 5872–5895. doi: 10.1111/mec.14350

Doi, H., Uchii, K., Takahara, T., Matsuhashi, S., Yamanaka, H., and Minamoto, T. (2015). Use of droplet digital PCR for estimation of fish abundance and biomass in environmental DNA surveys. *PLoS One* 10:e0122763. doi: 10.1371/journal.pone.0122763

Evans, N. T., Olds, B. P., Renshaw, M. A., Turner, C. R., Li, Y., Jerde, C. L., et al. (2016). Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Mol. Ecol. Resour.* 16, 29–41. doi: 10.1111/1755-0998.12433

Evans, N. T., Shirey, P. D., Wieringa, J. G., Mahon, A. R., and Lamberti, G. A. (2017). Comparative cost and effort of fish distribution detection via environmental DNA analysis and electrofishing. *Fisheries* 42, 90–99. doi: 10.1080/03632415.2017.1276329

Faraway, J. J. (2016). *Extending the Linear Model with R. Generalized Linear, Mixed Effects and Nonparametric Regression Models*, 2nd Edn. (New York, NY: Chapman and Hall). doi: 10.1201/9781315382722

Forstner, H. (1983). "An automated multiple-chamber intermittent-flow respirometer," in *Polarographic Oxygen Sensors*, eds D. E. Gnaiger and D. H. Forstner (Berlin: Springer), 111–126. doi: 10.1007/978-3-642-81863-9_12

Freyhof, J., and Kottelat, M. (2007). *Handbook of European Freshwater Fishes*. Berlin: Springer.

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., et al. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods Ecol. Evol.* 7, 1299–1307. doi: 10.1111/2041-210X.12595

Harrison, J. B., Sunday, J. M., and Rogers, S. M. (2019). Predicting the fate of eDNA in the environment and implications for studying biodiversity. *Proc. R. Soc. B Biol. Sci.* 286:20191409. doi: 10.1098/rspb.2019.1409

Hartig, F. (2020). *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models*. Available online at: https://cran.r-project.org/package=DHARMa (accessed January 5, 2021).

Helfman, G. S. (1986). "Fish behaviour by day, night and twilight," in *The Behaviour of Teleost Fishes*, ed. T. J. Pitcher (Boston, MA: Springer), 366–387. doi: 10.1007/978-1-4684-8261-4_14

Horiuchi, T., Masuda, R., Murakami, H., Yamamoto, S., and Minamoto, T. (2019). Biomass-dependent emission of environmental DNA in jack mackerel *Trachurus japonicus* juveniles. *J. Fish Biol.* 95:jfb.14095. doi: 10.1111/jfb.14095

Huerlimann, R., Cooper, M. K., Edmunds, R. C., Villacorta-Rath, C., Le Port, A., Robson, H. L. A., et al. (2020). Enhancing tropical conservation and ecology research with aquatic environmental DNA methods: an introduction for non-environmental DNA specialists. *Anim. Conserv.* 23, 632–645. doi: 10.1111/acv.12583

Jo, T., Arimoto, M., Murakami, H., Masuda, R., and Minamoto, T. (2020). Estimating shedding and decay rates of environmental nuclear DNA with relation to water temperature and biomass. *Environ. DNA* 2, 140–151. doi: 10.1002/edn3.51

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

Johnston, I. A., Camm, J. P., and White, M. (1988). Specialisations of swimming muscles in the pelagic antarctic fish *Pleuragramma antarcticum*. *Mar. Biol.* 100, 3–12. doi: 10.1007/BF00392949

Kassambara, A. (2019). *ggpubr: "ggplot2" Based Publication Ready Plots*. Available online at: https://cran.r-project.org/package=ggpubr (accessed September 15, 2020).

Killen, S. S., Atkinson, D., and Glazier, D. S. (2010). The intraspecific scaling of metabolic rate with body mass in fishes depends on lifestyle and temperature. *Ecol. Lett.* 13, 184–193. doi: 10.1111/j.1461-0248.2009.01415.x

Klymus, K. E., Richter, C. A., Chapman, D. C., and Paukert, C. (2015). Quantification of eDNA shedding rates from invasive bighead carp *Hypophthalmichthys nobilis* and silver carp *Hypophthalmichthys molitrix*. *Biol. Conserv.* 183, 77–84. doi: 10.1016/j.biocon.2014.11.020

Lacoursière-Roussel, A., Rosabal, M., and Bernatchez, L. (2016). Estimating fish abundance and biomass from eDNA concentrations: variability among capture methods and environmental conditions. *Mol. Ecol. Resour.* 16, 1401–1414. doi: 10.1111/1755-0998.12522

Leese, F., Altermatt, F., Bouchez, A., Ekrem, T., Hering, D., Meissner, K., et al. (2016). DNAqua-Net: developing new genetic tools for bioassessment and monitoring of aquatic ecosystems in Europe. *Res. Ideas Outcomes* 2:e11321. doi: 10.3897/rio.2.e11321

Littlefair, J. E., Hrenchuk, L. E., Blanchfield, P. J., Rennie, M. D., and Cristescu, M. E. (2020). Thermal stratification and fish thermal preference explain vertical eDNA distributions in lakes. *Mol. Ecol.* 1–14. doi: 10.1111/mec.15623

Lüdecke, D. (2020). *sjPlot: Data Visualization for Statistics in Social Science*. Available online at: https://cran.r-project.org/package=sjPlot (accessed September 15, 2020).

Maruyama, A., Nakamura, K., Yamanaka, H., Kondoh, M., and Minamoto, T. (2014). The release rate of environmental DNA from juvenile and adult fish. *PLoS One* 9:e0114639. doi: 10.1371/journal.pone.0114639

Mazerolle, M. J. (2020). *AICcmodavg: Model Selection and Multimodel Inference Based on (Q)AIC(c)*. Available online at: https://cran.r-project.org/package=AICcmodavg (accessed September 15, 2020).

McColl-Gausden, E., Weeks, A., Coleman, R., Robinson, K., Song, S., Raadik, T., et al. (2020). Multi-species models reveal that eDNA metabarcoding is more sensitive than backpack electrofishing for conducting fish surveys in freshwater streams. *Mol. Ecol.* 1–16. doi: 10.1111/mec.15644

Merkes, C. M., McCalla, S. G., Jensen, N. R., Gaikowski, M. P., and Amberg, J. J. (2014). Persistence of DNA in carcasses, slime and avian feces may affect interpretation of environmental DNA data. *PLoS One* 9:e113346. doi: 10.1371/journal.pone.0113346

Metcalfe, N. B., Van Leeuwen, T. E., and Killen, S. S. (2016). Does individual variation in metabolic phenotype predict fish behaviour and performance? *J. Fish Biol.* 88, 298–321. doi: 10.1111/jfb.12699

Minamoto, T., Miya, M., Sado, T., Seino, S., Doi, H., Kondoh, M., et al. (2020). An illustrated manual for environmental DNA research: water sampling guidelines and experimental protocols. *Environ. DNA* 3, 8–13. doi: 10.1002/edn3.121

Mizumoto, H., Urabe, H., Kanbe, T., Fukushima, M., and Araki, H. (2018). Establishing an environmental DNA method to detect and estimate the biomass of Sakhalin taimen, a critically endangered Asian salmonid. *Limnology* 19, 219–227. doi: 10.1007/s10201-017-0535-x

Muus, B. J., and Dahlström, P. (1968). *Süßwasserfische*. Munich: BLV Verlagsgesellschaft.

Nakagawa, S., and Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol. Evol.* 4, 133–142. doi: 10.1111/j.2041-210x.2012.00261.x

O'Shea, B., Mordue-Luntz, A. J., Fryer, R. J., Pert, C. C., and Bricknell, I. R. (2006). Determination of the surface area of a fish. *J. Fish Dis.* 29, 437–440. doi: 10.1111/j.1365-2761.2006.00728.x

Pilliod, D. S., Laramie, M. B., MacCoy, D., and Maclean, S. (2019). Integration of eDNA-Based Biological Monitoring within the U.S. Geological Survey's National Streamgage Network. *JAWRA J. Am. Water Resour. Assoc.* 55, 1505–1518. doi: 10.1111/1752-1688.12800

Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., et al. (2018). Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Sci. Rep.* 8:10361. doi: 10.1038/s41598-018-28424-8

Powell, M. J. D. (2009). *The BOBYQA Algorithm for Bound Constrained Optimization Without Derivatives. Rep. No. DAMTP 2009/NA06*. Available online at: http://www.damtp.cam.ac.uk/user/na/NA_papers/NA2009_06.pdf (accessed January 5, 2021).

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Available online at: https://www.r-project.org/ (accessed September 15, 2020).

Roberts, J. L. (1975). Active branchial and ram gill ventilation in fishes. *Biol. Bull.* 148, 85–105. doi: 10.2307/1540652

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., et al. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* 18:529. doi: 10.1186/s12859-017-1934-z

Sassoubre, L. M., Yamahara, K. M., Gardner, L. D., Block, B. A., and Boehm, A. B. (2016). Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. *Environ. Sci. Technol.* 50, 10456–10464. doi: 10.1021/acs.est.6b03114

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. doi: 10.1038/nmeth.2019

Shogren, A. J., Tank, J. L., Andruszkiewicz, E., Olds, B., Mahon, A. R., Jerde, C. L., et al. (2017). Controls on eDNA movement in streams: transport, Retention, and Resuspension. *Sci. Rep.* 7:5065. doi: 10.1038/s41598-017-05223-1

Sigsgaard, E. E., Carl, H., Møller, P. R., and Thomsen, P. F. (2015). Monitoring the near-extinct European weather loach in Denmark based on environmental DNA from water samples. *Biol. Conserv.* 183, 46–52. doi: 10.1016/j.biocon.2014.11.023

Spindler, T. (1997). *Fischfauna in Osterreich. Ökologie – Gefahrdung – Bioindikation – Fischerei – Gesetzgebung*. Vienna: Umweltbundesamt.

Strickler, K. M., Fremier, A. K., and Goldberg, C. S. (2015). Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biol. Conserv.* 183, 85–92. doi: 10.1016/j.biocon.2014.11.038

Svendsen, M. B. S., Bushnell, P. G., and Steffensen, J. F. (2016). Design and setup of intermittent-flow respirometry system for aquatic organisms. *J. Fish Biol.* 88, 26–50. doi: 10.1111/jfb.12797

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

Takeuchi, A., Iijima, T., Kakuzen, W., Watanabe, S., Yamada, Y., Okamura, A., et al. (2019). Release of eDNA by different life history stages and during spawning activities of laboratory-reared Japanese eels for interpretation of oceanic survey data. *Sci. Rep.* 9:6074. doi: 10.1038/s41598-019-42641-9

Thalinger, B., Rieder, A., Teuffenbach, A., Putz, Y., Schwerte, T., Wanzenböck, J., et al. (2020). The effect of activity, energy use, and species identity on environmental DNA shedding of freshwater fish. *bioRxiv* [Preprint].doi: 10.1101/2020.10.28.359653

Thalinger, B., Deiner, K., Harper, L., Rees, H., Blackman, R., Sint, D., et al. (2021a). A validation scale to determine the readiness of environmental DNA assays for routine species monitoring. *bioRxiv* [Preprint]. doi: 10.1101/2020.04.27.063990

Thalinger, B., Kirschner, D., Pütz, Y., Moritz, C., Schwarzenberger, R., Wanzenböck, J., et al. (2021b). Lateral and longitudinal fish environmental DNA distribution in dynamic riverine habitats. *Environ. DNA* 3, 305–318. doi: 10.1002/edn3.171

Thalinger, B., Oehm, J., Mayr, H., Obwexer, A., Zeisler, C., and Traugott, M. (2016). Molecular prey identification in Central European piscivores. *Mol. Ecol. Resour.* 16, 123–137. doi: 10.1111/1755-0998.12436

Tsuji, S., Ushio, M., Sakurai, S., Minamoto, T., and Yamanaka, H. (2017). Water temperature-dependent degradation of environmental DNA and its relation to bacterial abundance. *PLoS One* 12:e0176608. doi: 10.1371/journal.pone.0176608

Vanni, M. J., and McIntyre, P. B. (2016). Predicting nutrient excretion of aquatic animals with metabolic ecology and ecological stoichiometry: a global synthesis. *Ecology* 97, 3460–3471. doi: 10.1002/ecy.1582

Wegner, N. C., Sepulveda, C. A., Bull, K. B., and Graham, J. B. (2009). Gill morphometrics in relation to gas transfer and ram ventilation in high-energy demand teleosts: scombrids and billfishes. *J. Morphol.* 271, 36–49. doi: 10.1002/jmor.10777

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

Wilcox, T. M., McKelvey, K. S., Young, M. K., Sepulveda, A. J., Shepard, B. B., Jane, S. F., et al. (2016). Understanding environmental DNA detection probabilities: a case study using a stream-dwelling char *Salvelinus fontinalis*. *Biol. Conserv.* 194, 209–216. doi: 10.1016/j.biocon.2015.12.023

Yates, M. C., Glaser, D. M., Post, J. R., Cristescu, M. E., Fraser, D. J., and Derry, A. M. (2020). The relationship between eDNA particle concentration and organism abundance in nature is strengthened by allometric scaling. *Mol. Ecol.* 1–15. doi: 10.1111/mec.15543

Zhang, D. (2020). *rsq: R-Squared and Related Measures*. Available online at: https://cran.r-project.org/package=rsq (accessed January 5, 2021).

Conflict of Interest: MT is the co-founder of Sinsoma GmbH, a for profit company dedicated to DNA analyses in environmental studies.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 15 March 2021
doi: 10.3389/fevo.2021.574877

Exploring the Use of Environmental DNA (eDNA) to Detect Animal Taxa in the Mesopelagic Zone

Annette F. Govindarajan[1], Rene D. Francolini[1], J. Michael Jech[2], Andone C. Lavery[3], Joel K. Llopiz[1], Peter H. Wiebe[1] and Weifeng (Gordon) Zhang[3]*

[1] Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, United States, [2] National Oceanic and Atmospheric Administration (NOAA), Northeast Fisheries Science Center, Woods Hole, MA, United States, [3] Applied Ocean Physics and Engineering Department, Woods Hole Oceanographic Institution, Woods Hole, MA, United States

Animal biodiversity in the ocean's vast mesopelagic zone is relatively poorly studied due to technological and logistical challenges. Environmental DNA (eDNA) analyses show great promise for efficiently characterizing biodiversity and could provide new insight into the presence of mesopelagic species, including those that are missed by traditional net sampling. Here, we explore the utility of eDNA for identifying animal taxa. We describe the results from an August 2018 cruise in Slope Water off the northeast United States. Samples for eDNA analysis were collected using Niskin bottles during five CTD casts. Sampling depths along each cast were selected based on the presence of biomass as indicated by the shipboard Simrad EK60 echosounder. Metabarcoding of the 18S V9 gene region was used to assess taxonomic diversity. eDNA metabarcoding results were compared with those from net-collected (MOCNESS) plankton samples. We found that the MOCNESS sampling recovered more animal taxa, but the number of taxa detected per liter of water sampled was significantly higher in the eDNA samples. eDNA was especially useful for detecting delicate gelatinous animals which are undersampled by nets. We also detected eDNA changes in community composition with depth, but not with sample collection time (day vs. night). We provide recommendations for applying eDNA-based methods in the mesopelagic including the need for studies enabling interpretation of eDNA signals and improvement of barcode reference databases.

Keywords: environmental DNA, mesopelagic, biodiversity, metabarcoding, zooplankton

OPEN ACCESS

Edited by:
Hiroki Yamanaka, Ryukoku University, Japan

Reviewed by:
Bettina Thalinger, University of Guelph, Canada
Judith Van Bleijswijk, Royal Netherlands Institute for Sea Research (NIOZ), Netherlands

****Correspondence:***
Annette F. Govindarajan
afrese@whoi.edu

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *23 June 2020*
Accepted: *22 February 2021*
Published: *15 March 2021*

Citation:
Govindarajan AF, Francolini RD, Jech JM, Lavery AC, Llopiz JK, Wiebe PH and Zhang W (2021) Exploring the Use of Environmental DNA (eDNA) to Detect Animal Taxa in the Mesopelagic Zone. Front. Ecol. Evol. 9:574877. doi: 10.3389/fevo.2021.574877

INTRODUCTION

The ocean's mesopelagic zone is poorly explored, in large part due to the vastness of the habitat and the technological and logistical challenges in accessing it. Recently, it has been discovered that mesopelagic biomass is significantly greater than previously thought (Irigoien et al., 2014; St. John et al., 2016); however, there is significant uncertainty about the composition of mesopelagic biomass. Deep pelagic waters, including the mesopelagic, likely contain numerous undescribed species (Robison, 2004, 2009). Traditional sampling nets may miss important taxa such as delicate gelatinous species that fall apart when collected, or fish that avoid capture altogether. A recent genetic study suggested that species diversity is significantly underestimated (Sommer et al., 2017), and genetic analysis has shown that many nominal species may consist of multiple cryptic species, sometimes even belonging to different families (Lindsay et al., 2017).

Mesopelagic species are under immediate risk for exploitation as it is a potential source of fish meal and nutraceuticals (St. John et al., 2016; Hidalgo and Browman, 2019), but our lack of knowledge about their distribution, life history, and ecology impedes sustainable management (Webb et al., 2010; St. John et al., 2016; Glover et al., 2018; Hidalgo and Browman, 2019). The consequences of overharvesting are potentially severe and global in nature. Mesopelagic fish and invertebrates play a key role in the biological carbon pump, which transfers carbon from surface production to the deep sea where it is sequestered (Giering et al., 2014). Anthropogenic changes to mesopelagic biodiversity and biomass could alter this process. Thus, it is urgent to improve the understanding of the composition, distribution, and abundance of mesopelagic fauna before irreparable anthropogenically induced changes occur (St. John et al., 2016; Glover et al., 2018; Hidalgo and Browman, 2019).

New tools are emerging that will facilitate scientific study of this region. One of these is the analysis of environmental DNA (eDNA), which is DNA that is present in the environment that, for metazoans, is no longer associated with the organisms from which it originated (reviewed in Ruppert et al., 2019). In eDNA analyses, water is collected and filtered, the eDNA is extracted from the filter, and one or more DNA markers are analyzed, typically through qPCR, or metabarcoding. There are many potential benefits of eDNA animal biodiversity assessments relative to traditional assessments, including cost effectiveness (Sigsgaard et al., 2015; Thomsen and Willerslev, 2015; Evans et al., 2017) and improved detection of species that are rare or are difficult to sample by traditional means (Dejean et al., 2012; Piaggio et al., 2014; Kelly et al., 2017; Closek et al., 2019; Stat et al., 2019).

The use of eDNA analyses is still in its infancy and there is no standard protocol that is universally applicable to any given system (Deiner et al., 2017; Ruppert et al., 2019). eDNA sampling at meaningful spatial and temporal scales in mesopelagic waters, as opposed to coastal and freshwater systems where most eDNA work has been done, may be especially challenging due to the enormity of the environment and the difficulty in accessing it. In marine coastal regions, samples are taken using Niskin bottles, often mounted on a CTD rosette, and sample volumes are typically around one liter (Andruszkiewicz et al., 2017b; Djurhuus et al., 2017). For the mesopelagic, an adaptive sampling approach that targets sampling locations based on remotely sensed information from shipboard acoustics could be beneficial for targeting concentrations of mesopelagic organisms such as in deep scattering layers (Kinzer, 1969; Orlowski, 1990; Hazen and Johnston, 2010; D'Elia et al., 2016). Multi-frequency and broadband acoustic approaches would detect the greatest variety of organisms over a range of depths (Lavery et al., 2007; Davison et al., 2015; Bassett et al., 2020), as acoustic scattering by organisms is frequency-dependent and frequencies vary in the depth that they penetrate.

The goal of this work was to explore the collection and utility of eDNA for surveying and characterizing metazoan biodiversity in the mesopelagic zone, in order to better understand the composition of this region's biomass. We used an adaptive sampling approach based on acoustic backscatter (18, 38, 120, and 200 kHz frequencies) to target layers of organisms in daytime and nighttime depth-stratified sampling of eDNA (via Niskin bottles) and zooplankton (via MOCNESS nets). We conducted DNA metabarcoding on all samples using the 18S V9 barcode gene (Amaral-Zettler et al., 2009) to enable detection of a broad range of taxa above the species-level (i.e., genus or family), shedding light on the efficiency of both approaches for detecting animal taxa.

MATERIALS AND METHODS

Environmental DNA Sample Collection and Shipboard Processing

Seawater samples were collected during a cruise on the NOAA Ship *Henry B. Bigelow* between the 2,000 and 3,000 m isobaths off the shelf break south of the island of Martha's Vineyard in Massachusetts, United States (**Figure 1**). Samples were collected during casts of a Seabird 911 plus CTD mounted on a rosette frame with eight 5 l Niskin bottles. Niskin bottles were triggered to sample based on the layers identified in real time with the shipboard Simrad EK60 echosounder whose split-beam transducers (11° beam width for the 18 kHz and 7° beam width for the 38, 120, and 200 kHz) were mounted on a retractable keel on the vessel.

Three daytime and two nighttime CTD casts were completed (avoiding crepuscular periods when animals may be vertically migrating), yielding a total of 40 eDNA samples (**Table 1**) and consisting of 8 samples per cast. Niskin bottles were not cleaned between casts as the bottles were deployed open and thus exposed to the entire water column during the downcast, until they were triggered to close at the target depth on the upcast. In the first four casts, one sample was taken at each of 8 depths. On the final cast (Cast 10), two replicate samples were taken at each of 4 depths. Diel vertical migration was observed as the vertical shift in the position of the layers on the echosounder, and our water sampling for each cast coincided with these layers (**Figure 2** and **Supplementary Figures S1–S4**). In all of the CTD casts, the water column was stratified with clines in density, temperature, and salinity between approximately 25 and 50 m below the surface (**Figure 3** and **Supplementary Figures S5, S6**). Dissolved oxygen peaked in this region, before declining at the depth around 50–70 m. Oxygen minimums were found around approximately 250 m below surface in all casts (**Supplementary Figures S5, S6**).

Water from the Niskin bottles was filtered in a temperature-controlled cold room (maintained at 13–15°C) on the ship and began immediately upon retrieval of the CTD-rosette. The work area in the cold room was wiped down daily with bleach, followed by several rinsing wipes with Milli-Q water. Nitrile gloves were worn throughout the filtering protocol and were changed frequently. Three samples were processed at a time. Bottles were removed from the rosette, brought into the processing room, and secured on custom-built stands. The remaining bottles were wrapped in ice packs until they could be processed.

Three peristaltic pumps were used to pump water from the Niskin bottles through sterile encapsulated, single-use 0.2 μm PES Sterivex filters. The tubing was first cleaned by pumping

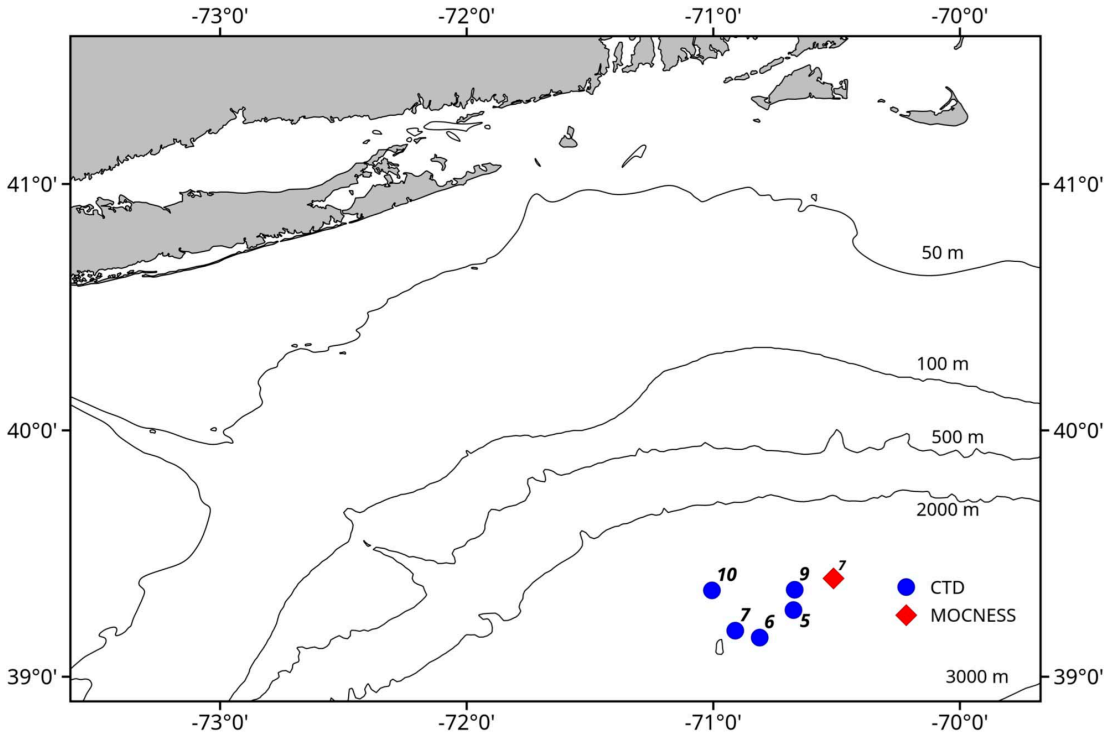


FIGURE 1 | Map of study area. Blue circles indicate locations of the CTD casts that sampled eDNA and the red diamond indicates the location of the MOCNESS tow.

with a minimum of 200 ml of Milli-Q water and then flushed with a minimum of 200 ml of sample water. The Sterivex filters were attached after cleaning. The remaining Niskin bottle water was filtered through the Sterivex and the volume of the flow-through was collected and measured (**Table 1**). Once filtration was complete, the Sterivex filter was removed, the ends of the filter were sealed with parafilm, the filter was placed in a sterile 50 ml Falcon tube, and the tube was stored at −80°C. A control sample was filtered along with each cast. The control samples consisted of approximately 4.8 l of Milli-Q water dispensed into pre-sterilized, single-use Whirlpak sample bags and pumped through Sterivex filters.

Zooplankton Sample Collection and Shipboard Processing

A 1 m^2 MOCNESS (Multiple opening and closing net and environmental sensing system; Wiebe et al., 1985) zooplankton tow was taken for comparison with the eDNA data from the CTD casts. The MOCNESS was equipped with nine nets (333 μm mesh), SeaBird temperature and conductivity probes, a pressure sensor, and a flowmeter. Eight of the MOCNESS nets sampled at discrete depth intervals ranging from the base of the mesopelagic zone (1,000 m) to the surface, and an additional net sampled the integrated water column (0–1,000 m). The MOCNESS tow was taken in the vicinity of the CTD casts (**Figure 1**). Temperature and salinity data were consistent with those from the CTD casts, indicating that MOCNESS and CTD sampled the same water mass (**Figure 3**). The nine nets sampled the following depth intervals: 0–1,000, 1,000–800, 800–600, 600–400, 400–200, 200–100, 100–50, 50–24, and 24–0 m, and tow sample volumes ranged from approximately 149 to 5,208 m^3 (**Table 2**). Upon retrieval, each net was washed and the cod end buckets were removed. The zooplankton from each net were split into four equal parts using a Folsom plankton splitter (McEwen et al., 1954). One of these splits was preserved in 95% ethanol for metabarcoding analysis. The ethanol preservative was replaced with fresh 95% ethanol approximately 24 h after collection.

Environmental DNA Extraction and Sequencing

Upon return to the laboratory, genomic DNA and controls were extracted using DNeasy Power Water Sterivex Kits (Qiagen). DNA extractions were performed in a dedicated, pre-amplification DNA work area. The work area was cleaned before use with 10% bleach followed by several rinses with Milli-Q water. Next Generation library preparation and sequencing were conducted at the Center for Genome Innovation at the University of Connecticut (Storrs, CT). Each extract was normalized to 5 ng/μl and amplified in duplicate using 2× KAPA HiFi HotStart ReadyMix (Kapa Biosciences, Wilmington, MA) following the 16S Metagenomic Sequencing Library Preparation guide (Illumina, San Diego, CA). The PCR was set up in a PCR hood and the surface area was first cleaned with 70% bleach and additionally decontaminated using UV light, and only sterile pipette tips with embedded aerosol filters were used for pipetting. 2.5 μl of sample was added to each PCR. Per sample and

TABLE 1 | Environmental DNA sampling summary.

Cast	Date	Deployment time (EDT)	Latitude	Longitude	Time of day	Sample depths (m)	Volume filtered (milliliters)
5	2018/08/14	11:53:59	39.2708	−70.6751	Day	30	3,580
						65	5,150
						200	5,230
						370	4,890
						460	5,180
						530	5,190
						610	5,220
						800	5,270
6	2018/08/15	02:49:59	39.1591	−70.8121	Night	25	5,440
						65	5,220
						150	5,350
						240	5,410
						535	5,160
						640	5,190
						700	5,110
						800	4,570
7	2018/08/15	14:49:59	39.1873	−70.9112	Day	15	5,360
						70	5,130
						325	5,390
						410	5,500
						530	5,220
						645	5,180
						725	5,230
						800	5,270
9	2018/08/16	01:20:59	39.3533	−70.6701	Night	15	5,060
						50	4,030
						220	5,430
						465	5,450
						550	5,180
						600	5,190
						640	5,230
						775	5,080
10	2918/08/19	11:19:52	39.3508	−71.0058	Day	375-1	5420
						375-2	5200
						450-1	5420
						450-2	5450
						585-1	4810
						585-2	5240
						800-1	5230
						800-2	5150

While the Niskin bottles were nominally 5 l bottles, the actual volume of seawater contained was greater than 5 l. In a small number of cases, less than 5 l was filtered due to apparent filter clogging.

filtration control two PCR replicates were run and pooled after the first PCR. The primers used were the 1380F and 1510R V9 primers (Amaral-Zettler et al., 2009) with Illumina adapters. The primer sequences with their adapters (in bold) were: 1380F 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCCTGC CHTTTGTACACAC-3′ and 1510R 5′-GTCTCGTGGGCTCGG AGATGTGTATAAGAGACAGCCTTCYGCAGGTTCACCTAC-3′. The PCR conditions for the first PCR were: 95°C for 3 min; 25 cycles of: 95°C for 30 s, 55°C for 30 s, 72°C for 30 s; 72°C for 5 min. Amplification success was assessed on the pooled PCRs using an Agilent 4200 TapeStation electrophoresis system using the High Sensitivity DNA D1000 assay (Agilent Technologies, Santa Clara, CA) and the products were purified using Agencourt AMPure XP Beads (Beckman Coulter, Brea, CA).

Index barcode sequences were incorporated into the purified amplicons by a second PCR using NexteraXT Index Kit v2 Sets C and D (Illumina, San Diego, CA) and 2× KAPA HiFi HotStart ReadyMix (Kapa Biosciences, Wilmington, MA). The second PCR protocol was: 95°C for 3 min; 8 cycles of: 95°C for 30 s, 55°C for 30 s, 72°C for 30 s; and 1 cycle of 72°C for 5 min. The indexed PCR products were purified using AMPure XP Beads (Agencourt) and assessed for quality and adapter removal using an Agilent 4200 TapeStation electrophoresis system (High Sensitivity DNA D1000 assay). Libraries were quantified using a Qubit 3.0 fluorometer (Life Technologies, Carlsbad, CA), normalized, pooled, and denatured according to Illumina MiSeq sample preparation for sequencing. 20–30% PhiX (Illumina, San Diego, CA) was added to our amplicons before running on the Illumina MiSeq using the 500 cycle v2 reagent kit. Our targeted sequencing depth was 200,000 reads per sample. The negative control samples were also sequenced. For pooling the negative controls (which did not contain detectable amounts of DNA), we added the maximum volume that was used for the samples.

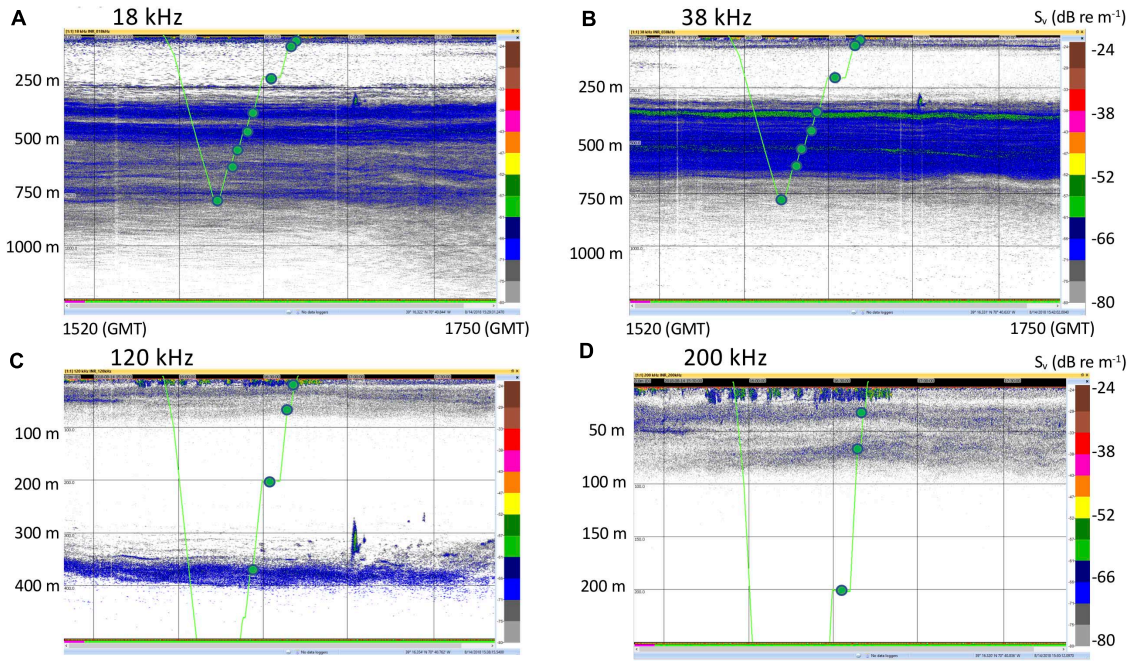


FIGURE 2 | Representative echosounder images with noise reduction applied, and the CTD track and sampling locations overlain in green. **(A)** 18 kHz. **(B)** 38 kHz. **(C)** 120 kHz. **(D)** 200 kHz. These images are for Cast 5 on 14 August 2018. Images for the rest of the casts are in **Supplementary Figures S1–S4**. Time stamps (GMT) indicate the start and end of the echograms and vertical bars are at 30 min intervals. Note change of vertical scale for the 120 kHz (500 m) and 200 kHz (250 m) echograms.

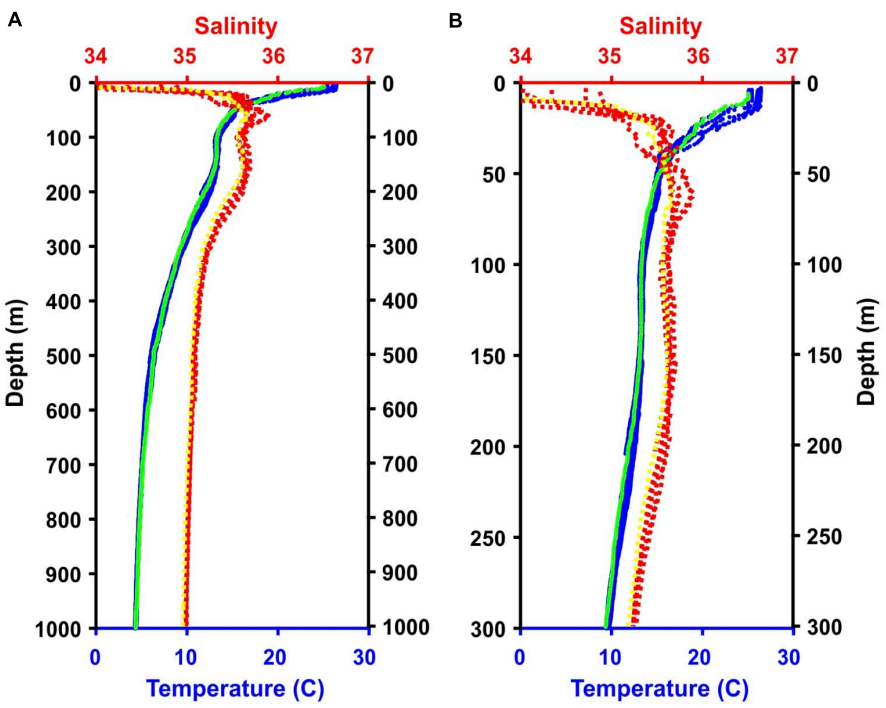


FIGURE 3 | Temperature and salinity data from the five CTD casts and the MOCNESS tow. Note, lines for all casts and the tow are present although they are not always visible because they overlap. **(A)** Data from 0 to 1,000 m. **(B)** Data from the top 300 m of the water column. Blue = CTD temperature data; Green = MOCNESS temperature data; Red = CTD salinity data; Yellow = MOCNESS salinity data.

TABLE 2 | MOCNESS tow summary.

Date	Time start/end (EDT)	Latitude (N) start/end	Longitude (W) start/end	Net	Depths sampled (m)	Volume filtered (m^3)
8/15/18	19:38:27	39.3993	−70.5127	0	0–998.7	5,208.2
	22:46:08	39.3578	−70.6418	1	998.7–798.4	1,283.4
				2	798.4–599.7	1,644.4
				3	599.7–400.3	1,306.9
				4	400.3–199.6	1,217.3
				5	199.6–99.4	764.1
				6	99.4–49.8	342.7
				7	49.8–23.6	149.2
				8	23.6–0	277.5

Date and time are based on local time (EDT). Total volume filtered was 12,193.7 m^3 (=12,193,700 L).

Zooplankton DNA Extraction and Sequencing

Zooplankton samples were sorted into small (333–1,000 μm) and large (>1,000 μm) size fractions before extraction in order to maximize the detection of low-biomass organisms and the wet weight of each was recorded (**Supplementary Table S1**). The samples were poured through stacked 150 and 1,000 μm sieves and the samples were rinsed with Milli-Q water, which also washed away any remaining ethanol. When present, exceptionally large animals or fragments were picked out of the 1,000 μm sieve and preserved for later individual (Sanger) sequencing. Sieve contents were rinsed into pre-weighed, sterile 50 ml Falcon tubes using as little water as possible and concentrated by spinning 1 min at 1,000 rpm. Residual water overlying the zooplankton was pipetted away.

Zooplankton samples were homogenized using a Benchmark D1000 homogenizer with a 10 mm sawtooth probe until the Falcon tube contents were visibly smooth and well-mixed. The homogenized tubes were centrifuged for 2 min at 5,000*g* and excess water overlying the pellets was decanted off. All tubes were weighed. The homogenizer was cleaned between samples by running it for 20 s in a 10% bleach solution followed by 3 additional runs in separate tubes of Milli Q water.

Two hundred mg of zooplankton biomass was transferred from each tube to a sterile 15 ml Falcon tube and DNA was extracted using DNeasy Blood and Tissue DNA extraction kits, using a slight modification of the manufacturer's protocol to accommodate the relatively large tissue biomass in the samples. Specifically, 1,800 μl of Buffer ATL and 200 μl of proteinase K were added to each tube, and tubes were incubated for 3 h at 56°C and vortexed periodically. After incubation, 200 μl of lysate was transferred to a new, sterile 1.5 ml tube. At that point, the Qiagen manufacturer's protocol was followed exactly, with a final elution into 200 μl Buffer AE. Aliquots of the extracted DNA were sent to the Center for Genomic Innovation at the University of Connecticut for library preparation and sequencing, following the same protocol as for the eDNA samples.

Genomic DNA of the large individuals that were set aside was extracted using the DNeasy Blood and Tissue Kit (Qiagen) following the manufacturer's protocol. The 18S V9 gene was amplified using the V9 primers (Amaral-Zettler et al., 2009) under the following conditions: 95°C for 3 min; 35 cycles of: 95°C for 30 s, 55°C for 30 s, 72°C for 30 s; and 1 cycle of 72°C for 5 min. PCR products were amplified on a 1.2% agarose gel and visualized with GelRed. No additional attempts were made to amplify samples that failed the first time, in order to be consistent with the metabarcoding protocol. Amplicons were purified with QiaQuick PCR Purification kits (Qiagen) according to the manufacturer's protocol, quantified using a Nanodrop, and sequenced in both directions at Eurofins Genomics[1].

Bioinformatics

The same bioinformatics workflow was applied to both the eDNA and the MOCNESS zooplankton samples, although each group was processed separately as they were sequenced on separate runs and thus differed in their sequencing error profiles. Demultiplexed paired-end sequence reads were processed using Quantitative Insights Into Microbial Ecology 2 (QIIME2) version 2019.1 (Sigsgaard et al., 2015; Thomsen and Willerslev, 2015; Bolyen et al., 2018). Forward primer sequences were trimmed, and forward and reverse reads were truncated based on quality plots of the data after base pair 120. Reverse primers were removed (if present) using the Cutadapt (Martin, 2011) Qiime2 plugin (v. 2020.11.1). DADA2 (Callahan et al., 2016) implemented through QIIME2 performed sequence quality control including error correction and chimera removal, and paired the reads to generate unique (i.e., 100% similarity) amplicon sequence variants (ASVs). The control samples contained very few reads (e.g., 0.3% of the metazoan-only dataset). For each ASV in the dataset that was present in both the samples and controls, the maximum number of reads found in any control was subtracted from every sample. ASVs that had a frequency of less than 10 (summed across all 40 samples) were deleted from the dataset.

Taxonomy was assigned using a naïve Bayesian classifier (Bokulich et al., 2018) that was trained on the Silva v. 132 99% database (Quast et al., 2013) using 18S sequences only for the 18S V9 gene region. The resulting dataset was filtered to generate a metazoan-only dataset that provided a high-level taxonomic classification for a broad range of animal species (Level 5 on the Silva database classification). We further resolved the most abundant categories based on Level 6 of the Silva database. Taxa

[1]https://www.eurofinsgenomics.com

comparisons between eDNA and MOCNESS were based on Level 7 (the deepest level) of the Silva database.

We also searched all ASVs against the GenBank nr database. Blast search parameters were set to exclude uncultured, unfiltered, and unclassified entries. We reported top hits when the percent identity between our ASV and the Genbank entry was 97% or greater. This threshold has been used in other studies (Pearman et al., 2014; Casas et al., 2017) and is effective in distinguishing genera and families in copepods (Wu et al., 2015). Note that even a perfect (100%) match does not necessarily indicate a species-level identification (Wu et al., 2015; Blanco-Bercial, 2020) and that, while better populated than the Silva database, representative V9 sequences for many mesopelagic animal species are not available on Genbank.

Rarefaction curves were generated on the eDNA and MOCNESS all-data and metazoan-only datasets to evaluate whether the obtained sequencing depths were sufficient to recover all taxa. Using datasets where sequencing depth in all samples was truncated to the lowest observed sequencing depth, we generated Jaccard (presence-absence) and Bray-Curtis indices in QIIME2, and used these to create non-metric multidimensional scaling (nMDS) plots with vegan 2.3_5 in the statistical package R (Oksanen et al., 2016). Data points were visualized relative to time of sampling (day or night) and sampling depth categories (0–100, 100–200, 200–500, 500–800 m). Functional regressions of the data points against each nMDS dimension were performed in Matlab to assess the significance of observed patterns (Ricker, 1973).

RESULTS

Environmental DNA Sequence Data Summary

A total of 13,561,867 sequences were obtained from 40 samples and 5 negative controls, with an average of 301,375 ± 22,404 (SE) reads per sample. The number of reads in the negative controls ranged from 550 to 6,899. A sixth negative control (from the PCRs) had no reads. In the DADA2 analysis, a total of 12,181,342 sequences (approximately 90%) remained after the filtering, denoising, merging, and chimera removal steps, with an average of 270,696 ± 20,331 (SE) sequences per sample (**Supplementary Table S2**). These sequences were classified into 10,543 unique amplicon sequence variants (ASVs). After removing rare ASVs (total frequency across samples < 10), 8,417 ASVs comprising 12,170,626 sequences remained. Of these, 351 ASVs comprising 1,421,650 sequences were metazoans. There were 65 ASVs that were present in both the negative controls and the samples (**Supplementary Table S3**). For each of these, we took the maximum number of reads found in any of the controls and subtracted that number from the corresponding ASV read count from every sample. Four additional ASVs that were discovered to be either misclassified by Silva as metazoans (based on our Blast results, below) or were unlikely mesopelagic inhabitants were also removed from our analyses.

The percentage of ASVs that were classified as metazoan taxa varied considerably between samples from different depths. The proportion of metazoan sequences in the results generally declined with depth beginning at 30 m (approximately coinciding with the pycnocline; **Supplementary Figure S6**), and in all cases except one, metazoans comprised 21% or less of the total number of obtained reads in samples collected below 200 m (**Figure 4**). Non-metazoan taxa (**Supplementary Table S4**) were not analyzed further.

A broad range of metazoan taxa were recovered in both daytime and nighttime casts (**Figure 5**). In Cast 10, where duplicate samples were taken at each depth, we observed considerable differences in the presence and abundance of taxa in the reads we obtained (**Figure 5C**). Overall, crustaceans (Copepoda, Eumalacostraca, Myodocopa) and medusozoans (Scyphozoa, Hydroidolina, Trachylina) were particularly well-represented. Crustacean ASVs comprised 53.7 ± 28.9% of the reads in mesopelagic samples (200–800 m) and 73.0 ± 24.2% of the reads in epipelagic samples (0–200 m). Conversely, medusozoan ASVs were generally a more prominent component of the reads (35.2 ± 24.8%) in mesopelagic depths than in epipelagic depths (16.6 ± 25.4%). The Silva database Level 6 analysis revealed that the major components of crustacean reads were classified as calanoid and cyclopoid copepods (**Figure 6**), and the major components of the obtained medusozoan reads were siphonopohores (especially at mesopelagic depths) and anthoathecate medusae (**Figure 7**). We also found that 150

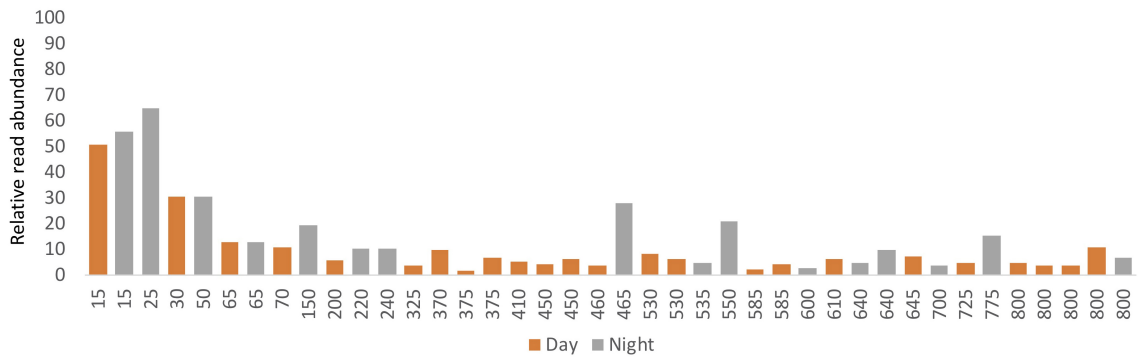


FIGURE 4 | Relative read abundances from ASVs classified as metazoans in the eDNA samples vs. depth.

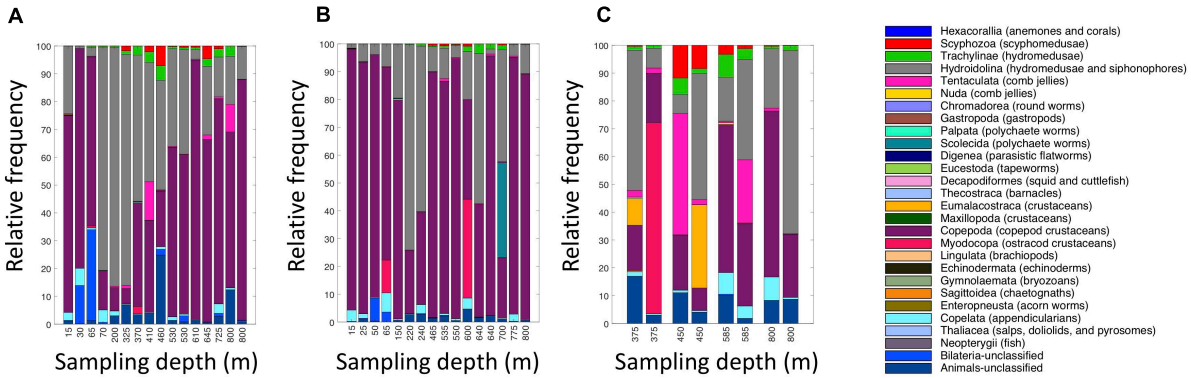


FIGURE 5 | Relative read abundances of metazoan taxa in the eDNA samples from **(A)** daytime casts (casts 5 and 7); **(B)** nighttime casts (casts 6 and 9); and **(C)** daytime cast with duplicates (cast 10). Taxon categories are based on the Silva database, Level 5. Different depths were sampled in each cast.

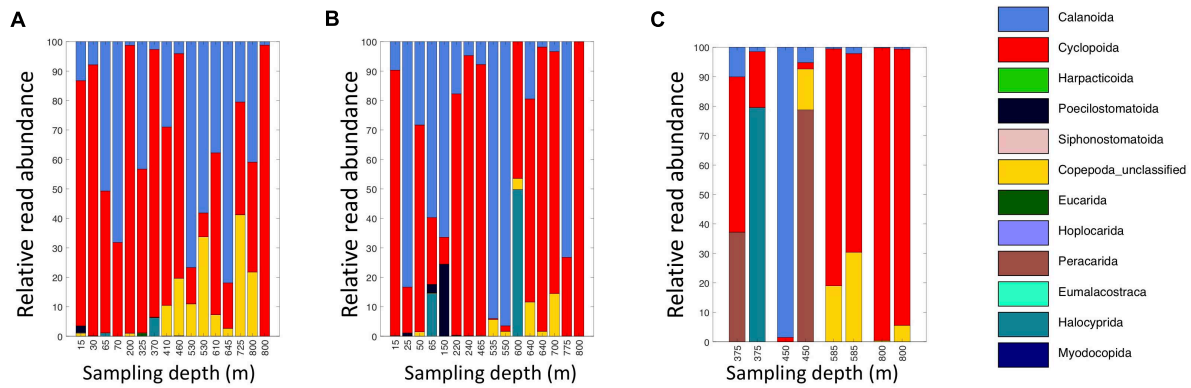


FIGURE 6 | Relative read abundances of classifications of eDNA Level 6 taxa identified as crustaceans (Copepoda, Eumalacostraca, Myodocopa, Thecostraca). **(A)** Daytime casts 5 and 8. **(B)** Nighttime casts 6 and 9. **(C)** Daytime cast 10 with duplicate sampling.

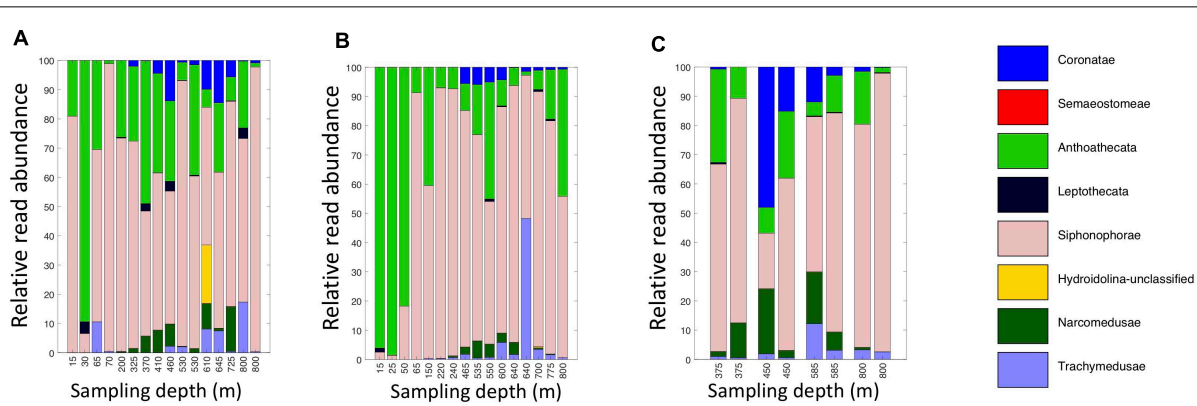


FIGURE 7 | Relative read abundances of eDNA Level 6 taxa identified as Medusozoan (Scyphozoa, Hydroidolina, Anthoatheca, Leptotheca, and Trachylinae). **(A)** Daytime casts 5 and 8. **(B)** Nighttime casts 6 and 9. **(C)** Daytime cast 10 with duplicate sampling.

of the 275 eDNA ASVs matched GenBank entries with 97% identity or greater (**Supplementary Table S5**). Consistent with the Silva classification, taxa with Genbank matches included many calanoid and cyclopoid copepods, siphonophores, and anthoathecate medusae, as well as other taxa such as trachymedusae, scyphomedusae, ctenophores, and other less commonly recovered taxa.

Rarefaction plots of all ASVs and the filtered metazoan-only dataset showed an initial increase in the number of ASVs with sequence subsampling depth and then leveled off,

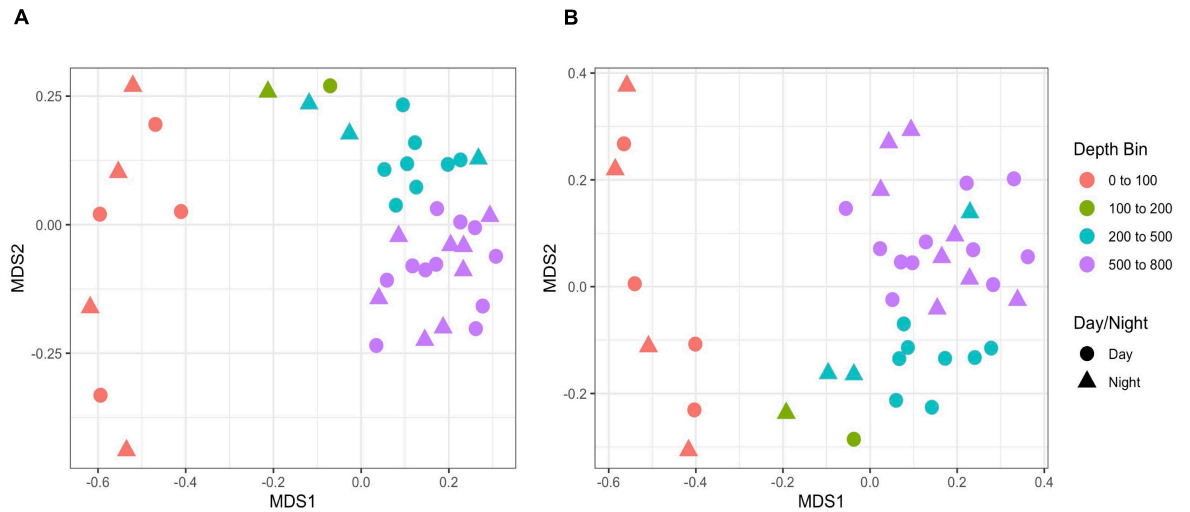


FIGURE 8 | NMDS plots of eDNA data based on **(A)** Jaccard and **(B)** Bray-Curtis indices. Bray-Curtis indices are based on log-transformed data. Shapes indicate day vs. night sampling and color indicates depth bins.

indicating that the sequencing depth was sufficient to capture ASV diversity in our samples (**Supplementary Figure S7**). Using the metazoan datasets truncated at the lowest sequence depth, nMDS plots revealed clustering was related to depth but not time of sampling (**Figure 8**). The structuring was more evident with the Jaccard indices (stress = 0.128) than with the Bray-Curtis indices (stress = 0.221). However when the data were log-transformed, the stress index for the Bray-Curtis analysis improved (stress = 0.154). Note also, unlike Jaccard similarity which is based on presence/absence of taxa, Bray-Curtis indices take into account read abundances, which in metabarcode data, are influenced by PCR biases (Kelly et al., 2018) and eDNA dynamics (Allan et al., 2020). A regression of the Jaccard data points against the nMDS axes revealed a significant relationship with dimension 1 (R^2 = 0.728, p < 0.001) but less so with dimension 2 (R^2 = 0.108, p = 0.0.38). Samples corresponding to the shallowest depth bin (0–100 m) appeared to be driving this relationship (**Figure 8**), so to assess the relationship with depth for the deeper samples, we repeated the regression analysis without the 0–100 m samples. We found significant correlations with both dimension 1 (R^2 = 0.34, p < 0.001) and dimension 2 (R^2 = 0.755, p = 0.001). There was no significant relationship with sampling time (day vs. night) for either dimension (dimension 1: R^2 = 0.025, p = 0.333; dimension 2: R^2 = 0.003, p = 0.744). The results were similar for the Bray-Curtis analysis based on the log-transformed data.

MOCNESS Zooplankton Sequence Data Summary and Comparison With eDNA

Altogether, the 9 MOCNESS nets sampled a combined volume of 12,193.7 m^3 (=12,193,700 l). From this, a total of 4,129,604 sequences were obtained from the large and small size fraction samples combined, with an average of 229,422 ± 8,025 (SE) reads per sample. A total of 3,629,395 sequences (approximately 87%) remained after the filtering, denoising, merging, and chimera removal steps, with an average of 201,633 ± 7,352 (SE) sequences per sample (**Supplementary Table S6**). These sequences were classified into 866 unique amplicon sequence variants (ASVs). After sequences with a frequency less than 10 were removed, 622 ASVs (3,628,313 sequences) were retained, 470 of which were metazoans that, like the eDNA results, comprised a diverse array of animal taxa (**Figure 9** and **Supplementary Table S7**). However, the number of ASV's obtained from the zooplankton dataset (455) was approximately two-thirds greater than that obtained from the eDNA dataset (275). Overall, medusozoan taxa were less abundant in the MOCNESS tow than in the eDNA samples (3.11 ± 3.35% in the large and 7.86 ± 10.34% in the small size fractions), but crustaceans were generally highly abundant (91.10 ± 4.06% in the large and 88.32 ± 11.93% in the small size fractions). Also in contrast to the eDNA results, the major components of the crustacea were calanoid copepods (but few cyclopoid copepods) and eumalacostracans (**Figure 10**). The medusozoan groups were comprised of primarily siphonophores, but other groups including narcomedusae, trachymedusae, and others were also detected (**Figure 11**). The Blast results included many 97% or greater matches with numerous calanoid copepods, and eumalacostracans including euphausiids, amphipods, and decapods. Other matches included a greater variety of non-crustacean taxa as compared to the eDNA results, and these included siphonophores and other medusazoan taxa as well as chaetognaths, polychaetes, fishes, salps, and others (**Supplementary Table S7**). As with the eDNA dataset, rarefaction curves showed that sequencing depth was sufficient to capture the ASV diversity in both the full data set and the metazoan-only dataset (**Supplementary Figure S8**).

Additionally, a total of 10 large individuals and organism fragments were removed from the MOCNESS samples before homogenizing. Six of these produced PCR bands, and 5 were sequenced successfully (**Supplementary Table S8**). Three of

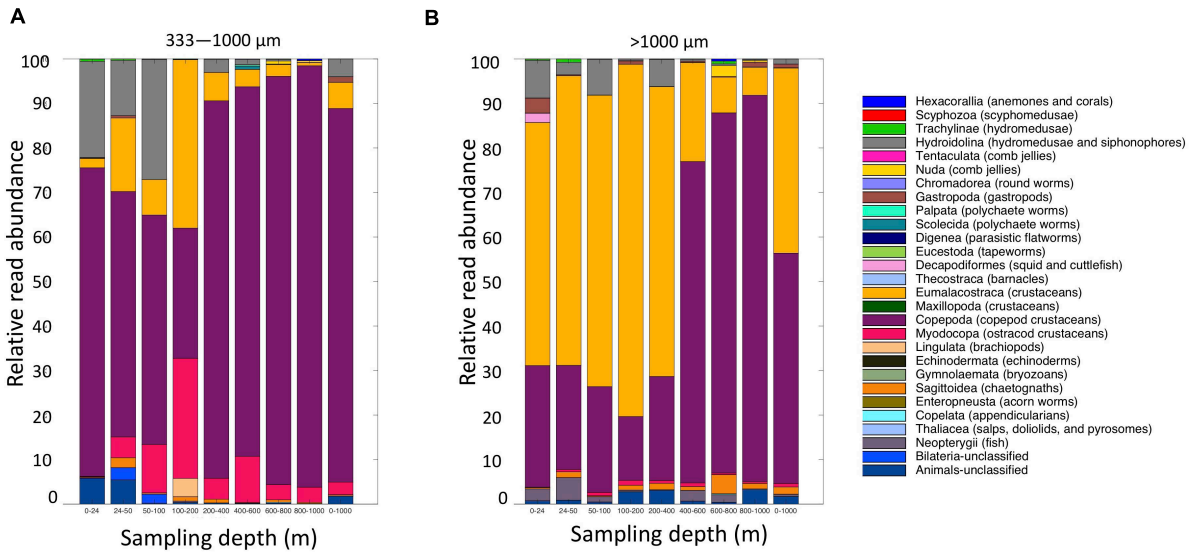


FIGURE 9 | Relative read abundances of metazoan taxa from the **(A)** small (333–1,000 µm); and **(B)** large (>1,000 µm) fractions of the MOCNESS tow. Taxon categories are based on the Silva database Level 5.

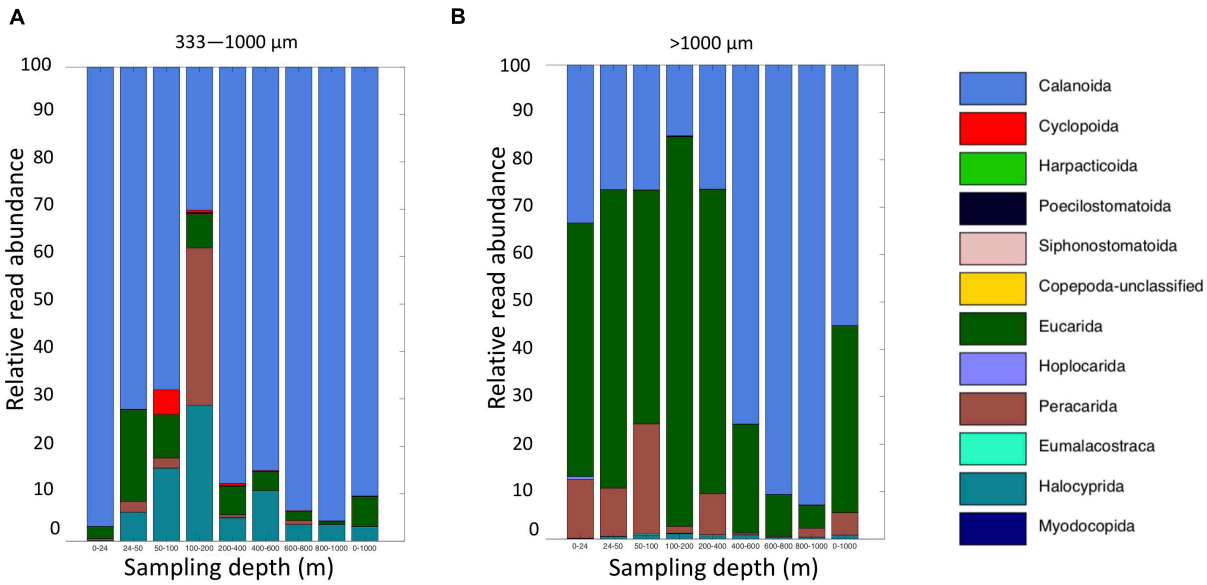


FIGURE 10 | Relative read abundances of classifications of Level 6 taxa identified as crustaceans (Copepoda, Eumalacostraca, Myodocopa, Thecostraca) **(A)** MOCNESS small size fraction and **(B)** MOCNESS large size fraction.

these were fish, one was a scyphozoan, and one was likely a eumalacostracan. All were identical to MOCNESS ASVs. The specimens that failed to be sequenced included 3 fish, one shrimp, and one unidentified gelatinous animal.

The MOCNESS sampling recovered a greater number of both metazoan ASVs and metazoan taxa than did sampling with the Niskin bottles (eDNA samples). Overall, we found 455 metazoan ASVs in the MOCNESS samples (metabarcoding ASVs from small and large fractions combined plus the one unique large-individual sequence), and 275 in the eDNA samples (all casts and samples combined). We compared the number of high-level animal taxa (as classified to Level 7 in the Silva database) and found that the MOCNESS sampling (all nets and fractions combined, including large individuals) detected 23 more taxa than the eDNA sampling. Thirty-nine taxa were detected in both approaches, 17 were found in eDNA only, and 40 were found in the MOCNESS only (**Supplementary Table S9**). Despite the greater number of metazoan taxa sampled by the MOCNESS, sampling efficiency, defined here as number of metazoan ASVs detected per liter of water sampled, was far greater for eDNA (275

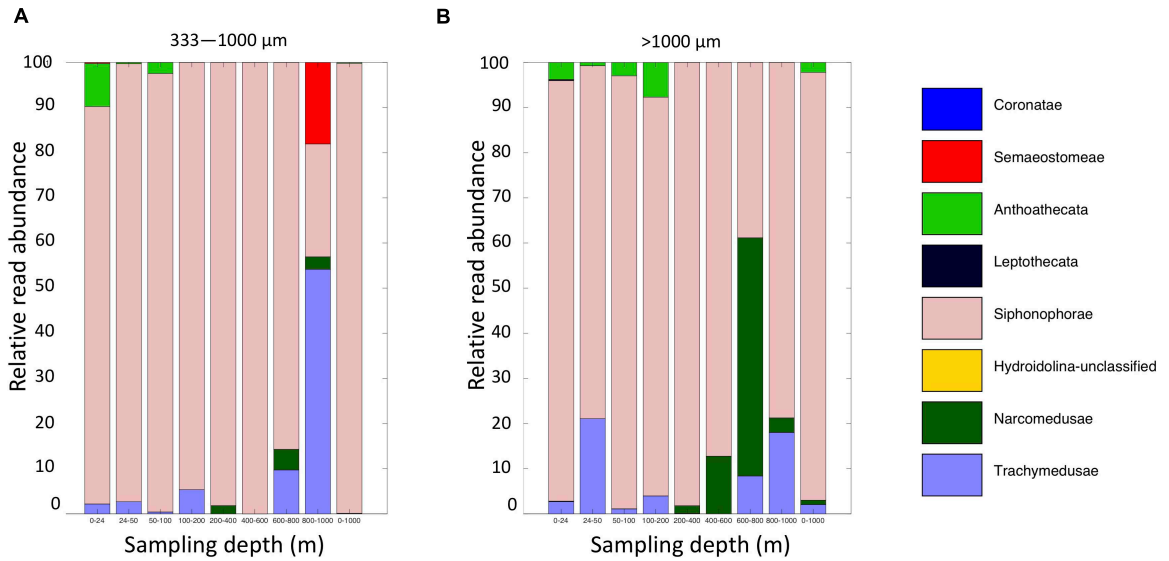


FIGURE 11 | Relative read abundances of Level 6 taxa identified as Medusozoan (Scyphozoa, Hydroidolina, Anthoatheca, Leptotheca, and Trachylinae). **(A)** MOCNESS small size fraction. **(B)** MOCNESS large size fraction.

ASVs/206 l = 1.335 ASVs/l; volume filtered from **Table 1**) than for the MOCNESS (455 ASVs/12,193,700 l = 0.00003731 ASVs/l).

DISCUSSION

Biodiversity Detection

Our results suggest that eDNA analysis is an informative and highly efficient approach for invertebrate biodiversity detection in the mesopelagic that complements net sampling. Our sequence results were based on the 18S V9 marker, which detects a comprehensive and phylogenetically diverse range of animals (Wu et al., 2015; Bucklin et al., 2019; Blanco-Bercial, 2020). The 18S V9 marker is well-suited for broadly focused eDNA surveys due to its relatively short length, as eDNA is likely degraded and thus comprised of relatively small fragments (Jo et al., 2017).

The animal component of our eDNA results declined with depth, and this result is consistent with Stefanoudis et al. (2019), who, in a survey extending to 800 m depth based on specimen identifications from net tows, showed that zooplankton abundance decreased substantially below 200 m. In our study, the observed decline was approximately coincident with the pycnocline (which was located in the 0–200 m depth interval in Stefanoudis et al. (2019)). While the pycnocline can be a barrier for animal dispersal (Suzuki et al., 2018), Closek et al. (2019) found no significant difference in fish taxa identified from eDNA above and below the pycnocline off of the central California coast. It will be interesting for future studies to examine whether or not the pycnocline presents a barrier to eDNA dispersal and how that impacts interpretation of eDNA signals.

The eDNA sequences originated primarily from calanoid and cyclopoid copepods, siphonophores, and other medusozoans. In contrast, the crustacean component of the MOCNESS sequences comprised primarily calanoid copepods (and relatively few cyclopoids) and eumalacostracans (e.g., euphausiids and amphipods). The medusozoan component of the MOCNESS sequences also included siphonophores (but with fewer ASVs), and the relative proportions of other medusuzoans were different than in the eDNA dataset.

Can Environmental DNA Can Add Substantially to Plankton Metabarcoding Analyses?

While the number and types of metazoan taxa detected in the genetic material collected by MOCNESS sampling were greater than those detected by the eDNA sampling, we detected 17 taxa via eDNA that were not detected by the MOCNESS. The volume of water sampled by Niskin bottles for eDNA analysis, however, was a tiny fraction of that sampled by the MOCNESS. Thus, in terms of the number of metazoan ASVs and taxa detected per volume of water sampled, eDNA from Niskin bottles outperformed MOCNESS sampling, although each approach detected taxa that the other did not. Given the close proximity of the CTD and MOCNESS sampling locations and times, and the temperature and salinity data that showed that the sampling locations belonged to the same water mass, we assume that the pool of biodiversity was similar for each sampling event and that the observed differences are due to sampling biases associated with each method. Our findings that many of the same taxa were recovered in both sampling approaches and that other taxa that were recovered in only one or the other approach, are also consistent with other studies that compare eDNA with other survey methods (Thomsen et al., 2016; Kelly et al., 2017; Sigsgaard et al., 2017; Closek et al., 2019; Liu et al., 2019).

The relative frequencies of medusozoan ASVs in our eDNA reads compared to our MOCNESS reads could indicate a

relatively high eDNA shedding rate or a relatively high biomass—or both. While there are few studies on eDNA shedding rates, the available data show that they are variable between taxa, between individuals within a species, and between different stages of an organism's life cycle (Sansom and Sassoubre, 2017). Most studies on eDNA shedding focus on coastal and freshwater fish (Takahara et al., 2012; Klymus et al., 2015; Sassoubre et al., 2016; Nevers et al., 2018). However, Minamoto et al. (2017) found that medusae shed eDNA at a significantly greater rate than fish (by 1–3 orders of magnitude). Allan et al. (2020) also found that medusae may shed eDNA at higher rates than other animal types. Animals with hard external surfaces (e.g., bivalves and crustaceans) may have lower shedding rates (Sansom and Sassoubre, 2017; Allan et al., 2020). Interestingly our eDNA data lacked eumalacostracan ASVs, which were common in the MOCNESS data. Eumalacostracans include euphausiids, which are a well-known component of mesopelagic biomass and so their paucity in the eDNA is notable. Additional research on shedding rates in mesopelagic organisms is critical for interpreting eDNA signals.

The biomass of medusozoans and other gelatinous animals in the mesopelagic is likely substantially underestimated due to difficulties in sampling them (Madin and Harbison, 1978; Larson et al., 1991; Robison, 2004, 2009). These animals are destroyed by traditional plankton net sampling, and many new forms were discovered only when human-occupied underwater vehicles (HOVs) and remotely operated vehicles (ROVs) came into use (Madin and Harbison, 1978; Larson et al., 1991; Robison, 2004, 2009). Our eDNA and MOCNESS results showing a greater presence of gelatinous animals in eDNA are consistent with the hypothesis that gelatinous animal biomass is under-sampled using net sampling and also demonstrate that analysis of eDNA provides another, complementary approach to detecting these fragile animals.

Biomass and eDNA Sampling Strategy

Our sampling strategy focused on acoustic scattering layers identified by the 18 and 38 kHz (deep scattering layers) and the 120 and 200 kHz (near-surface scattering layers) Simrad EK60 echosounders. These frequencies are typically used to detect biomass in fisheries surveys (Jech and Sullivan, 2014; Proud et al., 2019). We chose this approach as these layers likely had relatively high concentrations of organisms, although the types of taxa detected by acoustic backscatter is frequency-dependent (Lavery et al., 2007). The lower frequencies that we used (18 and 38 kHz) were more informative for our sampling as they penetrate to deeper depths. These frequencies also tend to detect gas-bearing organisms such as siphonophores and fish with swim bladders (Lavery et al., 2007) although deep scattering layers may contain a wide range of animal taxa (Stefanoudis et al., 2019). While the relationship between eDNA concentration and abundance or biomass is often not direct (Andruszkiewicz et al., 2017b), biomass indicated by acoustic backscatter has been correlated with eDNA signals. Yamamoto et al. (2016) found that in general, fish eDNA signals were correlated with biomass as determined from a 120 kHz echosounder in Maizuru Bay, a shallow semi-enclosed water body in the Sea of Japan. They noted, though, that confounding factors, such as the presence of exogenous eDNA sources, could disrupt this correlation.

The fact that we sampled in the deep-scattering layers where acoustic backscatter from ship-borne echosounders presumably indicates the presence of fish, yet did not detect fish eDNA, suggests more effort is needed for understanding how eDNA relates to biomass in the mesopelagic region. While 18S V9 is known to pick up a broad range of animals including fish (Blanco-Bercial, 2020), other taxa are potentially preferentially amplified (Sawaya et al., 2019). Fish were also only a very small component of the MOCNESS DNA results; however the failure to sequence three of our fish specimens from the MOCNESS sampling also suggests that some species may be missed with this marker. Analysis of eDNA samples with a marker such as 12S that targets vertebrates may be needed to uncover fish biodiversity (Andruszkiewicz et al., 2017b; Kelly et al., 2017). Additionally, the frequency-dependent scattering by organisms can bias our interpretation of the types of organisms present using acoustic echograms. This is especially true using ship-borne systems where the frequency content of the acoustic data decreases with range from the transducer—i.e., "higher" frequencies attenuate with depth so only "lower" frequency data are available for interpretation. Notably, our mesopelagic sampling was based on the two lower frequencies (18 and 38 kHz) and while we did not detect fish, we did observe a significant eDNA signal from siphonophores. Furthermore, our Blast search found that several siphonophore ASVs matched families in the suborder Physonectae, which contain gas-filled pneumatophores (Dunn et al., 2005), suggesting that they could be the source of those acoustic signals. eDNA studies that focus solely on 12S may miss these potentially important taxa.

In the near-surface scattering layers, the combination of all frequencies (18, 38, 120, and 200 kHz) provided additional information on a greater variety of animals and presumably a less biased interpretation (Davison et al., 2015; Bassett et al., 2020). Vertically migrating mesopelagic animals, including fish, crustacea, and gelatinous animals, comprise these near-surface scattering layers at night, and surface data collected at night may provide useful information on these mesopelagic species. Focusing on surface sampling during the night could thus be an efficient strategy; however, deep water sampling is necessary to identify non-migrating mesopelagic species.

Because our sampling depths were chosen based on real time echosounder data, sampling depths differed between casts and were not consistent with the MOCNESS intervals. To look for patterns associated with depth, we therefore combined our data from all casts into four 100–300 m depth bins and found that composition of our eDNA reads differed significantly among these depth bins. Thus, eDNA appears to be a suitable tool for examining biodiversity patterns at that vertical scale. Finer scale resolution of potential vertical patterns, such as those associated with acoustically detected layers, was limited by the number of Niskin bottles (and thus samples) available to us (8) given our goal to explore variation in the water column, but could be a direction for future research. Other studies have found that eDNA can resolve community differences between marine habitats, despite the biological and physical interconnectedness

of those habitats (Jeunen et al., 2019; Laroche et al., 2020; West et al., 2020).

Interpreting Mesopelagic eDNA Signals

Many biological and physical factors could have influenced our observed eDNA signals, including eDNA shedding rates, decay rates, and transport (Sansom and Sassoubre, 2017). As described above, eDNA shedding estimates are derived primarily from fish, but there is some evidence that medusae may shed eDNA at a substantially higher rate (Minamoto et al., 2017; Allan et al., 2020). This could potentially result in a disproportionately greater observed read abundance for medusozoan taxa. Shedding rates are also influenced by temperature (Lacoursière-Roussel et al., 2016), and constraining shedding rates of vertically migrating mesopelagic animals that experience a range of temperatures on a daily basis may be especially complex. eDNA decay estimates, which are also based primarily on fish studies, suggest that eDNA can be detected about 1–7 days after shedding (Thomsen et al., 2012; Sansom and Sassoubre, 2017). However, like shedding, the decay rate depends on a variety of environmental factors such as temperature. Cowart et al. (2018) reported a slower decay rate (with a half-life of 37.2 h) potentially allowing detection after 25 days from fish in subzero Antarctic waters. As physical and biogeochemical conditions in the mesopelagic zone are dramatically different from the surface layer of the ocean, eDNA decay rate derived for surface conditions might not be applicable to the mesopelagic zone. Interestingly, one seemingly relevant factor that varies with depth is sunlight, which does not appear to impact eDNA decay (Andruszkiewicz et al., 2017a).

Our eDNA samples may include signals from animals both in our sampling location as well as eDNA transported to that location from elsewhere. Water movement can transport eDNA downstream of its source. Thus, eDNA signals from flowing environments may actually indicate the presence of species in an upstream location (Wacker et al., 2019). Additionally, water mixing can cause the signal of eDNA released over a small spatial scale to quickly become diluted and undetectable (Pilliod et al., 2014; Stoeckle et al., 2017). Most studies on eDNA dispersal to date are from river and stream environments, where the flow is primarily unidirectional. Based on modeling in a coastal system (Monterey Bay, California), Andruszkiewicz et al. (2019) found that eDNA can be transported several 10 s of km horizontally by regional currents over the course of a few days. Open ocean environments are much more complex and subject to the influences of different flows, including mesoscale, sub-mesoscale, tidal, upwelling, and downwelling currents. In the Slope Water off the Northeast U.S. coast where this study occurred, flow can be impacted by the Gulf Stream, warm-core rings, subduction and upwelling along the ring periphery, slope current, tides, and internal tides (Joyce, 1984; Flagg et al., 2006; Kelly et al., 2016; Zhang and Partida, 2018). These flows could redistribute eDNA in the three-dimensional space.

Despite the potential for eDNA dispersal by ocean currents, some studies in coastal systems suggest that the eDNA signal can remain localized. In their survey of Maizuru Bay, Japan, Minamoto et al. (2017) found spatial and temporal correlation between eDNA concentrations of jellyfish from surface and subsurface samples and simultaneously performed visual surface-water observations. Kelly et al. (2018) found that tidal flow had minimal effect on characterizing local communities in tidally dynamic nearshore habitats. Similarly, Jeunen et al. (2019) found unique localized eDNA signatures from four coastal distinct, yet interconnected habitats that were within a few km of each other. Weak spatial dispersal of eDNA signal could result from either long residence time of the water in the eDNA source region or rapid decay of the eDNA materials. The relative importance of physical transport vs. biological degradation in eDNA signal dispersal is a question that calls for further study.

In the mesopelagic ecosystem, there is an additional complicating process not found in other aquatic habitats that may influence our interpretation of our eDNA signals. Diel vertical migration (DVM), or the movement of many mesopelagic animals to surface waters at night to feed and back down to deeper depths during the day, has been referred to as the largest migration on the planet (Hays, 2003; Sutton, 2013). Animals can travel vertical distances of hundreds of meters in each direction on a daily basis. However, we did not find a significant pattern associated with the time of sampling (day vs. night) in any depth bin as might be expected from vertically migrating animals. The lack of day/night signal suggests that the eDNA shedding rate might be very low or the time scale of diel vertical migration might be too short relative to eDNA transport and decay processes, both of which may limit our ability to detect vertical migration (Allan et al., 2020). Furthermore, the animals themselves may expedite the downward transport of eDNA originating from surface organisms through their feeding activities. For example, a migrating animal may feed at the surface at night, and release traces of eDNA from its surface prey through defecation at depth. Thus, while it is tempting to try to infer DVM-related (and other) changes in taxa with depth using eDNA, it is first necessary to understand the spatiotemporal scales of the relevant biological and physical processes in the surface through mesopelagic depths.

Lastly, we note that on our final cast where duplicate samples were taken at four depths, we observed considerable variation in read abundance between those duplicates. Other studies also show variation between replicates (Andruszkiewicz et al., 2017b). We do not currently understand the scale of eDNA distributions, which may be different for different species, and emphasize that the volume of our eDNA samples (5 l) is minimal relative to the scale of the habitat. More work should be done to optimize the scope and scale of eDNA sampling and replication to improve the accuracy of eDNA biodiversity estimates. Additionally, technological advances that allow for multiple, large-volume sample acquisition (Govindarajan et al., 2015; McQuillan and Robidart, 2017) may be especially useful for mesopelagic eDNA studies.

Taxonomy, Reference Libraries, and Marker Resolution

In our results, most of our ASVs did not have exact matches on GenBank. Possible explanations for lack of exact matches are

bioinformatics error and lack of available reference sequences. Metabarcoding studies typically uncover vastly more putative taxa than traditional morphological studies (Lindeque et al., 2013; Kelly et al., 2017; Sommer et al., 2017). While to some extent, this greater diversity may be real, it is also possible that the observed diversity is artifactual (Caron and Hu, 2019; Santoferrara, 2019). We feel this is unlikely in our case, as we employed an approach that assesses and corrects likely sequencing errors to produce "amplicon sequencing variants" (ASVs), which is thought to be more accurate than traditional clustering or "operational taxonomic unit" (OTU) approaches (Callahan et al., 2017; Macheriotou et al., 2019), especially for characterizing community biodiversity (Pauvert et al., 2019). We furthermore conservatively removed rare ASVs from our data analysis.

Identifying taxa in metabarcoding studies relies on accurate, well-populated reference libraries (Cristescu, 2014; Bucklin et al., 2016; Elbrecht et al., 2016; Porter and Hajibabaei, 2018). DNA sequences are matched against a reference database for taxon assignment; however, existing databases, including GenBank, are notoriously incomplete or contain errors (Leray and Knowlton, 2016; Lindsay et al., 2017; Santoferrara, 2019). Given the relative inaccessibility of deep oceanic waters, reference libraries may be particularly depauperate for animals in the mesopelagic. Our Genbank results showed matches for a variety of crustaceans and cnidarians; however we also noted an absence of V9 reference sequences on Genbank for many species of mesopelagic fishes (Govindarajan, pers. obs.). For some taxa, there is a mismatch between traditional barcode efforts that focus on the COI and 16S genes (Cristescu, 2014; Lindsay et al., 2015; Elbrecht et al., 2016; Leray and Knowlton, 2016), and metabarcoding analyses which often instead utilize 18S markers to capture a broad range of higher level taxa (de Vargas et al., 2015; Kelly et al., 2017; Djurhuus et al., 2018; Bucklin et al., 2019; Sawaya et al., 2019; Blanco-Bercial, 2020). Additionally, nominal species on Genbank may be misidentified, either due to mistakes in identification due to lack of taxonomic expertise or to the existence of cryptic species (Lindsay et al., 2015, 2017; Abad et al., 2017). Development of reference barcode libraries, grounded in taxonomic expertise (Wheeler, 2018; Pinheiro et al., 2019), for mesopelagic animals are essential for future applications of eDNA analyses in the mesopelagic.

Our goal, however, was primarily to identify higher level taxa rather than species, and V9 is better suited toward this task due to its relatively conserved nature compared to markers such as COI and 16S. Thus, an exact match to a reference V9 sequence might not necessarily indicate a species identification, because multiple closely related species could potentially share that sequence (Abad et al., 2017; Blanco-Bercial, 2020). On the other hand, because evolutionary rates are variable between lineages, a given percent similarity (e.g., 97%) might indicate different degrees of taxonomic classification depending on the lineage. In our Blast searches, we selected a 97% identity as our threshold for reporting Genbank results based on calibration of this marker in copepods (Wu et al., 2015). Similar calibration analyses should be conducted for other animal groups and markers as reference barcode libraries are expanded in order to better understand the taxonomic makeup of the sampled biodiversity.

CONCLUSION

More information on mesopelagic biodiversity is urgently needed as interest in harvesting these resources increases. eDNA analysis has tremendous potential to contribute to our understanding of this environment and fill important knowledge gaps, such as those related to the detection of animals that are poorly sampled by other methods. eDNA analysis is also highly efficient in terms of the number of taxa recovered per unit effort and can complement other traditional sampling methods. However, eDNA analysis is a new approach that to date has been primarily used in freshwater and coastal marine environments. Based on our results, we suggest the following future research directions for facilitating the application of eDNA analyses to the mesopelagic environment: (1) Experiments to determine eDNA shedding, decay, transport, and dispersal rates under mesopelagic physical and biological conditions to improve our interpretation of eDNA signals; (2) Assess relationships between eDNA and biomass; and (3) Populate genetic reference databases with sequences from mesopelagic species that have been morphologically identified by taxonomic experts for improved taxonomic assignments.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: Dryad repository (doi: 10.5061/dryad.4xgxd255j).

AUTHOR CONTRIBUTIONS

AG conceived the study and conducted the DNA sequence analyses and drafted the manuscript. AG, AL, JMJ, WZ, JL, and PW planned, prepared, and conducted the shipboard sampling. RF conducted the lab work and assisted with the acoustics plots. JMJ conducted the acoustics analyses. WZ and PW conducted the CTD analyses. All authors contributed to data interpretation, structuring of the manuscript, reviewed and edited the manuscript, and approved the final version for submission.

FUNDING

This research was part of the Woods Hole Oceanographic Institution's Ocean Twilight Zone Project, funded as part of The Audacious Project housed at TED. Funding for the NOAA Ship Henry B Bigelow was provided by NOAA's Office of Marine and Aviation Operations (OMAO).

ACKNOWLEDGMENTS

We thank the crew and scientists of the NOAA Ship Henry B Bigelow, Glenn McDonald (WHOI) and Helena McMonagle

(WHOI) for assistance with eDNA sampling, Bo Reese (UConn Center for Genomic Innovation) for assistance with library preparation and DNA sequencing, Dana Yoerger (WHOI) for assistance with plots, and Elizabeth Allan (WHOI) for helpful comments on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.574877/full#supplementary-material

REFERENCES

Abad, D., Albaina, A., Aguirre, M., and Estonba, A. (2017). 18S V9 metabarcoding correctly depicts plankton estuarine community drivers. *Mar. Ecol. Prog. Ser.* 584, 31–43. doi: 10.3354/meps12373

Allan, E. A., Zhang, W. G., Lavery, A. C., and Govindarajan, A. F. (2020). Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. *Environ. DNA* 2020:141. doi: 10.1002/edn3.141

Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W., and Huse, S. M. (2009). A Method for Studying Protistan Diversity Using Massively Parallel Sequencing of V9 Hypervariable Regions of Small-Subunit Ribosomal RNA Genes. *PLoS One* 4:e6372. doi: 10.1371/journal.pone.0006372

Andruszkiewicz, E. A., Koseff, J. R., Fringer, O. B., Ouellette, N. T., Lowe, A. B., Edwards, C. A., et al. (2019). Modeling Environmental DNA Transport in the Coastal Ocean Using Lagrangian Particle Tracking. *Front. Mar. Sci.* 6:.477. doi: 10.3389/fmars.2019.00477

Andruszkiewicz, E. A., Sassoubre, L. M., and Boehm, A. B. (2017a). Persistence of marine fish environmental DNA and the influence of sunlight. *PLoS One* 12:e0185043. doi: 10.1371/journal.pone.0185043

Andruszkiewicz, E. A., Starks, H. A., Chavez, F. P., Sassoubre, L. M., Block, B. A., and Boehm, A. B. (2017b). Biomonitoring of marine vertebrates in Monterey Bay using eDNA metabarcoding. *PLoS One* 12:e0176343. doi: 10.1371/journal.pone.0176343

Bassett, C., Lavery, A. C., Stanton, T. K., and Cotter, E. D. (2020). Frequency- and depth-dependent target strength measurements of individual mesopelagic scatterers. *J. Acoust. Soc. Am.* 148, EL153–EL158. doi: 10.1121/10.0001745

Blanco-Bercial, L. (2020). Metabarcoding Analyses and Seasonality of the Zooplankton Community at BATS. *Front. Mar. Sci.* 7:173. doi: 10.3389/fmars.2020.00173

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., et al. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6:90. doi: 10.1186/s40168-018-0470-z

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C., Al-Ghalith, G. A., et al. (2018). QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ* 2018:27295v2. doi: 10.7287/peerj.preprints.27295v2

Bucklin, A., Lindeque, P. K., Rodriguez-Ezpeleta, N., Albaina, A., and Lehtiniemi, M. (2016). Metabarcoding of marine zooplankton: prospects, progress and pitfalls. *J. Plankton Res.* 38, 393–400. doi: 10.1093/plankt/fbw023

Bucklin, A., Yeh, H. D., Questel, J. M., Richardson, D. E., Reese, B., Copley, N. J., et al. (2019). Time-series metabarcoding analysis of zooplankton diversity of the NW Atlantic continental shelf. *ICES J. Mar. Sci.* 76, 1162–1176. doi: 10.1093/icesjms/fsz021

Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* 11, 2639–2643. doi: 10.1038/ismej.2017.119

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869

Caron, D. A., and Hu, S. K. (2019). Are We Overestimating Protistan Diversity in Nature? *Trends Microbiol.* 27, 197–205. doi: 10.1016/j.tim.2018.10.009

Casas, L., Pearman, J. K., and Irigoien, X. (2017). Metabarcoding Reveals Seasonal and Temperature-Dependent Succession of Zooplankton Communities in the Red Sea. *Front. Mar. Sci.* 4:241. doi: 10.3389/fmars.2017.00241

Closek, C. J., Santora, J. A., Starks, H. A., Schroeder, I. D., Andruszkiewicz, E. A., Sakuma, K. M., et al. (2019). Marine Vertebrate Biodiversity and Distribution Within the Central California Current Using Environmental DNA (eDNA) Metabarcoding and Ecosystem Surveys. *Front. Mar. Sci.* 6:732. doi: 10.3389/fmars.2019.00732

Cowart, D. A., Murphy, K. R., and Cheng, C.-H. C. (2018). Metagenomic sequencing of environmental DNA reveals marine faunal assemblages from the West Antarctic Peninsula. *Mar. Genomics* 37, 148–160. doi: 10.1016/j.margen.2017.11.003

Cristescu, M. E. (2014). From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. *Trends Ecol. Evol.* 29, 566–571. doi: 10.1016/j.tree.2014.08.001

D'Elia, M., Warren, J. D., Rodriguez-Pinto, I., Sutton, T. T., Cook, A., and Boswell, K. M. (2016). Diel variation in the vertical distribution of deep-water scattering layers in the Gulf of Mexico. *Deep Sea Res. Part I Oceanogr. Res. Papers* 115, 91–102. doi: 10.1016/j.dsr.2016.05.014

Davison, P. C., Koslow, J. A., and Kloser, R. J. (2015). Acoustic biomass estimation of mesopelagic fish: backscattering from individuals, populations, and communities. *ICES J. Mar. Sci.* 72, 1413–1424. doi: 10.1093/icesjms/fsv023

de Vargas, C., de, Audic, S., Henry, N., Decelle, J., Mahé, F., et al. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science* 348:1261605. doi: 10.1126/science.1261605

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière—Roussel, A., Altermatt, F., et al. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Mol. Ecol.* 26, 5872–5895. doi: 10.1111/mec.14350

Dejean, T., Valentini, A., Miquel, C., Taberlet, P., Bellemain, E., and Miaud, C. (2012). Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog Lithobates catesbeianus. *J. Appl. Ecol.* 49, 953–959. doi: 10.1111/j.1365-2664.2012.02171.x

Djurhuus, A., Pitz, K., Sawaya, N. A., Rojas—Márquez, J., Michaud, B., Montes, E., et al. (2018). Evaluation of marine zooplankton community structure through environmental DNA metabarcoding. *Limnol. Oceanogr. Methods* 16, 209–221. doi: 10.1002/lom3.10237

Djurhuus, A., Port, J., Closek, C. J., Yamahara, K. M., Romero-Maraccini, O., Walz, K. R., et al. (2017). Evaluation of Filtration and DNA Extraction Methods for Environmental DNA Biodiversity Assessments across Multiple Trophic Levels. *Front. Mar. Sci.* 4:314. doi: 10.3389/fmars.2017.00314

Dunn, C. W., Pugh, P. R., and Haddock, S. H. D. (2005). Molecular Phylogenetics of the Siphonophora (Cnidaria), with Implications for the Evolution of Functional Specialization. *Syst. Biol.* 54, 916–935. doi: 10.1080/10635150500354837

Elbrecht, V., Taberlet, P., Dejean, T., Valentini, A., Usseglio-Polatera, P., Beisel, J.-N., et al. (2016). Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects. *PeerJ* 4:e1966. doi: 10.7717/peerj.1966

Evans, N. T., Li, Y., Renshaw, M. A., Olds, B. P., Deiner, K., Turner, C. R., et al. (2017). Fish community assessment with eDNA metabarcoding: effects of sampling design and bioinformatic filtering. *Can. J. Fish. Aquat. Sci.* 74, 1362–1374. doi: 10.1139/cjfas-2016-0306

Flagg, C. N., Dunn, M., Wang, D.-P., Rossby, H. T., and Benway, R. L. (2006). A study of the currents of the outer shelf and upper slope from a decade of shipboard ADCP observations in the Middle Atlantic Bight. *J. Geophys. Res. Oceans* 111:2005JC003116. doi: 10.1029/2005JC003116

Giering, S. L. C., Sanders, R., Lampitt, R. S., Anderson, T. R., Tamburini, C., Boutrif, M., et al. (2014). Reconciliation of the carbon budget in the ocean's twilight zone. *Nature* 507, 480–483. doi: 10.1038/nature13123

Glover, A. G., Wiklund, H., Chen, C., and Dahlgren, T. G. (2018). Managing a sustainable deep-sea 'blue economy' requires knowledge of what actually lives there. *eLife* 7:e41319. doi: 10.7554/eLife.41319

Govindarajan, A. F., Pineda, J., Purcell, M., and Breier, J. A. (2015). Species- and stage-specific barnacle larval distributions obtained from AUV sampling and genetic analysis in Buzzards Bay, Massachusetts, USA. *J. Exp. Mar. Biol. Ecol.* 472, 158–165. doi: 10.1016/j.jembe.2015.07.012

Hays, G. C. (2003). A review of the adaptive significance and ecosystem consequences of zooplankton diel vertical migrations. *Hydrobiologia* 503, 163–170. doi: 10.1023/B:HYDR.0000008476.23617.b0

Hazen, E. L., and Johnston, D. W. (2010). Meridional patterns in the deep scattering layers and top predator distribution in the central equatorial Pacific. *Fisher. Oceanogr.* 19, 427–433. doi: 10.1111/j.1365-2419.2010.00561.x

Hidalgo, M., and Browman, H. I. (2019). Developing the knowledge base needed to sustainably manage mesopelagic resources. *ICES J. Mar. Sci.* 76, 609–615. doi: 10.1093/icesjms/fsz067

Irigoien, X., Klevjer, T. A., Røstad, A., Martinez, U., Boyra, G., Acuña, J. L., et al. (2014). Large mesopelagic fishes biomass and trophic efficiency in the open ocean. *Nat. Commun.* 5:3271. doi: 10.1038/ncomms4271

Jech, J. M., and Sullivan, P. J. (2014). Distribution of Atlantic herring (Clupea harengus) in the Gulf of Maine from 1998 to 2012. *Fisher. Res.* 156, 26–33. doi: 10.1016/j.fishres.2014.04.016

Jeunen, G.-J., Knapp, M., Spencer, H. G., Lamare, M. D., Taylor, H. R., Stat, M., et al. (2019). Environmental DNA (eDNA) metabarcoding reveals strong discrimination among diverse marine habitats connected by water movement. *Mol. Ecol. Resour.* 19, 426–438. doi: 10.1111/1755-0998.12982

Jo, T., Murakami, H., Masuda, R., Sakata, M. K., Yamamoto, S., and Minamoto, T. (2017). Rapid degradation of longer DNA fragments enables the improved estimation of distribution and biomass using environmental DNA. *Mol. Ecol. Resour.* 17, 25–33e. doi: 10.1111/1755-0998.12685

Joyce, T. M. (1984). Velocity and Hydrographic Structure of a Gulf Stream Warm-Core Ring. *J. Phys. Oceanogr.* 14, 936–947. doi: 10.1175/1520-04851984014<0936:VAHSOA<2.0.CO;2

Kelly, R. P., Closek, C. J., O'Donnell, J. L., Kralj, J. E., Shelton, A. O., and Samhouri, J. F. (2017). Genetic and Manual Survey Methods Yield Different and Complementary Views of an Ecosystem. *Front. Mar. Sci.* 3:283. doi: 10.3389/fmars.2016.00283

Kelly, R. P., Gallego, R., and Jacobs-Palmer, E. (2018). The effect of tides on nearshore environmental DNA. *PeerJ* 6:e4521. doi: 10.7717/peerj.4521

Kelly, S. M., Lermusiaux, P. F. J., Duda, T. F., and Haley, P. J. (2016). A Coupled-Mode Shallow-Water Model for Tidal Analysis: Internal Tide Reflection and Refraction by the Gulf Stream. *J. Phys. Oceanogr.* 46, 3661–3679. doi: 10.1175/JPO-D-16-0018.1

Kinzer, J. (1969). On the quantitative distribution of zooplankton in deep scattering layers. *Deep Sea Res. Oceanogr. Abstracts* 16, 117–125. doi: 10.1016/0011-7471(69)90068-0

Klymus, K. E., Richter, C. A., Chapman, D. C., and Paukert, C. (2015). Quantification of eDNA shedding rates from invasive bighead carp Hypophthalmichthys nobilis and silver carp Hypophthalmichthys molitrix. *Biol. Conserv.* 183, 77–84. doi: 10.1016/j.biocon.2014.11.020

Lacoursière-Roussel, A., Rosabal, M., and Bernatchez, L. (2016). Estimating fish abundance and biomass from eDNA concentrations: variability among capture methods and environmental conditions. *Mol. Ecol. Resour.* 16, 1401–1414. doi: 10.1111/1755-0998.12522

Laroche, O., Kersten, O., Smith, C. R., and Goetze, E. (2020). Environmental DNA surveys detect distinct metazoan communities across abyssal plains and seamounts in the western Clarion Clipperton Zone. *Mol. Ecol.* 2020:15484. doi: 10.1111/mec.15484

Larson, R. J., Mills, C. E., and Harbison, G. R. (1991). Western Atlantic midwater hydrozoan and scyphozoan medusae: in situ studies using manned submersibles. *Hydrobiologia* 216, 311–317. doi: 10.1007/BF00026480

Lavery, A. C., Wiebe, P. H., Stanton, T. K., Lawson, G. L., Benfield, M. C., and Copley, N. (2007). Determining dominant scatterers of sound in mixed zooplankton populations. *J. Acoust. Soc. Am.* 122, 3304–3326. doi: 10.1121/1.2793613

Leray, M., and Knowlton, N. (2016). Censusing marine eukaryotic diversity in the twenty-first century. *Philosop. Trans. R. Soc. B Biol. Sci.* 371:20150331. doi: 10.1098/rstb.2015.0331

Lindeque, P. K., Parry, H. E., Harmer, R. A., Somerfield, P. J., and Atkinson, A. (2013). Next Generation Sequencing Reveals the Hidden Diversity of Zooplankton Assemblages. *PLoS One* 8:e81327. doi: 10.1371/journal.pone.0081327

Lindsay, D. J., Grossmann, M. M., Bentlage, B., Collins, A. G., Minemizu, R., Hopcroft, R. R., et al. (2017). The perils of online biogeographic databases: a case study with the 'monospecific' genus Aegina (Cnidaria, Hydrozoa, Narcomedusae). *Mar. Biol. Res.* 13, 494–512. doi: 10.1080/17451000.2016.1268261

Lindsay, D. J., Grossmann, M. M., Nishikawa, J., Bentlage, B., and Collins, A. G. (2015). DNA barcoding of pelagic cnidarians: current status and future prospects. *Bull. Plankton Soc. Japan* 62, 39–43. doi: 10.24763/bpsj.62.1_39

Liu, Y., Wikfors, G. H., Rose, J. M., McBride, R. S., Milke, L. M., and Mercaldo-Allen, R. (2019). Application of Environmental DNA Metabarcoding to Spatiotemporal Finfish Community Assessment in a Temperate Embayment. *Front. Mar. Sci.* 6:674. doi: 10.3389/fmars.2019.00674

Macheriotou, L., Guilini, K., Bezerra, T. N., Tytgat, B., Nguyen, D. T., Nguyen, T. X. P., et al. (2019). Metabarcoding free-living marine nematodes using curated 18S and CO1 reference sequence databases for species-level taxonomic assignments. *Ecol. Evol.* 9, 1211–1226. doi: 10.1002/ece3.4814

Madin, L. P., and Harbison, G. R. (1978). Bathocyroe fosteri gen.nov., sp.nov.: a mesopelagic ctenophore observed and collected from a submersible. *J. Mar. Biol. Associat. U K.* 58, 559–564. doi: 10.1017/S0025315400041217

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200

McEwen, G. F., Johnson, M. W., and Folsom, T. R. (1954). A statistical analysis of the performance of the folsom plankton sample splitter, based upon test observations. *Arch. Met. Geoph. Biokl. A* 7, 502–527. doi: 10.1007/BF02277939

McQuillan, J. S., and Robidart, J. C. (2017). Molecular-biological sensing in aquatic environments: recent developments and emerging capabilities. *Curr. Opin. Biotechnol.* 45, 43–50. doi: 10.1016/j.copbio.2016.11.022

Minamoto, T., Fukuda, M., Katsuhara, K. R., Fujiwara, A., Hidaka, S., Yamamoto, S., et al. (2017). Environmental DNA reflects spatial and temporal jellyfish distribution. *PLoS One* 12:e0173073. doi: 10.1371/journal.pone.0173073

Nevers, M. B., Byappanahalli, M. N., Morris, C. C., Shively, D., Przybyla-Kelly, K., Spoljaric, A. M., et al. (2018). Environmental DNA (eDNA): A tool for quantifying the abundant but elusive round goby (Neogobius melanostomus). *PLoS One* 13:e0191720. doi: 10.1371/journal.pone.0191720

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2016). *vegan: Community Ecology Package. R package version 2.4-3*. Vienna: R Foundation for Statistical Computing.

Orlowski, A. (1990). Hydroacoustic characteristics of scattering layers in the Northeastern Atlantic Ocean. *J. Acoust. Soc. Am.* 88, 298–309. doi: 10.1121/1.399952

Pauvert, C., Buée, M., Laval, V., Edel-Hermann, V., Fauchery, L., Gautier, A., et al. (2019). Bioinformatics matters: The accuracy of plant and soil fungal community data is highly dependent on the metabarcoding pipeline. *Fungal Ecol.* 41, 23–33. doi: 10.1016/j.funeco.2019.03.005

Pearman, J. K., El-Sherbiny, M. M., Lanzén, A., Al-Aidaroos, A. M., and Irigoien, X. (2014). Zooplankton diversity across three Red Sea reefs using pyrosequencing. *Front. Mar. Sci.* 1:27. doi: 10.3389/fmars.2014.00027

Piaggio, A. J., Engeman, R. M., Hopken, M. W., Humphrey, J. S., Keacher, K. L., Bruce, W. E., et al. (2014). Detecting an elusive invasive species: a diagnostic PCR to detect Burmese python in Florida waters and an assessment of persistence of environmental DNA. *Mol. Ecol. Resour.* 14, 374–380. doi: 10.1111/1755-0998.12180

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., and Waits, L. P. (2014). Factors influencing detection of eDNA from a stream-dwelling amphibian. *Mol. Ecol. Resour.* 14, 109–116. doi: 10.1111/1755-0998.12159

Pinheiro, H. T., Moreau, C. S., Daly, M., and Rocha, L. A. (2019). Will DNA barcoding meet taxonomic needs? *Science* 365, 873–874. doi: 10.1126/science.aay7174

Porter, T. M., and Hajibabaei, M. (2018). Automated high throughput animal CO1 metabarcode classification. *Sci. Rep.* 8, 1–10. doi: 10.1038/s41598-018-22505-4

Proud, R., Kloser, R. J., Handegard, N. O., Cox, M. J., and Brierley, A. S. (2019). Mapping the global prey-field: Combining acoustics, optics and net samples to reduce uncertainty in estimates of mesopelagic biomass. *J. Acoust. Soc. Am.* 146, 2898–2898. doi: 10.1121/1.5137055

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Ricker, W. E. (1973). Linear Regressions in Fishery Research. *J. Fisher. Board Can.* 30, 409–434. doi: 10.1139/f73-072

Robison, B. H. (2004). Deep pelagic biology. *J. Exp. Mar. Biol. Ecol.* 300, 253–272. doi: 10.1016/j.jembe.2004.01.012

Robison, B. H. (2009). Conservation of Deep Pelagic Biodiversity. *Conserv. Biol.* 23, 847–858. doi: 10.1111/j.1523-1739.2009.01219.x

Ruppert, K. M., Kline, R. J., and Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Glob. Ecol. Conserv.* 17:e00547. doi: 10.1016/j.gecco.2019.e00547

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

Santoferrara, L. F. (2019). Current practice in plankton metabarcoding: optimization and error management. *J. Plankton Res.* 41, 571–582. doi: 10.1093/plankt/fbz041

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

Sawaya, N. A., Djurhuus, A., Closek, C. J., Hepner, M., Olesin, E., Visser, L., et al. (2019). Assessing eukaryotic biodiversity in the Florida Keys National Marine Sanctuary through environmental DNA metabarcoding. *Ecol. Evol.* 9, 1029–1040. doi: 10.1002/ece3.4742

Sigsgaard, E. E., Carl, H., Møller, P. R., and Thomsen, P. F. (2015). Monitoring the near-extinct European weather loach in Denmark based on environmental DNA from water samples. *Biol. Conserv.* 183, 46–52. doi: 10.1016/j.biocon.2014.11.023

Sigsgaard, E. E., Nielsen, I. B., Carl, H., Krag, M. A., Knudsen, S. W., Xing, Y., et al. (2017). Seawater environmental DNA reflects seasonality of a coastal fish community. *Mar. Biol.* 164:128. doi: 10.1007/s00227-017-3147-4

Sommer, S. A., Woudenberg, L. V., Lenz, P. H., Cepeda, G., and Goetze, E. (2017). Vertical gradients in species richness and community composition across the twilight zone in the North Pacific Subtropical Gyre. *Mol. Ecol.* 26, 6136–6156. doi: 10.1111/mec.14286

St. John, M. A., Borja, A., Chust, G., Heath, M., Grigorov, I., Mariani, P., et al. (2016). A Dark Hole in Our Understanding of Marine Ecosystems and Their Services: Perspectives from the Mesopelagic Community. *Front. Mar. Sci.* 3:31. doi: 10.3389/fmars.2016.00031

Stat, M., John, J., DiBattista, J. D., Newman, S. J., Bunce, M., and Harvey, E. S. (2019). Combined use of eDNA metabarcoding and video surveillance for the assessment of fish biodiversity. *Conserv. Biol.* 33, 196–205. doi: 10.1111/cobi.13183

Stefanoudis, P. V., Rivers, M., Ford, H., Yashayaev, I. M., Rogers, A. D., and Woodall, L. C. (2019). Changes in zooplankton communities from epipelagic to lower mesopelagic waters. *Mar. Environ. Res.* 146, 1–11. doi: 10.1016/j.marenvres.2019.02.014

Stoeckle, M. Y., Soboleva, L., and Charlop-Powers, Z. (2017). Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. *PLoS One* 12:175186. doi: 10.1371/journal.pone.0175186

Sutton, T. T. (2013). Vertical ecology of the pelagic ocean: classical patterns and new perspectives. *J. Fish Biol.* 83, 1508–1527. doi: 10.1111/jfb.12263

Suzuki, K. S., Niida, Y., Tsubono, T., Takimoto, H., Kumakura, E., Ishii, H., et al. (2018). Mechanisms underlying heterogeneous distribution of moon jellyfish Aurelia aurita s.l. across a sharp pycnocline. *Mar. Ecol. Prog. Ser.* 591, 229–239. doi: 10.3354/meps12357

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

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

Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., et al. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Mol. Ecol.* 21, 2565–2573. doi: 10.1111/j.1365-294X.2011.05418.x

Thomsen, P. F., Møller, P. R., Sigsgaard, E. E., Knudsen, S. W., Jørgensen, O. A., and Willerslev, E. (2016). Environmental DNA from Seawater Samples Correlate with Trawl Catches of Subarctic, Deepwater Fishes. *PLoS One* 11:165252. doi: 10.1371/journal.pone.0165252

Wacker, S., Fossøy, F., Larsen, B. M., Brandsegg, H., Sivertsgård, R., and Karlsson, S. (2019). Downstream transport and seasonal variation in freshwater pearl mussel (Margaritifera margaritifera) eDNA concentration. *Environ. DNA* 1, 64–73. doi: 10.1002/edn3.10

Webb, T. J., Berghe, E. V., and O'Dor, R. (2010). Biodiversity's Big Wet Secret: The Global Distribution of Marine Biological Records Reveals Chronic Under-Exploration of the Deep Pelagic Ocean. *PLoS One* 5:e10223. doi: 10.1371/journal.pone.0010223

West, K. M., Stat, M., Harvey, E. S., Skepper, C. L., DiBattista, J. D., Richards, Z. T., et al. (2020). eDNA metabarcoding survey reveals fine-scale coral reef community variation across a remote, tropical island ecosystem. *Mol. Ecol.* 29, 1069–1086. doi: 10.1111/mec.15382

Wheeler, Q. (2018). Blank Canvas: The Case for Descriptive Taxonomy. *Integr. Comp. Biol.* 58, 1118–1121. doi: 10.1093/icb/icy067

Wiebe, P. H., Morton, A. W., Bradley, A. M., Backus, R. H., Craddock, J. E., Barber, V., et al. (1985). New development in the MOCNESS, an apparatus for sampling zooplankton and micronekton. *Mar. Biol.* 87, 313–323. doi: 10.1007/BF00397811

Wu, S., Xiong, J., and Yu, Y. (2015). Taxonomic Resolutions Based on 18S rRNA Genes: A Case Study of Subclass Copepoda. *PLoS One* 10:131498. doi: 10.1371/journal.pone.0131498

Yamamoto, S., Minami, K., Fukaya, K., Takahashi, K., Sawada, H., Murakami, H., et al. (2016). Environmental DNA as a 'Snapshot' of Fish Distribution: A Case Study of Japanese Jack Mackerel in Maizuru Bay, Sea of Japan. *PLoS One* 11:e0149786. doi: 10.1371/journal.pone.0149786

Zhang, W. G., and Partida, J. (2018). Frontal Subduction of the Mid-Atlantic Bight Shelf Water at the Onshore Edge of a Warm-Core Ring. *J. Geophys. Res. Oceans* 123, 7795–7818. doi: 10.1029/2018JC013794

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 16 March 2021
doi: 10.3389/fevo.2021.620715

Integrating Environmental DNA Results With Diverse Data Sets to Improve Biosurveillance of River Health

Adam J. Sepulveda[1], Andrew Hoegh[2], Joshua A. Gage[3], Sara L. Caldwell Eldridge[4], James M. Birch[5], Christian Stratton[2], Patrick R. Hutchins[1] and Elliott P. Barnhart[4]*

[1] U.S. Geological Survey, Northern Rocky Mountain Science Center, Bozeman, MT, United States, [2] Department of Mathematical Sciences, Montana State University, Bozeman, MT, United States, [3] Gage Cartographics, Bozeman, MT, United States, [4] U.S. Geological Survey, Wyoming-Montana Water Science Center, Helena, MT, United States, [5] Monterey Bay Aquarium Research Institute, Moss Landing, CA, United States

Autonomous, robotic environmental (e)DNA samplers now make it possible for biological observations to match the scale and quality of abiotic measurements collected by automated sensor networks. Merging these automated data streams may allow for improved insight into biotic responses to environmental change and stressors. Here, we merged eDNA data collected by robotic samplers installed at three U.S. Geological Survey (USGS) streamgages with gridded daily weather data, and daily water quality and quantity data into a cloud-hosted database. The eDNA targets were a rare fish parasite and a more common salmonid fish. We then used computationally expedient Bayesian hierarchical occupancy models to evaluate associations between abiotic conditions and eDNA detections and to simulate how uncertainty in result interpretation changes with the frequency of autonomous robotic eDNA sample collection. We developed scripts to automate data merging, cleaning and analysis steps into a chained-step, workflow. We found that inclusion of abiotic covariates only provided improved insight for the more common salmonid fish since its DNA was more frequently detected. Rare fish parasite DNA was infrequently detected, which caused occupancy parameter estimates and covariate associations to have high uncertainty. Our simulations found that collecting samples at least once per day resulted in more detections and less parameter uncertainty than less frequent sampling. Our occupancy and simulation results together demonstrate the advantages of robotic eDNA samplers and how these samples can be combined with easy to acquire, publicly available data to foster real-time biosurveillance and forecasting.

Keywords: climate, detection, molecular, occupancy analysis, river, salmon, streamgage

OPEN ACCESS

Edited by:
Hiroki Yamanaka, Ryukoku University, Japan

Reviewed by:
William Sutton, Tennessee State University, United States
Luca Carraro, Swiss Federal Institute of Aquatic Science and Technology, Switzerland

****Correspondence:***
Adam J. Sepulveda
asepulveda@usgs.gov

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *23 October 2020*
Accepted: *18 February 2021*
Published: *16 March 2021*

Citation:
Sepulveda AJ, Hoegh A, Gage JA, Caldwell Eldridge SL, Birch JM, Stratton C, Hutchins PR and Barnhart EP (2021) Integrating Environmental DNA Results With Diverse Data Sets to Improve Biosurveillance of River Health. Front. Ecol. Evol. 9:620715. doi: 10.3389/fevo.2021.620715

INTRODUCTION

Timely, up-to-date information concerning harmful invasive species and pathogens is critical for minimizing negative outcomes to ecosystem and human health (Stohlgren and Schnase, 2006; Bohan et al., 2017; Cordier et al., 2020). Assimilating this information has been challenging because the abiotic and biotic processes that drive invasive species and pathogen distributions

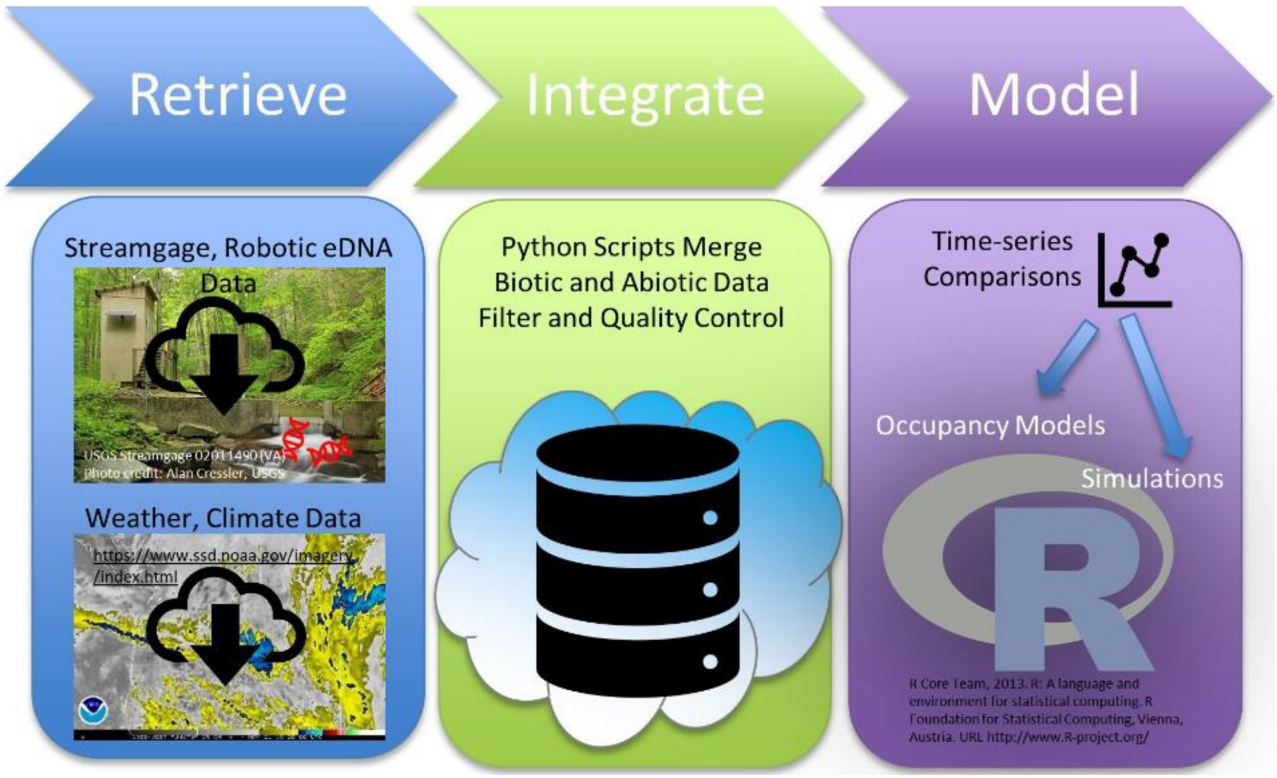


FIGURE 1 | Schematic of the data science workflow applied to our two eDNA case studies.

and abundances from benign to harmful levels often occur at different spatiotemporal scales (Collins et al., 2006; Gallien et al., 2010; Uden et al., 2015). These challenges make measurement, integration, and rapid analysis difficult (Michener and Jones, 2012). For example, marine harmful algal bloom alert bulletins and early warning systems require integrating information about phytoplankton, toxin concentrations within shellfish, water temperature and wind speeds, and ocean or lake circulation forecasts (e.g., Glibert et al., 2018). Automated sensor networks, such as the U.S. Geological Survey's (USGS) streamgage network and National Oceanic and Atmospheric Administration's (NOAA) weather station network, have made it much easier to track changing abiotic conditions with both high and broad spatiotemporal resolution (Sepulveda et al., 2015; Al-Chokhachy et al., 2017; Kovach et al., 2019), but automated, collection of comparable biological data remains a challenge (Sugai, 2020).

Environmental (e)DNA sample collection and subsequent analyses have revolutionized biosurveillance because inferences can be made about species occurrences sight-unseen (e.g., Cristescu and Hebert, 2018). More recently, autonomous robotic eDNA samplers have made it possible to make biological observations match the scale and quality of *in situ* physical and chemical measurements (Yamahara et al., 2019; Sepulveda et al., 2020). Autonomous robots placed within the environment can conduct high frequency (sub-daily) sampling, regardless of location, weather, or the availability of human resources. Satellite communication ports allow results to be uploaded to end-users. Hence, biological data collection at relevant scales is no longer the bottleneck.

One current challenge is rapidly integrating the high frequency data streams produced by autonomous robotic eDNA samplers with the high frequency data streams produced by automated sensor networks tracking abiotic conditions. This integration should enable timely, up-to-date information about biological hazards. Each data stream has nuances (e.g., unique attribute fields); however, the structure of eDNA data streams presents additional complications. This molecular method provides indirect inference about species presence, so occurrence probability must be modeled (Stratton et al., 2020). For eDNA analyses, multiple samples are collected at a location and multiple replicates from each sample are analyzed for the presence of the target organism DNA. Samples taken at the location occupied by a species may not necessarily contain DNA of that target species, just like replicates may lack target DNA even if the DNA is present in the sample (Darling, 2019; Sepulveda et al., 2020). Thus, each sample and replicate may detect the organism's DNA, if present, with some probability. Occupancy models provide a useful framework for the analysis of eDNA results (Stratton et al., 2020), in that these models account for the multiple nested levels of sampling that characterize eDNA surveys. Occupancy models also provide a useful framework for linking

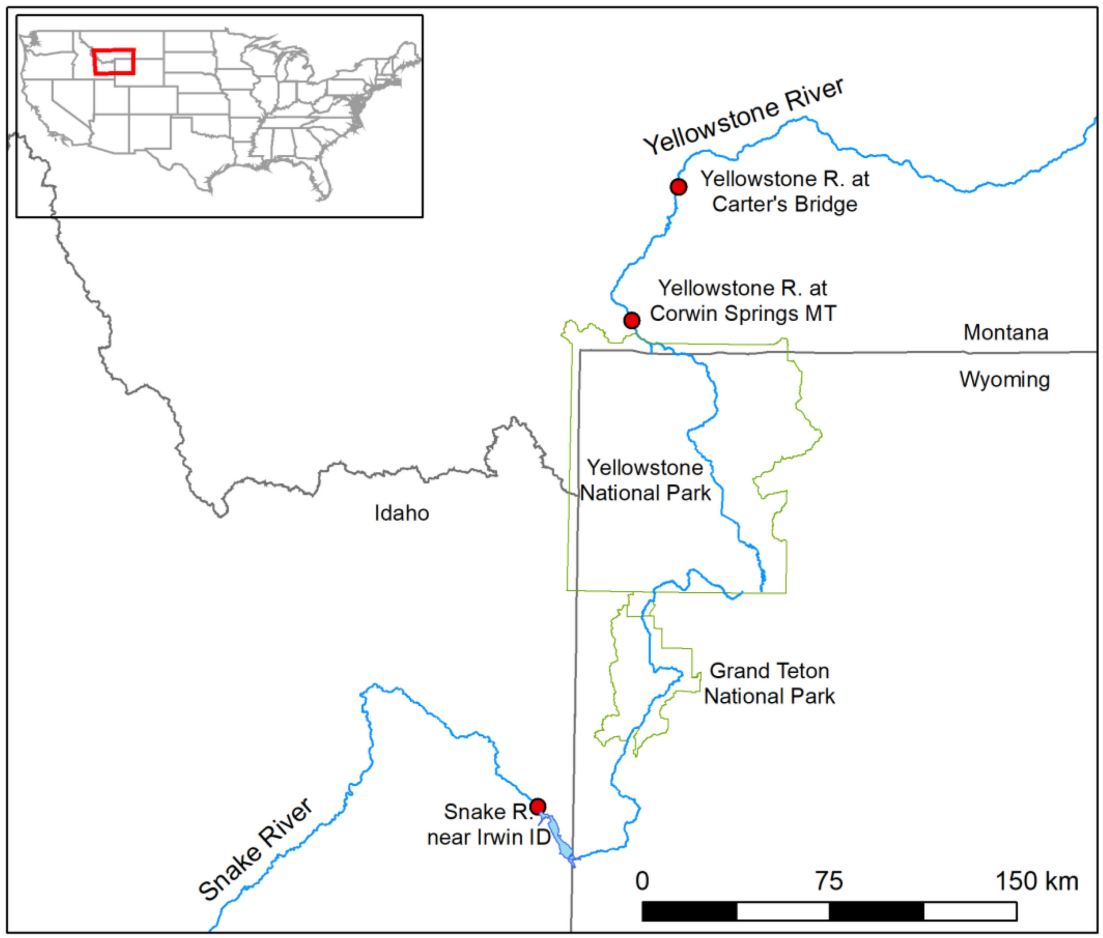


FIGURE 2 | Map of ESP water sampling locations. The red box in the inset map shows the location of the Upper Yellowstone River and Upper Snake River in the United States. The larger map shows the sample site locations (filled red circles) on each river relative to Yellowstone National Park and Grand Teton National Park (outlined in green).

data streams describing abiotic conditions to those describing biological data, such as target taxa DNA (Pilliod et al., 2019; Sepulveda et al., 2019).

Here, we present the constituent parts of a cloud-based, data science pipeline for using Bayesian hierarchical occupancy models to analyze relationships between the high frequency data streams produced by autonomous robotic eDNA samplers and the high frequency data streams produced by automated sensor networks tracking abiotic conditions (**Figure 1**). To demonstrate the general applicability of this workflow, we applied it to a dataset typical of an eDNA early detection program, where target taxa DNA was rarely detected, and to a dataset typical of an eDNA monitoring program, where target taxa DNA detections were more common. To evaluate the added value of high frequency, autonomous samples, we also evaluated how uncertainty in result interpretation changes with the frequency of autonomous robotic eDNA sample collection for each dataset. Our workflow is intended to serve as a prototype for how high-throughput eDNA data can be combined with easy to acquire, publicly available data to foster real-time biosurveillance and forecasting.

MATERIALS AND METHODS

eDNA Datasets

We applied the data science workflow to two datasets described in Sepulveda et al. (2020). In the first dataset used as an example of eDNA early detection programs, autonomous robotic eDNA samplers collected samples at two USGS streamgage sites in the Yellowstone River of Montana (United States): USGS 06191500 Yellowstone River at Corwin Springs MT and USGS 06192500 Yellowstone River near Livingston MT, described below as Corwin Springs and Carters Bridge, respectively (**Figure 2**). Robotic eDNA samplers were programmed to collect 1-L samples every 12 h, from July 24 to August 26, 2018, and every 3 h from August 27 to September 7, 2018. Samples were analyzed for DNA of the fish pathogen, *Tetracapsuloides bryosalmonae*, the causative agent of salmonid fish Proliferative Kidney Disease (PKD), which has resulted in large salmonid mortality events in this region and also in Europe (Hutchins et al., in press). Detections of *T. bryosalmonae* DNA were rare, as only 5 of 256 samples were scored as positive (Sepulveda et al., 2020).

In the second dataset used as an example of eDNA monitoring programs for more common species, an autonomous robotic eDNA sampler collected water samples at one USGS streamgage on the Snake River of Idaho (United States): USGS 13032500 Snake River near Irwin ID (**Figure 2**). This streamgage is 1.5 km downstream of Palisades Reservoir (WY/ID). The robotic eDNA sampler was programmed to collect 2-L samples every 12 h from July 17 to September 09 and then every 4 h from September 10 to October 01, 2019. Samples were analyzed for *Oncorhynchus nerka* (kokanee salmon) DNA; *O. nerka* occur upstream in Palisades Reservoir. Thirty-five of 128 samples were scored as positive for *O. nerka* DNA (Sepulveda et al., 2020).

The Monterey Bay Aquarium Research Institute's (MBARI) robotic instrument, the Environmental Sample Processor (ESP), was used to collect eDNA samples in both datasets. The ESP is a robotic device that can be programmed to automate water sample filtration and preservation of the captured material or homogenize it for immediate analyses *in situ*. Various iterations of the instrument have been realized over the past 25 years; eDNA samples in both datasets were collected using the "second-generation" (2G) ESP and its archival capabilities to filter water samples and preserve the collected material for later analysis in the laboratory (Scholin et al., 2017).

Details of the ESPs' operation and eDNA analysis methods used to generate the two data sets are summarized here as they were described in Scholin et al. (2017 and references therein) and Sepulveda et al. (2020). The ESP operated autonomously, needing only power, communications and fluid connections through a waterproof pressure housing. At the initiation of sampling, a small container (a "puck") loaded with 1.2-μm cellulose nitrate filter material was placed within a clamp. Valves opened to the outside, allowing a syringe to sequentially pull water through the puck. Once the target volume was filtered, or the filter was loaded with biomass (i.e., "clogged"), filtering stopped, and excess water was cleared. Five mL's of RNAlater preservative was then added to the puck, soaking the filter for 10 min before excess was evacuated and the puck was returned to storage. To reduce carry-over contamination, the sampling pump, tubing and external sampling module were flushed with river water for 10 min prior to every sample collection. The sampling port of the ESP itself was cleaned with 10% bleach and a 10% Tween-20 solution between samples. Two negative controls (1 L of molecular grade water) were run through each ESP prior to and at the conclusion of each deployment to assess for contamination. Metadata associated with each sample were communicated via telemetry after each sampling point.

At the end of each ESP deployment, pucks were removed, and filters were aseptically recovered into 2.0-mL screw cap centrifuge tubes, and then shipped frozen to the USGS Upper Midwest Environmental Science Center (La Crosse, WI, United States) for DNA extraction and quantitative PCR analyses. All samples were analyzed in four replicate 25-μL reactions and tested for PCR inhibition. Samples were scored as positive when one or more PCR replicates amplified for the target DNA. Field, extraction and PCR negative controls were analyzed as regular samples; no negative controls amplified.

TABLE 1 | Physical data available at daily time steps that were collated on a cloud-hosted database and then integrated with raw eDNA data in multi-scale occupancy models.

Data stream	Variable	Metric	Yellowstone	Snake
PRISM	Precipitation	Total	+	+
	Air temperature	Minimum, Maximum, Mean	+	+
	Vapor pressure deficit	Minimum, Maximum	+	+
	Dew point temperature	Mean	+	+
National Water Quality Portal	pH	Mean, Maximum, Minimum, St. Deviation, Count	–	–
	Dissolved calcium	Mean, Maximum, Minimum, St. Deviation, Count	–	–
	Water temperature	Mean, Maximum, Minimum, St. Deviation, Count	+	–
USGS National Water Information System	Water temperature	Mean, Maximum, Minimum	+	+
	Discharge	Mean, Maximum, Minimum	+	+

Abiotic Data Streams

We collated spatially and temporally explicit data from three data streams: (1) 800-m gridded climate data served from Oregon State University's PRISM, (2) water quality data served from the National Water Quality Portal, and (3) water quantity and quality data served from the USGS water services portal. Data attributes are described in **Table 1** and python scripts are available in the **Supplementary Material**. Spatial components of these data were delineated by the location of the eDNA sampling at USGS streamgage sites. Temporal components were reduced to daily time steps.

We processed and corrected the downloaded datasets. Sites with multiple observations for each day were aggregated, resulting in one average value for each site for each day. Additionally, some sites used different units of measurement such as temperature in Fahrenheit and Celsius so we standardized units across all sites and dates. In some cases, we removed white spaces from data entries to create usable numeric values. Additional columns not used in analyses were dropped from the datasets before importing data into the database.

Occupancy Analyses

We used Bayesian multi-scale occupancy models in the msocc package (Stratton et al., 2020; R version 3.5.2) to estimate the detection probability of *T. bryosalmonae* and *O. nerka* DNA, to gain insight about covariates associated with eDNA detection probability, and to estimate the effort needed for confident and high-probability detection of target eDNA. These models allow for the analysis of occupancy with three levels of hierarchical sampling, while also accounting for false negatives in detection.

TABLE 2 | Watanabe-Akaike information criterion (WAIC) scores of the Yellowstone and Snake river candidate models.

River	Model			WAIC	
	ψ	θ	p	t	t-1
Yellowstone	Site	Date		1085	1115
	Site	Site + Date		1102	1097
	Site	Water temperature		1129	1134
	Site	Dew point temperature mean		1131	1140
	Site	Sample volume		1135	–
	Site	Air temperature maximum		1138	1146
	Site	Air temp mean		1141	1148
	Site	Air temperature minimum		1142	1144
	Site	Discharge		1154	1148
	Site	Precipitation		1168	11111
	Site			1179	–
	Site	Site		1318	–
				1366	–
Snake		Sample volume		1433	–
		Discharge		1435	1443
		Date		1436	1440
		Dew point temperature mean		1436	1446
		Air temperature mean		1442	1446
		Air temperature minimum		1442	1448
		Precipitation		1446	1454
				1448	–
		Air temperature maximum		1458	1455

WAIC scores are show for models that included model covariates with (t-1) and without (t) a 1-day lag.

Models were used to estimate (1) ψ, the probability of occurrence of target eDNA at each of three streamgage sites; (2) θ the conditional probability of occurrence of target eDNA in each sample given that target eDNA was present at that site; and (3) p, the conditional probability of detection of target eDNA in each qPCR replicate of an eDNA sample given that target eDNA was present in the sample.

We used python scripts (described in the **Supplementary Material**) to merge our eDNA datasets with the abiotic data streams and to format and download the input data frames (detection data, site-level data with covariates, sample-level data with covariates, and replication-level data) required by the msocc package. Site-level covariates were streamgage site and date. Sample-level covariates were streamgage site, date, time, eDNA sample volume and the climate, water quality and quantity data listed in **Table 1**. We also added a 1-day lag to the variables listed in **Table 1** to assess if eDNA detections were associated with the prior day's conditions. We modeled p as constant since DNA extraction methods and laboratory analyses were the same for all samples.

We tested simple models that fit ψ and/or θ by each individual covariate or by combinations of covariates that were not correlated ($R < 0.7, p > 0.5$). We used the Watanabe-Akaike Information criterion (WAIC) to compare support for models fitted with and without covariates; models with lower WAIC values are favored (Gelman et al., 2014). We then computed estimates of the derived parameters ψ and θ for the most favored model. These estimates and their standard errors were computed using a Markov chain containing 10,000 iterations (excluding the first 1000 warm-up iterations).

Finally, we ran post-hoc power analyses to evaluate how sample size (i.e., the number of water samples and the number of PCR replicates) influenced the precision of θ estimates (msocc package, msocc_sim() as described in Stratton et al., 2020). We used the estimates of ψ, θ, and p from the most supported models of each dataset to simulate detection data. For these simulations, we varied the number of samples collected at each sampling event and the PCR replicates analyzed per sample. We then replicated this process 100 times and assessed whether the 95% credibility intervals for each parameter captured the value that generated the data. We also recorded the width of the credibility intervals. The sample sizes at which the proportion of the 95% credibility intervals that contained the original parameters stabilize and the average width of the credibility intervals stabilize provide insight about the point of diminishing returns, beyond which increasing sample size provides little benefit.

Simulations

We ran three types of simulations to assess how robotic eDNA sampling strategies can result in more detections of rare organisms and more precise estimates. In simulation 1 and 2, we explored whether high-frequency sampling can better detect target taxa DNA than lower-frequency sampling when

TABLE 3 | Posterior mean estimates (±95% CI) of ψ, θ, and p from the candidate models with the lowest WAIC scores.

River	Model	ψ		θ		p
Yellowstone	ψ(Site) θ(Date) p(.)	Carter's Bridge	0.83 (0.37–1.00)	Minimum	0.02 (0.01–0.04)	0.32 (0.14–0.53)
		Corwin Springs	0.87 (0.40–1.00)	Maximum	0.04 (0.01–0.08)	
Snake	ψ(.) θ(Date) p(.)		0.67 (0.22–0.98)	Minimum	0.25 (0.14–0.39)	0.39 (0.31–0.47)
				Maximum	0.41 (0.28–0.55)	
	ψ(.) θ(Discharge) p(.)		0.67 (0.23–0.97)	Minimum	0.25 (0.13–0.40)	0.39 (0.31–0.47)
				Maximum	0.46 (0.28–0.67)	
	ψ(.) θ(Sample volume) p(.)		0.66 (0.22–0.97)	Minimum	0.18 (0.08–0.33)	0.39 (0.30–0.47)
				Maximum	0.58 (0.35–0.80)	

Minimum and maximum estimates of θ are displayed to show the range of values associated with Date, Discharge or Sample volume covariates.

TABLE 4 | Estimates of the regression coefficients from the top models.

River	Parameter	Covariate	Mean	95% CI
Yellowstone	ψ	Site	1.128	−2.927 – 5.461
	θ	Date	−0.018	−0.038 – 0.002
Snake	θ	Date	0.012	−0.002 – 0.026
		Discharge	$-1.345e^{-4}$	$-3.178e^{-4}$ – $4.065e^{-5}$
		Sample volume	$-1.479e^{-3}$	-2.815 – $-2.255e^{-4}$

controlling for the total number of samples. High-frequency sampling consisted of daily samples for 8 weeks; whereas lower-frequency sampling consisted of a batch of seven samples on a single day once per week for 8 weeks. We modeled two ways to think about collecting a batch of seven samples on a single day. The first approach treats the batch of samples as subsamples; in other words, the organism is present, or not, and each subsample has a probability of detecting the organism, given that it is present. The second approach treats the seven samples as independent samples, where for each sample the organism is present or not, and the sample can be detected with a given probability of the organism being present. In practice, either scenario is plausible, but it likely depends on the underlying ecological processes and how samples are collected. If the sampling process involves collecting a set of water samples at an individual location and time point, then the first approach, subsampling, may be relevant. However, if samples could be spread out over the day, and potentially space, then using independent samples may be reasonable. Data were simulated from an occupancy model framework where,

$$Z_i \sim Bernoulli(\psi_i),$$

$$Y_{itj} \sim Binomial(n_j, z_i\rho_{it}),\ \rho_{it} = \frac{exp(\chi_\mu \mathrm{B})}{1 + exp(\chi_\mu \mathrm{B})},$$

where Z_i is the latent occupancy at site i, ψ_i is the probability that the species is present at site i, Y_{itj} is the observed occupancy at site i, time t, and for the jth replicate, and ρ_{it} is the probability that the species will be detected at site i and time t, given presence.

Each simulated data set was summarized by whether or not a rare species was detected at least once.

In simulation 1, we used a fixed occupancy probability (ψ) and fixed detection probability (ρ) and explored the impact of different levels of ψ = {0.05, 0.10, 0.15}, ρ = {0.05, 0.10, 0.15}, and the total number of samples on a high-frequency vs. lower-frequency sampling approach. Simulation 2 was related to simulation 1 but allowed the detection probability (ρ) to vary for subsamples collected on the same day. For these subsamples, ρ varied stochastically on a day-to-day basis. Formally, $X\mu\beta = \beta_0$ was simulated from a biased random walk such that ρ had a median value of either {0.05, 0.10, 0.15}. Code to recreate the simulations and figures is available in the

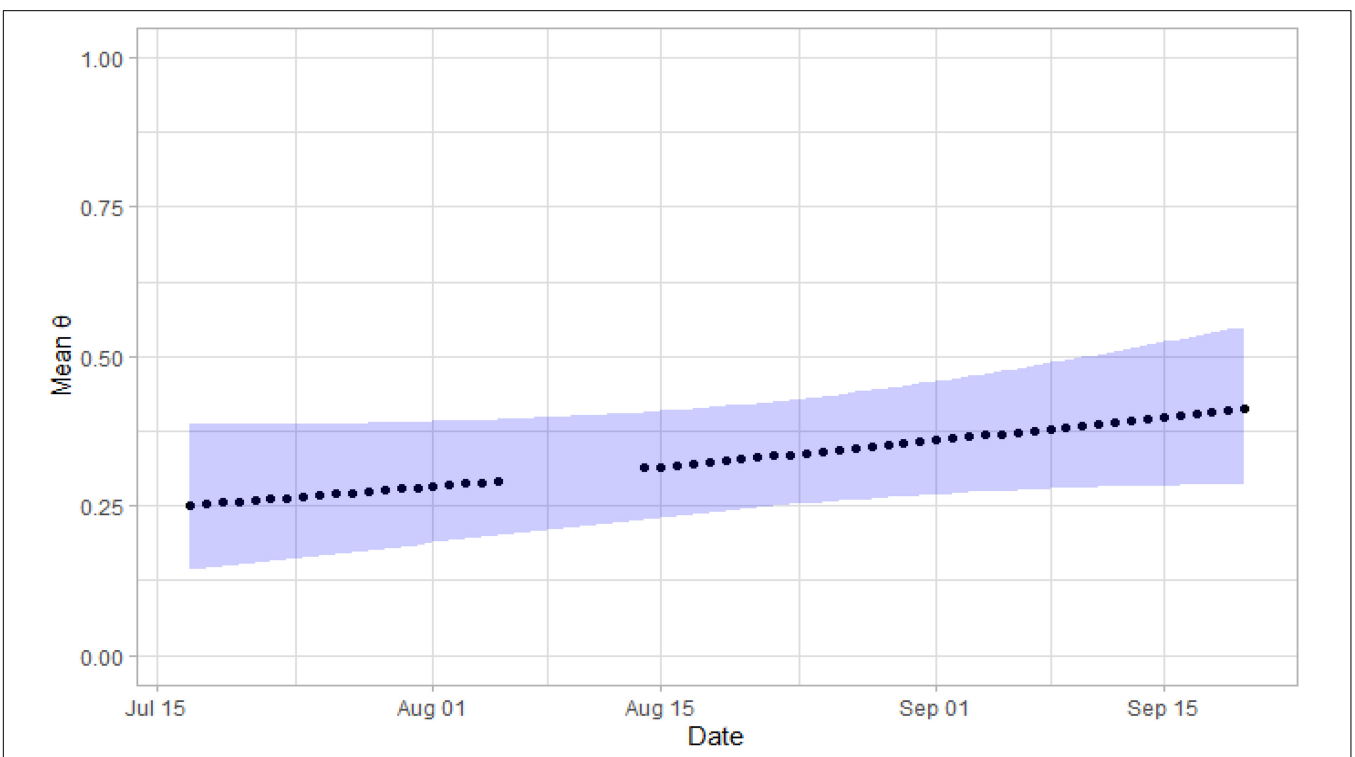


FIGURE 3 | Posterior mean estimates of θ relative to Julian date for *O. nerka* DNA in the Snake River, 2019. The purple band indicates the width of the 95% confidence intervals. No sampling occurred in early August because of ESP mechanical difficulties.

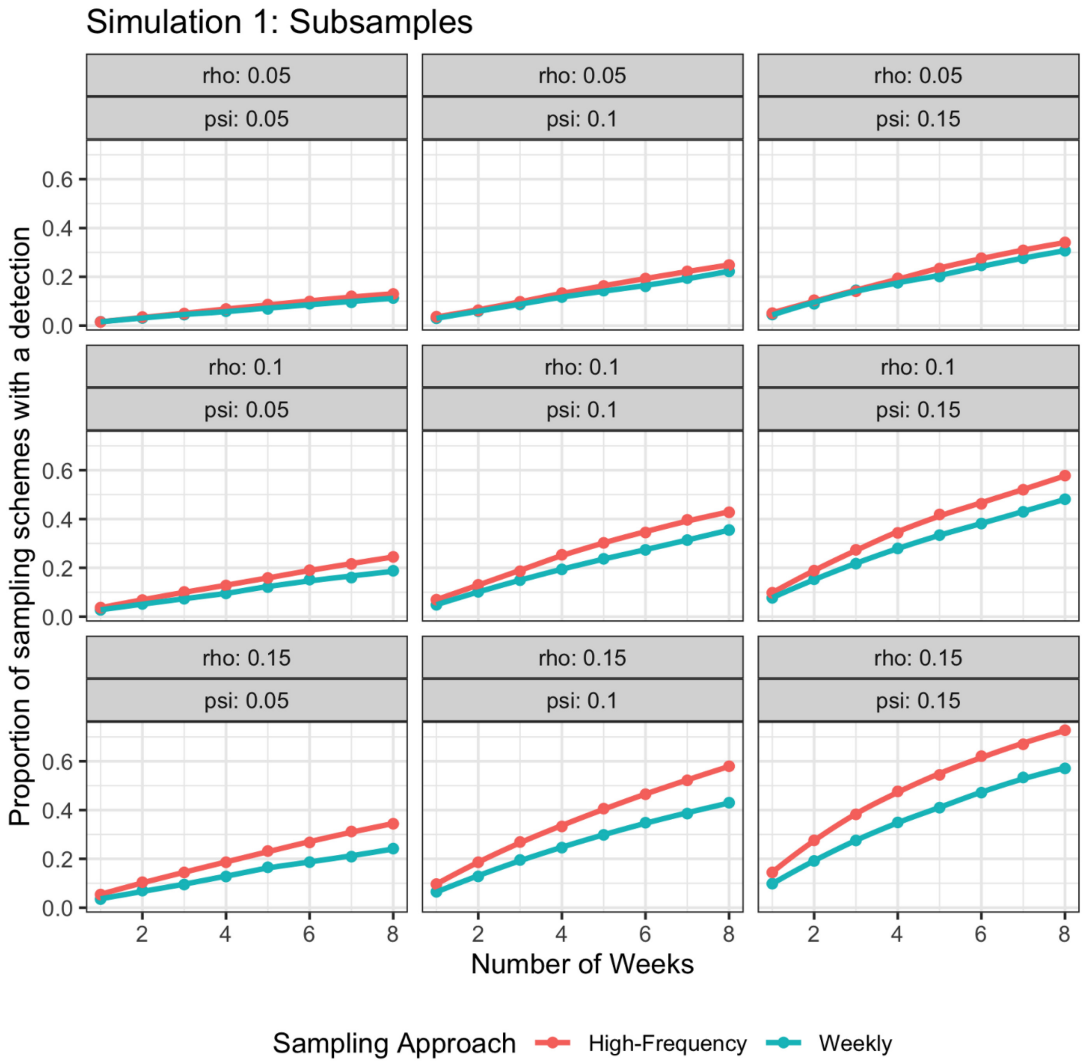


FIGURE 4 | Probability of detection as a function of *rho*, *psi*, and number of weeks; *rho* does not vary by time. The high-frequency approach collects samples daily and the weekly approach takes seven samples once a week. Weekly samples are modeled as non-independent subsamples.

Supplementary Material. The detection probability was the same for all independent samples within a day. We compared the total frequency of the sampling regimes, either daily samples for a certain number of weeks or seven samples collected for a certain number of weeks, that result in at least one detection in simulations 1 and 2.

In simulation 3, we randomly selected eDNA samples from each of the Yellowstone and Snake River datasets using the following temporal frequencies: 1 or 2 samples per day; 1, 2, or 3 samples per week; and 1 sample per month. The total number of samples differed for each time step. We replicated the random selection 100 times per time step. For each random selection, we used the msocc package to estimate θ using the null model without covariates: $\psi(.)$, $\theta(.)$, and $p(.)$. We compared the distributions of the 95% credibility interval widths of the θ estimates for each temporal frequency scheme.

RESULTS

Rare Species Detection Sites

Target DNA was detected in three of 128 ESP samples at Corwin Springs and in two of 128 ESP samples at Carters Bridge. Sampling date was significantly correlated with river discharge ($R = 0.90$, $p < 0.05$). Mean air temperature was significantly correlated with minimum and maximum air temperatures ($R = 0.9, p < 0.05$). Models lacking time lags were more supported than those with lags.

The most supported model was ψ(Site), θ(Date), $p(.)$ (**Table 2**). Posterior mean estimates of ψ were higher for Corwin Springs (0.87) than for Carter's Bridge (0.83), though confidence intervals had large overlap (**Table 3**). Raw estimates of the site coefficient were positive, but confidence intervals overlapped zero (**Table 4**). Posterior mean estimates of θ were near zero (0.02 – 0.04), but were highest on Aug 16 and lowest on Aug 08 (**Table 3**).

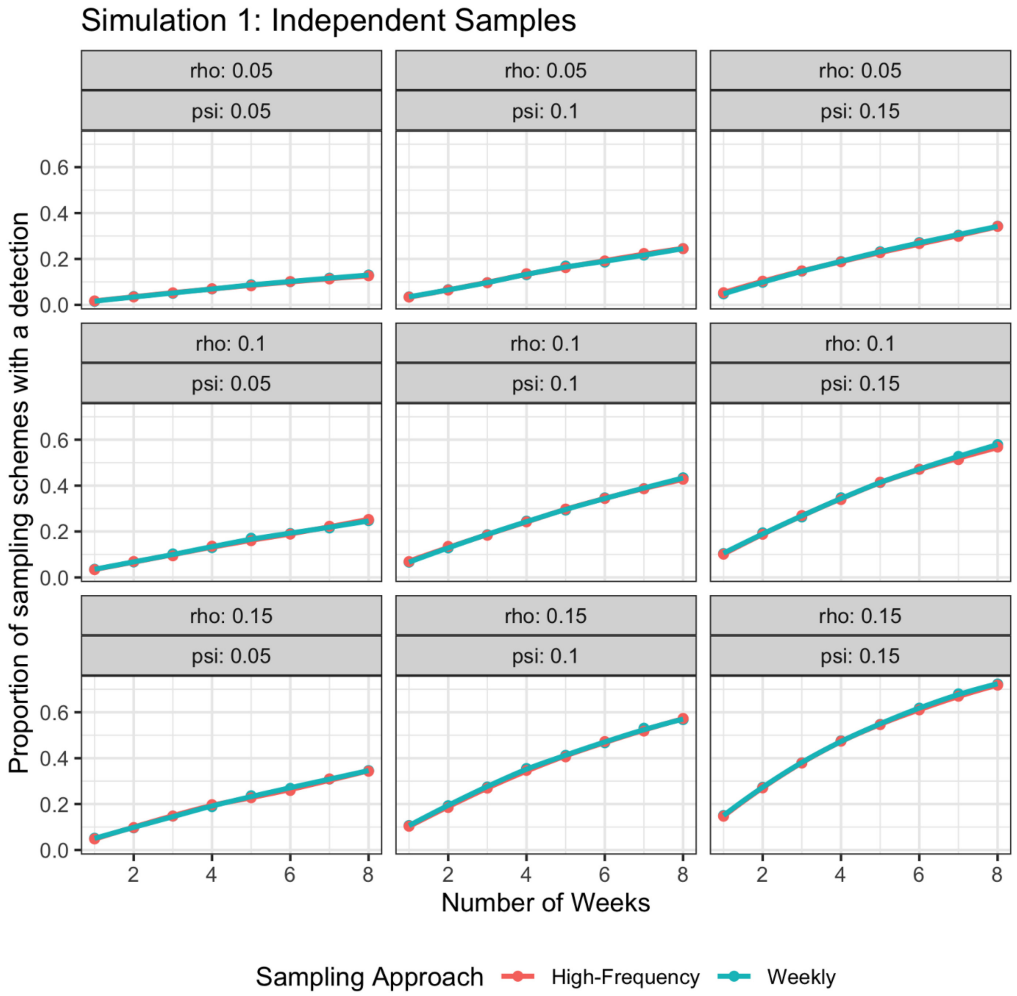


FIGURE 5 | Probability of detection as a function of *rho*, *psi*, and number of weeks; *rho* does not vary by time. The high-frequency approach collects samples daily and the weekly approach takes seven samples once a week. Weekly samples are modeled as independent samples.

There was no discernable temporal pattern in θ. Raw estimates of the date coefficient were negative, but confidence intervals overlapped zero (**Table 4**).

Power analyses indicated seven PCR replicates are required to reduce uncertainty in θ; the width of 95% credibility intervals for θ decreased from 0.59 at four PCR replicates to 0.05 at seven PCR replicates.

Common Species Detection Site

Target DNA was detected in 35 of 128 ESP samples. Sample date was significantly correlated with river discharge ($R = -0.8$, $P < 0.05$) and sample volume ($R = -0.8$, $P < 0.05$); river discharge and sample volume declined with sample date. River discharge and sample volume were also significantly correlated ($R = 0.8$, $P < 0.05$); smaller volumes of water were sampled when discharge was lower later in the study. Models with time lags were less supported than those without.

The most supported models included water sample volume, river discharge, or date as covariates of θ. These covariates were correlated with one another and could not be combined into a single model. Posterior mean estimates of θ increased, as water sample volume decreased, river discharge decreased, and date increased (**Figure 3**). Posterior mean estimates of θ ranged from 0.18 (0.08 – 0.33) to 0.58 (0.35 – 0.81) when sample volumes were 0.6 L. Raw estimates of the date coefficient were positive, while raw estimates of the discharge and sample volume coefficients were negative (**Table 4**). Only confidence intervals for the sample volume coefficient did not overlap zero (**Table 4**). Power analyses using parameters from the top three models indicated that more PCR replicates did not result in θ estimate precision gains.

Simulations

For Simulation 1, when the batch of weekly samples were considered as subsamples, high frequency sampling (i.e., daily samples) had a larger proportion of sampling schemes that resulted in at least one positive detection than lower frequency sampling (i.e., weekly samples) (**Figure 4**). The probability of at least one detection for daily samples was still $1 - (1 - \psi\rho)^{7n}$, but

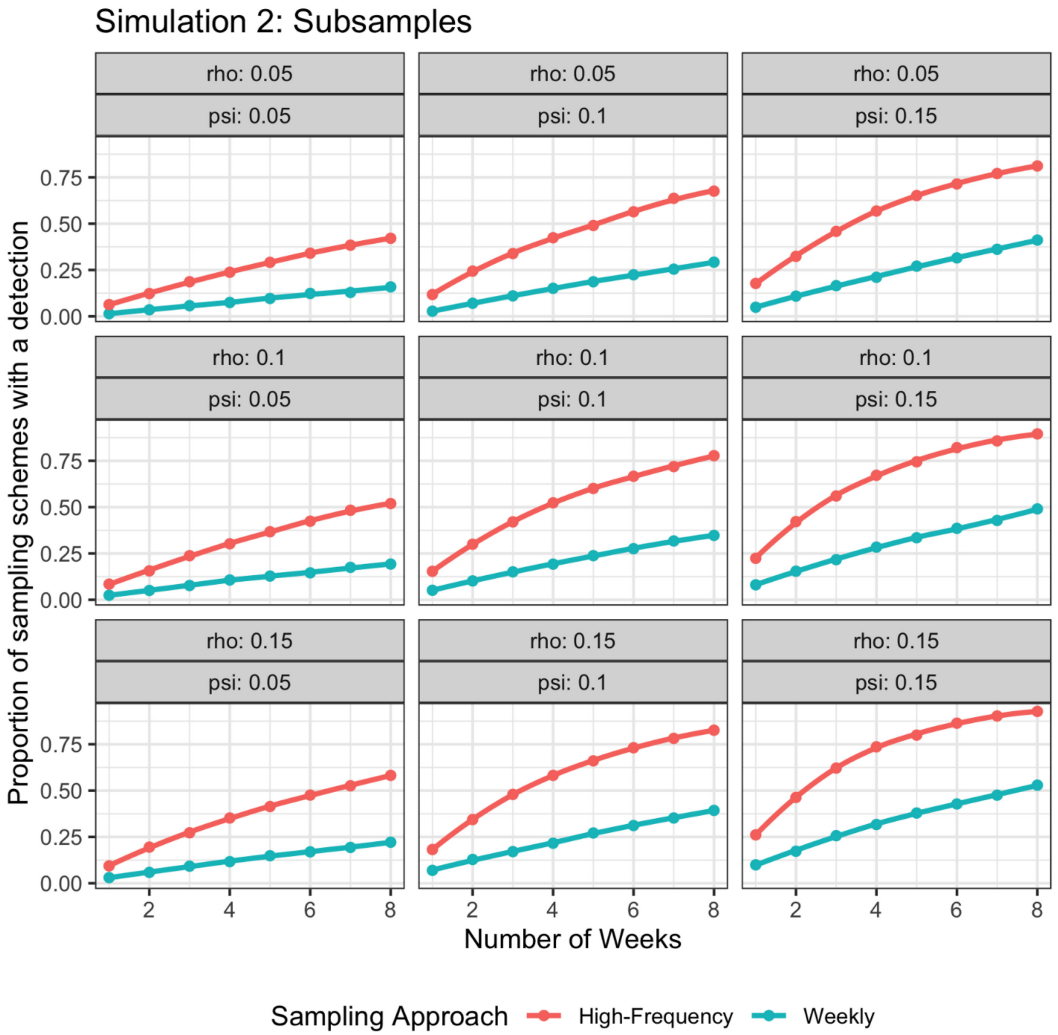


FIGURE 6 | Probability of detection as a function of *rho*, *psi*, and number of weeks; *rho* varies stochastically on a day-to-day basis. The high-frequency approach collects samples daily and the weekly approach takes seven samples once a week. Weekly samples are modeled as non-independent subsamples.

the probability of at least one detection for weekly samples is $1 - \left[1 - \psi\left[1 - (1-\rho)^7\right]\right]^n$. When the batch of weekly samples was considered independent, there were no discernable differences between the detection outcomes of higher vs. lower frequency sampling (**Figure 5**). For both daily samples and weekly batches of samples, the probability of no detection for a single sample is $1 - \psi\rho$; hence, the probability of at least one detection is $1 - (1 - \psi\rho)^{7n}$, where n is the number of weeks of sampling.

For simulation 2, when detection probability changed with time, the higher frequency sampling approach had a higher detection rate than the lower frequency approach regardless of if the batch of weekly samples were modeled as subsamples or independent (**Figures 6**, **7**).

For simulation 3, precision of the posterior mean θ estimates was highest (i.e., lower confidence interval width) when ESP samples were collected at least once per day in the Yellowstone River, where there were very few positive detections of the target (**Figure 8**). However, many of the precision values associated with these higher-frequency samples were still extremely low (i.e., higher confidence interval width), indicating that even high frequency sampling cannot reduce θ estimate uncertainty when target taxa are infrequently detected. In the Snake River, where positive detections of the target were more common, higher frequency sampling did increase precision (**Figure 8**). Sampling three times per week increased precision compared to sampling less frequently and sampling at least once per day increased precision compared to sampling three times per week. There were no gains in precision when sampling once vs. twice per day.

DISCUSSION

We used data from the Yellowstone River, Montana, and the Snake River, Idaho, to demonstrate how real-time data collected by autonomous samplers and automated sensors could be integrated into a data science pipeline to provide

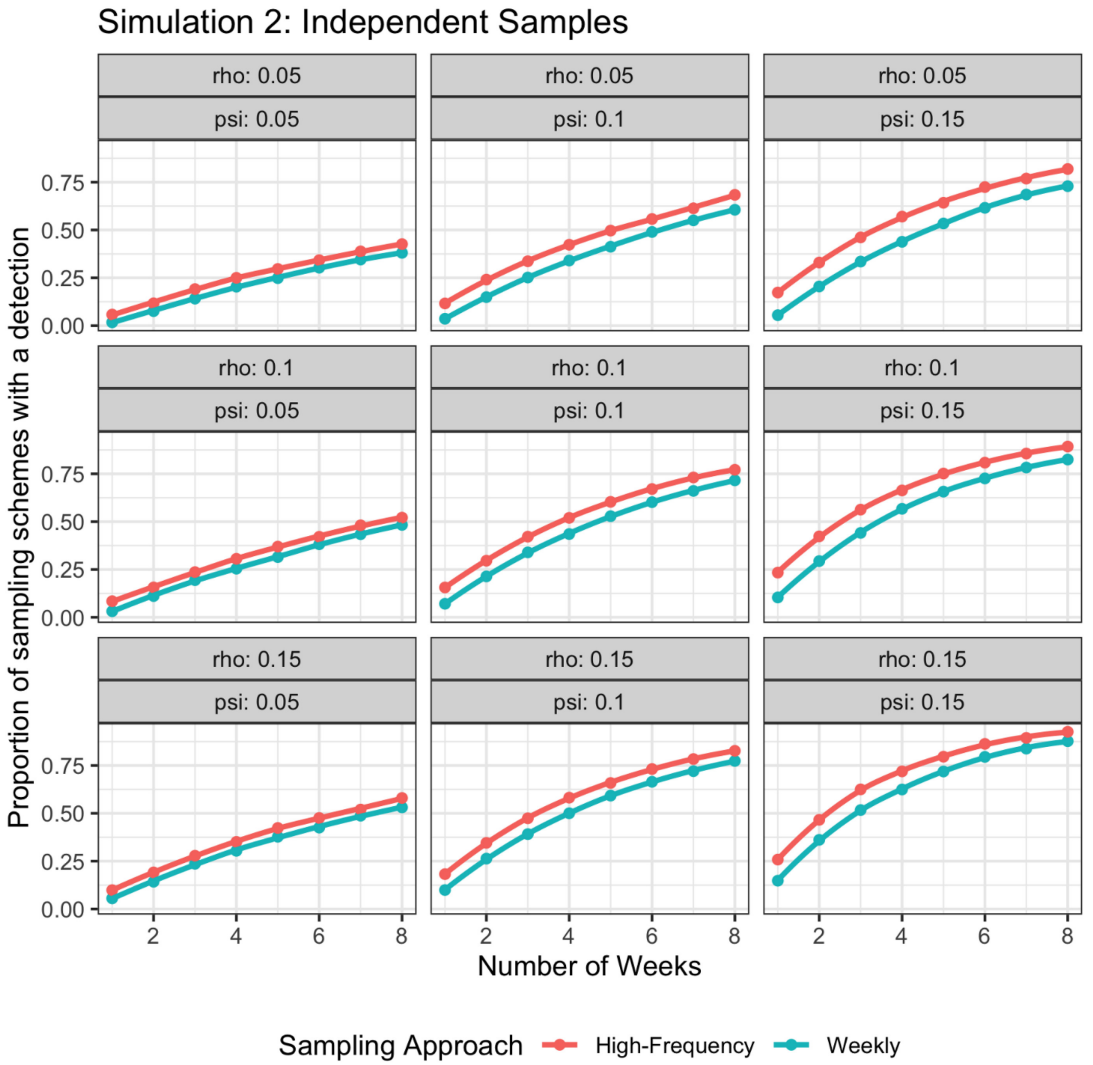


FIGURE 7 | Probability of detection as a function of *rho*, *psi*, and number of weeks; *rho* varies stochastically on a day-to-day basis. The high-frequency approach collects samples daily and the weekly approach takes seven samples once a week. Weekly samples are modeled as independent samples.

timely information about aquatic ecosystem health. Furthermore, through simulations, we showed that data collected from autonomous samplers that enable high frequency eDNA sampling improved the accuracy and precision of inferences. Taken together, our field results and simulations indicate that rare species can be difficult to detect via eDNA sampling, and considerably more sampling may be required for detections and strong inference. Additionally, automation of the eDNA collection process and analysis of the collected data simplifies the task of data collection and repeatability of a study. Our study underscores the promise of new technologies to deliver actionable information at scales and timeframes relevant to decision makers.

Components of our analytical workflow were eDNA detection data and publicly available water-quality and climate data that were collated on a cloud-hosted database and then downloaded into data frames easily processed by computationally efficient multi-scale occupancy models (**Figure 1**). In our case studies, inclusion of current and lagged water quality and climate covariates did not enhance understanding of target DNA detection probabilities. Rather, *T. bryosalmonae* DNA detection probabilities on the Yellowstone River were associated with date of collection, and *O. nerka* detection probabilities on the Snake River were associated with date, discharge or water sample volume (**Table 2**). These results do not mean that inclusion of sensor-derived physical data is not useful for other eDNA targets, it just means that the subset of physical data analyzed were not applicable for our case studies. Inclusion of additional or alternate data sets, such as stream temperature data (U.S. Forest Service NorWeST) or other water quality, biological, and physical data collected by the USEPA [STOrage and RETrieval (STORET) Data Warehouse] for example into the analytical workflow may reveal stronger relations between parameters, particularly when used to compare detection frequencies across space rather than time.

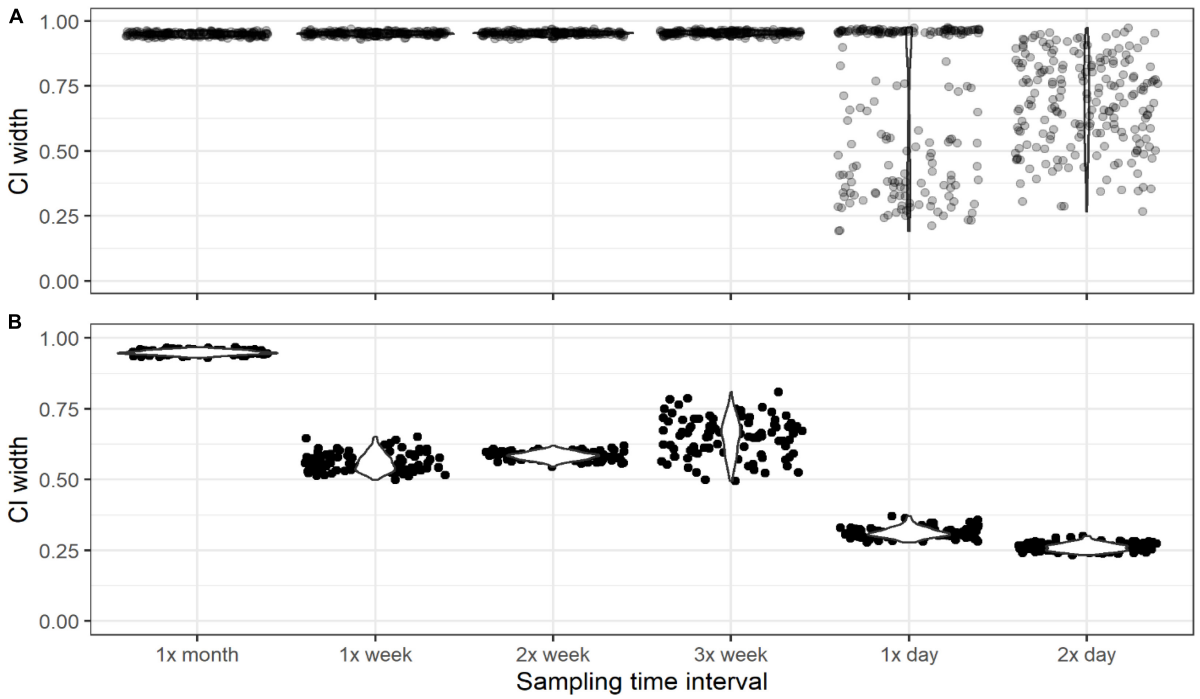


FIGURE 8 | The difference between the 95% and 5% confidence intervals (CI width) of the posterior mean of θ for simulations of Yellowstone River **(A)** and Snake River **(B)** targets when eDNA samples were collected at different daily and weekly frequencies.

Our multi-scale occupancy model results demonstrated that having more eDNA detections was necessary for detecting significant relations within our data. Target DNA of the fish parasite *T. bryosalmonae* was rarely detected in the Yellowstone River in 2018; it was likely at low abundance that year relative to previous years when PKD outbreaks were documented (Sepulveda et al., 2020; Hutchins et al., in press). When this target was detected in a sample, it was only detected in one to two of the four PCR replicates. The large uncertainty in θ probabilities made strong associations with covariates prohibitive. Increased sampling rates ($\geq$1 per day) did decrease confidence interval width but not to a magnitude that would be useful for confident decision making (**Figure 8**). This result is not a surprise nor is it unique to eDNA sampling because detection is strongly related to abundance and rare species usually have low abundance (Gaston et al., 2000). Nonetheless, this is a concerning result because eDNA sampling is championed as a superior technique for detecting rare species at low abundance, where amplification of a small subset of samples is normal (e.g., Strickland and Roberts, 2019). For example, 64 of 2822 samples tested positive for invasive Asian carp DNA (Jerde et al., 2013) and a later study found Asian carp eDNA sample detection probabilities as low as 0.04 (Mize et al., 2019). These low detection rates are the primary reason why eDNA surveys of rare species require careful consideration of field design and target species ecology (Mize et al., 2019; Strickland and Roberts, 2019). In the current study, robotic eDNA samplers were limited to the location of USGS streamgages, which were originally located for other specific purposes like recording water level. Locating robotic eDNA samplers in slow-moving waters where the *T. bryosalmonae* bryozoan primary host is more likely to occur, such as river pools and eddies (Wood, 2010), may have resulted in more *T. bryosalmonae* DNA detections. Using natural history insight of the target species to inform eDNA sampling location and timing has been recommended as a means to maximize eDNA detection probabilities of rare species (Dunker et al., 2016; Goldberg et al., 2016; Strickland and Roberts, 2019).

Comparison of multi-scale occupancy model results using the Snake River and Yellowstone River data showed that occupancy models provide the greatest benefit when sampling designs are informed by power analyses and are used in eDNA monitoring programs for established species rather than for rare species. Target DNA of *O. nerka* was more frequently detected in samples and replicates collected from the Snake River, especially in the fall when *O. nerka* was more reproductively active in the upstream reservoir (**Figure 3**). Collection date was correlated with river discharge and inversely correlated with sample volume. Consequently, we could not assume that date, as a surrogate for reproduction activity, was the primary driver of the eDNA detections, though this is a parsimonious explanation. Also, the Snake River streamgage is located directly below a hydropower dam at the mouth of the reservoir where *O. nerka* occur, so there was an increased likelihood for detection at the Snake River stream gage that was not present at the Yellowstone streamgage sites.

In addition to integrating real-time environmental data and abundant eDNA detections, our simulations illustrated that high

frequency eDNA sampling has a higher probability of detection and, under certain scenarios, can provide benefits above regular, low frequency sampling. Collecting one sample every day for 7 days each week for multiple weeks (i.e., high frequency) usually resulted in more detections than collecting seven samples one day per week for multiple weeks (i.e., low frequency), especially as the target became easier to detect and ρ varied stochastically (**Figures 4**-**7**). In particular, simulation 2 allows ρ to vary stochastically in time using a random walk behavior so that detection probabilities were similar day-to-day. This mimics the situation where there might be a short window when the detection probability is considerably higher than the remainder of the sampling period. This window is not missed by sampling daily. However, under other scenarios, where the detection probability is randomly selected for a given day, there may not be substantial benefits to the high frequency sampling when treating samples as independent. Simulations in which low-frequency samples were modeled as independent rather than dependent and ρ did not vary stochastically did not show any differences between low and high-frequency sampling. High frequency sampling designs (day or night) are more easily executed by robotic samplers, given that daily travel to field sites can be cost-prohibitive. But, when a high frequency sampling design is not feasible, sampling should be spread out over a temporal extent (hours – days) when ψ and θ probabilities are not expected to change. Sample sites should also be distributed to maximize independence, as is done in the sampling of several eDNA amphibian monitoring programs, where multiple water samples are taken from different sites around a pond or wetland (e.g., Rees et al., 2014; Bedwell and Goldberg, 2020). However, in cases where large numbers of PCR replicates and samples are required for detection and reliable parameter estimation, such as in lotic environments, it may be necessary to take multiple samples from the same site at a similar time (e.g., Erickson et al., 2019; Sepulveda et al., 2019; Woldt et al., 2019).

Combining autonomously collected molecular data with environmental data collected by sensor networks into an expedient data science pipeline is the next step in the evolution of effective biosurveillance programs. Overall, our study used data collected from the Snake River and Yellowstone River as an effective proof-of-concept for using high-throughput technologies and novel data synthesis and analysis to deliver actionable information to decision makers. We have established the initial steps in creating a flexible and customizable data science pipeline for automating the movement and transformation of data and the consolidation of data from multiple sources to be used more strategically. Multiple steps of this pipeline are still in developmental stages, such as *in situ* eDNA analyses (Ussler et al., 2013; Hansen et al., 2020; Sepulveda et al., 2020), and other steps require further refinement. The next steps will include improved and more customizable data collection by the addition of machine learning to the robot samplers to enable adjustments while *in situ,* based on ongoing power analysis simulations and the incorporation of the pipeline into a decision tree with explicit criteria for determining when stakeholders should be alerted. Continuation of this work will lead to the development of a more powerful, advanced data science pipeline with the potential to link current physical drivers to future biological responses, thereby enabling forecasting of environmental health and ultimately enhancing our understanding of ecological processes and stressors.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.sciencebase.gov/catalog/item/5e41ab1ce4b0edb47be63b4a.

AUTHOR CONTRIBUTIONS

AS, AH, JB, EB, and SC: study design and writing. AS, JB, EB, and PH: data collection. AS, AH, and CS: data analyses. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the USGS Ecosystem and Water Mission Areas, USGS National Innovation Center, and USGS Community for Data Integration.

ACKNOWLEDGMENTS

We thank Roman Marin III (MBARI), he was essential in getting the ESP devices to work in the USGS streamgage sites. He unexpectedly passed away prior to the completion of this project. We also thank Teton County Weed and Pest, the U.S. Bureau of Reclamation Palisades Dam, and hydro-technicians from the USGS Idaho Water Science Center for helping with robotic sampler logistics. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.620715/full#supplementary-material

REFERENCES

Al-Chokhachy, R., Sepulveda, A. J., Ray, A. M., Thoma, D. P., and Tercek, M. T. (2017). Evaluating species-specific changes in hydrologic regimes: an iterative approach for salmonids in the Greater Yellowstone Area (USA). *Rev. Fish Biol. Fish.* 27, 425–441. doi: 10.1007/s11160-017-9472-3

Bedwell, M. E., and Goldberg, C. S. (2020). Spatial and temporal patterns of environmental DNA detection to inform sampling protocols in

lentic and lotic systems. *Ecol. Evol.* 10, 1602–1612. doi: 10.1002/ece3.6014

Bohan, D. A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A. J., and Woodward, G. (2017). Next-generation global biomonitoring: large-scale, automated reconstruction of ecological networks. *Trends Ecol. Evol.* 32, 477–487. doi: 10.1016/j.tree.2017.03.001

Collins, S. L., Bettencourt, L. M., Hagberg, A., Brown, R. F., Moore, D. I., Bonito, G., et al. (2006). New opportunities in ecological sensing using wireless sensor networks. *Front. Ecol. Environ.* 4:402–407. doi: 10.1890/1540-929520064[402:NOIESU]2.0.CO;2

Cordier, T., Alonso-Sáez, L., Apothéloz-Perret-Gentil, L., Aylagas, E., Bohan, D. A., Bouchez, A., et al. (2020). Ecosystems monitoring powered by environmental genomics: a review of current strategies with an implementation roadmap. *Mol. Ecol.* doi: 10.1111/mec.15472 [Epub ahead of print],

Cristescu, M. E., and Hebert, P. D. (2018). Uses and misuses of environmental DNA in biodiversity science and conservation. *Annu. Rev. Ecol. Evol. Syst.* 49, 209–230. doi: 10.1146/annurev-ecolsys-110617-062306

Darling, J. A. (2019). How to learn to stop worrying and love environmental DNA monitoring. *Aquat. Ecosyst. Health Manag.* 22, 440–451. doi: 10.1080/14634988.2019.1682912

Dunker, K. J., Sepulveda, A. J., Massengill, R. L., Olsen, J. B., Russ, O. L., Wenburg, J. K., et al. (2016). Potential of environmental DNA to evaluate northern pike (*Esox lucius*) eradication efforts: an experimental test and case study. *PLoS One* 11:e0162277. doi: 10.1371/journal.pone.0162277

Erickson, R. A., Merkes, C. M., and Mize, E. L. (2019). Sampling designs for landscape-level eDNA monitoring programs. *Integr. Environ. Assess. Manag.* 15, 760–771. doi: 10.1002/ieam.4155

Gallien, L., Münkemüller, T., Albert, C. H., Boulangeat, I., and Thuiller, W. (2010). Predicting potential distributions of invasive species: where to go from here? *Divers. Distribut.* 16, 331–342. doi: 10.1111/j.1472-4642.2010.00652.x

Gaston, K. J., Blackburn, T. M., Greenwood, J. J., Gregory, R. D., Quinn, R. M., and Lawton, J. H. (2000). Abundance–occupancy relationships. *J. Appl. Ecol.* 37, 39–59. doi: 10.1046/j.1365-2664.2000.00485.x

Gelman, A., Hwang, J., and Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Stat. Comput.* 24, 997–1016. doi: 10.1007/s11222-013-9416-2

Glibert, P. M., Pitcher, G. C., Bernard, S., and Li, M. (2018). "Advancements and continuing challenges of emerging technologies and tools for detecting harmful algal blooms, their antecedent conditions and toxins, and applications in predictive models," in *Global Ecology and Oceanography of Harmful Algal Blooms*, Vol. 232, eds P. Glibert, E. Berdalet, M. Burford, G. Pitcher, and M. Zhou (Cham: Springer), 339–357. doi: 10.1007/978-3-319-70069-4_18

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., et al. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods Ecol. Evol.* 7, 1299–1307. doi: 10.1111/2041-210X.12595

Hansen, B. K., Jacobsen, M. W., Middelboe, A. L., Preston, C. M., Marin, R., Bekkevold, D., et al. (2020). Remote, autonomous real-time monitoring of environmental DNA from commercial fish. *Sci. Rep.* 10, 1–8. doi: 10.1038/s41598-020-70206-8

Hutchins, P. R., Sepulveda, A. J., Hartikainen, H. H., Staigmiller, K. D., Opitz, S., Yamamoto, R., et al. (in press). Exploration of the 2016 Yellowstone river fish kill and proliferative kidney disease in wild fish populations. *Ecosphere*.

Jerde, C. L., Chadderton, W. L., Mahon, A. R., Renshaw, M. A., Corush, J., Budny, M. L., et al. (2013). Detection of Asian carp DNA as part of a Great Lakes basin-wide surveillance program. *Can. J. Fish. Aquat. Sci.* 70, 522–526. doi: 10.1139/cjfas-2012-0478

Kovach, R. P., Dunham, J. B., Al-Chokhachy, R., Snyder, C. D., Letcher, B. H., Young, J. A., et al. (2019). An integrated framework for ecological drought across riverscapes of North America. *BioScience* 69, 418–431. doi: 10.1093/biosci/biz040

Michener, W. K., and Jones, M. B. (2012). Ecoinformatics: supporting ecology as a data-intensive science. *Trends Ecol. Evol.* 27, 85–93. doi: 10.1016/j.tree.2011.11.016

Mize, E. L., Erickson, R. A., Merkes, C. M., Berndt, N., Bockrath, K., Credico, J., et al. (2019). Refinement of eDNA as an early monitoring tool at the landscape-level: study design considerations. *Ecol. Appl.* 29:e01951. doi: 10.1002/eap.1951

Pilliod, D. S., Laramie, M. B., MacCoy, D., and Maclean, S. (2019). Integration of eDNA-based biological monitoring within the US Geological Survey's National Streamgage Network. *J. Am. Water Resour. Assoc.* 55, 1505–1518. doi: 10.1111/1752-1688.12800

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R., and Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. *J. Appl. Ecol.* 51, 1450–1459. doi: 10.1111/1365-2664.12306

Scholin, C. A., Birch, J., Jensen, S., Marin, R. III, Massion, E., Pargett, D., et al. (2017). The quest to develop ecogenomic sensors: a 25-year history of the Environmental Sample Processor (ESP) as a case study. *Oceanography* 30, 100–113. doi: 10.5670/oceanog.2017.427

Sepulveda, A. J., Birch, J. M., Barnhart, E. P., Merkes, C. M., Yamahara, K. M., Marin, R., et al. (2020). Robotic environmental DNA bio-surveillance of freshwater health. *Sci. Rep.* 10, 1–8. doi: 10.1038/s41598-020-71304-3

Sepulveda, A. J., Schmidt, C., Amberg, J., Hutchins, P., Stratton, C., Mebane, C., et al. (2019). Adding invasive species biosurveillance to the US Geological Survey streamgage network. *Ecosphere* 10:e02843. doi: 10.1002/ecs2.2843

Sepulveda, A. J., Tercek, M. T., Al-Chokhachy, R., Ray, A. M., Thoma, D. P., Hossack, B. R., et al. (2015). The shifting climate portfolio of the Greater Yellowstone Area. *PLoS One* 10:e0145060. doi: 10.1371/journal.pone.0145060

Stohlgren, T. J., and Schnase, J. L. (2006). Risk analysis for biological hazards: what we need to know about invasive species. *Risk Anal.* 26, 163–173. doi: 10.1111/j.1539-6924.2006.00707.x

Stratton, C., Sepulveda, A., and Hoegh, A. (2020). msocc: Fit and analyze computationally efficient multi-scale occupancy models in R. *Methods Ecol. Evol.* 11, 1113–1120. doi: 10.1111/2041-210X.13442

Strickland, G. J., and Roberts, J. H. (2019). Utility of eDNA and occupancy models for monitoring an endangered fish across diverse riverine habitats. *Hydrobiologia* 826, 129–144. doi: 10.1007/s10750-018-3723-8

Sugai, L. S. (2020). Pandemics and the need for automated systems for biodiversity monitoring. *J. Wildl. Manag.* doi: 10.1002/jwmg.21946 [Epub ahead of print],

Uden, D. R., Allen, C. R., Angeler, D. G., Corral, L., and Fricke, K. A. (2015). Adaptive invasive species distribution models: a framework for modeling incipient invasions. *Biol. Invasions* 17, 2831–2850. doi: 10.1007/s10530-015-0914-3

Ussler, W. III, Preston, C., Tavormina, P., Pargett, D., Jensen, S., Roman, B., et al. (2013). Autonomous application of quantitative PCR in the deep sea: in situ surveys of aerobic methanotrophs using the deep-sea environmental sample processor. *Environ. Sci. Technol.* 47, 9339–9346. doi: 10.1021/es4023199

Woldt, A., Baerwaldt, K., Monroe, E., Tuttle-Lau, M., Grueneis, N., Holey, M., et al. (2019). *Quality Assurance Project Plan: eDNA Monitoring of Bighead and Silver Carps*. Bloomington, MN: U.S. Fish and Wildlife Service.

Wood, T. S. (2010). "Bryozoans," in *Ecology and Classification of North American Freshwater Invertebrates*, eds J. H. Thorp and A. P. Covich (Amsterdam: Elsevier), 437–454. doi: 10.1016/B978-0-12-374855-3.00013-3

Yamahara, K. M., Preston, C. M., Birch, J. M., Walz, K. R., Marin, R. III, Jensen, S., et al. (2019). In-situ autonomous acquisition and preservation of marine environmental DNA using an autonomous underwater vehicle. *Front. Mar. Sci.* 6:373. doi: 10.3389/fmars.2019.00373

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 19 March 2021
doi: 10.3389/fevo.2021.623831

A Model and Simulation of the Influence of Temperature and Amplicon Length on Environmental DNA Degradation Rates: A Meta-Analysis Approach

Tatsuya Saito and Hideyuki Doi*

Graduate School of Simulation Studies, University of Hyogo, Kobe, Japan

Environmental DNA (eDNA) analysis can detect aquatic organisms, including rare and endangered species, in a variety of habitats. Degradation can influence eDNA persistence, impacting eDNA-based species distribution and occurrence results. Previous studies have investigated degradation rates and associated contributing factors. It is important to integrate data from across these studies to better understand and synthesize eDNA degradation in various environments. We complied the eDNA degradation rates and related factors, especially water temperature and amplicon lengths of the measured DNA from 28 studies, and subjected the data to a meta-analysis. In agreement with previous studies, our results suggest that water temperature and amplicon length are significantly related to the eDNA degradation rate. From the 95% quantile model simulation, we predicted the maximum eDNA degradation rate in various combinations of water temperature and amplicon length. Predicting eDNA degradation could be important for evaluating species distribution and inducing innovation (e.g., sampling, extraction, and analysis) of eDNA methods, especially for rare and endangered species with small population size.

Keywords: environmental DNA, polymerase chain reaction, degradation rate, quantile model, meta-analysis

OPEN ACCESS

Edited by:
Katy Klymus,
United States Geological Survey (USGS), United States

Reviewed by:
Catherine Richter,
Columbia Environmental Research Center, United States Geological Survey, United States
Sheena Feist,
Engineer Research and Development Center (ERDC), United States

****Correspondence:***
Tatsuya Saito
tatwoeight630@icloud.com

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *30 October 2020*
Accepted: *19 February 2021*
Published: *19 March 2021*

Citation:
Saito T and Doi H (2021) A Model and Simulation of the Influence of Temperature and Amplicon Length on Environmental DNA Degradation Rates: A Meta-Analysis Approach. Front. Ecol. Evol. 9:623831. doi: 10.3389/fevo.2021.623831

INTRODUCTION

Environmental DNA (eDNA) methods are innovative methods developed for monitoring macroorganisms, especially aquatic species (Ficetola et al., 2008; Minamoto et al., 2012; Taberlet et al., 2012; Takahara et al., 2012; Ushio et al., 2018; Kakuda et al., 2019; Tsuji et al., 2019). The eDNA method is used to investigate species distribution. It is less invasive to organisms, and is especially useful for rare and endangered species, which generally have low tolerance to sampling disturbance and may be difficult to detect. Consequently, eDNA methods have been used to detect rare and endangered species in various taxa, such as fish, salamander, and aquatic insects (Fukumoto et al., 2015; Sigsgaard et al., 2015; Pfleger et al., 2016; Doi et al., 2017; Sakata et al., 2017).

Environmental DNA, which is compromised of DNA fragments released by organisms into environments such as water or soil, is thought to be derived from mixtures of feces (Martellini et al., 2005), skin cells (Ficetola et al., 2008), mucus (Merkes et al., 2014), and secretions (Bylemans et al., 2018) of organisms. Previous studies have suggested that eDNA is mainly derived from fractions of cells or cellular organs (i.e., mitochondria and nuclei), but it can also be derived from fragmented DNA (degraded DNA) in the water (Turner et al., 2014; Minamoto et al., 2016).

Many points regarding the general behavior of eDNA in water (reviewed in Barnes and Turner, 2016) are still unclear, especially the state (fragment length) and degradation of eDNA (Turner et al., 2015; reviewed in Barnes and Turner, 2016). Understanding eDNA states and degradation is essential for the effective sampling and storage of eDNA, and may provide pertinent information to better interpret the results of species distribution and abundance and biomass estimations. This may be especially problematic for rare and endangered species, which are thought to have small populations and small amounts (or concentrations) of DNA (Fukumoto et al., 2015; Sigsgaard et al., 2015; Pfleger et al., 2016; Doi et al., 2017; Sakata et al., 2017). Both factors can influence eDNA persistence, potentially inducing false negatives which impact accuracy in occurrence and distribution data.

Many experiments have been conducted to reveal the detailed states and degradation rates of eDNA under various conditions (Thomsen et al., 2012; Barnes et al., 2014; Maruyama et al., 2014; Tsuji et al., 2017; Jo et al., 2019). In most cases, the eDNA degradation curves declined exponentially and quickly, often in less than a week (Thomsen et al., 2012; Barnes et al., 2014). Earlier meta-analyses for eDNA degradation (Collins et al., 2018) found that water conditions, such as salinity (Collins et al., 2018), water temperature (Tsuji et al., 2017; Jo et al., 2019), and pH (Barnes et al., 2014; Tsuji et al., 2017), influenced the eDNA degradation rate. In addition, the characteristics of DNA itself, such as its measured amplicon length, affected the eDNA degradation rate (Bylemans et al., 2018; Jo et al., 2019). From the data so far (temperature and amplification length), it seems possible to predict the approximate degradation rate and estimate the state of eDNA. Therefore, we conducted a novel meta-analysis to model the effects of water conditions and DNA amplicon length on the eDNA degradation rate using data generated in previous eDNA degradation studies. The previous meta-analysis (Collins et al., 2018) used the half-life of the degradation curve as an index of degradation. Although half-life has the advantage of being more intuitively meaningful, we instead used here the degradation rate constants "k" because our model uses the degradation rate, not half-life.

Using this approach, we aimed to evaluate the effects of water conditions (i.e., ecosystem, source, temperature, and pH), and target DNA region on eDNA degradation in previously published data. Also, we tested the relationship between DNA amplicon length and eDNA degradation because degradation may differ with amplicon length. Specifically, we conducted a simulation to predict the maximum degradation rate using quantile regression modeling with temperature and DNA amplicon length.

MATERIALS AND METHODS

Search Strategy

A Google Scholar search on September 9, 2020, using the search terms "eDNA" OR "environmental DNA" AND "degradation" OR "decay" OR "decomposition," returned 11,300 hits. The initial filtering of the articles was based on their title; any articles that obviously had no relevance to eDNA degradation were discarded. After title screening, 1,000 articles remained. After abstract screening, 42 articles remained. We manually inspected these remaining articles and selected papers describing the degradation rate of eDNA using experiments or field settings (**Supplementary Table 1**). Upon completion of the screening process, we obtained relevant eDNA data from 28 articles (**Table 1** and **Supplementary Table 1**) for the meta-analysis.

Data Extraction

From the selected publications, we assembled a list of factors for eDNA degradation (**Supplementary Table 1**). We collected the following factors and categories: "Ecosystem" was divided into marine and freshwater. "Source" was categorized into water sources (Freshwater: river, lake, well water, pond, tap water, and deionized water; Marine: marine and artificial seawater). "Temperature" and "pH" refer to the water temperature and pH of the water sample for each experiment, respectively. "Region" and "Amplicon length" refer to the amplified DNA region used for quantitative PCR (qPCR) and the number of amplified-DNA bases targeted by the qPCR reaction (bp). "Region" was divided into mtDNA (COI, CytB, 16s, 18s, D-loop, NADH, ND2, ND4), nuDNA (ITS), and RNA. "DNA type" was divided into spike (i.e., the DNA contained in the environment water) and organism. "Experiment type" was divided into "in tank" and "in field."

We extracted the simple exponential slope (hereafter referred to as "degradation rate") from the article contents and/or plots according to the simple exponential equation (Motulsky and Christopoulos, 2003) in each experiment:

$$C = C_o E^{kt}$$

where C_0 is the eDNA concentration at time 0 (i.e., the initial eDNA concentration), and k is the degradation slope (rate) constant per hour. We used the standardized degradation rate per hour. The degradation rate by day was divided by 24 to calculate the degradation rate per hour.

Statistical Analysis and Simulation

We performed the statistical analysis and graphics using R ver. 4.0.2 (R Core Team, 2020). We tested the differences in the eDNA degradation rate in measured DNA regions and water resources using a linear mixed-effect model (LMM) using "lme4" ver. 1.1.23 package with "lmerTest" ver. 3.1.2 package in R. We excluded data points without temperature information in the statistical analyses. We set each study as a random effect. Jo et al. (2020) compared the degradation of mtDNA and nuDNA and found the difference. However, we could not analyze mtDNA and nuDNA due to the limited data.

We performed quantile models (QM) for 0.1, 0.5, and 0.95 quantiles for the regression. By performing 0.95 and 0.1 quantiles for the regression, we evaluated the maximum and minimum degradation rate. The 0.5-quantile used median for the regression, so almost similar to simple linear regression. We employed the Bayesian mixed-effect quantile model using the "lqmm" function of "lqmm" package ver. 1.5.5 in R. In the QM, we set water temperature and amplicon length as explanatory effects and each study as the random effect. We performed the Nelder–Mead algorithm using 10000 MCMC permutations with

TABLE 1 | The organisms, ecosystem types (Ecosystem), water source (Source), and PCR-amplified DNA regions by quantitative PCR (Region) for all papers analyzed in this meta-analysis.

Organism	Ecosystem	Source	Region	References	Year	Study
Gasterosteus aculeatus	Marine	Marine	CytB	Thomsen et al.	2012	1
Platichthys flesus	Marine	Marine	CytB	Thomsen et al.	2012	1
Lepomis macrochirus	Freshwater	Tap	CytB	Maruyama et al.	2014	2
Cyprinus carpio	Freshwater	Well	CytB	Barnes et al.	2014	3
Lithobates catesbeianus	Freshwater	Tap	CytB	Strickler et al.	2015	4
Cyprinus carpio	Freshwater	Well	CytB	Eichmiller et al.	2016	5
Cyprinus carpio	Freshwater	Lake	CytB	Eichmiller et al.	2016	5
Engraulis mordax	Marine	Marine	D-loop	Sassoubre et al.	2016	6
Sardinops sagax	Marine	Marine	D-loop	Sassoubre et al.	2016	6
Scomber japonicus	Marine	Marine	COI	Sassoubre et al.	2016	6
Scomber japonicus	Marine	Marine	COI	Andruszkiewicz et al.	2017	7
Zearaja maugeana	Marine	Marine	ND4	Weltz et al.	2017	8
Chrysaora pacifica	Marine	Marine	COI	Minamoto et al.	2017	9
Trachurus japonicus	Marine	Marine	CytB	Jo et al.	2017	10
Plecoglossus altivelis	Freshwater	River	CytB	Tsuji et al.	2017	11
Cyprinus carpio	Freshwater	River	CytB	Tsuji et al.	2017	11
Margaritifera margaritifera	Freshwater	River	NADH	Sansom and Sassoubre	2017	12
Carcinus maenas	Marine	Marine	COI	Collins et al.	2018	13
Lipophrys pholis	Marine	Marine	COI	Collins et al.	2018	13
Hypophthalmichthys nobilis	Freshwater	Deionized	D-loop	Lance et al.	2017	14
Chionodraco rastrospinosus	Marine	Marine	ND2	Cowart et al.	2018	15
Carassius auratus	Freshwater	Tap	ITS	Bylemans et al.	2018	16
Neogobius melanostomus	Freshwater	Lake	COI	Nevers et al.	2018	17
Cyprinus carpio	Freshwater	River	CytB	Nukazawa et al.	2018	18
Grandidierella japonica	Marine	Artificial seawater	COI	Wei et al.	2018	19
Trachurus japonicus	Marine	Marine	CytB	Jo et al.	2019	20
Daphnia magna	Freshwater	Tap	COI	Moushomi et al.	2019	21
Daphnia magna	Freshwater	Tap	18S	Moushomi et al.	2019	21
cyanobacterial	Freshwater	Lake	16S	Zulkefli et al.	2019	22
Schistosoma mansoni	Freshwater	Tap	COI	Sengupta et al.	2019	23
Trachurus japonicus	Marine	Marine	CytB	Jo et al.	2020	24
Trachurus japonicus	Marine	Marine	ITS	Jo et al.	2020	24
Styela clava	Marine	Marine	COI	Wood et al.	2020	25
Spirographis spallanzani	Marine	Marine	COI	Wood et al.	2020	25
Styela clava	Marine	Marine	RNA	Wood et al.	2020	25
Spirographis spallanzani	Marine	Marine	RNA	Wood et al.	2020	25
Anguilla japonica	Freshwater	Tap	D-loop	Kasai et al.	2020	26
Rhinella marina	Freshwater	Tap	16S	Villacorta-Rath et al.	2020	27
Trachurus japonicus	Marine	Marine	CytB	Saito and Doi	2020	28
Cyprinus carpio	Freshwater	Pond	CytB	Saito and Doi	2020	28

the Gauss–Hermite quadrature approach. We set the statistical alpha as 0.05 for parameter evaluation. We did not find a significant interaction ($p > 0.1$) between water temperature and amplicon length, so we used the model excluding the interaction, i.e., eDNA degradation rate = water temperature + amplicon length. We evaluated the QM models using the Akaike information criteria (AIC), in which the best QM is identified by having the lowest AIC.

We simulated the combined effects of water temperature and amplicon length, using the obtained 0.95-quantile QM. We generated 100,000 random values for the combination of water temperature (ranging in published values from −1 to 35 °C; see the results) and amplicon length used for the experiments (ranging in published values from 70 to 719) using "runif" function in R, which generates a random number from the Mersenne-Twister method. We used 100,000 random values to predict the eDNA degradation rate from the 0.95-quantile QM (see results).

RESULTS

Degradation Rate Experiments From Literature

The number of obtained time points for the eDNA degradation data ranged from 3 to 25 (mean: 8.3, median: 8.0, **Supplementary Table 1**). Details of the sites are listed as water sources (**Table 1**). In total there were 21 marine sites, 1 artificial marine site, and 19 freshwater sites. Within the freshwater sites, there were 9 experiments that used tap or deionized water, 4 river sites, 3 lake

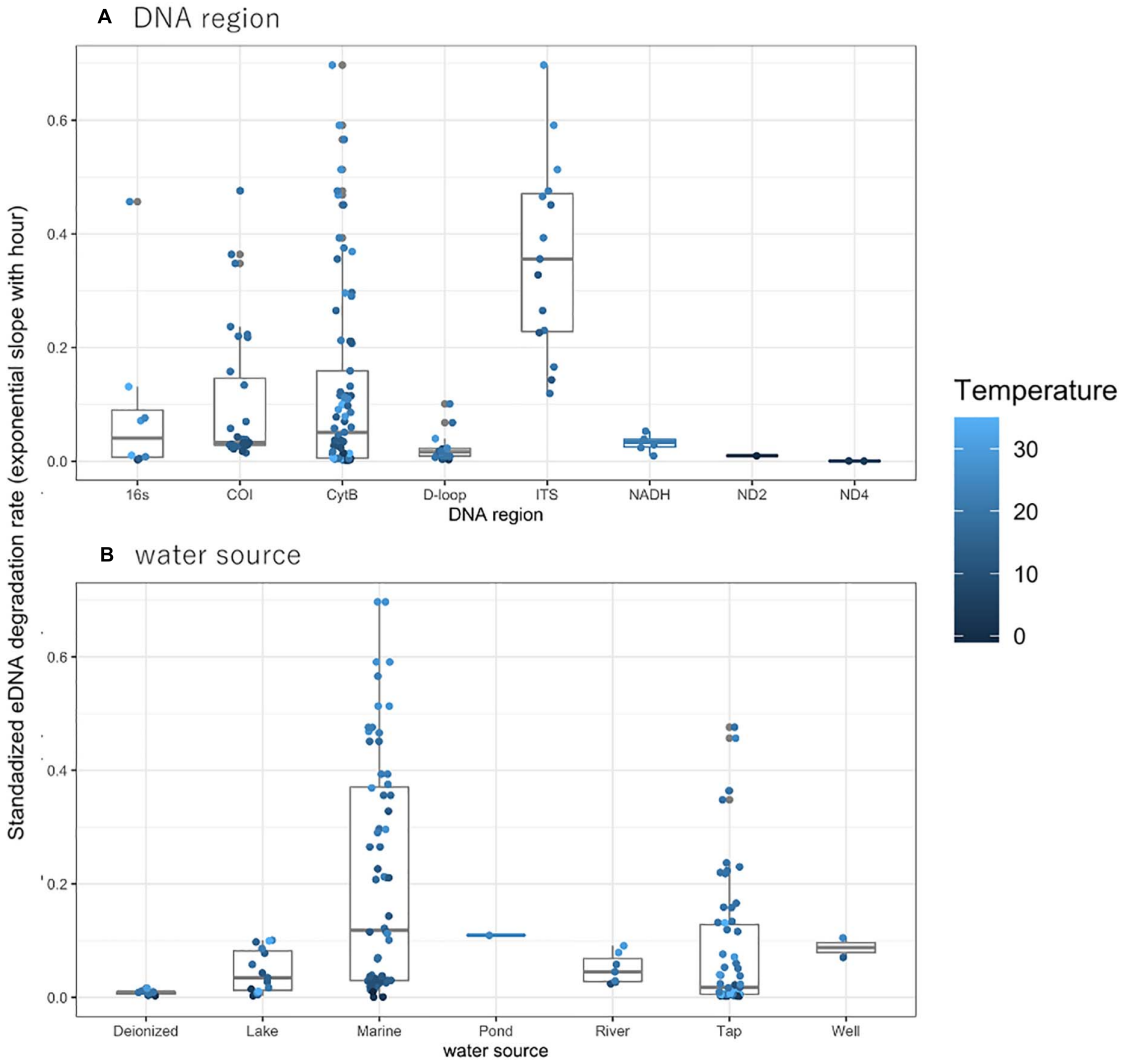


FIGURE 1 | The eDNA degradation rate (simple exponential slope) with **(A)** DNA region and **(B)** water source. The degradation rate without temperature data in the experiment were excluded in the plot. The dots indicate the individual eDNA degradation rate in each experiment in different ecosystems. The boxes and bars in the box plot indicate median ± inter-quartiles and ± 1.5 × inter-quartiles, respectively.

sites, 2 well water sites, and 1 pond site. The temperature for the experiments ranged from −1 to 35 °C (mean: 19, median: 20, **Supplementary Table 1**). The amplicon length used for the experiments ranged from 70 to 719 bp (mean: 150, median: 131, **Supplementary Table 1**), and the DNA fragment regions used were mainly Cyt B or COI regions in mtDNA (**Table 1**). Degradation experiments of nuDNA and RNA were very few data compared to mtDNA.

Degradation Rate

The observed degradation rate for the previously published eDNA data ranged from 0.0005 to 0.7010 (mean: 0.1317, median: 0.0440, **Supplementary Table 1**). Differences in PCR regions did not affect the rate of DNA degradation, nor did differences in water sources (**Figures 1A,B**). Although the degradation rates for Tap and Marine water sources appear much higher than that observed for other sources, there were no significant differences among water sources, nor among taxa or PCR regions (LMM, $t < 1.859$, $p > 0.07$, **Figure 1** and **Supplementary Figure 1**, respectively). With the limited data excluded, such as ND2, ND4 for PCR region and pond for water source, there were no significant differences among water sources (LMM, $t < 1.965$, $p > 0.06$, **Supplementary Figure 2**, respectively), but significant differences among PCR region (LMM, $t = -3.414$, $p = 0.002538$, **Supplementary Figure 2**).

Quantile Model for Temperature and Amplicon Length

The relationship between eDNA degradation rate and water temperature was significant in 0.95- quantile and showed that higher water temperatures accelerated eDNA degradation

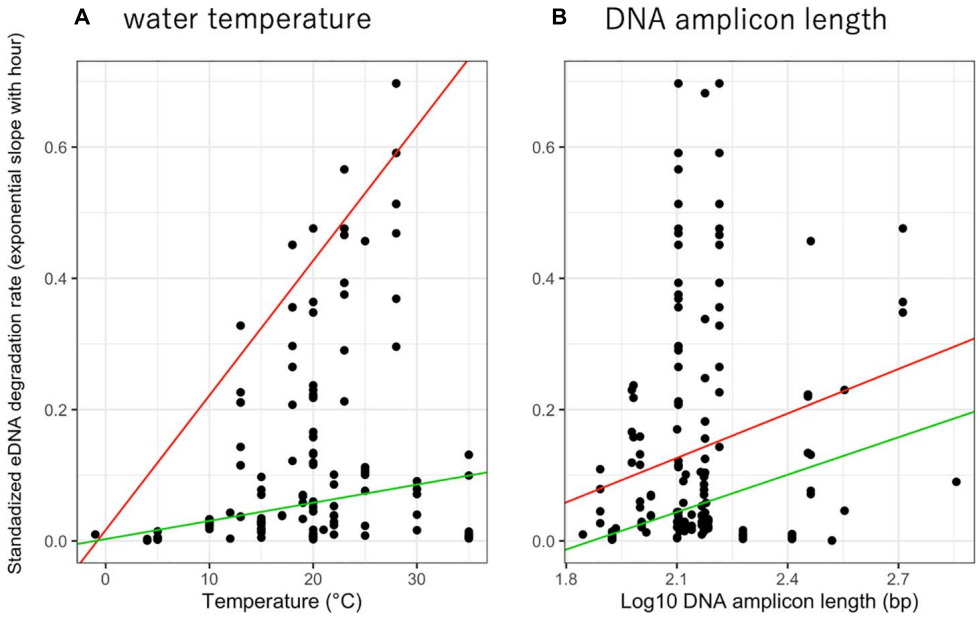


FIGURE 2 | The relationship between standardized eDNA degradation rate per hour (simple exponential slope) with **(A)** water temperature and **(B)** DNA amplicon length. The red and green lines show 0.95 and 0.5- quantile mixed-effect quantile models for each factor.

(**Figure 2A**, $p = 0.02004$ and 0.5761 for 0.95- and 0.5- quantiles, respectively). Upon comparing the QM of 0. 1-, 0. 5-, and 0.95- quantiles, the QM with 0.95-quantile was observed to have the lowest AIC value (0.1-quantile: 41.82, 0.5-quantile: −120.78, and 0.95-quantile: −161.26), indicating that the best model for the relationship. Therefore, we simulated these data using the QM with a 0.95-quantile with a positive slope (slope = 0.020, **Figure 2A**). The relationship between eDNA degradation rate and amplicon length suggests that longer amplicon length undergo greater eDNA degradation (**Figure 2B**). For amplicon length, as for water temperature, the QM with 0.95-quantile had the lowest AIC value (0.1-quantile: 155.1, 0.5-quantile: −110.2, and 0.95-quantile: −145.6). Therefore, we simulated and discussed these data using the QM with a 0.95-quantile with a positive slope (slope = 0.225). We also showed the categories of water temperature range (divided into four levels: −1, 0–10, 11–20, and > 21 °C) and amplicon length (divided into three levels: 0–100, 101–200, and > 201 bp) with eDNA degradation rate (**Supplementary Figures 3, 4** respectively) with similar trends of **Figure 2**.

eDNA Degradation Simulation

Our QM simulation lead to plotting the eDNA degradation on a matrix of water temperature and amplicon length (**Figure 3**), which showed that the water temperature had a great influence on the eDNA degradation rate. At lower (e.g., −1 to 5 °C) and higher (e.g., 15 to 35 °C) water temperatures, our model predicted that amplicon length would have a smaller effect on the eDNA degradation rate, while at moderate (e.g., 5 to 15 °C) water temperatures, our prediction more clearly showed that the longer amplicon length would have a faster degradation rate. Thus, at moderate water temperatures, the amplicon length should also be considered in evaluating eDNA degradation.

DISCUSSION

Our meta-analysis results showed that eDNA degradation was accelerated in higher water temperatures and in longer amplicon length. These generally supported the effect of water temperature on the eDNA degradation rate in previous hypotheses for each condition and species (e. g., Strickler et al., 2015; Eichmiller et al., 2016; Lance et al., 2017; Tsuji et al., 2017; Jo et al., 2019; Kasai et al., 2020). Previous studies have assumed that water temperature does not directly affect eDNA degradation, but indirectly affects it through enzymatic hydrolysis by microbes and extracellular nucleases (reviewed in Barnes and Turner, 2016). At high temperatures, with increasing activity of microorganisms and extracellular enzymes, the eDNA in water would degrade more quickly (reviewed in Barnes and Turner, 2016). Our meta-analysis results showed that there were no significant differences between laboratory water (purified or tap water) and environmental water (seawater or freshwater). This may indicate the enzymes and bacteria possessed by experimental organisms affected the eDNA degradation. In fact, the degradation experiment, which intracellular DNA and fragmented DNA were added to purified water, showed that intracellular and fragments DNA were not degraded in the water for a week (Saito and Doi, 2020).

Evidence from previous studies suggested that, in eDNA samples, long amplicon length are less likely to be detected than short amplicon (Jo et al., 2017). Our meta-analysis supports these previous results. A possible explanation is provided by Jo et al. (2017), in which it was suggested that the DNA degradation rate

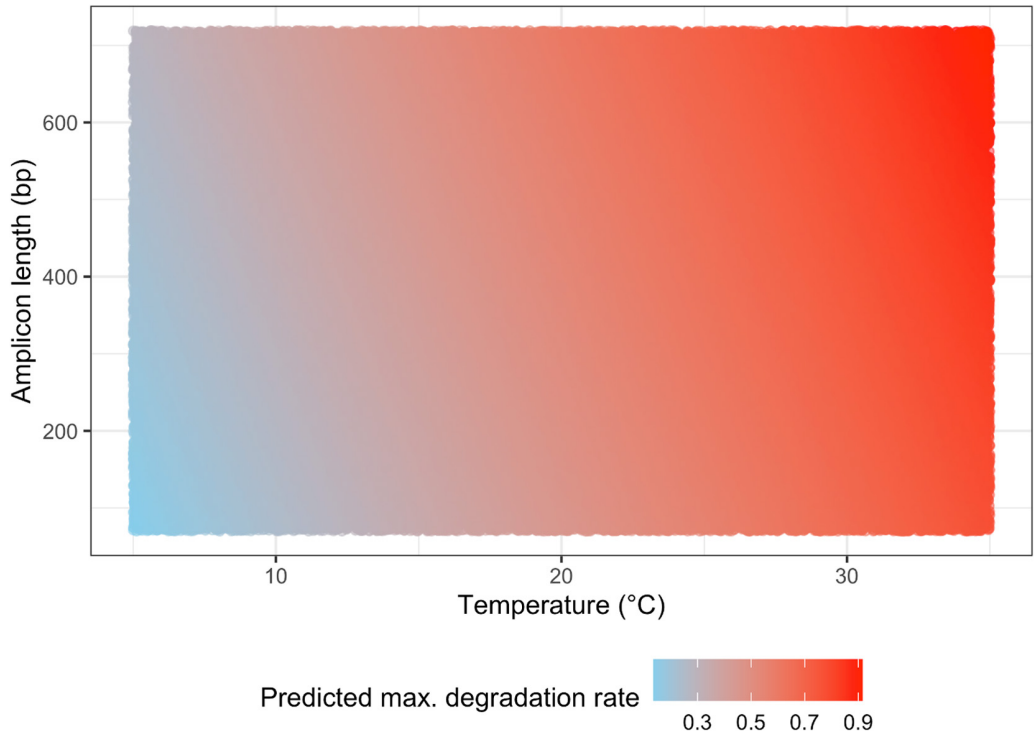


FIGURE 3 | The simulation result for predicting eDNA degradation rate on the matrix of water temperature and amplicon length.

was higher in longer amplicon length (719 bp) than in shorter amplicon (127 bp). Our simulation by QM indicated that shorter amplicon lengths were more likely to be detected when eDNA degradation was less affected by water temperature. When the eDNA degradation rates were very fast or very slow due to water temperature (e.g., 15 to 35 °C or 0 to 5°C, respectively), the amplicon length had a smaller effect on eDNA degradation than at other water temperature ranges. In higher temperatures, microbial activity that breaks down DNA is occurring fast on both large and short DNA fragments, such that both classes of fragments are not detectable by either a large or small fragment amplicon assay at a similar rate. Whereas in colder temperatures, both fragment classes are degraded at lower rates, and thus it is possible that the longer fragments are able to last longer than under warmer conditions, thus remaining detectable for longer (suggesting a slower decay rate).

In our meta-analysis, we evaluated amplicon lengths ranging from 70 to 719 bp, but there were no experiments in which longer amplicon were measured. Recently, however, long range PCR was used to amplify full mitogenomes from eDNA samples (Deiner et al., 2017a,b). Additional investigation is needed to better understand retention of such extremely long DNA (> 16,000 bps), and the role of degradation in these cases.

CONCLUSION

In conclusion, our meta-analysis results showed that eDNA degradation was accelerated in higher water temperatures and in longer DNA amplicon. We predicted the combined effects of water temperature and amplicon length on the maximum eDNA degradation rate. Our meta-analysis and simulation provided new insights for future eDNA studies. We should note the limitations: The number of papers used for our meta-analysis was limited to 28 studies, and the data was limited especially for other environmental factors, such as UV, pH, and salinity, which are important factors for eDNA degradation (Barnes et al., 2014; Lance et al., 2017; Tsuji et al., 2017; Collins et al., 2018; Mächler et al., 2018). When data such as UV, pH, and salinity are obtained in addition to water temperature, more complex phenomena can be evaluated to determine the eDNA degradation rate in water. A greater understanding and accumulation of eDNA degradation data would improve future eDNA methods.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

TS and HD designed the study, analyzed the data, interpreted the results, and wrote the manuscript. TS collected the data. Both authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Environment Research and Technology Development Fund (JPMEERF20204004).

ACKNOWLEDGMENTS

We thank the authors of all the articles used for the meta-analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.623831/full#supplementary-material

Supplementary Figure 1 | The eDNA degradation rate (simple exponential slope) with the targeted taxon group. The dots indicate the individual eDNA degradation rate in each experiment. The boxes and bars in the box plot indicate median ± inter-quartiles and ± 1.5 × inter-quartiles, respectively.

Supplementary Figure 2 | The eDNA degradation rate (simple exponential slope) with (A) DNA region and (B) water source. In the plot, the limited data were excluded; NADH, D-loop, ND2, and ND4 for PCR region and pond, well for water source. The dots indicate the individual eDNA degradation rate in each experiment in different ecosystems. The boxes and bars in the box plot indicate median ± inter-quartiles and ± 1.5 × inter-quartiles, respectively.

Supplementary Figure 3 | The eDNA degradation rate (simple exponential slope) with temperature category. The temperature categories are divided into four levels: −1, 0–10, 11–20, and > 21 °C. The dots indicate the individual eDNA degradation rate in each experiment in different ecosystems. The boxes and bars in the box plot indicate median ± inter-quartiles and ± 1.5 × inter-quartiles, respectively.

Supplementary Figure 4 | The eDNA degradation rate (simple exponential slope) with amplicon length category. The amplicon length categories are divided into three levels: 0–100, 101–200, and > 201. The dots indicate the individual eDNA degradation rate in each experiment in different ecosystems. The boxes and bars in the box plot indicate median ± inter-quartiles and ± 1.5 × inter-quartiles, respectively.

Supplementary Table 1 | All data analyzed in this meta-analysis.

REFERENCES

Andruszkiewicz, E. A., Sassoubre, L. M., and Boehm, A. B. (2017). Persistence of marine fish environmental DNA and the influence of sunlight. *PLoS One* 12:e0185043. doi: 10.1371/journal.pone.0185043

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

Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., and Lodge, D. M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environ. Sci. Technol.* 48, 1819–1827. doi: 10.1021/es404734p

Bylemans, J., Furlan, E. M., Gleeson, D. M., Hardy, C. M., and Duncan, R. P. (2018). Does size matter? An experimental evaluation of the relative abundance and decay rates of aquatic environmental DNA. *Environ. Sci. Technol.* 52, 6408–6416. doi: 10.1021/acs.est.8b01071

Collins, R. A., Wangensteen, O. S., O'Gorman, E. J., Mariani, S., Sims, D. W., and Genner, M. J. (2018). Persistence of environmental DNA in marine systems. *Commun. Biol.* 1, 1–11. doi: 10.1038/s42003-018-0192-6

Cowart, D. A., Murphy, K. R., and Cheng, C. H. C. (2018). Metagenomic sequencing of environmental DNA reveals marine faunal assemblages from the West Antarctic Peninsula. *Mar. Genom.* 37, 148–160. doi: 10.1016/j.margen.2017.11.003

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., et al. (2017a). Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Mol. Ecol.* 26, 5872–5895. doi: 10.1111/mec.14350

Deiner, K., Renshaw, M. A., Li, Y., Olds, B. P., Lodge, D. M., and Pfrender, M. E. (2017b). Long-range PCR allows sequencing of mitochondrial genomes from environmental DNA. *Methods Eco. Evol.* 8, 1888–1898. doi: 10.1111/2041-210X.12836

Doi, H., Katano, I., Sakata, Y., Souma, R., Kosuge, T., Nagano, M., et al. (2017). Detection of an endangered aquatic heteropteran using environmental DNA in a wetland ecosystem. *R. Soc. Open Sci.* 4:170568. doi: 10.1098/rsos.170568

Eichmiller, J. J., Best, S. E., and Sorensen, P. W. (2016). Effects of temperature and trophic state on degradation of environmental DNA in lake water. *Environ. Sci. Technol.* 50, 1859–1867. doi: 10.1021/acs.est.5b05672

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

Fukumoto, S., Ushimaru, A., and Minamoto, T. (2015). A basin-scale application of environmental DNA assessment for rare endemic species and closely related exotic species in rivers: a case study of giant salamanders in Japan. *J. Appl. Ecol.* 52, 358–365. doi: 10.1111/1365-2664.12392

Jo, T., Arimoto, M., Murakami, H., Masuda, R., and Minamoto, T. (2020). Estimating shedding and decay rates of environmental nuclear DNA with relation to water temperature and biomass. *Environ. DNA* 2, 140–151. doi: 10.1002/edn3.51

Jo, T., Murakami, H., Masuda, R., Sakata, M. K., Yamamoto, S., and Minamoto, T. (2017). Rapid degradation of longer DNA fragments enables the improved estimation of distribution and biomass using environmental DNA. *Mol. Ecol. Resour.* 17, e25–e33. doi: 10.1111/1755-0998.12685

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

Kakuda, A., Doi, H., Souma, R., Nagano, M., Minamoto, T., and Katano, I. (2019). Detection of environmental DNA of the invasive red-eared slider in ponds for evaluating their distribution with comparison to water quality. *PeerJ* 7:e8155. doi: 10.7717/peerj.8155

Kasai, A., Takada, S., Yamazaki, A., Masuda, R., and Yamanaka, H. (2020). The effect of temperature on environmental DNA degradation of Japanese eel. *Fish. Res.* 86, 465–471. doi: 10.1007/s12562-020-01409-1

Lance, R. F., Klymus, K. E., Richter, C. A., Guan, X., Farrington, H. L., Carr, M. R., et al. (2017). Experimental observations on the decay of environmental DNA from bighead and silver carps. *Manag. Biol. Invasions* 8:343. doi: 10.3391/mbi.2017.8.3.08

Mächler, E., Osathanunkul, M., and Altermatt, F. (2018). Shedding light on eDNA: neither natural levels of UV radiation nor the presence of a filter feeder affect eDNA-based detection of aquatic organisms. *PLoS One* 13:e0195529. doi: 10.1371/journal.pone.0195529

Martellini, A., Payment, P., and Villemur, R. (2005). Use of eukaryotic mitochondrial DNA to differentiate human, bovine, porcine and ovine sources in fecally contaminated surface water. *Water Res.* 39, 541–548. doi: 10.1016/j.watres.2004.11.012

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

Merkes, C. M., McCalla, S. G., Jensen, N. R., Gaikowski, M. P., and Amberg, J. J. (2014). Persistence of DNA in carcasses, slime and avian feces may affect interpretation of environmental DNA data. *PLoS One* 9:e0113346. doi: 10.1371/journal.pone.0113346

Minamoto, T., Fukuda, M., Katsuhara, K. R., Fujiwara, A., Hidaka, S., Yamamoto, S., et al. (2017). Environmental DNA reflects spatial and temporal jellyfish distribution. *PLoS One* 12:e0173073. doi: 10.1371/journal.pone.0173073

Minamoto, T., Naka, T., Moji, K., and Maruyama, A. (2016). Techniques for the practical collection of environmental DNA: filter selection, preservation, and extraction. *Limnology* 17, 23–32. doi: 10.1007/s10201-015-0457-4

Minamoto, T., Yamanaka, H., Takahara, T., Honjo, M. N., and Kawabata, Z. I. (2012). Surveillance of fish species composition using environmental DNA. *Limnology* 13, 193–197. doi: 10.1007/s10201-011-0362-4

Motulsky, H. J., and Christopoulos, A. (2003). *Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting*. San Diego: GraphPad Software, Inc. Available online at: www.graphpad.com

Moushomi, R., Wilgar, G., Carvalho, G., Creer, S., and Seymour, M. (2019). Environmental DNA size sorting and degradation experiment indicates the state of *Daphnia magna* mitochondrial and nuclear eDNA is subcellular. *Sci. Rep.* 9, 12500. doi: 10.1038/s41598-019-48984-7

Nevers, M. B., Byappanahalli, M. N., Morris, C. C., Shively, D., Przybyla-Kelly, K., Spoljaric, A. M., et al. (2018). Environmental DNA (eDNA): a tool for quantifying the abundant but elusive round goby (*Neogobius melanostomus*). *PLoS One* 13:e0191720. doi: 10.1371/journal.pone.0191720

Nukazawa, K., Hamasuna, Y., and Suzuki, Y. (2018). Simulating the advection and degradation of the environmental DNA of common carp along a river. *Environ. Sci. Technol.* 52, 10562–10570. doi: 10.1021/acs.est.8b02293

Pfleger, M. O., Rider, S. J., Johnston, C. E., and Janosik, A. M. (2016). Saving the doomed: using eDNA to aid in detection of rare sturgeon for conservation (*Acipenseridae*). *Glob. Ecol. Conserv.* 8, 99–107. doi: 10.1016/j.gecco.2016.08.008

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Saito, T., and Doi, H. (2020). Degradation modeling of water environmental DNA: experiments on multiple DNA sources in pond and seawater. *bioRxiv* [Preprint], doi: 10.1101/2020.04.03.023283

Sakata, M. K., Maki, N., Sugiyama, H., and Minamoto, T. (2017). Identifying a breeding habitat of a critically endangered fish, *Acheilognathus typus*, in a natural river in Japan. *Sci. Nat.* 104:100. doi: 10.1007/s00114-017-1521-1

Sansom, B. J., and Sassoubre, L. M. (2017). Environmental DNA (eDNA) shedding and decay rates to model freshwater mussel eDNA transport in a river. *Environ. Sci. Technol.* 51, 14244–14253. doi: 10.1021/acs.est.7b05199

Sassoubre, L. M., Yamahara, K. M., Gardner, L. D., Block, B. A., and Boehm, A. B. (2016). Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. *Environ. Sci. Technol.* 50, 10456–10464. doi: 10.1021/acs.est.6b03114

Sengupta, M. E., Hellström, M., Kariuki, H. C., Olsen, A., Thomsen, P. F., Mejer, H., et al. (2019). Environmental DNA for improved detection and environmental surveillance of schistosomiasis. *Proc. Natl. Acad. Sci. U.S.A.* 116, 8931–8940. doi: 10.1073/pnas.1815046116

Sigsgaard, E. E., Carl, H., Møller, P. R., and Thomsen, P. F. (2015). Monitoring the near-extinct European weather loach in Denmark based on environmental DNA from water samples. *Biol. Conserv.* 183, 46–52. doi: 10.1016/j.biocon.2014.11.023

Strickler, K. M., Fremier, A. K., and Goldberg, C. S. (2015). Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biol. Conserv.* 183, 85–92. doi: 10.1016/j.biocon.2014.11.038

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., and Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol. Ecol.* 21, 2045–2050. doi: 10.1111/j.1365-294X.2012.05470.x

Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., and Kawabata, Z. I. (2012). Estimation of fish biomass using environmental DNA. *PLoS One* 7:e035868. doi: 10.1371/journal.pone.0035868

Thomsen, P. F., Kielgast, J., Iversen, L. L., Møller, P. R., Rasmussen, M., and Willerslev, E. (2012). Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS One* 7:e41732. doi: 10.1371/journal.pone.0041732

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

Tsuji, S., Yamanaka, H., and Minamoto, T. (2017). Effects of water pH and proteinase K treatment on the yield of environmental DNA from water samples. *Limnology* 18, 1–7. doi: 10.1007/s10201-016-0483-x

Turner, C. R., Miller, D. J., Coyne, K. J., and Corush, J. (2014). Improved methods for capture, extraction, and quantitative assay of environmental DNA from Asian bigheaded carp (*Hypophthalmichthys* spp.). *PLoS One* 9:e114329. doi: 10.1371/journal.pone.0114329

Turner, C. R., Uy, K. L., and Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biol. Conserv.* 183, 93–102. doi: 10.1016/j.biocon.2014.11.017

Ushio, M., Murakami, H., Masuda, R., Sado, T., Miya, M., Sakurai, S., et al. (2018). Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. *Metabarcoding Metagenom* 2:e23297. doi: 10.3897/mbmg.2.23297

Villacorta-Rath, C., Adekunle, A. I., Edmunds, R. C., Strugnell, J. M., Schwarzkopf, L., and Burrows, D. (2020). Can environmental DNA be used to detect first arrivals of the cane toad, Rhinella marina, into novel locations? *Environ. DNA* 2, 635–646. doi: 10.1002/edn3.114

Wei, N., Nakajima, F., and Tobino, T. (2018). A microcosm study of surface sediment environmental DNA: decay observation, abundance estimation, and fragment length comparison. *Environ. Sci. Technol.* 52, 12428–12435. doi: 10.1021/acs.est.8b04956

Weltz, K., Lyle, J. M., Ovenden, J., Morgan, J. A., Moreno, D. A., and Semmens, J. M. (2017). Application of environmental DNA to detect an endangered marine skate species in the wild. *PLoS One* 12:e0178124. doi: 10.1371/journal.pone.0178124

Wood, S. A., Biessy, L., Latchford, J. L., Zaiko, A., von Ammon, U., Audrezet, F., et al. (2020). Release and degradation of environmental DNA and RNA in a marine system. *Sci. Total Environ.* 704:135314. doi: 10.1016/j.scitotenv.2019.135314

Zulkefli, N. S., Kim, K. H., and Hwang, S. J. (2019). Effects of microbial activity and environmental parameters on the degradation of extracellular environmental DNA from a eutrophic lake. *Int. J. Environ. Res. Public Health* 16:3339. doi: 10.3390/ijerph16183339

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 29 March 2021
doi: 10.3389/fevo.2021.622149

Spatiotemporal Changes of the Environmental DNA Concentrations of Amphidromous Fish *Plecoglossus altivelis altivelis* in the Spawning Grounds in the Takatsu River, Western Japan

Ryutei Inui[1], Yoshihisa Akamatsu[2], Takanori Kono[3], Minoru Saito[2], Seiji Miyazono[2] and Ryohei Nakao[2]*

[1] Faculty of Socio-Environmental Studies, Fukuoka Institute of Technology, Fukuoka, Japan, [2] Graduate School of Sciences and Technology for Innovation, Yamaguchi University, Yamaguchi, Japan, [3] Public Works Research Institute, National Research and Development Agency, Kakamigahara, Japan

OPEN ACCESS

Edited by:
Matthew A. Barnes, Texas Tech University, United States

Reviewed by:
Bettina Thalinger, University of Guelph, Canada
Takashi Yada, Japan Fisheries Research and Education Agency (FRA), Japan

***Correspondence:**
Ryutei Inui
inuiryutei@gmail.com

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *27 October 2020*
Accepted: *08 March 2021*
Published: *29 March 2021*

Citation:
Inui R, Akamatsu Y, Kono T, Saito M, Miyazono S and Nakao R (2021) Spatiotemporal Changes of the Environmental DNA Concentrations of Amphidromous Fish Plecoglossus altivelis altivelis in the Spawning Grounds in the Takatsu River, Western Japan. Front. Ecol. Evol. 9:622149. doi: 10.3389/fevo.2021.622149

The Ayu *Plecoglossus altivelis altivelis* is an amphidromous fish that is not only the most important commercial fishery species in Japanese rivers but also has a high economic value in recreational fishing. However, the degradation of its spawning grounds has caused a decrease in its abundance. In this study, we used environmental DNA (eDNA) to monitor the Ayu in the Takatsu River in Japan to (1) identify the spawning season in three known spawning grounds, (2) clarify changes in the main spawning grounds during the spawning season, and (3) discover unknown spawning grounds. We collected 1 L of the surface river water at three known spawning grounds nine times in 2018 and seven times in 2019 in the lower reaches of the Takatsu River. We also collected samples from seven unknown sites in 2018. The water samples were filtered through glass fiber filters. Total eDNA was extracted from each filtered sample and a Real-time quantitative PCR was performed with the specific primers and probe for Ayu. The results of the eDNA analyses showed that (1) the spawning season was in November in 2018 and in September in 2019. (2) One site was used as a spawning ground in both the early and the late spawning season, depending on the year. At the second site, the frequency of use changed year by year. The third site was the main spawning ground in the middle to late spawning season every year. From these results, we elucidated that some spawning grounds are used regularly every year, while the use of others varies year by year. (3) In five of the seven unknown sites, the nighttime eDNA concentrations were high at least once during the four surveys, suggesting that these sites may have functioned as spawning grounds. In particular, one site could be an important new spawning ground.

Keywords: *Plecoglossus altivelis*, amphidromous fish, fishery, river, spawning, eDNA, qPCR, Japan

INTRODUCTION

The Ayu *Plecoglossus altivelis altivelis* is an amphidromous fish with a body length of approximately 10–30 cm that inhabits rivers from Japan to Vietnam (Nishida, 2001). The adult fish spawn in the riffles of the lower reaches of rivers from evening to night in fall, and then die after spawning. The eggs adhere to the riverbed and hatch after approximately 10 to 14 days. The larvae migrate downstream to brackish waters and the sea immediately after hatching and live there during the winter. In spring (February to May in western Japan), juveniles migrate upstream and expand their distribution area further upstream as they grow. In fall, they mature and descend downstream to spawn. The Ayu lives for only 1 year (Nishida, 2001).

The Ayu is not only the most important commercial fishery species in Japanese rivers but also has a high economic value in recreational fishing (Takahashi and Azuma, 2006, 2016). However, catches of Ayu have decreased in recent years throughout Japan (Minister of Agriculture, Forestry and Fisheries, 2016), especially in western Japan (Takahashi, 2009). To use Ayu sustainably as a natural resource, it is necessary to understand the dynamics of the populations, appropriate catch limits, the location of suitable habitats for each life stage, and the best times and places for closed seasons.

However, it is not easy to monitor Ayu because they inhabit rapids in summer and migrate to the lower reaches of rivers in the fall, and they are very fast swimmers (Nishida, 2001; Doi et al., 2017). In addition, there are conflicts with commercial and recreational fishermen because the monitoring sites tend to overlap the fishing grounds.

Based on this background, we examined whether environmental DNA (eDNA) could be used for monitoring of Ayu spawning activity. In a previous eDNA study on Ayu, Yamanaka and Minamoto (2016) developed a species-specific primer for Ayu and used the primer to survey the occurrence of Ayu in a year-round eDNA water sampling regime at 15 sites on the Yodo River, Kinki region, Japan. They reported that Ayu DNA was detected at most of the sites in the freshwater area during the warm months. In contrast, in the coldest month of February, eDNA was only detected in the uppermost site (the southern tip of Lake Biwa). Doi et al. (2017) compared the eDNA concentrations of Ayu with the results of daytime snorkeling surveys in seven sites in the Saba River, Chugoku region, Japan. Across the 3 months (May, July, and October), there were significant correlations between the eDNA concentration of Ayu and the species abundance/biomass at study sites within the river.

In this study, we focused on the spawning periods for Ayu, because it has been suggested that the degradation of the spawning grounds (e.g., armor coating of the riverbed owing to dam release) is one of the factors causing the decrease in the abundance of Ayu (Takahashi and Azuma, 2006, 2016). Therefore, for sustainable Ayu stock management, it is important to develop an efficient method for monitoring their spawning grounds and identifying suitable spawning grounds. Another reason is that the "Ochi-ayu fishery" is carried out to collect spawning populations of Ayu near their spawning grounds during the spawning season in many rivers in Japan. If eDNA real time monitoring is feasible in rivers where the Ayu fisheries are conducted, fishermen in each river will know when and where spawning populations are present in real-time and the data can be used to set the appropriate fishing grounds and quotas for each year.

There are several studies using eDNA that focused on fish spawning (e.g., Erickson et al., 2016; Bylemans et al., 2017; Tillotson et al., 2018; Thalinger et al., 2019; Hayer et al., 2020; Yatsuyanagi et al., 2020). Thalinger et al. (2019) and Yatsuyanagi et al. (2020) showed that eDNA concentrations increased because of fish migration into rivers during the spawning season. Erickson et al. (2016); Tillotson et al. (2018), and Hayer et al. (2020) reported that fish spawning behavior in rivers could temporarily increase eDNA concentrations and the spawning behavior could be captured using eDNA. These studies show that eDNA may increase with fish spawning behavior, perhaps because DNA fragments from sperm and dead fish promote a temporary increase in the eDNA concentrations (Bylemans et al., 2017).

Several studies have examined Ayu spawning activity using eDNA. Yamanaka and Minamoto (2016) developed a specific primer for Ayu which all subsequent studies have used. Doi et al. (2017) found higher eDNA concentrations of Ayu in the lower reaches of the Saba River in October than those in May and July possibly owing to the spawning events. Kono et al. (2017) conducted eDNA water sampling during the day in the Takatsu and Saba rivers in the Chugoku region, Japan, six times in May, July, and October-November. These results showed that Ayu moved downstream for spawning in mid-November in both rivers. Inui et al. (2018) conducted eDNA water sampling during the spawning season of Ayu during the daytime and nighttime (3 h after sunset) at four sites: a downstream site near an artificial spawning ground and three control sites (upstream, main river, and downstream) on the Nahari River, Shikoku region, Japan twice in November. Visual diving surveys in the Nahari River over the past decade reported that most of the Ayu spawned in artificial spawning grounds. The results of eDNA analyses showed that the highest eDNA concentrations were found downstream of the artificial spawning grounds at night, and the eDNA concentrations were higher at night compared to the day. In the second half of November (the peak of spawning season), the eDNA concentration at night was 25 times higher than that in the day. Inui et al. (2019) conducted hourly eDNA water sampling from 15:00 to 22:00 on 3 days in November downstream of a site considered to be a major spawning ground in the Saba River. The results showed that the eDNA concentrations increased after sunset on all days, and peaked either 1 h after sunset or 3 h after sunset. Yoshida et al. (2019) conducted a survey of eggs and eDNA analysis (daytime and nighttime) at one known spawning site and two unknown spawning sites on the Asahikawa River, Chugoku region, Japan in October. These results showed that the known spawning sites had the highest eDNA concentrations at night and Ayu eggs were found in all surveys. No eggs were found at unknown sites; however, the eDNA concentrations increased at night in some sites, suggesting the possibility of spawning. Saito et al. (2020) conducted eDNA water sampling eight times from September to December during the day, 1 h after sunset, and 3 h after sunset at six sites in the Takatsu River, Chugoku

region, Japan. The results showed that the difference in the eDNA concentrations between 1 h after sunset and 3 h after sunset was small, but in a few sites the concentration was high only 1 h after sunset and decreased by 3 h after sunset. These results indicate that it is possible to check the spawning status both 1 and 3 h after sunset, but 1 h after sunset is ideal. Inui et al. (2020) sampled river water for an Ayu eDNA survey at three sites in the Shimanto River, Shikoku region, Japan and at three sites in the Takatsu River from November to February downstream of the known spawning sites (10 daytime samples in the Shimanto River, nine daytime samples in the Takatsu River, and one nighttime sample in November in both rivers) and measured water temperatures. The results showed that the spawning season was longer in the Shimanto River, which had a higher water temperature in December than the Takatsu River, and the daytime data indicated that the main spawning sites may have changed in both rivers.

These studies showed that the spawning migration of Ayu in rivers can be followed using eDNA. Furthermore, they demonstrated that the eDNA concentrations of Ayu spawning grounds increased at night during the spawning season. Thus, nighttime surveys can reveal the potential spawning grounds, and 1 to 3 h after sunset is an appropriate time for spawning surveys. Results from these methods suggest that Ayu migrate to the lower reaches of rivers in fall, that eDNA can be used to discover unknown spawning grounds, and that the main spawning grounds change during the spawning season. Previously, Kono et al. (2017) and Inui et al. (2020) studied the spawning season, spawning peak, and seasonal spawning grounds of Ayu in the Takatsu River, Chugoku region, Japan; however, these studies were mainly based on daytime sampling and used only a single year of data. In addition, a study searching for unknown spawning grounds was conducted by Yoshida et al. (2019); however, this was based on other rivers in Japan.

In this study, we clarify the generality of the spawning ecology of Ayu in the Takatsu River, Chugoku region, Japan, using daytime and nighttime eDNA data for two consecutive years (2018 and 2019). We aimed to: (1) identify the actual spawning season of Ayu, (2) clarify changes in the main spawning grounds during the spawning season, and (3) discover whether there were unknown spawning grounds.

MATERIALS AND METHODS

Field Surveys

The field surveys were conducted in the lower reaches of the Takatsu River, which runs through the western part of Shimane Prefecture in Japan (**Figure 1**). We collected 1 L of the surface river water at three known spawning grounds (Sites 1, 2, and 3), nine times between September and December in 2018, and seven times between September and December in 2019. In addition, we sampled at seven locations (Sites A to G) on October 24, and November 1, 8, and 14 in 2018 to discover unknown spawning grounds. The eDNA water sampling was conducted once during the day (2–5 h before sunset) and once during the night (3–4 h after sunset). Water sampling was conducted at two sampling points (one upstream and one downstream of clear riffles), because Ayu spawn upstream and/or within the riffles. Bottles for the water sampling were bleached with 10% bleach solution and washed with DNA-free deionized water in the laboratory. In the field, 1 mL of 10% benzalkonium chloride solution (Fujifilm Wako, Osaka, Japan) was added to each sample to suppress the degeneration of eDNA before filtering water samples (Yamanaka et al., 2017). The water samples were then transported to the laboratory in a cooler box at 4°C. To check for cross-contamination during sampling, each survey day 1 L of deionized water was taken to the site and 1 mL of 10% benzalkonium chloride solution was added in the field (hereafter called a cooler blank). A total of 16 cooler blanks were created in the present study.

Filtration, DNA Extraction, and Quantitative PCR

Water samples and cooler blanks were filtered through a GF/F glass fiber filter (normal pore size = 0.7 μm; diameter = 47 mm; Global Life Sciences Technologies Japan, Tokyo, Japan) within 24 h. To prevent cross-contamination among the water samples, the filter funnels and measuring cups were bleached after filtration with 10% bleach solution. Then, bleached funnels and measuring cups were rinsed with DNA free deionized water. In 2018, 1 L of deionized water was filtered in the same manner in each filtration step in the laboratory and used as filtration negative controls (hereafter called a room blank). A total of nine room blanks were adopted in the present study. All filtered samples were stored at −20°C in the freezer until the DNA extraction step.

Total eDNA was extracted from each filtered sample using the DNeasy Blood and Tissue Kit (QIAGEN, Hilden, Germany). Extraction methods were according to Doi et al. (2017). The filtered sample was placed in the upper part of a Salivette tube (Sarstedt, Nümbrecht, Germany) with 400 μL Buffer AL and 40 μL proteinase K solution and incubated at 56°C for 30 min. Afterward, the Salivette tubes with filters were centrifuged at 5,000 × *g* for 5 min and 220-μL Tris–EDTA (TE) buffer was added to the filtered samples and re-centrifuged at 5,000 × *g* for 1 min. Subsequently, 400 μL of ethanol was added to the collected solution, and the mixture was transferred to a spin column. Total eDNA was eluted in 100 μL buffer AE according to the manufacturer's instructions from the DNeasy Blood and Tissue Kit. All eDNA samples were stored at −20°C until the subsequent experiments.

A Real-time TaqMan quantitative PCR (qPCR) was performed on a PikoReal Real-Time System (Thermo Fisher Scientific, Waltham, MA, United States). The mitochondrial cytochrome b (cytb) gene fragments (131 bp) were amplified and quantified using the primers and TaqMan probe described in Yamanaka and Minamoto (2016): Paa-CytB-Forward: 5′-CCTAGTCTCCCTGGCTTTATTCTCT-3′, Paa-CytB-Reverse: 5′-GTAGAATGGCGTAGGCGAAAA-3′, and Paa-CytB-Probe: 5′-FAM-ACTTCACGGCAGCCAACCCCC-TAMRA-3′. Each 10 μL of PCR mixture contained 1 μL primer-probe mix (900 nM of primers and 125 nM probe), 5 μL TaqMan Environmental Master Mix 2.0 (Life Technologies, Carlsbad, CA, United States),

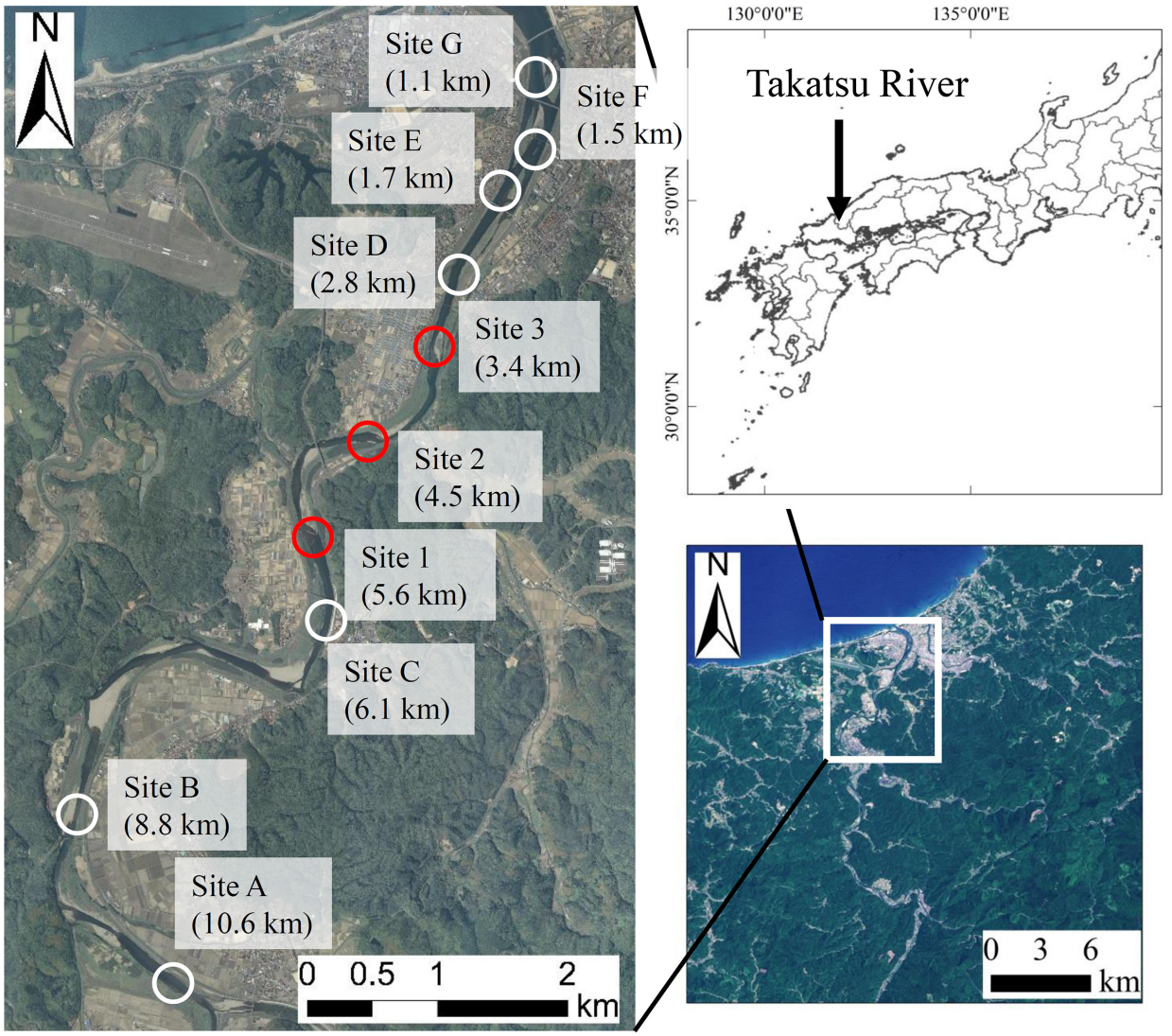


FIGURE 1 | Study sites in the Takatsu River, western Japan. The distance indicated at each point is the distance from the river mouth. Red circles (Sites 1, 2, and 3) indicate the known spawning grounds, and white circles (Sites A to G) indicate the potential spawning grounds.

0.1 μL AmpErase Uracil N-Glycosylase (UNG; Thermo Fisher Scientific, Waltham, MA, United States), 1.9 μL sterilized water (Nacalai tesque, Kyoto, Japan), and 2 μL of the eDNA solution. The qPCR conditions were as follows: UNG incubation step at 50°C for 2 min, denaturation at 95°C for 10 min, followed by 55 cycles of 95°C for 15 s, and 60°C for 60 s. In the 2018 and 2019 samples, room blank and ultrapure water instead of the template DNA water were used as the PCR blanks in each qPCR run. There were four PCR replicates of template DNA and negative controls in 2018 and three in 2019. The standard curve was constructed using a dilution series of 6.0×10^4, 10^3, 10^2, and 10^1 copies per PCR mixture using cloned plasmid DNA with the target sequence. Each PCR replicate of standard DNA was duplicated in 2018 and triplicated in 2019. The R^2 values for the standard curves of all qPCR ranged from 0.912 to 0.995 (**Supplementary Table 1**). In 2019, no DNA was detected in any of the blanks; however, in 2018, DNA was detected there. The DNA concentration of blank samples ranged from 0.00 to 3.57 copies/2 μL in the daytime and nighttime, respectively. Thus, the data of 2018 were corrected by calculating the average of the concentrations of the four blank wells analyzed on that day, excluding the non-detected wells, and subtracting them from the concentration at each site on the relevant day.

Data Analysis

The mean eDNA concentration per template DNA (copies/2 μL) was calculated using the raw data of the qPCR results. Then, eDNA concentration per sample (copies/L) and eDNA flux (copies/s) was also calculated to account for daily changes in

discharge or to make year to year comparisons. Discharge taken were the daily average discharge (m^3/s) estimated from the depth measured at three nearby gauge stations (Ministry of Land, Infrastructure, Transport, and Tourism, 2020) using the H-Q curve for each station.

The eDNA concentrations and fluxes of each site on each day were plotted as a scatter plot with nighttime concentrations/fluxes on the vertical axis and daytime concentrations/fluxes on the horizontal axis. The scatter plot was used to identify the actual spawning season of Ayu, clarify the changes in the main spawning grounds during the spawning season, and discover unknown spawning grounds. In all figures, each data point shows the average of PCR replicates from one water sample from one site, and the upstream site and downstream site of riffles are shown in the separate plots.

RESULTS

Spawning Season

We obtained eDNA concentrations for 132 samples, 16 cooler blanks and 16 PCR blanks from 28 qPCR runs. All the qPCR results are shown in **Supplementary Tables 2, 3**, while the eDNA concentrations and fluxes are shown in **Supplementary Table 4**. The data for Sites 1, 2, and 3 for 2019, and Sites 1, 2, and 3 for 2018 are taken from Inui et al. (2020) and Saito et al. (2020), respectively.

Figure 2 shows the seasonal changes in the daytime and nighttime eDNA concentrations at the three known spawning grounds (Sites 1, 2, and 3). **Figure 3** shows the seasonal changes of the daytime and nighttime eDNA fluxes at the three known spawning grounds. In each figure, the diagonal (i.e., the line $Y = X$) is shown as a reference, and the plots above this line indicate that the value of the eDNA concentration in the nighttime sample divided by the eDNA concentration in the daytime sample (abbreviated to "night/day value") is larger than 1, which could be an indicator of the spawning activity of Ayu.

In 2018, the daytime and nighttime eDNA concentrations and the eDNA fluxes were highest in November, intermediate in September and October, and lowest in December. The values above the diagonal were 67% in September, 50% in October, 83% in November, and 17% in December, respectively. In **Figure 3**, plots above the diagonal indicate the eDNA fluxes have increased in nighttime and those below the diagonal indicate the eDNA fluxes have decreased in nighttime. The distance from each plot to the diagonal in **Figure 3** was calculated and used as the indicator of the spawning activity (i.e., the positive value of the distance is indicative of the spawning activity). In 2018, the mean distances for September, October, November, and December were 188×10^6, 87×10^6, 2679×10^6, and -38×10^6, respectively.

In 2019, the nighttime and daytime eDNA concentrations were highest in September, intermediate in October and November, and lowest in December. The plots above the diagonal were one hundred percent, 58 and 83% of the points were above the diagonal in September, October, and December, respectively. In 2019, the mean distances for September, October, November, and December were $39{,}673 \times 10^6$, -547×10^6, 1327×10^6, and 1×10^6, respectively.

Changes in the Main Spawning Grounds

Figures 4–6 show the eDNA fluxes of the known spawning grounds in September, October, and November for each site.

The results for September are shown in **Figure 4**. In 2018, 50% of the samples in Site 1, 66.7% in Site 2, and 66.7% in Site 3 were above the diagonal. The nighttime eDNA fluxes were higher at Site 1 and Site 3 on September 26. In 2019, all samples at Site 1 and Site 3 were above the diagonal and had higher nighttime eDNA fluxes.

The results for October are shown in **Figure 6**. In 2018, few sites were above the diagonal although Site 3 had relatively high nighttime eDNA fluxes and night/day values. In 2019, Site 2 and Site 3 had high eDNA fluxes at nighttime, and the night/day values were mostly above 1.

The results for November are shown in **Figure 7**. In 2018, Site 3 had high nighttime eDNA fluxes and all the samples were above the diagonal. Site 2 had relatively high nighttime eDNA fluxes, with 62.5% of the values above the diagonal. Site 1 did not have high levels of either nighttime eDNA fluxes or night/day values. In 2019, most of the plots at the three sites were above the diagonal.

Searching for New Spawning Grounds

Figures 7–10 show the eDNA fluxes of the potential spawning grounds (Sites A to G). For comparison, the eDNA fluxes of the known spawning grounds (Sites 1, 2, and 3) on the same day are shown.

The results for October 24 are shown in **Figure 7**. The nighttime eDNA fluxes at Site D and Site E were higher than the average eDNA fluxes of the known spawning grounds, and the nighttime/daytime values of both sites were also higher.

On November 1, no sites exceeded the known spawning grounds on the nighttime eDNA fluxes; however, Sites B to E had relatively high nighttime eDNA fluxes and nighttime/daytime values (**Figure 8**).

The results on November 8 showed that no sites exceeded the nighttime eDNA fluxes of the known spawning grounds, but Sites D and F had relatively high nighttime eDNA fluxes and nighttime/daytime values (**Figure 9**).

The results on November 14 are shown in **Figure 10**. There were no sites that exceeded the nighttime eDNA fluxes of the known spawning grounds, but Sites C, D, and F had relatively high nighttime eDNA fluxes and nighttime/daytime values.

DISCUSSION

Spawning Season

The results shown in **Figures 2**, **3** suggest that the spawning season was mainly in November in 2018 and in September in 2019. The eDNA concentrations and fluxes in December were lower than those in September, October, and November in both years, suggesting that the spawning season was in the final stage

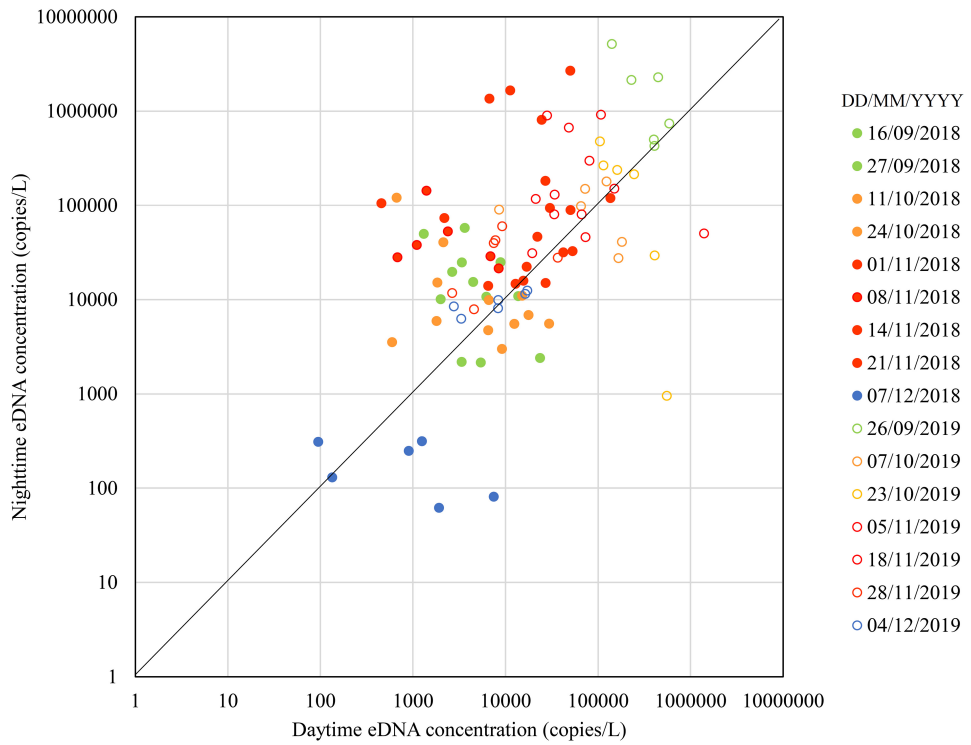


FIGURE 2 | Seasonal changes in daytime and nighttime environmental DNA concentrations in 2018 and 2019.

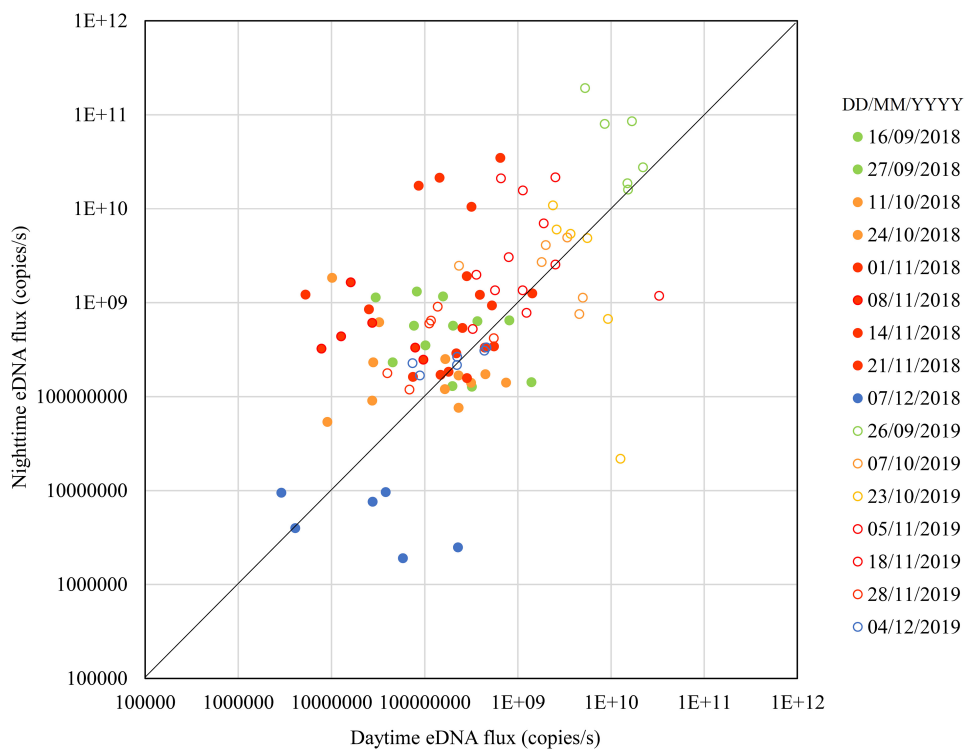


FIGURE 3 | Seasonal changes in daytime and nighttime environmental DNA fluxes in 2018 and 2019.

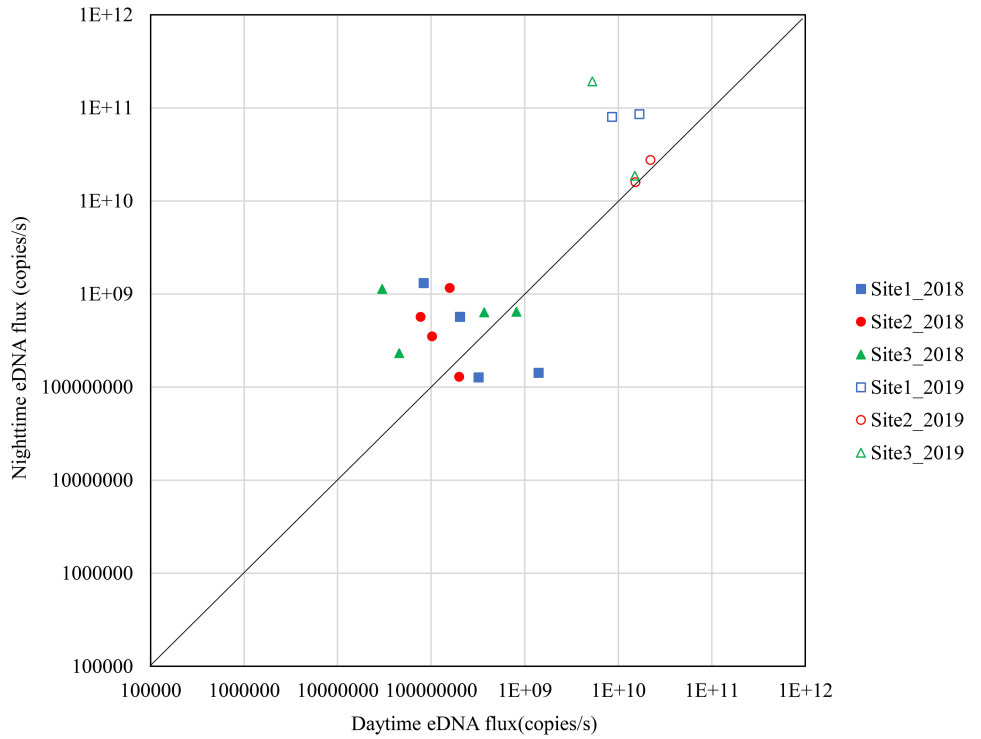


FIGURE 4 | Differences in daytime and nighttime environmental DNA fluxes at sites 1–3 in September 2018 and 2019.

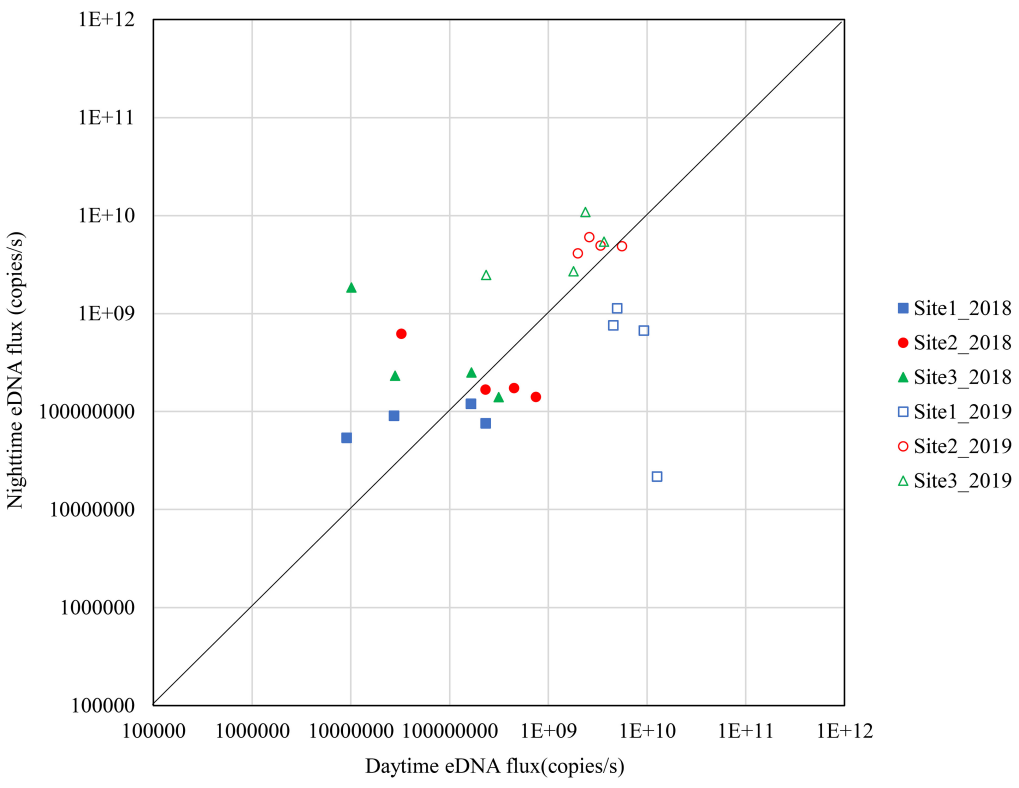


FIGURE 5 | Differences in daytime and nighttime environmental DNA fluxes at sites 1–3 in October 2018 and 2019.

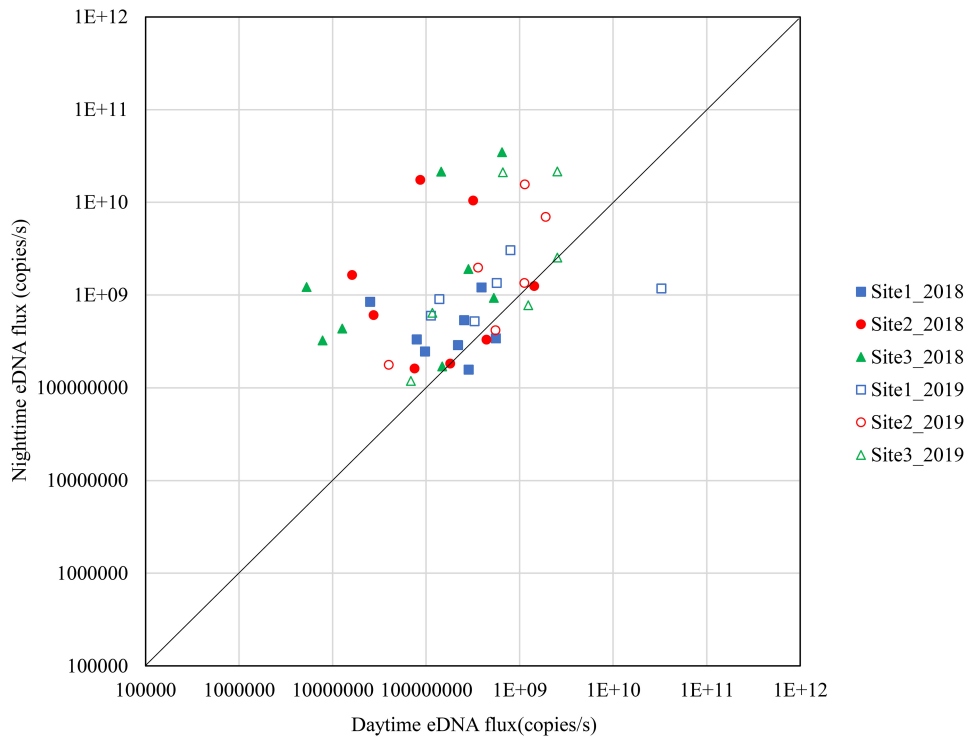


FIGURE 6 | Differences in daytime and nighttime environmental DNA fluxes at sites 1–3 in November 2018 and 2019.

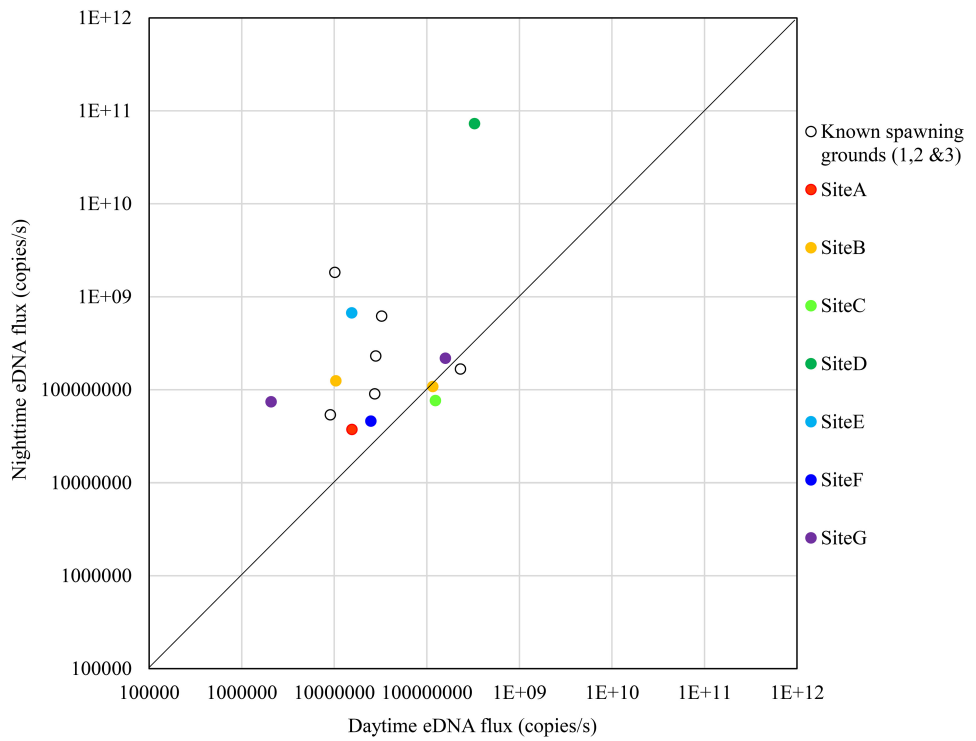


FIGURE 7 | Comparison of daytime and nighttime environmental DNA fluxes between the potential spawning grounds (Sites A to G) and the known spawning grounds (Sites 1, 2, and 3) on October 24, 2018.

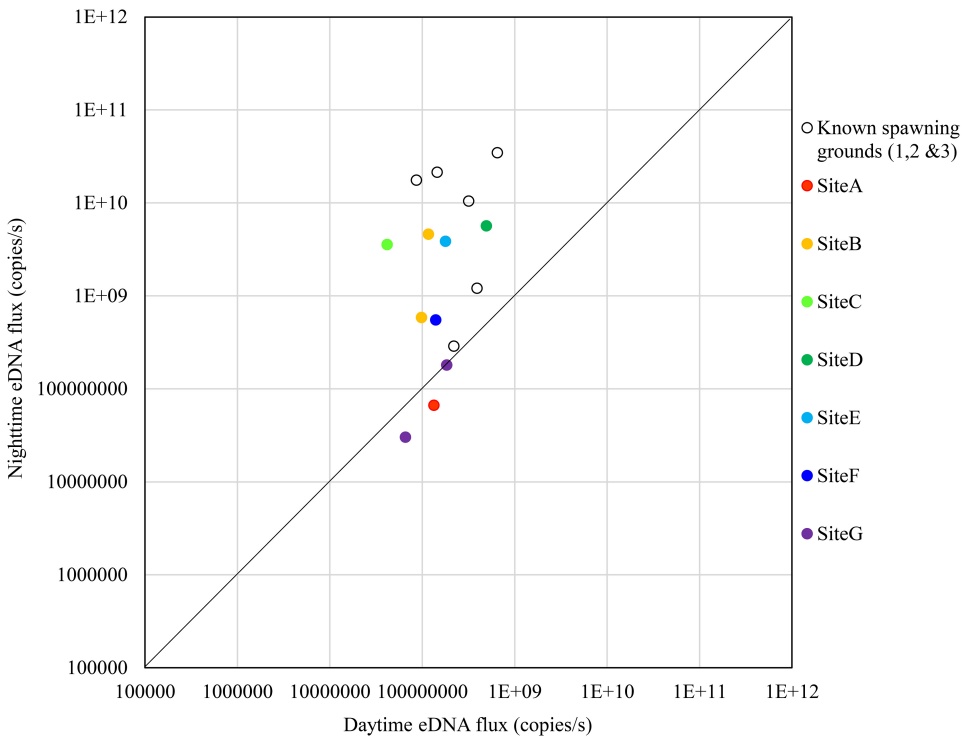


FIGURE 8 | Comparison of daytime and nighttime environmental DNA fluxes between the potential spawning grounds (Sites A to G) and the known spawning grounds (Sites 1, 2, and 3) on November 1, 2018.

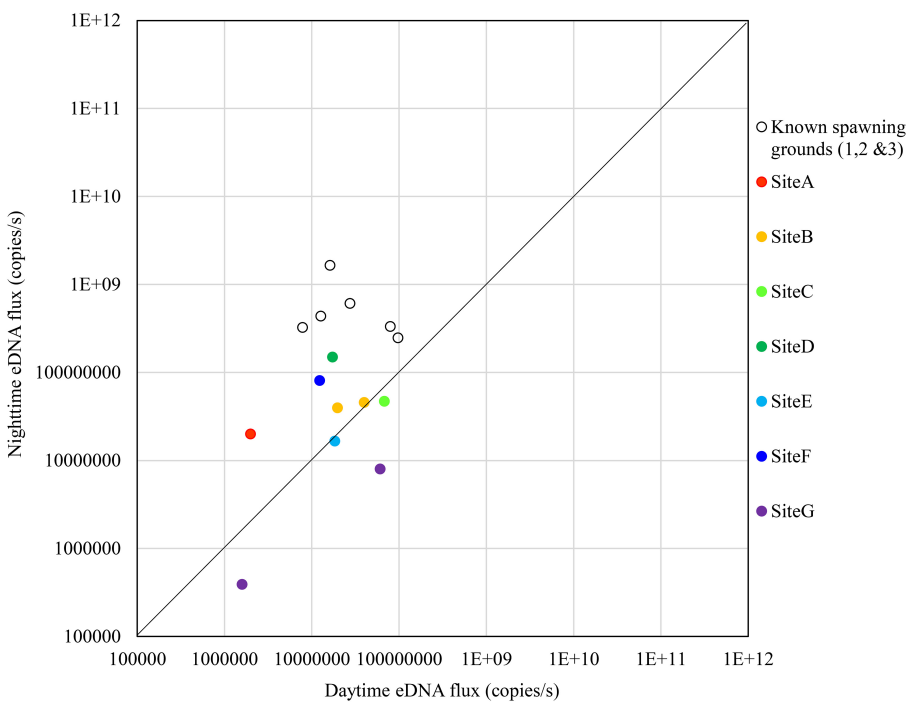


FIGURE 9 | Comparison of daytime and nighttime environmental DNA fluxes between the potential spawning grounds (Sites A to G) and the known spawning grounds (Sites 1, 2, and 3) on November 8, 2018.

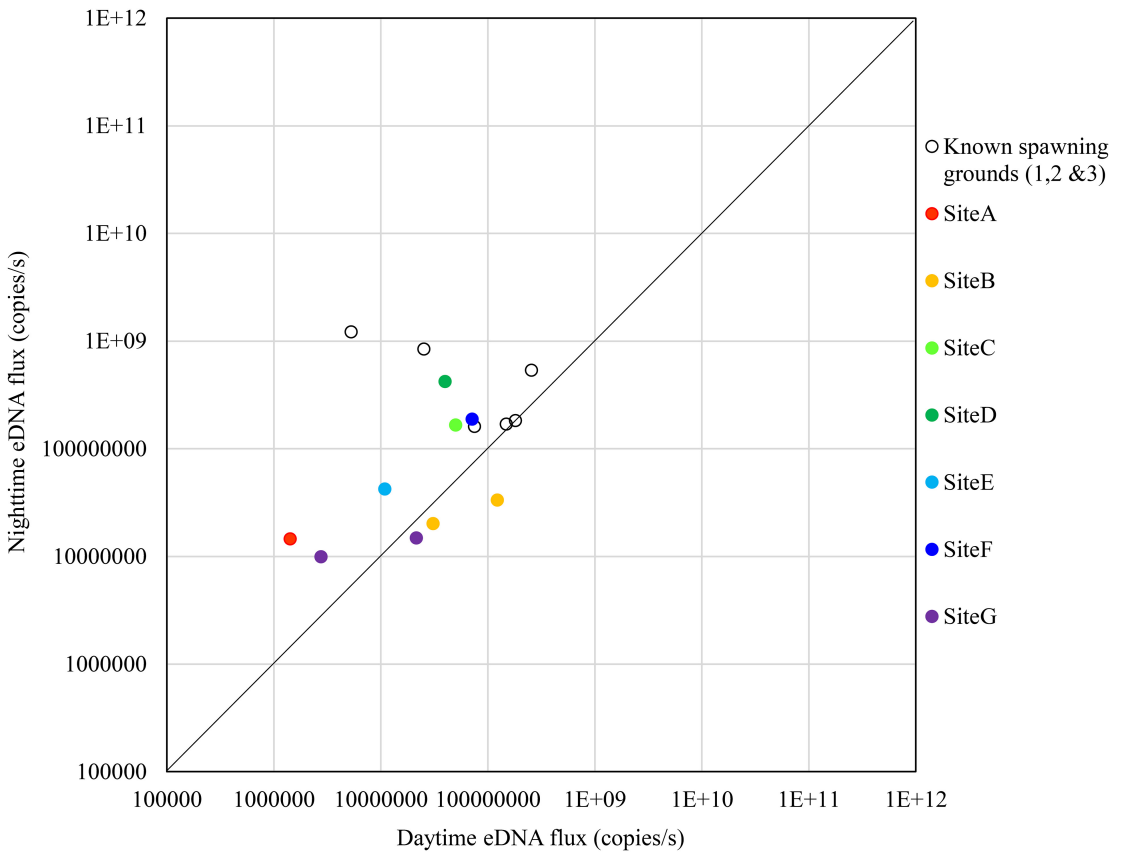


FIGURE 10 | Comparison of daytime and nighttime environmental DNA fluxes between the potential spawning grounds (Sites A to G) and the known spawning grounds (Sites 1, 2, and 3) on November 14, 2018.

or had ended by December. These results suggest that although the spawning season of Ayu varied from year to year, it is in the final stages or ended by December. Tillotson et al. (2018) reported that the contribution of corpses on eDNA concentration was greater than those of the living adults in a Salmonid, *Oncorhynchus nerka* in an Alaskan stream. Ayu have life history traits similar to the Salmonid that the majority of adults die soon after spawning (Nishida, 2001). During our surveys, dead bodies of Ayu were rarely found. Consequently, in the Takatsu River, dead Ayu after spawning have been likely eaten up by predators or transported downstream, and probably had less biomass of corpses remaining in the river than salmon. Therefore, it is possible that the end of the Ayu spawning season can be clearly identified with eDNA.

Yoshida et al. (2019) showed that eDNA concentrations increased at night in the riffles where eggs were found, and in our study nighttime eDNA concentrations also increased during the presumed spawning season. Kono et al. (2017) showed that the peak of downstream migration for spawning was in early November and that the spawning season had ended in late November in the Takatsu River based on temporal changes in eDNA concentrations in the day. Our study provides a clearer indication of the estimated peak date and the end of the spawning season by using eDNA concentrations at night.

Changes in the Main Spawning Grounds

From the results shown in **Figures 4**–**7**, in September, Site 1 and Site 3 were used for spawning in 2018, and all sites were used for spawning in 2019. In October, Site 3 was used for spawning in 2018, and Site 2 and Site 3 were used for the spawning in 2019. In November, Site 2 and Site 3 were used for spawning in 2018, and all sites were used for spawning in 2019. Based on these results, the three known spawning sites can be characterized as follows. Site 1 is a spawning ground in the early spawning season, but is it also used in the late spawning season depending on the year. In Site 2 the frequency of use changes from year to year. Site 3 is the main spawning ground in the middle to late spawning season every year. From the results of this study, we elucidated that there could be spawning grounds that are used every year, such as Site 3, and spawning grounds that are used differently from year to year, such as Site 2.

Suitable spawning grounds for Ayu are relatively soft riverbeds consisting of gravels and cobbles (Takahashi and Azuma, 2016). It is thought that flooding helps to create such riverbeds by decreasing the embeddedness of the riverbed

(Yoshida et al., 2018). Comparing the maximum annual outflows in 2018 and 2019, the former had an average daily flow of 1,253 m^3/s and the latter 1,040 m^3/s (Ministry of Land, Infrastructure, Transport, and Tourism, 2020). Therefore, there was no substantial difference in the maximum annual discharge between the 2 years. This suggests that the disturbance to the riverbed was not considerably different between the 2 years. Conversely, both the eDNA concentrations and fluxes were higher in 2019 than those in 2018, suggesting that the abundance of Ayu in 2019 may have been higher than in 2018 (**Figures 2**, **3**). Therefore, we expect that adult stocks in 2018 were lower than those in 2019, and the downstream sites which are more favorable for the drifting of larvae were mainly used for spawning in 2018.

Kono et al. (2017) showed that the daytime eDNA concentrations at Site 2 and Site 3 were the highest among the known spawning grounds. However, it was not possible to determine the location of major spawning grounds from those data because their study did not compare the daytime eDNA concentrations with the nighttime eDNA concentrations and the difference in daytime eDNA concentrations between sites was small compared to the difference between the study days. The comparison of the daytime and nighttime eDNA concentrations and fluxes allowed us to identify the changes in the main spawning grounds within the study days.

Searching for New Spawning Grounds

In Sites B, C, D, E, and F, the nighttime eDNA fluxes and night/day values were high at least once during the four surveys, suggesting that these sites may have functioned as the spawning grounds. In particular, Site D could be an important spawning ground because it had high nighttime eDNA fluxes and night/day values across all four surveys. Generally, Ayu use shallow riffles as their spawning grounds. Because Site D is approximately 1 m deep, it may not have been recognized as a spawning ground until this study. There is a high probability that the eggs of Ayu will be found in Site D, if a visual survey is carried out in the future. By using the eDNA method, we were able to identify a previously unknown potential spawning site. This was one of the major discoveries of our study.

By clarifying the eDNA concentrations and fluxes of day and night samples during the spawning season, as we did in this study, we will be able to discover unknown potential spawning grounds in rivers other than the Takatsu River. In addition, we will be able to narrow down the areas of the spawning grounds of Ayu in large river systems by using eDNA to guide visual surveys of the eggs. This survey scheme will enable us to discover important, previously unknown spawning sites and reassess the value of previously known spawning grounds in each river. However, there are some limitations. For instance, in small to medium scale rivers with multiple spawning grounds it could be difficult to distinguish eDNA signals from each site owing to the short distances among the spawning sites. Yamaguchi et al. (2018) showed that the influence range of eDNA from Ayu was between 1,000 and 2,000 m; however, the eDNA concentration was reduced by half during the 80 m flow down on experimental studies. This rapid depletion of eDNA concentrations during the downstream transportation is thought to be caused not only by the eDNA degradation but also by the eDNA sedimentation (Yamaguchi et al., 2018). If the sampling sites are close to each other and the eDNA reaches the next spawning site before the sedimentation, it will be difficult to identify the spawning site with eDNA. In addition, if large Ayu farms are located near the spawning grounds, point-source contamination of Ayu eDNA could muffle the eDNA signals produced during spawning activities although there are no large Ayu farms in the Takatsu River. Furthermore, in cases where Ayu spawning grounds are not clumped in the lower reaches, employing multiple water sampling teams would enable coverage of the entire span of possible spawning grounds within a short period of time. In such cases, collaboration with fishery cooperative association and/or with residents near the river may be important for monitoring.

This study revealed the spatiotemporal variation in the spawning season and suitable spawning grounds by examining the eDNA concentrations of Ayu sampled in the daytime and nighttime at the known spawning grounds in the Takatsu River in 2018 and 2019. In addition, examining the eDNA concentrations at multiple sites in 2018 at daytime and nighttime allowed us to find new potential spawning grounds. The method used in this study will enable us to quickly identify potential spawning grounds for Ayu and to minimize harmful effects to eggs caused by surveys. Applying eDNA methods to multiple rivers will also lead to more knowledge about suitable spawning environments. Furthermore, if this monitoring method is applied to a river where the "Ochi ayu" fishery is carried out, the eDNA concentration during the daytime will indicate the existence of the main population and the eDNA concentration at night will indicate the spawning grounds. This will be important information for setting the fishing locations and catch quota.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

Ethical review and approval was not required for the animal study, because we only analyzed the DNA released by the fish into the river water.

AUTHOR CONTRIBUTIONS

RI designed the plan of this study, conducted field survey, and analyzed all data. YA designed the study plan and obtained research funding. TK conducted field survey in 2018 and 2019. MS and SM conducted field survey and eDNA analysis in 2019. RN supervised the eDNA analysis of the entire study, checked the eDNA results, and instructed all co-authors on eDNA

techniques. All authors contributed to the article and approved the submitted version.

FUNDING

This study was conducted as a part of the research project, "A study on the impact of riverbed excavation on the river environment in the Takatsu River," commissioned by the Ministry of Land, Infrastructure, Transport, and Tourism in Japan. Takatsugawa Fishery Cooperative Association permitted us to conduct the survey, and the Chugoku Regional Development Bureau of the Ministry of Land, Infrastructure, and Transport provided us with support and valuable data on the river environment.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.622149/full#supplementary-material

Supplementary Table 1 | Results of real-time quantitative PCR quality for each experiment in the present study.

Supplementary Table 2 | Detailed descriptions of real-time quantitative PCR results in 2018 samples.

Supplementary Table 3 | Detailed descriptions of real-time quantitative PCR results in 2019 samples.

Supplementary Table 4 | Detailed results of environmental DNA concentrations and flux analysis for individual samples used in the present study.

REFERENCES

Bylemans, J., Furlan, M. E., Hardy, M. C., McGuffie, P., Lintermans, M., and Gleeson, M. D. (2017). An environmental DNA-based method for monitoring spawning activity: a case study, using the endangered *Macquarie* perch (*Macquaria australasica*). *Ecol. Evol.* 8, 646–655. doi: 10.1111/2041-210X.12709

Doi, H., Inui, R., Akamatsu, Y., Kanno, K., Yamanaka, H., Takahara, T., et al. (2017). Environmental DNA analysis for estimating the abundance and biomass of stream fish. *Freshw. Biol.* 62, 30–39. doi: 10.1111/fwb.12846

Erickson, A. R., Rees, B. C., Coulter, A. A., Merkes, M. C., McCalla, G. S., Touzinsky, F. K., et al. (2016). Detecting the movement and spawning activity of bigheaded carps with environmental DNA. *Mol. Ecol. Resour.* 16, 957–965. doi: 10.1111/1755-0998.12533

Hayer, C., Bayless, F. M., George, A., Thompson, N., Richter, A. C., and Chapman, C. D. (2020). Use of environmental DNA to detect grass carp spawning events. *Fishes* 5:27. doi: 10.3390/fishes5030027

Inui, R., Akamatsu, Y., Okada, S., Kono, T., and Nakao, R. (2020). A comparison of *Plecoglossus altivelis* spawning season in Shimanto River and Takatsu River using environmental DNA–The effect of water temperature. *Adv. Riv. Eng.* 26, 331–336.

Inui, R., Kono, T., Akamatsu, Y., Goto, M., and Yamaguchi, K. (2019). Examination of an appropriate method to monitor the spawning ground of *Plecoglossus altivelis* using environmental DNA – focusing on timely examination. *Adv. Riv. Eng.* 25, 429–434.

Inui, R., Takahashi, I., Goto, M., Akamatsu, Y., and Kawaguchi, Y. (2018). Monitoring the use of artificial spawning grounds for *Plecoglossus altivelis altivelis* at the Nahari River – focusing on the comparison between visual survey and environmental DNA analysis. *Adv. Riv. Eng.* 24, 333–338.

Kono, T., Akamatsu, Y., Goto, M., and Inui, R. (2017). Quantification of *Plecoglossus altivelis* using environmental DNA and trial of monitoring of downstream migration. *Adv. Riv. Eng.* 23, 669–674.

Minister of Agriculture, Forestry and Fisheries (2016). *Data From: Fishery and Aquaculture Production Statistics in 2014, National Production by Year and Fish Type From 2004 to 2014. Digital Repository of Minister of Agriculture, Forestry and Fisheries.* Available online at: https://www.maff.go.jp/j/tokei/kouhyou/kaimen_gyosei/ [accessed Oct. 1, 2020]

Ministry of Land, Infrastructure, Transport, and Tourism (2020). *Data From: Site Information for Takatsuno. Water Information System*. Available online at: http://163.49.30.82/cgi-bin/SiteInfo.exe?ID=307061287707040 [accessed Oct. 1, 2020]

Nishida, M. (2001). "Ayu," in *Freshwater Fishes of Japan*, eds H. Kawanabe, N. Mizuno, and K. Hosoya (Tokyo: Yama-Kei Press), 66–79.

Saito, M., Inui, R., Kono, T., Akamatsu, Y., Nakao, R., and Kobayashi, K. (2020). Examinations on appropriate time of a day for evaluating spawning intensity of *Plecoglossus altivelis altivelis* using environmental DNA in the Takatsu River, Western Japan. *J. Jpn. Soc. Civil Eng. Ser. B1 (Hydraul. Eng.)* 76, I_1315–I_1320.

Takahashi, I. (2009). *A River Where Natural Ayu Grow*. Tokyo: Tsukiji Shokan Press.

Takahashi, I., and Azuma, K. (2006). *The Book of Ayu*. Tokyo: Tsukiji Shokan Press.

Takahashi, I., and Azuma, K. (2016). *The Book of Natural Ayu*. Tokyo: Tsukiji Shokan Press.

Thalinger, B., Wolf, E., Traugott, M., and Wanzenböck, J. (2019). Monitoring spawning migrations of potamodromous fish species via eDNA. *Nature Sci. Rep.* 9:15388. doi: 10.1038/s41598-019-51398-0

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

Yamaguchi, K., Akamatsu, Y., Inui, R., Goto, M., Kono, T., and Kurita, Y. (2018). Fundamental study on the dynamics of materials containing eDNA in rivers. *J. Jpn. Soc. Civil Eng. Ser. B1 (Hydraul. Eng.)* 74, I_409–I_414.

Yamanaka, H., and Minamoto, T. (2016). The use of environmental DNA of fishes as an efficient method of determining habitat connectivity. *Ecol. Indic.* 62, 147–153. doi: 10.1016/j.ecolind.2015.11.022

Yamanaka, H., Minamoto, T., Matsuura, J., Sakurai, S., Tsuji, S., Motozawa, H., et al. (2017). A simple method for preserving environmental DNA in water samples at ambient temperature by addition of cationic surfactant. *Limnology* 18, 233–241. doi: 10.1007/s10201-016-0508-5

Yatsuyanagi, T., Ishida, R., Sakata, K. M., Kanbe, T., Mizumoto, H., Kobayashi, Y., et al. (2020). Environmental DNA monitoring for short-term reproductive migration of endemic anadromous species. *Shishamo* smelt (*Spirinchus lanceolatus*). EDNA. 2, 130–139. doi: 10.1002/edn3.50

Yoshida, K., Inui, R., Udagawa, R., Maeno, S., Akamatsu, Y., Nigo, S., et al. (2019). Spawning habitat potential evaluation for Ayu using an environmental DNA analysis in the lower Asahi River with degraded riverbed. *J. Jpn. Soc. Civil Eng. Ser. B1 (Hydraul. Eng.)* 75, 529–534. *vol,

Yoshida, K., Maeno, S., Takahashi, S., Nigo, S., Ogawa, S., and Akoh, R. (2018). Physical environmental evaluation of the spawning for Ayu based on estimation of flow field and sediment behavior using ALB data in the lower Asahi River. *J. Jpn. Soc. Civil Eng. Ser. B1 (Hydraul. Eng.)* 74, I_421–I_426.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 08 April 2021
doi: 10.3389/fevo.2021.650717

Spatial Heterogeneity of eDNA Transport Improves Stream Assessment of Threatened Salmon Presence, Abundance, and Location

Zachary T. Wood[1,2†], Anaïs Lacoursière-Roussel[3,5*†], Francis LeBlanc[4], Marc Trudel[3], Michael T. Kinnison[1,2], Colton Garry McBrine[5], Scott A. Pavey[5] and Nellie Gagné[4]*

[1] The University of Maine School of Biology and Ecology, Orono, ME, United States, [2] Maine Center for Genetics in the Environment, Orono, ME, United States, [3] St. Andrews Biological Station, Fisheries and Oceans Canada, St Andrews, NB, Canada, [4] Atlantic Science Enterprise Centre, Fisheries and Oceans Canada, Moncton, NB, Canada, [5] Biological Sciences, Canadian Rivers Institute, University of New Brunswick, Saint John, NB, Canada

OPEN ACCESS

Edited by:
Matthew A. Barnes, Texas Tech University, United States

Reviewed by:
Bryan Neff, University of Western Ontario, Canada
Didier Pont, University of Natural Resources and Life Sciences, Vienna, Austria

***Correspondence:**
Zachary T. Wood
zachary.t.wood@maine.edu
Anaïs Lacoursière-Roussel
anais.lacoursiere@dfo-mpo.gc.ca

†These authors share first authorship

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *07 January 2021*
Accepted: *01 March 2021*
Published: *08 April 2021*

Citation:
Wood ZT, Lacoursière-Roussel A, LeBlanc F, Trudel M, Kinnison MT, Garry McBrine C, Pavey SA and Gagné N (2021) Spatial Heterogeneity of eDNA Transport Improves Stream Assessment of Threatened Salmon Presence, Abundance, and Location. Front. Ecol. Evol. 9:650717. doi: 10.3389/fevo.2021.650717

The integration of environmental DNA (eDNA) within management strategies for lotic organisms requires translating eDNA detection and quantification data into inferences of the locations and abundances of target species. Understanding how eDNA is distributed in space and time within the complex environments of rivers and streams is a major factor in achieving this translation. Here we study bidimensional eDNA signals in streams to predict the position and abundance of Atlantic salmon (*Salmo salar*) juveniles. We use data from sentinel cages with a range of abundances (3–63 juveniles) that were deployed in three coastal streams in New Brunswick, Canada. We evaluate the spatial patterns of eDNA dispersal and determine the effect of discharge on the dilution rate of eDNA. Our results show that eDNA exhibits predictable plume dynamics downstream from sources, with eDNA being initially concentrated and transported in the midstream, but eventually accumulating in stream margins with time and distance. From these findings we developed a fish detection and distribution prediction model based on the eDNA ratio in midstream versus bankside sites for a variety of fish distribution scenarios. Finally, we advise that sampling midstream at every 400 m is sufficient to detect a single fish at low velocity, but sampling efforts need to be increased at higher water velocity (every 100 m in the systems surveyed in this study). Studying salmon eDNA spatio-temporal patterns in lotic environments is essential to developing strong quantitative population assessment models that successfully leverage eDNA as a tool to protect salmon populations.

Keywords: water eDNA, predictive model, quantitative distribution assessment, conservation, Atlantic salmon, lotic ecosystems, fish detection

INTRODUCTION

Recent advances in the collection and analysis of extra-organismal environmental DNA (eDNA) provide a novel indirect approach that can fill gaps in large-scale fish distribution assessments, complementing logistically difficult traditional methods. In addition to improving power of detection, eDNA promises to augment current fish abundance estimates, increasing their precision

and accuracy (Lacoursière-Roussel et al., 2015; Doi et al., 2017; Levi et al., 2019). Collecting water samples for eDNA surveys is non-invasive and, in contrast to direct organismal surveys, does not impose any stress on the studied organisms (Dolan and Miranda, 2004; Rummer and Bennett, 2005; Miranda and Kidwell, 2010). Using eDNA to detect and quantify aquatic populations has the power to drastically improve our knowledge of the large- and fine-scale spatial distribution of animals (Lacoursière-Roussel et al., 2016a; Yates et al., 2019). Although a number of studies have started to address the effects of environmental conditions [e.g., water temperature (Lacoursière-Roussel et al., 2016b), the ecology of species (e.g., life stage; Gibson et al., 2003), and eDNA hydrodynamics (Deiner and Altermatt, 2014; Jerde et al., 2016; Wood et al., 2020)], more work is needed to refine models and facilitate the integration of eDNA into conservation and fisheries management decision-making (Barnes et al., 2014; Barnes and Turner, 2016; Sepulveda et al., 2021).

Environmental DNA has three hierarchical potential uses for broadly examining single species: (1) *detection*, determining the large-scale spatial distribution of a species; (2) *quantification*, determining the population size in a system based on the eDNA concentration or detection rates, and (3) *quantitative distribution assessment*, localizing high or low concentrations of organisms to particular geographic locations based on eDNA variability. To date, most eDNA applications have focused on species detection, and few have focused on examining system-wide quantification (Yates et al., 2019). In lotic habitats (rivers and streams), fine-scale population quantification and quantitative distribution assessment are currently limited by our ability to translate eDNA distribution to upstream fish distributions. eDNA distribution is impacted by the physical properties of the stream, e.g., morphology and hydrodynamics (Dejean et al., 2011; Deiner and Altermatt, 2014; Jane et al., 2015). Despite the advantages of lotic eDNA surveys over traditional electric and net surveys in terms of person-power, cost, and potential harm to study organisms, implementation for management could be substantially improved by better characterizing of lotic eDNA dynamics that influence eDNA-based detection, quantification, and distribution assessment of aquatic species.

Examining eDNA in Streams

Despite the huge potential of eDNA for aquatic population management, there is no current model able to accurately and precisely predict upstream fish location or abundance based on lotic eDNA concentrations alone, particularly if eDNA is collected in a limited number of stream locations. In riverine systems, eDNA concentrations exhibit high variability in space and time as eDNA moves downstream from sources (Deiner et al., 2016; Shogren et al., 2016; Wilcox et al., 2016; Sansom and Sassoubre, 2017; Wood et al., 2020). Furthermore, due to the complexity of eDNA states (e.g., tissues, cells, and DNA fragments), eDNA movement is more complex than a conservative tracer or monodispersed solution (Wilcox et al., 2016; Shogren et al., 2017; Pont et al., 2018). Sansom and Sassoubre (2017) presented the first model of downstream eDNA transport based on a simplified longitudinal eDNA decay rate constant. In a marine system, Akatsuka et al. (2018) developed a numerical fate and transport simulation of eDNA and successfully paired the simulations with field sampling, while Andruszkiewicz et al. (2019) showed that an eDNA particle tracking model can be used to identify possible eDNA sources. Finally, Laporte et al. (2020) demonstrated that bidimensional hydrodynamic modeling including downstream advection and lateral dispersion predicted both detection and quantities of eDNA in a large estuarine system. Here, we attempt to build on this work by developing upstream quantitative population distribution models from downstream eDNA.

Identifying the timing and extent of hydrological and material isolation and connectivity between eDNA catch probability is necessary to calculate the ratio of transport and production rates (i.e., generalized Damköhler number) and is key to predicting the frequency and location of hot spots for detecting species using eDNA (Abbott et al., 2016). Early work treated eDNA downstream movement as exhibiting constant loss, based on a misperception that such systems are well mixed (e.g., Jane et al., 2015). Subsequent research suggests that eDNA is released from fish in plume and does not mix evenly downstream (Wood et al., 2020). This work hypothesizes that downstream of fish, eDNA is carried in the currents of the main channel. This eDNA "plume" disperses laterally (i.e., perpendicular to the predominant direction of flow) over time and distance into slow water margins, resulting in more evenly distributed but less concentrated eDNA farther downstream from the fish (**Figure 1**; Laporte et al., 2020). This plume poses several challenges for eDNA sampling. First, the plume leads to a methodological tradeoff, wherein sampling close to the fish

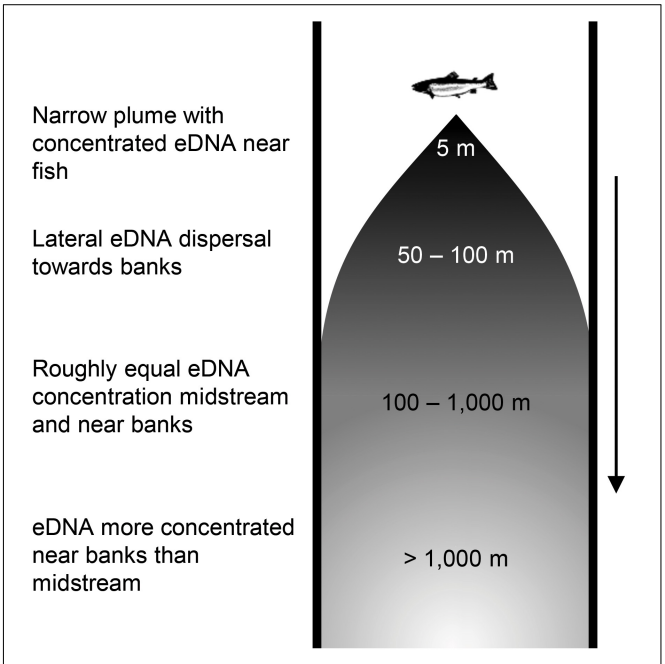


FIGURE 1 | Hypothesized eDNA plume in rivers and streams. See section "Results" for explanation of distance estimates and plume phases.

risks missing its very narrow plume, causing detection to be initially lower on average for non-targeted samples collected nearer the source (i.e., a breakout window), but sampling well downstream from the fish results in lower eDNA concentrations that could further bias detection or quantification of aquatic species. Second, the plume makes bankside sampling—which is more pragmatic in many streams and rivers—unpredictable for detection or abundance estimation depending on the source location. Finally, the progressive dispersion and dilution of eDNA as it moves downstream means that any particular eDNA sample will be an unequal integration of upstream fishes' eDNA, thus requiring careful calculus for accurate fish enumeration. While these challenges represent current hurdles to fish management with eDNA, all can be surmounted by calibrating quantitative population eDNA models based on experiments elucidating plume eDNA dynamics, as well as plume-conscious sampling designs.

Stream eDNA and Atlantic Salmon

Here we use Atlantic salmon in Bay of Fundy tributaries (New Brunswick, Canada) as a case study for examining spatial eDNA dynamics in streams. The abundance of Atlantic salmon has declined well below their conservation limits in the Bay of Fundy since the 1980s, with several stream-specific populations being extirpated from their native habitats (Jones et al., 2014; Fisheries and Oceans Canada, 2019). As they migrate back and forth between freshwater and marine environments to complete their life cycle, Atlantic salmon can be particularly vulnerable to a gauntlet of anthropogenic stressors along their migration corridor (Parrish et al., 1998; Cairns, 2001; Limburg and Waldman, 2009; Brown et al., 2013). As a result, fishery restrictions and recovery measures have been put in place to protect and recover these populations (Fisheries and Oceans Canada, 2019). Assessing their distribution and the habitat suitable for spawning, growth, and survival of juveniles in streams and rivers is thus essential to developing efficient protection and habitat restoration management strategies. Quantitative population size is also essential to evaluating the effectiveness of various recovery measures. However, current methods for assessing the distribution and abundance of Atlantic salmon in streams and rivers are time consuming, require considerable effort in the field, and risk inadvertent injury or mortality to salmon (Dolan and Miranda, 2004; Rummer and Bennett, 2005; Miranda and Kidwell, 2010). Spatial eDNA distribution models will lead to stronger salmon quantitative population estimates if we better understand how eDNA disperses and dilutes within the environment.

This project aims to develop a framework to assess quantitative population distribution in lotic ecosystems based on the spatial distribution of eDNA. First we evaluated spatial eDNA distributions and tested if the plume model is valid based on known quantities of salmon placed into sentinel cages in three salmon-free streams in Southwest New Brunswick. Second, we used the results from these sentinel cage experiments to build a model that uses eDNA dynamics to pinpoint and quantify upstream fish.

MATERIALS AND METHODS

Salmon Experiment and eDNA Capture

We placed 3, 10, 30, or 63 Atlantic salmons in sentinel cages in three streams in Southwest New Brunswick, Canada that contain little to no salmon (Jones et al., 2014): Dennis (45°15′13.32″ N, 67°16′2.639″ W), Waweigh (45°14′57.206″ N, 67°7′57.9″ W) and Digdeguash (45°23′6.151″ N, 67°8′52.065″ W). Atlantic salmon has not been detected during intensive electrofishing surveys conducted in the Dennis Stream and Waweig River in recent years, while typically a few have been caught each year in the Digdeguash River since 2009 (Graham Chafe, Atlantic Salmon Federation, personal communication). All three streams are shallow with rock bottoms and detailed physical conditions of each stream are presented in **Table 1**. The experiments were conducted in June (Dennis and Waweig: 3, 10, and 30 fish; Digdeguash: 10 and 63 fish) and in October (Dennis: 10 and 30 fish; Digdeguash: 10 fish). Each deployment consisted of adding a low abundance of fish (e.g., three fish) to the cage and collecting water samples 24 h after. Following water collection, fish were added to the cage to a higher desired abundance and water was again collected after 24 h. The cage (size: 4′ × 4′ × 2′) was made of 1/2″ hardware mesh with edge PVC pipe for reinforcement and a 2′ × 2′ opening panel on top. The cage was fixed to the

TABLE 1 | Environmental conditions within each stream during the experiment periods.

	Surveys		Biomass (g)	Water temperature (°C)	Total discharge (m^3/s)	Depth (m)	Velocity (m/s)
June	Dennis Stream	3 fish	105	13.4–17.7 (average = 15.16)	1.7	0.4	0.44
		10 fish	349				0.49
		30 fish	1,048				0.48
	Waweig River	3 fish	167	13.7–16.4 (average = 15.4)	0.4	0.4	0.15
		10 fish	555				0.31
		30 fish	1,665				0.27
	Digdeguash River	10 fish	610	18.2–20.6 (average = 19.9)	2.2	0.4	0.44
		63 fish	3,843				0.33
October	Dennis Stream	10 fish	1,497	12.4–12.6 (average = 12.5)	2.5	0.5	0.53
		30 fish	4,449				–
	Digdeguash River	10 fish	2,117	8.2–9.2 (average = 8.7)	2.9	0.5	0.40

stream bottom by adding rocks. Leaves and debris were retained by a 1/2″ hardware mesh installed about 20 m upstream of the cage. For each deployment and fish abundance, water samples (1 L) were collected from downstream to upstream from the cage at 1,600, 800, 400, 200, 100, 50, 5 m and approximately 50 m upstream. Additionally, Dennis Stream and Digdeguash River were also surveyed at 6,000 and 7,000 m, respectively. At each sampling distance, water samples were collected at each bankside and mid-stream for each lateral transects; a single 1 L sample was taken at each location (i.e., for a total of three samples per distance). To mitigate potential contaminations, 1 L Nalgene® bottles were shaken with 10% bleach solution (Clorox Javex® 12, 10.3% sodium hypochlorite) three times followed by five times with distilled water. We rinsed the bottles with river water three times prior to collecting the 1 L water samples to wash away any remaining bleach residue. For all surveys, sampling commenced at the most downstream location and moved upstream to reduce disturbance of eDNA that might be present in the riverbed. We collected surface water by fully submerging the bottles right below the surface while facing upstream. Nitrile gloves were worn and replaced between sampling sites and field controls to reduce contamination risks.

Water samples were kept on ice, and then at 4°C until filtration, which was performed in a dedicated filtration laboratory within 24 h of collection. All filtrations were done with 47 mm diameter 0.8 μm Whatman nylon membrane filters (GE Healthcare, IL, United States). Field blanks (tap water) were brought in the field during water sample collection and processed alongside stream samples for each sampling event. Lab filtration blanks, DNA extraction blanks and qPCR negative controls were also included during the processing and testing of samples. Furthermore, all reusable equipment (e.g., mason jars, forceps, and vacuum flasks) was soaked in a 1% bleach solution (i.e., 1 in 10 dilution prepared from 10% commercial bleach) for a minimum of 1 h.

Molecular Analyses

Atlantic Salmon qPCR Assay Design and Optimization

Atlantic salmon DNA barcode sequences from local specimens as well as sequences found in NCBI[1] and BOLD[2] were aligned in Geneious (version 9.1.4) along with DNA sequences from close relatives and other species with high nucleotide similarity (≥85%) and/or a similar geographic range. Specific primers [COI_82F_Ss (5′-TGGCGCCCTTCTGGGA-3′) and COI_276R_Ss (5′-AAGGAGGGAGGGAGAAGTCAAAAA-3′)] and probe [COI_194P_Ss (FAM – ATTAATTCCTCTTATAATCGGG – MGB)] were designed *in silico* to amplify a 195 base pairs (bp) region of the mitochondrial cytochrome c oxidase subunit 1 (CO1). To ensure species-specificity, the assay was designed with a high number of nucleotide polymorphisms between the targeted species and closely related and sympatric species. Primer-BLAST[3] was also used to ensure that primers were target-specific. The specificity of the qPCR assays was also tested *in vitro* using DNA from species that are closely related and/or potentially present in the studied environments : *Salmo trutta*, *Salvelinus fontinalis*, and *Morone saxatilis*. Serial genomic DNA dilutions were done to determine the efficiency [$E = -1 + 10^{(-1/slope)}$] and calculate the theoretical limit of detection (LOD) and limit of quantification (LOQ). Three serial dilutions from 10^0 to 10^{-8} were prepared and each serial dilution was tested in duplicate for a total of six qPCR threshold cycle (Ct) values. The theoretical LOD and LOQ was determined according to Klymus et al. (2020) using the discrete detection threshold approach. Non-target DNA normalized to 5 ng/μL was used as a background when preparing the serial dilutions to assess the efficiency of the assays under conditions similar to its prescribed usage.

DNA Extraction and Species-Specific qPCR Testing

DNA extraction from filters was conducted using half of each filter with the MN NucleoSpin Tissue Kit (Macherey-Nagel, PA, United States) following a modified protocol (LeBlanc et al., 2020). The resulting DNA extracts were stored at −20°C and the second half of the filter was kept as a back-up.

qPCR testing was done with the species-specific Atlantic salmon qPCR assay using the 2× TaqMan Gene Expression Kit (Thermo Fisher Scientific, MA, United States). Briefly, 3 μL of template DNA, 480 nM of each primer, 200 nM of the probe, 1 μL of 1% BSA, as well as 12.5 μL of master mix were used in 25 μL reactions. All qPCR tests were done in triplicate on a StepOnePlus qPCR platform (Thermo Fisher Scientific, MA, United States) using the following cycling parameters: initial hold at 50°C for 2 min, 95°C for 10 min, followed by 40 cycles at 95°C for 30 s, 60°C for 30 s and 72°C for 30 s, with fluorescence reading at the end of each elongation cycle. An exception to this was the samples from October, for which 50 qPCR cycles were used.

Sample Quality Control and Confirmation

To evaluate if PCR inhibitors were present in environmental samples, which could lead to potential false negative results, all samples (including blank controls) were spiked with an exogenous internal positive control (IPC) (linearized DNA plasmid containing a DNA sequence not found in the targeted environments) and tested using a qPCR assay specific to that IPC. Inhibition was considered present if a difference of more than 2 between the qPCR Ct of environmental samples and field blanks was observed. The IPC qPCR assay was done using the same parameters and reagents used for the species-specific Atlantic salmon qPCR assay.

To confirm the specificity of field results, sanger sequencing was performed on a subset of samples (6%). Briefly, PCR was performed using the optimized species-specific qPCR assay and the AmpliTaq Gold 360 PCR Master Mix (Thermo Fisher Scientific, MA, United States). PCR products were visualized on a 1.5% agarose gel followed by PCR product cleanup using ExoSAP-IT (Affymetrix, CA, United States) prior to being sent for Sanger sequencing at the Centre d'expertise et de services Génome Québec (Montréal, QC, Canada).

[1]https://www.ncbi.nlm.nih.gov/

[2]http://www.boldsystems.org/

[3]https://www.ncbi.nlm.nih.gov/tools/primer-blast/

The Atlantic salmon qPCR assay gave an efficiency of 95.9% and a theoretical LOD and LOQ of 0.36 pg (24 pg/L) and 2 pg (133 pg/L) of gDNA, respectively. The R^2 value of the assay was 0.992 and eDNA concentrations (pg/L) were calculated from the equation: $-3.4243[\log(x)] + 24.475$. Sanger sequencing on a subset of samples (6%), including some in October with Ct values >40 confirmed the assay as being specific to Atlantic salmon, with a few exceptions (mostly Ct > 44) which gave non-specific amplification; Ct values >40 have been kept to ensure capturing the full qualitative eDNA spatial distribution assessment.

Analyses

We conducted all analyses using R version 4.0.2 (R Core Team, 2020).

Data Availability

All data used in this study are available as **Supplementary Material**.

Exploring eDNA Spatial Distribution and Variation

First, we explored and described the spatial pattern of eDNA downstream of fish by examining the lateral and longitudinal distribution of eDNA. Second, we examined the effect of the water velocity and number of fish on spatial eDNA distribution and concentration. Finally, we examined the variability of eDNA detection and quantities across and within the three different streams.

Modeling eDNA Spatial Distribution and Variation

We built a model that reflected a combination of patterns in our data (**Figure 2**), trends from published studies, and first principle ecological assumptions, namely:

1. eDNA is more abundant when more fish are present (Pilliod et al., 2013; Doi et al., 2017; Wood et al., 2020)
2. eDNA decays or is lost as it moves downstream (Jerde et al., 2016; Wilcox et al., 2016; Laporte et al., 2020; Wood et al., 2020)
3. eDNA is less abundant when velocity is high (i.e., it is dissolved in a greater volume of water; Pilliod et al., 2014; Jane et al., 2015; Pont et al., 2018)
4. eDNA is initially entrained in the main flow of the stream, but disperses outward over time and distance toward the banks (hypothesized in Wood et al., 2020).

We tested these patterns and assumptions using likelihood ratio tests and relative likelihood (see below).

We modeled eDNA concentrations as an initial pulse in the center of the stream at the fish cage, with the pulse amount dependent on the number of fish. To account for lateral movement of eDNA within the stream (i.e., expansion of the "plume" perpendicular to the direction of water flow, **Figure 1**), we allowed the proportion of eDNA that was found at the banks to increase from 0 at the fish cage to a fixed (equilibrium) proportion moving downstream from fish. eDNA also was allowed to decay at midstream- and bankside-specific rates. eDNA detection was treated as a logistic function of predicted eDNA concentration. The following six equations represent sequential steps in one model that maximized the probability of obtaining our eDNA detection and quantity data (i.e., a maximum-likelihood model).

For the first step of the model, we modeled the proportion of eDNA found in the two bankside samples [$P_B(x)$]. We assumed lateral dispersal (square-root) dynamics (Wetzel, 2001), with an asymptote at a given proportion (P_{max}):

$$P_B(x) = P_{max}(1 - \exp(-\gamma\sqrt{x})) \quad (1)$$

P_{max} is the maximum proportion of eDNA found at the banks; γ is a lateral dispersal rate coefficient. For example, if $P_{max} = 0.4$, then 40% of eDNA will be found in the bankside samples (rather than the midstream samples) far downstream from fish.

For the second step of the model, we modeled the total amount of eDNA [$Q(x)$] across two bankside samples and one midstream sample at downstream distance x, incorporating midstream and bankside eDNA loss rates (r_m and r_b, respectively), number of fish (F), and velocity (V). Midstream- and bankside-specific decay rates were weighted by the proportion of eDNA found at each [$P_B(x)$; Eq. 1] and followed exponential decay over distance (x). eDNA quantities were modeled as proportional to the number of upstream fish (F), and inversely proportional to velocity (V):

$$Q(x) = \frac{\beta_0 F(P_B(x)(1 - r_b) + (1 - P_B(x))(1 - r_m))^x}{V^q} \quad (2)$$

β_0 is a stream-specific coefficient; F = the number of fish; $P_B(x)$ = proportion of eDNA in bankside samples (Eq. 1); r_b and r_m = eDNA decay rates at the banks and midstream, respectively; x = distance downstream; V = velocity; and q is a velocity scaling coefficient. We modeled β_0 specifically for each stream to examine the variation in β_0 for later analyses (see below).

For the third step of the model, we separated the quantity of eDNA found at a given distance [$Q(x)$] into its midstream [$M(x)$] and bankside components [$B(x)$] according to the proportion of eDNA expected at the midstream and banks [$P_B(x)$; Eq. 1]:

$$M(x) = Q(x)(1 - P_B(x)) \quad (3)$$

$$B(x) = \frac{P_B(x)\,Q(x)}{2} \quad (4)$$

For the fourth step of the model, we calculated eDNA detection rates [$R_M(x)$ and $R_B(x)$ for midstream and bankside samples, respectively] from eDNA quantities. We assumed that eDNA detection was logistically related to eDNA quantity (Klymus et al., 2020), i.e., low quantities of eDNA led to low detection rates and large quantities of eDNA led to high detection rates:

$$R_M(x) = \frac{e^{\alpha_0} M(x)^{\alpha_1}}{e^{\alpha_0} M(x)^{\alpha_1} + 1} \quad (5)$$

$$R_B(x) = \frac{e^{\alpha_0} B(x)^{\alpha_1}}{e^{\alpha_0} B(x)^{\alpha_1} + 1} \quad (6)$$

α_0 and α_1 are rate and shape parameters, respectively, for the detection rate ~ eDNA quantity relationship.

Finally, we calculated the likelihood of obtaining our eDNA detection and quantity data given the above model. We then

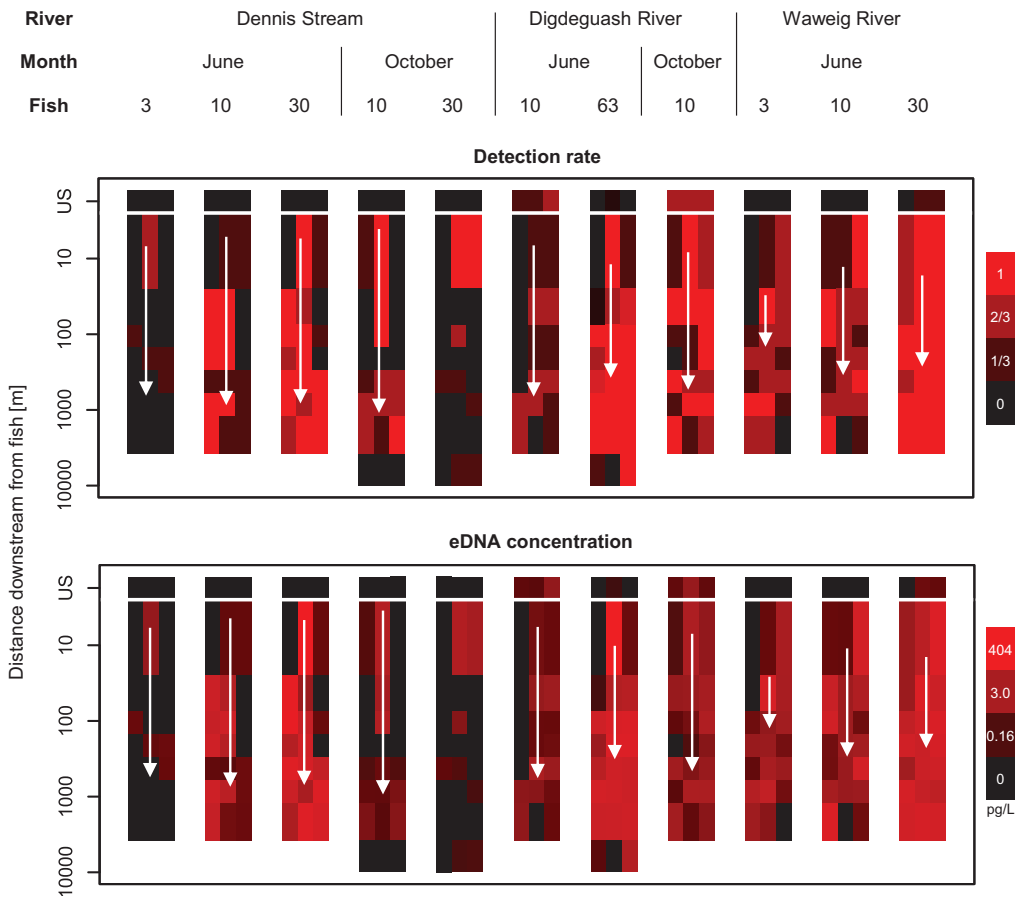


FIGURE 2 | Salmon eDNA detection rates and quantities for bankside and midstream samples in three New Brunswick streams. White arrows indicate average stream velocity, which ranged from 0.15 to 0.53 m/s. Velocity data was not available for the 30 fish treatment in Dennis Stream in October. Detection rate is defined as the number of qPCR replicates where salmon DNA was detected. US, upstream control samples (see section "Materials and Methods").

optimized the model, picking the set of parameters that maximized the likelihood (i.e., minimized the negative log likelihood) using the *mle* function in the *stats4* package, included in base R (R Core Team, 2020).

Model Testing – Parameters

We statistically tested the earlier described patterns and assumptions underlying the model. We tested the significance of several terms of interest (velocity, eDNA decay, inter-stream variation, and month-to-month variation) within the models using Type II likelihood ratio tests. We also tested the significance and precision of the eDNA quantity-detection rate relationship (i.e., Eqs 5 and 6) using a likelihood ratio test and a receiver operating curve (ROC).

Model Testing – Plume

To test for the presence of an eDNA plume (i.e., lateral diffusion of eDNA moving downstream from fish), we compared our model to several alternative models without lateral diffusion. These models are summarized in **Supplementary Table 1**. All alternative models have a fixed proportion of midstream versus bankside eDNA over all downstream distances from fish. The first alternative model assumes that eDNA is lost at a constant rate as it moves downstream from fish. The second assumes a lag period during which eDNA becomes easier to detect farther from fish, then is lost at a constant rate as it moves further downstream. The third assumes no changes in eDNA concentration moving downstream from fish.

As the four above models were not nested versions of each-other, we could not conduct likelihood ratio tests (Burnham and Anderson, 2003). Instead, we compared models using relative likelihood (Burnham and Anderson, 2004). We also calculated R^2 for all models for both our eDNA detection and quantity data. As we only sampled downstream >1,600 m for a small set of stream/fish/date combinations, we repeated the above analyses with only data for distances ≤1,600 m to remove the potential for bias.

Model Testing – Bankside eDNA Accumulation

We also examined the asymptotic proportion of eDNA found in the bankside samples far downstream from fish (P_{max}), as this parameter has important implications for accumulation or loss of eDNA near stream banks. We refit our model with P_{max} fixed at the range of values from 0 to 1 and examined the model

AIC. Again, as we only sampled downstream >1,600 m for a small set of stream/fish/date combinations, we repeated the above analyses with only data for distances ≤1,600 m to remove the potential for bias.

Optimal eDNA Sampling

We calculated detection power for two management scenarios: sampling downstream from a target reach (i.e., known or hypothesized fish location), and uniform sampling to detect fish in an unknown location. Significant variation exists across streams in eDNA detectability, even for the same number of fish and same volume of water (see section "Results"), likely due to the stochastic dispersion from stream-specific hydrodynamics and differences in stream morphology, as well as water chemistry (Dejean et al., 2011; Barnes et al., 2014; Deiner and Altermatt, 2014; Jerde et al., 2016; Shogren et al., 2016, 2018; Klymus et al., 2020). This variation is reflected in the stream- and date-specific β_0 parameter in Eq. 1. Thus, power analyses for eDNA detection must be sensitive to this variation, and generate predictions that work for *most* streams, rather than just the average stream. We used the distribution of the stream- and date-specific β_0 parameter to generate a "low detection" β_0 value corresponding to 5% quantile for the β_0 parameter distribution using the *quantile* function in R. As 95% of streams are expected to have higher detection than reflected in this "low detection" β_0 value, we expect the below sampling analyses to be a conservative estimate for nearly all streams similar in character to our study systems.

Sampling downstream from a target reach

We examined the power to detect fish in a single reach, i.e., where fish presence was known, hypothesized, or of potential management concern. We calculated the number of samples required for positive detection (S) for different numbers of fish (3, 10, and 30), downstream distance (1–10 km), velocity (0.15, 0.35, and 0.55 m/s), and bankside/midstream sampling:

$$S = \log_{1-R} A \tag{7}$$

R is detection rate (Eqs 5 and 6) and A is desired power (i.e., $A = 0.05$ for 95% chance of detection). We used model parameter values from the maximum likelihood model fitting described above, with the exception of β_0, for which we used the "low detection" value calculated above.

Constant-interval sampling to detect fish in an unknown location

We also tested the power of constant-interval sampling to detect fish presence anywhere in a stream, when potential fish location is unknown. This is the case for sampling studies looking to broadly describe species' distributions, in which simple presence/absence is sought in numerous streams. In this case, constant-interval sampling over the course of a stream is one method to determine whether fish are present (Wood et al., 2020). Following Wood et al. (2020), we simulated varying numbers of fish (1, 3, and 10) at a single, random location in a 100 km stream with different water velocities (0.15, 0.35, and 0.55 m/s). We then assumed a constant-interval (10–2,000 m), midstream sampling effort with three technical replicates. We examined fish detection rate with 10 replicate simulations per each fish, velocity, and sampling interval combination.

RESULTS

Environmental DNA concentrations downstream from fish were highly variable across and within streams, dates, and numbers of fish, ranging from no detections to 731 pg/L. The mean eDNA concentration for all samples with a positive detection was 25.6 pg/L and the lowest concentration detected was 0.01 pg/L. eDNA was detected in 56% of our samples, including 22% of samples taken 6 or 7 km downstream from caged salmon.

Similar to our expectation, a low concentration of eDNA was detected upstream of the cage deployment site in the Digdeguash (30.6% positive detection in the upstream samples, mean: 3.29 pg/L) and Waweig Rivers (7.4% positive detection in the upstream samples, mean: 3.78 pg/L). The various field and laboratory filtration blanks, as well as DNA extraction and qPCR negative controls included throughout this work showed minimal potential cross-contaminations, with only 1 out of 15 field blank (eDNA concentration = 4.19 pg/L) and 1 out of 28 DNA extraction blank (eDNA concentration = 2.1 pg/L) positive for salmon DNA.

eDNA Spatial Distribution and Variation

Environmental DNA concentrations and detections were highest midstream and lowest bankside shortly downstream from fish (**Figures 2**, **3**). Up to roughly 100 m downstream from fish, eDNA quantities and detection rates remained higher midstream, but became increasingly abundant bankside (**Figures 2–4**). Between 100 and 1,000 m, we observed roughly equal eDNA detection rates and quantities in midstream and bankside samples, followed by higher eDNA detection rates and quantities at banksides located >>1,000 m from fish (**Figures 2**, **3**). There was, however, significant variation within streams that added noise to—and sometimes masked—these general patterns. eDNA concentration variation was highest midstream closest to fish (5 m downstream), and lowest bankside closest to fish (**Figure 5**)—both for untransformed variation and variation of ln-transformed data, which removes right skew and assumes variance is proportional to the mean. Thus, the eDNA spatial distribution exhibited unique midstream and bankside patterns with increasing downstream distance from fish. Midstream, eDNA had initially high abundances, with relatively constant loss moving downstream (**Figure 6**). On the other hand, bankside eDNA was nearly undetectable close to the fish, but gradually rose in abundance, then fell again with increasing downstream distance from fish (**Figure 6** and **Supplementary Figure 1**). Thus, eDNA data showed a steady dispersal of eDNA from the midstream toward the banks over distance (**Figures 1, 3** and **Table 2**), which resulted in the eventual "buildup" of eDNA on the banks >1,000 m downstream from the fish (**Figures 2**, **3**).

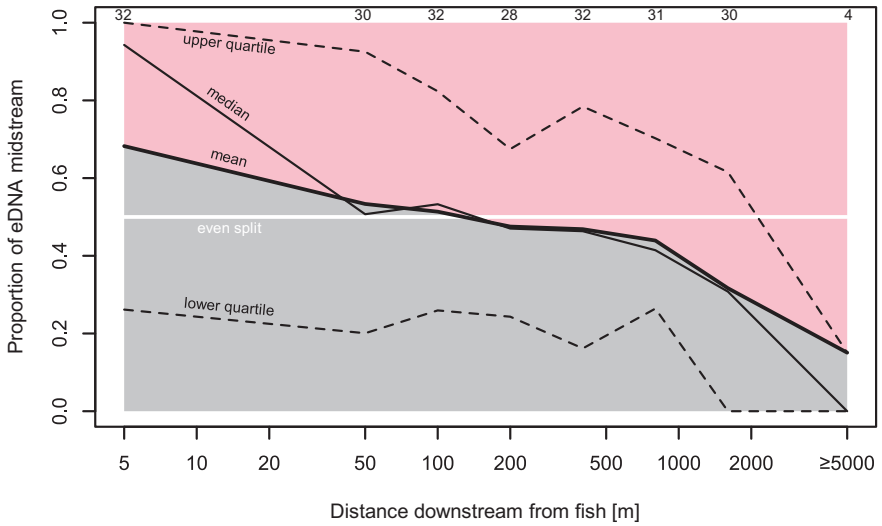


FIGURE 3 | Proportion of bankside versus midstream eDNA. Proportions here are calculated as the midstream eDNA quantity over the midstream eDNA quantity plus the average of both bankside quantities. All sampling transects in which at least one sample had positive detection are included here. Top numbers indicate number of transects included in proportion estimate.

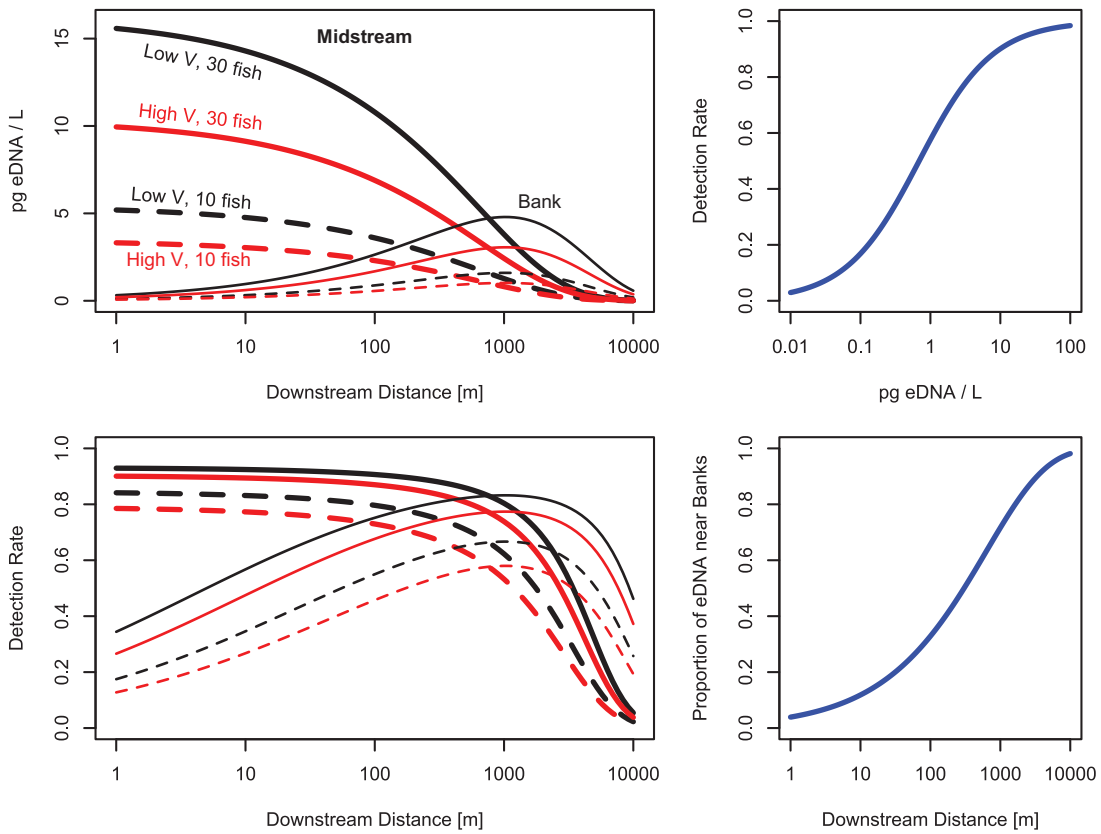


FIGURE 4 | Modeled salmon eDNA quantities and detection rates downstream from fish. **Top left:** eDNA is most abundant near the fish, then disperses to the banks, where eDNA is most abundant roughly 1,000 m downstream from the fish. Values shown are predicted values for an average stream. **Top right:** after roughly 1,000 m, more eDNA can be found in bankside samples than midstream samples. **Bottom left:** Midstream samples have higher detection rates near the fish, while bankside samples have higher detection rates roughly 1,000 m downstream from the fish. **Bottom right:** eDNA detection rates are greater than 50% when predicted eDNA concentrations are >1 pg/L; eDNA detection rates are roughly 90% when predicted eDNA concentrations are >10 pg/L.

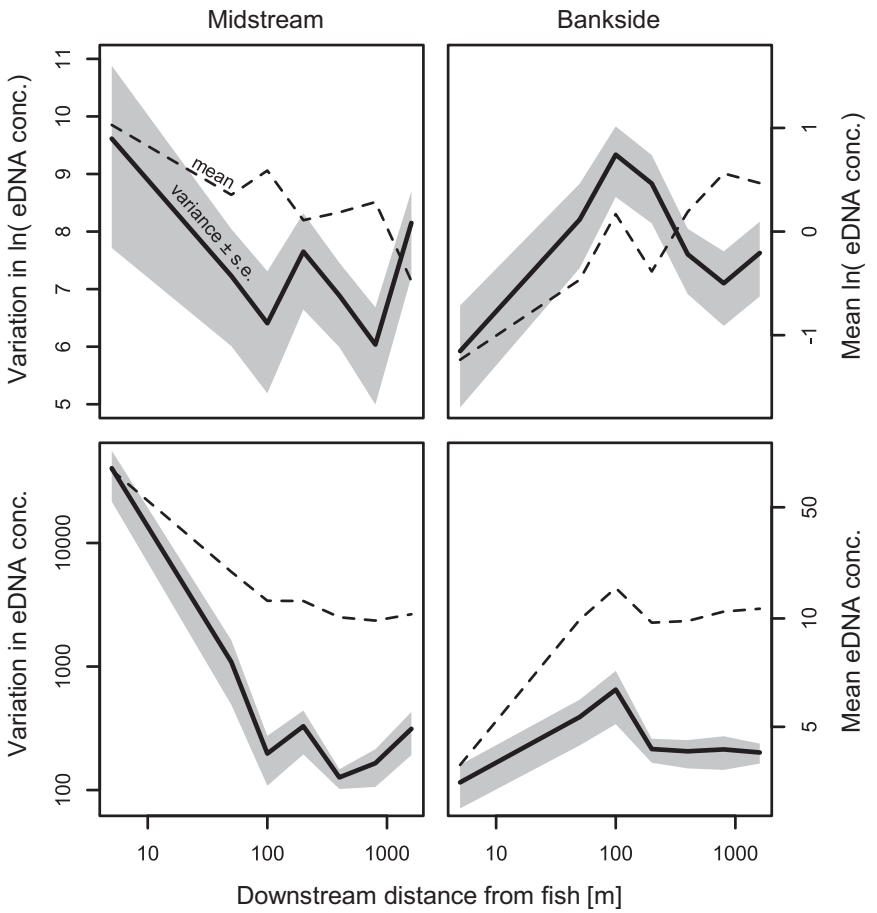


FIGURE 5 | Variation and mean eDNA abundance. Distance-specific variances for midstream and banks samples shown, pooled across all streams, dates, and fish abundances for each distance. Distances >1,600 m, which were not sampled for all streams, were excluded from this figure. The variance of ln-transformed data reduces the potential effect of variance scaling with the mean [i.e., ln($\mu\varepsilon$) becomes ln(μ) + ln(ε)] and is less sensitive to outliers.

Model Testing – Parameters

There was significant stream-level variation and temporal variation in eDNA detection and quantity (i.e., variation in β_0) that was not related to velocity or number of fish (**Table 3**). eDNA detection rates and quantities were significantly lower at higher velocity (**Table 3**). Our model showed significant eDNA decay or loss as it moved downstream; this loss occurred more rapidly at the banks according to our model, though this difference was not statistically significant (**Table 3**). Finally, our model showed a significant relationship between eDNA concentration and detection rate, with a relatively strong ability to discriminate between positive and failed detections (**Supplementary Figure 2**).

Model Testing – Plume

Our model strongly outperformed the three alternative models—all of which assumed no lateral dispersal of eDNA (i.e., no plume)—in both relative likelihood and R^2 (**Table 4**). The same pattern was apparent even when sampling distances >1,600 m were excluded from our analyses (**Supplementary Table 2**). Thus, lateral dispersal, rather than differential midstream/bankside eDNA decay rates, appeared to drive the differences in midstream and bankside eDNA dynamics in our model.

Model Testing – Bankside eDNA Accumulation

The optimal (in terms of AIC) value for P_{max}—the asymptotic proportion of eDNA found in bankside samples downstream from fish—was approximately 1.00 when all data were included and 0.95 when only distances ≤1,600 were included in our analyses (**Supplementary Figure 3**). Decreasing P_{max} to 0.67—the value expected if eDNA was evenly dispersed across one midstream and two bankside samples far downstream from fish—increased the model AIC >> 5 in both cases.

Optimal eDNA Sampling

We found that optimal sampling for eDNA is near the source for midstream samples, and roughly between 100 and 1,000 m downstream for bankside samples (**Supplementary Figure 4**). However, our midstream results are slightly misleading without context, as our midstream sampling was conducted with the

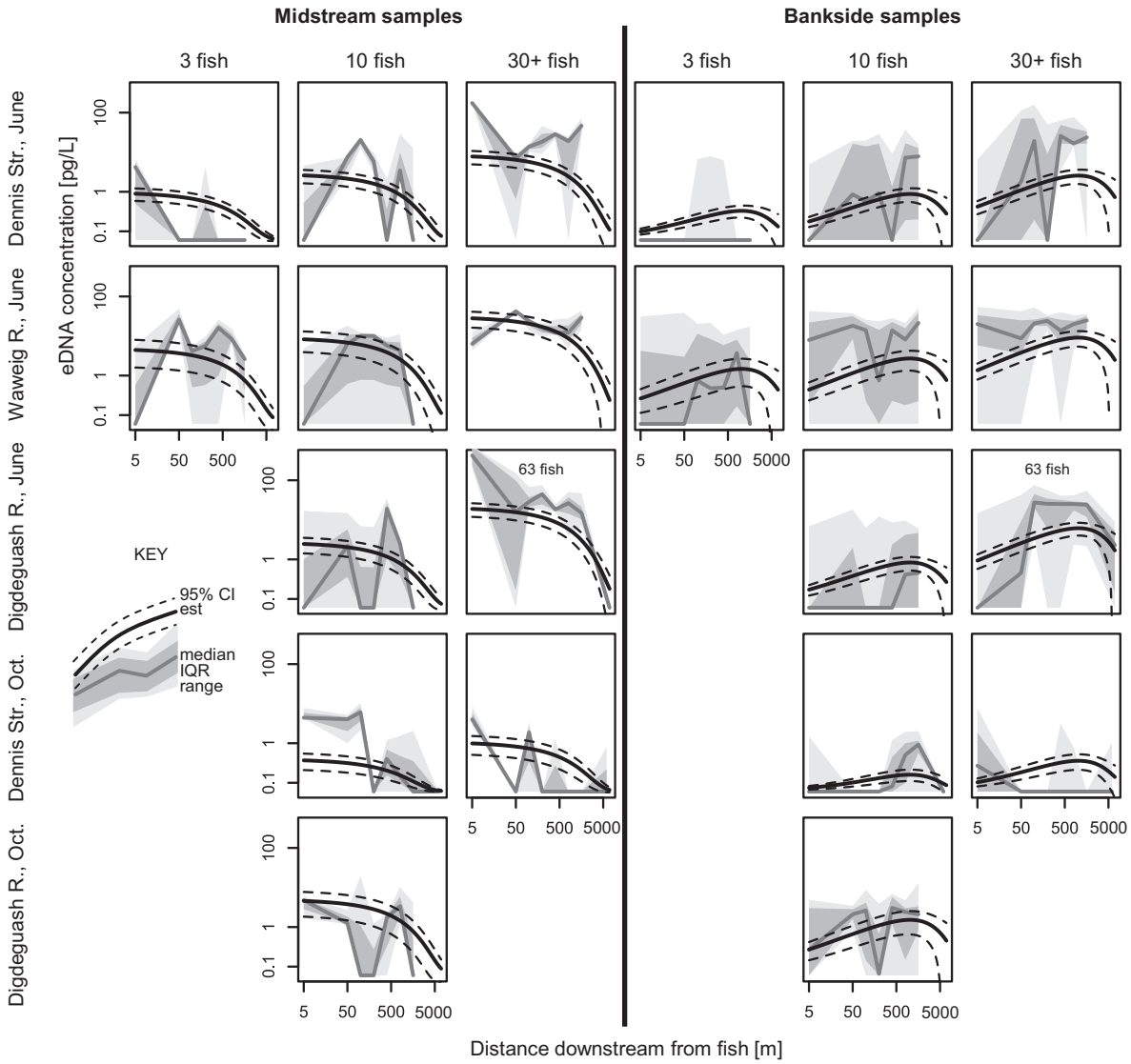


FIGURE 6 | eDNA concentration versus downstream distance from fish, separated for midstream and bankside samples. Data distributions, model estimates, and model estimates confidence intervals (CI) are shown. IQR, inter quartile range. For analogous eDNA detection rate data, see **Supplementary Figure 1**.

knowledge that our salmon were in the very middle of the stream. In reality, midstream sampling may have less optimal results shortly downstream from fish when the lateral positioning of the fish is not known. The optimal sampling distance that gave relatively high detection rates for *both* midstream and bankside sampling was roughly between 100 and 1,000 m downstream from fish.

Based on our constant-interval eDNA sampling power analysis, sampling midstream every 100 m (with three replicates) was sufficient to detect a single fish (>95% power) under high-velocity conditions in a low-detection (5% quantile β_0) scenario (**Supplementary Figure 5**). At low velocity, this sampling distance increases to about 400 m. With 10 fish present, sufficient sample spacing increases to 500 to > 1,000 m, depending on velocity (**Supplementary Figure 5**).

DISCUSSION

Consistent with other studies, our results indicate that eDNA is a reliable indirect approach for determining organism presence in streams. As expected, concentrations of eDNA increased predictably with the number of upstream fish (**Figures 2**, **6**). However, estimating fish distribution and abundance precisely require accounting for distance and flow. Our results demonstrate that a better understanding of spatial patterns of eDNA concentration, variation, and distribution is important for optimizing eDNA detection and essential to assessing quantitative population distribution in lotic environments. We examined the spatial dynamics of eDNA as it moves downstream from fish and found evidence of an eDNA plume—eDNA exhibiting a pattern initially concentrated in the

TABLE 2 | Parameter estimates and standard errors for eDNA detection rate and concentration models.

Parameter	Description	Estimate	Standard error
β_{DSJ}	Coefficient for Dennis Stream in June	0.16	0.05
β_{DSO}	Coefficient for Dennis Stream in October	0.021	0.007
β_{DRJ}	Coefficient for Digdeguash River in June	0.14	0.05
β_{DRO}	Coefficient for Digdeguash River in October	0.25	0.10
β_{WRJ}	Coefficient for Waweig River in June	0.37	0.20
r_b	Bankside-specific eDNA decay rate	0.00027	0.00007
r_m	Midstream-specific eDNA decay rate	0.00000	0.00002
γ	Lateral eDNA dispersal rate	0.040	0.004
q	Velocity scaling power	0.76	0.36
P_{max}	Theoretical maximum eDNA proportion at banks downstream from fish	1.00	0.01
α_0	Intercept for eDNA quantity to detection rate conversion logistic model	0.31	0.11
α_1	Slope for eDNA quantity to detection rate conversion logistic model	0.83	0.08

See Eqs 1–6 for parameter definitions.

midstream but widening as it travels downstream from fish (**Figure 1**). These dynamics were clear even despite a low level of upstream eDNA contamination in a few instances (see section "Results"). The detection rate and quantity of eDNA varied widely across rivers and sampling periods, but varied predictably with the number of upstream fish and velocity. Below, we discuss how physical factors can be gathered to predict quantitative population distribution in a single and multiple inhabited reaches based on the eDNA plume.

Characterizing the eDNA Plume

Results support that eDNA spreads from fish in the form of a plume beginning as concentrated large particles (e.g., tissue fragments and cells) close to the fish (Wilcox et al., 2015). Due to its state, eDNA concentrations near the source are highly variable with high upper concentration limits in the midstream (**Figures 2**, **5**). eDNA is then more evenly dispersed over distance due to the "breakout phase" processes wherein particle fragmentation and mixing result in smaller, more evenly distributed particles (**Figure 3**). In the breakout phase, eDNA becomes more equally abundant in midstream and bankside samples. Beyond this breakout zone, our results support that there is a more steady decrease in detection due to DNA degradation, dilution, and settlement (Barnes et al., 2014; Jerde et al., 2016; Shogren et al., 2017), but again this differs between midstream and bankside regions. Past the point of roughly equal eDNA in the midstream and bankside regions, eDNA persists at higher concentrations near the banks while eDNA drops off more in the midstream (**Figures 2**, **3**).

TABLE 3 | Likelihood ratio tests for model parameters.

Test	χ^2	df	*p*
Effect of month	87.35	2	< 0.001
Effect of stream	71.23	3	< 0.001
Effect of velocity	4.50	1	0.034
Effect of decay rates	14.4	2	< 0.001
Effect of different midstream/bankside decay rates	1.47	1	0.23
Effect of eDNA quantity on detection rate	221.92	1	< 0.001

Similar to fine particles, eDNA tends to accumulate in stream margins where the velocity is low. The shape of the plume process is likely to be driven by the hydrological water-bank interface. Based on fluid dynamic theory, stream velocity is greatest in the midstream near the surface and is slowest along the stream bed and banks due to friction. Faster flow tends to be turbulent, while slower flow tends to be laminar. Turbulent flow is more effective than laminar flow at keeping particles in suspension. Studies report a high degree of fine-particle retention within the streambed and banks (Skalak and Pizzuto, 2010 and Harvey et al., 2012). Banks can thus act as "sponges," catching and accumulating eDNA on sediments, complex debris, and in eddies while midstream eDNA is continually flushed out of the system or stochastically dispersed laterally. Other work has demonstrated net movement of organic matter out of the water column and into superficial sediments (Minshall et al., 2000), leading to higher eDNA concentrations in and near sediments (Turner et al., 2015). Thus, we hypothesize that bankside samples sufficiently downstream from fish have higher concentration than their respective midstream samples due to proximity to eDNA-rich sediments and eddies. Particularly, these observations suggest that eDNA does not behave strictly as "wash load" and instead depict the combination of downstream transport and transient retention influenced by stream geomorphology (Drummond et al., 2017; Phillips et al., 2019).

Overall, we observed the predictable pattern of decreasing eDNA concentration and detection rate with increasing velocity—a proxy for flow rate (**Figure 4** and **Table 3**). Similar to results from particle models (Andruszkiewicz et al., 2019), increasing the volume into which eDNA is diluted is at least partially responsible for this pattern. However, it remains to be seen whether increasing velocity—and thereby flow rate—leads to a greater transport distance for eDNA, shifting or stretching our eDNA detection curve longitudinally by modifying r_b, r_m, γ, or p_{max} (Pilliod et al., 2014; Pont et al., 2018). In midstream nearby the source (50–240 m), Jane et al. (2015) observed that,

TABLE 4 | Comparison of four eDNA dispersal models.

Model		Assumptions	df	AIC	Relative likelihood (%)	R^2	
						Detection	Abundance
Plume	eDNA proportion at banks increases over distance	13	4,332.41	>99.9	0.26	0.35
Universal decay	eDNA proportion at banks is fixed over distance	12	4,424.20	<0.1	0.22	0.32
Universal decay with lag	eDNA proportion at banks is fixed over distance; eDNA exhibits a lag phase in which detection is lower near fish	13	4,400.32	<0.1	0.23	0.32
Constant	eDNA quantities and proportion at banks are fixed over distance	10	4,427.65	<0.1	0.23	0.32

The first model, our plume model, assumes that the proportion of eDNA at banks can change (increase) over distance. The next two models do not have this assumption, and assume either decreasing or increasing then decreasing patterns of eDNA over distance. The final model assumes eDNA is constant throughout the stream. We did not conduct likelihood ratio tests for models, as models were not nested. For the line figures, thick lines represent midstream and thin lines represent bankside samples. Note that all models contain terms for stream, date, number of fish, midstream versus bankside, and water velocity regardless of eDNA dispersal model.

despite high variation between ecosystems, eDNA abundance was highest close to the source and quickly trailed off over distance at the lowest flows, whereas eDNA was relatively low both near and far from the source at the highest flows. Here, the relatively small number of distance sampling points, combined with our relatively high detection rates at 6 and 7 km downstream from fish, make this question better addressed by future studies.

Optimal Sampling

Determining the optimal eDNA sampling strategy depends on correctly and accurately quantifying the eDNA plume. When water samples are collected only a few meters downstream from the target organism, more replicates, water volume, or pooling water samples might be needed to overcome the high variation seen in samples near the target organisms (see section "eDNA Spatial Distribution and Variation"; also see Wood et al., 2020). However, as the lateral positioning of the target organism is generally unknown, many samples taken a few meters downstream from a target organism or reach are likely to miss the plume and have low eDNA concentrations or detection rates (as in the bankside samples, **Figures 2**, **5**). Therefore, while sampling in the plume immediately downstream from the fish would technically yield the highest quantity of eDNA, chances of successfully sampling in the plume close to the fish are lower. In the systems surveyed in this study, the highest eDNA detection probabilities and mean concentrations *across all lateral sampling positions* occurred between 100 and 1,000 m downstream of the source. Importantly, we do not recommend single-bankside eDNA sampling, due to the tendency of one bankside to be consistently biased compared to the other (see **Figure 2**). Instead, we recommend midstream sampling or combined sampling of *both* banks. The importance of field replicates is crucial in all locations (see **Figure 6**).

For studies that simply seek system-wide detection of rare taxa, even-interval eDNA sampling is a potential cost- and labor-saving method. Our simulations indicated that sampling every 100 m gives a 95% chance of detecting even a single fish in high velocity (0.55 m/s) conditions in nearly all streams of similar character to the study streams. Increasing the number of fish or decreasing velocity allows for significantly less sampling—to about 400 m for our low velocity (0.15) scenario or about 600 m for 10 fish (see **Supplementary Figure 4**). These estimates are likely conservative estimates, as our experiments—despite sampling to 7,000 m in some streams—were unable to find a downstream LOD. Thus, these estimates are slightly influenced by extrapolation from our model, which assumed that midstream eDNA detection was virtually impossible after 10,000 m.

Quantitative Population Distribution Assessment Using eDNA

Excitingly, the predictable patterns of midstream versus bankside eDNA transport should in principle allow one to estimate upstream fish distance and number using midstream and bankside water sampling, e.g., presented in **Figure 7**. In other words, eDNA heterogeneity across a stream channel can facilitate distributional assessment, the third operational goal for eDNA that has lagged detection and abundance quantification. Models can be solved for the proportion of eDNA in midstream samples and the total amount of eDNA in midstream and bankside samples to predict location and number of upstream fish. **Box 1** presents two different scenarios: a single inhabited upstream reach, and numerous inhabited upstream reaches. The first scenario is relatively straightforward, as it simply requires solving

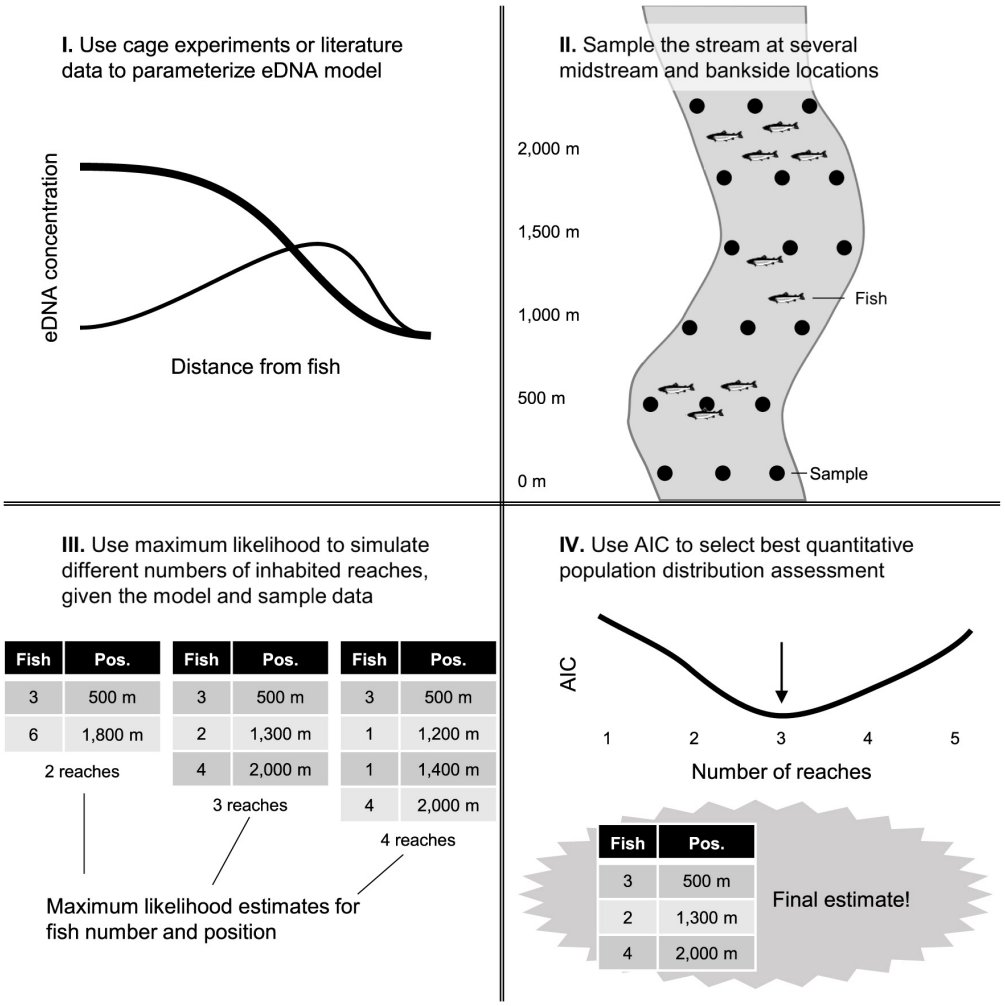


FIGURE 7 | Workflow for eDNA-based quantitative population distribution assessment in streams. Once a sufficient model of eDNA dynamics is constructed for a particular stream or stream type, sampling and maximum likelihood simulations can give estimates of fish numbers and locations.

of our models here for different variables. The second scenario is more realistic, but requires a layer of simulation and model fitting. Both scenarios assume fish are in midstream and in a similar type of stream as those selected in our study. We encourage stimulating datasets to calibrate and improve the performance of this predictive population model. Estimates of fish location can be confirmed with the expected spikes in midstream eDNA variation (**Figure 5**).

Next Steps—Environmental Covariates

The predictive models developed here are based on a large range of fish abundances and environmental conditions, at least within this study system of small shallow streams with rocky bottoms. However, here models assume that fish size is relatively constant and much variation in eDNA concentrations within and across streams and seasons remains—despite the general lack of PCR inhibition in our samples (**Figure 6**). This variation reflects complex interactions between the environment, organisms, and eDNA. More studies are needed to determine the probability that eDNA particles will go back to the mid-channel when the lateral positioning of the fish is not known, in order to decide where best to sample, i.e., mid-stream versus bank sides. There is an extensively growing literature about the effect of environmental conditions on eDNA dynamics, but a significant amount of effort is still needed to understand and account for these complex interactions. Quantifying the environmental parameters altering the eDNA breakout phase and the spatio-temporal variation of the eDNA plume should further improve the population predictive models included in this article—in particular by removing the need for a stream-specific parameter (β_0, Eq. 1) in our models, instead replacing it with a universal parameter and environmental covariates. To efficiently extrapolate our predictive model to all types of Atlantic salmon habitats, it will also be important to test and calibrate a tridimensional model in larger rivers where sinking, settlement, and resuspension processes can have a significant effect on eDNA. Ideally, a next step in integrating eDNA into salmon management will be a salmon population distribution model based on the combined effect of eDNA dilution rate, persistence, life stage parameters, and environmental covariates.

BOX 1 | A framework to infer a population estimate model based on the spatial eDNA concentration in lotic environments.

Single inhabited reach case

Assuming two bankside samples (B_1 and B_2) and one midstream sample (M), the proportion of eDNA at the bankside is obtained as:

$$P_B = \frac{B_1 + B_2}{B_1 + B_2 + M} \tag{8}$$

Then we can solve Eq. 2 (see section "Materials and Methods") for distance, x, assuming we have already estimated the parameters P_{max} and γ, the maximum eDNA proportion at banks and the lateral eDNA transport diffusion rate, respectively:

$$x = \frac{\ln^2\left(\frac{P_{max} - P_B}{P_{max}}\right)}{\gamma^2} \tag{9}$$

Now that we have distance, x, we can solve Eq. 1 for F, assuming we have already estimated the parameters q, β_0, r_b, and r_m, the velocity scaling power, stream-specific coefficient, and bankside- and midstream-specific decay rates, respectively:

$$F = \frac{QV^q}{\beta_0}\left(P_B\left(1 - r_b\right)\left(1 - P_B\right)\left(1 - r_m\right)\right)^x \tag{10}$$

Now we have successfully calculated estimates of F and x for the studied fish size—the number of fish and the distance upstream from sampling. A single stream location for such sampling is apt to result in considerable positional and abundance estimation error. However, this can be improved upon by averaging or model fitting the respective estimates of F and x for multiple cross-stream sampling locations would permit more precise, consensus estimates of where and how many fish are present.

Multiple inhabited reaches case

When there are multiple potential inhabited reaches, each eDNA sample becomes an unequally-weighted picture of *all* fish upstream of the sampling site. The simulation of a given number of reaches i, each with its own number of fish F_i can then be calculated from the expected amount of eDNA for each sample, j, using Eq. 1 for Q:

$$\widehat{Q}_j = \sum_{i}^{x_{ij}>0} Q(x_{ij}, F_i) \tag{11}$$

We can then estimate the values for x_i, F_i, and a standard deviation parameter s that maximize the likelihood, L (minimize the negative log likelihood) of generating our sample data, i.e., by maximum-likelihood model fitting:

$$-\ln\left(\hat{L}\right) = \min\left(\sum_j -\ln\left(p_N\left\{0, s^2\right\}\left(ln\left(Q_j\right) - ln\left(\widehat{Q}_j\right)\right)\right)\right) \tag{12}$$

Where p_N is the probability density estimate of a normal distribution.

The AIC is then obtained from:

$$AIC = 4i + 2 - 2\ln\left(\hat{L}\right) \tag{13}$$

If we plot AIC versus i (the number of simulated reaches), an upside-down hump-shaped curve is obtained, as adding additional values for x_i and F_i (reach distance upstream and number of fish in that reach, respectively) gives us diminishing marginal returns that are penalized when calculating AIC (via the $4i$ term). Thus, the maximum-likelihood estimates for x_i and F_i associated with the minimum AIC are the most reliable estimates for fish locations and abundances. This is a quantitative population distribution assessment (**Figure 7**). Note that numerous eDNA samples will need to be taken at various longitudinal locations within a stream to generate the necessary power to fit multiple x_i and F_i terms with any accuracy.

Such a model will be crucial to expand eDNA monitoring efforts to new streams.

CONCLUSION

Despite significant room for model development, eDNA is a powerful and growing tool for the conservation of riverine fish species. An essential component to using river eDNA to enumerate upstream stocks is an understanding of eDNA plume dynamics and its variation. As we have shown, plume dynamics may be leveraged to develop quantitative population distribution assessments in lotic environments that feed into management decision making. This spatial capacity would be a significant refinement to eDNA sampling which is currently largely used for detection and coarse abundance estimation. Further inclusion of source-specific eDNA release rates (e.g., fish life stages and metabolism) and environmental covariates will hopefully reduce much of the unexplained variability in eDNA data and generate eDNA models that are robust across study systems and management needs.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The "animal study was reviewed and approved by the Animal Use Protocol, Fisheries and Oceans Canada, following values and standards promoted by the Canadian Council on Animal Care".

AUTHOR CONTRIBUTIONS

ALR, MT, NG, FL, and MK conceived the study and contributed the resources. The sampling design has been developed in close consultation between all co-authors. CB, ALR, NG, and FL contributed to field coordination/data collection. FL and NG developed and led the laboratory analyses. ZW analyzed the data. ZW and ALR wrote the manuscript, while FL, NG, MT, and MK helped to draft and improve the manuscript. All authors edited the manuscript and approved the final version.

FUNDING

This work was supported by the Strategic Program for Ecosystem-Based Research and Advice (SPERA) and Genomics Research and Development Initiative Phase VII (GRDI) from Fisheries and Oceans Canada (DFO), the Maine Agricultural and Experiment Station, the University of Maine Janet Waldron Doctoral Research Fellowship, and the Maine-eDNA EPSCoR program (NSF EPSCoR: 11A-1849227).

ACKNOWLEDGMENTS

We especially thank Royce Steeves and Brad Erdman for their help to design this study. We thank the staff of the research group at the St Andrews Biological Stations (DFO). We also thank Ann Kinney, Steve Leadbeater, Matthew Black, Craig Smith, and David Needler for supporting the fish care and suggestions for animal health and wellbeing. Steve Neil, Vicky Merritt, John Reid, and Dheeraj Busawon helped us put in place the adequate animal care procedure. For the sample collection and technical support we thank Simon Clarke, Ann Kinney, Royce Steeves, Steve Neil, Kristine Gagnon, Michaela Harris, Chantal Gautreau, and Travis Melanson. Environmental factors could be monitored due to Andrew Cooper's help and support. Brent Wilson and Steve Leadbeater secured juvenile salmon from Cape d'Or Salmon (Nova Scotia, Canada).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.650717/full#supplementary-material

Supplementary Figure 1 | eDNA detection rate versus downstream distance from fish, separated for midstream and bankside samples.

Supplementary Figure 2 | Receiver Operating Characteristic (ROC) diagnostic curve for predicted eDNA detection. Gray line indicates null (random classifier) expectation.

Supplementary Figure 3 | AIC versus theoretical maximum eDNA proportion at banks downstream from fish (P_{max}) for the full and reduced dataset. The reduced dataset excludes distances >1,600 m, which were not sampled for all streams. Even lateral eDNA dispersion across a stream (i.e., across one midstream and both bankside samples) would result in an optimum P_{max} value of 0.67.

Supplementary Figure 4 | Number of samples necessary for 95% chance of salmon detection downstream from a target reach. Optimal sampling is roughly 1,000 m downstream of fish at banks and immediately downstream of fish in the midstream. Optimal sample numbers for fish detection decreases with increasing numbers of fish and decreasing velocity (V). Low, med, and high V indicate 0.15, 0.35, and 0.55 m/s, respectively. Predictions here are shown for 5% quantile detectability stream (i.e., low detectability).

Supplementary Figure 5 | Salmon detection rate based on uniform sample spacing. Optimal sample spacing for fish detection increases with increasing numbers of fish and decreasing velocity (V). Low, med, and high V indicate 0.15, 0.35, and 0.55 m/s, respectively. Detection rate assumes three independent technical replicates. Predictions here are shown for 5% quantile detectability stream (i.e., low detectability). Fish were simulated in a 100 km-long stream and detection rates estimated from uniform sample spacing.

Supplementary Table 1 | Formulas for four eDNA dispersal models.

Supplementary Table 2 | Comparison of four eDNA dispersal models with and without samples >1,600 m downstream included.

REFERENCES

Abbott, B. W., Baranov, V., Mendoza-Lera, C., Nikolakopoulou, M., Harjung, A., Kolbe, T., et al. (2016). Using multi-tracer inference to move beyond single-catchment ecohydrology. *Earth Sci. Rev.* 160, 19–42. doi: 10.1016/j.earscirev.2016.06.014

Akatsuka, M., Takayama, Y., and Ito, K. (2018). *Applicability of Environmental DNA Analysis and Numerical Simulation to Evaluate Seagrass Inhabitants in a Bay*. Cupertino, CA: International Society of Offshore and Polar Engineers.

Andruszkiewicz, E. A., Koseff, J. R., Fringer, O., Ouellette, N., Lowe, A. B., Edwards, C., et al. (2019). Modeling environmental DNA transport in the coastal ocean using lagrangian particle tracking. *Front. Mar. Sci.* 6:477. doi: 10.3389/fmars.2019.00477

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

Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., and Lodge, D. M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environ. Sci. Technol.* 48, 1819–1827. doi: 10.1021/es404734p

Brown, J. J., Limburg, K. E., Waldman, J. R., Stephenson, K., Glenn, E. P., Juanes, F., et al. (2013). Fish and hydropower on the U.S. Atlantic coast: failed fisheries policies from half-way technologies. *Conserv. Lett.* 6, 280–286. doi: 10.1111/conl.12000

Burnham, K. P., and Anderson, D. R. (2003). *Model Selection and Multimodel inference: A Practical Information-Theoretic Approach*. Berlin: Springer Science & Business Media.

Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Methods Res.* 33, 261–304. doi: 10.1177/0049124104268644

Cairns, D. K. (2001). *An Evaluation of Possible Causes of the Decline in Pre-fishery Abundance of North American Atlantic salmon. Canadian Technical Report of Fisheries and Aquatic Sciences, No. 2358*. Dartmouth, NS: Bedford Institute of Oceanography, 67.

Deiner, K., and Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. *PLoS One* 9:e88786. doi: 10.1371/journal.pone.0088786

Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J.-C., and Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nat. Commun.* 7:12544.

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., et al. (2011). Persistence of environmental DNA in freshwater ecosystems. *PLoS One* 6:e23398. doi: 10.1371/journal.pone.0023398

Doi, H., Inui, R., Akamatsu, Y., Kanno, K., Yamanaka, H., Takahara, T., et al. (2017). Environmental DNA analysis for estimating the abundance and biomass of stream fish. *Freshw. Biol.* 62, 30–39. doi: 10.1111/fwb.12846

Dolan, C. R., and Miranda, L. E. (2004). Injury and mortality of warmwater fishes immobilized by electrofishing. *North Am. J. Fish. Manag.* 24, 118–127. doi: 10.1577/m02-115

Drummond, J. D., Larsen, L. G., González−Pinzón, R., Packman, A. I., and Harvey, J. W. (2017). Fine particle retention within stream storage areas at base flow and in response to a storm event. *Water Resour. Res.* 53, 5690–5705. doi: 10.1002/2016wr020202

Fisheries and Oceans Canada (2019). "Action plan for the Atlantic Salmon (Salmo salar), inner Bay of Fundy population in Canada," in *Species at Risk Act Action Plan Series* (Ottawa: Fisheries and Oceans Canada), 61.

Gibson, A. J., Amiro, P. G., and Robichaud-LeBlanc, K. A. (2003). *Densities of Juvenile Atlantic Salmon, Salmo Salar, in Inner Bay of Fundy Rivers During 2000 and 2002 with Reference to Past Abundance Inferred from Catch Statistics and Electrofishing Surveys*. Ottawa ON: Fisheries & Oceans Canada.

Harvey, J. W., Drummond, J. D., Martin, R. L., McPhilipps, L. E., Packman, A. I., Jerolmack, D. J., et al. (2012). Hydrogeomorphology of the hyperheic zone: stream solute and fine particle interactions with a dynamic streambed. *J. Geophys. Res.* 117:G00N11.

Jane, S. F., Wilcox, T. M., McKelvey, K. S., Young, M. K., Schwartz, M. K., Lowe, W. H., et al. (2015). Distance, flow and PCR inhibition: eDNA dynamics in two headwater streams. *Mol. Ecol. Resour.* 15, 216–227. doi: 10.1111/1755-0998.12285

Jerde, C. L., Olds, B. P., Shogren, A. J., Andruszkiewicz, E. A., Mahon, A. R., Bolster, D., et al. (2016). Influence of stream bottom substrate on retention and transport of vertebrate environmental DNA. *Environ. Sci. Technol.* 50, 8770–8779. doi: 10.1021/acs.est.6b01761

Jones, R. A., Anderson, L., and Clarke, C. N. (2014). *Assessment of the Recovery Potential for the Outer Bay of Fundy Population of Atlantic Salmon (Salmo salar): Status, Trends, Distribution, Life History Characteristics and Recovery Targets. DFO Canadian Science Advisory Secretariat (CSAS), Research Document 2014/008*. Ottawa, ON: Fisheries and Oceans Canada, 94.

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

Lacoursière-Roussel, A., Côté, G., Leclerc, V., and Bernatchez, L. (2015). Quantifying relative fish abundance with eDNA: a promising tool for fisheries management. *J. Appl. Ecol.* 53, 1148–1157. doi: 10.1111/1365-2664.12598

Lacoursière-Roussel, A., Dubois, Y., Normandeau, E., and Bernatchez, L. (2016a). Improving herpetological surveys in eastern North America using the environmental DNA method. *Genome* 59, 991–1007.

Lacoursière-Roussel, A., Rosabal, M., and Bernatchez, L. (2016b). Estimating fish abundance and biomass from eDNA concentrations: variability among capture methods and environmental conditions. *Mol. Ecol. Resour.* 16, 1401–1414. doi: 10.1111/1755-0998.12522

Laporte, M., Bougas, B., Côté, G., Champoux, O., Paradis, Y., Morin, J., et al. (2020). Caged fish experiment and hydrodynamic bidimensional modeling highlight the importance to consider 2D dispersion in fluvial environmental DNA studies. *Environ. DNA* 2, 362–372. doi: 10.1002/edn3.88

LeBlanc, F., Belliveau, V., Watson, E., Coomber, C., Simard, N., DiBacco, C., et al. (2020). Environmental DNA (eDNA) detection of marine aquatic invasive species (AIS) in Eastern Canada using a targeted species-specific qPCR approach. *Manag. Biol. Invas.* 11, 201–217. doi: 10.3391/mbi.2020.11.2.03

Levi, T., Allen, J. M., Bell, D., Joyce, J., Russell, J. R., Tallmon, D. A., et al. (2019). Environmental DNA for the enumeration and management of Pacific salmon. *Mol. Ecol. Resour.* 19, 597–608. doi: 10.1111/1755-0998.12987

Limburg, K. E., and Waldman, J. R. (2009). Dramatic declines in North Atlantic diadromous fishes. *BioScience* 59, 955–965. doi: 10.1525/bio.2009.59.11.7

Minshall, G. W., Thomas, S. A., Newbold, J. D., Monaghan, M. T., and Cushing, C. E. (2000). Physical factors influencing fine organic particle transport and deposition in streams. *J. N. A. Benthol. Soc.* 19, 1–16. doi: 10.2307/1468278

Miranda, L. E., and Kidwell, R. H. (2010). Unintended effects of electrofishing on nongame fishes. *Trans. Am. Fish. Soc.* 139, 1315–1321. doi: 10.1577/t09-225.1

Parrish, D. L., Behnke, R. J., Gephard, S. R., McCormick, S. D., and Reeves, G. H. (1998). Why aren't there more Atlantic salmon (Salmo salar)? *Can. J. Fish. Aquat. Sci.* 55, 281–287.

Phillips, C. B., Dallmann, J. D., Jerolmack, D. J., and Packman, A. I. (2019). Fine−particle deposition, retention, and resuspension within a sand−bedded stream are determined by streambed morphodynamics. *Water Resour. Res.* 12, 10303–10318. doi: 10.1029/2019wr025272

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., and Waits, L. P. (2013). Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. *Can. J. Fish. Aquat. Sci.* 70, 1123–1130. doi: 10.1139/cjfas-2013-0047

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., and Waits, L. P. (2014). Factors influencing detection of eDNA from a stream-dwelling amphibian. *Mol. Ecol. Resour.* 14, 109–116. doi: 10.1111/1755-0998.12159

Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., et al. (2018). Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Sci. Rep.* 8:10361.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rummer, J. L., and Bennett, W. A. (2005). Physiological effects of swim bladder overexpansion and catastrophic decompression on red snapper. *Trans. Am. Fish. Soc.* 134, 1457–1470. doi: 10.1577/t04-235.1

Sansom, B. J., and Sassoubre, L. M. (2017). Environmental DNA (eDNA) shedding and decay rates to model freshwater mussel eDNA transport in a river. *Environ. Sci. Technol.* 51, 14244–14253. doi: 10.1021/acs.est.7b05199

Sepulveda, A. J., Al-Chokhachy, R., Laramie, M. B., Crapster, K., Knotek, L., Miller, B., et al. (2021). It's complicated. Environmental DNA as a predictor of trout and char abundance in streams. *Can. J. Fish. Squat. Sci.* doi: 10.1139/cjfas-2020-0182 [Epub ahead of print],

Shogren, A. J., Tank, J. L., Andruszkiewicz, E., Olds, B., Mahon, A. R., Jerde, C. L., et al. (2017). Controls on eDNA movement in streams: transport, retention, and resuspension. *Sci. Rep.* 7, 5065–5065.

Shogren, A. J., Tank, J. L., Andruszkiewicz, E. A., Olds, B., Jerde, C., and Bolster, D. (2016). Modelling the transport of environmental DNA through a porous substrate using continuous flow-through column experiments. *J. R. Soc. Interface* 13:20160290. doi: 10.1098/rsif.2016.0290

Shogren, A. J., Tank, J. L., Egan, S. P., August, O., Rosi, E. J., Hanrahan, B. R., et al. (2018). Water flow and biofilm cover influence environmental DNA detection in recirculating streams. *Environ. Sci. Technol.* 52, 8530–8537. doi: 10.1021/acs.est.8b01822

Skalak, K. J., and Pizzuto, J. E. (2010). The distribution and residence time of suspended sediment stored within the channel margins of a gravel-bad bedrock river. *Earth Surf. Process. Landforms* 35, 435–446.

Turner, C. R., Uy, K. L., and Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biol. Cons.* 183, 93–102. doi: 10.1016/j.biocon.2014.11.017

Wetzel, R. G. (2001). *Limnology: Lake and River Ecosystems*. San Diego, CA: Academic Press.

Wilcox, T. M., McKelvey, K. S., Young, M. K., Lowe, W. H., and Schwartz, M. K. (2015). Environmental DNA particle size distribution from brook trout. *Conserv. Genet. Resour.* 7, 639–641. doi: 10.1007/s12686-015-0465-z

Wilcox, T. M., McKelvey, K. S., Young, M. K., Sepulveda, A. J., Shepard, B. B., Jane, S. F., et al. (2016). Understanding environmental DNA detection probabilities: a case study using a stream-dwelling char Salvelinus fontinalis. *Biol. Conserv.* 194, 209–216. doi: 10.1016/j.biocon.2015.12.023

Wood, Z. T., Erdman, B. F., York, G., Trial, J. G., and Kinnison, M. T. (2020). Experimental assessment of optimal lotic eDNA sampling and assay multiplexing for a critically endangered fish. *Environ. DNA* 2, 407–417. doi: 10.1002/edn3.64

Yates, M. C., Fraser, D. J., and Derry, A. M. (2019). Meta-analysis supports further refinement of eDNA for monitoring aquatic species-specific abundance in nature. *Environ. DNA* 1, 5–13. doi: 10.1002/edn3.7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 10 May 2021
doi: 10.3389/fenvs.2021.617924

Diversity Metrics Are Robust to Differences in Sampling Location and Depth for Environmental DNA of Plants in Small Temperate Lakes

Jennifer A. Drummond[1], Eric R. Larson[2], Yiyuan Li[3], David M. Lodge[4,5], Crysta A. Gantz[3], Michael E. Pfrender[3], Mark A. Renshaw[3], Adrienne M. S. Correa[1] and Scott P. Egan[1]*

[1] Department of BioSciences, Rice University, Houston, TX, United States, [2] Department of Natural Resources and Environmental Sciences, University of Illinois, Champaign, IL, United States, [3] Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, United States, [4] Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, United States, [5] Cornell Atkinson Center for Sustainability, Cornell University, Ithaca, NY, United States

OPEN ACCESS

Edited by:
Hiroki Yamanaka, Ryukoku University, Japan

Reviewed by:
Inger Greve Alsos, Arctic University of Norway, Norway
Kevin Leempoel, Royal Botanic Gardens, Kew, United Kingdom

***Correspondence:**
Jennifer A. Drummond
jenn@rice.edu;
jenn.drummond@gmail.com

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Environmental Science

Received: *15 October 2020*
Accepted: *10 March 2021*
Published: *10 May 2021*

Citation:
Drummond JA, Larson ER, Li Y, Lodge DM, Gantz CA, Pfrender ME, Renshaw MA, Correa AMS and Egan SP (2021) Diversity Metrics Are Robust to Differences in Sampling Location and Depth for Environmental DNA of Plants in Small Temperate Lakes. Front. Environ. Sci. 9:617924. doi: 10.3389/fenvs.2021.617924

Environmental DNA (eDNA) analysis methods permit broad yet detailed biodiversity sampling to be performed with minimal field effort. However, considerable uncertainty remains regarding the spatial resolution necessary for effective sampling, especially in aquatic environments. Also, contemporary plant communities are under-investigated with eDNA methods relative to animals and microbes. We analyzed eDNA samples from six small temperate lakes to elucidate spatial patterns in the distributions of algae and aquatic and terrestrial plants, using metabarcoding of the Internal Transcribed Spacer-1 (ITS1) genomic region. Sampling locations were varied across horizontal and vertical space: sites in each lake included a mixture of nearshore and offshore positions, each of which was stratified into surface (shallow) and benthic (deep) samples. We detected the expected community variation (beta diversity) from lake to lake, but only small effects of offshore distance and sampling depth. Taxon richness (alpha diversity) was slightly higher in nearshore samples, but displayed no other significant spatial effects. These diversity metrics imply that plant eDNA is more evenly distributed than its generating organisms in these small lake environments. Read abundances were heavily weighted toward aquatic macrophytes, though taxon richness was greatest in the algae and other non-vascular plants. We also identified representatives of many phylogenetically and ecologically varied plant taxa, including terrestrial species from surrounding areas. We conclude that freshwater plant eDNA surveys successfully capture differences among lake communities, and that easily accessible, shore-based sampling may be a reliable technique for informing research and management in similar ecosystems.

Keywords: aquatic plants, ITS1, freshwater, biodiversity, monitoring, environmental DNA, metabarcoding

INTRODUCTION

Questions and experiments in ecology rely on knowledge of what organisms are present in a system, and how their patterns change over space and time. Moreover, the task of monitoring biodiversity, during an unprecedented rate of global extinction, is vital for all the life sciences and the quality of human life (e.g., Díaz et al., 2006). Biodiversity sampling has traditionally involved observation, trapping, and other direct census methods (Trolliet et al., 2014; Olinger et al., 2017; Lõhmus et al., 2018), which are biased toward organisms that are easy to locate and identify (Bosch et al., 2017; Wheeldon et al., 2019).

To support and complement traditional approaches to ecosystem monitoring, new tools have been developed over the last decade, including field detection technologies and lab-based high-throughput molecular techniques (Egan et al., 2013, 2015; Larson et al., 2020). One such technique is the analysis of DNA extsracted from bulk environmental samples of air, water, or soil, which is known as environmental DNA or eDNA. This analysis is currently accomplished either by high-throughput sequencing or targeted PCR, and can target narrow or broad ranges of organisms in many different habitats (Lodge et al., 2012; Creer et al., 2016; Deiner et al., 2017a; Cristescu and Hebert, 2018). However, in the aquatic medium, water and organism movement can confound eDNA-based efforts to determine the spatial distribution of organisms. This spatial uncertainty is one of the most significant remaining challenges for eDNA methods when they aim to characterize aquatic communities in detail (Barnes et al., 2014; Deiner et al., 2017a).

Lake habitats and their photosynthetic inhabitants are crucial elements of many ecosystems, and are both involved in and threatened by a multitude of human activities (Tickner et al., 2020). These factors make them candidates for improved monitoring methods (Mantzouki et al., 2018). This study constitutes one of only a few so far to systematically investigate community-level aquatic plant assemblages using eDNA analysis of easily accessible water rather than sediment (e.g., Alsos et al., 2018), even though ecological monitoring is known to be most effective when it considers multiple taxa (Oertli et al., 2005) and aquatic plants are known to be important proxies for lake health (Hatzenbeler et al., 2004).

Despite the foundational importance of freshwater plants to many ecosystems, aquatic eDNA work on them has been largely confined to PCR-based detection of invasive or elusive species (Matsuhashi et al., 2016; Kuehne et al., 2020) or attempts to measure the quantity and distribution of such species (Gantz et al., 2018; Chase et al., 2020; Kuehne et al., 2020). For instance, eDNA monitoring can detect low abundances of algal species responsible for harmful algal blooms (Keller et al., 2017). However, aquatic plants share many of the same detection challenges as taxa that have received more active attention from the eDNA analysis community: they can be rare, cryptic, and/or difficult to identify precisely by morphology alone (Fahner et al., 2016; Bolpagni et al., 2018).

The spatial aspects of aquatic eDNA sampling have been widely investigated, but thus far without conclusive and generalizable results. When sequencing entire (microscopic) organisms captured in an environmental sample, it is clear that they were present at the actual place and time of collection. However, much of the DNA recovered in aquatic environments from non-microscopic plants and animals is in the form of extraorganismal DNA: shed tissues, cells, or organelles (Turner et al., 2014; Deiner et al., 2017b; Lacoursière-Roussel and Deiner, 2019). This material travels freely through the environment in a manner similar to that of other inert small particulates (Turner et al., 2014; Kelly et al., 2018; Nguyen et al., 2020), limiting our ability to determine when, or even if, the source organism was actually present at the location of sampling. The relationship of eDNA sampling to physical space is highly system-specific, depending both on physical aspects of the system such as lotic or lentic hydrology (Civade et al., 2016; Bedwell and Goldberg, 2020) and on various properties of the targeted organisms (Deiner et al., 2015).

The primary goal of this study is therefore to determine how choices of sampling location and depth affect results of eDNA metabarcoding for photosynthetic communities in small temperate lakes. Because the distributions of many rooted plant species within lakes are complex and yet well-understood based on traditional sampling approaches (Sand-Jensen et al., 2019), we seek to compare the behavior of their eDNA to that of eDNA from non-sedentary organisms such as unicellular algae and photosynthetic protists. We sample a broad array of plants, including algae, from six well-studied natural research lakes in the northern United States. We divide each lake into four "compartments" defined by two spatial variables: surface and benthic sampling depths, and nearshore and offshore sampling coordinates. We apply novel eDNA primers targeting a broad range of internal transcribed spacer (ITS1) rDNA sequences for plant communities (**Supplementary Table 1**); statistically evaluate the relative roles of sampling depth and distance from shore as both between-lake and within-lake sampling factors; and investigate the effects of these variables on eDNA characterization of these aquatic and lakeshore plant communities. We additionally investigate the contribution of terrestrial plants to the reservoir of eDNA in these lakes, in part because recent studies have revealed the potential for widespread terrestrial biodiversity sampling based on the collection of eDNA in river and lake water and sediment (Giguet-Covex et al., 2014; Cannon et al., 2015; Deiner et al., 2016; Harper et al., 2019).

In this study, we define "plants" as green and golden algae, phytoplankton, and vascular and non-vascular macroscopic plants whether of terrestrial or aquatic habit. We expect that plant communities differ among the four different lake compartments: nearshore surface, nearshore benthic, offshore surface, and offshore benthic. Replicating the sampling design in multiple isolated bodies of water gives us the ability to distinguish overall spatial effects from local inter-lake variation. Specifically, we test three hypotheses: that (1) alpha diversity of plants is highest at the surface and near the shore due to enhanced light and resource availability; (2) spatial differentiation (sample beta diversity among compartments) is higher in larger lakes; and (3) nearshore and surface environments include a greater proportion of DNA from terrestrial plant species. An improved understanding of how sampling location influences the eDNA-based detection

of lake-associated plant communities will support large-scale biodiversity monitoring efforts and help minimize the costs associated with these efforts.

MATERIALS AND METHODS

Field Sampling for eDNA

We collected eDNA samples between July 8th and 13th, 2015, from six lakes at the University of Notre Dame Environmental Research Center (UNDERC) in the northern peninsula of Michigan, United States (**Figure 1A**). This study is part of a larger, 12-lake study that incorporated sampling for vertebrates and invertebrates as well as plants. For the plant data, we report here on the six-lake subset of samples that was amplified with our novel ITS1 primer set described below.

The six lakes sampled for plant communities range in size from 0.8 to 67.3 ha (approx. 2–167 acres), with maximum depths from 5.2 to 13.7 m (**Table 1**). These study lakes are generally considered mesotrophic, or moderately productive. The lakes vary in their stream inputs and outputs, and our random site selection was, in part, intended to reduce site-specific inflow/outflow effects. Water clarity as measured by Secchi disk values near the time of sampling ranged from 0.8 to 4.3 m, as many of the lakes have brown or stained waters due to contributions of tannins and humic acids from adjacent wetlands. Summer depths for the metalimnion (the thermocline between the two distinct temperature layers of a stratified lake) have been measured for these six lakes in recent years and range from roughly 0.75 to 3.2 m (**Table 1**). Further biotic and abiotic details on these study lakes at UNDERC are given by previous studies (Kelly et al., 2014; Craig et al., 2015).

At each lake, six sampling locations, divided into three nearshore and three offshore, were chosen randomly using ESRI ArcMap (Redlands, California, United States) (**Figure 1B**). All offshore sites were located a minimum of 20 m from the shoreline (**Table 1**), at an average distance of 46.2 m (range 20.4–108.4 m). Nearshore locations were constrained to be within 1 m of the shoreline to represent sampling by an individual from the shore, though for consistency, all samples for this study were taken from boats. At each sample site, one surface water sample and one benthic water sample were collected for eDNA filtration. Fresh nitrile gloves were used at each of the six sample collection sites in each lake.

Benthic water samples were taken within 1 m of the lake bottom. In all cases, benthic water samples at offshore locations were taken below the metalimnion. Many nearshore benthic samples were taken above the depth of the metalimnion (**Figure 2**), though four of six lakes had steep enough nearshore depth profiles for some or all of their nearshore benthic samples to be taken below the metalimnion. Benthic sample depths at offshore locations were reliably deeper than at nearshore locations for all lakes, with a mean of 5.3 m (range 3.4–7.7 m). Within each lake, all nearshore benthic sample depths were shallower than any of the offshore benthic sample depths.

Surface eDNA samples were taken directly into 250 mL bottles that had been previously sterilized via autoclave followed by an external overnight soak in 10% bleach solution and an external rinse in deionized water. Benthic water samples were taken using 2 L Van Dorn samplers lowered to within 1 m of the lake bottom based on a depth estimate from a hand-held digital sonar (HawkEye Handheld Sonar, Bass Pro Shops, Springfield, Missouri, United States). Subsamples of the water retrieved by Van Dorn samplers were immediately transferred into sterile 250 mL bottles. Three total Van Dorn samplers were used to collect the samples in this study. These three samplers were sterilized between study lakes, and between nearshore and

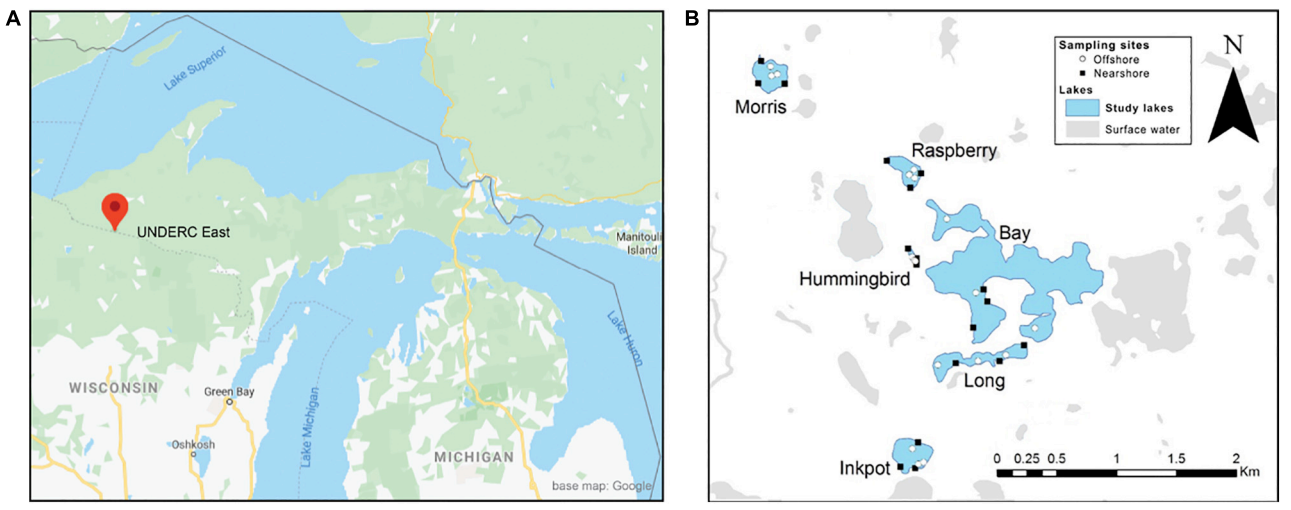


FIGURE 1 | Study lakes and sampling locations at the University of Notre Dame Environmental Research Center (UNDERC) in northern Michigan, United States. Detail map is centered near latitude 46.2336, longitude −89.5237 near the border between Wisconsin and the Upper Peninsula of Michigan. Nearshore and offshore sample locations are denoted within lakes by black squares and white circles, respectively.

TABLE 1 | Characteristics of lakes sampled at the University of Notre Dame Environmental Research Center (UNDERC), Michigan, United States, for environmental DNA metabarcoding during the summer of 2015.

					Deep samples below metalimnion					
Lake	**Area (ha)**	**Maximum depth (m)**	**Secchi depth (m)**	**Metalimnion depth (m)**	**Nearshore**			**Offshore**		
Bay	67.3	13.7	4.3	2.7–3.7	e	e	e	h	h	h
Long	7.9	14	3.6	1.3–1.8	e	m	h	h	h	h
Inkpot	6.6	5.2	1.1	2.3–2.7	e	e	e	h	h	h
Morris	5.9	6.7	1.6	1.1–1.2	e	e	h	h	h	h
Raspberry	4.6	6	2.8	1.7–1.8	e	e	h	h	h	h
Hummingbird	0.8	7	0.8	0.7–0.8	h	h	h	h	h	h

Information includes lake area, maximum lake depth, distance to visible Secchi disk (a measure of water clarity), and depth of metalimnion or thermocline (Kelly et al., 2014; Craig et al., 2015). The remaining columns indicate which of the six benthic samples were taken below the metalimnion, in contrast to samples at sites where the local lake bottom was itself above or within the lake's metalimnion.
e, epilimnion; m, metalimnion; h, hypolimnion.

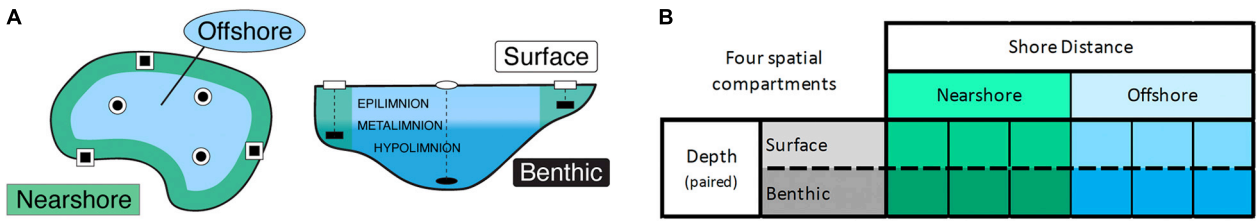


FIGURE 2 | Summary of sampling design for the six lakes in this study. Twelve samples were collected per lake based on depth and distance from shore to define four qualitative sample categories, herein termed compartments. **(A)** Generalized lake schematic showing nearshore and offshore zones (green and blue, respectively), and surface vs. benthic samples with respect to the metalimnion (see **Table 1**). **(B)** Table showing the same division of samples into three nearshore locations and three offshore locations, with a surface and a benthic sample taken at each location. The dashed line indicates that samples are paired vertically but not horizontally.

offshore sample locations within study lakes, by 15 min soaks in 50% bleach solution.

Sample bottles were kept on ice in a cooler until transported to a laboratory at UNDERC, where all samples for a given lake were immediately processed by filtration through 1.2 μm cellulose nitrate filters with the aid of an electric vacuum pump attached to side-arm flasks and filter funnels. In order to detect potential external contamination during handling and transport in the field, two filtration blanks per lake containing store-purchased bottled water were transported to and from the study sites along with the bottles for sample collection. Filtration blanks were filtered in the laboratory prior to sample filtration. Filters were immediately placed in 2 mL microcentrifuge tubes (United States Scientific, Ocala, Florida, United States) containing 700 μL of Longmire's buffer (Longmire et al., 1997). These tubes were kept at 4°C until being transported to the University of Notre Dame (Indiana, United States) for eDNA extraction. Fresh nitrile gloves were used during filtration of each individual sample.

ITS1 Primer Design

We chose to develop a primer specific to the study area, as we were not aware at the time of a published ITS1 primer with good specificity for the Great Lakes region. Previous experience with taxon-specific plant assays (i.e., *Elodea* and *Hydrilla*; Gantz et al., 2018) highlighted the ITS1 region as an ideal candidate for unique IDs among closely related plant species over relatively short fragments of DNA. Sequences for the ITS1 genomic region were downloaded from GenBank for a list of 119 locally relevant plant species (**Supplementary Table 2**). Sequence alignments were built and assays were designed on areas of the ITS1 region that were conserved across the list of plant taxa, based on recommendations from Primer3 (Untergasser et al., 2012). The final selected primer set, designated ITS1-F/-R3 (**Supplementary Table 1B**), produced a range of amplicons approximately 240–500 bp in size. Primers targeting matK, rbcL, and ITS2 were also developed and evaluated, and the ITS1 primers were found to be the best match for the local lake environment. The primers used are thus not intended to be widely employed outside the area of this study.

DNA Extraction and Target Amplification

DNA extractions followed a modified chloroform-isoamyl alcohol (24:1, Amresco) extraction and isopropanol precipitation protocol as outlined in Renshaw et al. (2015). Following the extraction process, rehydrated DNA pellets were treated with the OneStep™ PCR Inhibitor Removal Kit (Zymo Research, Irvine, California, United States).

A two-step PCR-based method was used for preparation of sequencing libraries (Olds et al., 2016). In the first step, locus-specific PCR amplicons were generated from each sample

using the novel primer set designed to amplify the intergenic spacer between the 18S and 5.8S ribosomal RNAs in plants: ITS1-F (5′-GTCGTAACAAGGTTTCCGTAGGT-3′) and ITS1-R3 (5′-GATATCCGTTGCCGRGAGTC-3′). The ITS1 primer set included an overhang on the 5′-end to allow for indexing and Illumina adapter addition during the two-step PCR-based Illumina library preparation (**Supplementary Table 1**).

A 50 μL first step PCR reaction was used with the following recipe: 29.5 μL sterile water, 10 μL 5× HF buffer, 1 μL 10 mM dNTPs, 1.5 μL 50 mM $MgCl_2$, 1.25 μL 10 μM ITS1-F + PrefixNX/F primer, 1.25 μL 10 μM ITS1-R3 + PrefixNX/R primer, 0.5 μL 2 U/μL iProof High-Fidelity DNA Polymerase (Bio-Rad, Hercules, California, United States), and 5 μL DNA. Temperature cycling conditions were: initial denaturation at 98°C for 2 min; 25 cycles: denaturation at 98°C for 10 s, annealing at 55°C for 20 s, extension at 72°C for 30 s; and final extension at 72°C for 10 min.

First step PCR products were run through a 2% agarose gel, stained with ethidium bromide, and visualized on a UV light platform. Amplified products were manually cut from the gels with single-use razor blades, cleaned with the QIAquick Gel Extraction Kit (Qiagen, Venlo, Netherlands), and eluted from spin columns with 30 μL of Buffer EB. The DNA concentration of each elution was quantified with the Qubit dsDNA HS Assay (Life Technologies, Carlsbad, California, United States).

Library Preparation and Sequencing

To complete the addition of the Illumina adapter and dual-indexing barcode, a 50 μL PCR reaction was used for the second step. The PCR mix consisted of 22 μL sterile water, 10 μL 5× HF buffer, 1 μL 10 mM dNTPs, 1.5 μL 50 mM $MgCl_2$, 5 μL 10 μM Nextera Index Primer 1 (N701-N712), 5 μL 10 μM Nextera Index Primer 2 (S502-S508 and S517), 0.5 μL 2 U/μL iProof High-Fidelity DNA Polymerase (Bio-Rad, Hercules, California, United States), and 5 μL DNA. Sequences for the Nextera Index Primers 1 and 2 are given in **Supplementary Table 1B**. All PCR primers, for both the first step and second step, were synthesized by Integrated DNA Technologies (Coralville, Iowa, United States).

Temperature cycling conditions for the second-step PCR consisted of an initial denaturation step at 98°C for 2 min; followed by 8 cycles of denaturation at 98°C for 10 s, annealing at 55°C for 20 s, and extension at 72°C for 30 s; followed by a final extension step at 72°C for 10 min. PCR clean-up was performed with Agencourt AMPure XP magnetic beads (Beckman-Coulter, Indianapolis, Indiana, United States) and DNA concentrations were quantified with the Qubit dsDNA HS Assay. Amplicon sizes were verified within each library on a Bioanalyzer DNA 7500 chip (Agilent Technologies, Santa Clara, California, United States). Sequencing was performed on an Illumina MiSeq sequencer at the University of Notre Dame's Genomics and Bioinformatics Core Facility[1] with a MiSeq Reagent Kit v3 (600-cycle; Illumina, San Diego, California, United States).

Over the course of the entire study, 98 total samples were processed, including 72 field samples (12 per lake), 12 field collection blanks (2 per lake), 8 no-template control PCR blanks (one per lake and one per sequencing run), 6 extraction blanks (1 per set of 18 eDNA extractions), and one blank each for the laboratory's DI water and tap water.

The extraction blanks included all the reagents used in the DNA extraction process and were processed alongside eDNA samples, monitoring for contamination that might occur during the DNA extraction step. Although none of the no-template controls visibly amplified, a band was cut out of the agarose gel at the expected size for each NTC and carried through the remaining library prep for subsequent Illumina sequencing.

Samples were spread across two runs on the Illumina MiSeq as part of a larger pooled eDNA study including vertebrate and general eukaryotic eDNA sequencing. Specific to this plant eDNA study, we included 36 field samples (3 lakes) per run, plus 12 associated control samples (6 field, 2 extraction, and 4 NTC). The tap and DI water blanks were amplified and sequenced on a separate run, associated with the larger project.

Bioinformatic Analysis

A summary of the informatic workflow, with full parameters and software versions, is available as **Supplementary Figure 1**. Raw MiSeq forward and reverse reads were assessed with FastQC (Andrews, 2010) and adapter-trimmed with Trimmomatic ILLUMINACLIP (Bolger et al., 2014) at a simple clip threshold of 6 bp. A sliding-window quality check was performed simultaneously with adapter removal: bases were trimmed when their windowed average fell below a quality score of 20. A custom Perl script (as used in Olds et al., 2016) performed demultiplexing to separate ITS-primer-specific reads from the other simultaneously sequenced amplicon sets. Forward and reverse reads were merged, and then quality-filtered with USEARCH (Edgar and Bateman, 2010) functions fastq_mergepairs and fastq_filter to a maximum expected error rate of 0.5 and a maximum N count of 1. Filtered merged reads were collapsed to unique sequences, then chimera-filtered and clustered with a 97% similarity threshold using USEARCH cluster_otus and usearch_global.

Taxonomic Assignment

Identification of OTUs was exclusively by BLASTn (Camacho et al., 2009), against the NCBI nt online database as of 12/11/2017. An initial assignment of OTUs to NCBI taxa was performed without identity thresholds, and any completely unassigned OTUs were discarded. Automated resolution of ambiguous BLAST results was then conducted according to a defined set of rules (see **Supplementary Figure 1**). The assigned OTUs were further filtered and combined as described below.

OTU Filtering

Details of OTU resolution and filtering are given in **Supplementary Figure 1**. The BLAST criteria for a fully passing OTU taxon assignment were a global percent identity of 97 or higher (as in Olds et al., 2016) and a bitscore of 100 or greater (similar to the approach in Shaw et al., 2016), corresponding to an *e*-value of approximately $2E^{-17}$. The overall

[1] genomics.nd.edu

set of these "fully passing" OTU assignments was taken as the working set of taxa presumed to be present in the study.

Ambiguity Resolution

Some OTUs were assigned by BLAST to more than one NCBI taxon with equal confidence (i.e., identical bitscore) using the NCBI BLAST taxonomy[2]. Within each set of ambiguous assignments for a given OTU, a subset of the most informative assignments was selected for final resolution of the OTU. The "most informative" assignments were considered to be, in order of priority: (1) species or subspecies records, (2) genus records with no species specified, and (3) records labeled as "uncultured" members of specific higher taxa. Records labeled as "environmental samples," which are unidentified sequences reported by other environmental studies, were always excluded in favor of more informative assignments, and an OTU was eliminated from the study altogether if an "uncultured" or "environmental" sample was the most informative assignment that appeared in the list of top hits. See the **Code Supplement** for further details of this step.

In most cases, this process successfully produced a single species or subspecies assignment. Where it did not, the most specific common genus, family, or order was used as the identification, as defined by the NCBI taxonomy[3]. Strong sub-threshold matches to known taxonomic groups indicated that some sequences were legitimate examples of as-yet unsequenced organisms rather than sequencing errors or chimeras.

Statistical Analyses

We calculated Shannon's alpha diversity index (H') and the inverse Simpson's diversity index (1/D) on raw passing taxon read totals within each lake, and across spatial variables, using the packages *phyloseq* version 1.30.0 (McMurdie and Holmes, 2013) and *vegan* version 2.5.7 (Oksanen et al., 2019) in the statistical program R (version 3.4.2). Differences in means across sets of variables were assessed with Mann-Whitney tests.

We tested for an effect of eDNA sample location and depth on plant communities using permutational multivariate analysis of variance (PERMANOVA) using the *adonis* function in the R package *vegan*. Due to the sensitivity of multivariate statistics to outliers, we omitted the rarest taxa (found in < 5 field samples total) from these beta-diversity analyses. A total of 47 vascular plants and 65 algal taxa were excluded from beta diversity analysis as a result of this filtering. We then $\log + 1$ transformed the remaining reads (Anderson et al., 2006) prior to calculating pairwise Bray-Curtis distances between our samples. In each PERMANOVA run on individual taxonomic groups, we considered sample location (nearshore or offshore) and depth (surface or benthic) as factors, with lake identity as a stratum or group. Inclusion of lakes as strata allowed us to partition the variation in communities among these lakes owing to inherent differences in biotic or abiotic factors between sites, as opposed to differences within lakes caused by sample site locations and depths. We also considered an interaction between location and depth, because offshore benthic water samples were deeper than nearshore benthic water samples in all cases, making nearshore benthic and offshore benthic samples potentially distinct. We attempted to perform independent analysis on the vascular-plant subportion of results, but total read abundance was too low to provide sufficient statistical power.

[2] ftp://ftp.ncbi.nlm.nih.gov/blast/db/taxdb.tar.gz, retrieved 12/2/2017

[3] ftp://ftp.ncbi.nih.gov/pub/taxonomy/taxdump.tar.gz, retrieved 12/2/2017

RESULTS

Sequencing

Two multiplexed 2×300 paired-end Illumina MiSeq runs yielded 32M reads, including 3.2M ITS1 amplicon sequences for this study. Among the 72 field samples, total per-sample plant-specific read counts ranged from 1.1 to 376K with a mean of 44.0K reads. After quality filtering, forward/reverse read merging, and chimera removal, per-sample plant-specific read counts for field samples totaled 2.4M reads, with per-sample mean of 33K reads. The laboratory controls, and 18 of the 24 field controls, yielded fewer than 10 plant reads apiece. Most reads in the remaining field controls were from plant taxa that appeared only in controls, and are therefore presumed to result from transport contaminants rather than field or laboratory cross-contamination.

OTU Clustering and Filtering

Counts of distinct sequences averaged 1,115 per sample, and distinct non-singleton sequences averaged 341 per sample. These non-singletons were clustered into 580 OTUs and identified with BLAST. Of these, 376 OTU identifications met thresholds of 97% identity and bitscore of 100 and were considered "passing" identifications. Of those, 50% matched with 100% identity. The majority had unambiguous BLAST assignments to a single NCBI species or subspecies. The remaining OTUs had identical sequence identities and bitscores for multiple possible taxa. These OTUs were resolved by a custom algorithm to species or subspecies level (7 OTUs); resolved to genus or a higher taxonomic level (21 OTUs); or discarded as unresolvable (3 OTUs). 5 OTUs were eliminated because they belonged to fungi or other non-target taxa. The resulting OTUs represented 248 distinct taxa, 63 of which were vascular plants and 185 of which were algae or other non-vascular plant taxa. See Methods and the Python code in the supplement for further details of this step.

Correspondence With Known Area Species

A list provided by UNDERC of known area vascular plants includes 609 species, mostly terrestrial, and approximately 25 aquatic. Of those species, only 159 have ITS1 records available at NCBI that would enable identification of corresponding OTUs in the eDNA results. The exact species overlaps between eDNA results and the area list numbered 21 (14 wetland or upland and 7 aquatic). This relatively low correspondence at the species level reflects both the absence of many species from NCBI ITS1 records and the fact that the plant lists cover a much wider area than is represented by these specific study lakes.

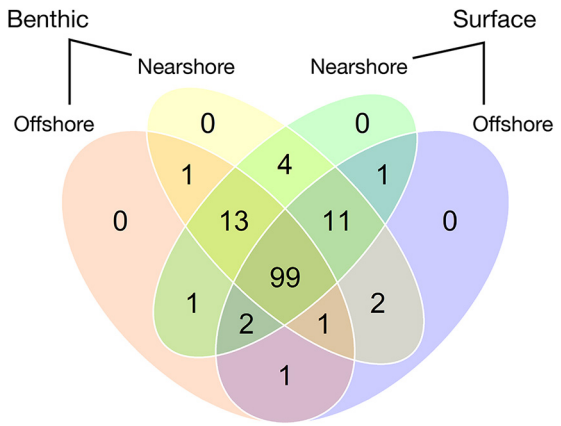


FIGURE 3 | Numbers of major taxa common to the four spatial compartments of eDNA sampling across six lakes. Major taxa are defined as the set of 136 identified plant and algal species (or, rarely, higher taxa) that were used for community analysis after the rarest taxa were excluded as described in Statistical Analyses, above.

At the genus level, however, the UNDERC area list included 117 total vascular plant genera with at least some available ITS1 species records. The eDNA results contained 47 vascular genera, with 41 genera in common between the two lists. Of the six "false positive" genera found in the eDNA results that do not appear on the area species list, three are crop or widespread weedy species (e.g., guar, soybean, ragweed) that may well contribute material to the area without appearing on local species inventories (**Supplementary Table 2**). The remaining OTUs belong to algae and other non-vascular plants, including microscopic taxa.

Our attempts to compare our lake-specific macrophyte results to previous, unpublished plant surveys[4] yielded mixed results, as might be expected given the age of the surveys. No more recent lake-specific plant surveys were available to us for comparison.

Spatial Taxon Distribution

Spatial distributions of the taxa identified by eDNA sequencing show that 150 of the 209 major taxa are found in all four distance-depth compartments of the dataset (**Figure 3**). Major taxa are those appearing in five or more samples, as defined as above for purposes of beta diversity analysis. None of these major taxa are unique to a single compartment.

Figure 4 contrasts numbers of taxa in various NCBI "taxon categories" with their read prevalence. This set of NCBI taxon categories are occasionally non-intuitive due to the presence of generic-seeming basal groups such as "flowering plants"; nevertheless, these high-level categorizations are a useful rough overview of the distribution of taxon types in the dataset. A majority of total taxa identified are algae and phytoplankton (**Figure 4A**), but those categories are less dominant in number of reads (**Figure 4B**).

[4]https://underc.nd.edu/assets/216433/fullsize/francl1996.pdf; https://underc.nd.edu/assets/215506/fullsize/brown1997.pdf

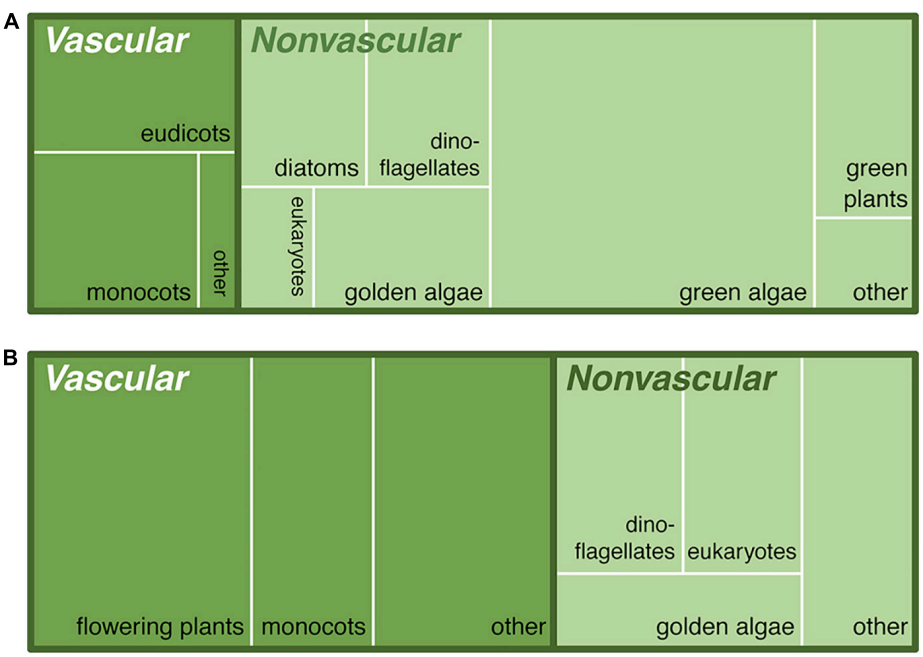


FIGURE 4 | Study taxa by NCBI taxonomic categories identified via ITS1 barcode sequencing in six lakes. These treemaps (Shneiderman, 1992) are proportional representations of vascular plants (darker green) and non-vascular plants/algae (lighter green) by **(A)** numbers of taxa identified vs. **(B)** read abundance. Note the variation in dominance in subcategories like "eukaryotes" (generally photosynthetic protists) and "flowering plants" (largely aquatics in the Nymphales and Ceratophyllales, predating the monocot/dicot split). The "green plants" in the non-vascular category are in the Zygnematophyceae, or conjugating algae, a close outgroup to the land plants.

TABLE 2 | Taxon and read abundances of vascular plants by wetland status.

(A)

Wetland indicator status	Taxa	Reads
Aquatic	18	668,608
Obligate wetland	9	7,461
Facultative wetland	4	53,098
Facultative	9	37,594
Facultative upland	22	18,074
Upland	1	28

(B)

Aggregated status	Taxa	Reads
Aquatic	18	668,608
Wetland (hydrophytes)	22	98,153
Upland (non-hydrophytes)	23	18,102

Wetland indicator statuses follow U. S. Army Corps of Engineers (USACoE) designations (U.S. Army Corps of Engineers, 2018). Hydrophyte and non-hydrophyte statuses correspond to official USDA designations. See https://plants.sc.egov.usda.gov/wetinfo.html for definitions of both schemata (retrieved as of 6/2018). (a) Read counts by detailed wetland indicator statuses, plus an "Aquatic" category selected from within the Obligate Wetland plants and encompassing submerged, floating, and emergent aquatic macrophytes. See ***Supplementary Table 3*** *for details. (b) Aggregations into USDA categories as used in* ***Figure 5****.*

Vascular Plant Taxa by Wetland Status

Of the 63 individual vascular plant taxa detected via eDNA sequencing, only 63% (40 taxa) are aquatic species or terrestrial species classified as hydrophytes according to United States federal wetland delineation guidelines (U.S. Army Corps of Engineers, 2018; **Table 2** and **Figure 5**). The other, substantial, portion of taxa (23) occupies non-aquatic classifications, indicating that a wide range of species contribute DNA to the lake environments. However, aquatic species heavily dominate the read abundances, and hydrophyte species also have read abundances disproportionately higher than their taxon representation would indicate (**Table 2**). The absolute number of sequencing reads from vascular plants was insufficient for statistical testing of vascular-specific diversity differences within or among lakes: 793K of 1.9M total reads, spread across 72 samples, with two aquatic plant taxa accounting for over half the vascular plant reads.

Alpha Diversity

Alpha diversity among identified taxa varied substantially among the six lakes, with mean Shannon index values ranging from 1.01 to 1.83 (**Figure 6A**). The inverse Simpson's diversity index (1/D; not shown) was consistent with the Shannon results (Spearman correlation $\rho = 0.943$). Mean alpha diversities of surface vs. benthic samples showed little difference (**Figure 6B**). Mean nearshore diversity was slightly though significantly higher than offshore diversity by both indices, with a p-value of 0.0026 for the difference in Shannon-index means (**Figure 6C**).

Beta Diversity and Community Composition

Community characterization by PERMANOVA was performed on the data from six lakes, across each spatial dimension, using lake identity as a blocking factor. Simultaneous analysis of all factors showed a dominant effect of lake identity as the primary explanatory variable for community composition ($R^2 = 0.5109$, $p < 0.001$). This substantial effect was followed by small effects of distance and depth ($R^2 = 0.0277$ and $R^2 = 0.0373$ respectively, with $p < 0.001$). The distance-depth interaction showed an additional small yet statistically significant effect ($R^2 = 0.0137$, $p = 0.014$) (**Table 3**).

Four of the six individual lakes also showed moderately-sized internal patterns with regard to depth and shore distance, with statistically significant R^2-values of up to 0.206 for depth and 0.414 for shore distance (**Table 4**). Effect sizes in each lake showed no significant correlations with per-lake alpha diversity, nor with physical lake characteristics such as longitude, latitude, water clarity, or maximum lake depth, as determined by Spearman correlation (**Supplementary Figure 2**).

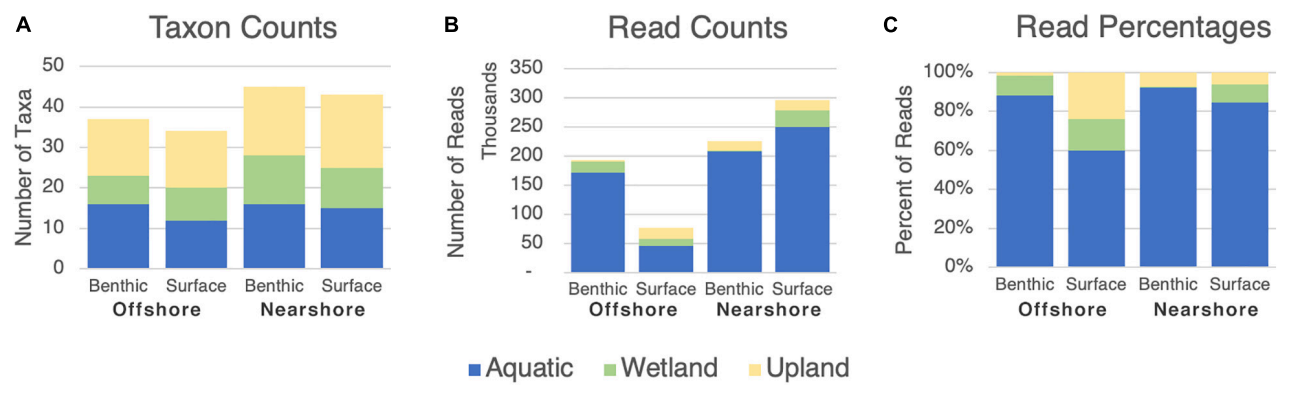


FIGURE 5 | Prevalence of wetland designations among vascular plant taxa categories identified by eDNA sampling in four spatial compartments. "Aquatic," "Wetland," and "Upland" correspond to the three aggregated statuses in **Table 2B** above. Category labels are the four spatial sampling compartments (depth/distance combinations). **(A)** Numbers of taxa from each broad category represented in eDNA results. **(B)** Numbers of reads represented in eDNA results. **(C)** Percentage of reads represented in eDNA results.

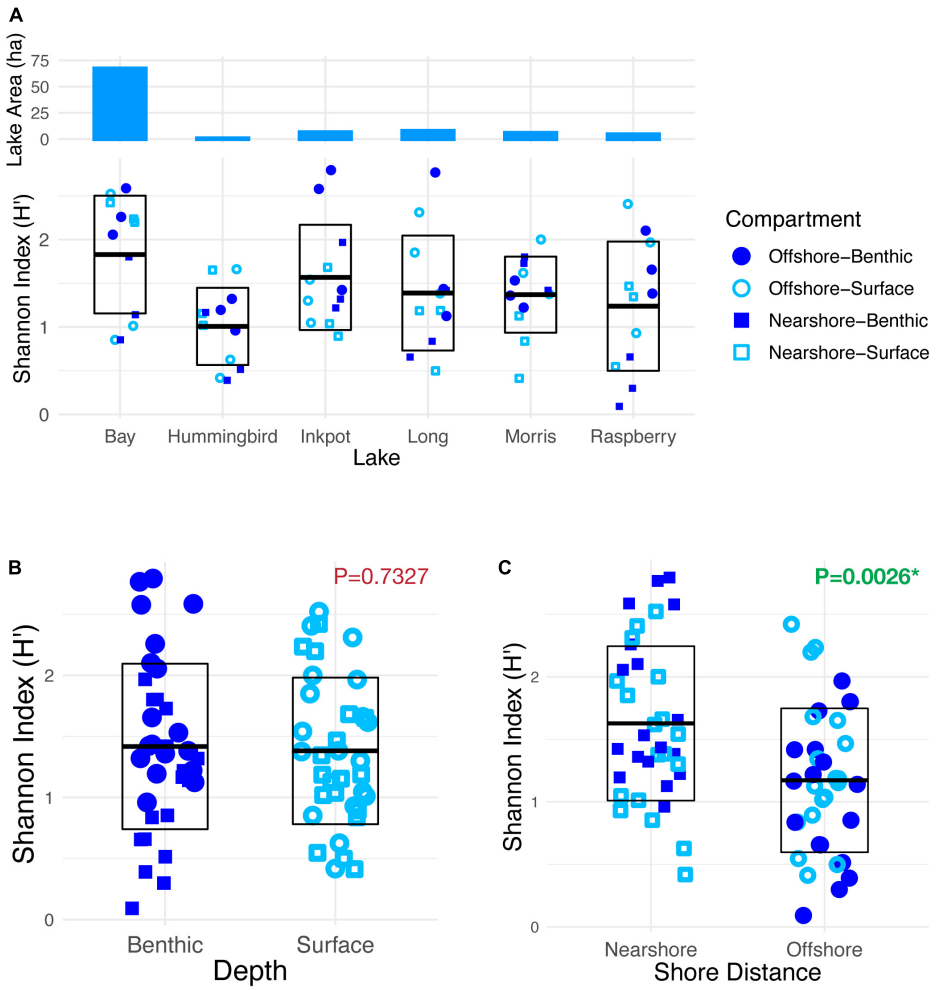


FIGURE 6 | Alpha diversity of combined plant and algal taxa detected by eDNA sampling in six lakes. **(A)** Distributions of Shannon's diversity index (H') for the 12 samples within each of six lakes, showing mean index value with one standard deviation above and below. Surface area of lakes is overlaid above (blue bars), showing a correspondence between lake size and mean alpha diversity (Spearman correlation: $\rho = 0.94$). **(B)** Shannon diversity for all surface vs. all benthic samples. Mean values are not significantly different. **(C)** Shannon diversity for all nearshore vs. all offshore samples. Mean nearshore alpha diversity is significantly higher than mean offshore diversity ($p < 0.005$).

TABLE 3 | Spatial effects by PERMANOVA across six lakes.

	Df	SumsOfSqs	MeanSqs	F.Model	R^2	Pr (>F)
Lake	5	7.7803	1.5561	15.6883	0.5109	0.001***
Distance	1	0.4214	0.4214	4.2489	0.0277	0.001***
Depth	1	0.5680	0.5680	5.7271	0.0373	0.001***
Distance:Depth	1	0.2090	0.2090	2.1068	0.0137	0.014*
Residuals	63	6.2487	0.0992		0.014	
Total	71	15.2274			1	

Columns are: sources of variation, degrees of freedom, sums of squares, mean squares, F statistics, partial R^2 and P-values, and zero, one, two, or three asterisks as visual indicators of P-values below 0.05, 0.01, and 0.001, respectively.

NMDS plots visually confirmed the beta diversity patterns, showing lakes scattered across the plot but distances and depths largely overlapping with only small differences in the overall shapes of the clusters (**Figure 7**).

DISCUSSION

Environmental DNA (eDNA) metabarcoding has grown rapidly as a novel but repeatedly proven tool for species surveillance (Rees et al., 2014; Thomsen and Willerslev, 2015). Our study is among the first to assess freshwater plant community diversity using eDNA metabarcoding and to address clear hypotheses about eDNA sampling location in lake community surveillance (Valentini et al., 2016). Even though plant community ecology is a fundamental field in biology, most plant-targeted eDNA studies have focused on a single species, largely in the context of invasives detection (Scriver et al., 2015; Fujiwara et al., 2016; Matsuhashi et al., 2016; Gantz et al., 2018). We report here on

TABLE 4 | Spatial effects by PERMANOVA within individual lakes.

	Spatial factors					
	Distance		Depth		Depth × Distance	
Lake	R^2	Pr (>F)	R^2	Pr (>F)	R^2	Pr (>F)
Bay	0.17420	0.006**	0.18897	0.003**	0.12178	0.111
Hummingbird	0.06885	0.248	0.41425	0.001***	0.06119	0.296
Inkpot	0.15671	0.075	0.06622	0.615	0.10142	0.284
Long	0.16211	0.053	0.15636	0.063	0.09853	0.253
Morris	0.18892	0.007**	0.25137	0.001***	0.07271	0.334
Raspberry	0.20165	0.003**	0.22658	0.002**	0.17080	0.111

PERMANOVA among the 12 samples in each lake with R^2-values (effect size) and posterior probabilities. One, two, or three asterisks are visual indicators of probabilities below 0.05, 0.01, and 0.001, respectively.

the success of a community inventory and its implications for future study design.

Spatial Distribution of Taxa Within and Among Lakes

The largest community assemblage differences in this study appear in the whole-community comparisons among lakes, which is consistent with traditional sampling approaches for lakes and ponds that find environmental heterogeneity to be the strongest predictor of beta diversity (Alahuhta et al., 2017). However, small beta diversity signals are also detectible in the spatial variables combined across the six lakes (**Table 3**). Moreover, the community compositions internal to 4 of the 6 lakes showed significant small or moderate differentiation in one or both spatial dimensions within the individual lake communities themselves (**Table 4**). The depth component of the signal is as strong as it is expected to be at any point in the year, according to an eDNA study on fish in a large English lake which confirmed the presumption that vertical habitat differentiation is strongest in the summer for temperate stratified lakes (Lawson Handley et al., 2019).

Due to the small but significant absolute magnitude of the spatial patterns (< 5%), we surmised that they might relate to subtle ecological effects, either from inherent characteristics of each lake or the effect of those characteristics on eDNA transport and degradation patterns. Accordingly, we investigated potential relationships between various characteristics of each lake and the sizes of distance and depth effects in that lake, testing for correlation between lake characteristics (position, surface area, maximum depth, and water clarity) and sequencing-based diversity metrics (alpha diversity, and magnitude and significance of depth and distance effects). No correlation between a lake characteristic and an alpha or beta diversity metric reached statistical significance either before or after correction for multiple testing. The lack of relationships between lake size and strength of spatial patterns causes us to reject our hypothesis that larger lakes will have clearer spatial structure in eDNA profiles.

We believe that our single-point sampling was sufficient to discover the existing spatial patterns in these lakes, given that several effect sizes were small but nevertheless highly

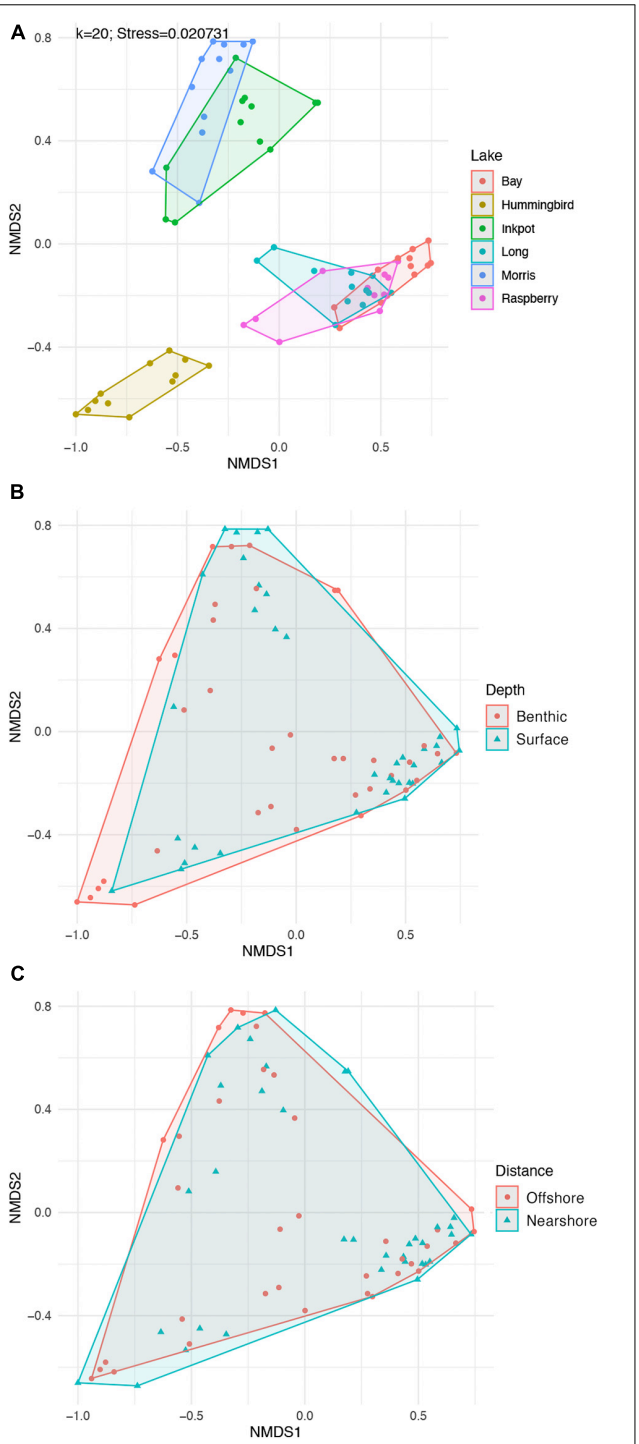


FIGURE 7 | Visualizations of diversity among lake communities based on 12 spatially varied eDNA samples from each of six lakes. Each point represents the read-abundance-based community of one sample from the indicated lake. All plots are based on NMDS ordination of Bray-Curtis distances calculated from log + 1-transformed read abundances. **(A)** Diversity among the six lakes, showing substantial clustering of points within lakes, corresponding to the R^2-value of 0.5109 (**Table 3**). The bottom two plots are the same points grouped **(B)** by depth category and **(C)** by shore distance category, showing the substantial overlap between the 36 sampling locations in one category vs. the other.

significant ($p < 0.001$). Regardless, in the future, we recommend taking multiple technical replicates and/or performing replicate PCRs for each sample to reduce the PCR-related false negative rate. Field technical replicates and PCR replicates would also increase the overall likelihood of detecting rare taxa that would be especially informative about site differences (Leray and Knowlton, 2017).

Our investigation of distance and depth patterns among our unreplicated lake eDNA point samples demonstrates two subtle and opposing effects. First is that low levels of unique taxa are present in each compartment, and therefore that point sampling may not generate an exhaustive species inventory where one is needed. Secondly, however, is an overall indication that much of the eDNA originating in each compartment is well-distributed throughout a small lake, which indicates that point sampling may be adequate for many other purposes. Several individual taxa per lake would have been missed had sampling been limited to one of the four spatial compartments, and therefore specific applications such as time-series comparisons would clearly benefit from consistency in sampling location. Alternatively, for basic lake-wide biological inventories, pooling samples from multiple location types may help to increase overall numbers of species detected without increasing the amount of laboratory work needed. Finally, the spatial information itself may be useful for certain types of detailed investigation such as those of ecological gradients or localized populations. However, the effects of sampling position were small in magnitude, and many taxa were found in common among spatial compartments, which indicates that community composition results were reasonably robust to the choice of sampling position. Our resulting recommendations are summarized in **Table 5**. Ideally, eDNA sampling projects of this type should perform pilot studies to determine the relative contributions and importance of spatial variation to the local eDNA environment.

Terrestrial and Aquatic Plant Diversity and Distribution

Broad airborne distribution of terrestrial plant DNA has been demonstrated to occur, largely mediated by particles other than pollen (Parducci et al., 2017; Sjögren et al., 2017; Johnson et al., 2019). Of the 77 individual vascular plant taxa detected by eDNA analysis in this study, fewer than half are classified as obligate or facultative upland species according to United States federal wetland delineation guidelines (U.S. Army Corps of Engineers, 2018). At this qualitative level, we therefore conclude that the vascular plant eDNA in these lakes is not a general homogenous sampling of the entire surrounding terrestrial area, but rather that the lakes' eDNA content is enriched in DNA from their local aquatic and shoreline species. However, our hypothesis that nearshore and surface environments will contain more terrestrial plant eDNA is not entirely supported because there is no marked increase in the relative numbers of upland species near the shore (**Figure 5C**). In fact, upland species are proportionately best represented in the offshore-surface compartment, perhaps owing in part to the lower overall read abundance in those samples (**Figure 5B**). Another possible explanation for the offshore-surface concentration is that windborne plant detritus settles relatively evenly across the surface of a lake, but once it sinks it is bound to or buried in sediment and less available for sampling in the water column.

The distribution of alpha diversity (**Figure 6**) across the two spatial dimensions is also somewhat unexpected, showing a small increase in diversity for the nearshore samples as a whole, but no distinction between surface and deep samples. This result partially rejects our hypothesis that diversity is greater both near shore and in shallower water due to light and habitat availability. However, it is conceivable that surface-area eDNA is harder to detect and identify, since ultraviolet light exposure may speed the breakdown of eDNA (Barnes and Turner, 2016).

Overall Taxon Distributions by Lake and Sample Location

Because taxa differ in their DNA shedding rates due to differences in their habitats and physical characteristics (Barnes and Turner, 2016), numbers of sequencing reads recovered from eDNA may not be representative of the absolute numbers or biomass of their associated taxa (Buxton et al., 2017; Fonseca, 2018). The most conservative approach for interpreting highly varied sets of taxa sampled by eDNA methods is therefore to convert read abundances to presences and absences (Ransome et al., 2017). However, we retained the read-number information in our analyses, because relative abundances of a given taxon can be compared among comparably-handled samples (Grey et al., 2018) as long as care is taken not to draw inferences across taxa. Regardless of organism-specific detection and quantification biases, comparable sets of eDNA results still generate "fingerprints" of communities that can be usefully compared and contrasted at a statistical or even ecological level (Leray and Knowlton, 2015).

Among the six studied lakes, three have their largest numbers of reads corresponding to *Nuphar variegata* (yellow or variegated pond-lily), an emergent aquatic plant. The other

TABLE 5 | eDNA sampling recommendations based on this study of small stratified north temperate lakes.

Study type	PCR replicates?	Field point sample replicates?	Samples per lake
Inter-community comparisons	Yes	As desired	Several samples, pooled or individual, at same depth/distance
Intra-community time series	Yes	As desired	Several samples, pooled or individual, at same depth/distance
Biological inventory	Yes	As desired	Multiple depths/distances, pooled
Detection of spatial patterns	Yes	Yes	Multiple depths/distances, analyzed individually

Our study used single field samples with no PCR replication and a variety of sample depths and distances. Based on the results we achieved with this design, we suggest these optimal designs for future studies in similar systems.

three are dominated numerically by algae, either single-celled (e.g., *Peridinium*) or colonial (e.g., Chrysophyceae, *Synura*). As mentioned above, read-number dominance does not indicate that pond-lilies or algae are absolutely dominant in area or biomass in these respective lakes; but the relative differences illustrate the inter-lake variability that gives rise to our ability to characterize one sample's community as different from another. Future work should look for correspondences between lake characteristics and read dominance by vascular plants vs. algae, to determine whether these read number relationships are accurately describing lakes with more or less emergent vegetation or are correlated with any known lake characteristics. Such correspondences could give rise to methods for interrogating other important lake characteristics such as primary production and eutrophication (Hilt et al., 2017; Poikane et al., 2018).

When a study's primary goal is to develop a reasonably complete catalog of an area's biodiversity, it is currently recommended to combine eDNA surveys with traditional sampling (Cowart et al., 2015; Olds et al., 2016; Shaw et al., 2016). The two methods have some complementary biases, and each will likely miss a significant number of taxa that are captured by the other (Nguyen et al., 2020). This two-pronged approach may be difficult in environments with barriers to observational sampling, such as remote or marine environments or environments that contain a large proportion of undescribed taxa. However, where the goal is, as in this study, to investigate broad patterns rather than individual species, sampling with consistent eDNA methods can provide useful information even without detailed comparison to conventionally established "ground truth." When necessary, fingerprints based simply on the presence and relative abundance of genotypes can be usefully compared among sample sites at an ecological or functional level, even when species lists are not complete and sequences are not fully identified (Leray and Knowlton, 2016).

Comparisons With Known Species Inventories

UNDERC maintains lists of terrestrial and aquatic plants found in the general area covered by the research center, and we compared these catalogs with the vascular-plant portion of our eDNA results. Fewer than a third of the vascular plants on the UNDERC area list have species-level ITS1 reference sequence records at NCBI and many taxa were therefore unavailable for identification using the ITS1 marker in our study. Of the many OTUs that were discarded due to identification failure, many presumably belong to these taxa, or to undescribed or unsequenced algae. However, at the genus level, the plants that were identified by eDNA are largely a subset of the listed area genera, indicating that detection and identification of local plants was generally successful. Only six taxa are left as false positives relative to the area species list, and most have reasonable explanations such as being common food or industrial species (guar, soybeans) or species found in the surrounding area (mulberry) (**Supplementary Table 3**). No lake-specific plant inventories are available from the year in which the eDNA was sampled. Comparisons with earlier macrophyte inventories from 1996 to 1997 were mixed, but this could reflect differences in timing and sampling locations as well as the passage of time. See **Supplementary Tables 4, 5** for presence/absence information per lake and taxon.

One clear result is that the eDNA results contain no trace of four of Michigan's most common non-indigenous invasive lake plants, despite the ITS1 sequences of those plants being available and susceptible to capture by the ITS1-F/-R3 primer set *in silico*. These species are common reed (*Phragmites australis*), purple loosestrife (*Lythrum salicaria*), curly-leaf pondweed (*Potamogeton crispus*), and Eurasian water-milfoil (*Myriophyllum spicatum*) (O'Neal and Soulliere, 2006). These species have, in fact, not been historically observed at UNDERC, probably due to strict enforcement of boat transfer and other measures for preventing invasive species transport. This finding establishes what may be a useful historical baseline for future sampling of these frequently researched lakes, and is a demonstration that careful management of lake use can prevent establishment of exotic species (Trebitz et al., 2017).

Performance of eDNA Sampling Methods in This Study

Overall, the study supports the premise that eDNA metabarcoding is accurate and specific, having captured a substantial number of the local area's known plant genera. Many area plants were not detected, representing false negatives that are largely due to reference database gaps, but could also potentially have been caused by primer biases, sampling methods, and/or seasonality. Specifically, it is possible that our broad assay preferentially amplified the generally shorter ITS1 regions of algae at the expense of vascular plants. On the other side, however, false positives were low. Because our main focus is on diversity metrics rather than species inventory, we retained some doubtful or out-of-range individual species as placeholders for close relatives. Only six genera remained outstanding as false positives relative to the list of known species (**Supplementary Table 3**). Species accumulation curves (**Supplementary Figure 3**) show a moderate approach to the asymptote, indicating that many though not all of the available species were detected.

An advantage of the eDNA sampling approach is that difficulties such as reference database incompleteness are not permanent ones, because methods in the capture and analysis of eDNA are rapidly improving (Lacoursière-Roussel and Deiner, 2019). Sequences generated years previously can be reanalyzed and reidentified as often as necessary to provide comparison datasets and historical information. Statistical methods are continually advancing, such that analysis may be repeated with better error correction, more sophisticated algorithms, and more complete reference databases, or simply to ensure analytical consistency with more recent studies.

CONCLUSION AND FUTURE WORK

We rejected part or all of our initial hypotheses about the spatial distribution of plant eDNA in these lakes, and therefore conclude that sampling of plants in small temperate lakes

of this type is robust to choice of sampling location. Given the subtle differences within lakes, but broad differences among them, the methods employed in this study should perform well for many purposes such as community fingerprinting and time-series monitoring even when limited to easily accessible shore sampling sites. A comprehensive inventory of a given small-lake community, with the intent of capturing the largest possible number of taxa, may nevertheless benefit from spatial diversity in sampling, especially when combined with adequate technical replicates; pooling samples across depth and shore distance may be the best strategy for maximizing taxon recovery. Further work is suggested to quantify the role of small lakes as significant repositories for terrestrial as well as aquatic plant DNA, and to characterize the functional significance of the patterns found in this and similar studies.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories: NCBI PRJNA693787.

AUTHOR CONTRIBUTIONS

EL, DL, CG, MP, and SE: concept and designing. EL, CG, and MR: field sampling and lab working. JD, AC, SE, EL, and YL: informatics and analysing. JD, AC, and SE: writing 1st draft. All authors writing final version.

FUNDING

This work was supported by the Biotechnology Risk Assessment Grant program, Grant No. 2016-33522-25629, from the U. S. Department of Agriculture, and was also supported by the U. S. Environmental Protection Agency Grant EPA-R5-GL2012-1. JD was supported by a Faculty Initiative Fund and an IDEA grant from Rice University to AC and SE.

ACKNOWLEDGMENTS

JD wishes to thank the two non-coauthor members of her master's thesis committee: Dr. Evan Siemann (Rice University, BioSciences) and Dr. Todd Treangen (Rice University, Computer Science).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenvs.2021.617924/full#supplementary-material

Supplementary Figure 1 | Details of OTU clustering and identification pipeline. Software versions of third-party software and descriptions of custom scripts for clustering and BLAST pipeline, including ambiguity resolution.

Supplementary Figure 2 | Spearman correlations with diversity indices among lake covariates. The only factors that display substantial correlation across lake characteristics and biodiversity metrics are the area of the lake and the strength of its depth effect. However, the effect is not statistically significant ($P = 0.195$), nor are any other correlations across lake characteristics and diversity metrics.

Supplementary Figure 3 | Rarefaction and species accumulation curves. **(A)** Rarefaction curves by read number and per lake. **(B)** Species accumulation curve by sample over entire study.

Supplementary Table 1 | Adapter and primer sequences. **(A)** Illumina Nextera primers used for all libraries. **(B)** Novel primers for plant and algal ITS1 regions. **(C)** Adapters, second round of PCR. **(D)** Barcode index sequences.

Supplementary Table 2 | Species and sequences used in novel ITS1 primer design.

Supplementary Table 3 | Potential false-positive plant taxa. These six records represent the only vascular genera in the eDNA dataset not previously recorded at UNDERC.

Supplementary Table 4 | Macrophyte presences by lake. Wetland statuses follow USDA/USACoE conventions, plus an additional category "AQU" for obligate wetland species that are fully aquatic, i.e., floating and submerged macrophytes.

Supplementary Table 5 | Phytoplankton presences by lake. Bold/highlighted items are the top ten phytoplankton taxa overall by read abundance.

Code Supplement | See supplementary file, or github.com/jennformatics/ITS-eDNA-2021.

REFERENCES

Alahuhta, J., Kosten, S., Akasaka, M., Auderset, D., Azzella, M. M., Bolpagni, R., et al. (2017). Global variation in the beta diversity of lake macrophytes is driven by environmental heterogeneity rather than latitude. *J. Biogeogr.* 44, 1758–1769. doi: 10.1111/jbi.12978

Alsos, I. G., Lammers, Y., Yoccoz, N. G., Jørgensen, T., Sjögren, P., Gielly, L., et al. (2018). Plant DNA metabarcoding of lake sediments: how does it represent the contemporary vegetation. *PLoS One* 13:e0195403. doi: 10.1371/journal.pone.0195403

Anderson, M. J., Ellingsen, K. E., and McArdle, B. H. (2006). Multivariate dispersion as a measure of beta diversity. *Ecol. Lett.* 9, 683–693. doi: 10.1111/j.1461-0248.2006.00926.x

Andrews, S. (2010). FastQC: a Quality Control Tool for High Throughput Sequence Data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

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

Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., and Lodge, D. M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environ. Sci. Technol.* 48, 1819–1827. doi: 10.1021/es404734p

Bedwell, M. E., and Goldberg, C. S. (2020). Spatial and temporal patterns of environmental DNA detection to inform sampling protocols in lentic and lotic systems. *Ecol. Evol.* 10, 1602–1612. doi: 10.1002/ece3.6014

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinfomatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bolpagni, R., Laini, A., Stanzani, C., and Chiarucci, A. (2018). Aquatic plant diversity in Italy: distribution, drivers and strategic conservation actions. *Front. Plant Sci.* 9:116.

Bosch, N. E., Gonçalves, J. M. S., Erzini, K., and Tuya, F. (2017). "How" and "what" matters: sampling method affects biodiversity estimates of reef fishes. *Ecol. Evol.* 7, 4891–4906. doi: 10.1002/ece3.2979

Buxton, A. S., Groombridge, J. J., Zakaria, N. B., and Griffiths, R. A. (2017). Seasonal variation in environmental DNA in relation to population size and environmental factors. *Sci. Rep.* 7:46294.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinform.* 10:421. doi: 10.1186/1471-2105-10-421

Cannon, M., Hester, J., Shalkhauser, A., Chan, E. R., Logue, K., Small, S. T., et al. (2015). Deep sequencing of environmental DNA isolated from the Cuyahoga river highlights the utility of river water samples to query surrounding aquatic and terrestrial biodiversity. *BioRxiv* 3, 77–90.

Chase, D. M., Kuehne, L. M., Olden, J. D., and Ostberg, C. O. (2020). Development of a quantitative PCR assay for detecting *Egeria densa* in environmental DNA samples. *Conserv. Genet. Resour.* 12, 545–548. doi: 10.1007/s12686-020-01152-w

Civade, R., Dejean, T., Valentini, A., Roset, N., Raymond, J. C., Bonin, A., et al. (2016). Spatial representativeness of environmental DNA metabarcoding signal for fish biodiversity assessment in a natural freshwater system. *PLoS One* 11:e0157366. doi: 10.1371/journal.pone.0157366

Cowart, D. A., Pinheiro, M., Mouchel, O., Maguer, M., Grall, J., Miné, J., et al. (2015). Metabarcoding is powerful yet still blind: a comparative analysis of morphological and molecular surveys of seagrass communities. *PLoS One* 10:e0117562. doi: 10.1371/journal.pone.0117562

Craig, N., Jones, S. E., Weidel, B. C., and Solomon, C. T. (2015). Habitat, not resource availability, limits consumer production in lake ecosystems. *Limnol. Oceanogr.* 60, 2079–2089. doi: 10.1002/lno.10153

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., et al. (2016). The ecologist's field guide to sequence-based identification of biodiversity. *Methods Ecol. Evol.* 7, 1008–1018. doi: 10.1111/2041-210x.12574

Cristescu, M. E., and Hebert, P. D. N. (2018). Uses and misuses of environmental DNA in biodiversity science and conservation. *Annu. Rev. Ecol. Evol. Syst.* 49, 209–230. doi: 10.1146/annurev-ecolsys-110617-062306

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., et al. (2017a). Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Mol. Ecol.* 26, 5872–5895. doi: 10.1111/mec.14350

Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J. C., and Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nat. Commun.* 7:12544.

Deiner, K., Renshaw, M. A., Li, Y., Olds, B. P., Lodge, D. M., and Pfrender, M. E. (2017b). Long-range PCR allows sequencing of mitochondrial genomes from environmental DNA. *Methods Ecol. Evol.* 8, 1888–1898. doi: 10.1111/2041-210x.12836

Deiner, K., Walser, J.-C., Mächler, E., and Altermatt, F. (2015). Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biol. Conserv.* 183, 53–63. doi: 10.1016/j.biocon.2014.11.018

Díaz, S., Fargione, J., Chapin, F. S., and Tilman, D. (2006). Biodiversity loss threatens human well-being. *PLoS Biol.* 4:e277. doi: 10.1371/journal.pbio.0040277

Edgar, R. C., and Bateman, A. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

Egan, S. P., Barnes, M. A., Hwang, C.-T., Mahon, A. R., Feder, J. L., Ruggiero, S. T., et al. (2013). Rapid invasive species detection by combining environmental DNA with light transmission spectroscopy. *Conserv. Lett.* 6, 402–409. doi: 10.1111/conl.12017

Egan, S. P., Ragland, G. J., Assour, L., Powell, T. H. Q., Hood, G. R., Emrich, S., et al. (2015). Experimental evidence of genome-wide impact of ecological selection during early stages of speciation-with-gene-flow. *Ecol. Lett.* 18, 817–825. doi: 10.1111/ele.12460

Fahner, N. A., Shokralla, S., Baird, D. J., and Hajibabaei, M. (2016). Large-scale monitoring of plants through environmental DNA metabarcoding of soil: recovery, resolution, and annotation of four DNA markers. *PLoS One* 11:e0157505. doi: 10.1371/journal.pone.0157505

Fonseca, V. G. (2018). Pitfalls in relative abundance estimation using eDNA metabarcoding. *Mol. Ecol. Resour.* 18, 923–926. doi: 10.1111/1755-0998.12902

Fujiwara, A., Matsuhashi, S., Doi, H., Yamamoto, S., and Minamoto, T. (2016). Use of environmental DNA to survey the distribution of an invasive submerged plant in ponds. *Freshw. Sci.* 35, 748–754. doi: 10.1086/685882

Gantz, C. A., Renshaw, M. A., Erickson, D., Lodge, D. M., and Egan, S. P. (2018). Environmental DNA detection of aquatic invasive plants in lab mesocosm and natural field conditions. *Biol. Invasions* 20, 2535–2552. doi: 10.1007/s10530-018-1718-z

Giguet-Covex, C., Pansu, J., Arnaud, F., Rey, P. J., Griggo, C., Gielly, L., et al. (2014). Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nat. Commun.* 5:4211.

Grey, E. K., Bernatchez, L., Cassey, P., Deiner, K., Deveney, M., Howland, K. L., et al. (2018). Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. *Sci. Rep.* 11:8843.

Harper, L. R., Lawson Handley, L., Carpenter, A. I., Ghazali, M., Di Muri, C., Macgregor, C. J., et al. (2019). Environmental DNA (eDNA) metabarcoding of pond water as a tool to survey conservation and management priority mammals. *Biol. Conserv.* 238:108225. doi: 10.1016/j.biocon.2019.108225

Hatzenbeler, G. R., Kampa, J. M., Jennings, M. J., and Emmons, E. E. (2004). A comparison of fish and aquatic plant assemblages to assess ecological health of small Wisconsin lakes. *Lake Reserv. Manag.* 20, 211–218. doi: 10.1080/07438140409354245

Hilt, S., Brothers, S., Jeppesen, E., Veraart, A. J., and Kosten, S. (2017). Translating regime shifts in shallow lakes into changes in ecosystem functions and services. *BioScience* 67, 928–936. doi: 10.1093/biosci/bix106

Johnson, M. D., Cox, R. D., and Barnes, M. A. (2019). The detection of a non-anemophilous plant species using airborne eDNA. *PLoS One* 14:e0225262. doi: 10.1371/journal.pone.0225262

Keller, S. R., Hilderbrand, R. H., Shank, M. K., and Potapova, M. (2017). Environmental DNA genetic monitoring of the nuisance freshwater diatom, *Didymosphenia geminata*, in eastern North American streams. *Divers. Distrib.* 23, 381–393. doi: 10.1111/ddi.12536

Kelly, P. T., Solomon, C. T., Weidel, B. C., and Jones, S. E. (2014). Terrestrial carbon is a resource, but not a subsidy, for lake zooplankton. *Ecology* 95, 1236–1242. doi: 10.1890/13-1586.1

Kelly, R. P., Gallego, R., and Jacobs-Palme, E. (2018). The effect of tides on nearshore environmental DNA. *PeerJ* 6:e4521. doi: 10.7717/peerj.4521

Kuehne, L. M., Ostberg, C. O., Chase, D. M., Duda, J. J., and Olden, J. D. (2020). Use of environmental DNA to detect the invasive aquatic plants *Myriophyllum spicatum and Egeria densa* in lakes. *Freshw. Sci.* 39.

Lacoursière-Roussel, A., and Deiner, K. (2019). Environmental DNA is not the tool by itself. *J. Fish Biol.* 98, 383–386. doi: 10.1111/jfb.14177

Larson, E. R., Graham, B. M., Achury, R., Coon, J. J., Daniels, M. K., Gambrell, D. K., et al. (2020). From eDNA to citizen science: emerging tools for the early detection of invasive species. *Front. Ecol. Environ.* 18, 194–202. doi: 10.1002/fee.2162

Lawson Handley, L., Read, D. S., Winfield, I. J., Kimbell, H., Johnson, H., Li, J., et al. (2019). Temporal and spatial variation in distribution of fish environmental DNA in England's largest lake. *Environ. DNA* 1, 26–395. doi: 10.1002/edn3.5

Leray, M., and Knowlton, N. (2015). DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. *Proc. Natl. Acad. Sci. USA* 112, 2076–2081. doi: 10.1073/pnas.1424997112

Leray, M., and Knowlton, N. (2016). Censusing marine eukaryotic diversity in the twenty-first century. *Philos. Trans. R. Soc. B Biol. Sci.* 371:20150331. doi: 10.1098/rstb.2015.0331

Leray, M., and Knowlton, N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. *PeerJ* 5:e3006. doi: 10.7717/peerj.3006

Lodge, D. M., Turner, C. R., Jerde, C. L., Barnes, M. A., Chadderton, L., Egan, S. P., et al. (2012). Conservation in a cup of water: estimating biodiversity and population abundance from environmental DNA. *Mol. Ecol.* 21, 2555–2558. doi: 10.1111/j.1365-294x.2012.05600.x

Lõhmus, A., Lõhmus, P., and Runnel, K. (2018). A simple survey protocol for assessing terrestrial biodiversity in a broad range of ecosystems. *PLoS One* 13:e0208535. doi: 10.1371/journal.pone.0208535

Longmire, J. L., Maltbie, M., and Baker, R. J. (1997). "Use of "lysis buffer" in DNA isolation and its implication for museum collections," in *Occasional Papers, Museum of Texas Tech University*, (Lubbock, TX: Museum of Texas Tech University).

Mantzouki, E., Beklioglu, M., Brookes, J. D., de Senerpont Domis, L. N., Dugan, H. A., Doubek, J. P., et al. (2018). Snapshot surveys for lake monitoring, more than a shot in the dark. *Front. Ecol. Evol.* 6:201.

Matsuhashi, S., Doi, H., Fujiwara, A., Watanabe, S., and Minamoto, T. (2016). Evaluation of the environmental DNA method for estimating distribution and

biomass of submerged aquatic plants. *PLoS One* 11:e0156217. doi: 10.1371/journal.pone.0156217

McMurdie, P. J., and Holmes, S. (2013). phyloseq: an r package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. doi: 10.1371/journal.pone.0061217

Nguyen, B. N., Shen, E. W., Seemann, J., Correa, A. M. S., O'Donnell, J. L., Altieri, A. H., et al. (2020). Environmental DNA survey captures patterns of fish and invertebrate diversity across a tropical seascape. *Sci. Rep.* 10:6729.

O'Neal, R. P., and Soulliere, G. J. (2006). *Conservation Guidelines for Michigan Lakes and Associated Natural Resources*. Available online at: http://www.michigan.gov/dnr/ (accessed June 2018).

Oertli, B., Joye, D. A., Castella, E., Juge, R., Lehmann, A., and Lachavanne, J. B. (2005). PLOCH: A standardized method for sampling and assessing the biodiversity in ponds. *Aquatic Conserv. Mar. Freshw. Ecosyst.* 15, 665–679. doi: 10.1002/aqc.744

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). *Vegan: Community Ecology Package*. Available online at: https://cran.r-project.org/package=vegan.

Olds, B. P., Jerde, C. L., Renshaw, M. A., Li, Y., Evans, N. T., Turner, C. R., et al. (2016). Estimating species richness using environmental DNA. *Ecol. Evol.* 6, 4214–4226.

Olinger, L. K., Heidmann, S. L., Durdall, A. N., Howe, C., Ramseyer, T., Thomas, S. G., et al. (2017). Altered juvenile fish communities associated with invasive *Halophila stipulacea* seagrass habitats in the U. S. Virgin Islands. *PLoS One* 12:e0188386. doi: 10.1371/journal.pone.0188386

Parducci, L., Bennett, K. D., Ficetola, G. F., Alsos, I. G., Suyama, Y., Wood, J. R., et al. (2017). Ancient plant DNA in lake sediments. *New Phytol.* 214, 924–942. doi: 10.1111/nph.14470

Poikane, S., Portielje, R., Denys, L., Elferts, D., Kelly, M., Kolada, A., et al. (2018). Macrophyte assessment in European lakes: diverse approaches but convergent views of 'good' ecological status. *Ecol. Indic.* 94, 185–197. doi: 10.1016/j.ecolind.2018.06.056

Ransome, E., Geller, J. B., Timmers, M., Leray, M., Mahardini, A., Sembiring, A., et al. (2017). The importance of standardization for biodiversity comparisons: a case study using autonomous reef monitoring structures (ARMS) and metabarcoding to measure cryptic diversity on Mo'orea coral reefs, French Polynesia. *PLoS One* 12:e0175066. doi: 10.1371/journal.pone.0175066

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R. M., and Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. *J. Appl. Ecol.* 51, 1450–1459. doi: 10.1111/1365-2664.12306

Renshaw, M. A., Olds, B. P., Jerde, C. L., Mcveigh, M. M., and Lodge, D. M. (2015). The room temperature preservation of filtered environmental DNA samples and assimilation into a phenol-chloroform-isoamyl alcohol DNA extraction. *Mol. Ecol. Resour.* 15, 168–176. doi: 10.1111/1755-0998.12281

Sand-Jensen, K., Andersen, M. R., Martinsen, K. T., Borum, J., Kristensen, E., and Kragh, T. (2019). Shallow plant-dominated lakes - extreme environmental variability, carbon cycling and ecological species challenges. *Ann. Bot.* 124, 355–366. doi: 10.1093/aob/mcz084

Scriver, M., Marinich, A., Wilson, C., and Freeland, J. (2015). Development of species-specific environmental DNA (eDNA) markers for invasive aquatic plants. *Aquat. Bot.* 122, 27–31. doi: 10.1016/j.aquabot.2015.01.003

Shaw, J. L. A., Clarke, L. J., Wedderburn, S. D., Barnes, T. C., Weyrich, L. S., and Cooper, A. (2016). Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biol. Conserv.* 197, 131–138. doi: 10.1016/j.biocon.2016.03.010

Shneiderman, B. (1992). Tree visualization with Tree-Maps: 2-d space-filling approach. *ACM Trans. Graph.* 11, 92–99. doi: 10.1145/102377.115768

Sjögren, P., Edwards, M. E., Gielly, L., Langdon, C. T., Croudace, I. W., Merkel, M. K. F., et al. (2017). Lake sedimentary DNA accurately records 20th Century introductions of exotic conifers in Scotland. *New Phytol.* 213, 929–941. doi: 10.1111/nph.14199

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

Tickner, D., Opperman, J. J., Abell, R., Acreman, M., Arthington, A. H., Bunn, S. E., et al. (2020). Bending the curve of global freshwater biodiversity loss: an emergency recovery plan. *BioScience* 70, 330–342. doi: 10.1093/biosci/biaa002

Trebitz, A. S., Hoffman, J. C., Darling, J. A., Pilgrim, E. M., Kelly, J. R., C1, E. A., et al. (2017). Early detection monitoring for aquatic non-indigenous species: optimizing surveillance, incorporating advanced technologies, and identifying research needs. *J. Environ. Manag.* 202, 299–310. doi: 10.1016/j.jenvman.2017.07.045

Trolliet, F., Huynen, M.-C., Vermeulen, C., and Hambuckers, A. (2014). Use of camera traps for wildlife studies. a review. *Biotechnol. Agron. Soc. Environ.* 18, 446–454.

Turner, C. R., Barnes, M. A., Xu, C. C. Y., Jones, S. E., Jerde, C. L., and Lodge, D. M. (2014). Particle size distribution and optimal capture of aqueous macrobial eDNA. *Methods Ecol. Evol.* 5, 676–684. doi: 10.1111/2041-210x.12206

U. S. Army Corps of Engineers (2018). *National Wetland Plant List, version 3.4*.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., et al. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Res.* 40:e115. doi: 10.1093/nar/gks596

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., et al. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Mol. Ecol.* 25, 929–942. doi: 10.1111/mec.13428

Wheeldon, A., Mossman, H. L., Sullivan, M. J. P., Mathenge, J., and de Kort, S. R. (2019). Comparison of acoustic and traditional point count methods to assess bird diversity and composition in the Aberdare National Park, Kenya. *Afr. J. Ecol.* 57, 168–176. doi: 10.1111/aje.12596

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

BRIEF RESEARCH REPORT
published: 13 May 2021
doi: 10.3389/fevo.2021.613200

Detection of Jaguar (*Panthera onca*) From Genetic Material in Drinking Water

Taylor M. Wilcox[1], Anthony Caragiulo[2], Joseph C. Dysthe[1], Thomas W. Franklin[1], Daniel H. Mason[1], Kevin S. McKelvey[1], Katherine E. Zarn[1] and Michael K. Schwartz[1]*

[1] National Genomics Center for Wildlife and Fish Conservation, Rocky Mountain Research Station, USDA Forest Service, Missoula, MT, United States, [2] Sackler Institute for Comparative Genomics, American Museum of Natural History, New York, NY, United States

Jaguar (*Panthera onca*) are of conservation concern and occur at very low densities in the northern portion of their range in northern Mexico and the southwestern United States. Environmental DNA sampling to detect genetic material from drinking water may be an effective approach for jaguar detection in these arid landscapes. Here we develop a qPCR assay for the detection of jaguar mitochondrial DNA, show that large quantities of DNA (mean 66,820 copies/L) can be found in the drinking water of captive animals, and observe detectable levels of DNA (80 copies/L) in a wild habitat with known jaguar populations. We suggest that environmental DNA sampling may represent a useful, complementary sampling tool for detection of rare jaguars, although effective application would require careful consideration of DNA persistence time in the environment.

Keywords: environmental DNA, eDNA, non-invasive genetics, wildlife, monitoring

OPEN ACCESS

Edited by:
Matthew A. Barnes, Texas Tech University, United States

Reviewed by:
Antoinette J. Piaggio, National Wildlife Research Center, Animal and Plant Health Inspection Service (USDA), United States
Allan D. McDevitt, University of Salford, United Kingdom
Rafael Valentin, Princeton University, United States

****Correspondence:***
Taylor M. Wilcox
taylor.wilcox@usda.gov

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *01 October 2020*
Accepted: *21 April 2021*
Published: *13 May 2021*

Citation:
Wilcox TM, Caragiulo A, Dysthe JC, Franklin TW, Mason DH, McKelvey KS, Zarn KE and Schwartz MK (2021) Detection of Jaguar (Panthera onca) From Genetic Material in Drinking Water. Front. Ecol. Evol. 9:613200. doi: 10.3389/fevo.2021.613200

INTRODUCTION

Jaguar (*Panthera onca*) are of conservation concern, particularly in the northern portion of their range in northern Mexico and the southwestern United States (Brown and González, 2000). Within the United States, the species is Federally protected (U.S. Federal Register 37 FR 6476) and occurs at extremely low abundances, with individuals rarely being detected. Current jaguar monitoring in the United States primarily uses camera traps and genetic testing of scat samples which may be located with the use of detection dogs (Culver, 2016). These approaches can be labor intensive, involving hundreds of cameras and tens of thousands of images (Culver, 2016). Additional sampling tools could help build a better understanding of jaguar distributions at the northern margin of their range and enable more effective protection of rare individuals and management of their habitat.

Environmental DNA (eDNA) sampling—the inference of species presence from genetic material in the environment—has been rapidly adopted for rare aquatic species sampling (e.g., Cristescu and Hebert, 2018; Sepulveda et al., 2020). Recent studies suggest that under the right circumstances, eDNA sampling of water may also be an effective approach for the detection of terrestrial species (e.g., Williams et al., 2018; Franklin et al., 2019). Harper et al. (2019) and Sales et al. (2020) documented the ability to detect terrestrial species from water samples, in some cases with detection rates comparable to camera trapping. Although there are likely multiple routes of transmission, DNA deposited when drinking is likely a major contributor of terrestrial species eDNA in aquatic habitats (Rodgers and Mock, 2015). In arid landscapes, drinking water may be scarce and represent relatively concentrated sources of jaguar eDNA.

Here we describe a hydrolysis assay developed for detection of jaguar mitochondrial DNA in environmental samples, then demonstrate this tool on water samples taken from known and suspected jaguar drinking water sources.

METHODS

Using sequence data from the NCBI GenBank database and the R (R Core Development Team, 2020) package *DECIPHER* (Wright, 2016), we designed and tested candidate quantitative PCR (qPCR) primers *in silico*, targeting jaguar mitochondrial DNA to the exclusion of 14 other felid species that either potentially co-occur with jaguar in North America or are closely related to jaguar (**Table 1**). Although only four jaguar mitochondrial genome sequences were included in this initial screen, there are low levels of genetic diversity and structure within the northern range of this species (e.g., Wultsch et al., 2016). We then selected one of these primer sets within the mitochondrial gene *ATP6* for hydrolysis probe development using PrimerExpress software (Thermo Fisher Scientific). These primers were selected based on a scan of the entire mitochondrial genome, but this same locus has also been found to have particularly good species discriminatory power across Carnivora (Chaves et al., 2012). We also conducted an *in silico* analysis of primer specificity through a BLAST (Atschul et al., 1990) search against GenBank to identify any unexpected cross-amplification with non-felids, but we did not attempt to use an *in silico* approach to evaluate potential cross-amplification of other rare felids which are not expected to be found in North America. We sourced the primers from Integrated DNA Technologies and obtained a FAM-labeled minor groove-binding (MGB) non-fluorescent quencher (NFQ) probe from Thermo Fisher Scientific (**Table 2**).

We validated assay specificity *in vitro* by testing tissue-extracted DNA from felids which may be found in the southwestern U.S. (*Felis catus, Puma concolor, Lynx rufus, Leopardus pardalis, Puma yagouaroundi*) or are closely related to jaguar [*Acinonyx jubatus* (n = 2), *Leptailurus serval* (n = 2), *Panthera leo* (n = 3), *Panthera tigris* (n = 3)]. These samples were sourced from the collections of partners for other projects and were collected in accordance with any relevant animal care guidelines. We extracted DNA using the DNeasy Blood and Tissue Kit (QIAGEN), quantified DNA using a Qubit fluorometer (Invitrogen), and diluted extracts to approximately 0.1 ng genomic DNA per microliter (0.4 ng gDNA per reaction), then analyzed with qPCR as described below. There was low-level amplification in one African lion sample. Low-level contamination of tissue-derived DNA samples that is only detected when used for eDNA-type applications is common (Rodgers, 2017) and was suspected in this case because there are many basepair mismatches between African lion and the jaguar assay (**Table 1**). To confirm this, we ensured that the assay did not amplify a synthetic gene fragment with the same sequence as African lion (gBlock; Integrated DNA Technologies), diluted to 6,250 and 1,250 copies per reaction.

We tested the ability of the assay to amplify jaguar DNA by analyzing DNA extracted from 10 jaguar scats that were collected in Belize and included in Menchaca et al. (2019; AC; Sackler Institute for Comparative Genomics). We tested the ability to detect jaguar DNA from drinking water by analyzing DNA extracted from six water samples provided to captive animals at the Phoenix Zoo in Phoenix, Arizona (n = 5; **Figure 1**) and Banana Bank Lodge in Belize (n = 1, **Table 3**). This sampling required minimal animal disturbance and sampling at the Phoenix Zoo was approved by the Arizona Center for Nature Conservation Research Committee. Finally, we analyzed five water samples from ponds in Belize where wild jaguar are known to occur regionally, but whose recent use of these habitats is unknown. Water samples were collected as described in Carim et al. (2016b). Briefly, for each sample, 5 L of water was drawn through a 47 mm diameter, 1.5 micron pore size glass microfiber filter paper using an electric peristaltic pump at the sampling site. The filter paper was then stored in silica desiccant until received at the lab (<2 weeks) when they were archived at −20°C until extraction. Filters were handled with sterile forceps. Filter cups, forceps, and all other sampling supplies were prepared in a dedicated, restricted-access room at the National Genomics Center for Wildlife and Fish Conservation and were sterilized using a 50% household bleach solution and ultrapure water (4% hypochlorite solution).

We then extracted DNA from half filters using DNeasy Blood and Tissue Kit (QIAGEN) with modifications as described in

TABLE 1 | List of species and GenBank accession numbers for sequences used in assay development.

Species	Latin	GenBank accession number(s)	Assay mismatches
Jaguar	*Panthera onca*	KM236783, NC022842, KF483864, KP202264	0
Domestic cat	*Felis catus*	U20753.1	14
Cougar	*Puma concolor*	NC016470	13
Bobcat	*Lynx rufus*	NC014456	16
Ocelot	*Leopardus pardalis*	NC028315	15
Margay	*Leopardus wiedii*	NC028318	16
Andean mountain cat	*Leopardus jacobita*	NC028322	15
Geoffroy's cat	*Leopardus geoffroyi*	NC028320	12
Kodkod	*Leopardus guigna*	NC028321	15
Oncilla	*Leopardus tigrinus*	NC028317	15
Pampas cat	*Leopardus colocolo*	NC028314	16
Jaguarundi	*Puma yagouroundi*	NC028311	13
Tiger	*Panthera tigris*	KP202268	13
Leopard	*Panthera pardus*	KP001507	12
African lion	*Panthera leo*	KP001506	12

Assay mismatches indicate number of base pair differences between sequence(s) and jaguar assay (primers and probe).

TABLE 2 | Assay oligonucleotide sequences.

Oligo	Sequence
Forward	5′-AACAATCGTCTAATCTCACTCCAACAG-3′
Reverse	5′-CCAACAGGTTTGTTGATCCAATG-3′
Probe	5′-FAM-CTTGGGCTCTAATACTC-MGB-NFQ-3′

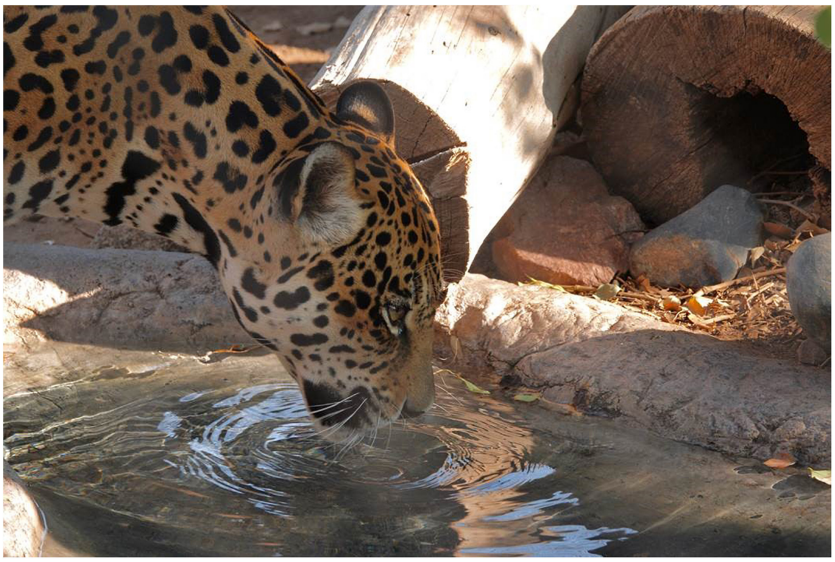


FIGURE 1 | Image of a jaguar drinking at the Phoenix Zoo, AZ, United States. Photo credit: David Bissegger.

Carim et al. (2016a). In steps 1 and 2, we doubled the amount of ATL and proteinase K, and incubated samples for 48 h. We doubled the amount of AL buffer in step 3 and added 400 µl of ethanol simultaneously. We repeated step 4 loading using a single spin column for each sample until all elution for a given sample had been processed through the spin column. Additionally, we loaded each filter using sterile forceps into a QIAshredder spin column and centrifuged for 2 min a 20,000 × g. The elution from the QIAshredder was also loaded and processed through the corresponding spin column for that sample. Between steps 5 and 6, we added an additional wash of 500 µl ethanol and centrifuged for 2 min at 20,000 × g. In step 6, we increased spin time to 4 min. In step 7, we eluted DNA in 100 µl of 70°C TE and allowed to incubate at room temperature for 10 min before the final centrifuge step. All environmental samples were extracted in a dedicated space where no high concentration sources of DNA are handled.

Quantitative PCRs contained 7.5 µl TaqMan Environmental Master Mix 2.0 (Thermo Fisher Scientific), 900 nM each primer, 250 nM hydrolysis probe (also known as "taqman" probe), 4 µl template DNA, and molecular grade water to a total volume of 15 µl. Environmental and scat samples (described below) also contained an internal positive control template and assay to test for the presence of PCR inhibitors (indicated by a > 1 C_T shift in amplification relative to the control samples; TaqMan Exogenous Internal Positive Control Kit from Thermo Fisher Scientific). On the PCR plate with environmental samples, a triplicate no-template control was included to test for contamination. All reactions were run in triplicate. We optimized primer concentrations by testing all possible combinations of forward and reverse primers at 100, 300, 600, and 900 nM concentration (n = 16 combinations) and selecting the combination with the lowest C_T value and highest end-point fluorescence (900:900 nM combination).

We also quantified jaguar mitochondrial DNA based on comparison with a standard curve. The standard curve was a dilution series (2, 10, 50, 250, 1,250, 6,250, and 31,250 copies per reaction) of a synthetic jaguar gene covering the target amplicon (gBlock; Integrated DNA Technologies) which was quantified on a Qubit 2.0 fluorometer (dsDNA Broad Range Kit; Invitrogen). Each dilution level was run in six replicates.

TABLE 3 | Jaguar mtDNA concentration for five samples collected from drinking water sources provided to captive animals.

Samples	Site	Jaguar mtDNA copies/L	Notes
1,2,3	Phoenix Zoo (concrete-lined pond)	33,990–40,260	Water changed 24 h prior to sample collection
4,5	Phoenix Zoo (metal water trough)	123,540–152,140	Water changed 24 h prior to sample collection
6	Belize Jungle Lodge (concrete-lined pond)	32,630	Water not recently changed

RESULTS AND DISCUSSION

In silico and *in vitro* testing showed our assay to be highly specific against potential non-target taxa in North America. Additional specificity testing of the assay may be necessary prior to application in the southern portion of the species' range. Although there was low-level amplification in one African lion tissue sample, we were able to verify that this was due to sample contamination using a synthetic gene covering the target amplicon. In our *in silico* assessment, the closest non-felid present in North America was *Equus caballus* (GenBank

accession# AY584828) with a total of 12 bp mismatches (8 primer mismatches, including the 3′ end of the forward primer, and 4 probe mismatches). Thus, non-felids are unlikely to cross-amplify. The assay also amplified all ten jaguar scat samples from Belize and had 100% amplification of the standard curve down to 10 copies/reaction (2/6 amplifications at a concentration of two copies/reaction). The standard curve slope implied an amplification efficiency of 85.3% (r^2 = 0.996). Based on this information, we estimated the Limit of Detection (LOD) to be 10 copies/reaction (minimum concentration with a 95% amplification rate) and the Limit of Quantification (LOQ) to be 50 copies/reaction (minimum concentration with a coefficient of variation < 35%) as described in Klymus et al. (2020).

All drinking water samples were strongly positive with a mean of 6,682 copies/reaction (range 3,251–14,254), or 66,820 copies/L of water. One of the five unknown Belize pond water samples was also positive, but at much lower concentration (8 copies/reaction or 80 copies/L of water), which is in line with concentrations of other terrestrial species detections from eDNA (e.g., Williams et al., 2018). Negative control reactions all showed no amplification.

The high concentrations of jaguar DNA in captive animal drinking water sources (over 30,000 copies/L) suggest that eDNA sampling may have reasonable detection probabilities in more natural settings, as long as sampling has occurred soon after the site was visited by an animal. Harper et al. (2019) and Sales et al. (2020) used eDNA sampling for detection of terrestrial mammals using a community-wide, metabarcoding approach. They found that low-density, wide-ranging taxa like jaguar tend to have lower detection rates than common and evenly distributed taxa. Aquatic organisms, even at low abundances, provide a constant input source of eDNA. In contrast, terrestrial species deposit their DNA at a drinking site over a very brief period of time. Thus, detection probably relies on sampling within hours or days of visitation—before the DNA has degraded beyond detectable levels.

In the studies described by Harper et al. (2019) and Sales et al. (2020), drinking water sources were abundant on the landscape. For sampling low-density jaguars, a key landscape characteristic influencing detection might be the number of drinking water sources within a single home-range. When there are only several water sources, the mean frequency with which each water source is visited is relatively high. When there are many, it could be days or weeks between visits to any one site. Hence, we see eDNA sampling for rare terrestrial species as being particularly promising in arid landscapes such as the southwestern United States and northern Mexico.

Camera trapping, in contrast to eDNA sampling, provides a more continuous view of habitat use. However, the initial equipment investment and ongoing maintenance costs may limit sampling effort over large scales. Other authors have suggested that eDNA sampling might be particularly useful in lower probability areas, such as at the presumed range periphery of a taxon, where investment in camera traps is impractical (Harper et al., 2019). Environmental DNA sampling might also become more useful as recent advances in continuous or autonomous eDNA sampling are refined and provide a more continuous view of DNA inputs (e.g., Yamahara et al., 2019; Kirtane et al., 2020).

We imagine eDNA sampling being a useful complement to existing sampling approaches and conveniently added to other monitoring efforts on the landscape. For example, technicians performing vegetation surveys in a remote area could also opportunistically collect water samples without much additional effort (e.g., <20 min per water sample in this study). Much of the cost of eDNA sampling, particularly in remote areas, comes from travel time (Smart et al., 2016). In the arid southwestern United States the relative cost for sample collection versus analysis might be even greater. Often in these remote and arid landscapes, the hydroperiod of ephemeral springs and tanks are unknown. There is a high level of risk during dry seasons of hiking into a remote habitat only to find that there is no water. This risk is mitigated when technicians are performing other fieldwork simultaneously.

Further savings might be gained when the genetic analyses for jaguar is conducted on eDNA samples initially collected to survey for the presence of other species. In this case, the cost of sample collection and DNA extraction is already covered. For example, the jaguar qPCR assay described here might complement a more community-wide survey effort with metabarcoding. Metabarcoding tends to be less sensitive than single-species qPCR and even reasonably abundant taxa might be missed (Harper et al., 2019; Sales et al., 2020). Targeted analysis used in conjunction with metabarcoding has been found to be a useful approach for picking up rare-but-important taxa (Simmons et al., 2016). Alternatively, an analysis for jaguar eDNA could even be included as part of a larger high throughput qPCR (HT-qPCR) panel which includes assays for detection of various aquatic and terrestrial species (Wilcox et al., 2020). In this case, the expense invested in analyzing samples for the presence of jaguar is likely <10% of the overall laboratory cost. Generally, our hope is that this new molecular tool can be leveraged in creative ways to build cost-effective study designs composed of multiple, complementary sampling approaches.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

ETHICS STATEMENT

The animal study was reviewed and approved by Arizona Center for Nature Conservation Research Committee.

AUTHOR CONTRIBUTIONS

TW, AC, KM, KZ, and MS conceived the study design. TW, JD, TF, DM, and KZ conducted sampling and analyses. All authors contributed to writing of the manuscript.

FUNDING

This work was funded by the SERDP (RC18-1348).

ACKNOWLEDGMENTS

We thank the Phoenix Zoo for assistance with sampling and the Banana Bank Lodge for allowing sample collection on their property. Samples collected in central Belize as part of the Valley of Peace Archaeology Project directed by Dr. Lisa J. Lucero with permission of the Institute of Archaeology, National Institute of Culture and History, and the Government of Belize. We thank the Mills Lab at the University of Montana and Bill Murphy at Texas A&M for sharing DNA samples from various non-target felids. This manuscript was improved by suggestions made by John Kronenberger and Torrey Rodgers.

REFERENCES

Atschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.

Brown, D. E., and González, C. A. L. (2000). Notes on the occurrences of jaguars in Arizona and New Mexico. *Southwestern Nat.* 45, 537–542. doi: 10.2307/3672607

Carim, K. J., Dysthe, J. C. S., Young, M. K., McKelvey, K. S., and Schwartz, M. K. (2016a). An environmental DNA assay for detecting Arctic grayling in the upper Missouri River basin, North America. *Conserv. Genet. Resour.* 8, 197–199. doi: 10.1007/s12686-016-0531-1

Carim, K. J., McKelvey, K. S., Young, M. K., Wilcox, T. M., and Schwartz, M. K. (2016b). *A Protocol for Collecting Environmental DNA Samples from Streams. Gen. Tech. Rep. RMRS-GTR-355*. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station, 18.

Chaves, P. B., Graeff, V. G., Lion, M. B., Oliveira, L. R., and Eizirik, E. (2012). DNA barcoding meets molecular scatology: short mtDNA sequences for standardized species assignment of carnivore noninvasive samples. *Mol. Ecol. Resour.* 12, 18–35. doi: 10.1111/j.1755-0998.2011.03056.x

Cristescu, M. E., and Hebert, P. D. N. (2018). Uses and misuses of environmental DNA in biodiversity science and conservation. *Annu. Rev. Ecol. Evol. Syst.* 49, 209–230. doi: 10.1146/annurev-ecolsys-110617-062306

Culver, M. (2016). *Jaguar Surveying and Monitoring in the United States (No. 2016-1095)*. Reston, VA: US Geological Survey. doi: 10.3133/ofr20161095

Franklin, T. W., McKelvey, K. S., Golding, J. D., Mason, D. H., Dysthe, J. C., Pilgrim, K. L., et al. (2019). Using environmental DNA methods to improve winter surveys for rare carnivores: DNA from snow and improved noninvasive techniques. *Biol. Conserv.* 229, 50–58. doi: 10.1016/j.biocon.2018.11.006

Harper, L. R., Handley, L. L., Carpenter, A. I., Ghazali, M., DiMuri, C., Mcgregor, C. J., et al. (2019). Environmental DNA (eDNA) metabarcoding of pond water as a tool to survey conservation and management priority mammals in the UK. *Biol. Conserv.* 238:108225. doi: 10.1016/j.biocon.2019.108225

Kirtane, A., Atkinson, J. D., and Sassoubre, L. (2020). Design and validation of passive environmental DNA samplers using granular activated carbon and Montmorillonite Clay. *Environ. Sci. Technol.* 54, 11961–11970. doi: 10.1021/acs.est.0c01863

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

Menchaca, A., Rossi, N. A., Froidevaux, J., Dias-Freedman, I., Caragiulo, A., Wultsch, C., et al. (2019). Population genetic structure and habitat connectivity for jaguar (*Panthera onca*) conservation in Central Belize. *BMC Genet.* 20:100. doi: 10.1186/s12863-019-0801-5

R Core Development Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rodgers, T. (2017). Proper fin-clip sample collection for molecular analyses in the age of eDNA. *J. Fish Biol.* 91, 1265–1267. doi: 10.1111/jfb.13485

Rodgers, T., and Mock, K. E. (2015). Drinking water as a source of environmental DNA for the detection of terrestrial wildlife species. *Conserv. Genet. Resour.* 7, 693–696. doi: 10.1007/s12686-015-0478-7

Sales, N. G., McKenzie, M. B., Drake, J., Harper, L. R., Browett, S. S., Coscia, I., et al. (2020). Fishing for mammals: landscape-level monitoring of terrestrial and semi-aquatic communities using eDNA from riverine systems. *J. Appl. Ecol.* 57, 707–716. doi: 10.1111/1365-2664.13592

Sepulveda, A. J., Nelson, N. M., Jerde, C. L., and Luikart, G. (2020). Are environmental DNA methods ready for aquatic invasive species management? *Trends Ecol. Evol.* 35, 668–678. doi: 10.1016/j.tree.2020.03.011

Simmons, M., Tucker, A., Chadderton, W. L., Jerde, C. L., and Mahon, A. R. (2016). Active and passive environmental DNA surveillance of aquatic invasive species. *Can. J. Fish. Aquat. Sci.* 73, 76–83. doi: 10.1139/cjfas-2015-0262

Smart, A. S., Weeks, A. R., van Rooyen, A. R., Moore, A., McCarthy, M. A., and Tingley, R. (2016). Assessing the cost-efficiency of environmental DNA sampling. *Methods Ecol. Evol.* 7, 1291–1298. doi: 10.1111/2041-210x.12598

Wilcox, T. M., McKelvey, K. S., Young, M. K., Engkjer, C., Lance, R. F., Lahr, A., et al. (2020). Parallel, targeted analysis of environmental samples via high-throughput quantitative PCR. *Environ. DNA* 2, 544–553. doi: 10.1002/edn3.80

Williams, K. E., Huyvaert, K. P., Vercauteren, K. C., Davis, A. J., and Piaggio, A. J. (2018). Detection and persistence of environmental DNA from an invasive, terrestrial mammal. *Ecol. Evol.* 08, 688–695. doi: 10.1002/ece3.3698

Wright, E. S. (2016). Using DECIPHER v2.0 to analyze big biological sequence data in R. *R J.* 8, 352–359. doi: 10.32614/rj-2016-025

Wultsch, C., Caragiulo, A., Dias-Freedman, I., Quigley, H., Rabinowitz, S., and Amato, G. (2016). Genetic diversity and population structure of Mesoamerican jaguars (*Panthera onca*): implications for conservation and management. *PLoS One* 11:e0162377. doi: 10.1371/journal.pone.0162377

Yamahara, K. M., Preston, C. M., Birch, J., Walz, K., Marin, R., Jensen, S., et al. (2019). In situ autonomous acquisition and preservation of marine environmental DNA using an autonomous underwater vehicle. *Front. Mar. Sci.* 6:373. doi: 10.3389/fmars.2019.00373

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 17 May 2021
doi: 10.3389/fevo.2021.612107

Environmental DNA Metabarcoding for Simultaneous Monitoring and Ecological Assessment of Many Harmful Algae

Emily Jacobs-Palmer[1], *Ramón Gallego*[1,2], *Kelly Cribari*[1], *Abigail G. Keller*[1] *and Ryan P. Kelly*[1]*

[1] University of Washington, Seattle, WA, United States, [2] Northwest Fisheries Science Center, National Oceanic and Atmospheric Administration, Seattle, WA, United States

Harmful algae can have profound economic, environmental, and social consequences. As the timing, frequency, and severity of harmful algal blooms (HABs) change alongside global climate, efficient tools to monitor and understand the current ecological context of these taxa are increasingly important. Here we employ environmental DNA metabarcoding to identify patterns in a wide variety of potentially harmful algae and associated ecological communities in the Hood Canal of Puget Sound in Washington State, USA. Tracking trends of occurrence in a series of water samples over a period of 19 months, we find algal sequences from genera with harmful members in a majority of samples, suggesting that these groups are routinely present in local waters. We report patterns in variants of the economically important genus *Pseudo-nitzschia* (of which some members produce domoic acid; family Bacillariaceae), as well as multiple potentially harmful algal taxa previously unknown or poorly documented in the region, including a cold-water variant from the genus *Alexandrium* (of which some members produce saxitoxin; family Gonyaulacaceae), two variants from the genus *Karlodinium* (of which some members produce karlotoxins; family Kareniaceae), and one variant from the parasitic genus *Hematodinium* (family Syndiniaceae). We then use data on environmental variables and the biological community surrounding each algal taxon to illustrate the ecological context in which they are commonly found. Environmental DNA metabarcoding thus simultaneously (1) alerts us to potential new or cryptic occurrences of algae from harmful genera, (2) expands our knowledge of the co-occurring conditions and species associated with the growth of these organisms in changing marine environments, and (3) suggests a pathway for multispecies monitoring and management moving forward.

Keywords: harmful algal bloom, environmental DNA, ecological model, metabarcoding, *Alexandrium*, *Pseudo-nitzchia*, *Hematodinium*, *Karlodinium*

OPEN ACCESS

Edited by:
Hiroki Yamanaka, Ryukoku University, Japan

Reviewed by:
Nansheng Chen, Institute of Oceanology, Chinese Academy of Sciences (CAS), China
Laurence J. Clarke, Antarctic Climate and Ecosystems Cooperative Research Centre (ACE CRC), Australia

***Correspondence:**
Ryan P. Kelly
rpkelly@uw.edu

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *30 October 2020*
Accepted: *12 April 2021*
Published: *17 May 2021*

Citation:
Jacobs-Palmer E, Gallego R, Cribari K, Keller AG and Kelly RP (2021) Environmental DNA Metabarcoding for Simultaneous Monitoring and Ecological Assessment of Many Harmful Algae. Front. Ecol. Evol. 9:612107. doi: 10.3389/fevo.2021.612107

1. INTRODUCTION

Harmful algae and associated blooms create environmental, health, and economic challenges at a global scale, causing mass die-offs in ecosystems from de-oxygenation (Gobler, 2020; Griffith and Gobler, 2020), multiple types of poisoning in humans (Trainer et al., 2013), and significant losses of revenue for the aquaculture industry (Trainer and Yoshida, 2014; Diaz et al., 2019). For

these reasons, local, national, and international governing bodies organize and fund monitoring programs to track HABs and identify the conditions that lead to their occurrence (Graneli and Lipiatou, 2002; Trainer, 2002; Lopez et al., 2008; Moestrup et al., 2020). In addition, changing marine environments appear to be causing increases in the duration, frequency, and severity of HABs globally in association with rising temperatures and declining pH (Gattuso et al., 2015a; Gobler et al., 2017; Gobler, 2020).

Hood Canal, a natural glacial fjord within the Puget Sound of Washington, USA, is a useful natural system in which to study the ecology of harmful algae and likely future changes to their patterns of occurrence. Surface temperatures of the region have risen 1.0°C since the 1950s, dissolved oxygen levels are below 5 mg/L in deeper sections of the sound, and pH has dropped by 0.05–0.15 units since pre-industrial era (~1750) (Feely et al., 2010; Busch et al., 2013; Mauger et al., 2015). Warmer temperatures and longer durations of warm conditions will create larger windows of growth for some HABs moving forward (Moore et al., 2011; Mauger et al., 2015; Brosnahan et al., 2020), with ocean acidification potentially exacerbating the impacts of these blooms by further increasing the toxicity and growth of some harmful algal species (Fu et al., 2012; Field et al., 2014; Raven et al., 2020).

Harmful algae fall into four primary categories: diatoms, dinoflagellates, haptophytes, and raphidophytes. Of particular concern locally are select diatoms from the genus *Pseudo-nitzschia* and dinoflagellates from the genera *Alexandrium*, *Gonyaulax*, and *Protoceratium*, each of which contain species that produce toxins capable of accumulating in shellfish grazers (Shimizu et al., 1975; Satake et al., 1998; Cembella et al., 2000; Trainer et al., 2009, 2016). When consumed by humans, the toxins then cause symptoms ranging from amnesia to paralysis, and can be deadly (Ferrante et al., 2013; Grattan et al., 2016). Additional harmful algae of concern are fish-killing species such as the diatom *Chaetoceros concavicornis* and the raphidophyte *Heterosigma akashiwo* (Yang and Albright, 1994; Khan et al., 1997). There is no current ensemble testing protocol for all of these local problematic algae, and both human-mediated transport and warming-related range shifts are likely to introduce additional taxa. For example, there is recent evidence that the toxic dinoflagellate *Karenia mikimotoi* (family Kareniaceae; described from Japan and also occurring in the North Atlantic) is now present along the west coast of North America, specifically off of Alaska and California (National Centers for Coastal Ocean Science., 2014).

Because the effects of harmful algae are wide-ranging and potentially devastating (Lewitus et al., 2012; Moore et al., 2019), monitoring these organisms and the environmental conditions with which they are associated has long been a public health priority in Puget Sound and in many locations around the world. Efforts to track blooms of harmful algae have historically relied on the work of skilled taxonomists using microscopic visual analysis of cells to identify species (e.g., Yang et al., 2000; Lapworth et al., 2001). More recently, satellite spectrographic data (Tomlinson et al., 2004; Ahn et al., 2006), molecular assays for toxin and cell detection (Pierce and Kirkpatrick, 2001; Groben and Medlin, 2005; Töbe et al., 2010; Murray et al., 2011), and flow cytometry coupled with machine learning (Campbell et al., 2010) have been employed to detect and track HAB taxa.

Adding to the list of technological advances for monitoring are two types of genetic techniques that rely on environmental DNA (eDNA) present in the water to classify and assess the abundance of harmful algae: quantitative Polymerase Chain Reaction (qPCR) and DNA metabarcoding (Al-Tebrineh et al., 2010; Erdner et al., 2010; Antonella and Luca, 2013; Grzebyk et al., 2017; Ruvindy et al., 2018; Jacobs-Palmer et al., 2020). The former method tracks known taxa individually and requires substantial sequence data to design species-specific primers and/or probes; the latter method involves PCR with less-specific primers to generate amplicons from a broad swath of taxa at a common locus. This second method, metabarcoding, allows detection and even quantification of many taxa simultaneously.

Because the specific target organisms need not be chosen *a priori*, metabarcoding may uncover taxa unexpected in the study region, and in addition, can reveal a cross-section of the biological community surrounding any particular group of interest (Deiner et al., 2017; Taberlet et al., 2018). Previous work has described the ecological context of harmful algae and their blooms using environmental covariates (e.g., Wells et al., 2015; Banerji et al., 2019), as well as assessments of bloom-associated taxa (typically other microorganisms or viruses) (e.g., Buskey, 2008; Loureiro et al., 2011), although the labor required for traditional survey methodology has limited the breadth of contextual taxa included in these studies.

Here, we couple environmental observations with eDNA metabarcoding to examine the ecological and biological context of taxa from dozens of harmful algae-containing genera across 19 months of sampling at 10 locations in Hood Canal. Specifically, we use a single primer set amplifying Cytochrome Oxidase I (COI) to detect focal algal taxa from within the economically and socially relevant genera *Alexandrium* (dinoflagellate family Gonyaulacaceae), *Hematodinium* (dinoflagellate family Syndiniaceae), *Karlodinium* (dinoflagellate family Kareniaceae), and *Pseudo-nitzschia* (diatom family Bacillariaceae), and to simultaneously survey hundreds of other eukaryotic taxa comprising their biological milieu. We note that our choice of locus and primers (Leray et al., 2013) suitable to amplify and identify a broad swath of eukarya to the level of genus allows us to cast a wide net in our search for organisms co-occurring with these algae but not to resolve individual species; alternative metabarcoding assays will yield different but equally valid views of the community, depending on survey goals. With these sequences in hand, we create exploratory models of the associations of individual algal lineages with key environmental variables, and subsequently improve the predictive value of these models by including co-occurring (both algal and non-algal) lineages as possible indicator taxa. Given the relatively small sample sizes in the current dataset, we see these models as demonstrating the power of community-based genetic monitoring to better understand harmful algal dynamics, rather than as an end-product immediately useful for broad application. In summary, while individual management teams must determine parameters such as the choice of metabarcoding primer set(s), depth of sampling, and model sets to test based on

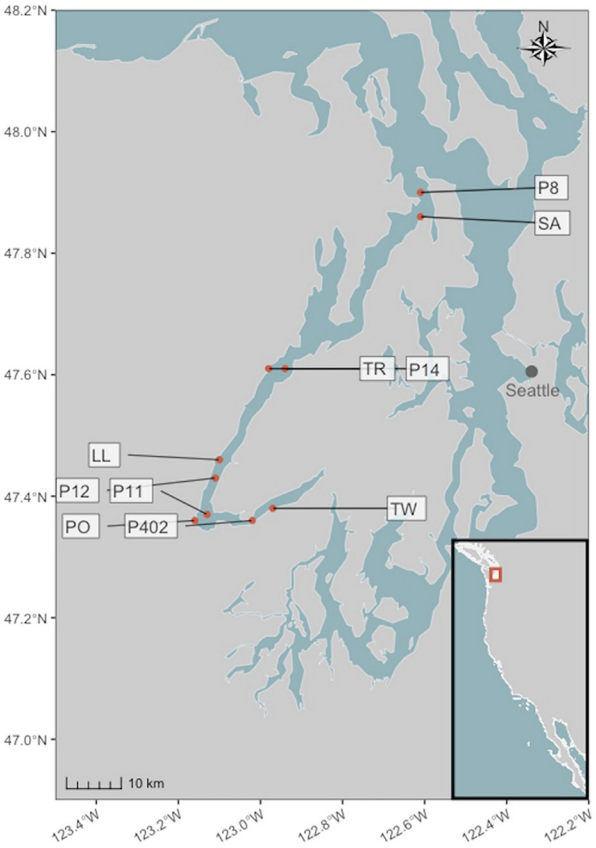


FIGURE 1 | Intertidal and nearshore sampling locations in Hood Canal, Washington, USA. Site abbreviations are described in the text, and coordinates are given in **Supplementary Table 1**. Inset map shows the Pacific coast of the continental United States.

local needs and knowledge, we provide a basic framework for the use of eDNA metabarcoding to improve our understanding and management of harmful algae, both within the Puget Sound and globally.

2. MATERIALS AND METHODS

2.1. Environmental DNA Sampling and Measuring Environmental Variables

We sampled seawater for eDNA from 10 sites within Hood Canal, a natural glacial fjord in Puget Sound, Washington, USA, and identified a broad range of potential harmful algal taxa and simultaneously survey the surrounding biological community. Five sampling sites were intertidal, and five were nearby nearshore locations at the approximate center of the fjord.

For each site/date combination at intertidal locations (Salisbury Point County Park (SA), Triton Cove State Park (TR), Lilliwaup Tidelands State Park (LL), Potlatch State Park (PO), and Twanoh State Park (TW); see **Figure 1**; **Supplementary Table 1** for location coordinates), we collected three 1 L samples of water (biological replicates) from immediately below the surface using a bleach-cleaned (10% bleach for 10 min) plastic bottle held at the end of a 1.7 m pole. We sampled intertidal locations every 1–2 months between March 2017 and July 2018 (see **Supplementary Table 1** for sampling dates). At these same stations and simultaneous with eDNA sampling, we collected one 120 ml water sample from each site and poisoned it with 0.1 ml of saturated $HgCl_2$ for carbonate chemistry analysis including pH (Dickson et al., 2007). We also collected *in situ* measurements of temperature and salinity using a handheld multiprobe (Hanna Instruments, USA) and a portable refractometer. We characterized sample carbonate chemistry by measuring Total Alkalinity (TA; open-cell automated titration based on a Dosimat plus (Metrohm AG) as part of a custom system assembled by Andrew Dickson (University of California San Diego) and used in the laboratory of Alex Gagnon at the University of Washington) and Dissolved Inorganic Carbon (DIC; Apollo Instruments, USA; CO2 extraction system with 10% (v/v) phosphoric acid). Both measurements were calibrated and validated with certified reference material from the Scripps Oceanographic Institute. Using DIC and TA, we calculated pH and the remaining carbonate system parameters with the R package "seacarb" (Gattuso et al., 2015b), removing a single outlier sample from the dataset used for environmental modeling (see below) due to an unreasonably low pH value (<7.5).

For nearshore locations, we sampled a selection of stations (P8, P14, P12, P11, P402) surveyed by the Washington Ocean Acidification Center during triannual cruises (see **Figure 1**; **Supplementary Table 1** for location coordinates). The samples used here were collected in September 2017 (P12 and P402), April 2018 (P8, P12, and P402), and September 2018 (P14, P12, and P11); see **Supplementary Table 1**. At each station, a CTD was deployed with twelve Niskin bottles, and collected data on temperature, salinity, and pH (Alin et al., 2019a,b,c) in addition to water for a single eDNA sample from immediately below the surface.

For both intertidal and nearshore locations, we filtered 500 mL of each water sample for eDNA with a cellulose acetate filter (47 mm diameter, 0.45 μm pore size), and preserved this filter in Longmire buffer until DNA extraction (Renshaw et al., 2015). Many unmeasured variables influence planktonic communities (e.g., nutrients, sunlight, and wave energy); nevertheless the minimal set of parameters we analyzed here clearly distinguished communities and was adequate for the purposes of assessing temporal and spatial trends. Our purpose was to describe patterns of focal algal taxa over space and time, along with the environmental and ecological contexts in which they occurred, rather than to test any particular mechanism by which these algae might respond to different environmental parameters.

2.2. Extraction, Amplification, and Sequencing

To extract DNA from sample filters, we used a phenol:chloroform:isoamyl alcohol protocol (modified from Renshaw et al., 2015). To maximize extraction efficiency and minimize co-extraction of inhibitors, we incubated filter

membranes at 56°C for 30 min before adding 900 μL of phenol:chloroform:isoamyl alcohol and shaking vigorously for 60 s. We conducted two consecutive chloroform washes by centrifuging at 14,000 rpm for 5 min, transferring the aqueous layer to 700 μL chloroform, and shaking vigorously for 60 s. After a third centrifugation, we transferred 500 μL of the aqueous layer to tubes containing 20 μL 5 molar NaCl and 500 μL 100% isopropanol, and froze these at −20°C for approximately 15 h. Finally, we centrifuged samples at 14,000 rpm for 10 min, poured off or pipetted out any remaining liquid, and dried in a vacuum centrifuge at 45°C for 15 min. We resuspended the eluate in 200 μL water, and used 1 μL of diluted DNA extract (between 1:10 and 1:400) as template for PCR.

To identify a wide variety of metazoan taxa, including algal genera with harmful members and their surrounding biological communities using eDNA, we amplified a ~315 base pair segment of the Cytochrome Oxidase I (COI) using primers described in Leray et al. (2013) (although these primers were designed for metazoa, they consistently amplify a broad swath of eukaryotic phyla; (see, e.g., Gallego et al., 2020; Jacobs-Palmer et al., 2020). To distinguish technical from biological variance and to quantify each, we ran and sequenced in triplicate PCR reactions from each of the samples (i.e., individual bottles of water). For multiplex sequencing on an Illumina MiSeq, we followed a two-step PCR protocol (O'Donnell et al., 2016) with redundant 3' and 5' indexing. In the first step, we used a PCR reaction containing 1X HotStar Buffer, 2.5 mM MgCl2, 0.5 mM dNTP, 0.3 μM of each primer, and 0.5 units of HotStar Taq (Qiagen Corp., Valencia, CA, USA) per 20 μL reaction. The PCR protocol for this step consisted of 40 cycles, including an annealing touchdown from 62 to 46°C (−1°C per cycle), followed by 25 cycles at 46°C. In the second step, we added identical 6 base-pair nucleotide tags to both ends of our amplicons, with unique index sequences for each individual PCR reaction. We allowed for no sequencing error in these tags; only sequences with identical tags on both the forward and reverse read-directions survived quality control. This gave us high confidence in assigning amplicons back to individual field samples.

We generated amplicons from extracted tissue with the same replication scheme for positive control kangaroo (genus *Macropus*), selected because this genus is absent from the sampling sites and common molecular biology reagents, but amplifies well with the primer set (Leray et al., 2013) used in this study. We could therefore use positive control samples to identify possible cross-contamination: reads from other taxa that appear in these samples allow us to estimate and account for the proportion of sequences that are present in the incorrect PCR reaction (see section 2.3 below). We also amplified negative controls (molecular grade water) in triplicate alongside environmental samples and positive controls, and verified by gel electrophoresis and fluorometry that these PCR reactions contained no visible band of DNA (see Kelly et al., 2018 for a discussion of the merits of sequencing positive and not negative controls).

To prepare libraries of replicated, indexed samples and positive controls, we followed manufacturers' protocols (KAPA Biosystems, Wilmington, MA, USA; NEXTflex DNA barcodes; BIOO Scientific, Austin, TX, USA). We then performed sequencing on an Illumina MiSeq (250–300 bp, paired-end) platform in seven different sets of samples: six for the intertidal dataset and one for the nearshore dataset.

2.3. Bioinformatics

We followed updated versions of previously published procedures for bioinformatics, quality control, and decontamination (Kelly et al., 2018). This protocol uses a custom Unix-based script (available at https://github.com/ramongallego/demultiplexer_for_DADA2) calling third-party programs to perform initial quality control on sequence reads from all seven runs combined, demultiplexing sequences to their sample of origin and clustering unique variants into amplicon sequence variants (ASVs) (Martin, 2011; Callahan et al., 2016).

Specifically, to address possible cross-sample contamination (see Schnell et al., 2015), we subtracted the maximum proportional representation of each ASV across all positive control samples (see section 2.2 above) from the respective ASV in field samples. We estimated the probability of ASV occurrence by performing occupancy modeling (Royle and Link, 2006; Lahoz-Monfort et al., 2016). Following Lahoz-Monfort et al. (2016) and using the full Bayesian code for package rjags (Plummer et al., 2016) provided by those authors, we modeled the probability of occupancy (i.e., true presence) for each of the unique sequence variants in our dataset. We treated replicate PCR reactions of each water bottle as independent trials, estimating the true-positive rate of detection (P_{11}), false-positive rate (P_{10}), and occupancy probability (psi, ψ) in a binomial model. We then used these parameters to estimate the overall likelihood of occupancy (true presence) for each ASV; those with low likelihoods (<80%) were deemed unlikely to be truly present in the dataset, and therefore culled. We removed samples whose PCR replicates were highly dissimilar by calculating the Bray-Curtis dissimilarity amongst PCR replicates from the same bottle of water and discarding those with distance to the sample centroid outside a 95% confidence interval. The result was a dataset of $3.98 * 10^8$ reads from 5,275 unique ASVs. Lastly, to collapse variants likely due to PCR error, we converted ASVs to operational taxonomic units (OTUs) by clustering with SWARM (Mahe et al., 2015) with a radius of 1. All bioinformatic and analytical code is included in a GitHub repository (https://github.com/ramongallego/Harmful.Algae.eDNA), including the details of parameter settings in the bioinformatics pipelines used. Sequence and annotation information are included as well, and the former are deposited and publicly available in GenBank Project ID: PRJNA699686; individual FASTQ files: SRR13659629-639 & SRR13685891-947.

2.4. Taxonomy

We performed the taxonomic identification using a CRUX-generated database for the Leray fragment of the COI gene (see section 2.2 above), querying that database with a Bowtie2 algorithm (as described in Curd et al., 2019). The algorithm classifies the query sequence to the last common ancestor of ambiguously classified sequences. Only matches with a bootstrap support greater than 90% were kept. Here, we assigned taxonomy

at the level of genus, rather than species, for two main reasons. First, for some taxa, variation may not be sufficient to distinguish species within a genus, and second, representation of local species in the databases used may not be complete, leading to the mis-assignment of sequences to their nearest represented neighbor. We denoted different lineages within genera using three-character abbreviation derived from the sequence variants themselves. Full sequences for each variant are provided in **Supplementary Table 2**, and within-genera DNA sequence alignments (Sievers and Higgins, 2014) are provided in **Supplementary Table 3**. To assess similarity of algal lineages from genera with harmful members, we translated nucleotide sequences with the ExPASy Bioinformatics Resource Portal Translate tool using the mold, protozoan, and coelenterate mitochondrial, mycoplasma/spiroplasma genetic code (Gasteiger et al., 2003), and aligned both nucleotide and amino acid sequences within genera using the Clustal Omega Multiple Sequence Alignment tool (Sievers and Higgins, 2014).

2.5. Taxon Distributions and Environmental and Biological Context

We plotted detection/non-detection for each focal algal taxon across all sampling events from both intertidal and nearshore eDNA collections in our time series between March 2017 and September 2018. We note that while detection almost certainly indicates true presence of a taxon, given our stringent standards for sequence quality and taxonomic assignment, non-detection does not absolutely confirm absence; however, given amplicon counts in the hundreds to tens of thousands for focal sequences, detection/non-detection should at the very least track biomass (see Kelly et al., 2019), and therefore be a useful metric for modeling. To explore the ways in which environmental variables were associated with the detection or non-detection of our focal harmful algal taxa, we compared logistic-regression models using taxon detection as outcome, and combinations of three environmental variables (temperature, pH, and salinity) as predictors. We also fit a variety of models in a Bayesian hierarchical framework, where the slopes of predictors and intercepts could vary by season; for purposes of the models, we designated April through September as being the "warm" season, and other months the "cool" season, based on the environmental data and the approximate congruence of these seasons' boundaries with the fall and spring equinoxes. Rather than mechanically testing all possible combinations of models, we proposed models that were reasonable given the observed patterns of occurrences; in total, this resulted in between five and 17 models per taxon. Given many possible predictor variables, developing a useful model without overfitting can be a challenge. To combat this, we compared models using the widely applicable information criterion (WAIC) (Watanabe, 2010), which makes no assumptions about the shape of the posterior probability distribution and – like information criteria in general – penalizes more complex models. Moreover, WAIC quickly approximates the results of leave-one-out cross-validation (McElreath, 2020) to estimate out-of-sample model performance. Following model selection using WAIC, we reported the in-sample model accuracy for reference.

We further note the difficulty of predicting relatively rare events with logistic regression; even using information criteria such as WAIC does not completely guard against overfitting, and where the number of candidate predictor variables approaches the number of observations of an event (say, the occurrence of an algal taxon), parameter estimates become unreliable (Peduzzi et al., 1996). For our multivariate models, we use a lasso prior to penalize these parameters and shrink them toward zero, making the estimate of their effect more conservative; this approach has shown good success in taming rare-events logistic models (Pavlou et al., 2016; van Smeden et al., 2019). Even so, recognizing the limited sample size with which we are working here, we regard these models as exploratory rather than as definitive descriptions of associations in nature. As more data of this kind becomes available and the number of observations of any given taxon grows, we can develop more robust models.

To determine the species most closely associated with our focal algal taxa, we performed canonical analysis of principal coordinates (CAP) for each focal variant by implementing the capscale function in vegan (Oksanen et al., 2013), which revealed the degree to which other taxa in the surrounding biological communities could be associated with detection of the focal algal taxon. Using this ordination technique avoids the problem of testing each co-occurring taxon for significant associations with our focal algae, thereby removing the need to statistically correct for multiple comparisons. We then used these putative indicator taxa as predictors in a second round of logistic regressions, adding only the single most-strongly associated taxon as a separate term to the best-fit environmental models for each of our focal lineages (above). Such contextual ecological information is useful to the extent that it helps to predict the occurrence of our focal algal taxa.

3. RESULTS

3.1. Taxonomy

Environmental DNA metabarcoding of 63 samples from five intertidal and five nearshore locations in Hood Canal, Washington, United States revealed a total of 605 unique operational taxonomic units (OTUs) for which we were able to assign taxonomy to 262 distinct genera. Of these, exactly 100 OTUs were assigned to genera that are known to contain harmful algae (Horner and Postel, 1993; Horner et al., 1997; Trainer et al., 2016; Moestrup et al., 2020). These taxa are members of four main groups—diatoms, dinoflagellates, haptophytes, and raphidophytes—and represent seventeen genera (diatoms: *Chaetoceros*, *Nitzschia*, *Pseudo-nitzschia*; dinoflagellates: *Alexandrium*, *Dinophysis*, *Gonyaulax*, *Gymnodinium (Akashiwo)*, *Hematodinium*, *Heterocapsa*, *Karlodinium*, *Prorocentrum*, *Woloszynskia*; haptophytes: *Chrysocromulina*, *Phaeocystis*; raphidophytes: *Chattonella*, *Heterosigma*, *Pseudochattonella*; See **Supplementary Table 2** for a complete list of taxonomic assignments and COI sequences, as well as total read counts for each variant from an algal genus with harmful members).

Taxa that occur only in small numbers of samples lack sufficient observations to allow robust tests for association with environmental variables. Consequently, we focus hereafter on the OTUs detected in at least 10 percent of samples (minimum 7 occurrences out of 63 samples), an adequate sample size to compare with environmental variables and biological context. This subset of sequences included 191 total variants, 37 of which belong to genera containing harmful algal members (**Table 1**), and the rest to other organisms in the biological community. These potentially harmful algal variants belong to 12 genera containing differing degrees of sequence variation, with some such as *Hematodinium* represented by a single DNA and protein sequence, and others such as *Nitzschia* represented by a much larger number of DNA (10) and amino acid (5) variants. Members of each of the algal genera represented here exhibit varying degrees and types of toxicity or harm, ranging from physical irritation of fish gill tissue to production of toxins dangerous to human health (**Table 1**; Simonsen and Moestrup, 1997; Lindberg et al., 2005; Stentiford and Shields, 2005; Kotaki et al., 2006; Peperzak and Poelman, 2008; Skjelbred et al., 2011; Place et al., 2012; Trainer et al., 2016; Cho et al., 2017).

Amplicon sequences from environmental samples cannot be matched directly with phenotypes, by definition, and taxonomic annotations of those sequences depend upon adequate reference material. Acknowledging both the intra-specific variation that exists at the COI locus and the incompleteness of the GenBank reference database for many of these groups, we treat polymorphism within a putative genus as being ambiguous: these variants may be intra-specific, or they may represent distinct evolutionary lineages. For these reasons, we conservatively perform analyses on the sequence variants themselves (denoted with their genus names and a three-character code that abbreviates the hash of the unique nucleotide sequence) rather than making assumptions regarding their status as haplotypes vs. species.

The variants within a few genera containing harmful members are of particular interest, due to the nature of their toxicity (*Alexandrium*), to their unexpected presence in the study region (*Hematodinium* and *Karlodinium*), or to their potential economic impact (*Pseudo-nitzschia*). For these reasons, we chose to examine aspects of their taxonomy, distribution, and ecology in greater detail. We first examined COI sequences for these taxa from our original metabarcoding effort, noting that both *Alexandrium* and *Karlodinium* genera were each represented by two sequence variants, *Pseudo-nitzschia* by three sequence variants, and *Hematodinium* by a single sequence variant (**Table 1**). Amino acid translation revealed that the two *Alexandrium* OTUs differed by a single amino acid substitution, the two *Karlodinium* OTUs differed by five substitutions, and although two of the three *Pseudo-nitzschia* sequences (Pseudonitzschia_4e5 and Pseudonitzschia_d36) were identical in amino acid sequence, they differed from the third (Pseudonitzschia_d40) by two substitutions. The results below focus on these eight sequence variants, which we hereafter refer to as our "focal lineages."

3.2. Taxon Distributions in Space and Time

To identify the seasonal and spatial distributions of taxa from our eight focal lineages, we next visualized their patterns of detection and non-detection in time and space (**Figure 2**). The variants assigned to *Alexandrium*, Alexandrium_3fc and Alexandrium_2b2, had completely non-overlapping distributions in space and time, never appearing in the same sampling event. Alexandrium_3fc appeared solely in the warm (April-September) season (25 of 43 warm season samples vs. 0 of 20 cool season samples; $p < 0.001$) whereas Alexandrium_2b2 appeared primarily in the cool (October-March) season (1 of 43 warm season samples vs. 7 of 20 cool season samples; $p < 0.001$). In contrast, the single variant assigned to *Hematodinium*, Hematodinium_449, was not significantly seasonal (9 of 43 warm season samples and 3 of 20 cool season samples; $p = 0.742$); neither were the two variants assigned to *Karlodinium*, Karlodinium_8ed and Karlodinium_a27 (Karlodinium_8ed: 14 of 43 warm season samples and 7 of 20 cool season samples, $p \sim 1$; Karlodinium_a27: 6 of 43 warm season samples and 5 of 20 cool season samples, $p = 0.322$). One of the three variants assigned to *Pseudo-nitzschia* occurred significantly more frequently in the warm than in the cool season (Pseudonitzchia_d36: 10 of 43 warm season samples vs. 0 of 20 cool season samples, $p = 0.023$), while the others did not (Pseudonitzchia_4e5: 8 of 43 warm season samples and 1 of 20 cool season samples, $p = 0.244$; Pseudonitzchia_d40, 7 of 43 warm season samples and 4 of 20 in cool season samples; $p = 0.737$).

All together, we detected at least one of the eight focal sequence variants in 51 out of 63 sampling events (81%), indicating that these algal taxa are present more often than not in local waters.

3.3. Environmental Context

The above results suggest both *Alexandrium* lineages and at least one of the three *Pseudo-nitzschia* lineages are associated with environmental conditions that change seasonally, while the others are more stochastic in space and time. For each focal taxon, we fit a series of logistic-regression models (see section 2) describing taxon occurrence as a function of sea-surface temperature, pH, and salinity, both with and without a global intercept term (see **Supplementary Table 4** for a complete list of models tested, by taxon). A subset of our models was also hierarchical, allowing slopes to vary according to season, and we used WAIC to identify the best-fit models of those tested (**Table 2**).

Three of these models involve multiple environmental parameters, making them difficult to adequately visualize in two dimensions. Nevertheless, plotting the probability of taxon detection as a function of the single most-influential environmental variable and capturing seasonal variation in slope when models are hierarchical illustrates the degree to which the models do (or do not) explain the observed variance in focal algal taxa (**Figure 3**).

Although environmental covariates sea-surface temperature, pH, and salinity are associated with the detection of our eight focal lineages, accuracy of these models varies widely (**Table 2**).

TABLE 1 | Potential harmful algal taxa identified by eDNA in at least ten percent of samples from in Hood Canal, WA.

Type	Genus	DNA variants	Protein variants	Toxicity (target)	Sampling location
Diatom	*Chaetoceros*	8	5	Gill irritation (finfish)	Intertidal; Nearshore
Diatom	*Nitzschia*	10	5	Domoic acid and derivatives (human)	Intertidal
Diatom	*Pseudo-nitzschia*	3	2	Domoic Acid (human)	Intertidal; Nearshore
Dinoflagellate	*Alexandrium*	2	2	Saxitoxin (human)	Intertidal; Nearshore
Dinoflagellate	*Hematodinium*	1	1	Parasitism (crab)	Intertidal
Dinoflagellate	*Heterocapsa*	2	2	Haemolysis (shellfish)	Intertidal; Nearshore
Dinoflagellate	*Karlodinium*	2	2	Karlotoxin (human)	Intertidal; Nearshore
Dinoflagellate	*Woloszynskia*	1	1	Reddening of water (general)	Intertidal
Haptophyte	*Chrysochromulina*	3	3	Haemolysis (shellfish)	Intertidal; Nearshore
Haptophyte	*Phaeocystis*	2	2	Oxygen depletion (general)	Intertidal; Nearshore
Raphidophyte	*Chattonella*	2	1	Reactive oxygen species (finfish)	Intertidal; Nearshore
Raphidophyte	*Pseudochattonella*	1	1	Gill irritation (finfish)	Intertidal

Type of harmful algae and genus are given, as well as the number of DNA and protein variants, toxicity, and sampling location(s) for member(s) of that genus.

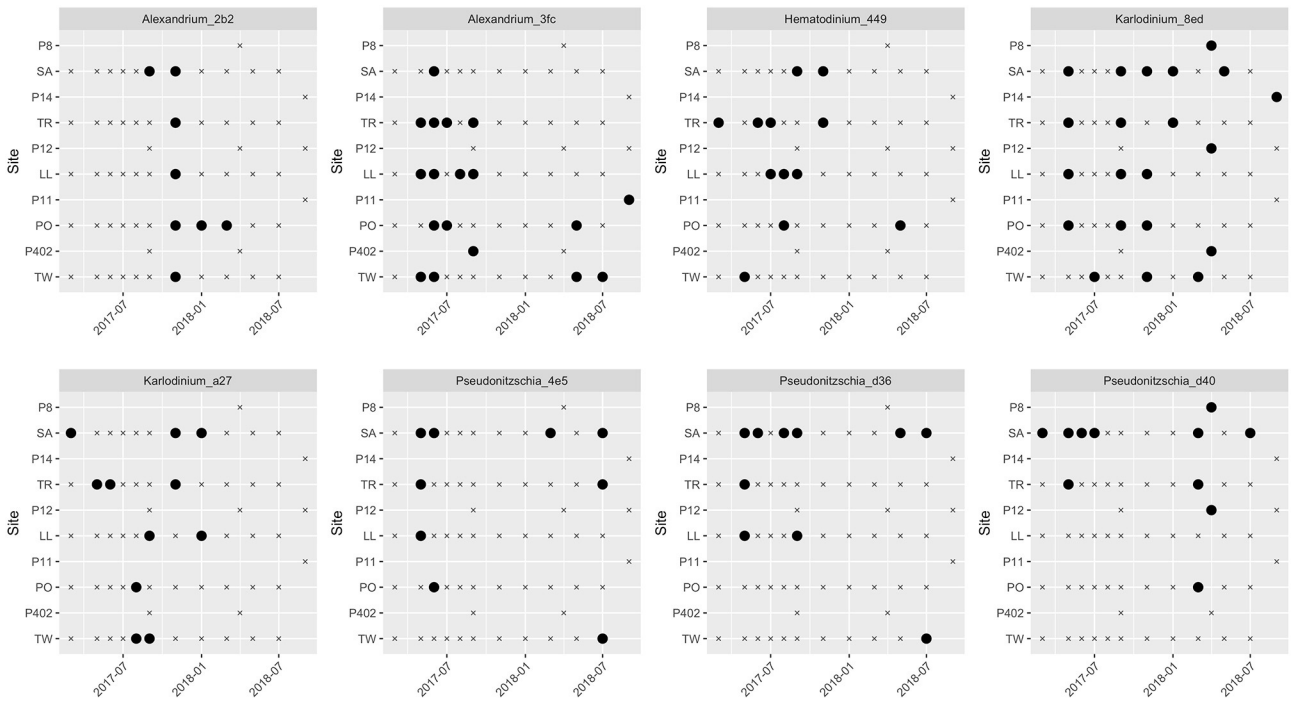


FIGURE 2 | Spatial and temporal distribution of eight focal algal lineages across time and space. Black circles indicate detection; "x" symbol indicates non-detection at that site/date. Site abbreviations are as in **Figure 1** and **Supplementary Table 1**.

Notably, these covariates alone poorly predict occurrences for all taxa (**Table 2**, true positive rate). Consequently, we use eDNA metabarcoding data from the communities surrounding our focal lineages for potentially helpful information about the ecology of these algae.

3.4. Biological Context

To identify the biological community associated with our focal lineages, we searched for co-occurring taxa using a canonical analysis of principal coordinates (CAP) (Anderson and Willis, 2003). Constraining this multivariate analysis according to the detection of each focal algal variant revealed no striking patterns of association across taxa (**Supplementary Tables 5–12**), but rather helped to identify individual community members particularly likely or unlikely to co-occur with our focal lineages. These associated community members were those with the strongest deviations from 0 on the CAP1 axis (**Table 3**), and included lineages of *Ditylum*, a centric diatom, *Prasinoderma*, a non-harmful green algae, *Saxidomus*, a clam, *Balanus*, a barnacle, and *Calanus*, a copepod.

For seven of our eight focal lineages, adding the most closely associated predictor taxon improved model fit even after

TABLE 2 | Best-fit models of environmental covariates for eight focal algal lineages.

Taxon	Environmental model	Accuracy	True positive rate	True negative rate
Alexandrium_2b2	(Intercept + pH | Season)	0.92	0.29	1
Alexandrium_3fc	(Intercept + Temperature | Season)	0.73	0.33	0.89
Hematodinium_449	Intercept + pH + Temperature	0.4	0	0.45
Karlodinium_8ed	(Temperature | Season)	0.76	0.45	0.9
Karlodinium_a27	(Intercept + Salinity | Season)	0.85	0.36	0.96
Pseudonitzschia_4e5	Intercept + Salinity + Temperature	0.43	0	0.46
Pseudonitzschia_d36	Salinity + (Intercept | Season)	0.92	0.5	1
Pseudonitzschia_d40	Intercept + Salinity	0.82	0.27	0.94

Models shown as conditional on season have terms that vary by warmer or colder season, as described in the text.

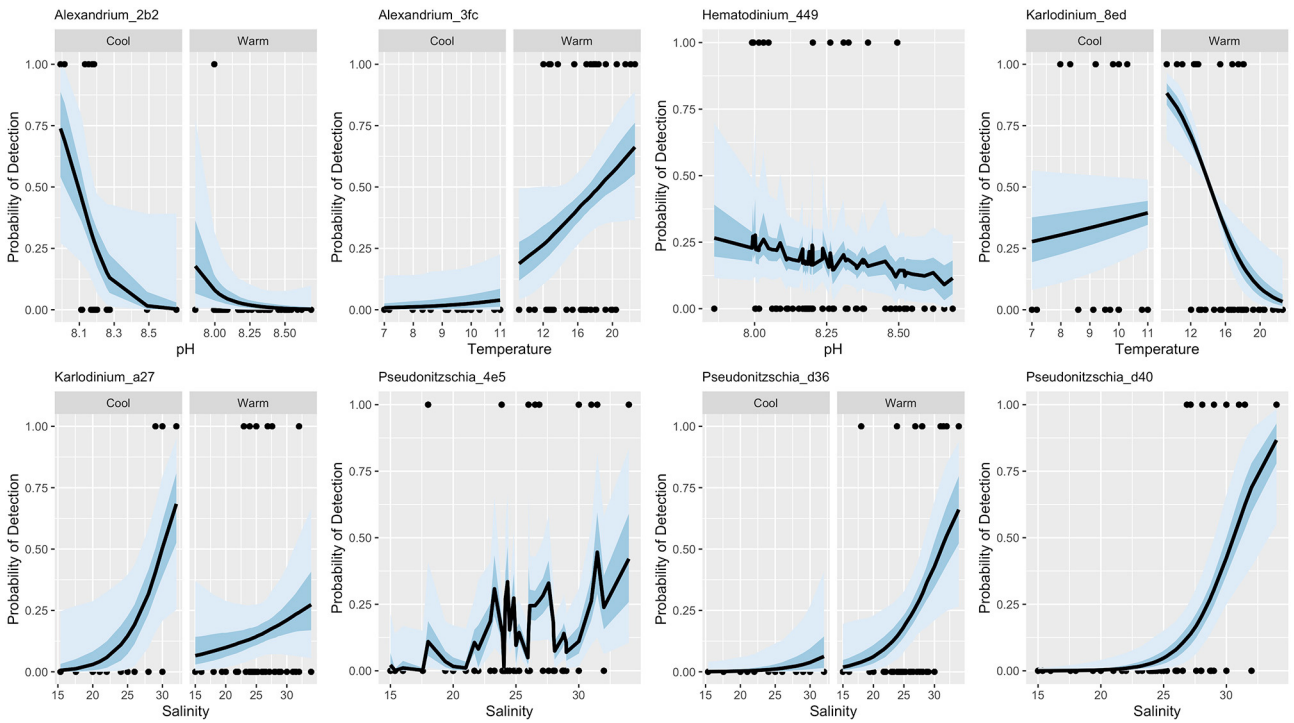


FIGURE 3 | Best-fit logistic models for eight focal algal lineages. Here, probability of detection is shown as a function of the single most influential environmental variable in the model, along with overall model means (lines) and 50 and 95% credibility intervals (shaded areas). Where two-panel figures are shown for a taxon, the best-fit model included a slope term that varied by season.

accounting for the additional model complexity (**Table 3**). Thus, including information about co-occurring organisms alongside baseline environmental covariates substantially increased our ability to predict the detection of these algal taxa within the scope of our sampling. For example, Hematodinium_449 occurs somewhat stochastically in space and time (**Figure 2**) and is not strongly associated with environmental covariates (**Table 2**). However, the CAP analysis revealed that a haplotype from the clam genus *Saxidomus* (likely the species *S. giganteus* (butter clam), given the sampling location), was routinely found in samples in which *Hematodinium* also occurred (**Table 3**). Adding *Saxidomus* as a term in the previous best-fit model more than doubles the model's overall accuracy (overall accuracy 0.4 vs. 0.87; changes in true and false detection rates shown in **Figure 4**); adding this biological variable provides a better prediction than the measured environmental variables alone ($\Delta\text{WAIC} = -9.76$).

Performing the same analysis for each of our focal lineages yields improvements across most taxa, although these are more modest (**Figure 4**), demonstrating an overall model accuracy above environmental covariates alone for most algal variants. Specifically, adding these candidate indicator taxa improved the true-positive rate of detection for six of the eight models (Karlodinium_8ed and Pseudonitzschia_d36 are the exceptions), which accounted for the increase in overall accuracy across models.

TABLE 3 | Terms of the best-fit models combining environmental variables and most closely associated biological taxon for eight focal algal lineages.

Taxon	Combined_environmental and biological model	Delta WAIC
Alexandrium_2b2	*Ditylum* presence + (Intercept + pH | Season)	-7.63
Alexandrium_3fc	*Prasinoderma* presence + (Intercept + Temperature | Season)	0.12
Hematodinium_449	*Saxidomus* presence + pH + Temperature	-9.76
Karlodinium_8ed	*Ditylum* presence + (Temperature | Season)	-3.22
Karlodinium_a27	*Balanus* presence + (Intercept + Salinity | Season)	-7.66
Pseudonitzschia_4e5	*Calanus* presence + Salinity + Temperature	-9.6
Pseudonitzschia_d36	*Calanus* presence + Salinity + (Intercept | Season)	-9.71
Pseudonitzschia_d40	*Ditylum* presence + Salinity	-9.57

Delta WAIC reflects the change in the Widely Applicable Information Criterion for the combined model, relative to the best model using only environmental data; negative values indicate improved model fit.

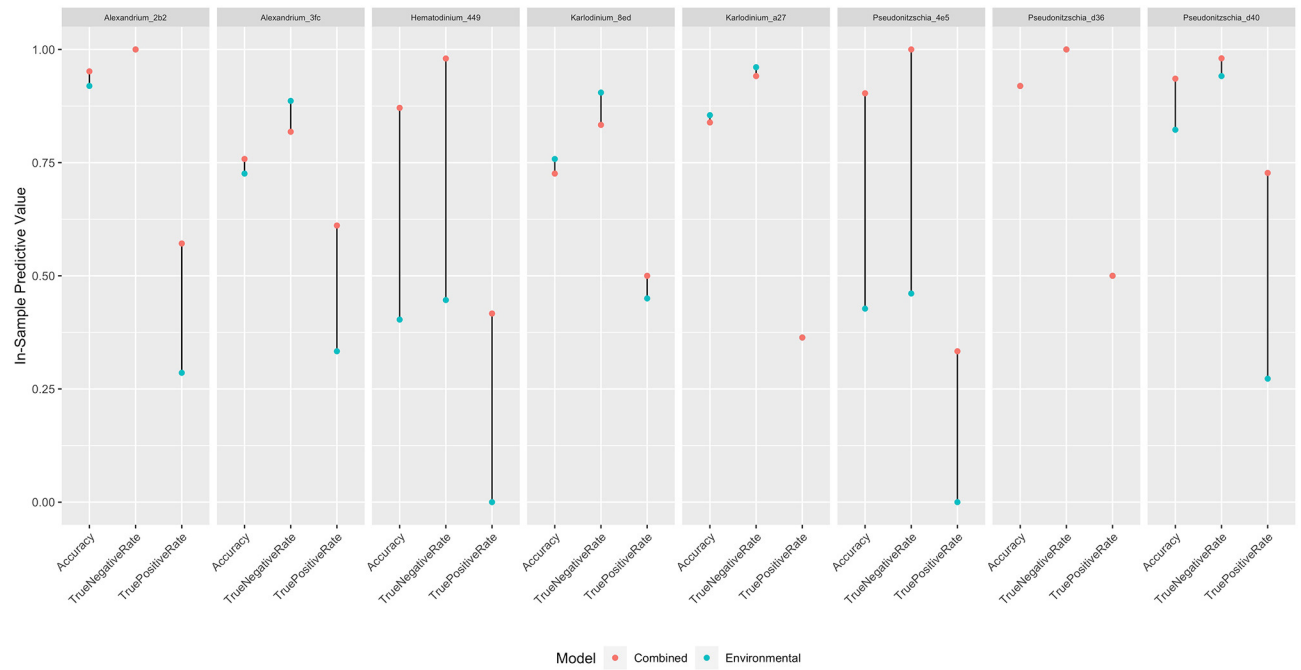


FIGURE 4 | In-sample predictive value of best-fit models for eight focal algal lineages, as measured by accuracy, true-negative rate, and true-positive rate. Red dots indicate values for models combining environmental information with a single associated predictor taxon; blue dots indicate values for models with environmental information alone. Where only a single dot is visible, models produced equivalent results.

4. DISCUSSION

Here, we use genetic monitoring to highlight variants from harmful algae-containing genera within a larger survey of several hundred taxa in intertidal and nearshore marine habitats. Within these focal algal groups, we find several variants from lineages that are unexpected in the study area (*Karlodinium*, *Hematodinium*), in addition to an apparently cryptic lineage of *Alexandrium*, and multiple variants of economically important taxa (e.g., *Pseudo-nitzschia*). Our time-series sampling indicates different seasonal patterns and attendant associations of sea-surface temperature, pH, and salinity for some of these lineages, but on the whole, models using purely environmental covariates offer poor predictive value. We therefore use a constrained ordination to identify taxa (both algal and non-algal) that commonly occur in association with our focal lineages; adding only a single prospective biological indicator taxon improved most of the predictive models.

4.1. Detecting Expected and Unexpected Algal Taxa

eDNA metabarcoding has a number of distinct advantages relative to common techniques currently used to identify harmful algae. First, this technique can reveal a diversity of potentially harmful algal variants present, rather than targeting specific species. In our survey of intertidal and nearshore communities using broad-spectrum mitochondrial COI primers

(Leray et al., 2013), the taxa we identified were largely consistent with what we expected *a priori*, in that we found dozens of variants with excellent overlap from records of known local algal genera containing harmful members (**Table 1**; Horner and Postel, 1993; Horner et al., 1997; Trainer et al., 2016; Moestrup et al., 2020) such as *Alexandrium*, *Chaetoceros*, *Heterocapsa*, *Nitzschia*, *Phaeocystis*, and *Pseudo-nitzschia* (though it is important to note that confirming the harmful nature of individual variants requires identification to species level and a toxicity assay, which are beyond the scope of this study). While confirming expectations demonstrates the reliability of eDNA metabarcoding for identification, detecting unexpected taxa underscores the ability of this technique to reveal novel lineages, range-shifts, or nascent invasions with potentially profound ecological and economic consequences. For example, the genus *Alexandrium*, though known to cause paralytic shellfish poisoning in the region (Trainer et al., 2016), was not previously understood to have two distinct seasonal forms (see below). Additionally, members of *Karlodinium* produce karlotoxins responsible for fish kills in the United States and globally (e.g., *Karlodinium veneficum*, Place et al., 2012), yet this genus is not reported from Puget Sound in the peer-reviewed scientific literature, despite having been noted by a local cell-based monitoring program (Kolb et al., 2016). Similarly, member(s) of the genus *Hematodinium*, which have caused massive losses for the tanner crab (*Chionoecetes bairdi*) and snow crab (*C. opilio*) fisheries in the United States (Meyers et al., 1987, 1996; Wood et al., 2017) and among other species worldwide (Stentiford and Shields, 2005), have also not previously been reported within Puget Sound in the peer-reviewed scientific literature.

Additionally, eDNA metabarcoding can help to standardize data collection and analysis across studies. Here we employ samples collected with two distinct methods by two groups: intertidal water was gathered on foot and by hand, whereas nearshore water was gathered by boat on the Washington Ocean Acidification Cruise in a more routinized process. Nonetheless, taxa from harmful algae-containing genera identified in nearshore surface samples were all found within the larger intertidal dataset as well, suggesting excellent agreement despite differences in methods and personnel. Such congruence between studies also arises from how the data are analyzed: eDNA sequences from multiple studies can consistently identify cryptic taxa by sequence (e.g., Thomsen and Willerslev, 2015; Uchii et al., 2016), and can undergo identical taxonomic analyses that are not subject to differences in interpretation via morphology (Proschold and Leliaert, 2007). However, we note that the success of eDNA studies rests heavily on the shoulders of expert taxonomists: without their contributions to the identification of specimens with sequences in databases, it is impossible to link a fragment of DNA found in the water to an organism (Manoylov, 2014; Zimmermann et al., 2014). Furthermore, the presence-absence eDNA results we treat here do not reflect an absolute quantification of algal cells in a water sample, and do not reveal the toxicity of those cells (or lack thereof). We thus present our monitoring strategy not as a stand-alone method, but as an important complement to existing practices in the field (e.g., Yang et al., 2000; Groben and Medlin, 2005; Ahn et al., 2006; Campbell et al., 2010; Töbe et al., 2010; Murray et al., 2011).

4.2. Taxon Distributions in Space and Time

Our spatial and temporal data indicate that variants from harmful algae-containing genera are constitutively present in the intertidal and nearshore environment, or at the very least are routinely detectable (**Table 1**; **Figure 2**). Although challenges in relating sequence counts to absolute organism abundances limit the utility of eDNA metabarcoding for precise measurement of bloom intensity, the ability of eDNA to reveal harmful algal taxa even when cell counts are much lower than bloom conditions can be advantageous. For example, we detect *Alexandrium* variants year-round in Hood Canal, including winters, when recorded *Alexandrium* blooms are rare, but when fisheries closures due to presence of paralytic shellfish toxin do exist (Trainer et al., 2003; Moore et al., 2011). When paired with more traditional methods, this tool therefore provides another layer of information regarding the behavior of potentially harmful algae, and at the very least can indicate a temporal and spatial starting place for more time-, labor-, and taxonomic expertise-intensive identification and counting strategies.

Sampling many taxa over time and space additionally facilitates important within- and cross-species comparisons (**Figure 2**). Here, such comparisons reveal two lineages of *Alexandrium* with different temporal patterns, and completely non-overlapping distributions. Although taxonomic revisions of the *Alexandrium tamarense* species complex – and competing classifications of local taxa (Lilly et al., 2007; John et al., 2014) – make it impossible to identify the variants present in our survey without additional information, amino acid differences in the COI sequence of the two lineages alongside temporal distribution information suggest that they may represent distinct species, rather than haplotypes. Regardless, recognizing two distinct *Alexandrium* lineages with opposite seasonal dynamics is likely to be important to local monitoring and research programs aiming to identify risk of saxitoxin poisoning (e.g., Kolb et al., 2016; Trainer et al., 2016). Previous work on *Alexandrium* within the Puget Sound found a lower limit for toxic bloom events at 13°C (Nishitani and Chew, 1984), with recent work identifying higher growth of cells above 17-18°C (Bill et al., 2016), and more frequent blooms accompanying warmer air and sea surface temperatures over multiple decades (Moore et al., 2009, 2011). Alexandrium_2b2, detected nearly exclusively in cool season samples, does not match the profile expected given these studies, suggesting either that it does not bloom frequently and/or that it is a recent introduction to the local algal community, whose role is not yet appreciated.

Nucleotide variants from the genus *Pseudo-nitzschia* identified here are not unexpected; multiple species from this genus have been described locally using both visual (e.g., Trainer et al., 2016) and molecular tools (e.g., Hubbard et al., 2014). In the Western Pacific, clades of a single *Pseudo-nitzschia* species (*P. pungens*) with distinct ecological niches and hybrid zones have been documented (Kim et al., 2018); it is interesting to note that here we similarly find two *Pseudo-nitzschia* variants

with distinct nucleotide but identical amino acid sequences at the COI locus. Based on their overlap in time and space, it is likely these represent haplotype lineages that have begun to diverge but are not yet distinct species. Additionally, the specific local antecedents of toxicity in *Pseudo-nitzschia* are still under study (e.g., Zhu et al., 2017; Trick et al., 2018), with production of domoic acid historically limited to the outer coast of Washington (Trainer et al., 2017), but recently moving into Puget Sound (Trainer et al., 2007). Revealing the diversity and pattern of variants in this genus across time and space by eDNA metabarcoding might thus support efforts to better characterize the causes underlying dangerous and costly *Pseudo-nitzschia* bloom events.

4.3. Environmental and Biological Context

Quantitative models of each focal lineage with respect to environmental variables (**Table 2**), motivated by taxon-specific patterns in space and time (**Figure 2**), yield a synoptic view of the occurrence of many algal taxa in the region – a perspective that is otherwise not easily achievable, though many detailed quantitative models have been built for individual harmful algal species (e.g., Hubbard et al., 2014; Moore et al., 2015). Across variants, we note that values expected with global climate change (higher sea surface temperature) and ocean acidification (lower pH) are typically associated with increases in the occurrence of our focal lineages, such as Alexandrium_2b2, Alexandrium_3fc, Hematodinium_449, and Karlodinium_a27. These results in sum align with other studies of local harmful algal taxa, suggesting future scenarios will involve greater seasonal windows of opportunity for toxic bloom events (e.g., Moore et al., 2011; Trainer et al., 2020).

Overall, however, the quantitative models we have built with environmental variables alone have low accuracy, and universally fail to predict a majority of occurrences for our focal lineages (**Table 2**). The difficulty of predicting the presence of harmful algal taxa and their blooms based on a limited number of environmental covariates is not atypical; building accurate models of these species' ecology has long been a challenge for the field (Flynn and McGillicuddy, 2018), even when many more environmental covariates are considered. In this study, Hematodinium_449 and Pseudonitzschia_4e5 are extreme representatives of this challenge; our environmental models do not predict their detection correctly even once (**Table 2**), though they appear in 12 and 9 of 63 samples gathered, respectively. These models are consequently worse than uniformative; they can be actively misleading by inaccurately predicting the absence of harmful species.

Fortunately, eDNA metabarcoding provides an additional layer of information regarding the context of algal variants: the surrounding biological community members (both algal and non-algal). Specifically, the choice of primers amplifying a common molecular marker (mitochondrial COI; Leray et al., 2013) allows us to identify over 600 taxa in total, from more than 250 genera. These data enable us to associate the detection of individual focal variants with a wide range of eukaryotes by CAP (**Table 3**; **Supplementary Tables 5–12**). Studies such as this one that examine individual taxa from harmful algae-containing genera within the breadth of their biological communities are rare, but provide an opportunity to improve prediction (e.g., Banerji et al., 2019). Here, we see that adding only the single best predictor taxon to our quantitative models of focal lineages generally improves their utility (**Table 3**), justifying added model complexity. Although it may appear circular to identify co-occurring species in the dataset and subsequently add them into the model of the same dataset, use of WAIC allows us to assess the value of additional information for future out-of-sample data, generating testable hypotheses for indicator species. As an example, adding the detection of an easily surveyed taxon, a *Saxidomus* clam, improves the prediction of Hematodinium_449 dramatically, including the true-positive measure essential to management (**Figure 4**). Likewise, the majority of focal lineages examined here show an improvement in model accuracy with addition of biological information. These improvements are driven by increases in the true positive rate, which are particularly notable for taxa in which the environmental variables surveyed had low predictive power.

5. CONCLUSION

In this study, we employ COI primers (Leray et al., 2013) to simultaneously identify dozens of algal variants from genera with harmful members, as well as hundreds of other local taxa comprising the biological context for these algae. The broad nature of eDNA metabarcoding surveys allows us to track both expected and unexpected taxa, and the distribution in time and space of eight focal variants from the genera *Alexandrium*, *Hematodinium*, *Karlodinium*, and *Pseudo-nitzschia* suggests the constitutive presence of taxa from these groups in the study region, as well the possibility of nascent range shifts, invasions, and/or ongoing evolutionary divergence. Building individual quantitative models for each of eight focal lineages, we find that many variants are likely to become more common under conditions of higher sea surface temperature and ocean acidification, but note that models using environmental covariates alone have low explanatory power. Adding even a single associated member of the biological community, however, improves most models, and in particular boosts the true positive rates useful for prediction of these algal taxa in the field. eDNA metabarcoding is hence an opportunity to reveal potentially harmful algae outside of bloom events and expected ranges, to map phylogenetic complexity underlying HAB dynamics, to interrogate the relevant environmental context in an era of global change, and to improve models of algal prediction with inclusion of the biological milieu.

DATA AVAILABILITY STATEMENT

The fastq files have been deposited in NCBI - The Project ID is PRJNA699686, and individual FASTQ files are SRR13659629 - 639 & SRR13685891-947.

AUTHOR CONTRIBUTIONS

EJ-P, RG, KC, and RK collected the samples and performed the computational analysis. EJ-P, RG, KC, and AK performed the laboratory analysis. EJ-P and RK wrote the manuscript. EJ-P, RG, KC, AK, and RK edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The Washington Ocean Acidification Center supported this work by purchasing a computer for data analysis. The Washington State Department of Natural Resources supported the salaries of EJP and RPK through Interagency Agreement 93-099546.

ACKNOWLEDGMENTS

We thank the Washington Ocean Acidification Center (WOAC) National Oceanic and Atmospheric Administration Harmful Algal Blooms (NOAA HABs) program cruise, and in particular Simone Alin for her provision of samples and environmental data, as well as Terrie Klinger and Jan Newton for their expert comments on the manuscript. We thank our colleagues at the NOAA Northwest Fisheries Science Center, including the Conservation Biology Molecular Genetics Laboratory for their use of the Illumina MiSeq, as well as Linda Rhodes, Stephanie Moore, Vera Trainer, and Nic Adams for their direction of the analysis in its early stages. We thank Micah Horwith and the Washington State Department of Natural Resources for field assistance. Finally, we thank Alex Gagnon, in whose lab we performed carbonate chemistry analysis for the intertidal dataset, and the lab groups in the University of Washington's Center for Environmental Genomics, with whom we share a work space.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.612107/full#supplementary-material

Supplementary Table 1 | Sample location, date, and read count information.

Supplementary Table 2 | Taxonomic assignments and COI sequences with total read counts for each OTU within genera containing harmful algal species.

Supplementary Table 3 | Within-genera COI sequence alignments for each genus containing harmful algal species (produced with Clustal Omega).

Supplementary Table 4 | Complete list of logistic-regression models tested describing focal algal taxon occurrence as a function of sea-surface temperature, pH, and salinity.

Supplementary Table 5 | Top 10 taxa associated with Alexandrium_2b2 by CAP, with association strength and direction.

Supplementary Table 6 | Top 10 taxa associated with Alexandrium_3fc by CAP, with association strength and direction.

Supplementary Table 7 | Top 10 taxa associated with Hematodinium_449 by CAP, with association strength and direction.

Supplementary Table 8 | Top 10 taxa associated with Karlodinium_8ed by CAP, with association strength and direction.

Supplementary Table 9 | Top 10 taxa associated with Karlodinium_a27 by CAP, with association strength and direction.

Supplementary Table 10 | Top 10 taxa associated with Pseudonitzschia_4e5 by CAP, with association strength and direction.

Supplementary Table 11 | Top 10 taxa associated with Pseudonitzschia_d36 by CAP, with association strength and direction.

Supplementary Table 12 | Top 10 taxa associated with Pseudonitzschia_d40 by CAP, with association strength and direction.

REFERENCES

Ahn, Y-H., Shanmugam, P., Ryu, J-H., and Jeong, J-C. (2006). Satellite detection of harmful algal bloom occurrences in Korean waters. *Harmful Algae* 5, 213–231. doi: 10.1016/j.hal.2005.07.007

Alin, S. R., Newton, J., Greeley, D., Curry, B., Herndon, J., Ostendorf, M. L., et al. (2019a). *Dissolved Inorganic Carbon (DIC), Total Alkalinity (TA), Temperature, Salinity, Oxygen, Nutrient, and CTD Data Collected From Discrete Profile Measurements During Puget Sound Cruise CAB1079 (EXPOCODE 33CB20170911) on R/V Clifford A. Barnes from 2017-09-11 to 2017-09-15 (NCEI Accession 0206674)*. NOAA National Centers for Environmental Information.

Alin, S. R., Newton, J., Greeley, D., Curry, B., Herndon, J., Ostendorf, M. L., et al. (2019b). *Dissolved Inorganic Carbon (DIC), Total Alkalinity (TA), Temperature, Salinity, Oxygen, Nutrient, and CTD Data Collected from Discrete Profile Measurements During Puget Sound Cruise RC001 (EXPOCODE 33IY20180407) on R/V Rachel Carson from 2018-04-07 to 2018-04-11 (NCEI Accession 0206802)*. NOAA National Centers for Environmental Information.

Alin, S. R., Newton, J., Greeley, D., Curry, B., Herndon, J., Ostendorf, M. L., et al. (2019c). *Dissolved Inorganic Carbon (DIC), Total Alkalinity (TA), Temperature, Salinity, Oxygen, Nutrient, and CTD Data Collected from Discrete Profile Measurements During Puget Sound Cruise RC007 (EXPOCODE 33IY20180911) on R/V Rachel Carson from 2018-09-11 to 2018-09-15 (NCEI Accession 0206804)*. NOAA National Centers for Environmental Information.

Al-Tebrineh, J., Mihali, T. K., Pomati, F., and Neilan, B. A. (2010). Detection of saxitoxin-producing Cyanobacteria and *Anabaena circinalis* in environmental water blooms by quantitative PCR. *Appl. Environ. Microbiol.* 76, 7836–7842. doi: 10.1128/AEM.00174-10

Anderson, M. J., and Willis, T. J. (2003). Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. *Ecology* 84, 511–525. doi: 10.1890/0012-9658(2003)084[0511:CAOPCA]2.0.CO;2

Antonella, P., and Luca, G. (2013). The quantitative real-time PCR applications in the monitoring of marine harmful algal bloom (HAB) species. *Environ. Sci. Pollut. Res.* 20, 6851–6862. doi: 10.1007/s11356-012-1377-z

Banerji, A., Bagley, M. J., Shoemaker, J. A., Tettenhorst, D. R., Nietch, C. T., Allen, H. J., et al. (2019). Evaluating putative ecological drivers of microcystin spatiotemporal dynamics using metabarcoding and environmental data. *Harmful Algae* 86, 84–95. doi: 10.1016/j.hal.2019.05.004

Bill, B. D., Moore, S. K., Hay, L. R., Anderson, D. M., and Trainer, V. L. (2016). Effects of temperature and salinity on the growth of *Alexandrium* (Dinophyceae) isolates from the Salish Sea. *J. Phycol.* 52, 230–238. doi: 10.1111/jpy.12386

Brosnahan, M. L., Fischer, A. D., Lopez, C. B., Moore, S. K., and Anderson, D. M. (2020). Cyst-forming dinoflagellates in a warming climate. *Harmful Algae* 91:101728. doi: 10.1016/j.hal.2019.101728

Busch, D. S., Harvey, C. J., and McElhany, P. (2013). Potential impacts of ocean acidification on the Puget Sound food web. *ICES J. Mar. Sci.* 70, 823–833. doi: 10.2307/4451315

Buskey, E. J. (2008). How does eutrophication affect the role of grazers in harmful algal bloom dynamics? *Harmful Algae* 8, 152–157. doi: 10.1016/j.hal.2008.08.009

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13:581. doi: 10.1038/nmeth.3869

Campbell, L., Olson, R. J., Sosik, H. M., Abraham, A., Henrichs, D. W., Hyatt, C. J., et al. (2010). First harmful *Dinophysis* (Dinophyceae, Dinophysales) bloom in the U.S. is revealed by automated imaging flow cytometry. *J. Phycol.* 46, 66–75. doi: 10.1111/j.1529-8817.2009.00791.x

Cembella, A. D., Lewis, N. I., and Quilliam, M. A. (2000). The marine dinoflagellate *Alexandrium ostenfeldii* (Dinophyceae) as the causative organism of spirolide shellfish toxins. *Phycologia* 39, 67–74. doi: 10.2216/i0031-8884-39-1-67.1

Cho, K., Kasaoka, T., Ueno, M., Basti, L., Yamasaki, Y., Kim, D., et al. (2017). Haemolytic activity and reactive oxygen species production of four harmful algal bloom species. *Eur. J. Phycol.* 52, 311–319. doi: 10.1080/09670262.2017.1286525

Curd, E. E., Gold, Z., Kandlikar, G. S., Gomer, J., Ogden, M., O'Connell, T., et al. (2019). Anacapa Toolkit: an environmental DNA toolkit for processing multilocus metabarcode datasets. *Methods Ecol. Evol.* 10, 1469–1475. doi: 10.1111/2041-210X.13214

Deiner, K., Bik, H. M., Mochler, E., Seymour, M., Lacoursiere-Roussel, A., Altermatt, F., et al. (2017). Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Mol. Ecol.* 26, 5872–5895. doi: 10.1111/mec.14350

Diaz, P. A., Alvarez, A., Varela, D., Perez-Santos, I., Diaz, M., Molinet, C., et al. (2019). Impacts of harmful algal blooms on the aquaculture industry: Chile as a case study. *Perspect. Phycol.* 6, 39–50. doi: 10.1127/pip/2019/0081

Dickson, A. G., Sabine, C. L., and Christian, J. R. (2007). *Guide to Best Practices for Ocean Co2 Measurements.* North Pacific Marine Science Organization.

Erdner, D. L., Percy, L., Keafer, B., Lewis, J., and Anderson, D. M. (2010). A quantitative real-time PCR assay for the identification and enumeration of Alexandrium cysts in marine sediments. *Deep Sea Res. Top. Stud. Oceanogr.* 57, 279–287. doi: 10.1016/j.dsr2.2009.09.006

Feely, R. A., Alin, S. R., Newton, J., Sabine, C. L., Warner, M., Devol, A., et'al. (2010). The combined effects of ocean acidification, mixing, and respiration on pH and carbonate saturation in an urbanized estuary. *Estuar. Coast. Shelf Sci.* 88, 442–449. doi: 10.1016/j.ecss.2010.05.004

Ferrante, M., Conti, G. O., Fiore, M., Rapisarda, V., and Ledda, C. (2013). Harmful algal blooms in the Mediterranean sea: effects on human health. *EuroMediterranean Biomed. J.* 8, 25–34. doi: 10.3269/1970-5492.2013.8.6

Field, C. B., Barros, V. R., Mastrandrea, M. D., Mach, K. J., Abdrabo, M. A.-K., Adger, W. N., et al. (2014). *Summary for Policy Makers: Working Group 11 Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Intergovernmental panel on Climate Change.

Flynn, K. J., and McGillicuddy, D. J. (2018). "Modeling marine harmful algal blooms: Current status and future prospects," in *Harmful Algal Blooms: A Compendium Desk Reference*, eds S. E. Shumway, J. M. Burkholder, and S. L. Morton (New York, NY: Wiley Science Publishers), 115–134.

Fu, F. X., Tatters, A. O., and Hutchins, D. A. (2012). Global change and the future of harmful algal blooms in the ocean. *Mar. Ecol. Prog. Ser.* 470, 207–233. doi: 10.3354/meps10047

Gallego, R., Jacobs-Palmer, E., Cribari, K., and Kelly, R. P. (2020). Environmental DNA metabarcoding reveals winners and losers of global change in coastal waters. *Proc. R. Soc. B* 287:20202424. doi: 10.1098/rspb.2020.2424

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784–3788. doi: 10.1093/nar/gkg563

Gattuso, J.-P., Epitalon, J.-M., Lavigne, H., Orr, J., Gentili, B., Hofmann, A., et al. (2015b). *seacarb: Seawater Carbonate Chemistry. R package version.*

Gattuso, J.-P., Magnan, A., Bille, R., Cheung, W. W. L., Howes, E. L., Joos, F., et al. (2015a). Contrasting futures for ocean and society from different anthropogenic CO2 emissions scenarios. *Science* 349:aac4722. doi: 10.1126/science.aac4722

Gobler, C. J. (2020). Climate change and harmful algal blooms: insights and perspective. *Harmful Algae* 91:101731. doi: 10.1016/j.hal.2019.101731

Gobler, C. J., Doherty, O. M., Hattenrath-Lehmann, T. K., Griffith, A. W., Kang, Y., and Litaker, R. W. (2017). Ocean warming since 1982 has expanded the niche of toxic algal blooms in the North Atlantic and North Pacific oceans. *Proc. Natl. Acad. Sci. U.S.A.* 114, 4975–4980. doi: 10.1073/pnas.1619575114

Graneli, E., and Lipiatou, E. (2002). *EUROHAB, Part B, Research and Infrastructural Needs, National European and International Programmes.* Luxembourg: Office for Official Publications of the European Communities.

Grattan, L. M., Holobaugh, S., and Morris, J. G. Jr. (2016). Harmful algal blooms and public health. *Harmful Algae* 57, 2–8. doi: 10.1016/j.hal.2016.05.003

Griffith, A. W., and Gobler, C. J. (2020). Harmful algal blooms: a climate change co-stressor in marine and freshwater ecosystems. *Harmful Algae* 91:101590. doi: 10.1016/j.hal.2019.03.008

Groben, R., and Medlin, L. (2005). "*In situ* hybridization of phytoplankton using fluorescently labeled rRNA probes," in *Methods in Enzymology*, eds E. A. Zimmer and E. H. Roalson (Elsevier), 299–310.

Grzebyk, D., Audic, S., Lasserre, B., Abadie, E., Vargas, C., and Bec, B. (2017). Insights into the harmful algal flora in northwestern Mediterranean coastal lagoons revealed by pyrosequencing metabarcodes of the 28S rRNA gene. *Harmful Algae* 68, 1–16. doi: 10.1016/j.hal.2017.06.003

Horner, R. A., Garrison, D. L., and Plumley, F. G. (1997). Harmful algal blooms and red tide problems on the US west coast. *Limnol. Oceanogr.* 42, 1076–1088. doi: 10.4319/lo.1997.42.5_part_2.1076

Horner, R. A., and Postel, J. R. (1993). "Toxic diatoms in western Washington waters (US west coast)," in *Twelfth International Diatom Symposium* (Renesse: Springer), 197–205.

Hubbard, K. A., Olson, C. E., and Armbrust, E. V. (2014). Molecular characterization of *Pseudo-nitzschia* community structure and species ecology in a hydrographically complex estuarine system (Puget Sound, Washington, USA). *Mar. Ecol. Prog. Ser.* 507, 39–55. doi: 10.3354/meps10820

Jacobs-Palmer, E., Gallego, R., Ramon-Laca, A., Kunselman, E., Cribari, K., Horwith, M., et al. (2020). A halo of reduced dinoflagellate abundances in and around eelgrass beds. *PeerJ* 8:e8869. doi: 10.7717/peerj.8869

John, U., Litaker, R. W., Montresor, M., Murray, S., Brosnahan, M. L., Anderson, D. M. (2014). Formal revision of the *Alexandrium tamarense* species complex (Dinophyceae) taxonomy: the introduction of five species with emphasis on molecular-based (rDNA) classification. *Protist* 165, 779–804. doi: 10.1016/j.protis.2014.10.001

Kelly, R. P., Gallego, R., and Jacobs-Palmer, E. (2018). The effect of tides on nearshore environmental DNA. *PeerJ* 2018. doi: 10.7717/peerj.4521

Kelly, R. P., Shelton, A. O., and Gallego, R. (2019). Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. *Sci. Rep.* 9:12133. doi: 10.1038/s41598-019-48546-x

Khan, S., Arakawa, O., and Onoue, Y. (1997). Neurotoxins in a toxic red tide of *Heterosigma akashiwo* (Raphidophyceae) in Kagoshima Bay, Japan. *Aquaculture Res.* 28, 9–14. doi: 10.1046/j.1365-2109.1997.t01-1-00823.x

Kim, J. H., Wang, P., Park, B. S., Kim, J.-H., Patidar, S. K., and Han, M.-S. (2018). Revealing the distinct habitat ranges and hybrid zone of genetic sub-populations within *Pseudo-nitzschia pungens* (Bacillariophyceae) in the West Pacific area. *Harmful algae* 73, 72–83. doi: 10.1016/j.hal.2018.01.007

Kolb, A., Hannach, G., and Swanson, L. (2016). *Marine Phytoplankton Monitoring Program Sampling and Analysis Plan*. Seattle, WA: King County Department of Natural Resources; Parks.

Kotaki, Y., Furio, E. F., Bajarias, F. A., Satake, M., Lundholm, N., Koike K., et al. (2006). New stage of the study on domoic acid-producing diatoms–a finding of *Nitzschia navis-varingica* that produces domoic acid derivatives as major toxin components. *Coast. Mar. Sci.* 30, 116–120.

Lahoz-Monfort, J. J., Guillera-Arroita, G., and Tingley, R. (2016). Statistical approaches to account for false-positive errors in environmental DNA samples. *Mol. Ecol. Resour.* 16, 673–685. doi: 10.1111/1755-0998.12486

Lapworth, C., Hallegraeff, G. M. J., and Ajani, P. A. (2001). "Identification of domoic-acid-producing *Pseudo-nitzschia* species in Australian waters," in *Harmful Algal Blooms 2000 Proceedings of the Ninth International Conference on Harmful Algal Blooms* (Paris), 38–41.

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., et al. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Front. Zool.* 10:34. doi: 10.1186/1742-9994-10-34

Lewitus, A. J., Horner, R. A., Caron, D. A., Garcia-Mendoza, E., Hickey, B. M., Hunter, M., et al. (2012). Harmful algal blooms along the North American west

coast region: History, trends, causes, and impacts. *Harmful Algae* 19, 133–159. doi: 10.1016/j.hal.2012.06.009

Lilly, E. L., Halanych, K. M., and Anderson, D. M. (2007). Species boundaries and global biogeography of the *Alexandrium tamarense* complex (Dinophyceae) 1. *J. Phycol.* 43, 1329–1338. doi: 10.1111/j.1529-8817.2007.00420.x

Lindberg, K., Moestrup, E., and Daugbjerg, N. (2005). Studies on woloszynskioid dinoflagellates i: *Woloszynskia coronata* re-examined using light and electron microscopy and partial lsu rDNA sequences, with description of *tovellia* gen. nov. And *jadwigia* gen. nov. (Tovelliaceae fam. Nov.). *Phycologia* 44, 416–440.

Lopez, C. B., Jewett, E. B., Dortch, Q., Walton, B. T., and Hudnell, H. K. (2008). *Scientific Assessment of Freshwater Harmful Algal Blooms*. Interagency Working Group on Harmful Algal Blooms, Hypoxia, and Human Health.

Loureiro, S., Rene, A., Garces, E., Camp, J., and Vaque, D. (2011). Harmful algal blooms (HABs), dissolved organic matter (DOM), and planktonic microbial community dynamics at a near-shore and a harbour station influenced by upwelling (SW Iberian Peninsula). *J. Sea Res.* 65, 401–413. doi: 10.1016/j.seares.2011.03.0040

Mahe, F., Rognes, T., Quince, C., Vargas, C., and Dunthorn, M. (2015). Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ* 3:e1420. doi: 10.7717/peerj.1420

Manoylov, K. M. (2014). Taxonomic identification of algae (morphological and molecular): species concepts, methodologies, and their implications for ecological bioassessment. *J. Phycol.* 50, 409–424. doi: 10.1111/jpy.12183

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17:10. doi: 10.14806/ej.17.1.200

Mauger, G., Casola, J., Morgan, H., Strauch, R., Jones, B., Curry, B., et al. (2015). *State of Knowledge: Climate Change in Puget Sound*. NOAA and the University of Washington Climate Impacts Group.

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course With Examples in R and Stan*. Boca Raton, FL: CRC Press.

Meyers, T., Morado, J., Sparks, A., Bishop, G., Pearson, T., Urban, D., et al. (1996). Distribution of bitter crab syndrome in Tanner crabs (*Chionoecetes bairdi, C. opilio*) from the Gulf of Alaska and the Bering Sea. *Dis. Aquat. Organ.* 26, 221–227.

Meyers, T. R., Koeneman, T. M., Botelho, C., and Short, S. (1987). Bitter crab disease: a fatal dinoflagellate infection and marketing problem for Alaskan Tanner crabs *Chionoecetes bairdi*. *Dis. Aquat. Organ.* 3, 195–216.

Moestrup, E., Akselmann-Cardella, R., Churro, C., Fraga, S., Hoppenrath, M., Iwataki, M., et al. (2020). IOC-UNESCO taxonomic reference list of harmful micro algae. doi: 10.14284/362

Moore, S. K., Cline, M. R., Blair, K., Klinger, T., Varney, A., and Norman, K. (2019). An index of fisheries closures due to harmful algal blooms and a framework for identifying vulnerable fishing communities on the US West Coast. *Marine Policy* 110:103543. doi: 10.1016/j.marpol.2019.103543

Moore, S. K., Johnstone, J. A., Banas, N. S., Salathe, E. P. Jr. (2015). Present-day and future climate pathways affecting Alexandrium blooms in Puget Sound, WA, USA. *Harmful Algae* 48, 1–11. doi: 10.1016/j.hal.2015.06.008

Moore, S. K., Mantua, N. J., Hickey, B. M., and Trainer, V. L. (2009). Recent trends in paralytic shellfish toxins in Puget Sound, relationships to climate, and capacity for prediction of toxic events. *Harmful Algae* 8, 463–477. doi: 10.1016/j.hal.2008.10.003

Moore, S. K., Mantua, N. J., and Salathe, E. P. Jr. (2011). Past trends and future scenarios for environmental conditions favoring the accumulation of paralytic shellfish toxins in puget sound shellfish. *Harmful Algae* 10, 521–529. doi: 10.1016/j.hal.2011.04.004

Murray, S. A., Wiese, M., Stuken, A., Brett, S., Kellmann, R., Hallegraeff, G., et al. (2011). sxtA-based quantitative molecular assay to identify saxitoxin-producing harmful algal blooms in marine waters. *Appl. Environ. Microbiol.* 77, 7050–7057. doi: 10.1128/AEM.05308-11

National Centers for Coastal Ocean Science. (2014). *Working With State to Document First Occurrence of Harmful Karenia mikimotoi Algae in Alaskan Waters*.

Nishitani, L., and Chew, K. K. (1984). Recent developments in paralytic shellfish poisoning research. *Aquaculture* 39, 317–329.

O'Donnell, J. L., Kelly, R. P., Lowell, N. C., and Port, J. A. (2016). Indexed PCR primers induce template-specific bias in large-scale DNA sequencing studies. *PLoS ONE* 11:e0148698. doi: 10.1371/journal.pone.0148698

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., et al. (2013). *Package "vegan". Community Ecology Package, version 2*.

Pavlou, M., Ambler, G., Seaman, S., De Iorio, M., and Omar, R. Z. (2016). Review and evaluation of penalised regression methods for risk prediction in low-dimensional data with few events. *Stat. Med.* 35, 1159–1177. doi: 10.1002/sim.6782

Peduzzi, P., Concato, J., Kemper, E., Holford, TR., and Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *J. Clin. Epidemiol.* 49, 1373–1379.

Peperzak, L., and Poelman, M. (2008). Mass mussel mortality in the Netherlands after a bloom of *Phaeocystis globosa* (prymnesiophyceae). *J. Sea Res.* 60, 220–222. doi: 10.1016/j.seares.2008.06.001

Pierce, R. H., and Kirkpatrick, G. J. (2001). Innovative techniques for harmful algal toxin analysis. *Environ. Toxicol. Chem. Int. J.* 20, 107–114. doi: 10.1897/1551-5028(2001)020<0107:itfhat>2.0.co;2

Place, A. R., Bowers, H. A., Bachvaroff, T. R., Adolf, J. E., Deeds, J. R., and Sheng, J. (2012). *Karlodinium veneficum*—The little dinoflagellate with a big bite. *Harmful Algae* 14, 179–195. doi: 10.1016/J.HAL.2011.10.021

Plummer, M., Stukalov, A., Denwood, M., and Plummer, M. M. (2016). *Package 'rjags'*. Vienna.

Proschold, T., and Leliaert, F. (2007). Systematics of the green algae: Conflict of classic and modern approaches. *Syst. Assoc. Spcl. Vol.* 75:123. doi: 10.1201/9780849379901.ch7

Raven, J. A., Gobler, C. J., and Hansen, P. J. (2020). Dynamic Co2 and pH levels in coastal, estuarine, and inland waters: theoretical and observed effects on harmful algal blooms. *Harmful Algae* 91:101594. doi: 10.1016/j.hal.2019.03.012

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

Royle, J. A., and Link, W. A. (2006). Generalized site occupancy models allowing for false positive and false negative errors. *Ecology* 87, 835–841. doi: 10.1890/0012-9658(2006)87[835:gsomaf]2.0.co;2

Ruvindy, R., Bolch, C. J., MacKenzie, L., Smith, K. F., and Murray, S. A. (2018). qPCR assays for the detection and quantification of multiple paralytic shellfish toxin-producing species of *Alexandrium*. *Front. Microbiol.* 9:3153. doi: 10.3389/fmicb.2018.03153

Satake, M., MacKenzie, L., and Yasumoto, T. (1998). Identification of *Protoceratium reticulatum* as the biogenetic origin of yessotoxin. *Nat. Toxins* 5, 164–167.

Schnell, I. B., Bohmann, K., and Gilbert, M. T. P. (2015). Tag jumps illuminated–reducing sequence-to-sample misidentifications in metabarcoding studies. *Mol. Ecol. Resour.* 15, 1289–1303. doi: 10.1111/1755-0998.12402

Shimizu, Y., Alam, M., Oshima, Y., and Fallon, W. E. (1975). Presence of four toxins in red tide infested clams and cultured *Gonyaulax tamarensis* cells. *Biochem. Biophys. Res. Commun.* 66, 731–737. doi: 10.1016/0006-291X(75)90571-9

Sievers, F., and Higgins, D. G. (2014). "Clustal Omega, accurate alignment of very large numbers of sequences," in: *Multiple Sequence Alignment Methods*, ed D. J. Russell (New York, NY: Springer), 105–116.

Simonsen, S., and Moestrup, E. (1997). Toxicity tests in eight species of *Chrysochromulina* (Haptophyta). *Can. J. Bot.* 75, 129–136.

Skjelbred, B., Horsberg, T. E., Tollefsen, K. E., Andersen, T., and Edvardsen, B. (2011). Toxicity of the ichthyotoxic marine flagellate *Pseudochattonella* (Dictyochophyceae, Heterokonta) assessed by six bioassays. *Harmful Algae* 10, 144–154. doi: 10.1016/j.hal.2010.08.007

Stentiford, G. D., and Shields, J. D. (2005). A review of the parasitic dinoflagellates *Hematodinium* species and *Hematodinium*-like infections in marine crustaceans. *Dis. Aquat. Organ.* 66, 47–70. doi: 10.3354/dao066047

Taberlet, P., Bonin, A., Coissac, E., and Zinger, L. (2018). *Environmental DNA: For Biodiversity Research and Monitoring*. Oxford University Press.

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

Töbe, K., Kulis, D., Anderson, D. M., Gladstone, M., and Medlin, L. K. (2010). Detecting intact algal cells with whole cell hybridisation assays. *Microsc. Mol. Methods Quant. Phytoplankton Anal.* 55.

Tomlinson, M. C., Stumpf, R. P., Ransibrahmanakul, V., Truby, E. W., Kirkpatrick, G. J., Pederson, B. A., et al. (2004). Evaluation of the use of SeaWiFS imagery for detecting *Karenia brevis* harmful algal blooms in the eastern Gulf of Mexico. *Remote Sens. Environ.* 91, 293–303. doi: 10.1016/j.rse.2004.02.014

Trainer, V., King, T., Bill, B., and Runyan, J. (2016). *SoundToxins Manual: Puget Sound Harmful Algal Bloom Monitoring Program.* NOAA/NMFS/NWFSC Marine Biotoxins Program for Washington Sea Grant.

Trainer, V., Moore, L., Bill, B., Adams, N., Harrington, N., Borchert, J., da Silva, D., et al. (2013). Diarrhetic shellfish toxins and other lipophilic toxins of human health concern in Washington State. *Mar. Drugs* 11, 1815–1835. doi: 10.3390/md11061815

Trainer, V. L. (2002). "Harmful algal blooms on the US west coast," in *Harmful Algal Blooms in the PICES Region of the North Pacific*, eds F. J. R Taylor and V. L. Trainer (Sidney, BC), 89–118.

Trainer, V. L., Adams, N. G., Bill, B. D., Ayres, D. L., Forster, Z. R., Odell, A., et al. (2017). *Pseudo-nitzschia* blooms in the northeastern Pacific Ocean. *PICES Sci. Rep* 53, 37–48.

Trainer, V. L., Cochlan, W. P., Erickson, A., Bill, B. D., Cox, F. H., Borchert, J. A., et al. (2007). Recent domoic acid closures of shellfish harvest areas in Washington state inland waterways. *Harmful Algae* 6, 449–459. doi: 10.1016/j.hal.2006.12.001

Trainer, V. L., Eberhart, B.-T. L., Wekell, J. C., Adams, N. G., Hanson, L., Cox, F., et al. (2003). Paralytic shellfish toxins in Puget Sound, Washington state. *J. Shellfish Res.* 22, 213–223.

Trainer, V. L., Moore, S. K., Hallegraeff, G., Kudela, R. M., Clement, A., Mardones, J. I., and Cochlan, W. P. (2020). Pelagic harmful algal blooms and climate change: lessons from nature's experiments with extremes. *Harmful Algae* 91:101591. doi: 10.1016/j.hal.2019.03.009

Trainer, V. L., Wells, M. L., Cochlan, W. P., Trick, C. G., Bill, B. D., Batgh, K. A., et al. (2009). An ecological study of a massive bloom of toxigenic *Pseudo-nitzschia* cuspidata off the Washington State coast. *Limnol. Oceanogr.* 54, 1461–1474. doi: 10.4319/lo.2009.54.5.1461

Trainer, V. L., and Yoshida, T. (2014). *Proceedings of the Workshop on Economic Impacts of Harmful Algal Blooms on Fisheries and Aquaculture*. North Pacific Marine Science Organization (PICES).

Trick, C. G., Trainer, V. L., Cochlan, W. P., Wells, M. L., and Beall, B. (2018). The successional formation and release of domoic acid in a *Pseudo-nitzschia* bloom in the Juan de Fuca Eddy: a drifter study. *Harmful Algae* 79, 105–114. doi: 10.1016/j.hal.2018.08.007

Uchii, K., Doi, H., and Minamoto, T. (2016). A novel environmental dna approach to quantify the cryptic invasion of non-native genotypes. *Mol. Ecol. Resour.* 16, 415–422. doi: 10.1111/1755-0998.12460

van Smeden, M., Moons, KG., Groot, J. A., Collins, G. S., Altman, D. G., Eijkemans, M. J., et al. (2019). Sample size for binary logistic prediction models: beyond events per variable criteria. *Stat. Methods Med. Res.* 28, 2455–2474. doi: 10.1177/0962280218784726

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J. Mach. Learn. Res.* 11, 3571–3594.

Wells, M. L., Trainer, V. L., Smayda, T. J., Karlson, B. S. O., Trick, C. G., Kudela, R. M., et al. (2015). Harmful algal blooms and climate change: learning from the past and present to forecast the future. *Harmful Algae* 49, 68–93. doi: 10.1016/j.hal.2015.07.009

Wood, K., Stratman, J. P., Messmer, A., and Palof, K. J. (2017). *Annual Management Report for the 2016/2017 Southeast Alaska and Yakutat Tanner Crab Fisheries*. Alaska Department of Fish; Game, Divisions of Sport Fish; Commercial Fisheries.

Yang, C. Z., and Albright, L. J. (1994). The harmful phytoplankter *Chaetoceros concavicornis* causes high mortalities and leucopenia in chinook salmon (*Oncorhynchus tshawytscha*) and coho salmon (*O. kisutch*). *Can. J. Fish. Aquat. Sci.* 51, 2493–2500.

Yang, Z. B., Takayama, H., Matsuoka, K., and Hodgkiss, I. J. (2000). *Karenia digitata* sp. nov. (Gymnodiniales, Dinophyceae), a new harmful algal bloom species from the coastal waters of west Japan and Hong Kong. *Phycologia* 39, 463–470. doi: 10.2216/i0031-8884-39-6-463.1

Zhu, Z., Qu, P., Fu, F., Tennenbaum, N., Tatters, A. O., and Hutchins, D. A. (2017). Understanding the blob bloom: Warming increases toxicity and abundance of the harmful bloom diatom *Pseudo-nitzschia* in California coastal waters. *Harmful Algae* 67, 36–43. doi: 10.1016/j.hal.2017.06.004

Zimmermann, J., Abarca, N., Enk, N., Skibbe, O., Kusber, W.-H., and Jahn, R. (2014). Taxonomic reference libraries for environmental barcoding: a best practice example from diatom research. *PLoS ONE* 9:e108793. doi: 10.1371/journal.pone.0114758

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 28 May 2021
doi: 10.3389/fevo.2021.630560

Unearthing the Potential of Soil eDNA Metabarcoding—Towards Best Practice Advice for Invertebrate Biodiversity Assessment

Ameli Kirse[1], Sarah J. Bourlat[1], Kathrin Langen[1] and Vera G. Fonseca[1,2]**

[1] Centre for Biodiversity Monitoring, Zoological Research Museum Alexander Koenig, Bonn, Germany, [2] Centre for Environment Fisheries and Aquaculture Science (Cefas), Weymouth, United Kingdom

Metabarcoding has proven to be a powerful tool to assess ecological patterns and diversity from different habitats. Terrestrial invertebrate diversity is frequently based on bulk samples, which require comparatively high sampling effort. With environmental DNA (eDNA) metabarcoding, field sampling effort can be reduced while increasing the number of recovered organism groups. However, a proof of concept is missing for several invertebrate groups, hampering the development of best-practice advice for these groups. This study aims to provide recommendations on key aspects for the processing of soil samples, from sampling effort to choice of DNA extraction method and marker genes. This study uses eDNA metabarcoding as a tool for assessing invertebrate biodiversity in soil samples, specifically comparing two DNA extraction methods (with and without a lysis step) and two genes, 18S and COI markers. The results show that the choice of marker and DNA extraction method (including a lysis step) significantly affect species detection rates and concomitantly observed invertebrate community composition. Combining methods, by using larger amounts of starting material and including a lysis step resulted in an increase of invertebrate species numbers. Together, these methods improved the detection of species with known lower population densities and allowed the assessment of temporary mesofauna. Furthermore, the choice of marker significantly influenced the diversity levels found. The 18S marker allowed the detection of a higher number of annelid and nematode OTUs, while the COI marker was more suitable for detecting changes in arthropod community structure, especially at the species level. This study makes significant advances to the field of invertebrate biodiversity assessment, particularly using metabarcoding tools by addressing several methodological considerations that are key for accurate ecological appraisals.

Keywords: eDNA extraction, extracellular DNA, intracellular DNA, invertebrates, metabarcoding, metazoa, soil biodiversity, eDNA metabarcoding

OPEN ACCESS

Edited by:
Katy Klymus, United States Geological Survey (USGS), United States

Reviewed by:
Stefano Mariani, Liverpool John Moores University, United Kingdom
Tobias Guldberg Frøslev, University of Copenhagen, Denmark

****Correspondence:***
Ameli Kirse
A.Kirse@leibniz-zfmk.de
Vera G. Fonseca
vera.fonseca@cefas.co.uk

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *17 November 2020*
Accepted: *05 May 2021*
Published: *28 May 2021*

Citation:
Kirse A, Bourlat SJ, Langen K and Fonseca VG (2021) Unearthing the Potential of Soil eDNA Metabarcoding—Towards Best Practice Advice for Invertebrate Biodiversity Assessment. Front. Ecol. Evol. 9:630560. doi: 10.3389/fevo.2021.630560

INTRODUCTION

Despite the indisputable fact that the fertility of soil is directly linked to existing fauna and flora (Delgado-Baquerizo et al., 2017) little is known about soil biodiversity. To prevent the ongoing loss of biodiversity leading to soil degradation processes, entailing annual costs of several billion dollars (Kuhlman et al., 2010), it is of uttermost importance to develop timely and cost-efficient assessment strategies.

In particular, environmental DNA (eDNA) metabarcoding appears to be a promising tool for filling-in the knowledge gap on soil biodiversity (Oliverio et al., 2018). As it is unnecessary to collect species to detect their presence, eDNA metabarcoding is a non-invasive approach, which limits the sampling effort to a minimum while retrieving unparalleled diversity information from any habitat with reduced costs (Deiner et al., 2017). Several studies have already shown the applicability of eDNA metabarcoding for the assessment of soil invertebrate diversity (Bienert et al., 2012; Zinger et al., 2019) even to study past ecosystems (Epp et al., 2012). Few methodological studies exist to date for soil arthropod detection despite in particular soil arthropods can be used as key indicators of faunal community structure (Neher et al., 2012).

The choice of method and protocol often has a direct influence on the assessed community composition (Alberdi et al., 2018; Dopheide et al., 2019). For microbial community studies it has been observed that DNA extraction methods (Delmont et al., 2011) and sample size (Kang and Mills, 2006) have an effect on the community composition found. Invertebrates, which have heterogenous morphologies, sizes and abundances (Taberlet et al., 2012; Dopheide et al., 2019), will likely require tailored DNA extraction steps as most available commercial kits are optimized for microbial diversity assessment (Zinger et al., 2016). Furthermore, many invertebrates undergo several life stages incorporating inactive phases (e.g., pupal or dormant stage) which might only be detected through a lysis step (Pietramellara et al., 2009). The amount of source material (e.g., grams of soil) used for DNA extraction and the inclusion of biological replicates can be crucial for maximum detection of soil arthropod species richness (Porter et al., 2019) but also marker choice can significantly influence the composition of the recovered community (Giebner et al., 2020). For the phylogenetically diverse soil invertebrates, it remains unclear which marker is most suitable or if a one-fits all marker exists. Previously, the 16S and 18S markers have been used to assess soil arthropod communities (Epp et al., 2012; Yang et al., 2014), but more recent studies also utilized the CO1 marker (Oliverio et al., 2018; Porter et al., 2019).

This study aims to provide recommendations on key aspects for processing soil samples, from sampling effort to choice of DNA extraction methods and marker genes. Here two well-known soil DNA extraction methods (with and without a lysis step) are compared to evaluate their suitability for invertebrate mesofauna community diversity assessment from forest soil samples. To our knowledge this is the first study investigating the direct effect of the application of a lysis step on soil eDNA metabarcoding targeting invertebrate taxa.

MATERIALS AND METHODS

Sample Collection and Processing

To monitor changes in soil biodiversity over a period of 12 months, we sampled each season between summer 2016 and spring 2017 (**Supplementary Table 2**). In summer 2016, sample collection was conducted at 12 sites located in the Eifel National Park, in south-western Germany (**Supplementary Figure 1** and **Supplementary Table 1**). In autumn, winter and spring, sample collection was conducted at 14 sites comprising the 12 sampling sites being sampled in summer (**Supplementary Table 2** and **Supplementary Figure 2**). At each sampling site, three soil samples were collected approximately 4–5 m apart from each other from the top 10 cm of soil, using a hand-held metal corer of 4.4 cm diameter × 10 cm length. A total of 162 soil samples were collected and kept in individual 250 ml containers which were transported to the laboratory shortly after sampling and stored at −20°C until further processing.

For this study, a forest conversion gradient from a Norway spruce (*Picea abies*) monoculture to a european beech forest (*Fagus sylvatica*) was sampled. The four forest types sampled differed in tree species composition, degree of anthropogenic influence and in the approximate ages of the trees. The pure beech (120 years old) and pure spruce (*Picea abies*) (60 years old) sampling sites were located in monoculture stands. At the young beech sampling sites 60 year old spruce stands had only recently been underplanted with young beeches which had not yet reached 3 m in size at the time of sampling. At the so-called old beech sampling sites, 60 year old spruces were underplanted with beeches several years ago. At the time of sampling the beeches had already reached a height of more than 3 m and actions to remove spruces from the forest had already been undertaken.

DNA Extraction

Soil samples were removed from the −20°C chamber approximately 12 h before DNA extraction and stored at + 4°C overnight. The next morning, each sample was thoroughly homogenized by gently swirling the 250 ml container. Two different DNA extraction methods were used, one using the silica membrane based NucleoSpin Soil kit (Macherey-Nagel) (MN kit herein) with a lysis step and the other using the phosphate buffer protocol (PB herein) from Taberlet et al. (2012). For the first method, 0.5 g of soil was used per sample to extract DNA from the 162 soil samples using the MN kit, following the manufacturer's protocol. The Taberlet et al. (2012) method consists of using a saturated phosphate buffer to desorb DNA fragments from sediment particles, whereby extracellular DNA is recovered using also the NucleoSpin Soil kit (Macherey-Nagel) but skipping the lysis step and following manufacturer's instructions. Briefly, in the PB method, DNA was extracted from ca. 100 g of soil using a phosphate buffer-based solution (Na_2HPO_4; 0.12 M; pH 8) following the Taberlet et al. (2012) protocol. For the latter, soil samples were saturated in the phosphate buffer solution and placed in an orbital shaker at 120 rpm for 15 min. Subsequently, two 2 ml Eppendorf safe lock tubes were filled with 1.7 ml of the resulting mixture and centrifuged for 10 min at 10,000 g. Afterward, 400 µl of the resulting supernatant were transferred to a new 2 ml collection tube to which 200 µl of SB binding buffer of the NucleoSpin® Soil kit (Macherey-Nagel) was added. Supernatants from duplicate samples were loaded onto NucleoSpin® Soil Column and centrifuged at 10,000 g for 1 min. The remaining extraction steps followed the standard manufacturer's protocol from the NucleoSpin® Soil kit starting from step 8 (excluding the lysis step). All DNA extracts were

eluted with 50 μl of SE Buffer. Ten microliters of the resulting elution step were combined with 90 μl pure H_2O (Sigma), followed by DNA purification using the PowerClean® Pro DNA Clean-Up Kit (MO Bio Laboratories, Inc.) following the manufacturer's protocol. DNA extracts obtained with either of the two methods were subsequently quantified using the Quantus Fluorometer (Promega).

Choice of Primers and Library Preparation

Amplicon library preparation was conducted following a two-step PCR approach (Fonseca and Lallias, 2016). For library preparation of soil samples, two primer pairs targeting the COI and 18S markers respectively were used. A 380 bp fragment of the V4 region of the nuclear 18S rRNA was amplified using the following forward primer TAReuk454FWD1 *(5′-* ACACTCTTTCCCTACACGACGCTCTTCCGATCT ***CCAGCASCYGCGGTAATTCC****-3′)* combined with the reverse primer TAReukREV3r *(5′*-GTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCT ***ACTTTCGTTCTTGATYRA****-3′*) (Stoeck et al., 2010). The mitochondrial COI primer pair consisted of the forward primer mlCOIintF (*5′*-ACACTCTTTCCCTACACGACGCTCTTCCGATCT **GGWA *CWGGWTGAACWGTWTAYCCYCC****-3′)* and the reverse primer dgHCO2198 *(5′*- GTGACTGGAGTTCAGACGTGTGC TCTTCCGATCT ***TAAACTTCAGGGTGACCAAARAAYCA****-3′)* (Leray et al., 2013), targeting a 313 bp region of the 658 bp long barcoding COI gene.

Approximately 10 ng of template DNA was used for all PCR reactions. For PCR1 the mastermix consisted of 7.5 μl Q5 Hot Start High-Fidelity 2X Master Mix (New England BioLabs), 5 μl Sigma H_2O, 0.5 μl of each primer (10 μM), 0.5 μl Bovine Serum Albumin (Thermoscientific) and 1 μl template DNA, making up a final volume of 15 μl. The first PCR (PCR1) consisted in an initial denaturation for 2 min at 98°C, followed by 20 cycles with 40 s at 98°C, 40 s at 45°C, 30 s at 72°C (COI), or 20 cycles with 40 s at 98°C, 40 s at 55°C, 30 s at 72°C (18S), and a final extension of 3 min at 72°C. PCR1 products were then purified using 4 μl HT ExoSAP-IT^{TM} (Applied Biosystems) per 10 μl of PCR1 product, following the manufacturer's protocol. For PCR2, the purified PCR1 products were split into two PCR tubes. Each tube contained 12.5 μl Q5 Hot Start High-Fidelity 2X MasterMix (New England BioLabs), 3 μl Sigma H_2O, 1.2 μl of forward index primer (10 μM) (AATGATACGGCGACCACCGAGATCTACAC NNNNNNNN ACACTCTTTCCCTACACGACGCTC) and 1.2 μl of reverse index primer (10 μM) (CAAGCAGAAGACGGCATACGAGAT NNNNNNNN GTGACTGGAGTTCAGACGTGTGCTC), and 8 μl purified PCR1 product. The PCR2 conditions consisted of an initial denaturation of 2 min at 98°C, followed by 20 cycles with 40 s at 98°C, 30 s at 55°C, 30 s at 72°C, and a final extension of 3 min at 72°C. PCR2 products were visualized by gel electrophoresis and purified using the QIAquick Gel Extraction Kit (Qiagen), according to manufacturer's instructions. All final purified amplicons (PCR2) were quantified using the Quantus Fluorometer (Promega) and normalized to the same concentration (3 ng/μl) before being pooled together to create two amplicon libraries (18s and COI). The resulting purified amplicon library pools were sequenced on two runs on the Illumina Miseq (2 × 300 bp) sequencing platform at the Centre for Genomic Research (CGR, Liverpool University).

Bioinformatics and Data Analysis

Initial quality check of raw sequences at CGR comprised trimming of fastq sequences for the presence of Illumina adapter sequences using Cutadapt version 1.2.1. Afterward, sequences were trimmed using Sickle version 1.200 with a minimum window quality score of 20. Only reads longer than 19 bp were kept for further analysis.

The fastq sequences were checked for the presence of the COI and 18S primers with Cutadapt version 1.18 (Martin, 2011) using the following settings: maximum error rate (−e): 0.1, minimum overlap (−O): 20, minimum sequence length (−m): 50. Sequences lacking either forward or reverse primers were removed from the dataset. Subsequently, paired-end reads were merged with vsearch version 2.7.0 (Rognes et al., 2016). Merged sequences with a length of 360–400 bp for the 18S and 293–333 bp for the COI dataset respectively were retained for further analysis and filtered with a maxEE threshold of 1.0 using vsearch (version 2.7.0) (Rognes et al., 2016). Afterward, fastq sequences were demultiplexed using the script split_libraries_fastq.py with a phred quality threshold of 19 implemented in QIIME1 (Caporaso et al., 2010). Dereplicating, size sorting, *de novo* chimera detection and Operational Taxonomic Units (OTUs) clustering at 97% cutoff was conducted with vsearch 2.7.0 (Rognes et al., 2016). An OTU table was built using the –usearch_global function in vsearch 2.7.0 (Rognes et al., 2016) and the python script "uc2otutab.py" written by Robert Edgar[1]. Resulting OTU tables for both markers were further curated using the LULU algorithm, known to decrease erroneous OTUs (Frøslev et al., 2017). Curation started with an initial blasting of OTU representative sequences against each other using blastn (version 2.9.0) with "query coverage high-scoring sequence pair percent" (-qcov_hsp_perc) set to 80 and minimum percent identity (-perc_identity) set to 84 and a customized output format defined by the –outfmt setting "6 qseqid sseqid pident." Subsequently, the resulting filtered OTU match list was uploaded into R (version 3.5) (R Core Team, 2013), where the R-package "lulu" (version 0.1.0) (Frøslev et al., 2017) was used to perform post clustering curation using standard settings.

For taxonomy assignment, the COI dataset was blasted against the BOLD database (downloaded on May 5th 2019) using blastn 2.9.0 + (Altschul et al., 1990). As the BOLD database is strongly limited in number of bacterial sequences and barcodes of many eukaryotic species outside Metazoa, a second database was downloaded on February 27th 2020 from GenBank using the following search criteria: [COI(All Fields) OR COX1(All Fields)] OR CO1(All Fields) AND [fungi(filter) OR protists(filter) OR bacteria(filter) OR archaea(filter)]. All sequences not assigned to Metazoa when blasted against the downloaded BOLD database were compared to the above GenBank reference database. For

[1]https://drive5.com/python/uc2otutab_py.html

taxonomy assignment of the 18S dataset all sequences were blasted against a customized reference database downloaded on February 27th 2020 from GenBank according to the following criteria: [(18S) OR V4 AND animals(filter) OR fungi(filter) OR plants(filter)]. Sequences without assignment were in a second step blasted against the newly released SILVA132 release[2]. Raw sequence data for this project are deposited in NCBI's SRA database under accession number PRJNA681091.

Statistical and Ecological Analysis

The resulting OTU tables (**Supplementary Tables 3, 4**) were loaded into Excel and formatted for upload into R studio v3.5 (R Core Team, 2013). For statistical analysis, several R packages were used. For data visualization we used ggplot2 (Wickham, 2016) and for data manipulation dplyr version 0.8.3 (Wickham et al., 2015). To further visualize shared and unique OTU numbers per marker, phylum and season between the different methods we used VennDiagram version 1.6.20 (Chen and Boutros, 2011).

[2]www.arb-silva.de/silva-license-information

Pairwise dissimilarities between the two methods on OTU presence-absence matrices based on Jaccard similarity index were performed for incidence data of detected OTUs with a 90% BlastID to Eukaryota, using the R-package betapart version 1.5.1 (Baselga and Orme, 2012). Sample completeness curves and sample-size-based R/E curve with extrapolations of Hill numbers for incidence data based on a combined dataset of both markers, encompassing all OTUs assigned to Arthropoda with a blastID of at least 99% (removed duplicate assignments) were prepared using the R-package iNEXT (Hsieh et al., 2016) at default settings (40 knots, 95% confidence intervals generated by the bootstrap procedure (50 bootstraps).

To visualize and analyze community dissimilarities between methods, PCoAs and statistical tests based on Jaccard similarity index for incidence data of detected eukaryote OTUs with a 90% BlastID were performed using betapart v1.5.1. The betadisper test was performed to test for homogeneity between samples followed by PERMANOVA (adonis) to further test for differences in community composition depending on the DNA extraction method and marker used.

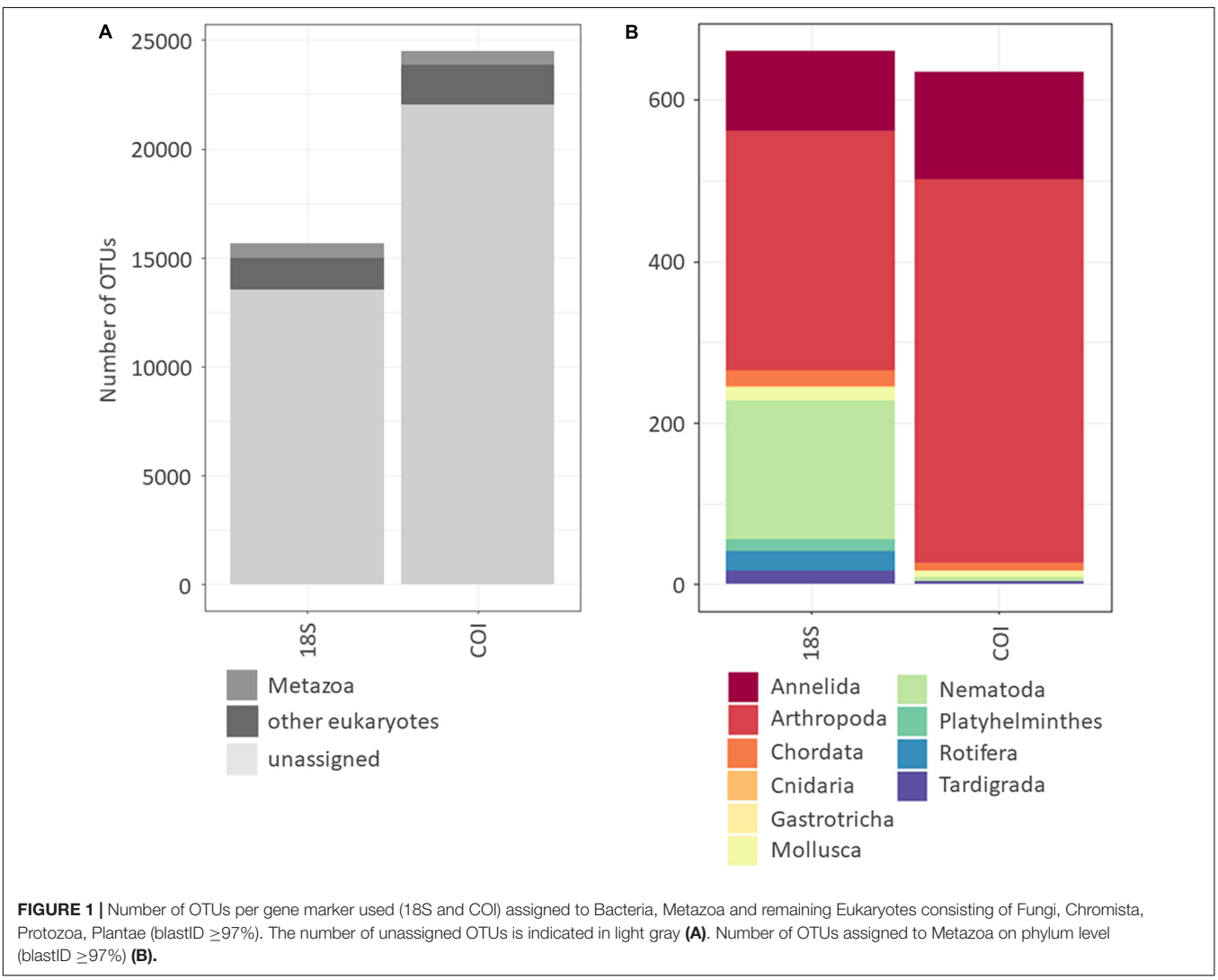


FIGURE 1 | Number of OTUs per gene marker used (18S and COI) assigned to Bacteria, Metazoa and remaining Eukaryotes consisting of Fungi, Chromista, Protozoa, Plantae (blastID ≥97%). The number of unassigned OTUs is indicated in light gray **(A)**. Number of OTUs assigned to Metazoa on phylum level (blastID ≥97%) **(B).**

To identify the insect species primarily contributing to community dissimilarities between extraction methods depending on season a SIMPER analysis (Gibert and Escarguel, 2019) was performed in R. The SIMPER analysis was done using COI OTUs assigned to Insecta at the species level with a blastID of at least 99%.

RESULTS

Amplification of the COI marker resulted in the detection of 25,036,251 high quality-filtered reads, which were subsequently clustered into 31,781 OTUs. When amplifying the V4 region of the 18S marker a total of 22,036,784 high quality filtered reads were obtained, which were clustered into 33953 OTUs. After Lulu curation, the total number of OTUs was 23,004 OTUs for the COI dataset (72.4%) and 15,650 OTUs for the 18S dataset (46%).

The complete COI dataset showed a lower assignment rate compared to the 18S dataset (**Figure 1**). Based on a blast sequence identity cutoff (blastID) of at least 97%, 13.48% of all retrieved 18S OTUs matched an entry in the reference databases, whereas for the COI it was 10.08% (**Figure 1A**). At the kingdom level, 31.48% of the taxonomically identified 18S OTUs (664 OTUs) and 25.72% of the COI OTUs (635 OTUs) were assigned to Metazoa, respectively (**Figure 1A**). For both marker datasets it was found that with 68.42% (1443 OTUs) and 74.20% (1832 OTUs) the lion share of assigned OTUs accounted for eukaryotes outside of Metazoa. Additionally, two OTUs of each marker dataset were assigned to Bacteria. Within the Metazoa, the 18S marker identified ten phyla: Annelida, Arthropoda, Chordata, Cnidaria, Gastrotricha, Mollusca, Nematoda, Platyhelminthes, Rotifera, Tardigrada (**Figure 1B**), while the COI marker, identified six phyla (Annelida, Arthropoda, Chordata, Mollusca, Nematoda, Tardigrada). Collapsing all OTUs with the same taxonomic annotation with a blastID of at least 99% a total of 12 annelid species were identified. Out of them, eight were exclusively detected with the 18S rRNA gene, while the remaining four species were only found with the mitochondrial marker (**Figure 2A**). Out of the 208 detected arthropod species, 146 (70%) were exclusively found by the COI marker, whereas the 18S retrieved additional 57 species. The two marker datasets shared five arthropod species (**Figure 2B**). For the Insecta, 96 species were identified using the COI with a BlastID of at least 99%, and six species using the 18S marker (**Figure 2C**). No insect species was detected with both markers.

The number of OTUs did not vary substantially between extraction methods but more so between markers within each dataset representing either of the two extraction methods (**Figure 3**). For the Machery Nagel kit (MN kit) a total of 18,439 COI OTUs and 13,164 18S OTUs were found, while a total of 17,329 COI OTUs and 13,034 18S OTUs were identified with the Phosphate buffer (PB) (**Figure 3**). Followed by Metazoa, several OTUs were assigned to other eukaryotic taxa (for simplification herein referred to as "other Eukaryota," mainly dominated by Fungi and protists). The 18S marker retrieved ca. seven times more OTUs assigned to "other eukaryotes" than COI, with a slightly higher number of 18S OTU numbers when using the MN kit (+1,173 OTUs) than the PB (+1,093 OTUs). For the Metazoa, the amplification of the 18S marker led to a slightly higher OTU yield when the PB was used for DNA extraction. The opposite was the case for the COI dataset, where an increase in the number of OTUs was associated with the use of the MN kit (**Figure 3**).

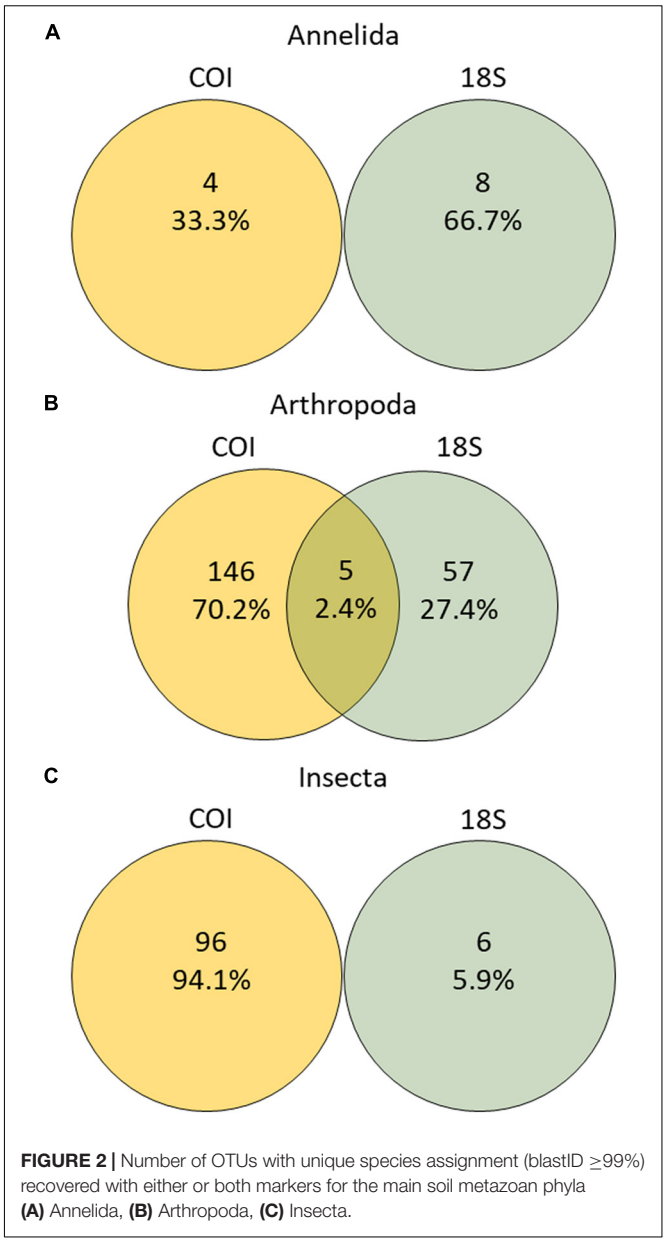


FIGURE 2 | Number of OTUs with unique species assignment (blastID ≥99%) recovered with either or both markers for the main soil metazoan phyla **(A)** Annelida, **(B)** Arthropoda, **(C)** Insecta.

Accumulation (**Figure 4A**) and sampling effort curves (**Figure 4B**) from a total of 162 soil samples did not reach a plateau. An extrapolation indicated that at least 400 samples must be processed with each of the two extraction methods to cover total existing diversity in our sampled environments (**Figure 4A**).

A Principal Coordinate Analysis (PCoAs) indicated major differences in the eukaryotic communities between the different extraction methods, although there was more overlap between the methods in the COI dataset. These differences between

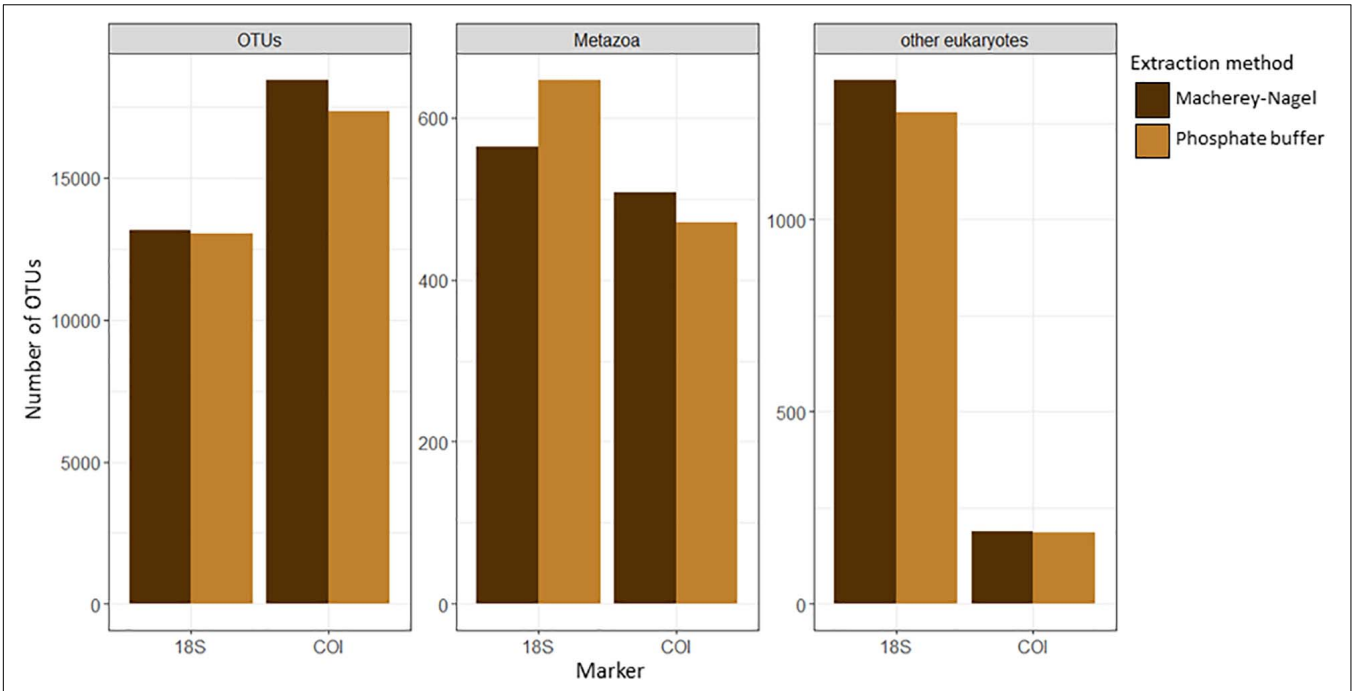


FIGURE 3 | Total number of OTUs and number of OTUs assigned (blastID ≥97%) to Metazoa and other eukaryotes (consisting of Fungi, Chromista, Protozoa, Plantae) for the two extraction methods and markers used (18S and COI).

DNA extraction methods (beta diversity) were subsequently statistically confirmed (PERMANOVA, COI: $F_{323} = 11.26$, $p < 0.001$; 18S: $F_{323} = 43.92$, $p < 0.001$) (**Figures 5A,B**). However, for the COI as well as for the 18S dataset, a betadisper-test indicated a very heterogeneous dispersion within samples of each extraction and marker group (COI: $F_1 = 31.12$, $p = 0.001$; 18S: $F_1 = 3.65$, $p < 0.05$), highlighting the importance of replicates (**Figures 5C,D**).

When using the 18S marker, both extraction methods shared 8 and 47 species of annelids and arthropods, respectively. No arthropod and annelid species (blastID ≥99%) were exclusively identified with the MN kit, while one annelid and 10 arthropod species were unique to the PB method (**Figure 6**). When using the COI marker no differences were observed in the number of annelid and arthropod species between the two methods. The same four annelid species were identified with both extraction methods (**Figure 6**). From a total of 107 arthropod species, 68 were uncovered by both extraction methods and 39 species were unique to each method (**Figure 6**).

Based on the complete dataset, seasonal differences were observed between the two DNA extraction methods. The summer season retrieved the highest number of arthropod species when using the MN kit (36) as opposed to the PB (23). The MN kit also showed a peak in Diptera diversity during summer (20) as opposed to the autumn (6), winter (3), and spring (4) seasons (**Figure 7**). The number of insect species identified during autumn was 29 for the MN kit and 31 for the PB, respectively, with 19 species uncovered by both methods. For winter and spring seasons the PB uncovered more insect species, with 36 species in each season, while using the MN kit 26 and 27 species were found, respectively (**Figure 7**). While the MN kit resulted in the detection of a higher number of dipteran species in summer, in each season more coleopteran species were identified by the PB (summer: + 2; autumn: + 5; winter: + 7; spring: + 8). When considering data from all seasons and forest types 17 coleopteran species were exclusively detected with the phosphate buffer, while the MN kit exclusively revealed the presence of five coleopteran species. For the dipterans, extraction with the MN kit resulted in the exclusive detection of 18 species but the same method left seven species undetected which were found by the PB. Depending on the dipteran family, differences in relative species count were observed between the two extraction methods. Based on the COI marker more species for the families Sciaridae (+2) Mycetophilidae (+3), Limoniidae (+3), and Phoridae (+4) were retrieved with the MN kit (**Supplementary Figure 4**).

SIMPER analysis revealed that *Athous subfuscus* (10.64%), *Ctenosciara lutea* (6.42%), *Corynoptera globiformis* (4.30%), *Cratyna perplexa* (4.00%), and *Strophosoma melanogrammum* (3.82%) had a key influence on the assessed community differences between extraction methods in the summer season. In autumn, next to *C. globiformis* (7.10%) and *Athous subfuscus* (9.75%) also *Tipula limbata* (10.23%), *Othius myrmecophilus* (8.52%), and *Barypeithes araneiformis* (5.97%) were significantly contributing to the observed differences. In winter, again *A. subfuscus* (12.19%), *C. globiformis* (7.75%), *O. myrmecophilus* (6.36%), and *T. limbata* (6.01%) were under the five species responsible for observed differences. Additionally, a high contribution (%) was also observed for *Corynoptera minima* (6.83%). In winter season, *A. subfuscus* (12.73%), *T. limbata* (8.64%), *C. globiformis*, *S. melanogrammum,* and *Berypeithes*

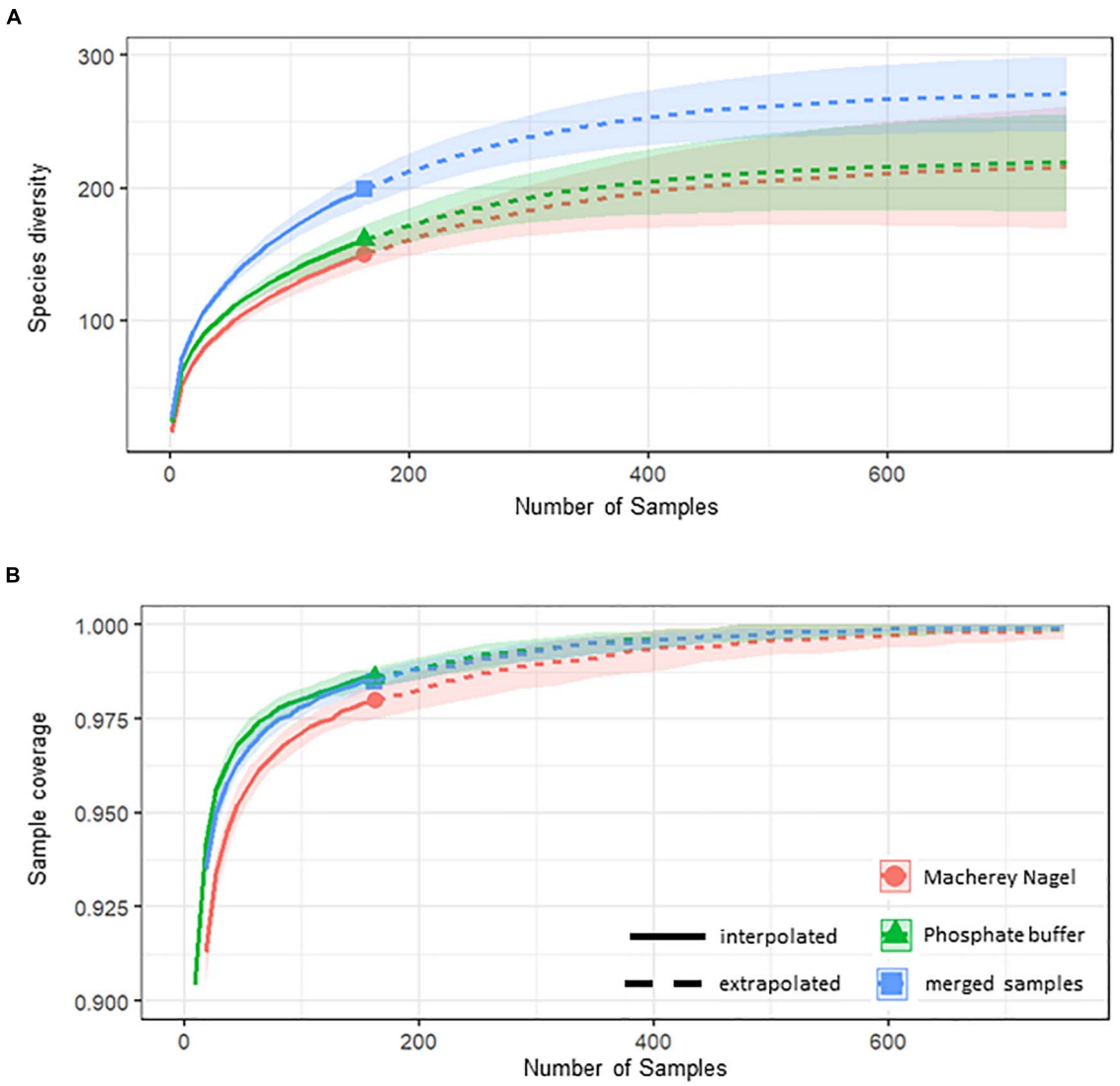


FIGURE 4 | Sample completeness curves (solid line) with extrapolation (dashed line) showing number of detected arthropod species (blastID ≥99%) with both markers per number of soil samples taken for each extraction method **(A)**. Additionally, sample coverage curves were calculated based on sample coverage per number of soil samples **(B)**.

pellucidus were the five species identified as most relevant for observed differences between extraction methods (**Table 1**).

DISCUSSION

This study demonstrates that extraction methods can greatly influence the levels of diversity and species uncovered in a specific location and this is even more significant depending on the targeted taxa and gene used. Many factors can influence eDNA yields from soil samples, namely organic content and humic substances, choice of buffer and purification steps utilized (Frostegård et al., 1999), and thereby the completeness of species lists retrieved. Dopheide et al. (2019) found a correlation between the amount of source material and the number of species retrieved, a finding which is partly contradicting the results of this study. Here we found that the amount of starting material did not significantly influence the number of species retrieved, but more so the taxonomic composition and representativeness of the sampled area. However, it cannot be excluded that the detection of several species exclusively found by the phosphate buffer were also associated with the larger amount of source material used (Taberlet et al., 2012; Dopheide et al., 2019). However, we argue that the taxonomic differences found between the DNA extraction methods are partly inherent to the specificities of the protocols. Up to 44% of species identified were unique to each method showing that half of the species would not have been identified if using one extraction method only, disproving a positive effect of sample size on completeness of community composition.

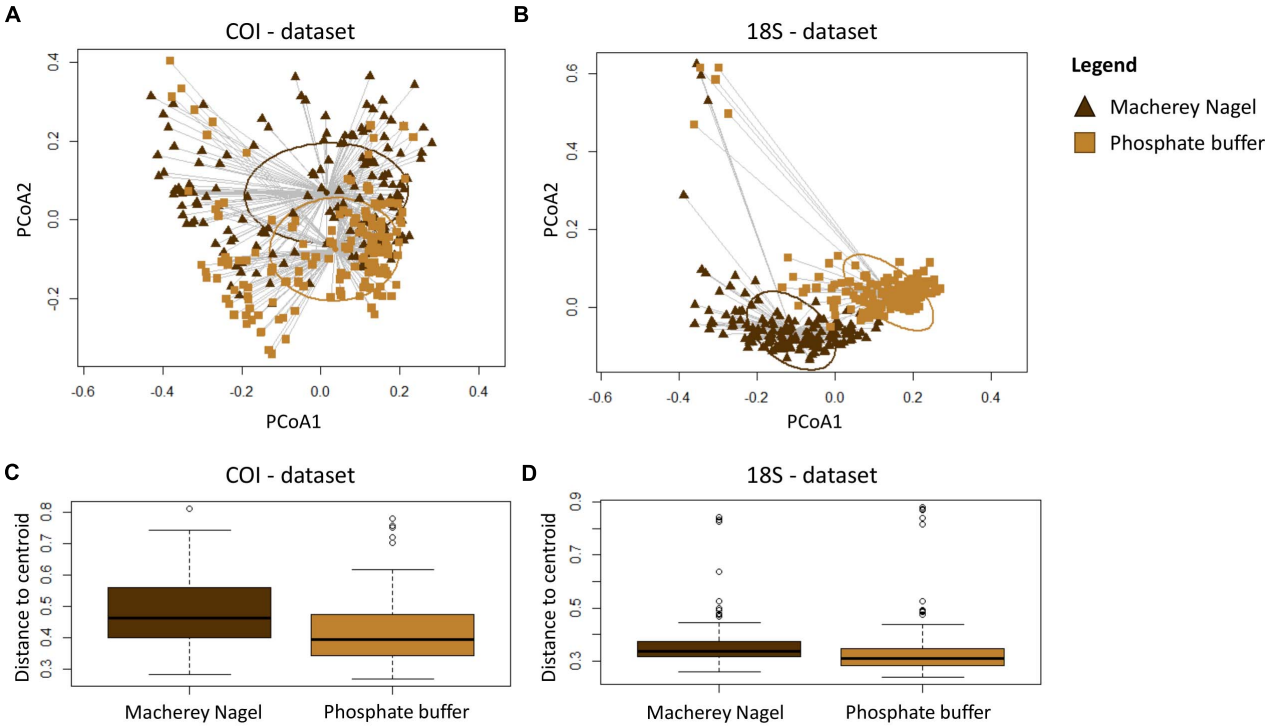


FIGURE 5 | Eukaryotic community composition using Principal Coordinate Analysis (PCoA) based on a "jaccard" dissimilarity index for the **(A)** COI and **(B)** 18S dataset for incidence data of detected OTUs with a 90% BlastID to Eukaryota. Group dispersion plots indicate homogeneity of community composition within all soil samples using the same DNA extraction method for the **(C)** COI dataset and **(D)** 18S dataset. Light brown indicates community composition assessed using the phosphate buffer; Dark brown indicates community composition assessed with the Macherey Nagel kit.

Both markers recovered a high number of OTUs assigned to groups outside the Metazoa, which supports the understanding of significant non-targeted amplification (Yu et al., 2012; Yang et al., 2014; Giebner et al., 2020). The COI barcode is especially limited when working with eDNA due to the vast diversity of the DNA mixture (Deagle et al., 2014). This marker is known to fail to amplify some groups of arthropods (Marquina et al., 2019a), especially in eDNA samples where primers are rarely universal and have different amplification efficiencies. In this study a high proportion of the COI OTUs found could not be taxonomically assigned, probably because the COI marker is less well-used outside Metazoa (Kress and Ericksoneds, 2012) leading to incomplete databases. Additionally, the use of lower blast thresholds (sequence% ID $\leq$97%) and the presence of a consensus blast could have allowed more assignments and greater confidence in the assigned taxonomies, but such parameters were not tested in this study. While the phosphate buffer exclusively allowed the extraction of extracellular DNA, the Macherey Nagel kit included a lysis step, additionally enabling the extraction of intracellular DNA. As the highest amount of intracellular DNA in soil usually originates from microbial organisms (Taberlet et al., 2012) the application of a lysis step is expected to lead to an accumulation of microbial DNA in the DNA extract. Nonetheless, we observed that lysis also allowed the detection of specific invertebrate groups, namely temporary mesofauna (e.g., transient life stages). In the summer, extraction with the Macherey Nagel kit indicated a peak in dipteran diversity, in particular for the families Sciaridae, Mycetophilidiae, Liimonidae, and Phoridae (**Supplementary Figure 5**) which are known to have larval stages developing in the soils (Barnard, 2011; Disney, 2012; Jakovlev, 2012). From the ten species identified as primarily contributing to the observed community dissimilarities between extraction methods in summer three were members of the dipteran family Sciaridae. This highlights the direct effect of choice of extraction method on the composition of the dipteran diversity found.

As the proportion with which a species contributes to the DNA mixture directly influences its detection probability (Elbrecht et al., 2017, 2019), lysis can facilitate the detection of transient species, but for the costs of a lower detection probability of DNA traces. Although little is known about natural eDNA release processes in soil and how they might vary between species it can be expected that detection rate is affected by population density, whereby highly abundant species together with high primer affinities will likely be PCR amplified more efficiently, with concomitantly higher amplification success and more reads (Hajibabaei et al., 2011; Brandon-Mong et al., 2015). Former studies indicated that annelids can reach abundances of up to 134,000 specimens per m^2 (Coleman et al., 2004), with fecal pellets up to 29% of the volume of the higher soil

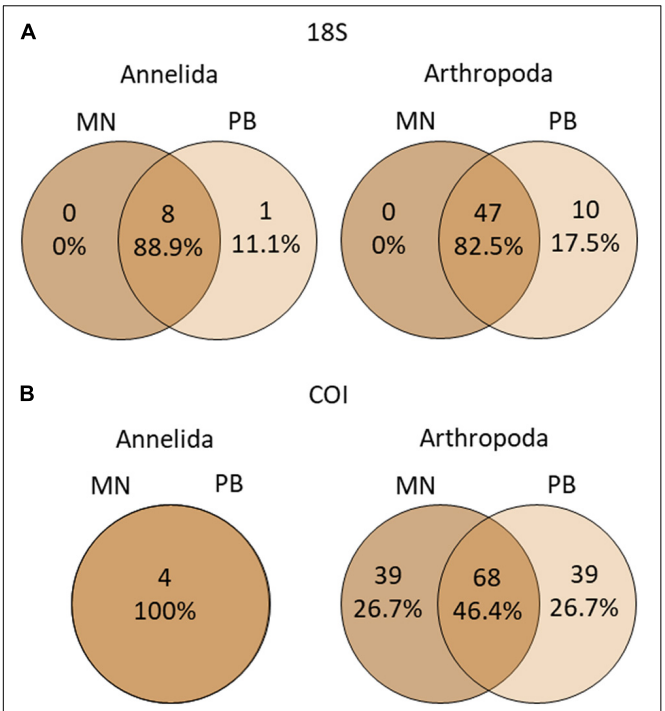


FIGURE 6 | Venn diagram showing number of shared and unique Annelida and Arthropoda species (blastID ≥99%) per extraction method for the **(A)** 18S and **(B)** COI marker.

A- horizon (Davidson et al., 2002). Here, with only 0.5 g of soil using the Macherey Nagel kit, we captured exceptionally high levels of oligochaete Enchytraeid DNA, but both extraction methods captured the same species, probably due to their high abundance and biomass in soils. Although the number of dipteran species exclusively recovered with the Macherey-Nagel kit exceeded the number detected with the phosphate buffer, a high number of small sized dipteran Sciaridae species was also recovered with both extraction methods. In dipterans up to 14,500 larvae can accumulate on very narrow areas (Altmüller, 1977), which can result in an accumulation of DNA traces detectable with both methods.

Soils are heterogenous and stratified, either horizontally or vertically thus sampling larger quantities of soil will allow a better representativeness and homogeneity between replicates. Consequently, size and replication will be key when targeting larger organisms, such as meso- and macrofauna. Here, we observed that soil communities were indeed taxonomically more similar between sample replicates when using the phosphate buffer (**Supplementary Figures 3A,B**). Such findings corroborate the idea that using larger amounts of soil for DNA extraction will increase the chances to assess a more complete picture of the existing invertebrate diversity. Similarly, the rarefaction curves evidenced the need to increase the sampling effort and combine different methods whenever possible, since a total of 162 soil samples did not reach a plateau and at least 400 samples would be needed from each extraction method to cover the existing diversity in our sampled environments. By doing so, we would have been able to assess the arthropod diversity at a given area, as shown by the leveled sampling and species efforts curves. The relatively high percentage of species exclusively recovered from one sampling site using either extraction method substantiates the fact that even at small scales there is a large variation in soil community composition. A more extensive sampling and the combination of different extraction methods can therefore lead to higher local-diversity levels (alpha-diversity) which are commonly found in soils (Nielsen et al., 2010).

Both COI and 18S markers showed non-targeted amplification, but for Metazoa, the 18S gene identified three times more phyla than the COI from the forest soil eDNA samples. This is mainly due to the highly conserved priming sites in 18S, that allow amplification across broader taxonomic groups (Hebert et al., 2003; Zhang et al., 2018). Due to COI marker having a higher taxonomic resolution for Metazoa and especially Arthropoda, more OTUs had an assignment to these phyla relative to the 18S marker. Whereas, the lower number of 18S OTUs assigned to Metazoa were likely because some sequences originating from different species/genus are merged into the same OTU due to the limited species-level resolution in the 18S marker (Potter et al., 2017). However, it must be noted that many COI OTUs did not get a taxonomic assignment mainly because available COI databases can still be fragmentary for some taxonomic groups (Clarke et al., 2017). When focusing on the two main metazoan phyla Annelida and Arthropoda, we observed that it was mainly the marker used that influenced the number OTUs retrieved per phyla. While with the 18S marker more arthropod and annelid species (blastID ≥99%) were detected when the extraction was conducted with the phosphate buffer. Conversely, when using the COI marker the extraction method did not influence the number of arthropod and annelid species identified. As previously mentioned, the 18S marker at a 99% nucleotide divergence threshold is prone to underestimate the real diversity of several metazoans at lower taxonomic levels, namely the Arthropoda (Tang et al., 2012; Drummond et al., 2015). While the primer binding sites of the 18S marker are more conserved (Clarke et al., 2017), its species-level resolution is strongly hampered by the lack of variability within the discriminative region (Tang et al., 2012; Yang et al., 2013). However, due to its low variability in primer sites, amplification success of the 18S marker might be less influenced by the complexity and composition of the DNA mixture, as opposed to COI, since primer affinities are substantially more similar for the majority of taxa.

So far, no primer or single gene region has been identified that will amplify all taxa in eDNA samples and assessments of complete biodiversity are nearly impossible. The combination of several genetic markers can allow better estimates of biodiversity in a given habitat (Drummond et al., 2015; Zhang et al., 2018; Marquina et al., 2019a,b), especially when looking at different phyla or samples with high taxonomic diversity. For example the COI marker is not suitable to identify nematodes and the 18S marker alone would not be suitable to target specific arthropods, due to the specificities of the markers (impairing higher taxon delineation) and available databases. In fact, a recent study found that the combination of at least two markers can improve

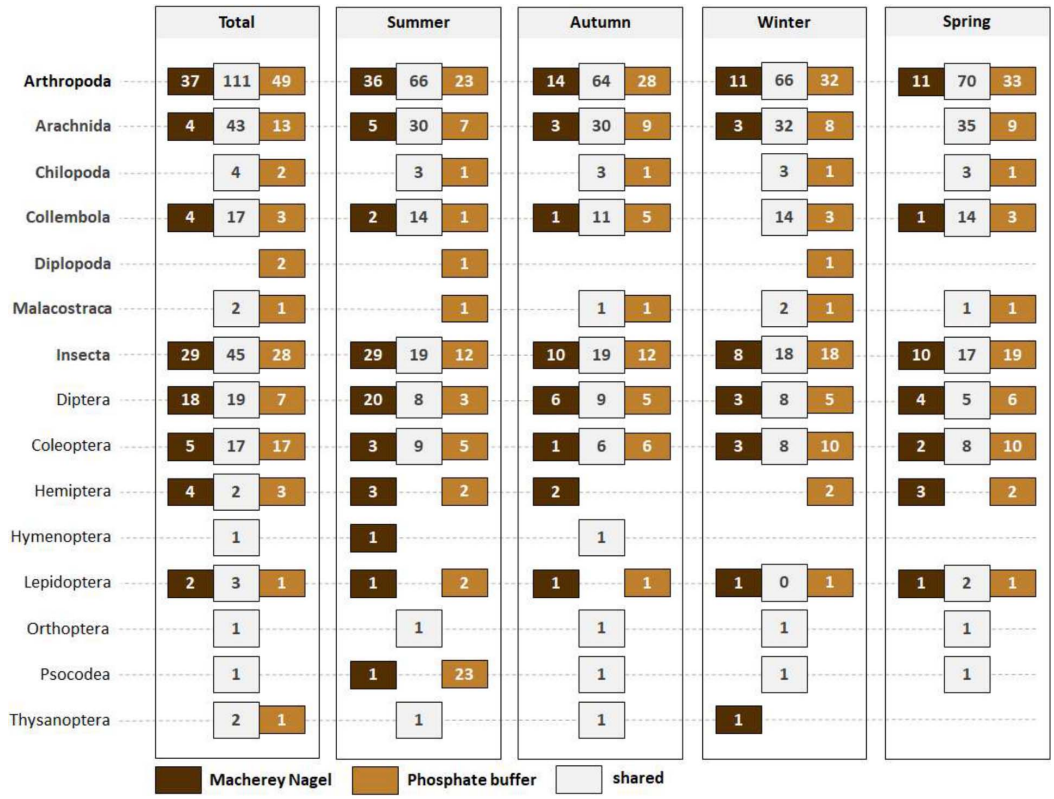


FIGURE 7 | Number of shared and unique arthropod species (blastID ≥99%) found between extraction methods for each season using both markers. The number of species per arthropod class and insect order recovered with either one or both of the two extraction methods is shown. Dark brown: OTUs from Macherey Nagel kit; Light brown: OTUs from phosphate buffer; White box: shared OTUs between Macherey Nagel kit and phosphate buffer.

taxonomic resolution by up to 10% (Marquina et al., 2019a) and can significantly increase the number of target invertebrate taxa. Notwithstanding some studies targeting arthropods using multiple COI primers suggest that when targeting taxonomic groups with limited diversity the use of multiple primer sets could represent unnecessary costs with no substantial improvement in taxon detection (Elbrecht et al., 2019), allowing maximum richness but not affecting beta diversity (Hajibabaei et al., 2019). Despite the fact that the COI barcode covers up to 95% of several groups of organisms (Hajibabaei et al., 2007), it is not an all-purpose answer as its taxonomic resolution and coverage is limited for many invertebrate taxa (Kvist, 2014; Creer et al., 2016). Due to the absence of a COI barcoding gap for earthworms (Bienert et al., 2012; Kvist, 2014) and the low taxonomic resolution of the 18S marker (Tang et al., 2012), none of the 12 annelid species identified were simultaneously retrieved by both markers. This demonstrates how complementary nuclear and mitochondrial markers can be (Drummond et al., 2015; Giebner et al., 2020) and how incorporating these strategies can impact further biodiversity and ecological assessment on a given habitat.

TABLE 1 | SIMPER analysis showing the contribution (%) of Insect OTUs assigned on species level with a blastID of at least 99% to the differences between extraction methods depending on season.

	Summer		Autumn		Winter		Spring	
Species	**%**	**Rank**	**%**	**Rank**	**%**	**Rank**	**%**	**Rank**
Athous subfuscus	10.64	1	9.75	2	12.19	1	12.73	1
Ctenosciara lutea	6.42	2	3.34	8	5.08	6	3.97	8
Corynoptera globiformis	4.30	3	7.10	4	7.75	2	6.10	4
Cratyna perplexa	4.00	4		–		–		–
Strophosoma melanogrammum	3.82	5	4.98	6	5.06	7	5.32	5
Tipula limbata	3.78	6	10.23	1	6.01	5	8.64	2
Tipula scripta	3.13	8	0.85	25		–		–
Othius myrmecophilus	2.83	9	8.52	3	6.36	4	4.24	7
Barypeithes araneiformis	2.71	10	5.97	5	3.23	10	0.97	20
Corynoptera minima	1.90	17	3.05	10	6.83	3	3.69	9
Barypeithes pellucidus	0.43	51	1.88	13	4.12	8	6.47	3

The rank highlights the relative contribution of the corresponding species within the indicated season.

The results presented here highlight that prior knowledge about the target group and an understanding of the methodological trade-offs is required to allow for decisions that can significantly improve taxon detection. Based on our results, we suggest the following recommendations for invertebrate biodiversity assessment from forest soil samples: (1) Choice of marker should be carefully considered based on target groups (e.g., CO1 for arthropods, 18S for nematodes,

platyhelminthes, rotifers, and tardigrades); (2) The use of a phosphate buffer is suitable for the detection of eDNA traces from macro invertebrates which actively interact with their habitat; (3) The use of a lysis based extraction method is more suitable for the detection of micro-invertebrates as well as other life stages of macro invertebrates such as eggs and larvae; (4) Sampling effort can be maximized by combining several DNA extraction methods, but this will add to cost (5) the use of a multi-marker approach (markers or primer pairs, depending on the study objectives) will improve taxon recovery in environmental samples with high taxonomic diversity and concomitantly better reflect biodiversity levels, but this will add to cost; (7) Sampling effort to cover mesofaunal diversity in the forest ecosystem under study should be high (ca. 500 forest soil samples using both extraction methods).

This study adds recommendations on key aspects for processing soil samples, from sampling effort to the importance of the DNA extraction method chosen and the use of a multiplex marker approach, which will allow a better assessment of diversity levels in one of the most species-rich habitats, the soils. We show that eDNA is an effective tool that allows diversity assessments of soil invertebrate communities, but its efficacy relies (but not only) on a combined effect of the method used, the development of specific primer pairs or multiplex approach and the completeness of public databases.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI BioProject, accession no: PRJNA681091.

AUTHOR CONTRIBUTIONS

VF conceived the original idea. VF and SB supervised the project. AK and KL carried out the field and lab work. AK performed the analysis and wrote the manuscript with support from SB and VF. All authors provided critical feedback and helped shape the research, analysis and manuscript.

FUNDING

This study was partly funded by the German Federal Ministry of Education and Research, through the project German Barcode of Life (GBOL1, FKZ01LI1101 and GBOL2, FKZ01LI1501).

ACKNOWLEDGMENTS

We thank Sönke Twietmeyer from the Eifel Nationalpark for his support during the field work phase. We would also like to thank all helpers who went with us to the field and helped us to collect soil samples. This study contains material that has previously formed part of my Ph.D. thesis which will be published according to the requirements of the institution awarding the qualification (University of Bonn).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.630560/full#supplementary-material

Supplementary Figure 1 | Location of the sampling sites. The area highlighted in purple corresponds to the Eifel National Park. Maps were downloaded from https://www.geoportal.nrw/ on 20th of November 2019 and modified according to our purposes using Microsoft Power Point 2016.

Supplementary Figure 2 | Number of collected soil samples per forest type and season. In summer three sampling sites per forest type were sampled. At each sampling site three replicates were taken. In the remaining four seasons triplicates were taken at 14 sampling sites. While three sampling sites were located at the pure beech and young beech sites, respectively, (9 samples per forest type and season), at the pure spruce and old beech forests triplicates were taken at four sampling sites (12 samples were forest type and season). In total 162 samples were collected.

Supplementary Figure 3 | (A) Number of unique and shared species between sampling sites (cardinals) depending on forest type (columns) and season (rows) using the Macherey-Nagel kit. The data shown here comprises all arthropod species detected with either one or both of the two used markers (18S and COI). Only species detected with a BlastID of at least 99% to the reference databases are considered. **(B)** Number of unique and shared species between sampling sites (cardinals) depending on forest type (columns) and season (rows) based on the extraction with the phosphate buffer. The data shown here comprises all arthropod species detected with either one or both of the two used markers (18S and COI). Only species detected with a blastID of at least 99% to the reference databases are considered.

Supplementary Figure 4 | Number of detected dipteran species (blastID $\geq$ 99%) per family depending on extraction method.

Supplementary Figure 5 | Number of detected dipteran species (blastID 99%) per family in summer season depending on extraction method.

Supplementary Table 1 | Geographical location and ecological characteristics of the 14 sampling sites. For each sampling site, the coordinates (altitude N and latitude E) and the associated forest type are specified.

Supplementary Table 2 | Time of sampling. For each sampling period the season, associated time of the year, number of samples taken and sampling dates (d,m,y) are specified.

Supplementary Table 3 | Presence/absence OTU Table for the COI marker. The table contains information on OTU occurrence at each sampling site. Each column (2–324) represents one sample. The name of each sample (XXYYrZSX) indicates extraction method (XX), sampling site (YY), replicate number (rZ) and collection season (SX). Taxonomic assignment (Blast ID%) for assigned OTUs is indicated.

Supplementary Table 4 | Presence/absence OTU Table for the COI marker. The table contains information on OTU occurrence at each sampling site. Each column (2–324) represents one sample. The name of each sample (XXYYrZSX) indicates extraction method (XX), sampling site (YY), replicate number (rZ) and collection season (SX). Taxonomic assignment (Blast ID%) for assigned OTUs is indicated.

Supplementary Table 5 | presence/absence list of OTUs assigned at species level (blastID $\geq$ 90%) with either of the two extraction methods and marker. The table contains information on species occurrence at each sampling site. Each column (2–324) represents one sample. The name of each sample (XXYYrZSX) indicates extraction method (XX), sampling site (YY), replicate number (rZ), and collection season (SX).

REFERENCES

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., and Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods Ecol. Evol.* 9, 134–147. doi: 10.1111/2041-210x.12849

Altmüller, R. (1977). "Ökoenergetische untersuchungen an dipterenpopulationen im buchenwald," in *Verhandlungen der Gesellschaft für Ökologie, Göttingen 1976: 6. Jahresversammlung vom 20. bis 24. September 1976 in Göttingen*, ed. P. Müller, (Dordrecht: Springer Netherlands), 133–138. doi: 10.1007/978-94-011-5957-9_13

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.

Barnard, P. C. (2011). *The Royal Entomological Society Book of British Insects*. New York, NY: John Wiley & Sons.

Baselga, A., and Orme, C. D. L. (2012). betapart: an R package for the study of beta diversity. *Methods Ecol. Evol.* 3, 808–812. doi: 10.1111/j.2041-210x.2012.00224.x

Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., Brun, J., et al. (2012). Tracking earthworm communities from soil DNA. *Mol. Ecol.* 21, 2017–2030. doi: 10.1111/j.1365-294x.2011.05407.x

Brandon-Mong, G.-J., Gan, H.-M., Sing, K.-W., Lee, P.-S., Lim, P.-E., and Wilson, J.-J. (2015). DNA metabarcoding of insects and allies: an evaluation of primers and pipelines. *Bull. Entomol. Res.* 105, 717–727. doi: 10.1017/s0007485315000681

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Chen, H., and Boutros, P. C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinform.* 12:35. doi: 10.1186/1471-2105-12-35

Clarke, L. J., Beard, J. M., Swadling, K. M., and Deagle, B. E. (2017). Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. *Ecol. Evol.* 7, 873–883. doi: 10.1002/ece3.2667

Coleman, D. C., Crossley, D. A. Jr., and Hendrix, P. F. (2004). *Fundamentals of Soil Ecology*. London: Academic Press.

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., et al. (2016). The ecologist's field guide to sequence-based identification of biodiversity. *Methods Ecol. Evol.* 7, 1008–1018. doi: 10.1111/2041-210x.12574

Davidson, D., Grieve, I., and Young, I. (2002). Impacts of fauna on an upland grassland soil as determined by micromorphological analysis. *Appl. Soil Ecol.* 20, 133–143. doi: 10.1016/s0929-1393(02)00017-3

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., and Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biol. Lett.* 10:20140562. doi: 10.1098/rsbl.2014.0562

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., et al. (2017). Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Mol. Ecol.* 26, 5872–5895. doi: 10.1111/mec.14350

Delgado-Baquerizo, M., Powell, J. R., Hamonts, K., Reith, F., Mele, P., Brown, M. V., et al. (2017). Circular linkages between soil biodiversity, fertility and plant productivity are limited to topsoil at the continental scale. *New Phytol.* 215, 1186–1196. doi: 10.1111/nph.14634

Delmont, T. O., Robe, P., Cecillon, S., Clark, I. M., Constancias, F., Simonet, P., et al. (2011). Accessing the soil metagenome for studies of microbial diversity. *Appl. Environ. Microbiol.* 77, 1315–1324. doi: 10.1128/aem.01526-10

Disney, H. (2012). *Scuttle Flies: The Phoridae*. Berlin: Springer Science & Business Media.

Dopheide, A., Xie, D., Buckley, T. R., Drummond, A. J., and Newcomb, R. D. (2019). Impacts of DNA extraction and PCR on DNA metabarcoding estimates of soil biodiversity. *Methods Ecol. Evol.* 10, 120–133. doi: 10.1111/2041-210x.13086

Drummond, A. J., Newcomb, R. D., Buckley, T. R., Xie, D., Dopheide, A., Potter, B. C., et al. (2015). Evaluating a multigene environmental DNA approach for biodiversity assessment. *GigaScience* 4:s13742-15.

Elbrecht, V., Braukmann, T. W., Ivanova, N. V., Prosser, S. W., Hajibabaei, M., Wright, M., et al. (2019). Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ* 7:e27801v1.

Elbrecht, V., Peinert, B., and Leese, F. (2017). Sorting things out: assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecol. Evol.* 7, 6918–6926. doi: 10.1002/ece3.3192

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., et al. (2012). New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Mol. Ecol.* 21, 1821–1833. doi: 10.1111/j.1365-294x.2012.05537.x

Fonseca, V. G., and Lallias, D. (2016). Metabarcoding marine sediments: preparation of amplicon libraries. *Methods Mol Biol.* 1452, 183–196. doi: 10.1007/978-1-4939-3774-5_12

Frøslev, T. G., Kjøller, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., et al. (2017). Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nat. Commun.* 8, 1–11. doi: 10.1038/s41467-017-01312-x

Frostegård, Å, Courtois, S., Ramisse, V., Clerc, S., Bernillon, D., Le Gall, F., et al. (1999). Quantification of bias related to the extraction of DNA directly from soils. *Appl. Environ. Microbiol.* 65, 5409–5420. doi: 10.1128/aem.65.12.5409-5420.1999

Gibert, C., and Escarguel, G. (2019). PER-SIMPER—a new tool for inferring community assembly processes from taxon occurrences. *Glob. Ecol. Biogeogr.* 28, 374–385. doi: 10.1111/geb.12859

Giebner, H., Langen, K., Bourlat, S., Kukowka, S., Mayer, C., Astrin, J., et al. (2020). Comparing diversity levels in environmental samples: DNA sequence capture and metabarcoding approaches using 18S and COI genes. *Mol. Ecol. Resourc.* 20, 1333–1345. doi: 10.1111/1755-0998.13201

Hajibabaei, M., Porter, T. M., Wright, M., and Rudar, J. (2019). COI metabarcoding primer choice affects richness and recovery of indicator taxa in freshwater systems. *PLoS One* 14:e0220953. doi: 10.1371/journal.pone.0220953

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. A., and Baird, D. J. (2011). Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS One* 6:e17497. doi: 10.1371/journal.pone.0017497

Hajibabaei, M., Singer, G. A., Clare, E. L. and Hebert, P. D. N. (2007). Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. *BMC Biol.* 5:24. doi: 10.1186/1741-7007-5-24

Hebert, P. D. N., Cywinska, A., Ball, S. L., and deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 270, 313–321. doi: 10.1098/rspb.2002.2218

Hsieh, T., Ma, K., and Chao, A. (2016). iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). *Methods Ecol. Evol.* 7, 1451–1456. doi: 10.1111/2041-210x.12613

Jakovlev, J. (2012). Fungal hosts of mycetophilids (Diptera: Sciaroidea excluding Sciaridae): a review. *Mycology* 3, 11–23.

Kang, S., and Mills, A. L. (2006). The effect of sample size in studies of soil microbial community structure. *J. Microbiol. Methods* 66, 242–250. doi: 10.1016/j.mimet.2005.11.013

Kress, W. J., and Erickson, D. L. eds (2012). "DNA barcodes: methods and protocols," in *DNA Barcodes*, (Berlin: Springer), 3–8. doi: 10.1007/978-1-61779-591-6_1

Kuhlman, T., Reinhard, S., and Gaaff, A. (2010). Estimating the costs and benefits of soil conservation in Europe. *Land Use Policy* 27, 22–32. doi: 10.1016/j.landusepol.2008.08.002

Kvist, S. (2014). Does a global DNA barcoding gap exist in Annelida? *Mitochondr. DNA Part A* 27, 2241–2252.

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., et al. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Front. Zool.* 10:34. doi: 10.1186/1742-9994-10-34

Marquina, D., Andersson, A. F., and Ronquist, F. (2019a). New mitochondrial primers for metabarcoding of insects, designed and evaluated using in silico methods. *Mol. Ecol. Resourc.* 19, 90–104. doi: 10.1111/1755-0998.12942

Marquina, D., Esparza-Salas, R., Roslin, T., and Ronquist, F. (2019b). Establishing arthropod community composition using metabarcoding: surprising inconsistencies between soil samples and preservative ethanol and homogenate from Malaise trap catches. *Mol. Ecol. Resourc.* 19, 1516–1530. doi: 10.1111/1755-0998.13071

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Neher, D., Weicht, T., and Barbercheck, M. E. (2012). Linking invertebrate communities to decomposition rate and nitrogen availability in pine forest soils. *Appl. Soil Ecol.* 54, 14–23. doi: 10.1016/j.apsoil.2011.12.001

Nielsen, U. N., Osler, G. H., Campbell, C. D., Neilson, R., Burslem, D. F., and Van der Wal, R. (2010). The enigma of soil animal species diversity revisited: the role of small-scale heterogeneity. *PLoS One* 5:e11567. doi: 10.1371/journal.pone.0011567

Oliverio, A. M., Gan, H., Wickings, K., and Fierer, N. (2018). A DNA metabarcoding approach to characterize soil arthropod communities. *Soil Biol. Biochem.* 125, 37–43. doi: 10.1016/j.soilbio.2018.06.026

Pietramellara, G., Ascher, J., Borgogni, F., Ceccherini, M., Guerri, G., and Nannipieri, P. (2009). Extracellular DNA in soil and sediment: fate and ecological relevance. *Biol. Fertil. Soils* 45, 219–235. doi: 10.1007/s00374-008-0345-8

Porter, T. M., Morris, D. M., Basiliko, N., Hajibabaei, M., Doucet, D., Bowman, S., et al. (2019). Variations in terrestrial arthropod DnA metabarcoding methods recovers robust beta diversity but variable richness and site indicators. *Sci. Rep.* 9, 1–11.

Potter, C., Tang, C. Q., Fonseca, V., Lallias, D., Gaspar, J. M., Thomas, K., et al. (2017). De novo species delimitation in metabarcoding datasets using ecology and phylogeny. *PeerJ* 5:e3121v1.

R Core Team, (2013). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584. doi: 10.7717/peerj.2584

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D., Breiner, H., et al. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* 19, 21–31. doi: 10.1111/j.1365-294x.2009.04480.x

Taberlet, P., PRUD'HOMME, S. M., Campione, E., Roy, J., Miquel, C., Shehzad, W., et al. (2012). Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. *Mol. Ecol.* 21, 1816–1820. doi: 10.1111/j.1365-294x.2011.05317.x

Tang, C. Q., Leasi, F., Obertegger, U., Kieneke, A., Barraclough, T. G., and Fontaneto, D. (2012). The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16208–16212. doi: 10.1073/pnas.1209160109

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Berlin: Springer.

Wickham, H., Francois, R., Henry, L., and Müller, K. (2015). *dplyr: A Grammar of Data Manipulation. R Package Version 0.8.3* 3.

Yang, C., Ji, Y., Wang, X., Yang, C., and Douglas, W. Y. (2013). Testing three pipelines for 18S rDNA-based metabarcoding of soil faunal diversity. *Sci. China Life Sci.* 56, 73–81. doi: 10.1007/s11427-012-4423-7

Yang, C., Wang, X., Miller, J. A., de Blécourt, M., Ji, Y., Yang, C., et al. (2014). Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator. *Ecol. Indicat.* 46, 379–389. doi: 10.1016/j.ecolind.2014.06.028

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., et al. (2012). Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods Ecol. Evol.* 3, 613–623. doi: 10.1111/j.2041-210X.2012.00198.x

Zhang, G. K., Chain, F. J., Abbott, C. L., and Cristescu, M. E. (2018). Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities. *Evol. Appl.* 11, 1901–1914. doi: 10.1111/eva.12694

Zinger, L., Chave, J., Coissac, E., Iribar, A., Louisanna, E., Manzi, S., et al. (2016). Extracellular DNA extraction is a fast, cheap and reliable alternative for multi-taxa surveys based on soil DNA. *Soil Biol. Biochem.* 96, 16–19. doi: 10.1016/j.soilbio.2016.01.008

Zinger, L., Taberlet, P., Schimann, H., Bonin, A., Boyer, F., De Barba, M., et al. (2019). Body size determines soil community assembly in a tropical forest. *Mol. Ecol.* 28, 528–543. doi: 10.1111/mec.14919

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 04 June 2021
doi: 10.3389/fevo.2021.632973

The Application of PMA (Propidium Monoazide) to Different Target Sequence Lengths of Zebrafish eDNA: A New Approach Aimed Toward Improving Environmental DNA Ecology and Biological Surveillance

*Takaya Hirohara[1], Kenji Tsuri[1], Koichi Miyagawa[1], Robert T. R. Paine[2] and Hiroki Yamanaka[1,3]**

[1] Department of Environmental Solution Technology, Faculty of Science and Technology, Ryukoku University, Otsu, Japan, [2] Cooperative Fishery Research Unit, Tennessee Technological University, Cookeville, TN, United States, [3] Center for Biodiversity Science, Ryukoku University, Otsu, Japan

OPEN ACCESS

Edited by:
Richard Lance, U.S. Army Engineer Research and Development Center, United States

Reviewed by:
Chris Wilson, Ontario Ministry of Natural Resources and Forestry, Canada
Tomasz Suchan, Władysław Szafer Institute of Botany, Polish Academy of Sciences (PAN), Poland

***Correspondence:**
Hiroki Yamanaka
yamanaka@rins.ryukoku.ac.jp

Specialty section:
This article was submitted to Conservation and Restoration Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *24 November 2020*
Accepted: *16 April 2021*
Published: *04 June 2021*

Citation:
Hirohara T, Tsuri K, Miyagawa K, Paine RTR and Yamanaka H (2021) The Application of PMA (Propidium Monoazide) to Different Target Sequence Lengths of Zebrafish eDNA: A New Approach Aimed Toward Improving Environmental DNA Ecology and Biological Surveillance. Front. Ecol. Evol. 9:632973. doi: 10.3389/fevo.2021.632973

Environmental DNA (eDNA) analysis has enabled more sensitive and efficient biological monitoring than traditional methods. However, since the target species is not directly observed, interpretation of results cannot preclude process Type I errors. Specifically, there may be a spatial or temporal gap between the target eDNA and the eDNA source in the sampled area. Moreover, eDNA surveillance lacks the ability to distinguish whether eDNA originated from a living or non-living source. This kind of Type I error is difficult to control for, in part, because the relationship between the state of eDNA (i.e., intracellular or extracellular) and the degradation rate is still unclear. Here, we applied PMA (Propidium monoazide) to eDNA analysis which enabled us to differentiate "intact cells" from "disrupted cells." PMA is a dye that has a high affinity for double-stranded DNA and forms a covalent bond with double-stranded DNA and inhibits amplification of the bonded DNA molecules by PCR. Since PMA is impermeable to the cell membrane, DNA protected by an intact cell membrane can be selectively detected. In this study, we investigated the workability of PMA on vertebrate eDNA using zebrafish, *Danio rerio*. Aquarium water was incubated for 1 week to monitor the eDNA degradation process of both intracellular and extracellular eDNA. We developed ten species-specific quantitative PCR assays for *D. rerio* with different amplification lengths that enabled independent quantification of total eDNA (sum of intracellular and extracellular eDNA, commonly measured in other studies) and intracellular eDNA (DNA in intact cells) and allow for analyses of sequence length-dependent eDNA degradation in combination with PMA. We confirmed that PMA is effective at differentiating "intact" and "disrupted" fish cells. We found that total eDNA and intracellular eDNA have different degradation processes

that are dependent on the length of target sequence. For future conservation efforts using eDNA analyses, it is necessary to increase the reliability of the analysis results. The research presented here provides new analysis tools that expand our understanding of the ecology of eDNA, so that more accurate and reliable conclusions can be determined.

Keywords: environmental DNA, decay rate, propidium monoazide, quantitative PCR, zebrafish

INTRODUCTION

Aquatic ecosystems are disproportionately affected by anthropogenic influences, such as pollution, habitat degradation, introduction of invasive species, and overuse of natural resources (Abell et al., 2008; Strayer and Dudgeon, 2010; Collen et al., 2014). Conservation efforts that are used to help mitigate the damage caused to aquatic ecosystems have to be empirically monitored to determine which strategies are effective. One way to monitor the effectiveness of various conservation strategies is to survey the biodiversity within a system. Surveillance of biodiversity includes monitoring the distribution of species of interest (e.g., alien or endangered species), estimating population abundance or biomass of targeted species, estimating occupancy at a site, or assessing the presence or absence of species in targeted systems (Lodge et al., 2012; Pilliod et al., 2013, 2014; Thompson, 2013). Traditional monitoring of the activity ranges and habitat usages of aquatic organisms typically includes a variety of surveillance approaches (e.g., direct capture or visual surveys), which employ different gears and techniques that can be difficult to learn and standardize, and require a considerable amount of labor and cost (Hayes et al., 1996; Rees et al., 2014; Hajibabaei et al., 2016; Evans et al., 2017).

There has been growing interest over the last decade in eDNA (environmental DNA) surveillance as a biological monitoring method for aquatic species, due to several advantages over traditional methods. Costs, labor, and environmental disturbance associated with eDNA surveillance are often smaller compared to direct capture or visual surveys because eDNA surveys only require small volumes of water to be collected at targeted sites (Jerde et al., 2011; Takahara et al., 2013; Smart et al., 2015, 2016; Evans et al., 2017). Furthermore, eDNA monitoring has also been demonstrated to have higher detection sensitivity compared to traditional methods, especially at low target species densities (Dejean et al., 2012; Pilliod et al., 2013; Takahara et al., 2013). Since the target species is not directly observed, however, the results of eDNA analyses can potentially include errors in ecological interpretation (process Type I errors) (Darling and Mahon, 2011; Rees et al., 2014; Barnes and Turner, 2016; also see Taberlet et al., 2018). For example, in both natural and experimental systems, legacy eDNA of the target species can still be accurately amplified and detected even though there are no living individuals of the target species in the system (Merkes et al., 2014; Lance et al., 2017; Tsuji et al., 2017; Kamoroff and Goldberg, 2018). Simply, positive detection of legacy eDNA reported as "presence" of the target species are not analytical false-positives (method-based Type I error), but ecological misinterpretations (*sensu* Darling and Mahon, 2011). While method-based Type I errors are fairly easy to control for in eDNA designs (e.g., inclusion of negative controls during all field, extraction, and PCR processes, multiple biological and technical replicates) (Champlot et al., 2010; Darling and Mahon, 2011; Ficetola et al., 2015), there are no empirically tested protocols that allow eDNA surveys to reduce or eliminate process Type I errors, aside from tandem traditional surveillance methods.

Our understanding of eDNA ecology regarding the existing states and degradation rates that influence eDNA detection is limited, in part, due to the inability to demarcate spatial or temporal relationships between eDNA and the source of eDNA. For example, positive eDNA detection could result from DNA derived from dead individuals or transported from another site via bird droppings (Merkes et al., 2014). Furthermore, it is unknown when the detected eDNA was released from the target species or if the individual is still in the immediate surveyed area (Barnes and Turner, 2016; Jo et al., 2017). Several studies using experimental systems have demonstrated that eDNA signals can be detected many days (17–25) after the target organism is removed from the system (Dejean et al., 2011, 2012; Goldberg et al., 2013). Environmental DNA from common carp (*Cyprinus carpio*) remained detectable in sediment over 132 days (Turner et al., 2015). Moreover, eDNA that has an anthropogenic origin could be a source of process Type I error as shown in a reported case where effluent from fish markets along the Maizuru Bay area (Japan) yielded positive eDNA detection for Japanese jack mackerel (*Trachurus japonicus*) (Yamamoto et al., 2016). Clearly, a contributing factor of process Type I error in eDNA analysis and interpretation is the limited spatio-temporal nature of eDNA caused by various aspects associated with the origin, state, and fate of eDNA (Barnes and Turner, 2016).

In addition to the aforementioned issues related to eDNA decay, various biotic/abiotic factors associated with the state of eDNA also impact the detectability and persistence of eDNA in various aquatic environments (Thomsen et al., 2012a,b; Maruyama et al., 2014; Pilliod et al., 2014; Lance et al., 2017; Mächler et al., 2018; Murakami et al., 2019). Turner et al. (2014) suggested that eDNA is released from organisms as relatively large particles (1–10 μm), indicating that eDNA shed from fish is likely within cells and mitochondria, at least at the time point of release. In other words, eDNA is released as intracellular eDNA, which are relatively large particles, that undergo degradation and change their physical state and structure to become smaller (extracellular) particles (Jo et al., 2019). Moreover, Jo et al. (2017) clarified that long eDNA fragments degraded faster than shorter eDNA fragments using two different qPCR assays for Japanese jack mackerel and suggested the potential of longer eDNA fragments as a better proxy for the presence of the target fish. The concentrations of longer eDNA fragments also gave better correlation with estimations of fish distribution/biomass (based

on quantitative echo intensity) than shorter eDNA fragments (Jo et al., 2017). These conclusions suggest that the length of a target sequence in eDNA analysis is a key factor for reliable detection and determination of true presence for the target organism. Thus, the degradation process as it relates to DNA fragment length requires further experimental testing and clarification.

The eDNA degradation patterns observed so far have been measured as the degradation of total eDNA (DNA of target sequence in samples commonly measured in other studies) containing both extra- and intracellular eDNA. It is suggested that DNA molecules within intact cell membranes (intracellular DNA) would be less vulnerable to the attacks from microbes and extracellular enzymes in the environment than extracellular DNA (Levy-Booth et al., 2007). Therefore, the independent examination of both intra- and extracellular eDNA would also be beneficial to understanding the mechanisms of the eDNA degradation process. Clarification of the state of eDNA and the dynamics of degradation will contribute to improve the reliability of eDNA analysis and interpretations for more accurate decision making in future conservation efforts.

Here, we applied PMA (propidium monoazide) to eDNA analysis to differentiate "intact cells" from "disrupted cells," which has been mainly used in microbial research. PMA is a photoreactive dye that has a high affinity for double-stranded DNA and forms a covalent bond with double-stranded DNA when exposed to strong visible light, and inhibits amplification of the bonded DNA molecules by PCR (Nocker et al., 2007). In addition, since PMA is impermeable to the cell membrane, DNA protected by an intact cell membrane can be selectively detected (**Figure 1**; García-Fontana et al., 2016; Emerson et al., 2017). Although there are a few research examples which used PMA on other taxa such as phytoplankton (*Microcystis aeruginosa*, *Anabaena* sp., *Aphanizomenon* sp., *Synechocystis* sp., *Cryptomonas ovata*, *Scenedesmus obliquus*, and *Nitzschia apiculata*; Joo et al., 2019) and shellfish (*Dreissena polymorpha*; Lance and Carr, 2012), to the best of our knowledge, there are no research examples which applied PMA on vertebrates such as fish.

In this study, we investigated the workability of PMA on vertebrate eDNA using zebrafish, *Danio rerio*. Water samples collected from aquaria were boiled to disrupt the membrane of fish cells in the water to make mock "damaged" samples. PMA was applied to each of the boiled samples and the non-boiled samples, and detectable eDNA copy numbers were quantified using quantitative real-time PCR (qPCR). Next, water samples were collected from the aquaria and incubated for 1 week to monitor the eDNA degradation process between intracellular and extracellular eDNA. For examining the length-dependent degradation of eDNA, we developed 10 species-specific qPCR assays for *D. rerio* with different amplification lengths. The combined usage of PMA and the 10 qPCR assays enabled independent quantification of total eDNA (commonly measured in other studies) and intracellular eDNA (DNA in intact cells). We believe that clarification on the eDNA state (i.e., eDNA from intact vs. disrupted cells) and their respective degradation rates will help provide a basis for reducing process Type I errors and the continual improvement of eDNA techniques.

MATERIALS AND METHODS

In this study, two types of experiments were conducted using *D. rerio* as a model organism: (1) confirming if PMA could be effective for preventing amplification of extracellular DNA, and limiting amplification to DNA from intact cells or organelles, (2)

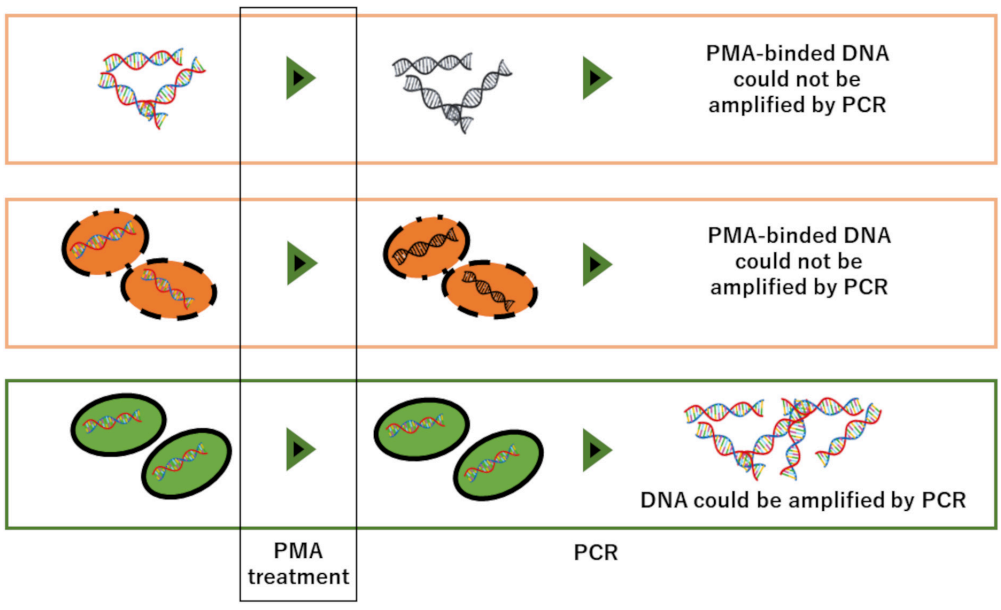


FIGURE 1 | Functional mechanism of PMA-DNA interaction: PMA-bound DNA cannot be amplified by PCR. The cell membrane is impermeable against PMA, thus DNA inside of intact cell membranes are protected from PMA-binding and can be selectively detected in PCR test.

determining the degradation rate for total eDNA (intracellular and extracellular) and for eDNA in intact cells only, using PMA. In experiment 1, a control experiment was performed to quantify eDNA in PMA-treated and non-treated samples by qPCR. In experiment 2, the time-dependent degradation of eDNA was measured for total eDNA and intracellular eDNA by analyzing PMA-treated and non-treated eDNA samples from *D. rerio* tank water. Water samples obtained from the *D. rerio* tank were stored in a water bath and incubated at the same temperature as the *D. rerio* tank. Thereafter, the incubated water samples were sampled in a time series (0, 1, 2, 3, 5, and 7 days). After filtration, extraction and purification of eDNA from the sample, DNA copy numbers were quantified by qPCR using 10 newly developed qPCR assays with different amplification lengths which share the same probe and reverse primer but have 10 different forward primers. The DNA concentrations from all samples were analyzed to compare the degradation rate and degradation pattern of each treatment. Copy numbers of intracellular eDNA and total eDNA (intra- and extracellular eDNA) were all measured as the copy numbers of target sequence fragments in the eDNA samples by qPCR using the new assays.

Assay Design

Sequences of the mitochondrial genes including cyt*b* (cytochrome *b*), tRNA-Glu and ND6 of the target species *D. rerio* as well as the other non-target species were downloaded from NCBI (the National Center for Biotechnology Information)[1]. The accession numbers for both target and non-target species used are shown in **Supplementary Table 1**. The non-target species were *Cyprinus carpio, Carassius auratus auratus, Opsariichthys platypus,* which are kept in our fish keeping facility on a regular basis and belong to the same family (*Cyprinidae*) as *D. rerio*. The sequences were aligned using MAFFT v. 7[2] with default settings. We searched for base pair differences between *D. rerio* and other non-target species in the 3′ end of each primer to generate species-specific assays. Expected amplicon lengths generated by each assay were 132, 225, 333, 430, 529, 621, 715, 823, 935, and 1,021 bp, respectively (**Table 1**). All primers had a minimum of 2 bp differences between the target species and all non-target species, with at least 1 bp difference in the last 5 bases of the 3′ end, except the forward primer for 529 bp assay (**Supplementary Figure 1**). The probe had 2 bp mismatches between the target species and other non-target species. The probe was designed between the primers for the shortest target sequence length (132 bp) and therefore can be used for all assays (**Supplementary Figure 1**). We confirmed the specificity of the primer pairs *in silico* using the Primer-BLAST with default settings[3]. The workability of the assays was tested on the target *in vitro* by qPCR in triplicate with a 10-fold dilution series generated from an artificially synthesized DNA fragment (gBlocks^{TM}: Integrated DNA Technologies, Inc., Coralville, IA, United States). We defined the limit of detection (LOD) as the lowest number of copies that could be detected in one of the three qPCR replicas and the limit of quantification (LOQ) as the lowest number of copies that could be detected in all of the three qPCR replicas (**Table 1**). The details of qPCR settings are shown in a section "qPCR Conditions."

Experiment 1: PMA Confirmation Test

An acrylic 54 L aquarium was filled with 20 L of dechlorinated tap water and used as an experimental tank and *D. rerio* was held at a density of 5 individuals · L^{-1}. Water temperature was adjusted to 25°C and a photoperiod was set to light: dark = 12:12 h. An air stone and a heater were installed, and water sampling was performed after fish were acclimatized for 3 days. All equipment, including the aquarium used for the sampling, were all decontaminated with 10% bleach solution prior to use.

A disposable plastic cup was used to transfer 6 L of water from the aquarium into a 10 L plastic container. The 6 L water sample was thoroughly mixed by shaking, and was divided into two 3 L samples (contained in 4 L plastic containers). One 3 L sample

[1]https://www.ncbi.nlm.nih.gov

[2]https://mafft.cbrc.jp/alignment/server/index.html

[3]https://www.ncbi.nlm.nih.gov/tools/primer-blast/

TABLE 1 | Sequences of primers and a probe used in this study.

Target species	Primer/Probe	Name of primer/probe	Oligo name	Sequence (5′–>3′)	Amplicon length	LOD (Copies)	LOQ (Copies)
Danio rerio	Forward	Dre-cytb-V132F	F1	CTTACGTGGGAGATACCCTAGTG	132 bp	30	30
	Forward	Dre-cytb-V225F	F2	TAATAATAACAGCTTTTGTGGGCTACG	225 bp	3	3
	Forward	Dre-cytb-V333F	F3	CTTCCTTCTTCTTCATCTGCCTG	333 bp	30	30
	Forward	Dre-cytb-V430F	F4	TACACCTCAGACATCTCAACAGC	430 bp	3	3
	Forward	Dre-cytb- V529F	F5	CCAACGCCACTAAATATTTCAGCG	529 bp	30	30
	Forward	Dre-tRNA-Glu-V621F	F6	TGTTGTAGTTCAACTACAAGAACTGC	621 bp	3	3
	Forward	Dre-ND6-V715F	F7	CATACCCCCAACTAGAGCTGC	715 bp	3	30
	Forward	Dre-ND6-V823F	F8	AGACAAAAATGAACCCCCATAACTAAC	823 bp	3	3
	Forward	Dre-ND6-V935F	F9	GTCAAAACACCACACGGTCAC	935 bp	3	3
	Forward	Dre-ND6-V1021F	F10	AGCATCAACCGATATTAATAAACCAGTG	1021 bp	1	30
	Reverse	Dre-cytb-VR	R	GCAAGTGTAGAATAACTATGGCGATG			
	Probe	Dre-cytb-VPr-FAMZEN	Pr	FAM-ACAATGCAA-ZEN-CCCTTACACGATTCTT CGCATTCC-3IABkFQ			

was boiled for 10 min to disrupt the membrane of fish cells in the water (Martin et al., 2013). The other 3 L sample was left at room temperature while the boiling treatment was performed. Both of the 3 L samples were divided into six subsamples with 500 mL each in new disposable plastic cups, and filtered through Sterivex cartridge filters (pore size 0.45 μm; Merck, Darmstadt, Germany) using a sterile 50 mL luer-lock syringe (SS-50LZ; Terumo Co., Tokyo, Japan). The water was aspirated into the syringe from the disposable cup and then the Sterivex filter is connected to the syringe. The water was then slowly forced through the filter that was inside the Sterivex filter. The water inside the Sterivex filter was removed by pushing air into the Sterivex filter three times using the syringe. The inlet and outlet port on the Sterivex filter were capped (VRMP6 and VRSP6, respectively, ISIS Co., Osaka, Japan). This operation was repeated until all water samples were processed. The Sterivex filters were all stored at room temperature until all filtrations were complete.

PMA treatment (see section "PMA Treatment" for details) was performed on 3 of 6 samples for each of the boiled and non-boiled controls following the procedure shown in **Figure 2A**. The remaining three samples from each of the boiled and non-boiled controls were not treated with PMA. All DNA extractions from all samples were performed just after the PMA processing and stored at −20°C until qPCR testing. The details of PMA treatment and DNA extraction are shown in sections "PMA Treatment" and "DNA Extraction" respectively. **Figure 2A** shows the flowchart of experiment 1.

Experiment 2: Elucidation of Degradation Process of Intracellular/Extracellular eDNA

Aquarium conditions for experiment 2 matched those in experiment 1 (section "Experiment 1: PMA Confirmation Test") with two exceptions. The volume of water was 30 L, and the fish were kept at a density of 3 individuals $\cdot$ L^{-1}. We reduced the density of fish in experiment 2 based on the results of experiment 1, because a sufficient DNA copy number was expected to be obtained for quantification by qPCR.

Using a disposable plastic cup, 18 L of water was transferred from the aquarium into a 20 L plastic container. The 18 L water sample was thoroughly mixed by shaking and 500 mL samples (n = 36) were dispensed into disposable plastic packs (DP16-TN1000; Yanagi Co., Nagoya, Japan) and stored in a water bath maintained at 25.14 ± 0.12°C using a heater for incubation. We randomly selected six samples and one FNC (filtration negative control: 500 mL ultrapure water) to process for each point in our time-series (day 0, 1, 2, 3, 5, and 7). All six samples and one FNC for each time-point were filtered through a Sterivex filter. The filtration was performed following Miya et al. (2016) with minor modifications. Briefly, the lid of the plastic pack was replaced with a rubber cap before filtration. A plastic needle with a silicon tube (4987458150067; NIPRO Co., Osaka, Japan) was put into the rubber cap and the end of the other side of the tube was connected to the inlet port of Sterivex filter via plastic luer connector (VR306; ISIS Co.). The outlet port of the Sterivex filter was connected to a vacuum tube with a 10 μL pipette tip. The inlet and outlet ports of the Sterivex filters were capped with VRMP6 and VRSP6, respectively, after filtration and stored at room temperature until all filtrations were completed. PMA treatment was performed on three of the six subsamples from each time-point. The remaining three subsamples from each time-point were not treated with PMA for comparison. DNA was extracted from all samples just after the PMA treatment (see section "PMA Treatment" for details) and stored at −20°C until qPCR testing. **Figure 2B** shows the flowchart of experiment 2.

PMA Treatment

Two milliliters of 50 μM PMA (PMAxx^{TM}: Biotium, Fremont, CA, United States; adjusted to the concentration with ultrapure water) was added into each Sterivex filter, which was assigned to PMA treatment, from the inlet port using a thin pipette tip. Each Sterivex filter was incubated at room temperature for 10 min while covered by aluminum foil to avoid premature exposure to the light. Then, each Sterivex filter assigned to PMA treatment was exposed to strong visible light (465–470 nm wavelength) for 15 min, activating PMA to form a covalent bond with double-stranded DNA, while non-PMA samples (including FNC) were continuously covered by aluminum foil to avoid premature exposure to the light and stored at room temperature during the PMA treatment. We confirmed that there was not a significant difference in DNA copy numbers between the controls kept in dark and light during non-PMA treatment in a preliminary experiment. Adequate wavelength and sufficient luminance were attained using LED bulbs that met the requirements for the PMA reaction based on information provided from the manufacturer (Biotium). The functional mechanism of the PMA-eDNA interaction is shown in **Figure 1**. As all the residues containing eDNA are trapped on the surface of the tubular filter in the Sterivex cartridge, we performed PMA treatment by irradiating the filter surface with visible light from all directions from the outside of the cartridge, referring to Ribeiro et al. (2019) that performed PMA treatment on the eDNA samples trapped on sieve clothes.

DNA Extraction

DNA was extracted from all Sterivex filters using a commercial DNA extraction kit (DNeasy Blood and Tissue Kit: Qiagen, Hilden, Germany) following Miya et al. (2016) with minor modifications. Briefly, we removed the cap from the outlet port of the Sterivex filter and inserted the outlet port into a 3 mL test tube (cat.No.5821-255, WATSON Co., Tokyo, Japan). The Sterivex filter was connected in the test tube using surgical grade tape (1530SP-1, 3M Japan Limited, Tokyo, Japan). The combined Sterivex-test tube unit was centrifuged at 4,000 × g for 2 min to remove PMA solution from PMA-treated samples and the remaining water from non-PMA treated samples. The 3 mL test tube was removed from Sterivex filter and discarded, and the outlet port of Sterivex filter was recapped. Then, the inlet port cap was removed, and 20 μL of proteinase-K and 200 μL of Buffer AL was added to Sterivex filter using a thin pipette tip. The inlet port of Sterivex filter was recapped and the Sterivex filter was incubated for 20 min at 56°C on a shaker at 20 rpm (ROLLER6 digital, IKA, Staufenberg, Germany). After the incubation, the

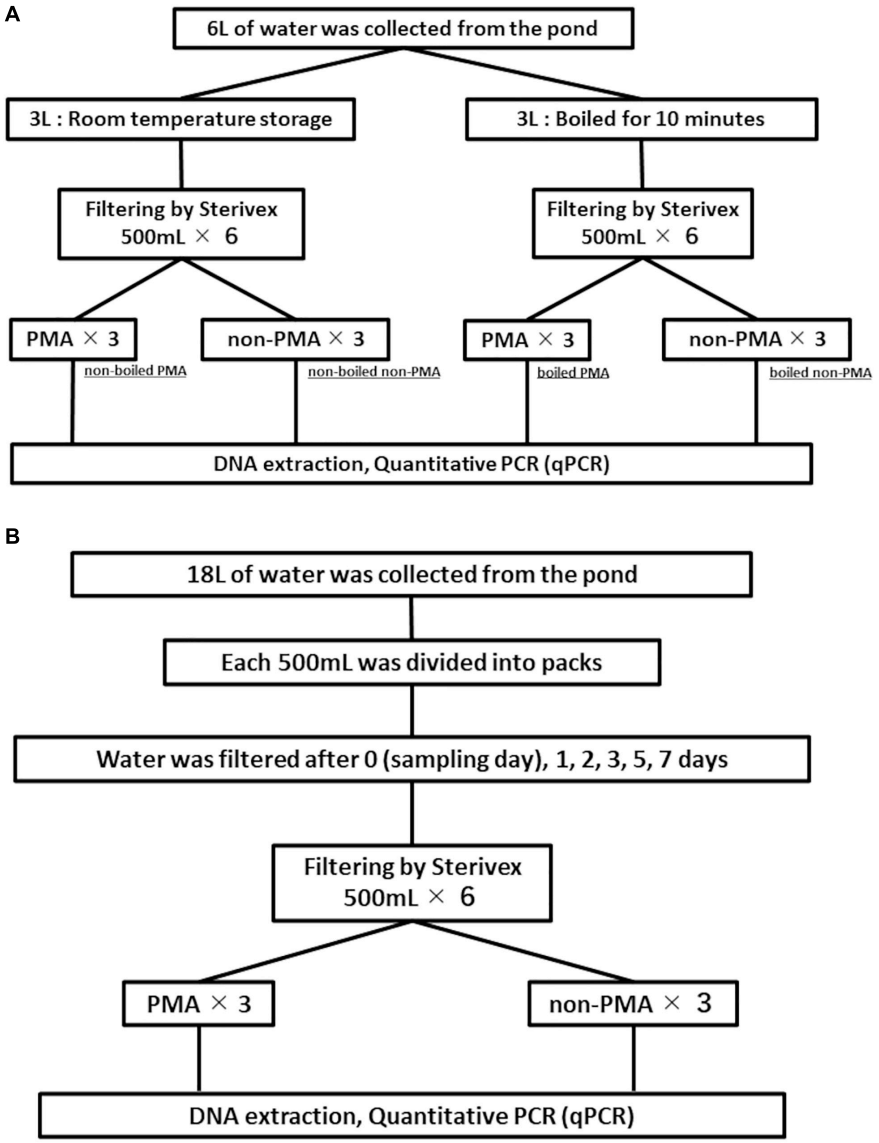


FIGURE 2 | Flowcharts for experiment 1 **(A)** and experiment 2 **(B)**. Filtration negative controls are not included in the figure.

inlet cap was removed and the inlet port side of the Sterivex filter was connected to a 3 mL test tube. The Sterivex-test tube unit was tightly connected using surgical grade tape and centrifuged at 4,000 × *g* for 2 min to collect the DNA solution from Sterivex filter. The Sterivex filter was discarded and 200 μL of absolute ethanol was added to the filtrate in the 3 mL test tube. The DNA in the mixture was purified using a DNeasy kit following the manufacturer's instruction. During the final elution step, DNA trapped on the silica membrane of the spin column was eluted with 100 μL of Buffer AE. The buffers (Buffer AL and AE) and the proteinase K were provided from the DNeasy kit.

qPCR Conditions

We used StepOnePlus® Real-Time PCR System (Life Technologies, Carlsbad, CA, United States) for qPCR. All qPCR reactions were performed in a total volume of 12 μL, which included 6-μL of 2 × TaqPath^{TM} qPCR Master Mix (Thermo Fisher Scientific, Waltham, United States), 900 nM of each primer and 125 nM of a probe at final concentrations, a 2 μL DNA template and ultrapure water to adjust the total volume. All qPCR was conducted as singleplex with triplicated technical replications. A 1,085 bp fragment of the *D. rerio* mitochondrial DNA sequence was synthesized as gBlocks^{TM} Gene Fragments (Integrated DNA Technologies Inc.) and used to develop a standard curve for DNA quantification. The synthesized fragment crosses small regions of the three genes (cytb, tRNA-Glu, and ND6) and includes the 10 qPCR assays described in section "Experiment 1: PMA Confirmation Test." The standards were adjusted to the copy numbers of 3×10^1 – 3×10^4 copies per reaction and were included in triplicate in each qPCR run.

Thermal conditions were as follows: 2 min at 50°C and 10 min at 95°C followed by 55 cycles of 95°C for 15 s and 64°C for 90 s. The cycling temperature was determined based on a preliminary experiment to enable all assays to work properly under the same conditions. Negative controls (NTC: non-template control) were conducted in triplicate in all qPCR assays for all qPCR runs, to assess the occurrence of unintended cross contamination using ultrapure water instead of the DNA template. The r^2 values of the standard curve for qPCR exceeded 0.98 in all runs in experiment 1 and 2. In addition, the average slope and y-intercept of the standard curve for qPCR were -3.54 ± 0.16 and 37.28 ± 1.38 (average $\pm$ standard deviation), respectively.

Statistical Analysis

The eDNA copy number was calculated by averaging technical replicates for each sample. If any of the triplicate reactions for any sample had no amplification, then the copy number for that reaction was regarded as zero and included into the calculation of the average (Ellison et al., 2006). Also, data below the LOQ were excluded from the following analysis. We used R version 3.6.1 for all the statistical analyses (R Core Team, 2019). For experiment 1, analysis of variance (ANOVA) was conducted to test the effect of boiling treatment, PMA treatment, and target sequence length on the eDNA copy number. All of the interactions among the factors were also included to the test as factors.

For experiment 2, we performed multivariate analysis of variance (MANOVA) and *post hoc* ANOVAs to test the effect of sampling time and target sequence length on the eDNA copy numbers in the non-PMA and PMA treated samples. MANOVA can simultaneously evaluate the effects of each factor on multiple response variables and reduce the likelihood of Type I errors and increase the statistical power (Warne, 2014). A general linear model (GLM) was used to examine how eDNA copy number changes depending on the target sequence length. Furthermore, in order to analyze the eDNA degradation, the degradation constant and the half-life were calculated as indices instead of the time to be undetected which should depend on the initial copy number of eDNA, allowing comparisons between studies (Maruyama et al., 2014; Sassoubre et al., 2016). We assumed that DNA degrades at a constant rate over time, exhibiting exponential degradation, which can be modeled by the following equation:

$$N(t) = N_0 * e^{-\lambda t}$$

where N(t) is the estimated number of copies of eDNA at time t, N_0 is the number of copies of eDNA at time zero, λ is the decay rate (sometimes denoted as β; Thomsen et al., 2012b; Maruyama et al., 2014), and t is time (i.e., the number of days). Using the nls function in R, the eDNA copy number obtained over the time of each target sequence length in experiment 2 was fitted to an exponential decay curve to calculate the decay rate. The half-life was calculated as follows:

$$t_{1/2} = \frac{\ln(2)}{\lambda}$$

RESULTS

Primer-Probe Design

Both qPCR and agarose gel electrophoresis results indicated that *D. rerio* DNA was successfully amplified by all 10 assays. However, *in silico* tests suggested that a single *D. nigrofasciatus* sequence was potentially amplified with the primer sets F6–F10 (Accession number: KR606519.1). This species had never been kept at our facility nor is it common as an experimental organism, so it was determined that it did not affect the results of this study.

Experiment 1: PMA Confirmation Test

DNA of *D. rerio* was detected by all assays (F1–F10) in non-boiled-non-PMA, non-boiled-PMA and boiled-non-PMA treatments. Environmental DNA was detected and quantified with only two assays (F1 and F2; 132 and 225 bp) in boiled-PMA samples, while the eDNA copy numbers for assays F3 – F10 (333 bp or longer) fell below the LOQ (**Tables 2**, **3** and **Figure 3**). Samples that fell below the LOQ were excluded from further analyses. For non-boiled-PMA treated samples, the percentage decrease in mean eDNA copy number ranged from (95.04–99.17%), where greater decreases in copy number corresponded with increasing target sequence length (**Table 2** and **Figure 3**). The mean eDNA copy number for non-boiled-non-PMA treated samples did not significantly change as target sequence length increased (**Table 2** and **Figure 3**). Comparatively, boiled-PMA treated samples had a greater decrease in mean eDNA copy number, 98.72% (132 bp) and 99.55% (225 bp), while the remaining assays (333 bp or longer) fell below the LOQ. Similar to the non-boiled-non-PMA samples, the mean eDNA copy number for boiled-non-PMA treated samples also did not statistically change (**Figure 3**) across all target sequence lengths (**Table 2** and **Figure 3**), although the overall mean eDNA copy number was lower. The ANOVA test indicated that eDNA copy number was significantly affected by boiling treatment, where both total eDNA (non-PMA) and intracellular eDNA (PMA) copy number decreased significantly after boiling (all $p < 0.05$; **Table 3** and **Figure 3**). All FNCs and NTCs were negative for the amplification of *D. rerio* DNA. All of the raw data obtained by qPCR are shown in **Supplementary Table 3**.

Experiment 2: Elucidation of the Degradation Process of Intracellular eDNA

eDNA copy number data was used as a response to assess the effects that sampling time and target sequence length had on PMA and non-PMA treated samples. Time and target sequence length had significant effects on eDNA copy number (MANOVA, all $p < 0.05$ **Table 4**). *Post hoc* test confirmed that eDNA copy number was significantly affected by target sequence length only in PMA treated samples (intracellular eDNA; ANOVA, $p < 0.05$ **Table 4**). While eDNA copy number gradually decreased in both PMA and non-PMA treatments as target sequence length increased (**Figure 4**), intracellular eDNA copy numbers (PMA treated samples) were significantly lower in 225, 333, 430, 621,

TABLE 2 | Percentage decrease of average eDNA copy number (shown as Δ%) for each assay by PMA treatment in non-boiled and boiled treatments.

	Non-boiled			Boiled		
	Avg. eDNA copy number			Avg. eDNA copy number		
Amplicon length	**Non-PMA**	**PMA**	**Δ%**	**Non-PMA**	**PMA**	**Δ%**
132 bp	15637.78	859.678	94.50	5994.187	103.558	98.27
225 bp	15520.94	651.186	95.80	5950.311	35.873	99.4
333 bp	16203.333	390.317	97.59	4899.785	0	–
430 bp	16144.768	317.031	98.04	4873.065	0	–
529 bp	19064.72	336.607	98.23	5205.062	0	–
621 bp	17562.675	346.196	98.03	4818.851	0	–
715 bp	17110.21	318.856	98.14	4551.071	0	–
823 bp	19574.029	322.742	98.35	4779.743	0	–
935 bp	18788.451	287.817	98.47	4270.684	0	–
1021 bp	19991.794	237.982	98.81	4568.729	0	–

The average eDNA copy number was calculated by averaging the copy number generated during qPCR amplification from each of three technical replicates from three biological replicates. For 333 bp and above, eDNA was below the LOQ in boiled-PMA control and they were excluded from the analyses. Percentage decrease (Δ%) was calculated as the reduction from eDNA copies detected in non-PMA treatment to the one detected in PMA treatment.

and 823 bp assay when using eDNA copy number from 132 bp as a baseline (GLM, **Table 5**). As for the total eDNA copy number (non-PMA samples), no significant difference was found among the target sequence lengths (MANOVA, $p > 0.05$ **Table 4**), so GLM was performed only on the result for intracellular eDNA. The decay rates and the half-life calculations are shown in **Supplementary Table 2** and **Figure 5**, respectively. The half-life of the total eDNA was almost unchanged with respect to the target sequence length (To be exact, it decreased with the target sequence length, but the *p*-value was not significant; MANOVA, $p > 0.05$ **Table 4**), whereas the half-life of intracellular eDNA tended to decrease with increasing target sequence length (**Figure 5**). The calculated half-life was longer for intracellular eDNA at all (testable) target sequence lengths, and nearly double at 132 bp (**Figure 5**). In addition, as target sequence length increased, the difference in half-life tended to decrease. All of FNC and NTC were negative for the amplification of *D. rerio* DNA. All of the raw data obtained by qPCR are shown in **Supplementary Table 4**.

TABLE 3 | ANOVA results for the difference in eDNA copy number as a response to boiled and non-boiled treatment.

ANOVAs		
Response	**Factor**	***P*-value**
eDNA conc.	Boiled	0.0000
	PMA	0.0000
	Assay length	0.0003
	Boiled: PMA	0.0000
	Boiled: assay length	0.0000
	PMA: assay length	0.0000
	Boiled: PMA: assay length	0.0000

Let $p < 0.05$ be the threshold of significance. All eDNA copy numbers were log-transformed.

DISCUSSION

Experiment 1: PMA Confirmation Test

For the first time, we have shown that PMA is effective for selectively amplifying DNA from intact vertebrate cells. However, while PMA could not completely exclude extracellular and disrupted cell-derived DNA from PCR amplification, we confirmed that PMA could work on damaged fish cells to cause a significant reduction in the detected amount of DNA (**Tables 2**, **3** and **Figure 3**). As we expected, there was a drastic difference in the mean eDNA copy number between the boiled and non-boiled, PMA treated samples. Boiling disrupted cell membranes thereby allowing PMA to bind to DNA and either prevent amplification all together or reduce amplification below the LOQ (**Figure 3**) for all the assays, except the two shortest target sequence lengths (132 and 225 bp). The significant decrease in eDNA copy number of the boiled and non-boiled samples after PMA treatment, and the (expected) small decrease of eDNA copy number for non-PMA treated samples (**Figure 3**) suggests that boiling caused DNA degradation. It is probable that longer fragments in our samples that were boiled (100°C) for 10 min would have been degraded, which would cause an increase in smaller fragments (i.e., < 333 bp). Studies looking at the effects of thermal degradation on DNA for plants and bacteria indicate that prolonged exposure to high temperatures (50–200°C) will cause longer fragments to degrade, but smaller amplifiable fragments can persist (Zhang and Wu, 2005; Hrnčírová et al., 2008; Lo et al., 2015; Bitskinashvili et al., 2018). Given the short thermal exposure time in our experiment, any DNA degradation caused by high temperature is likely small and probably associated with other degradation processes (e.g., lipid peroxidation; sensu Zhang and Wu, 2005); however, this mechanism is outside the scope of our study, and our hypotheses are limited. Furthermore, there is evidence to suggest that PMA cannot completely suppress PCR amplification in cases where the target fragment is too short (e.g., 190 bp) (Luo et al., 2010). Given the increase in smaller fragments

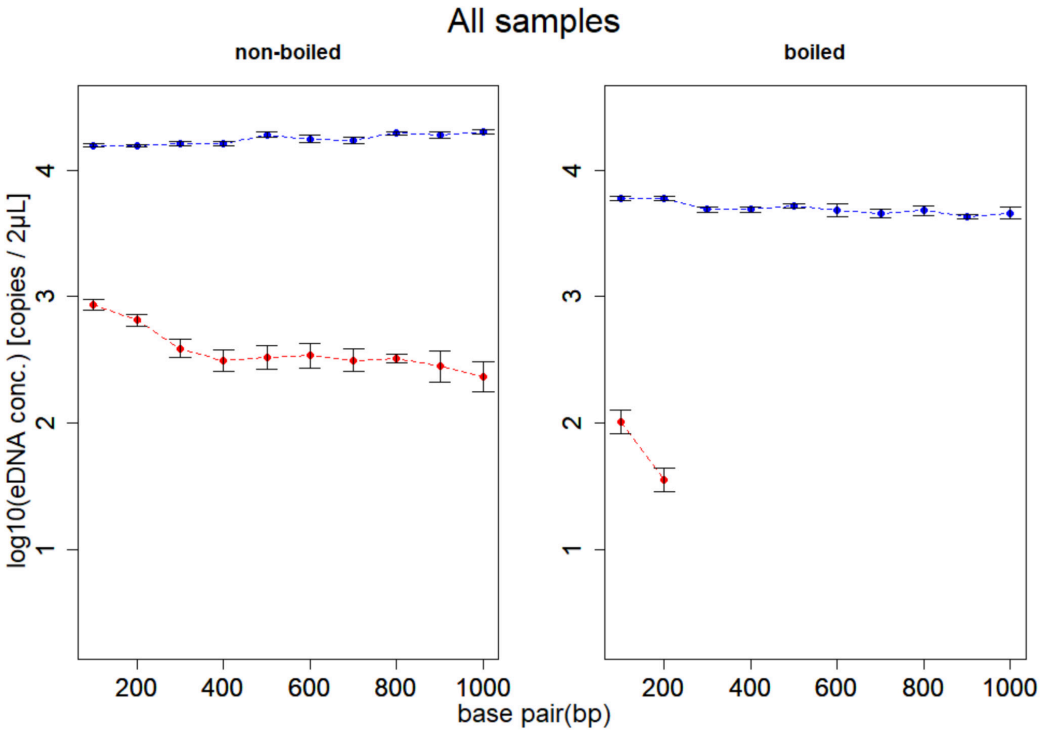


FIGURE 3 | Results of experiment 1. Mean eDNA copy number of samples for boiling and PMA treatments at each of 10 target sequence lengths (132, 225, 333, 440, 529, 621, 715, 832, 935, and 1,021 bp). The left panel shows the non-boiled samples, and the right panel shows the boiled samples. Dashed lines show change in total eDNA copy number (blue; non-PMA) and intracellular eDNA copy number (red; PMA). The total eDNA copy number (blue line) did not change even if the target sequence length increased, and the intracellular eDNA (red line) tended to decrease as the target sequence length increased. In addition, no eDNA above 333 bp was detected in the boiled PMA sample. The effects of the boiling treatment are found on both total and intracellular eDNA, but intracellular eDNA is more significantly affected.

TABLE 4 | Results of MANOVA (upper) and *post hoc* test (lower) for the relationships between eDNA copy numbers at each treatment (PMA/intracellular eDNA vs. non-PMA/total eDNA) and each factor (days and assay length).

MANOVA				
Response	**Factor**		***F*-value**	***P*-value**
eDNA conc.(Including total eDNAand intracellular eDNA)	Days		715.33	0.0000
	Assay length		63.21	0.0000
	Days: assay length		9.20	0.0002
Response	**Factor**	**Sum Sq**	***F*-value**	***P*-value**
eDNA conc. (Total eDNA)	Days	68.881	1420.364	0.0000
	Assay length	0.160	3.3052	0.0708
	Days: assay length	0.025	0.5151	0.4739
	Residuals	8.535		
eDNA conc. (Intracellular eDNA)	Days	86.146	363.504	0.0000
	Assay length	21.218	89.534	0.0000
	Days: assay length	4.127	17.416	0.0000
	Residuals	41.710		

Let $p < 0.05$ be the threshold of significance. All the eDNA copy numbers were log-transformed.

as extracellular DNA caused by boiling and the inability of PMA to complete inhibit PCR for short fragments, this may explain why amplification was still observed with the 132 and 225 bp fragments after PMA treatment was applied.

Experiment 2: Elucidation of Degradation Process of Intracellular eDNA

Also, for the first time, we were able to selectively amplify and quantify intracellular eDNA, and model the associated

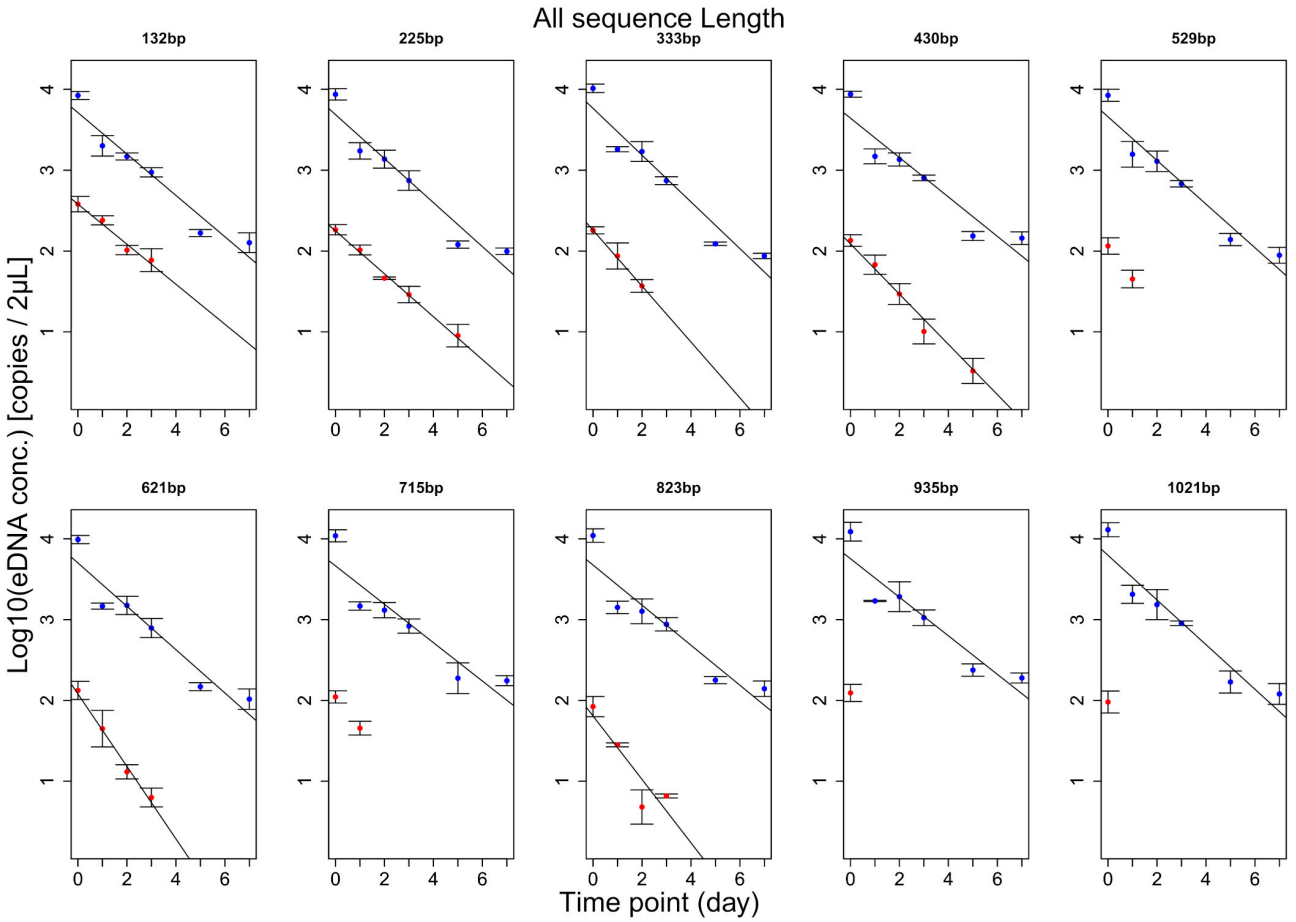


FIGURE 4 | Results of experiment 2. The change over time for each amplicon length is shown. The blue line shows the change in total eDNA copy number, and the red line shows the change in intracellular eDNA copy number. For 529, 715, 935, and 1,021 bp, only two or less time points were detected, so statistical processing was not performed. The total eDNA has the same result regardless of the target sequence length, but the intracellular eDNA tends to have a slight change in the regression line as the target sequence length increases.

degradation rate. The MANOVA results suggested that the total eDNA present in the sample maintained comparable copy numbers between different target sequence lengths, even in the longest sequence (1,021 bp), and this pattern was maintained during the 7-day incubation period. Conversely, the intracellular eDNA copy number showed significant decreases as assay length increased over the 7-day incubation period. An interesting pattern that was observed was the significant interactions between the incubation time and three of the target sequence lengths, 430, 621, 823 bp for PMA treated samples (MANOVA and GLM; **Tables 4**, **5**). There is a distinct difference in the slopes for these three assays that indicates a drastic decrease in mean eDNA copy number with increased time. However, the cause of the drastic slope of these three assays may affect the decay rates of smaller fragments (132, 225, and 333 bp). At any given timepoint, a leptokurtic pattern is expected regarding the number of fragments, such that there will be a significantly larger number of small fragments compared to a smaller number of larger fragments (*sensu* Deagle et al., 2006). Specifically, when larger fragments degrade, they are broken into smaller fragments and these smaller fragments add to the overall quantity of smaller fragments, which lend themselves to "slowing" the decay rate for smaller fragments.

Our results for total eDNA follow similar patterns seen in other studies that assessed target sequence length and time as a factor of eDNA degradation (Sassoubre et al., 2016; Lance et al., 2017; Bylemans et al., 2018; Jo et al., 2019). GLM was not performed on total eDNA, as no effect of target sequence length was observed (**Table 4**), indicating that the degradation process of total eDNA might be similar regardless of the target sequence length, at least in the aquarium setup of the present study. Alternatively, our results demonstrate that intracellular eDNA may have a different degradation process from that of total eDNA (**Table 4** and **Figures 4**, **5**). The half-life calculated for each target sequence length was greater for intracellular eDNA, but was negatively correlated with increasing target sequence length (**Figure 5** and **Supplementary Table 2**). While lack of amplification (or amplification meeting the LOQ) prevented statistical analyses of half-life indices for larger fragment sizes in intracellular eDNA, it might be assumed that total and

TABLE 5 | GLM analysis results for testing response of intracellular (PMA) eDNA copy numbers to days and assay length.

Response	Explanatory	Estimate	SE	*P*-value
eDNA conc.	Intercept	2.5986	0.0703	0.0000
(Intracellular eDNA)	Days	−0.2591	0.0198	0.0000
	Assay length (225 bp)	−0.2917	0.0986	0.0041
	Assay length (333 bp)	−0.4232	0.1038	0.0001
	Assay length (430 bp)	−0.4871	0.1038	0.0000
	Assay length (621 bp)	−0.6180	0.1070	0.0000
	Assay length (823 bp)	−0.7901	0.1135	0.0000
	Days: assay length (225 bp)	−0.0426	0.0268	0.1153
	Days: assay length (333 bp)	−0.0454	0.0338	0.1821
	Days: assay length (430 bp)	−0.0723	0.0338	0.0353
	Days: assay length (621 bp)	−0.1025	0.0407	0.0138
	Days: assay length (823 bp)	−0.1331	0.0551	0.0182

Let $p < 0.05$ be the threshold of significance. All the eDNA copy numbers were log-transformed. Responses of each assay length and interactions was calculated using 132 bp as a baseline.

intracellular eDNA decay rates are similar at larger fragments sizes based on regression (**Figure 5**). This is, in part, supported by our inability to detect larger fragments after the initial sampling (Day 0) which indicates a short half-life (**Figure 4**). Other studies have reported that longer amplicons amplified by PCR have lower quantity and faster degradation when analyzing total eDNA (Deagle et al., 2006; Jo et al., 2017), although the fraction of intracellular eDNA as part of the total eDNA was much smaller in the case of our aquarium water compared to the larger systems used in Deagle et al. (2006) and Jo et al. (2017). Furthermore, the interpretation of eDNA degradation results could be better supported if accurate estimates of intact vertebrate cells in total eDNA could be calculated.

Synthesis

In this study, PMA was confirmed to be effective to differentiate "intact cells" from "disrupted cells" in vertebrate eDNA analysis, enabling the independent detection of intracellular eDNA among the total eDNA. It has been suggested that intracellular DNA represents a specific fraction of total eDNA (Turner et al., 2014; Jo et al., 2019) but this study is the first to present a method to quantify the fraction more clearly using PMA. Our results suggested that the decay rate is dependent on both the existing state of eDNA, i.e., intracellular vs. extracellular, and the length of the target sequence. The process of eDNA degradation may differ between experimental conditions and in the natural environment, and thus it is important to confirm the degradation pattern of intracellular eDNA in environmental water samples under various environmental conditions (e.g., pH, trophic state, temperature and biomass; Jo et al., 2019). For example, intracellular DNA is suggested to degrade much less efficiently than extracellular DNA, by extracellular enzymes,

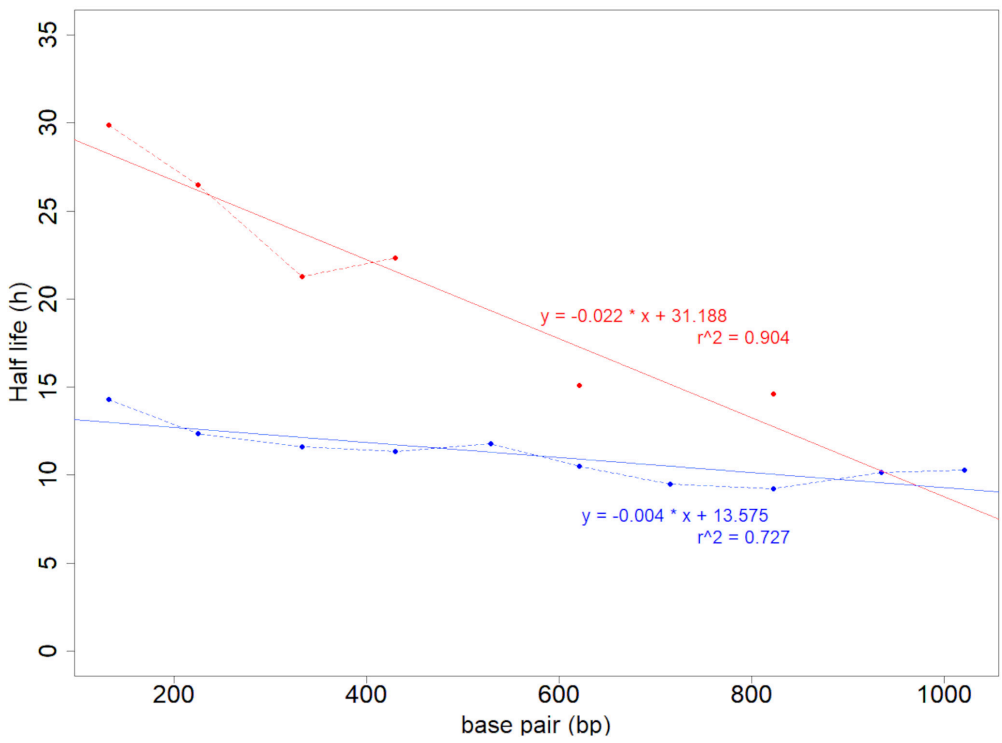


FIGURE 5 | Half-life results calculated from the decay rate. The vertical axis shows the half-life (hours), and the horizontal axis shows the sequence lengths. The blue line shows the change in total eDNA copy number, and the red line shows the change in intracellular eDNA copy number. At 529, 715, 935, and 1,021 bp of intracellular eDNA (red line), the half-life could not be statistically calculated because the eDNA copy number was obtained only at 2 time points or less.

due to the presence of cell membranes (Paul et al., 1987; Levy-Booth et al., 2007; Corinaldesi et al., 2008). Our results from experiment 2, where the half-life of intracellular eDNA was longer than the half-life of total eDNA, suggest that intracellular eDNA is protected from external degradation factors by the cell membrane. Many cells begin to degrade by apoptosis before shedding from the individual (Hotchkiss et al., 2009). However, during normal apoptotic shedding of epithelial cells, intact mitochondria may be released from the cell and mtDNA may be protected from endonuclease degradation (Murgia et al., 1992; Tepper and Studzinski, 1993). In aquatic animals, this process releases the entire mitochondria into the water column, where the bilayer resists degradation and mitochondrial nucleoids further protect mtDNA (Rickwood and Chambers, 1981). The mitochondrial bilayer and tissue cell membrane could provide a greater level of protection against DNA degradation that is not seen with just one cell membrane alone (Foran, 2006).

Conversely, while protected from external decomposition factors, internal degradation factors may act on eDNA before it is released from the target species. In Experiment 1, the total eDNA copy number of the non-boiled samples did not change as the target sequence length increased, but the intracellular eDNA copy number changed. Although intracellular eDNA is protected from external degradation factors in the environment and has a longer half-life, there is a possibility that DNA fragmentation is progressing due to internal degradation factors associated with apoptotic processes (Nagata, 2005; Hotchkiss et al., 2009). This indicates that extracellular eDNA (free eDNA), damaged intracellular eDNA and intracellular eDNA may be subject to separate degradation processes. In the future, in order to further clarify mechanisms driving eDNA degradation rates and processes, it is necessary to study the dynamics of each existing state, as well as, total eDNA. This will ultimately provide a basis for developing more robust analyses that help limit or prevent process Type I and II errors.

PMA has been effective as a molecular surveillance tool for human health issues, but has a great potential to assist with future conservation efforts of aquatic systems. Removing the "noise," such as eDNA resuspended from sediments or transported from upstream sites, can increase the accuracy eDNA surveillance by allowing researchers to detect contemporary signals from of targeted species. As discussed earlier, process Type I errors in current eDNA analyses may results due to the (unintentional) misinterpretation of the eDNA sources (e.g., the transport of eDNA by predator droppings, boats, or wetland birds) (Deiner and Altermatt, 2014; Merkes et al., 2014). In cases where rapid and accurate biosurveillance information is needed, like determining if an invasive species has entered a new body of water, combined PMA-treatment with multi-assay surveillance (especially assays with different amplicon lengths) may help avoid misinterpretation and determine whether the signal obtained by eDNA analysis is new or old. To this end, resource management agencies deploy various eradication measures to remove invasive species from invaded systems (Bonar et al., 2009; Booy et al., 2017). PMA-eDNA surveillance would allow researchers to discern between intact and degraded cells, to identify the effectiveness of eradication strategies, i.e., detection of intracellular eDNA after a set time period might indicate that an eradication strategy has failed. The timing of seasonal migratory patterns of fish (long-term stabilization by sediment adsorption) could be more easily and reliably assessed (Maruyama et al., 2014). Critical habitat and spawning habitat (and season) are unknown for many aquatic organisms, and thus PMA-eDNA can help researchers hone in on specific areas of an aquatic environment without the need visual surveys that require many hours and resources. To realize these applied ideas in future research, we hope that the combination of PMA with multiple-sized amplicons for target species will facilitate new research for clarifying other mechanistic underpinnings of eDNA ecology.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

Ethical review and approval was not required for the animal study in accordance with the local legislation and institutional requirements, but we followed the instructions suggested by Japanese animal welfare regulations and the "Act on Welfare and Management of Animals" (Ministry of Environment of Japan) for husbandry and handling of fish.

AUTHOR CONTRIBUTIONS

TH, RP, and HY: concept and design and data analysis. TH, KT, and KM: lab work. All authors: writing final version.

FUNDING

This work received financial support from the Joint Research Center for Science and Technology Fund of Ryukoku University (2019) and (2018) Ryukoku University Science and Technology Fund. The study was partly supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (26840152 and 20H03326).

ACKNOWLEDGMENTS

We thank Dr. Carl O. Ostberg (Western Fisheries Research Center, USGS, United States) for discussions on the interpretation of the results.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.632973/full#supplementary-material

REFERENCES

Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., et al. (2008). Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. *BioScience* 58, 403–414. doi: 10.1641/B580507

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

Bitskinashvili, K., Gabriadze, I., Kutateladze, T., Vishnepolsky, B., Mikeladze, D., and Datukishvili, N. (2018). Effects of thermal-acid treatment on degradation and amplification of wheat and maize DNA. *J. Food Nutr. Res.* 57, 242–251.

Bonar, S. A., Hubert, W. A., and Willis, D. W. (2009). *Standard Methods for Sampling North American Freshwater Fishes*. Bethesda, MA: American Fisheries Society.

Booy, O., Mill, A. C., Roy, H. E., Hiley, A., Moore, N., Robertson, P., et al. (2017). Risk management to prioritise the eradication of new and emerging invasive non-native species. *Biol. Invasions* 19, 2401–2417. doi: 10.1007/s10530-017-1451-z

Bylemans, J., Furlan, E. M., Gleeson, D. M., Hardy, C. M., and Duncan, R. P. (2018). Does size matter? An experimental evaluation of the relative abundance and decay rates of aquatic environmental DNA. *Environ. Sci. Technol.* 52, 6408–6416. doi: 10.1021/acs.est.8b01071

Champlot, S., Berthelot, C., Pruvost, M., Bennett, E. A., Grange, T., and Geigl, E.-M. (2010). An efficient multistrategy DNA decontamination procedure of PCR reagents for hypersensitive PCR applications. *PLoS One* 5:e13042. doi: 10.1371/journal.pone.0013042

Collen, B., Whitton, F., Dyer, E. E., Baillie, J. E., Cumberlidge, N., Darwall, W. R., et al. (2014). Global patterns of freshwater species diversity, threat and endemism. *Glob. Ecol. Biogeo.* 23, 40–51. doi: 10.1111/geb.12096

Corinaldesi, C., Beolchini, F., and Dell'Anno, A. (2008). Damage and degradation rates of extracellular DNA in marine sediments: implications for the preservation of gene sequences. *Mol. Ecol.* 17, 3939–3951. doi: 10.1111/j.1365-294X.2008.03880.x

Darling, J. A., and Mahon, A. R. (2011). From molecules to management: adopting DNA-based methods for monitoring biological invasions in aquatic environments. *Environ. Res.* 111, 978–988. doi: 10.1016/j.envres.2011.02.001

Deagle, B. E., Eveson, J. P., and Jarman, S. N. (2006). Quantification of damage in DNA recovered from highly degraded samples – a case study on DNA in faeces. *Front Zool* 3:11. doi: 10.1186/1742-9994-3-11

Deiner, K., and Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. *PLoS One* 9:e88786. doi: 10.1371/journal.pone.0088786

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., et al. (2011). Persistence of environmental DNA in freshwater ecosystems. *PLoS One* 6:e23398. doi: 10.1371/journal.pone.0023398

Dejean, T., Valentini, A., Miquel, C., Taberlet, P., Bellemain, E., and Miaud, C. (2012). Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog *Lithobates catesbeianus*: alien invasive species detection using eDNA. *J. Appl. Ecol.* 49, 953–959. doi: 10.1111/j.1365-2664.2012.02171.x

Ellison, S. L., English, C. A., Burns, M. J., and Keer, J. T. (2006). Routes to improving the reliability of low-level DNA analysis using real-time PCR. *BMC Biotechnol.* 6:33. doi: 10.1186/1472-6750-6-33

Emerson, J. B., Adams, R. I., Román, C. M. B., Brooks, B., Coil, D. A., Dahlhausen, K., et al. (2017). Schrödinger's microbes: tools for distinguishing the living from the dead in microbial ecosystems. *Microbiome* 5:86. doi: 10.1186/s40168-017-0285-3

Evans, N. T., Shirey, P. D., Wieringa, J. G., Mahon, A. R., and Lamberti, G. A. (2017). Comparative cost and effort of fish distribution detection via environmental DNA analysis and electrofishing. *Fisheries* 42, 90–99. doi: 10.1080/03632415.2017.1276329

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., et al. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Mol. Ecol. Resour.* 15, 543–556. doi: 10.1111/1755-0998.12338

Foran, D. R. (2006). Relative degradation of nuclear and mitochondrial DNA: an experimental approach. *J. Forensic Sci* 51, 766–770. doi: 10.1111/j.1556-4029.2006.00176.x

García-Fontana, C., Narváez-Reinaldo, J. J., Castillo, F., González-López, J., Luque, I., and Manzanera, M. (2016). A new physiological role for the DNA molecule as a protector against drying stress in desiccation-tolerant microorganisms. *Front. Microbiol.* 7:2066. doi: 10.3389/fmicb.2016.02066

Goldberg, C. S., Sepulveda, A., Ray, A., Baumgardt, J., and Waits, L. P. (2013). Environmental DNA as a new method for early detection of New Zealand mudsnails (*Potamopyrgus antipodarum*). *Freshw. Sci.* 32, 792–800. doi: 10.1899/13-046.1

Hajibabaei, M., Baird, D. J., Fahner, N. A., Beiko, R., and Golding, G. B. (2016). A new way to contemplate Darwin's tangled bank: how DNA barcodes are reconnecting biodiversity science and biomonitoring. *Philos. Trans. R. Soc. Lond. B* 371:20150330. doi: 10.1098/rstb.2015.0330

Hayes, D. B., Ferreri, C. P., and Taylor, W. W. (1996). "Active fish capture methods," in *Fisheries Techniques*, 2nd Edn, eds B. R. Murphy and D. W. Willis (Bethesda, MD: American Fisheries Society), 193–220.

Hotchkiss, R. S., Strasser, A., McDunn, J. E., and Swanson, P. E. (2009). Cell death. *N. Engl. J. Med.* 361, 1570–1583. doi: 10.1056/NEJMra0901217

Hrnčírová, Z., Bergerová, E., and Siekel, P. (2008). Effects of technological treatment on DNA degradation in selected food matrices of plant origin. *J. Food Nutr. Res* 47, 23–28.

Jerde, C. L., Mahon, A. R., Chadderton, W. L., and Lodge, D. M. (2011). "Sight-unseen" detection of rare aquatic species using environmental DNA: EDNA surveillance of rare aquatic species. *Conserv. Lett.* 4, 150–157. doi: 10.1111/j.1755-263X.2010.00158.x

Jo, T., Arimoto, M., Murakami, H., Masuda, R., and Minamoto, T. (2019). Particle size distribution of environmental DNA from the nuclei of marine fish. *Environ. Sci. Technol.* 53, 9947–9956. doi: 10.1021/acs.est.9b02833

Jo, T., Murakami, H., Masuda, R., Sakata, M. K., Yamamoto, S., and Minamoto, T. (2017). Rapid degradation of longer DNA fragments enables the improved estimation of distribution and biomass using environmental DNA. *Mol. Ecol. Resour,* 17, e25–e33. doi: 10.1111/1755-0998.12685

Joo, S., Park, P., and Park, S. (2019). Applicability of propidium monoazide (PMA) for discrimination between living and dead phytoplankton cells. *PLoS One* 14:e0218924. doi: 10.1371/journal.pone.0218924

Kamoroff, C., and Goldberg, C. S. (2018). An issue of life or death: using eDNA to detect viable individuals in wilderness restoration. *Freshw. Sci.* 37, 685–696. doi: 10.1086/699203

Lance, R., Klymus, K., Richter, C., Guan, X., Farrington, H., Carr, M., et al. (2017). Experimental observations on the decay of environmental DNA from bighead and silver carps. *Manag. Biol. Invasions* 8, 343–359. doi: 10.3391/mbi.2017.8.3.08

Lance, R. F., and Carr, M. R. (2012). *Detecting eDNA of Invasive Dreissenid Mussels: Report on Capital Investment Project.* ERDC/TN ANSRP-12-2. Vicksburg, MS: US Army Engineer Research & Development Center, doi: 10.13140/2.1.1265.0887

Levy-Booth, D. J., Campbell, R. G., Gulden, R. H., Hart, M. M., Powell, J. R., Klironomos, J. N., et al. (2007). Cycling of extracellular DNA in the soil environment. *Soil Biol. Biochem.* 39, 2977–2991. doi: 10.1016/j.soilbio.2007.06.020

Lo, Y.-T., Li, M., and Shaw, P.-C. (2015). Identification of constituent herbs in ginseng decoctions by DNA markers. *Chin. Med.* 10:1. doi: 10.1186/s13020-015-0029-x

Lodge, D. M., Turner, C. R., Jerde, C. L., Barnes, M. A., Chadderton, L., Egan, S. P., et al. (2012). Conservation in a cup of water: Estimating biodiversity and population abundance from environmental DNA. *Mol. Ecol.* 21, 2555–2558. doi: 10.1111/j.1365-294X.2012.05600.x

Luo, J. F., Lin, W. T., and Guo, Y. (2010). Method to detect only viable cells in microbial ecology. *Appl. Microbiol. Biotechnol.* 86, 377–384. doi: 10.1007/s00253-009-2373-1

Mächler, E., Osathanunkul, M., and Altermatt, F. (2018). Shedding light on eDNA: Neither natural levels of UV radiation nor the presence of a filter feeder affects eDNA-based detection of aquatic organisms. *PLoS One* 13:e0195529. doi: 10.1371/journal.pone.0195529

Martin, B., Raurich, S., Garriga, M., and Aymerich, T. (2013). Effect of amplicon length in propidium monoazide quantitative PCR for the enumeration of viable cells of *Salmonella* in cooked ham. *Food Anal. Methods* 6, 683–690. doi: 10.1007/s12161-012-9460-0

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

Merkes, C. M., McCalla, S. G., Jensen, N. R., Gaikowski, M. P., and Amberg, J. J. (2014). Persistence of DNA in carcasses, slime and avian feces may affect interpretation of environmental DNA data. *PLoS One* 9:e113346. doi: 10.1371/journal.pone.0113346

Miya, M., Minamoto, T., Yamanaka, H., Oka, S., Sato, K., Yamamoto, S., et al. (2016). Use of a filter cartridge for filtration of water samples and extraction of environmental DNA. *J. Vis. Exp.* 117:54741. doi: 10.3791/54741

Murakami, H., Yoon, S., Kasai, A., Minamoto, T., Yamamoto, S., Sakata, M. K., et al. (2019). Dispersion and degradation of environmental DNA from caged fish in a marine environment. *Fish. Sci.* 85, 327–337. doi: 10.1007/s12562-018-1282-6

Murgia, M., Pizzo, P., Sandoná, D., Zanovello, P., Rizzut, R., and Virgilio, F. D. (1992). Mitochondrial DNA is not fragmented during apoptosis. *J. Biol. Chem.* 276, 10939–10941.

Nagata, S. (2005). DNA degradation in development and programmed cell death. *Annu. Rev. Immunol.* 23, 853–875. doi: 10.1146/annurev.immunol.23.021704.115811

Nocker, A., Sossa-Fernandez, P., Burr, M. D., and Camper, A. K. (2007). Use of propidium monoazide for live/dead distinction in microbial ecology. *Appl. Environ. Microbiol.* 73, 5111–5117. doi: 10.1128/AEM.02987-06

Paul, J. H., Jeffrey, W. H., and DeFlaun, M. F. (1987). Dynamics of extracellular DNA in the marine environment. *Appl. Environ. Microbiol.* 53, 170–179. doi: 10.1128/AEM.53.1.170-179.1987

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., and Waits, L. P. (2013). Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. *Can. J. Fish. Aquat. Sci.* 70, 1123–1130. doi: 10.1139/cjfas-2013-0047

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., and Waits, L. P. (2014). Factors influencing detection of eDNA from a stream-dwelling amphibian. *Mol. Ecol. Resour.* 14, 109–116. doi: 10.1111/1755-0998.12159

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Core Team.

Rees, H. C., Bishop, K., Middleditch, D. J., Patmore, J. R. M., Maddison, B. C., and Gough, K. C. (2014). The application of eDNA for monitoring of the great crested newt in the UK. *Ecol. Evol.* 4, 4023–4032. doi: 10.1002/ece3.1272

Ribeiro, H., Martins, A., Gonçalves, M., Guedes, M., Tomasino, M. P., Dias, N., et al. (2019). Development of an autonomous biosampler to capture in situ aquatic microbiomes. *PLoS One* 14:e0216882. doi: 10.1371/journal.pone.0216882

Rickwood, D., and Chambers, J. A. A. (1981). Evidence for protected regions of DNA in the mitochondrial nucleoid of *Saccharomyces cerevisiae*. *FEMS Microbiol. Lett.* 12, 187–190. doi: 10.1111/j.1574-6968.1981.tb07639.x

Sassoubre, L. M., Yamahara, K. M., Gardner, L. D., Block, B. A., and Boehm, A. B. (2016). Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. *Environ. Sci. Technol.* 50, 10456–10464. doi: 10.1021/acs.est.6b03114

Smart, A. S., Tingley, R., Weeks, A. R., van Rooyen, A. R., and McCarthy, M. A. (2015). Environmental DNA sampling is more sensitive than a traditional survey technique for detecting an aquatic invader. *Ecol. Appl.* 25, 1944–1952. doi: 10.1890/14-1751.1

Smart, A. S., Weeks, A. R., van Rooyen, A. R., Moore, A., McCarthy, M. A., and Tingley, R. (2016). Assessing the cost-efficiency of environmental DNA sampling. *Methods Ecol. Evol.* 7, 1291–1298. doi: 10.1111/2041-210X.12598

Strayer, D. L., and Dudgeon, D. (2010). Freshwater biodiversity conservation: recent progress and future challenges. *J. N. Amer. Benthol. Soc.* 29, 344–358. doi: 10.1899/08-171.1

Taberlet, P., Bonin, A., Zinger, L., and Coissac, E. (2018). *Environmental DNA – For Biodiversity Research and Monitoring*. Oxford: Oxford University Press, doi: 10.1093/OSO/9780198767220.001.0001

Takahara, T., Minamoto, T., and Doi, H. (2013). Using environmental DNA to estimate the distribution of an invasive fish species in ponds. *PLoS One* 8:e56584. doi: 10.1371/journal.pone.0056584

Tepper, C. G., and Studzinski, G. P. (1993). Resistance of mitochondrial DNA to degradation characterizes the apoptotic but not the necrotic mode of human leukemia cell death. *J. Cell. Biochem.* 52, 352–361. doi: 10.1002/jcb.240520311

Thompson, W. (Ed.) (2013). *Sampling Rare or Elusive Species: Concepts, Designs, and Techniques for Estimating Population Parameters*. Washington, D.C: Island Press.

Thomsen, P. F., Kielgast, J., Iversen, L. L., Møller, P. R., Rasmussen, M., and Willerslev, E. (2012a). Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS One* 7:e41732. doi: 10.1371/journal.pone.0041732

Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., et al. (2012b). Monitoring endangered freshwater biodiversity using environmental DNA. *Mol. Ecol.* 21, 2565–2573. doi: 10.1111/j.1365-294X.2011.05418.x

Tsuji, S., Ushio, M., Sakurai, S., Minamoto, T., and Yamanaka, H. (2017). Water temperature-dependent degradation of environmental DNA and its relation to bacterial abundance. *PLoS One* 12:e0176608. doi: 10.1371/journal.pone.0176608

Turner, C. R., Barnes, M. A., Xu, C. C. Y., Jones, S. E., Jerde, C. L., and Lodge, D. M. (2014). Particle size distribution and optimal capture of aqueous macrobial eDNA. *Methods Ecol. Evol.* 5, 676–684. doi: 10.1111/2041-210X.12206

Turner, C. R., Uy, K. L., and Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biol. Conserv.* 183, 93–102. doi: 10.1016/j.biocon.2014.11.017

Warne, R. T. (2014). A primer on multivariate analysis of variance (MANOVA) for behavioral scientists. *Pract. Assess. Res. Eval.* 19:10. doi: 10.7275/sm63-7h70

Yamamoto, S., Minami, K., Fukaya, K., Takahashi, K., Sawada, H., Murakami, H., et al. (2016). Environmental DNA as a 'Snapshot' of fish distribution: a case study of Japanese Jack Mackerel in Maizuru Bay, sea of Japan. *PLoS One* 11:e0149786. doi: 10.1371/journal.pone.0149786

Zhang, L., and Wu, Q. (2005). Single gene retrieval from thermally degraded DNA. *J. Biosci.* 30, 599–604. doi: 10.1007/BF02703559

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership