THE DEVELOPMENT AND APPLICATION OF eDNA METABARCODING TO IDENTIFY CRITICAL HABITAT CHARACTERISTICS FOR THREATENED AND ENDANGERED HERPETOFAUNA OF THE GREAT LAKES REGION

By

Olivia M. Ruppert

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Fisheries and Wildlife – Master of Science

2023

ABSTRACT

Information pertaining to relationships between habitat characteristics and species distributions is critical to guide conservation strategies for imperiled species. Little is currently known about the habitat features required by many threatened or endangered reptiles and amphibians of the Great Lakes region. Environmental DNA (eDNA) metabarcoding can be used to detect species otherwise not easily detected by other sampling methods. Through analysis of species-specific DNA sequences collected from filtered water, species presence and relative sequence abundance can be estimated. Our objectives for this research were to develop and optimize eDNA sampling protocols for wetland habitats, implement a regional eDNA survey to identify hotspots of wetland biodiversity and update species distribution data, and use species distribution and occupancy models to characterize habitat and climate associations for numerous species of conservation concern. In Years 1 and 2 of the project, we detected 29 and 27 herpetofauna species, respectively. We observed no difference in the number of species detected between water samples collected at a single point and across a 5m transect (p = 0.52). No difference in species richness was observed between community sequence data obtained from 12S and 16S mitochondrial DNA markers (p = 0.96). We found a significant difference in the number of species detected between late and early sampling periods (p = 7e-7). Occupancy modeling results suggest that the number of wetland types sampled (*Thamnophis unclassified*: Bayesian estimate coefficient = 0.04) and the percentage of developed land cover (*Ambystoma texanum*: coefficient = 0.20; *Emydoidea blandingii*: coefficient = -0.21) had no significant effect on patterns of site occupancy.

ACKNOWLEDGEMENTS

This project was funded by the Great Lakes Fish and Wildlife Restoration Act (USFWS). Thanks to our sponsor, Erin Johnson. A special thank you to committee members Dr. John Robinson, Dr. Kim Scribner, Dr. Jen Moore, and Dr. Jen Owen. Thank you to other major project contributors including Dr. Jared Homola, Dr. Jeannette Kanefsky and Dr. Yu Man Lee. Additionally, this project was supported through use of resources at the Institute for Cyber-Enabled Research and the Genomics Core of the Research Technology Support Facility at Michigan State University. Thank you to the countless managers, biologists, and researchers across the Midwest who contributed to this project, as well as the entire Scribner/Robinson Lab.

TABLE OF CONTENTS

CHAPTER 1: AN OVERVIEW OF ENVIRONMENTAL DNA METABARCODING: PAST, PRESENT, AND FUTURE APPLICATIONS IN WETLANDS 1
ABSTRACT 1
INTRODUCTION 2
OBJECTIVES 5
LITERATURE CITED 9

CHAPTER 2: DEVELOPMENT AND OPTIMIZATION OF WETLAND ENVIRONMENTAL DNA METABARCODING PROTOCOLS TO SURVEY HERPETOFAUNA IN THE GREAT LAKES REGION 13
ABSTRACT 13
INTRODUCTION 14
METHODS 16
Field Collections 16
Filter Processing and DNA Extraction 19
PCR and High-Throughput Sequencing 19
Mock Communities 21
Taxonomic Database Development 23
Data Analysis 28
RESULTS 29
Overview 29
Comparisons of Sample Collection Methodology 32
Comparisons Between 12S and 16S Markers 34
Comparisons Between Seasonal Samples and Wetland Types 36
DISCUSSION 39
Point and Transect Sampling within Wetland Systems 39
Metabarcoding Marker Comparisons 41
Number of Species Detected Between Sampling Periods and Between Wetland Types 42
CONCLUSIONS AND REFLECTIONS 43
LITERATURE CITED 45

CHAPTER 3: PATTERNS OF BIODIVERSITY ACROSS GREAT LAKES REGION WETLANDS AND HABITAT ASSOCIATIONS FOR HERPETOFAUNA OCCUPANCY 50
ABSTRACT 50
INTRODUCTION 51
METHODS 52
Site Selection and Field Collections 52
Filter Processing and DNA Extraction 58
PCR and High-Throughput Amplicon Sequencing 58
Data Analysis 59
Modeling Analyses 60
RESULTS 66
Alpha Diversity Measures and Generalized Linear Modeling 66
Occupancy Models 73

DISCUSSION 77
Alpha Diversity Measures and Generalized Linear Modeling 77
Species Distribution Modeling 78
Occupancy Models 78
CONCLUSIONS AND REFLECTIONS 83
LITERATURE CITED 85
APPENDIX 90

CHAPTER 1: AN OVERVIEW OF ENVIRONMENTAL DNA METABARCODING: PAST, PRESENT, AND FUTURE APPLICATIONS IN WETLANDS

ABSTRACT

Characterizing associations between habitat characteristics and species distributions is critical to inform conservation strategies for imperiled species. Essential ecological information is lacking for many species of threatened or endangered reptiles and amphibians in wetlands of the Great Lakes region. The scarcity of information is due in part to sampling difficulties due to low species abundance, cryptic behavior, and ephemeral use of wetlands. Environmental DNA (eDNA) metabarcoding is a novel survey tool that can be used to detect cryptic species otherwise not detected easily using traditional sampling methods (Taberlet et al., 2012; Valentini et al., 2009). Environmental DNA has been implemented in previous studies for single-species surveys (Adams et al., 2019; Ficetola et al., 2008; Hunter et al., 2015) and widely applied in aquatic systems since its introduction (Euclide et al., 2021; Pukk et al. 2021). Environmental DNA has also been applied within wetland systems (Goldberg et al., 2018; McKee et al., 2015) but in comparatively fewer studies than in lake or river systems. Our study's overarching goal was to design and implement a community-based herpetofauna metabarcoding survey using eDNA collections within wetlands across the Great Lakes region. Surveys were designed to address challenges experienced in these diverse systems. Our objectives included development and optimization of eDNA sampling protocols for wetland habitats, implementation of a regional eDNA survey to identify hotspots of wetland diversity and to update species distribution data, and the use of species distribution and occupancy models to characterize habitat requirements for species of conservation concern. Through eDNA protocol optimization, we sought to identify best practices for eDNA sampling in future monitoring and research applications in wetlands.

Occupancy models used in this project also informed future survey efforts by providing estimates of species' occupancy based on site-specific physical environmental conditions. By combining eDNA metabarcoding and species distribution modeling, this project helped identify environmental characteristics that predict wetland biodiversity, addressing knowledge gaps concerning habitat requirements for threatened wetland herpetofauna.

INTRODUCTION

Wetlands are important landscape features that provide ecosystem services to the communities around them (Constanza et al., 2014; De Groot et al., 2012; Mitsch & Gosselink, 2000). Sediment stabilization and water filtration are two of the key services provided by wetland systems (Wolfson et al., 2002). In addition, wetlands are critical for many facultative and obligate aquatic species across the continental United States, including many threatened and endangered reptile and amphibian species. Herpetofauna are often overlooked species when considering implications of climate change, yet they are at risk across the globe (Gibbons et al., 2000). While wetlands and the species that inhabit them are critical to the function of an area's greater ecosystem, wetlands have and continue to experience large amounts of fragmentation and degradation (Millennium Ecosystem Assessment, 2005). In Michigan alone, approximately 4 million acres of wetland habitat has been lost since European settlement in the 1800s (Fizzel et al., 2015). Anthropogenic impacts have drastically changed the structure, floral and faunal species composition and diversity, and habitat connectivity in the Great Lakes region. Consequently, the distribution and abundance of species and the habitat features associated with these systems have likewise changed. Threats to wetlands combined with a lack of information about distribution and habitat requirements for many herpetofauna species of conservation

concern make management approaches difficult. This lack of information is due, in part, to challenges associated with sampling wetland habitats.

Traditional sampling methods such as observational or call sampling may be insufficient to characterize entire wetland herpetofauna communities. Certain species could be missed in calling surveys if targeted at particular times of the breeding season, if they have quiet calls, or call less frequently than other species (Čeirāns et al., 2020; Crouch & Paton, 2002). In addition, life history patterns and ephemeral use of wetland systems mean timing of observational or call surveys are critical (Williams et al., 2013). Sampling within such short windows (a few hours during two to three days) could potentially leave room for false negatives for presence/absence reports. In addition, many herpetofauna species are difficult to find and observe using sampling methods that may lack the breadth necessary to properly characterize entire communities. Recent studies suggest that the use of non-invasive genetic sampling, such as environmental DNA (eDNA) metabarcoding, can improve wetland examination and biodiversity estimates, leading to better management (McKee et al., 2015; Saenz-Agudelo et al., 2002).

Environmental DNA is DNA obtained from an environmental sample collected without direct contact with a target organism or the organism being present. Sources of eDNA include skin cells, feces, and extracellular DNA shed into the environment. (Taberlet et al., 2012). eDNA can be taken from water, air, or soil samples. Using eDNA with on-site and remote sensing of wetland habitat features, we can improve our understanding of the influences of environmental characteristics on species distributions. Single species eDNA surveys allow individual species detection in herpetofauna communities using species-specific primers and quantitative PCR (Adams et al., 2017; Ficetola et al., 2008). More recently, eDNA metabarcoding has allowed for non-invasive examination of entire communities (Andruszkiewicz et al., 2017; Deiner et al.,

2017; Kelly et al., 2014). Given the ability to simultaneously detect multiple species of interest, eDNA metabarcoding provides an efficient and low-cost alternative to multi-species capture or call-based surveys, with great potential for implementation in wetland systems. Species detections from eDNA metabarcoding surveys can also be used with species distribution and occupancy modeling to identify environmental characteristics associated with species presence/absence and wetland biodiversity (Muha et al., 2017). The utility of eDNA metabarcoding data for estimating relative abundance has been debated. Previous studies suggest that eDNA metabarcoding data can provide information on the relative abundance of species (Hänfling et al., 2016; Sard et al., 2019), but there are many biotic (life stage, reproductive status) and abiotic (temperature, pH, flow) factors that may influence eDNA production and decay rates natural systems. Other studies such as Yates et al., 2019 suggest that while in laboratory settings, abundance and eDNA data show correlation, it can be difficult to replicate in natural systems and suggest refining techniques to accurately describe abundance using eDNA. In small closed systems like wetlands, it is not clear whether species sequence counts can be attributed to an individual organism's time spent in the environment (ephemeral wetland use, different life stage detections) or species biomass. The physical complexity of wetland systems and the unique ecologies of herpetofauna species make the environmental DNA metabarcoding an attractive monitoring and research tool for wetlands community analyses.

Here, we apply eDNA metabarcoding to inland wetland systems across three states in the Great Lakes region. The primary goal of this research was to identify habitat characteristics that were predictive of site occupancy of threatened and endangered herpetofauna. This regional study also characterized wetland biodiversity and identified associations between species richness and environmental variables associated with wetland systems.

The need for additional information on threatened and endangered herpetofauna in the Great Lakes is great (Hecnar et al., 2004). This thesis reports on interdisciplinary approaches to herpetofauna assessment based on molecular methods, quantitative assessment of alternate field sampling practices, on-site and remote characterizations of wetland habitat features during two years and field sampling periods within each year, and species distribution/occupancy modeling. The breadth of information will inform management strategies for ecologically diverse species of conservation concern.

OBJECTIVES

1. Compare and evaluate the efficacy of eDNA sampling protocols for wetland habitats, focusing on three aspects of the survey: water sampling, survey timing, and eDNA marker choice.

We conducted a pilot study in 10 Michigan wetlands to evaluate eDNA sampling protocols in diverse wetland environments for reptiles and amphibians. Surveys allowed comparisons of three aspects of the experimental design: water sampling approach (point versus transect sampling), the timing of sampling (differences associated with season), and the locus targeted for eDNA metabarcoding (12S versus 16S rRNA genes from the mitochondrial genome). Other markers have been previously developed for amphibians and reptiles separately (Lacoursière-Roussel et al., 2016), but the two loci chosen for this study are able to simultaneously detect both reptiles and amphibians due to sequence conservation and previous development of universal vertebrate primers (Deagle et al., 2009; Harper et al., 2019; Pukk et al., 2021; Riaz et al., 2011). Additionally, the availability of sequences for North American herpetofauna in online repositories made 12S and 16S mitochondrial markers suitable for this study. To facilitate future wetland eDNA surveys, our baseline sequence database (comprised of available sequences from

international sequence repositories; Sayers et al., 2020) was supplemented by sequencing tissue samples obtained from museum collections or provided by colleagues. These additional data improved the representation of Great Lakes herpetofauna species. Comparisons between point and transect sampling were made, as well as assessments of the differences in species compositional estimates between seasons. We hypothesized that point samples would provide fewer sequence reads and detect fewer species when compared to transect samples due to the increased area sampled during transect sampling. Further, we hypothesized that patterns of detection for taxonomic groups would reflect the time wetlands were occupied (based on requirements of different life stages of different species: e.g., detection of amphibians earlier in the year than reptiles).

2. Identify environmental attributes that coincide with a) site occupancy for 22 species of conservation concern and b) wetland biodiversity across Michigan, Ohio, and Indiana.

Using optimal sampling methods identified in Chapter 2, we surveyed ~50 wetland sites across Michigan, Indiana, and Ohio. To account for ephemeral habitat use, our surveys included early and late summer collections but focused on a single mitochondrial DNA marker and water sampling method. We combined species detections from our eDNA biodiversity surveys with on-site and remotely sensed environmental variables in species distribution and occupancy models to identify the local and landscape-scale habitat characteristics predictive of species presence and biodiversity of the wetland herpetofauna. We hypothesized that anthropogenic influences such as development and agricultural land use would negatively impact biodiversity of wetland communities. We also hypothesized that occupancy of amphibian species would be affected by wetland type, specifically degree of ephemerality and environmental conditions such as pH and predator presence/absence.

In accordance with the Great Lakes Fish and Wildlife Restoration Act of 1990, as reauthorized in 2006 (GLFWRA, 2006) the goal of this work was to gain an understanding of wetland communities that include state and federally-listed species. Despite the ecological importance and rapid habitat degradation and fragmentation of wetlands, these systems remain understudied. The information gathered from this project has potential to inform the management of wetland habitats in the Great Lakes region. Information will assist managers in identifying hotspots of wetland herpetofauna biodiversity in the region and will characterize associations between biodiversity and the landscape and habitat features in the sites they manage. Additionally, this study contributed updated occurrence records for several threatened and endangered species. This project also identified best sampling practices for future wetland eDNA surveys and improved the representation of Great Lakes herpetofauna in DNA sequence databases. Collectively, this project provided an outreach opportunity to connect with agency cooperators, stakeholders, and other groups interested in incorporating eDNA into wetland research efforts.

Prior to the data collection and analysis, we expected to see relationships between habitat connectivity, habitat quality, and species richness across all species of conservation concern. We anticipated that variables such as wetland ephemerality would dictate seasonal occupancy for a variety of species in different capacities. From data collected, we expected that the implementation of environmental DNA would be feasible in wetlands, with some constraints due to the diverse nature of these ecosystems. We expected to see differences in habitat requirements between taxonomic groups depending on their space use, seasonal occupancy, and life history traits. Further, we expected to identify associations between species presence/absence and site

characteristics including abiotic (temperature, salinity, pH) and biotic factors (predator presence) using occupancy models.

The standardized wetland eDNA assessment protocols we provided in the second chapter of this thesis will enable management agencies to conduct efficient and accurate wetland herpetofauna community assessments. The third chapter of this thesis identified areas of biodiversity across Great Lakes wetlands that will assist managers in the prioritization of conservation and restoration efforts. By combining the optimization of eDNA sampling protocols, implementation of a regional eDNA survey, and the use of species distribution and occupancy models, we sought to identify environmental characteristics that predict wetland biodiversity and address knowledge gaps concerning habitat requirements for threatened herpetofauna.

LITERATURE CITED

Adams, C.I., Hoekstra, L.A., Muell, M.R., & Janzen, F.J. 2019. A Brief Review of Non-Avian Reptile Environmental DNA (eDNA), with a Case Study of Painted Turtle (Chrysemys picta) eDNA Under Field Conditions. *Diversity*, *11*(4), 50. doi:10.3390/d11040050

Andruszkiewicz, E.A., Starks, H.A., Chavez, F.P., Sassoubre, L.M., Block, B.A., & Boehm, A.B. 2017. Biomonitoring of marine vertebrates in Monterey Bay using eDNA metabarcoding. *PLOS ONE*, *12*(4), e0176343. doi:10.1371/journal.pone.0176343

Čeirāns, A., Pupina, A., & Pupins, M. 2020. A new method for the estimation of minimum adult frog density from a large-scale audial survey. *Scientific Reports*, *10*(1), 1-12. doi:10.1038/s41598-020-65560-6

Costanza, R., De Groot, R., Sutton, P., Van Der Ploeg, S., Anderson, S.J., Kubiszewski, I., Farber, S., Turner, R.K. 2014. Changes in the global value of ecosystem services. *Global Environmental Change*, 26:152–158. doi:10.1016/j.gloenvcha.2014.04.002

Crouch, W.B., & Paton, P.W.C. 2002. Assessing the Use of Call Surveys to Monitor Breeding Anurans in Rhode Island. *Journal of Herpetology*, *36*(2), 185–192. doi:10.2307/1565990

Deagle, B.E., Kirkwood, R., & Jarman, S.N. 2009. Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular Ecology*, *18*(9), 2022– 2038.

De Groot, R., Brander, L., Van Der Ploeg, S., Costanza, R., Bernard, F., Braat, L, Christie, M., Crossman, N., Ghermandi, A., Hein, L. 2012. Global estimates of the value of ecosystems and their services in monetary units. *Ecosystem Services*, 1:50–61. doi:10.1016/j.ecoser.2012.07.005

Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., Pfrender, M.E., & Bernatchez, L. 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872-5895. doi:10.1111/mec.14350

Euclide, P.T., Lor, Y., Spear, M.J. 2021. Environmental DNA metabarcoding as a tool for biodiversity assessment and monitoring: reconstructing established fish communities of north-temperate lakes and rivers. *Diversity and Distributions,* 2021;00:1–15. doi:10.1111/ddi.13253

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

Fizzel, C. 2015. Status and trends of Michigan's wetlands: pre-European settlement to 2005.

Michigan Department of Environment, Great Lakes, and Energy. https://www.michigan.gov/media/Project/Websites/egle/Documents/Programs/WRD/Wetlands/Status-Trends-of-Michigans-Wetlands-PreEuropean-Settlement-to-2005.pdf?rev=30194c230acd42d98c2d92624ab19a6c

Gibbons, J.W., Scott, D.E., Ryan, T.J., Buhlman, K.A., Tuberville, T.D., Metts, B.S., Green, J.L., Mills, T., Leiden, Y., Poppy, S., Winne, C.T. 2000. The global decline of reptiles, déjà vu amphibians. *Bioscience,* 50, 653. doi:10.1641/00063568(2000)050[0653:tgdord]2.0.co;2

Goldberg, C.S., Strickler, K.M., & Fremier, A.K. 2018. Degradation and dispersion limit environmental DNA detection of rare amphibians in wetlands: Increasing efficacy of sampling designs. *Science of The Total Environment*, 633, 695-703. doi:10.1016/j.scitotenv.2018.02.295

Great Lakes Fish and Wildlife Restoration Act. S. Pub. L. 109–326, § 120 Stat. 1761 (2006).

Hänfling, B., Handley, L.L., Read, D.S., Hahn, C., Li, J., Nichols, P., Blackman, R.C., Oliver, A., Winfield, I.J. 2016. Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, *25*(13), 3101-3119. doi:10.1111/mec.13660

Harper, L.R., Lawson Handley, L., Hahn, C., Boonham, N., Rees, H.C., Lewis, E., Adams, I.P., Brotherton, P., Phillips, S., Hänfling, B. 2019. Generating and testing ecological hypotheses at the pondscape with environmental DNA metabarcoding: A case study on a threatened amphibian. *Environmental DNA.* 1–16. doi:10.1002/edn3.57

Hecnar, S.J. 2004. Great Lakes wetlands as amphibian habitats: A review. *Aquatic Ecosystem Health & Management, 7*(2): 289–303. doi:10.1080/14634980490461542

Hunter, M.E., Oyler-McCance, S.J., Dorazio, R.M., Fike, J.A., Smith, B.J., Hunter, C.T., Reed, R.N., & Hart, K.M. 2015. Environmental DNA (eDNA) sampling improves occurrence and detection estimates of invasive Burmese pythons. *PLOS ONE*, *10*(4), e0121655. doi:10.1371/journal.pone.0121655

Kelly, R.P., Port, J.A., Yamahara, K.M., & Crowder, L.B. 2014. Using environmental DNA to census marine fishes in a large mesocosm. *PLOS ONE*, *9*(1), e86175. doi:10.1371/journal.pone.0086175

Lacoursière-Roussel, A., Dubois, Y., Normandeau, E., Bernatchez, L., 2016. Improving herpetological surveys in eastern North America using the environmental DNA method. *Genome.* 59, 991–1007. doi:/10.1139/gen-2015-0218

McKee, A.M., Calhoun, D.L., Barichivich, W.J., Spear, S.F., Goldberg, C.S., Glenn, T.C. 2015.

Assessment of environmental DNA for detecting presence of imperiled aquatic amphibian species in isolated wetlands. *Journal of Fish and Wildlife Management, 6*(2): 498–510. doi:10.3996/042014-JFWM-034

Millennium Ecosystem Assessment. 2005. Ecosystems and human well-being: Wetlands and Water Synthesis. World Resources Institute, Washington, DC.

Mitsch, W.J., Gosselink, J.G. 2000. The value of wetlands: importance of scale and landscape setting. *Ecological Economics,* 35:25–33. doi:10.1016/S0921-8009(00)00165-8

Muha, T.P., Rolla, M., & Tricarico, E. 2017. Using environmental DNA to improve species distribution models for freshwater invaders. *Frontiers in Ecology and Evolution*, 5, 312785. doi:10.3389/fevo.2017.00158

Pukk, L., Kanefsky, J., Heathman, A.L., Weise, E.M., Nathan, L.R., Herbst, S.J., Robinson, J.D. 2021. eDNA metabarcoding in lakes to quantify influences of landscape features and human activity on aquatic invasive species prevalence and fish community diversity. *Diversity and Distributions*, 27, 2016–2031. doi:10.1111/ddi.13370

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., Coissac, E., 2011. EcoPrimers: Inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, 39, 1–11. doi:10.1093/nar/gkr732

Saenz-Agudelo, P., Delrieu-Trottin, E., DiBattista, J. D., Martínez-Rincon, D., Morales-González, S., Pontigo, F., Ramírez, P., Silva, A., Soto, M., & Correa, C. 2022. Monitoring vertebrate biodiversity of a protected coastal wetland using eDNA metabarcoding. *Environmental DNA*, *4*(1), 77-92. doi:10.1002/edn3.200

Sard, N., Scribner, K.T., Robinson, J.D., Herbst, S., Kanefsky, J., 2019. Comparison of fish detections, community diversity and relative abundance using environmental DNA metabarcoding and traditional gears. *Environmental DNA*, *1*(4), 368-384. doi:10.1002/edn3.38

Sayers, E.W., Cavanaugh, M., Clark, K., Ostell, J., & Pruitt, K. D. 2020. GenBank. *Nucleic Acids Research*, *48*(D1), D84-D86. doi:10.1093/nar/gkz956

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. 2012. Environmental DNA. *Molecular Ecology*, *21*(8), 1789-1793. doi:10.1111/j.1365-294X.2012.05542.x

Valentini, A., Pompanon F., Taberlet, P. 2009. DNA barcoding for ecologists. *Trends in Ecology & Evolution*, 24, 110–117. doi:10.1016/j.tree.2008.09.011

Williams, P.J., Engbrecht, N.J., Robb, J.R., Terrell, V.C.K., Lannoo, M.J. 2013. Surveying a threatened amphibian species through a narrow detection window. *Copeia*, 3, 552–561.

Wolfson, L., Mokma, D., Schultink, G. and Dersch, E. 2002. Development and use of a wetlands information system for assessing wetland functions. *Lakes & Reservoirs: Research & Management*, 7, 207-216. doi:10.1046/j.1440-1770.2002.00189.x

Yates, M.C., Fraser, D.J., & Derry, A.M. 2019. Meta-analysis supports further refinement of eDNA for monitoring aquatic species-specific abundance in nature. *Environmental DNA*, *1*(1), 5-13. doi:10.1002/edn3.7

CHAPTER 2: DEVELOPMENT AND OPTIMIZATION OF WETLAND ENVIRONMENTAL DNA METABARCODING PROTOCOLS TO SURVEY HERPETOFAUNA IN THE GREAT LAKES REGION

ABSTRACT

Wetlands are important systems for many species of reptiles and amphibians (herpetofauna), including many threatened and endangered species. Despite the critical importance of wetlands to herpetofauna, wetlands are being degraded and lost at a high rate, which poses a threat to the persistence of resident species. Characterization of herpetofauna community diversity and relative species abundance and distribution in different wetland types and in different geographic areas is an important requirement to guide conservation strategies. However, sampling using traditional methods that may fall within a small temporal window relative to periods and life stages of species occupancy may fail to accurately characterize some components of wetland communities. In contrast, environmental DNA has been shown to effectively survey entire aquatic communities. Applications in wetlands, particularly those that host threatened herpetofauna have been more limited. Questions about how to best implement eDNA sampling in structurally heterogeneous wetlands like fens, bogs, and marshes require further research. The overall objective of this study was to design and optimize eDNA sampling and laboratory protocols for wetland herpetofauna surveys. Protocols evaluated included different water sampling approaches (point versus transect sampling), seasonality of sampling periods (early versus late summer), and the molecular marker used for metabarcoding (mitochondrial 12S versus 16S rDNA). Sample collection was conducted at 10 sites across the southern portion of the Lower Peninsula of Michigan. We documented 18 amphibian and 11 reptile species, including four species of conservation concern. We observed no difference in the number of species detected between water samples collected at a single point and across a 5m transect (p =

0.52), but point sampling required significantly less time ($p = 0.04$) and allowed significantly larger volumes of water to be filtered ($p = 0.0005$). No difference in species richness was observed between community sequence data obtained from 12S and 16S mitochondrial DNA markers ($p = 0.61$). However, examination of reference sequence databases indicated a greater number of taxa were identifiable at the species level when using the 16S locus. There was a significant different in the number of species detected between late and early sampling periods ($p = 7e\text{-}7$). However, some species were found in only the early sampling period, so sample collection should occur in both periods. In addition to comparisons of field protocols, a taxonomic database of 65 species was developed as a resource for both the 12S and 16S markers to identify species in the Great Lakes region.

INTRODUCTION

Wetlands are crucial habitats for many aquatic species across the midwestern United States, including many threatened and endangered reptile and amphibian (herpetofauna) species (Gibbons, 2003, Sierszen et al., 2012). Wetlands are experiencing habitat fragmentation and degradation at a high rate (Dahl et al., 2000; Davidson, 2014), leading to species extirpation and declines in distribution and abundance. Due to low abundance and varying life history traits, species vary in periods of wetland occupancy during and across different life stages. Thus, threatened and endangered herpetofauna may be less often observed within structurally heterogeneous wetland habitats based on surveys conducted using traditional sampling methods. Human encroachment and habitat loss within wetlands are also intensified by a lack of information about distribution and habitat requirements for many herpetofauna species of conservation concern. Differences in wetland environmental features including water and air temperature, pH, predator presence/absence and type, structural habitat complexity, and

hydroperiod can influence wetland suitability for a specific species or taxonomic group (Davis et al., 2017; Korfel et al., 2010; Skidds & Golet, 2005). Additionally, many reptiles and amphibians use wetland habitats facultatively or for only a portion of the year, which varies depending on species' ecology. The combination of wetland ephemerality, differences in habitat use between species, and wetland physical characteristics can complicate sampling efficiency. Thus, surveying wetlands can be challenging when attempting to characterize entire communities.

Traditional sampling methods such as observations, pitfall traps, or call surveys alone are insufficient to characterize entire herpetofauna communities. Call surveys are not able to be implemented for species such as snakes, salamanders, and turtles that do not call. Other issues that could arise with call surveys are observer bias, weather, temporal components, or misidentification (Lotz & Allen, 2007; Mazerolle et al., 2005; Schmidt et al., 2005). Other traditional sampling methods may consist of capture-based surveys which are time-consuming. Life history traits such as breeding seasonality, ephemeral habitat use, and call volume/frequency limit the utility of "snapshot" surveys conducted at a single point in time (Čeirāns et al., 2020; Crouch & Paton, 2002), thus motivating the need for development of more inclusive and accurate survey techniques. Recent studies suggest that non-invasive genetic sampling, such as environmental DNA (eDNA) surveys, can improve efforts to characterize wetland biodiversity (McKee et al., 2015; Saenz-Agudelo et al., 2002; Wikston et al., 2023) and have demonstrated greater sensitivity of detecting species (Evans et al., 2017). Additionally, eDNA surveys may mitigate sampling bias of certain taxa by using deposited DNA in the study system (Cilleros et al., 2019; Deiner et al., 2017).

Single-species eDNA surveys are a useful tool to detect individual species presence/absence within aquatic habitats, and have been utilized for herpetofauna (Adams et al.,

2019; Dejean et al., 2012; Hunter et al., 2015; Piaggio et al., 2014). Additionally, quantitative PCR and the combination of multiple taxon-specific primers allow for high sensitivity when detecting a species using this approach (Brys et al., 2023). More recently, eDNA metabarcoding has allowed for the characterization of entire aquatic communities (Pukk et al., 2021; Taberlet et al., 2014; Valentini et al., 2016). Due to the ability to detect multiple species of interest and rapid advances in high-throughput sequencing technologies (Deiner et a., 2017), eDNA metabarcoding provides an efficient and low-cost alternative to capture or call-based surveys, with great potential for implementation in diverse wetland systems. Environmental DNA sampling approaches are an enticing method to improve species detection and distribution patterns (Bohmann et al., 2014; Thomsen and Willerslev, 2015). However, different methods of eDNA sampling in wetland systems must be tested and optimized to ensure efficiency and accuracy in data collection for biodiversity monitoring and research.

Our objectives were to optimize field sampling and laboratory protocols for eDNA metabarcoding surveys targeting threatened and endangered herpetofauna inhabiting diverse wetland habitats. Protocol evaluations focused on three comparisons: i) physical water sampling methods (point versus transect sample collection), ii) choice of metabarcoding marker marker for the detection of reptile and amphibian species (mitochondrial 12S versus 16S rDNA), and iii) timing of sampling within a field season (early versus late summer). Results from this project will provide critical information for the design of future wetland eDNA surveys.

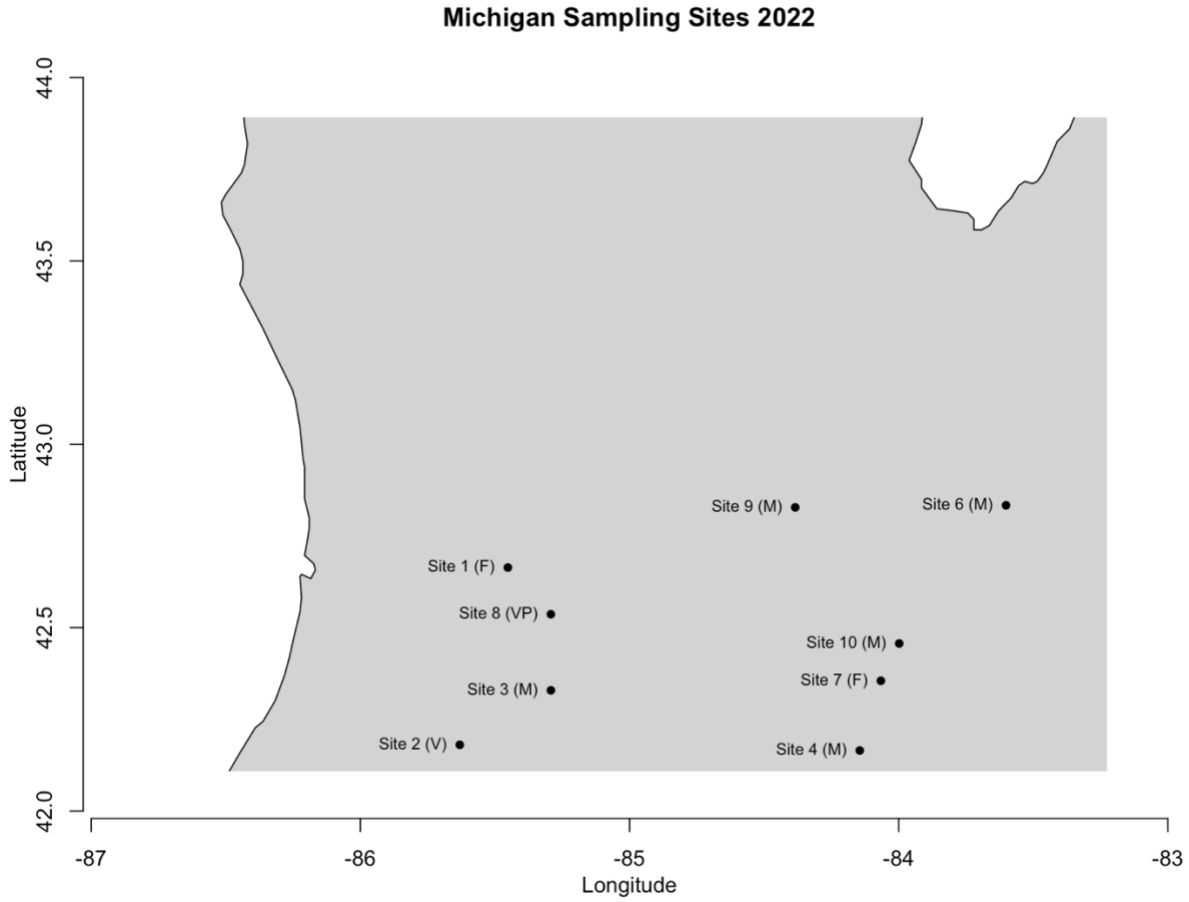
METHODS

Field Collections

Sampling sites for this study were selected across the southern Lower Peninsula of Michigan (Figure 1.1). We attempted to collect 10 samples in one wetland, at each of 10 selected sites,

during each of two time periods. All 10 samples were collected at the same wetlands within a site in each time period. The first time period was May 22nd, 2021 to June 23rd, 2021, and the second sampling occurred between July 21st, 2021 to August 4th, 2021. Due to the lack of water and general ephemerality of wetlands, some wetlands did not have enough water to properly sample in the second (later) period. For all project sampling, 1L water samples were collected and filtered using single-use, sterile filter packs (polyethersulfone (PES) filters, 1.2 µm porosity) and a Smith-Root ANDe eDNA backpack filtering device (Smith-Root, Vancouver, WA). Point samples were collected by filtering 1L of water from a single location in the wetland. Transect samples were collected (in the same areas) by moving the eDNA boom in an arc motion of ~5m during water filtration. Because shallow water levels and sedimentation often affected filtering efficiency, the duration of filtering and total water volume filtered were recorded. After filtering 1L of water, the filter was removed from the housing using antiseptic techniques and placed in 95% ethanol in 5 ml tubes for preservation and transportation to the laboratory. All equipment used for sampling was single-use (filters) or decontaminated with 20% bleach solution following eDNA best sampling practices (Dickie et al., 2018; Prince and Andrus, 1992). Water temperature was recorded at each site during sampling using a digital thermometer. Air temperature was recorded from local weather stations. After all samples from a site were filtered and preserved, one liter of distilled water was filtered at each site and sampling time period for use as a negative (no DNA) field control.

Figure 1.1. A site map for Michigan sampling sites and their wetland classification[1] for the 2022 field season.



[1] F = fen, M = marsh, VP = vernal pool

Filter Processing and DNA Extraction

After transportation to the laboratory, samples were stored in 95% ethanol at room temperature. For processing, filters were removed from storage ethanol to be dried and cut into multiple pieces. After drying, DNA extractions were performed for all cut filters using Qiagen Qia-shredder and DNeasy Blood and Tissue kits (QIAGEN, Venlo, NL) using the protocol described in Laramie et al. (2015). Zymo One-Step columns (Zymo Research, Irvine, CA) were used to remove PCR inhibitors. DNA was eluted in a final volume of 100 µL in Buffer AE (10 mM Tris-Cl. 0.5 mM EDTA; pH 9.0) (QIAGEN, Venlo, NL) and stored at -20 °C for later processing. One sterile filter was dried and extracted as per the above methods to serve as a negative (no DNA) control to quantify levels of contamination during the DNA extraction stages.

PCR and High-Throughput Sequencing

Following DNA extraction and inhibitor removal, two loci (12S and 16S rDNA mitochondrial rDNA gene regions) were amplified using previously published protocols from work in our lab (Pukk et al., 2021; Sard et al., 2019) and primer sets for both 12S (Riaz et al., 2011) and 16S (Deagle et al., 2009). For both loci, the following reagents were combined in a sterile tube (25 µl per reaction): 1X AmpliTaq Gold PCR Buffer II (no Mg^{2+}), 1 µg/µl of BSA, 1.25 U AmpliTaq Gold DNA polymerase, and 0.32 µM of forward and reverse primer. The forward and reverse primers included CS1 and CS2 oligo tails (Fluidigm). The 12S and 16S assays also included 0.24 and 0.32 mM dNTPs, and 2.0 and 2.5 mM $MgCl_2$ (respectively). For the 16S assays, samples were incubated at 95° C for 10 minutes, followed by 35 cycles of 95° C for 30 seconds, 57° C for 45 seconds, and 72° C for 45 seconds, with a final elongation step at 72° C for 5 minutes. For the 12S assays, samples were incubated at 95° C for 10 minutes, followed by 35 cycles of 95° C for 30 seconds, 57° C for 30 seconds, and 72° C for 45 seconds, with a final elongation step at 72° C

for 5 minutes. PCR products were cleaned using a Qiagen Qiaquick PCR Purification kit and concentrations were determined using a NanoDrop 1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE). One PCR negative was added to each plate that contained deionized water and all PCR reagents to serve as the negative (no DNA) sample during the PCR stage.

Following amplification, samples were individually barcoded using i7 (plate) and i5 (sample) primers added to amplified PCR products. This mixture included the desired i7 index (10 μM) and 2X Qiagen Plus Master Mix (QIAGEN; #206152) for a reaction volume of 6 μl. The i5 indices (5 μM) were added to each well of the plate. For barcoding, sample plates were placed in the thermocycler and incubated at 95° C for 15 minutes, followed by 10 cycles of 95° C for 10 seconds, 65° C for 30 seconds, and 72° C for 30 seconds, with a final elongation step at 72° C for 5 minutes. Once each sample had been assigned an i5 and i7 primer, distinguishing them from other samples, they were pooled for bead size selection by combining all samples from a plate into one 1.5 ml tube. Once pooled, size selection was performed using AMPure XP beads (Beckman Coulter, Indianapolis, IN). To remove long sequences, we performed a 0.5X (1:2 ratio) reaction of beads to pooled library volume. Short sequences were removed using a 1.2X reaction of beads to pooled library volume. To quantify the concentration of each library, samples were analyzed using a Qubit fluorometer (Life Technologies, ThermoFisher Scientific, USA). Due to the high concentration of eluted DNA in each library, it was required to dilute each one to a 1:10 ratio. Once diluted, libraries were prepared for Qubit analysis. Diluted

libraries were sent to be processed on the TapeStation (assay High Sensitivity D1000 ScreenTape®) to confirm sample concentrations and base-pair lengths. Once returned, fully eluted libraries were sequenced on an Illumina MiSeq lane (generating 150bp paired-end reads) at Michigan State University's Research Technology Support Facility.

Mock Communities

In addition to the eDNA samples collected in the field, we sequenced artificial mixtures with known amounts of genomic DNA from multiple species ("mock communities") to assess amplification bias among targeted herpetofauna species. This was done by obtaining genomic DNA from tissue samples of known herpetofauna species in the lab that represented the range of taxonomic diversity expected to be encountered across surveyed sites. We created three separate communities using different relative volumes of species DNA. For Mock Community 1 (MC1), genomic DNA from ten species was mixed in equal concentrations. MC2 and MC3 were mixtures of genomic DNA from the same ten species, but in different proportions (ranging from 25% to 1% of the total gDNA in the sample; Table 1.1). To assess amplification bias in mock communities, we tested for correlation between the relative concentration of each species' genomic DNA in the mixture and the relative abundance of sequencing reads from the sample for both the 12S and 16S markers.

Table 1.1. Observed and expected sequence read proportions for three mock communities containing genomic DNA from 10 species. Expected proportion is the proportion of known genomic DNA placed in the sample and observed proportion is the proportion of sequencing reads from the sample attributed to the species.

Mock community	**Species DNA Mixture MC #1**				**Species DNA Mixture MC #2**				**Species DNA mixture MC #3**			
	12S		**16S**		**12S**		**16S**		**12S**		**16S**	
	Obs.	**Exp.**	**Obs.**	**Exp.**	**Obs.**	**Exp.**	**Obs.**	**Exp.**	**Obs.**	**Exp.**	**Obs.**	**Exp.**
Species												
Rana sylvatica	0.13	0.10	0.12	0.10	0.30	0.25	0.23	0.25	0.01	0.01	0.01	0.01
Pseudacris crucifer	0.07	0.10	0.08	0.10	0.13	0.20	0.17	0.20	0.02	0.03	0.03	0.03
Bufo americanus	0.08	0.10	0.22	0.10	0.12	0.15	0.26	0.15	0.04	0.04	0.12	0.04
Sistrurus catenatus	0.00	0.10	0.05	0.10	0.00	0.12	0.05	0.12	0.00	0.05	0.04	0.05
Crotalus horridus	0.00	0.10	0.32	0.10	0.00	0.09	0.23	0.09	0.00	0.06	0.29	0.06
Ambystoma texanum	0.06	0.10	0.00	0.10	0.04	0.06	0.00	0.06	0.05	0.09	0.00	0.09
Plethodon cinereus	0.07	0.10	0.13	0.10	0.04	0.05	0.04	0.05	0.08	0.12	0.26	0.12
Sternotherus odoratus	0.00	0.10	0.04	0.10	0.00	0.04	0.01	0.04	0.00	0.15	0.12	0.15
Terrapene carolina carolina	0.54	0.10	0.00	0.10	0.34	0.03	0.00	0.03	0.71	0.20	0.00	0.20
Plestiodon fasciatus	0.02	0.10	0.03	0.10	0.00	0.01	0.00	0.01	0.06	0.25	0.13	0.25
Other/Unidentified	0.03	0.00	0.01	0.00	0.02	0.00	0.01	0.00	0.03	0.00	0.01	0.00

Taxonomic Database Development

While existing sequence data were available for many threatened and endangered herpetofauna, the taxonomic databases used for our metabarcoding analyses (available from: https://github.com/rupperto1/Herp-eDNA-Taxonomic-Database-2020-2023) did have notable species gaps. We started with sequences for both the 12S and 16S loci sourced from National Center for Biotechnology Information's GenBank repository (Sayers et al., 2020) for 73 species. To expand representation in our database, we generated new sequence data for the 16S locus for 53 herpetofauna species using tissues loaned from The Field Museum of Natural History, The University of Michigan Museum of Zoology, the United States Fish & Wildlife Service, The Chicago Academy of Sciences/Peggy Notebaert Nature Museum, Grand Valley State University, University of Wisconsin- Stevens Point, and Michigan State University.

After transportation to the laboratory, samples were stored long-term in 90% ethanol at room temperature. Once ready for processing, tissue samples were removed from ethanol and DNA was extracted using Qiagen DNeasy Blood & Tissue Kits (QIAGEN, Germantown, MD), following manufacturer protocols. DNA concentrations were quantified using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific; Waltham, MA), and targeted gene regions were amplified via the polymerase chain reaction. For the 16S locus, the same PCR recipe used for the metabarcoding was used. The thermal cycling protocol included initial denaturation at 95° C for 3 minutes, followed by 30 cycles of 95° C for 45 seconds, 57° C for 45 seconds, and 72° C for 45 seconds, with a final elongation step at 72° C for 5 minutes. PCR products were cleaned using a Qiagen Qiaquick PCR Purification kit and concentrations were determined using a NanoDrop 1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE). The cleaned products

were diluted to 10 ng/ul and sent for cycle sequencing at the Michigan State University Research Technology Support Facility.

Resulting sequences were aligned to an existing reference database in the software MEGA (MEGA11: Molecular Evolutionary Genetics Analysis; Tamura et al., 2021) (Table 1.2) and later used to match sample sequences from eDNA metabarcoding of water samples. The final database contains 77 individual sequences for the 12S locus and 283 individual sequences for the 16S locus. Of the total 360 sequences in this taxonomic database, our laboratory contributed 180. Pairwise base pair divergence between sequences in the taxonomic database (excluding alignment gaps) were assessed to determine the expected resolving power of the two markers. This analysis used functions from *seqinr* and *ape* R packages (Charif & Lobry, 2007; Paradis & Schliep, 2019).

Table 1.2. Final sequence counts in reference taxonomic databases used for eDNA metabarcoding with 12S and 16S markers. Combined, the two databases include representation for 70 reptile and amphibian species.

		Number of sequences	
Common name	**Scientific name**	**12S (Riaz)**	**16S (Deagle)**
Northern ribbon snake	*Thamnophis sauritus septentrionalis*	0	5
Blanchard's cricket frog	*Acris crepitans blanchardi*	1	0
Blue-spotted salamander	*Ambystoma laterale*	1	6
Spotted salamander	*Ambystoma maculatum*	0	4
Marbled salamander	*Ambystoma opacum*	0	2
Smallmouth salamander	*Ambystoma texanum*	1	5
Tiger salamander	*Ambystoma tigrinum*	1	5
Midland smooth softshell	*Apalone mutica*	1	0
Spiny softshell turtle	*Apalone spinifera*	1	4
Six-lined Racerunner	*Aspidoscelis sexlineata*	0	1
American toad	*Bufo americanus*	2	8
Fowler's toad	*Bufo fowleri*	3	5
Common snapping turtle	*Chelydra serpentina*	1	2
Painted turtle	*Chrysemys picta*	1	1
Spotted turtle	*Clemmys guttata*	0	5
Kirtland's snake	*Clonophis kirtlandii*	1	5
Black racer	*Coluber constrictor*	2	5
Timber rattlesnake	*Crotalus horridus*	1	3
Ring-necked snake	*Diadophis punctatus*	0	7
Blanding's turtle	*Emydoidea blandingii*	0	2
Northern two-lined salamander	*Eurycea bislineata*	1	3
Southern two-lined salamander	*Eurycea cirrigera*	1	6
Long-tailed salamander	*Eurycea longicauda*	0	4
Wood turtle	*Glyptemys insculpta*	1	6
Northern map turtle	*Graptemys geographica*	1	0
False map turtle	*Graptemys pseudogeographica*	1	0

Table 1.2 (cont'd)

Spring salamander	*Gyrinophilus porphyriticus*	6	5
Four-toed salamander	*Hemidactylium scutatum*	1	3
Eastern hognose snake	*Heterodon platirhinos*	1	5
Cope's gray treefrog	*Hyla chrysoscelis*	1	2
Gray treefrog	*Hyla versicolor*	1	6
Eastern milksnake	*Lampropeltis triangulum*	1	5
Wood frog	*Rana sylvaticus*	2	7
Common mudpuppy	*Necturus maculosus*	1	2
Plain-bellied water snake	*Nerodia erythrogaster*	1	2
Copperbelly water snake	*Nerodia erythrogaster neglecta*	0	7
Common watersnake	*Nerodia sipedon*	1	5
Eastern newt	*Notophthalmus viridescens*	1	6
Rough greensnake	*Opheodrys aestivus aestivus*	0	2
Smooth green snake	*Opheodrys vernalis*	0	5
Slender glass lizard	*Ophisaurus attenuatus*	1	3
Eastern foxsnake	*Pantherophis gloydi*	0	4
Gray rat snake	*Pantherophis spiloides*	0	5
Black rat snake	*Pantherophis obsoletus*	1	2
Western foxsnake	*Pantherophis vulpinus*	2	3
Common five-lined skink	*Plestiodon fasciatus*	1	5
Red-backed salamander	*Plethodon cinereus*	1	5
Spring peeper	*Pseudacris crucifer*	3	4
Slimy salamander	*Plethodon glutinosus*	2	2
Boreal chorus frog	*Pseudacris maculata*	2	7
Western chorus frog	*Pseudacris triseriata*	3	8
American bullfrog	*Rana catesbeiana*	1	7
Plains leopard frog	*Rana blairi*	1	1
Green frog	*Rana clamitans*	1	5
Pickerel frog	*Rana palustris*	1	2
Northern leopard frog	*Rana pipiens*	1	5
Mink frog	*Rana septentrionalis*	1	1

Table 1.2 (cont'd)

Queen snake	*Regina septemvittata*	1	5
Lesser siren	*Siren intermedia*	4	7
Western lesser siren	*Siren intermedia nettingi*	1	1
Eastern massasauga	*Sistrurus catenatus*	2	6
Eastern musk turtle	*Sternotherus odoratus*	2	4
De Kay's snake or brown snake	*Storeria dekayi*	1	6
Redbellied snake	*Storeria occipitomaculata*	1	4
Eastern box turtle	*Terrapene carolina carolina*	0	5
Ornate box turtle	*Terrapene ornata ornata*	0	5
Butler's garter snake	*Thamnophis butleri*	1	0
Plains garter snake	*Thamnophis radix*	1	4
Common garter snake	*Thamnophis sirtalis*	1	5
Red-eared slider turtle	*Trachemys scripta*	1	6

Data Analysis

Once eDNA water samples were processed and sequenced, bioinformatic analyses were conducted using the Mothur pipeline (Schloss, 2009) from previous eDNA metabarcoding research in our lab (e.g., Pukk et al., 2021; Sard et al., 2019). Briefly, raw sequences were demultiplexed and filtered for length and quality (12S: minimum length = 83, maximum length= 117, maximum homopolymers = 6; 16S: minimum length = 35, maximum length = 95, maximum homopolymers = 13) using the *make.contigs* function with the Mothur pipeline. This command extracts the sequence and quality score data from raw data files and reads the complemented and paired sequences into the pipeline. Mothur ver. 1.39.5 was used to group similar sequences into operational taxonomic units (OTUs; 99% pre-clustering threshold) and classify OTUs to the lowest taxonomic level possible, based on comparisons to marker-specific reference databases. The median number of reads for detections in negative controls was used as a threshold for detections in water samples associated with our project. Individual detections based on fewer reads than this threshold were removed from the community matrix prior to further analysis.

The R package *vegan* (Oksanen et al., 2020) was used to calculate alpha and beta diversity measures and analyze herpetofauna communities across the surveyed wetlands. ANOVAs were used to compare species richness, water volume filtered, and filtering duration from alternative water sampling methods (point vs. transect). Paired t-tests evaluated differences in species richness between loci. When quantifying the number of species detected per season, a linear model was used to compare estimates of species diversity between seasons and among wetland type. Models included the following variables: sample type, season, locus, wetland, and the interaction between wetland sample type. Metabarcoding markers were also compared on

their expected resolving power (number of taxa that are identifiable to the species level) and the number of species represented in our taxonomic databases.

RESULTS

Overview

A total of 1,053,508 sequence reads were generated for 177 water samples collected with an average of 6,543 (± 10,772) reads per sample. A total of 29 herpetofauna species were detected across 10 sampling sites. For each surveyed wetland (Figure 1), we calculated the Shannon diversity (Table 1.3) for both the 12S and 16S markers. The Kalamazoo County (1) site held the highest level of diversity when analyzed using a Shannon diversity calculation. For the 12S marker, Kalamazoo County (1) was also identified as having the highest level of diversity. For the 16S marker, the Barry County site was ranked as the site with the highest levels of species diversity using either diversity index. This may be influenced by the ability of the 16S to identify additional reptile and amphibian species that were not detected by the 12S marker. Measures of species alpha diversity (Shannon diversity) for the markers combined ranged from a value of 0.5758 to 1.4665 ($\mu = 0.54 \pm 0.44$). Results for point versus transect, marker comparisons, and seasonal differences are below. Detections based on fewer than two sequence reads (the median number of reads for detections in negative controls) were removed from the community matrix prior to further analysis.

Table 1.3. The number of samples taken per sampling period for A) point sampling and B) transect sampling, wetland type for each site, means, and standard deviations for water volumes and durations, and the number of species detected for both mitochondrial markers (12S and 16S).

A.

		Point Sampling									
		No. Samples		Water volume (L)		Duration of sampling (sec)		No. of species (12S gene)		No. of species (16S gene)	
		Period 1	Period 2	Period 1	Period 2	Period 1	Period 2	Period 1	Period 2	Period 1	Period 2
Site	Wetland Type			Mean (SD)		Mean (SD)					
Barry 1	Vernal pool	5	5	0.84 (0.30)	1.00 (0.13)	NA	211.00 (209.59)	4	2	6	1
Kalamazoo 1	Fen	5	5	0.89 (0.34)	0.98 (0.16)	NA	191.60 (171.70)	3	5	3	4
Kalamazoo 2	Marsh	4	3	1.08 (0.07)	0.96 (0.15)	NA	239.00 (240.76)	5	0	2	1
Jackson 1	Marsh	5	6	0.89 (0.34)	0.95 (0.25)	NA	133.40 (104.86)	3	4	4	3
Hillsdale	Fen	4	5	0.75 (0.35)	0.90 (0.28)	NA	169.75 (97.41)	0	0	1	1
Oakland	Marsh	3	3	1.10 (0.07)	1.05 (0.02)	NA	89.33 (9.61)	1	1	2	2
Jackson 1	Fen	6	3	0.64 (0.18)	1.02 (0.01)	NA	83.00 (5.66)	3	1	3	2
Barry 2	Vernal pool	3	3	1.02 (0.03)	1.07 (0.02)	NA	168 (52.09)	5	2	6	4
Clinton	Marsh	6	5	0.66 (0.41)	0.84 (0.23)	272.4 (70.71)	228.25 (136.72)	4	3	2	3
Livingston 1	Marsh	5	5	0.91 (0.12)	0.81 (0.25)	362.6 (234.09)	312.00 (117.75)	4	1	5	2
Livingston 2	Marsh	1	1	1.00 (NA)	1.01 (NA)	424.00 (NA)	243.00 (NA)	2	1	2	2
Livingston 3	Marsh	1	1	1.00 (NA)	1.03 (NA)	545.00 (NA)	90.00 (NA)	2	0	1	2

Table 1.3 (cont'd)

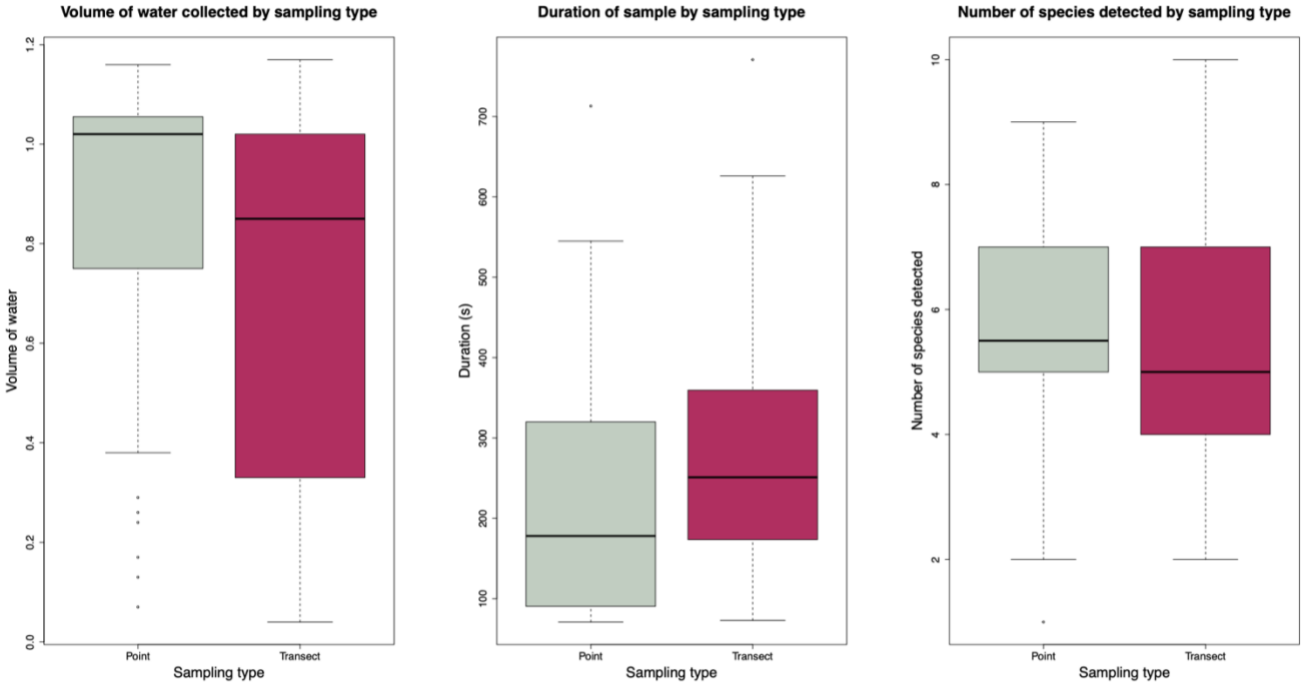
B.

		Transect Sampling									
		No. Samples		Water volume (L)		Duration of sampling (sec)		No. of species (12S gene)		No. of species (16S gene)	
		Period 1	Period 2	Period 1	Period 2	Period 1	Period 2	Period 1	Period 2	Period 1	Period 2
Site	Wetland Type			Mean (SD)		Mean (SD)					
Barry 1	Vernal pool	5	5	0.63 (0.15)	1.06 (0.08)	NA	279.33 (208.10)	4	3	6	2
Kalamazoo 1	Fen	5	5	0.98 (0.12)	0.83 (0.36)	NA	287.00 (179.32)	3	4	3	2
Kalamazoo 2	Marsh	4	3	0.81 (0.24)	0.75 (0.44)	NA	308.33 (137.00)	7	0	3	0
Jackson 1	Marsh	5	5	0.61 (0.41)	0.86 (0.28)	NA	347.60 (221.47)	4	4	3	3
Hillsdale	Fen	4	5	0.60 (0.50)	0.78 (0.31)	NA	188.25 (132.04)	2	1	1	1
Oakland	Marsh	3	3	1.04 (0.03)	1.03 (0.06)	NA	240.00 (174.23)	2	0	2	1
Jackson 1	Fen	5	3	0.31 (0.20)	0.71 (0.51)	NA	113.00 (45.64)	4	0	4	1
Barry 2	Vernal pool	3	3	0.70 (0.43)	0.51 (0.43)	NA	275.00 (63.38)	5	4	6	4
Clinton	Marsh	5	4	0.50 (0.40)	0.55 (0.34)	244.50 (81.13)	263.00 (93.32)	3	3	3	1
Livingston 1	Marsh	5	5	0.89 (0.19)	0.56 (0.41)	422.40 (268.91)	280.40 (66.4)	3	1	5	3
Livingston 2	Marsh	1	2	0.75 (NA)	0.66 (NA)	275.00 (NA)	324.00 (NA)	0	3	1	2
Livingston 3	Marsh	1	1	0.37 (NA)	1.05 (NA)	NA (NA)	140.00 (NA)	2	1	2	1

Comparisons of Sample Collection Methodology

When comparing the type of sampling used (point or transect) using a fully parameterized ANOVA, we detected no significant difference in species richness ($p = 0.52$, MS = 0.44, F-value = 0.41; Figure 1.2). There was a significant difference in the water collection duration ($p = 0.04$, F-value = 4.19) and total volume of water filtered ($p = 0.0005$, F-value = 1.25) for point and transect samples (Figure 1.2). On average, water collected using transects took longer to filter than point samples (253.83 ± 0.27 s versus 179.86 ± 0.13 s) and sampled a smaller volume of water (0.75 ± 0.25L versus 0.93 ± 0.15L) prior to filter clogging. Mean water volumes and durations can be found in Table 1.3 for point and transect samples.

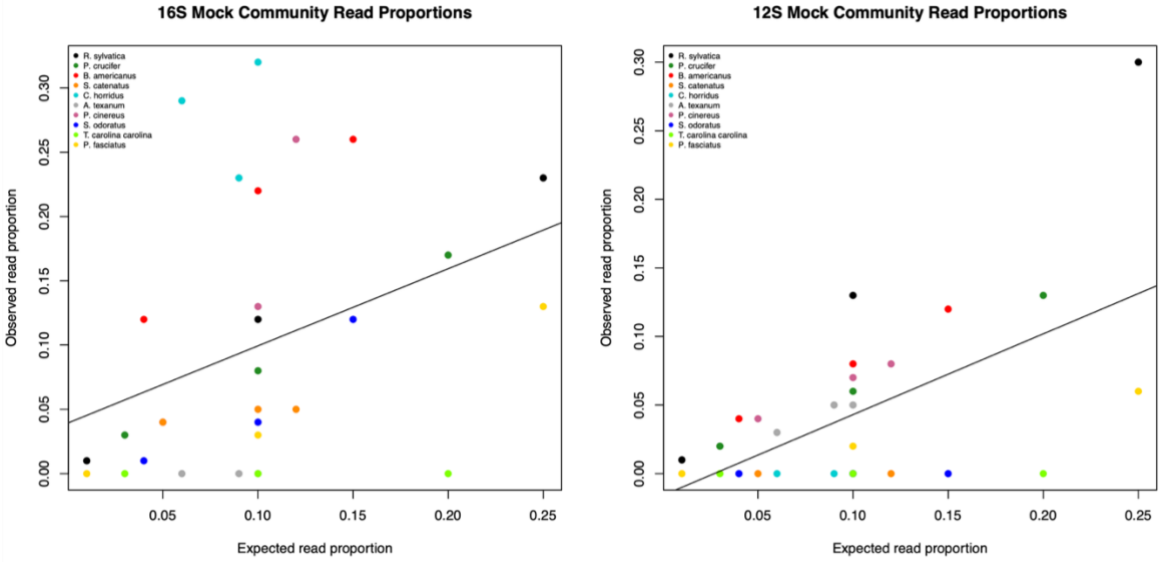
Figure 1.2. Boxplots comparing sampling type (point and transect) and their relationship to volume of water collected while sampling, the duration of a sample, and the number of species detected by each sampling type.



Comparisons Between 12S and 16S Markers

Using a paired t-test on sample-level data, we determined there was no significant difference in the number of species detected per sample between the two loci (12S: $\mu = 1.57, \pm 1.23$; 16S: $\mu = 1.58, \pm 1.25$; $t = -0.05$, $p = 0.96$). We did observe significant correlation between relative genomic DNA concentrations and sequence read counts across three mock communities analyzed for 12S and 16S (12S: $p = 0.02$, $t = 2.36$, correlation = 0.39; 16S: $p = 0.01$, $t = 2.66$, correlation = 0.43). However, for the 12S marker the snake and turtle species chosen for the mock community were not detected (Figure 1.3). In side-by-side comparisons between markers, interspecific sequence differences suggested improved taxonomic resolution was achieved using the 16S metabarcoding marker. Using a threshold sequence divergence of 2 base pairs for reliable assignment to the species level, 16S had three indistinguishable species pairs. The 12S region could not distinguish 8 species pairs from one another across all species in our sequence baseline database.

Figure 1.3. Mock community read proportions for 12S and 16S markers for ten species found in Great Lakes region wetlands. Expected read proportions are based on the known relative concentration of gDNA for each species in the three mock community samples.



Comparisons Between Seasonal Samples and Wetland Types

ANOVA results showed significant differences in species richness characterizing sample collections made early and later in the summer (MS = 33.06, F-value = 30.51, $p < 7e\text{-}7$; Figure 1.4) and samples collected in different wetland types (MS = 52.56, F-value = 48.51, $p < 2e\text{-}16$). Three wetland types were included in the ten sites sampled (marsh, vernal pool, and fen). The mean number of species detected for fens was 2.67 (± 1.53), 3.8 (± 1.10) for marshes, and 6.5 (± 2.12) for vernal pools. On average, samples collected earlier in the summer detected more species than those collected later in the summer (early: $\mu = 5.6 \pm 2.17$ species; late: $\mu = 3.50 \pm 1.65$ species). Using the 12S marker, more amphibians and reptiles were detected in the first sampling period versus the second (Table 1.4). Furthermore, several amphibian species such as *Pseudacris crucifer crucifer*, *Acris crepitans*, and *Notophthalmus viridescens* were detected only in the first sampling period. For the 16S marker, more reptiles were detected in the second period while the number of amphibians detected was approximately equal across sampling seasons. Additionally, no species was exclusively detected in one sampling period compared to the other for the 16S marker.

Figure 1.4. A boxplot depicting the number of species detected (species richness) by sampling period or “season”.

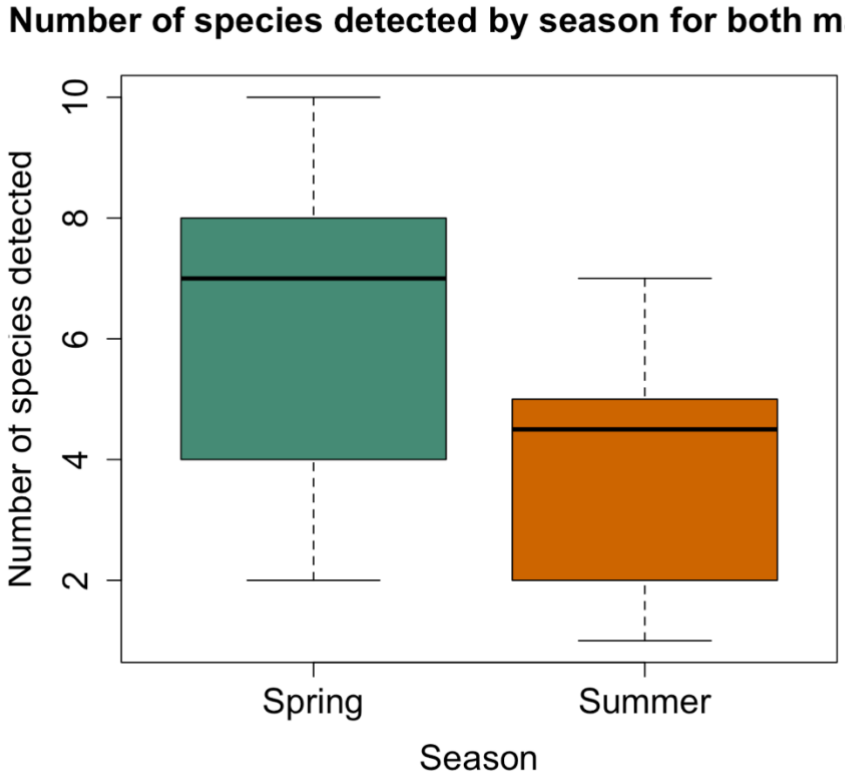


Table 1.4. The number of species found at each of 10 Michigan sampling sites. The number of amphibian and reptile species is listed for each site and marker.

			No. of species (12S gene)						No. of species (16S gene)					
Site (County)	**Wetland Type**	**Shannon Diversity**	**Period 1 Total**	**Period 2 Total**	**Period 1 Amphibians**	**Period 2 Amphibians**	**Period 1 Reptiles**	**Period 2 Reptiles**	**Period 1 Total**	**Period 2 Total**	**Period 1 Amphibians**	**Period 2 Amphibians**	**Period 1 Reptiles**	**Period 2 Reptiles**
Barry 1	Vernal pool	1.078	4	3	4	3	0	0	6	3	4	2	2	1
Kalamazoo 1	Fen	1.466	3	5	1	3	2	2	3	4	1	3	2	1
Kalamazoo 2	Marsh	1.093	7	0	5	0	2	0	3	1	3	1	0	0
Jackson 1	Marsh	0.576	4	4	2	2	2	2	4	3	2	2	2	1
Hillsdale	Fen	0.471	0	1	0	1	0	0	1	1	1	1	0	0
Oakland	Marsh	0.782	2	1	1	1	1	0	2	2	2	2	0	0
Jackson 1	Fen	0.613	4	1	2	1	2	0	5	3	3	1	2	2
Barry 2	Vernal pool	0.896	5	4	3	2	2	2	6	7	0	5	0	2
Clinton	Marsh	1.024	4	3	2	1	2	2	3	3	2	2	1	1
Livingston 1	Marsh	1.051	4	1	2	1	2	0	5	3	3	2	2	1
Livingston 2	Marsh	-	2	1	2	1	0	0	2	2	2	2	0	0
Livingston 3	Marsh	-	3	1	1	1	2	0	2	2	2	2	2	0

DISCUSSION

The methodological optimization and subsequent recommendations provide a framework for future eDNA metabarcoding analyses and wetland herpetofauna studies. The future surveys that this chapter facilitates are important for improving our understanding of the environmental factors and wetland features that may be contributing to declines in distribution and relative abundance for many reptile and amphibian species. Point sampling maximized water volume in this study, which could also lead to greater eDNA concentrations and a higher detection probability for targeted species. Our evaluation of two potential eDNA metabarcoding markers, both designed to universally amplify vertebrates (Deagle et al., 2009; Riaz et al., 2011), favored the 16S locus. While there was no difference in species richness between markers, 16S provided increased taxonomic resolution at the species level, with fewer indistinguishable species pairs based on pairwise sequence differences among taxa in the reference database. In addition, when conducting mock community trials, there was a notable absence of sequence reads for turtle and snake species in the 12S dataset.

Point and Transect Sampling within Wetland Systems

This study determined that point sampling was an effective method for collecting eDNA samples within wetlands using the Smith-Root ANDe eDNA backpack filtering system. This question can be considered specific to using the Smith-Root backpack apparatus, but there are alternative eDNA sampling options such collection of surface water in sterilized plastic bottles. We note that our results comparing point and transect sampling may not apply to alternative collection methods. Herpetofauna community diversity did not vary greatly as a function of the physical sampling approach applied. Our working hypothesis was that by sampling a larger area (5 m), the transect method would result in higher species richness. However, the number of species

detected in point and transect samples did not differ. Duration of filtration was longer for transect sampling than for point sampling. Further, smaller volumes of water were filtered for transect samples than point samples on a filter-by-filter basis (Table 1.3) prior to the filter clogging. While 1L of water was the initial desired volume per sample, many samples required additional filters to be used to reach the full 1L of water, or the desired volume was not reached. It is possible that, if the same volume of water was filtered for both point and transect methods, we may have seen a difference in the number of species detected for a sample taken over a larger (5 m transect) area.

There are several factors that contributed to duration and volume differences between point and transect samples. The first is physical characteristics of the system that was sampled. Wetlands are often heavily vegetated and characterized by shallow water depths. In warmer months (later sampling period), or in locations experiencing a lack of precipitation, accessible water is difficult to sample. Wetlands contain large amounts of loose sediments or materials that can be easily disturbed. Filters used during transect sampling were often covered in sediment or small pieces of aquatic vegetation that prevented water from moving through the metal screen and filter. Access to wetland systems without risking user contamination is also more difficult when using transect sampling relative to point sampling. Additionally, maintaining a consistent arc motion across a 5 m transect was challenging and introduced a higher chance of sediment fouling the filter. For these reasons and given the lack of significant differences in species richness for point and transect samples, we recommend a point sampling approach when collecting water for eDNA metabarcoding studies in wetlands. This study utilized the Smith-Root backpack filtering system and disposable filter units, but alternative approaches (e.g., water

collection in sterile bottles with subsequent filtering or centrifugation in the lab) could be implemented in diverse wetland systems.

Metabarcoding Marker Comparisons

The species used in the mock communities (artificial mixtures of extracted genomic DNA) were chosen to represent species expected to be encountered in herpetofauna communities in the Great Lakes region (Harding & Mifsud, 2017). Taxa included common and threatened species that inhabit different wetlands to determine if our sampling would be able to detect them in the field and if detected, to assess the potential for PCR amplification bias. Each of three mock communities was evaluated to account for variation in relative concentration of DNA across species (Schloss et al., 2011) and for amplification bias. For example, PCR amplification bias could be more consequential for estimates of species presence/absence or diversity when DNA is present in low abundance (simulated by low relative DNA concentration). When genomic DNA from eDNA samples was amplified and sequenced, we were able to compare the known relative concentration of genomic DNA from the mixtures with the relative sequence abundances from eDNA metabarcoding. In mock community trials for the 12S marker, species from the families Emydidae, Kinosternidae, and Viperidae were not detected, though known amounts of genomic DNA were included in the samples (representing up to 30% of the mock community). It is noted that detecting reptilian DNA from samples taken from the environment has proven difficult in past studies (Adams et al., 2019; Kucherenko et al., 2018), especially in comparison to amphibians (Raemy & Ursenbacher, 2018). This may provide an explanation for our higher amphibian detections in comparison to reptiles for both the 12S and 16S markers used for this project.

The use of multiple markers may increase sensitivity for single species and metabarcoding eDNA projects (Brys et al., 2023; Stefanni, 2018). Our results indicate that of the two loci evaluated, the 16S locus provides the ability to survey heterogeneous herpetofauna communities. Comparisons among sequences in our reference databases suggest that 16S could identify a greater proportion of taxa to the species level than could 12S. Using the 12S marker, 8 species pairings were indistinguishable (1 or fewer base pairs differences) versus only 3 indistinguishable species pairs for 16S. Due to the metabarcoding and community-focused aspect of our study, similar to previous metabarcoding studies (Andruszkiewicz et al., 2017; Deiner et al., 2017; Kelly et al., 2014), species-level identification is critical to properly characterizing the sampling sites and habitat characteristics shared by a suite of species. While we did not detect statistically significant differences between the number of species detected by the two markers, these key differences in taxonomic resolution and representation in our mock community data lead us to favor the 16S locus. Other studies comparing the 12S and 16S markers in fish have found similar results in the identification of target species (Morey et al., 2020).

Number of Species Detected Between Sampling Periods and Between Wetland Types

Timing of sampling is important for detecting wetland-obligate reptile and amphibian species. Obligatory wetland occupancy is often associated with a particular life stage, typically reproduction and early (egg and larval) life stages, the timing of which is difficult to anticipate. We sampled twice in a field season, once in early summer and once in late summer, to attempt to detect as many herpetofauna species as possible when considering differing life histories. For many species, visual or auditory detection can be difficult, but residual eDNA could be available for collection beyond the period of wetland occupancy. However, the rate of eDNA degradation is expected to vary across habitats due to a variety of environmental factors, such as UV

exposure, pH, and water temperature, among others (Strickler et al., 2015). Based on our findings, it is important to sample across the entire summer. To properly capture the full diversity of a wetland (particularly when sampling with 12S), it is recommended to sample during both early and late summer. An organism's biology also plays an important role in deciding when to sample for eDNA. Genera such as the Anurans could be found more often in the early summer/spring sampling seasons when vernal pools are more likely to be present and larvae or juveniles are utilizing these pools. This approach to wetland sampling may not be achievable for all studies, which is why the consideration of the target species' biology is critical to guiding sampling efforts. We noticed a marginal difference in number of species detected between wetland types sampled. Species richness in vernal pools was higher, which may relate to the ephemeral and smaller nature of vernal pools and the obligate amphibian use during breeding periods (Gibbs, 1993; Semlitsch & Bodie, 1998), in comparison to a marsh or fen which may be larger and/or more permanent staples of the wetland system.

CONCLUSIONS AND REFLECTIONS

Our study has empirically evaluated the efficacy of alternative eDNA sampling methods for reptiles and amphibians between time periods and across several wetland types in the Great Lakes region. Of the two mtDNA gene regions evaluated, the 16S rDNA locus appears best suited to identify herpetofauna taxa at the species level and to identify a greater taxonomic breadth of species. Due to specific ecological difficulties in timing and life stage of occupancy, multiple sampling periods may be required to accurately characterize the entire herpetofauna community in wetlands. Detections of amphibians during early sampling periods and their absence in samples taken in later periods indicate if a robust survey is required for research purposes, it is recommended to sample across multiple sampling periods. Environmental DNA

metabarcoding provides a cost and time-efficient means for characterizing wetland herpetofauna communities. It is critical to survey these communities to better understand patterns of habitat use and species distributions, given the consistent and pertinent threats posed by wetland degradation and fragmentation. The optimization of these methods contributed to the understanding of environmental DNA sampling in wetland systems and will inform future research applications over a more geographically expansive and heterogeneous landscape.

When reflecting on what we did not include in this study, we arrived at three major conclusions. The first is the further exploration of sampling techniques used for this project. While we focused on eDNA, an interesting component that could have been introduced is pairing eDNA metabarcoding alongside traditional sampling methods to compare both sampling efficacy and species detected among methods. Second, the interrogation into different spatial sampling techniques would have been useful to identify how to sample in these diverse wetland systems when considering structural complexity, eDNA distribution, and species wetland use. Third, further examination into amplification biases for amphibians and reptiles would have been useful to better quantify the performance of 12S and 16S mtDNA markers for future research, as both have their respective characteristics that may be best suited for a more taxonomic group-specific research project.

LITERATURE CITED

Adams, C.I., Hoekstra, L.A., Muell, M.R., & Janzen, F.J. 2019. A Brief Review of Non-Avian Reptile Environmental DNA (eDNA), with a Case Study of Painted Turtle (*Chrysemys picta*) eDNA Under Field Conditions. *Diversity, 11*(4), 50. doi:10.3390/d11040050

Andruszkiewicz, E.A., Starks, H.A., Chavez, F.P., Sassoubre, L.M., Block, B.A., & Boehm, A.B. 2017. Biomonitoring of marine vertebrates in Monterey Bay using eDNA metabarcoding. *PLOS ONE*, *12*(4), e0176343. doi:10.1371/journal.pone.0176343

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Yu, D.W., de Bruyn, M. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution,* 29, 358–367. doi:10.1016/j.tree.2014.04.003

Brys, R., Halfmaerten, D., Everts, T., Driessche, C.V., & Neyrinck, S. 2023. Combining multiple markers significantly increases the sensitivity and precision of eDNA-based single-species analyses. *Environmental DNA*. doi:10.1002/edn3.420

Charif, D. and Lobry, J.R. 2007. R package: seqinr.

Čeirāns, A., Pupina, A., & Pupins, M. 2020. A new method for the estimation of minimum adult frog density from a large-scale audial survey. *Scientific Reports*, *10*(1), 1-12. doi:10.1038/s41598-020-65560-6

Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., Iribar, A., Taberlet, P., Vigouroux, R., Brosse, S. 2019. Unlocking biodiversity and conservation studies in high-diversity environments using environmental DNA (eDNA): A test with Guianese freshwater fishes. *Molecular Ecology Resources,* 19, 27–46. doi:10.1111/1755-0998.12900

Crouch, W.B., & Paton, P.W.C. 2002. Assessing the Use of Call Surveys to Monitor Breeding Anurans in Rhode Island. *Journal of Herpetology*, *36*(2), 185–192. doi:10.2307/1565990

Dahl, T.E. 2000. Status and trends of wetlands in the conterminous United States 1986 to 1997. U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C. 82 pp

Davidson, N.A. 2014. How much wetland has the world lost? Long-term and recent trends in global wetland area. *Marine and Freshwater Research*, 2014, 65, 934–941 doi:10.1071/MF14173

Davis, C.L., Miller, D.A.W., Walls, S.C., Barichivich, W.J., Riley, J.W., & Brown, M.E. 2017. Species interactions and the effects of climate variability on a wetland amphibian metacommunity. *Ecological Applications, 27*(1), 285–296. http://www.jstor.org/stable/44132599

Deagle, B.E., Kirkwood, R., & Jarman, S.N. 2009. Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular Ecology*, *18*(9), 2022– 2038.

Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., de Vere, N., Pfrender, M.E., Bernatchez, L., 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, 26, 5872–5895. doi:10.1111/mec.14350

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

Dickie, I.A., Boyer, S., Buckley, H.L., Duncan, R.P., Gardner, P.P., Hogg, I.D., Holdaway, R.J., Lear, G., Makiola, A., Morales, S.E., Powell, J.R., Weaver, L. 2018. Towards robust and repeatable sampling methods in eDNA-based studies. *Molecular Ecology Resources*, 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., Jerde, C.L., Lodge, D.M., Lamberti, G.A., Pfrender, M.E. 2017. Fish community assessment with eDNA metabarcoding: Effects of sampling design and bioinformatic filtering. *Canadian Journal of Fisheries and Aquatic Sciences*, 74, 1362–1374. doi:10.1139/cjfas-2016-0306

Gibbons, J.W. 2003. Terrestrial habitat: A vital component for herpetofauna of isolated wetlands. *Wetlands*, 23, 630–635. doi:10.1672/0277-5212(2003)023[0630:THAVCF]2.0.CO;2

Gibbs, J.P. 1993. Importance of small wetlands for the persistence of local populations of wetland associated animals. *Wetlands*, 13, 25–31. doi:10.1007/BF03160862

Harding, J.H. and Misfud, D.A. 2017. Amphibians and reptiles of the Great Lakes region. Revised.

Hunter, M.E., Oyler-McCance, S.J., Dorazio, R.M., Fike, J.A., Smith, B.J., Hunter, C.T., Reed, R.N., & Hart, K.M. 2015. Environmental DNA (eDNA) Sampling improves occurrence and detection estimates of invasive Burmese pythons. *PLOS ONE*, *10*(4), e0121655. doi:10.1371/journal.pone.0121655

Kelly, R.P., Port, J.A., Yamahara, K.M., & Crowder, L.B. 2014. Using environmental DNA to census marine fishes in a large mesocosm. *PLOS ONE*, *9*(1), e86175. doi:10.1371/journal.pone.0086175

Korfel, C.A., Mitsch, W.J., Hetherington, T.E., & Mack, J.J. 2010. Hydrology, physiochemistry, and amphibians in natural and created vernal pool wetlands. *Restoration Ecology*, *18*(6), 843-854. doi:10.1111/j.1526-100X.2008.00510.x

Kucherenko, A., Herman, J.E., Iii, E. M.E., & Urakawa, H. 2018. Terrestrial snake environmental DNA accumulation and degradation dynamics and its environmental application. *Herpetologica,* 74, 38– 49. doi:10.1655/Herpetologica-D-16-00088

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

Lotz, A., & Allen, C.R. 2007. Observer bias in Anuran call surveys. USGS Staff -- Published Research. 13. https://digitalcommons.unl.edu/usgsstaffpub/13

Mazerolle, M.J., Desrochers A., Rochefort L. 2005. Landscape characteristics influence pond occupancy by frogs after accounting for detectability. *Ecological Applications*, 15, 824–834. doi: 10.1890/04-0502

McKee, A.M., Calhoun, D.L., Barichivich, W.J., Spear, S.F., Goldberg, C.S., Glenn, T.C. 2015. Assessment of environmental DNA for detecting presence of imperiled aquatic amphibian species in isolated wetlands. *Journal of Fish and Wildlife Management, 6*(2): 498–510.

Morey, K.C., Bartley, T.J., & Hanner, R.H. 2020. Validating environmental DNA metabarcoding for marine fishes in diverse ecosystems using a public aquarium. *Environmental DNA, 2*(3), 330-342. doi:10.1002/edn3.76

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., and Wagner, H. 2020. vegan: Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Paradis E. & Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35: 526-528.

Piaggio, A.J., Engeman, R.M., Hopken, M.W., Humphrey, J.S., Keacher, K.L., Bruce, W.E., Avery, M.L., 2014. Detecting an elusive invasive species: A diagnostic PCR to detect Burmese python in Florida waters and an assessment of persistence of environmental DNA. *Molecular Ecology Resources*, 14, 374–380. doi:10.1111/1755-0998.12180

Prince, A.M., Andrus, L., 1992. PCR: How to kill unwanted DNA. Biotechniques 12, 358–360.

Pukk, L., Kanefsky, J., Heathman, A.L., Weise, E.M., Nathan, L.R., Herbst, S.J., Robinson, J.D. 2021. eDNA metabarcoding in lakes to quantify influences of landscape features and human activity on aquatic invasive species prevalence and fish community diversity. *Diversity and Distributions*, 27, 2016–2031. doi:10.1111/ddi.13370

Raemy, M., & Ursenbacher, S. 2018. Detection of the European pond turtle (Emys orbicularis) by environmental DNA: Is eDNA adequate for reptiles? *Amphibia-Reptilia*, 39, 135–143. doi:10.1163/15685381-17000025

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., Coissac, E., 2011. EcoPrimers: Inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, 39, 1–11. doi:10.1093/nar/gkr732

Saenz-Agudelo, P., Delrieu-Trottin, E., DiBattista, J. D., Martínez-Rincon, D., Morales-González, S., Pontigo, F., Ramírez, P., Silva, A., Soto, M., & Correa, C. 2022. Monitoring vertebrate biodiversity of a protected coastal wetland using eDNA metabarcoding. *Environmental DNA*, *4*(1), 77-92. doi:10.1002/edn3.200

Sard, N., Scribner, K.T., Robinson, J.D., Herbst, S., Kanefsky, J. 2019. Comparison of fish detections, community diversity and relative abundance using environmental DNA metabarcoding and traditional gears. *Environmental DNA*, 1– 31. doi:10.1111/j.1365-2958.2003.03935.x

Sayers, E.W., Cavanaugh, M., Clark, K., Ostell, J., & Pruitt, K.D. 2020. GenBank. *Nucleic Acids Research*, *48*(D1), D84-D86. doi:10.1093/nar/gkz956

Schloss, P.D., Gevers, D., Westcott, S.L. 2011. Reducing the effects of PCR amplification and sequencing Artifacts on 16s rRNA-based studies. *PLOS ONE*, 6. doi:10.1371/journal.pone.0027310

Schloss, P.D. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology 75:7537–7541.

Schmidt B. R. 2005. Monitoring the distribution of pond-breeding amphibians when species are detected imperfectly. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 15, 681–692. doi: 10.1002/aqc.740

Semlitsch, R.D., & Bodie, J.R. 1998. Are Small, Isolated Wetlands Expendable? *Conservation Biology*, *12*(5), 1129-1133. doi:10.1046/j.1523-1739.1998.98166.x

Sierszen, M., Morrice, J., Trebitz, A., Hoffman, J. 2012. A review of selected ecosystem services provided by coastal wetlands of the Laurentian Great Lakes. *Aquatic Ecosystem Health & Management*, 15, 92-106. doi:10.1080/14634988.2011.624970.

Sierszen, M., Morrice, J., Trebitz, A., Hoffman, J. 2012. A review of selected ecosystem services provided by coastal wetlands of the Laurentian Great Lakes. *Aquatic Ecosystem Health & Management*, 15, 92-106. doi:10.1080/14634988.2011.624970.

Skidds, D.E., Golet, F.C. 2005. Estimating hydroperiod suitability for breeding amphibians in southern Rhode Island seasonal forest ponds. *Wetlands Ecology and Management*, 13, 349–366. doi:10.1007/s11273-004-7527-4

Stefanni, S., Stanković, D., Borme, D., de Olazabal, A., Juretić, T., Pallavicini, A., & Tirelli, V. 2018. Multi-marker metabarcoding approach to study mesozooplankton at basin scale. *Scientific Reports*, *8*(1), 1-13. doi:10.1038/s41598-018-30157-7

Strickler, K.M., Fremier, A.K., & Goldberg, C.S. 2015. Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biological Conservation*. 183, 85-92. doi:10.1016/j.biocon.2014.11.038

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., Willerslev, E. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045-2050. doi:10.1111/j.1365-294X.2012.05470.x

Tamura, K., Stecher, G., Kumar, S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Molecular Biology and Evolution*, 38:3022-3027

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

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G. H., Geniez, P., Pont, D., Argillier, C., Baudoin, M., . . . Dejean, T. 2016. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, *25*(4), 929-942. doi:10.1111/mec.13428

Wikston, M., Breton, B., Vilaça, S. T., Bennett, A. M., Kyle, C. J., Beresford, D. V., Lesbarrères, D., Wilson, C. C., Green, D. M., Fortin, M., & Murray, D. L. 2023. Comparative efficacy of eDNA and conventional methods for monitoring wetland anuran communities. *Frontiers in Ecology and Evolution*, *11*, 1179158. doi:10.3389/fevo.2023.1179158

CHAPTER 3: PATTERNS OF BIODIVERSITY ACROSS GREAT LAKES REGION WETLANDS AND HABITAT ASSOCIATIONS FOR HERPETOFAUNA OCCUPANCY

ABSTRACT

Information pertaining to relationships between habitat characteristics and species site occupancy and distribution is critical to guide conservation strategies for imperiled species. Little is known about the habitat features associated with species occupancy for many threatened or endangered reptiles and amphibians (herpetofauna) of the Great Lakes region. Species-specific environmental (e)DNA sequences collected from filtered water were used to detect species presence, quantify relative sequence abundance, and estimate measures of species diversity. eDNA detections from this project were used to assess the habitat features and environmental characteristics that contributed to site occupancy of herpetofauna. Joint application of eDNA metabarcoding data and generalized linear modeling identified habitat characteristics predictive of wetland herpetofauna biodiversity such as the number of wetland types sampled and the percentage of developed landcover, but coefficient estimates from these models were not significant. Occupancy modeling results favored models relating occupancy to the number of wetland types sampled (*Thamnophis unclassified*: Bayesian estimate coefficient = 0.04) and the degree of development surrounding wetlands (*Ambystoma texanum*: coefficient = 0.20; *Emydoidea blandingii*: coefficient = -0.21), but credible intervals for model coefficients suggest uncertainty in the direction of effects for these covariate. Collectively, these findings on species occurrence and habitat requirements can inform wetland management and restoration strategies throughout the Great Lakes region by identifying habitat characteristics that best support threatened and endangered herpetofauna.

INTRODUCTION

Wetlands are crucial habitats across the midwestern United States, representing important water sources for obligate and facultatively aquatic species, including threatened and endangered reptiles and amphibians (jointly, herpetofauna; Gibbons, 2003). Wetlands are being degraded at a high rate (Dahl et al., 2000; Davidson, 2014), leading to declines in species abundance and distribution. While habitat loss within wetlands has been widely described (Kingsford et al., 2016; Quesnelle et al., 2013), management for resident species is confounded by the lack of information about distribution and habitat requirements for many herpetofauna species.

Recent studies suggest that non-invasive genetic sampling, such as environmental DNA (eDNA) surveys can improve efforts to survey wetland biodiversity (McKee et al., 2015; Saenz-Agudelo et al., 2002; Wikston et al., 2023) and have demonstrated higher sensitivity in detecting multiple species (Evans et al., 2017) in comparison to traditional sampling methods. Single-species eDNA surveys have been used to detect individual species' presence/absence in herpetofauna communities (Dejean et al., 2012; Piaggio et al., 2014). However, single-species studies are not able to characterize the species composition or diversity of the wetland herpetofauna community as a whole. Recently, eDNA metabarcoding, made possible by rapid advances in high-throughput sequencing (Deiner et al., 2017), has been developed to permit characterization of entire communities, including residents of aquatic habitats (Taberlet et al., 2014; Valentini et al., 2016). Environmental DNA metabarcoding provides an alternative to traditional sampling methods such as capture-based or call-based community surveys, with the potential for implementation in wetland systems. eDNA sampling shows great potential to improve multi-species detections and characterizations of species distributions (Bohmann et al., 2014; Thomsen and Willerslev, 2015).

Species distribution models (SDMs) and occupancy models (OMs) have also been useful in identifying habitat and land use patterns and species distribution. SDMs allow for the prediction of species presence at a site alongside the required habitat characteristics conducive to supporting a species (Franklin, 2013; Guisan and Thuiller, 2005). Occupancy models also identify key habitat characteristics and estimate species occurrence, while accounting for imperfect detection (Mackenzie and Royle, 2005). Using these two modeling techniques together as in Peterman et al. (2013), will provide information about both the habitat characteristics associated with site occupancy and areas where targeted species may persist.

In this study we combine eDNA metabarcoding data with SDMs and OMs. Together, these tools can contribute to the collection of data about target herpetofauna species in wetlands within the Great Lakes region. In addition to contributing to updated occurrence data, we can also learn about community composition and environmental factors that could influence habitat use (Muha et al., 2017). Our analyses considered the following questions: What habitat features are associated with wetland occupancy of threatened and endangered herpetofauna, and can eDNA be an effective tool to monitor these imperiled species? Specifically, objectives for this project were to conduct a survey of herpetofauna biodiversity across three states in the Great Lakes region, identify areas of high wetland biodiversity and the environmental features associated with biodiversity hotspots, and characterize influences of environmental variables on site occupancy for threatened and endangered species.

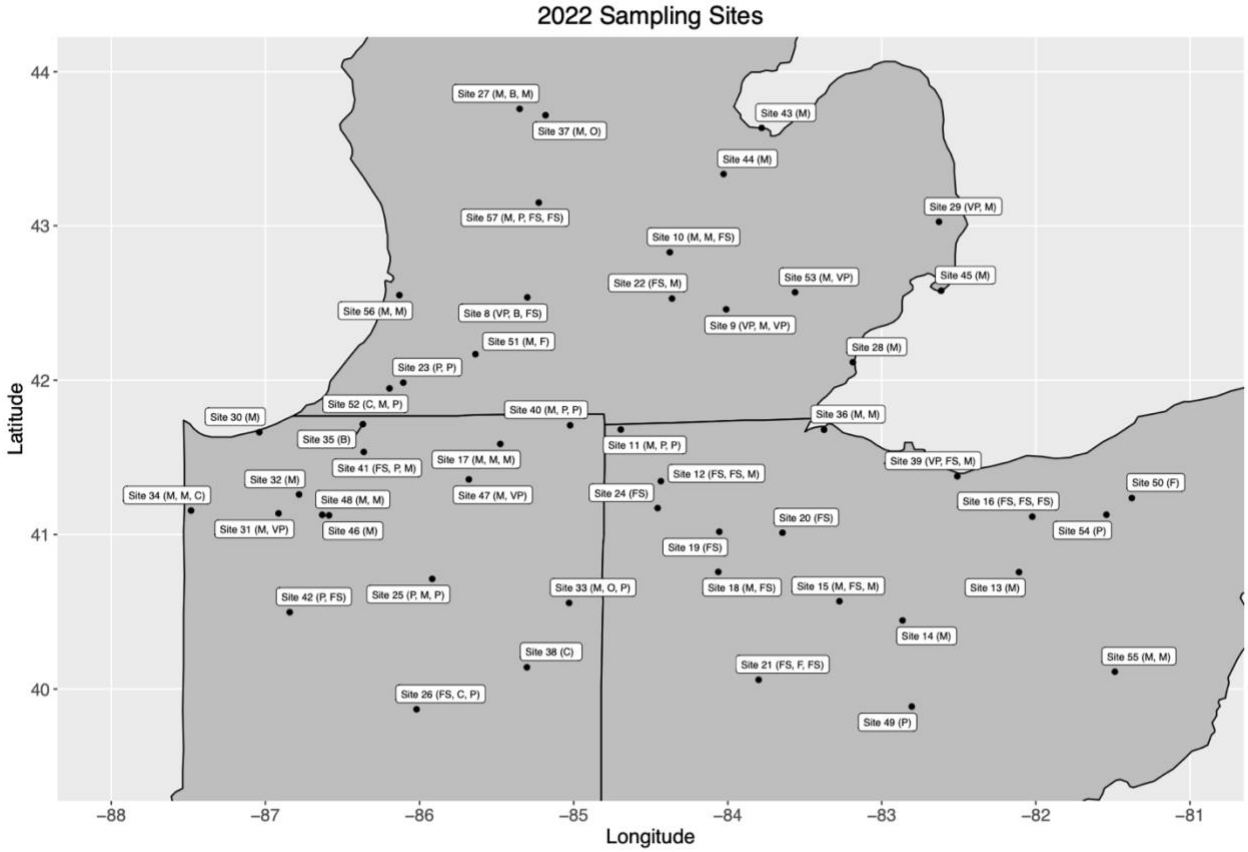
METHODS

Site Selection and Field Collections

Sampling sites were distributed across southern Michigan, northern Indiana, and northern Ohio (Figure 2.1), including a total of 50 sites and 104 wetlands. Sampling covered approximately

60,000 square miles across the study region. Site selection was guided by preliminary composite SDMs for 22 species of conservation concern (Table 2.1; Figure 2.2). SDMs were developed using occurrence records from online repositories GBIF (GBIF; gbif.org) and VertNet (VertNet; vertnet.org). Environmental variables for preliminary SDMs were sourced from WorldClim (Fick & Hijmans, 2017). The early-stage SDMs were developed using the R package *enmSdm* (Smith, 2022). The climate raster that described the model environment used environmental variables sourced from WorldClim (Fick & Hijmans, 2017). Random background sites were selected from our pre-defined study area to serve in place of species absences. For each species, we used the *maxnet* R package (Phillips et al., 2023) to construct SDMs using the MaxEnt algorithm (Phillips et al. 2006) with auto-selected predictors. Records were filtered to exclude occurrences that were outside of our study region and to retain only occurrences reported between 1970-2020. In conjunction with stakeholder engagement, early-stage SDMs were used to identify and prioritize wetland sites for sampling. We engaged stakeholders from numerous organizations across the study region including state, federal, and tribal agencies, county and local park systems, and various non-profits. All sampling that occurred in 2022 followed optimized protocols from previous work.

Figure 2.1. Year 2 sampling sites and wetlands sampled at each site[1] across Michigan, Ohio, and Indiana.

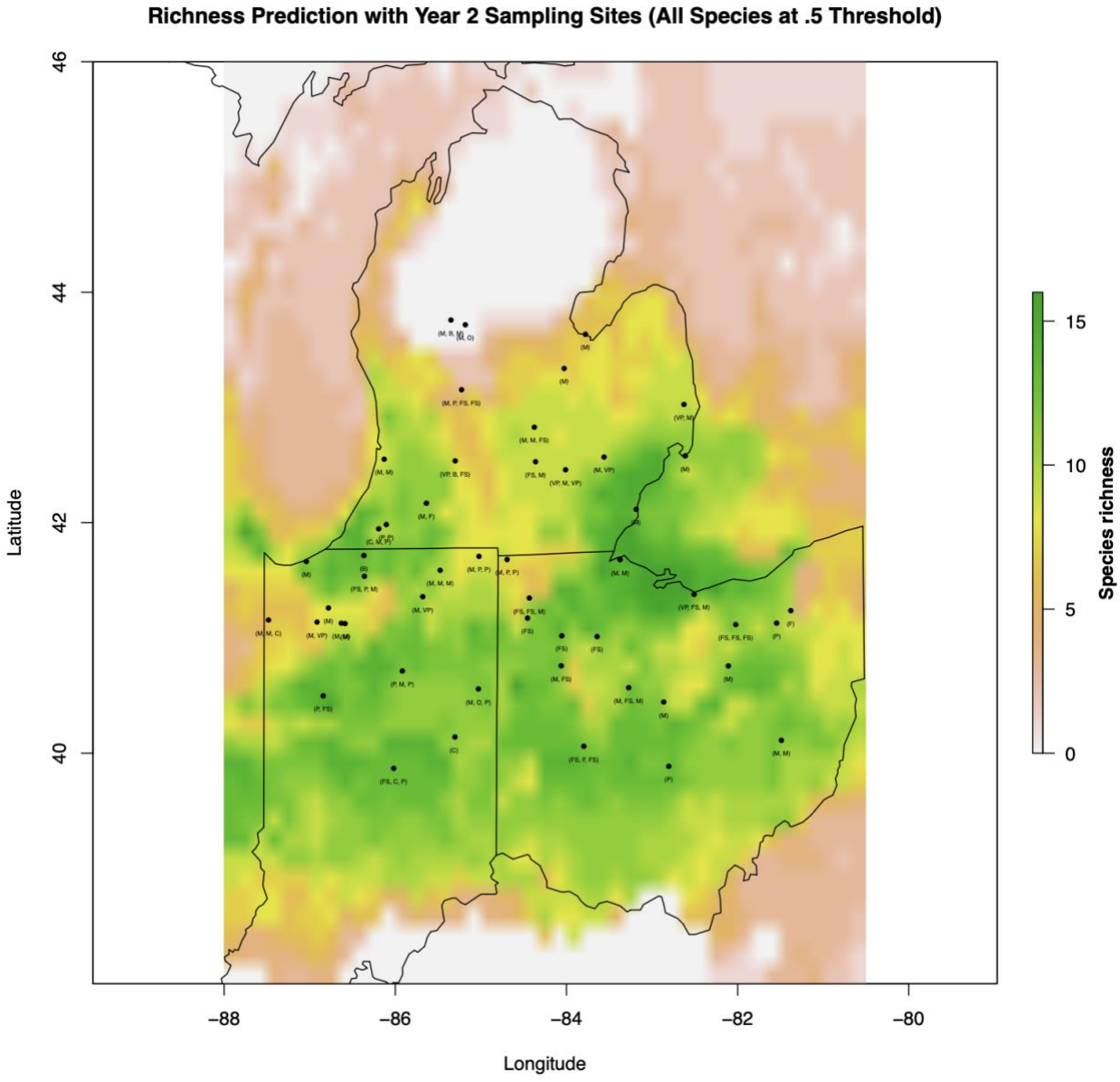


[1] VP = vernal pool, M = marsh, FS = forested shrub, F = fen, P = pool/pond, C = creek/drain, B = bog, O = other

Table 2.1. Species of conservation concern targeted in wetlands in the Great Lakes region. Species in bold font were detected during the biodiversity assessment conducted in Michigan, Ohio, and Indiana in Spring/Summer 2022.

Common name	Scientific name
Blanchard's cricket frog	*Acris crepitans blanchardi*
Marbled salamander	*Ambystoma opacum*
Smallmouth salamander	***Ambystoma texanum***
Midland smooth softshell	*Apalone mutica*
Spotted turtle	*Clemmys guttata*
Kirtland's snake	*Clonophis kirtlandii*
Northern ring-necked snake	*Diadophis punctatus edwardsii*
Blanding's turtle	***Emydoidea blandingii***
Northern map turtle	*Graptemys geographica*
Copperbelly water snake	*Nerodia erythrogaster neglecta*
Rough green snake	*Opheodrys aestivus aestivus*
Smooth green snake	*Opheodrys vernalis*
Eastern fox snake	*Pantherophis gloydi*
Gray rat snake	*Pantherophis spiloides*
Pickerel frog	***Rana palustris***
Queen snake	*Regina septemvittata*
Western lesser siren	*Siren intermedia nettingi*
Eastern massasauga	*Sistrurus catenatus*
Eastern musk turtle	*Sternotherus odoratus*
Eastern box turtle	*Terrapene carolina carolina*
Ornate box turtle	*Terrapene ornata ornata*
Butler's garter snake	*Thamnophis butleri*
Plains garter snake	***Thamnophis radix***
Northern ribbon snake	*Thamnophis sauritus septentrionalis*

Figure 2.2. A cumulative species richness prediction plot using 20 individual species distribution models for species of conservation concern (See Table 1). Black points on the map represent 50 sampling sites surveyed in 2022. Labels indicated wetlands sampled at each site.



We attempted to collect six point samples (see Chapter 2) per site, during each of two time periods. Samples were collected from multiple wetland types per site, when possible. The following types of wetlands were sampled over the course of the 2022 field season: marsh, fen, bog, vernal pool, forested shrub/scrub, pond, creek/drain, and other (Appendix, Table 1). The first sampling period occurred between May 5th and June 6th, 2022. The second sampling period occurred between July 8th and July 30th, 2022. One liter water samples were collected and filtered in the field using single-use, sterile filter packs (polyethersulfone (PES) filters, 1.2 µm porosity) and a Smith-Root ANDe eDNA backpack filtering device (Smith-Root, Vancouver, WA). Because shallow water levels and sedimentation often affected filtering efficiency (Chapter 2), the duration of filtering and total volume filtered were both recorded. After filtering, the filter was removed from the housing using antiseptic techniques including latex gloves and sterile forceps, and placed in 95% ethanol in 5 ml tubes for preservation and transportation to the laboratory. All equipment used for sampling was single-use (filters) or decontaminated with 20% bleach solution following eDNA best sampling practices (Dickie et al., 2018; Prince & Andrus, 1992). One liter of distilled water was filtered at each site and sampling period for use as a negative (no DNA) field sampling control.

Several environmental variables were recorded at each site to serve as covariates in models relating wetland characteristics to site occupancy and/or measures of species diversity, including water temperature, air temperature, salinity, and pH. Water temperature, salinity, and pH were recorded using an ExStik® II pH/conductivity meter (EXTECH, Nashua, NH), while air temperature was recorded from local weather stations.

Filter Processing and DNA Extraction

After transportation to the laboratory, samples were stored in 95% ethanol at room temperature. For DNA extraction, filters were removed from ethanol, dried, and cut into multiple pieces. After drying, DNA extractions were performed for all cut filters using Qiagen Qia-shredder and DNeasy Blood and Tissue kits (QIAGEN, Venlo, NL) using protocols described in Laramie et al. (2015). Zymo One-Step columns (Zymo Research, Irvine, CA) were used to remove PCR inhibitors. DNA was eluted in a final volume of 100 µL in Buffer AE (10 mM Tris-Cl. 0.5 mM EDTA; pH 9.0) (QIAGEN, Venlo, NL) and stored at -20 °C for later processing.

PCR and High-Throughput Amplicon Sequencing

Following DNA extraction and inhibitor removal, the 16S rDNA region of the mitochondrial genome was amplified using published protocols (Pukk et al., 2021; Sard et al., 2019). PCR reagents were combined in a sterile tube in the following concentrations (25 µl total volume): 1X AmpliTaq Gold PCR Buffer II (no Mg^{2+}), 1 µg/µl of BSA, 1.25 U AmpliTaq Gold DNA polymerase, 0.32 mM dNTPs, 0.32 µM of the forward and reverse primer (Deagle et al., 2009), and 2.5 mM $MgCl_2$. The forward and reverse primers included CS1 and CS2 oligo tails (Fluidigm). Samples were incubated at 95° C for 10 minutes, followed by 35 cycles of 95° C for 30 seconds, 57° C for 45 seconds, and 72° C for 45 seconds, with a final elongation step at 72° C for 5 minutes. PCR products were cleaned using a Qiagen Qiaquick PCR Purification kit, and concentrations were determined using a NanoDrop 1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE). We performed two PCR replicate reactions for each sample. One PCR negative control, containing deionized water and all PCR reagents, was included on each plate.

Following amplification, samples were individually barcoded using i7 (plate) and i5 (sample) primers added to amplified PCR products. Barcoding reactions included the desired i7 index (10 μM) and 2X Qiagen Plus Master Mix (QIAGEN; #206152) for a reaction of 6 μl. The i5 indices (5 μM) were added to each well of the plate. For barcoding, sample plates were placed in the thermocycler and incubated at 95° C for 15 minutes, followed by 10 cycles of 95° C for 10 seconds, 65° C for 30 seconds, and 72° C for 30 seconds, with a final elongation step at 72° C for 5 minutes. Once both the i5 and i7 primers were adhered to each sample, allowing them to be distinguished from other samples, they were pooled for bead size selection by combining all samples from a plate into one 1.5 ml tube. Once samples were pooled, size selection was performed using AMPure XP beads (Beckman Coulter, Indianapolis, IN). To remove long sequences, we performed a 0.5X (1:2 ratio) reaction of beads to pooled library volume. Short sequences were removed using a 1.2X reaction of beads to pooled library volume. To quantify the concentration of each library, samples were analyzed using a Qubit fluorometer (Life Technologies, ThermoFisher Scientific, USA). Due to the high concentration of eluted DNA in each library, it was required to dilute each one to a 1:10 ratio. Once diluted, libraries were prepared for Qubit analysis. Diluted libraries were sent to be processed on the TapeStation (assay High Sensitivity D1000 ScreenTape®) to confirm sample concentrations and base-pair lengths. Once returned, fully eluted libraries were sequenced on the Illumina MiSeq platform (generating 150bp paired-end reads) in three lanes at Michigan State University's Research Technology Support Facility.

Data Analysis

Bioinformatic analyses of high-throughput sequencing data were conducted using a Mothur (ver. 1.39.5) pipeline (Schloss et al., 2009) developed for previous eDNA metabarcoding research in

our lab (e.g., Sard et al., 2019; Pukk et al., 2021). Raw sequences were demultiplexed and filtered for length and quality (minimum length = 35, maximum length = 95, maximum homopolymers = 13) using the *make.contigs* function from Mothur. This command extracted the sequence and quality score data from raw data files and read the complemented and paired sequences into the pipeline. Mothur was used to group similar sequences into operational taxonomic units (OTUs; 99% sequence homology pre-clustering threshold) and to classify OTUs to the lowest taxonomic level possible, based on comparisons to a reference 16S sequence database (see Chapter 2, https://github.com/rupperto1/Herp-eDNA-Taxonomic-Database-2020-2023). The taxonomic database was developed using sequence data from GenBank (Sayers et al., 2020) and sequences from tissues loaned from project collaborators (Chapter 2). The final database included 65 target species and contained 165 individual sequences for the 16S locus.

Once sequences were grouped into OTUs, they were used to create community matrices to identify species found in each sample. When controlling for contamination, the median number of sequence reads for detections in negative control samples was calculated and used as a threshold for detections in water samples for the project. Species detections that fell below the threshold were removed from the community matrices prior to further analysis. The R package *vegan* (Oksanen et al., 2020) was used to calculate alpha diversity measures and analyze herpetofauna communities across the surveyed wetlands.

Modeling Analyses

Generalized linear models were used to assess relationships between species richness of the wetland herpetofauna communities at each site and a variety of environmental variables, including National Land Cover Database (NLCD) measures of land use (Dewitz, 2021) within 1 km and mean annual temperature and precipitation between 1991 and 2020 (PRISM Climate

Group; https://prism.oregonstate.edu). Candidate models representing different combinations of these variables were fit to the data and model selection analyses used corrected Akaike Information Criterion (AICc; Akaike, 1974, Hurvich and Tsai, 1989) to evaluate relative support for each model.

Following bioinformatic analysis of eDNA metabarcoding data, occupancy models were used to identify habitat characteristics of wetlands and surrounding uplands that are associated with site occupancy for species of conservation concern that were detected in the eDNA metabarcoding survey. These species of concerns and their wetland use patterns can be found in Table 2.2.

Table 2.2. Species of conservation concern, their wetland use patterns, and the number of samples in which they were detected.

Common name	Terrestrial/facultative wetland use/obligate?	Sample detections
Blanchard's cricket frog	Wetland obligate	0
Marbled salamander	Wetland obligate	0
Smallmouth salamander	Wetland obligate	**17**
Midland smooth softshell	Terrestrial	0
Spotted turtle	Terrestrial	0
Kirtland's snake	Terrestrial	0
Northern ring-necked snake	Terrestrial	0
Blanding's turtle	Terrestrial	**1**
Northern map turtle	Wetland facultative	0
Copperbelly water snake	Wetland facultative	0
Rough green snake	Terrestrial	0
Smooth green snake	Terrestrial	0
Eastern fox snake	Terrestrial	0
Gray rat snake	Terrestrial	0
Pickerel frog	Wetland obligate	**3**
Queen snake	Terrestrial	0
Western lesser siren	Wetland obligate	0
Eastern massasauga	Terrestrial	0
Eastern musk turtle	Wetland facultative	0
Eastern box turtle	Terrestrial	0
Ornate box turtle	Terrestrial	0
Butler's garter snake	Terrestrial	0
Plains garter snake	Terrestrial	**23**
Northern ribbon snake	Terrestrial	0

Occupancy models were developed and analyzed using the R package *eDNAoccupancy* (Dorazio & Erickson, 2017). We created a list of 17 candidate models for each species of conservation concern detected in our eDNA survey (Table 2.3). Independent variable models included the percentage of forested land cover, the percentage developed land cover, the percentage of wetland land cover, canopy cover, fish predator presence, sampling season, mean annual precipitation and temperature, and the number of wetlands sampled. Land cover variables were considered based on previous literature identifying forested and developed landcover as crucial variables with differing effects on amphibian and reptile distributions, identifying these groups as vulnerable to land use change (Marsh et al., 2016; Martinuzzi et al., 2015). Dertien et al. (2020) identified positive relationships for multiple groups of organisms (including mammals, birds, amphibians, and reptiles) with wetland land cover using National Wetland Inventory classifications. The utilization of fish predator presence is based on literature that identifies highly ephemeral wetland use by amphibians due to lack of predation (Holbrook & Dorn, 2016) and avoidance of wetlands that support fish predators by adult amphibians (Binckley & Resetarits, 2003). Mean annual precipitation and temperature were chosen as general climactic variables that may have an effect on overall species richness in a given geographical area. We ran a null model (intercept only), a "global" model that included all variables represented in other candidate models, and multiple single-variable models. Additionally, we also included three models representing hypotheses about regional drivers of biodiversity: a "habitat" model, a "landcover" model, and a "climate" model. The habitat model included features that we hypothesized may have an influence on species occupancy including land cover within 1km of the sampled wetland (derived from NLCD classifications, percent forested, canopy cover, percent wetland) and the number of wetlands sampled at a site. Our land cover model includes

all the land cover variables sourced from NLCD including percentage developed, wetland, and forested (within 1km of the sampled wetland). Finally, our climate model included mean annual temperature, mean annual precipitation, and latitude. Due to lack of sample or occurrence data, not every model was run for all detected species. Each model was run for 20,000 iterations, with a burn-in of 1000 iterations. Convergence was assessed using trace plots of each model parameter (Appendix, Figures 1-4). Occupancy models were compared using widely applicable information criterion (WAIC; Watanabe, 2013).

Table 2.3. Occupancy models analyzed for four species of conservation concern. Occupancy models are described at three levels: site (ψ), sample (θ), and detection probability (p).

Model	Description
MODEL 1 (NULL)	ψ(·), θ(·), p(·)
MODEL 2	ψ(·), θ(predators), p(·)
MODEL 3	ψ(# wetland type), θ(·), p(·)
MODEL 4	ψ(% wetland), θ(·), p(·)
MODEL 5	ψ(% urban), θ(·), p(·)
MODEL 6	ψ(% forest), θ(season), p(·)
MODEL 7	ψ(% forest), θ(·), p(·)
MODEL 8	ψ(% urban), θ(season), p(·)
MODEL 9	ψ(season), θ(season + predators), p(·)
MODEL 10	ψ(·), θ(season + predators), p(·)
MODEL 11	ψ(# wetland type), θ(season), p(·)
MODEL 12	ψ(% wetland), θ(season), p(·)
MODEL 13	ψ(MAT + MAP), θ(·), p(·)
MODEL 14	ψ(% wetland + % developed + % forest + canopy cover), θ(·), p(·)
MODEL 15	ψ(MAT + MAP), θ(season), p(·)
MODEL 16	ψ(% wetland + % developed + % forest + canopy cover), θ(season), p(·)
MODEL 17	ψ(% wetland + % developed + % forest + canopy cover + # wetlands + MAP + MAT), θ(% wetland + % developed + % forest + canopy cover + # wetlands + MAP + MAT + predators), p(·)

RESULTS

A total of 28,009,686 vertebrate sequences were generated from collected project samples. Prior to filtering a total of 2,399,964 herpetofauna sequences were generated from 536 samples collected from 50 sites and over two sampling periods with a mean of 2,246 reads per sample (± 5,486). A total of 2,372,333 herpetofauna sequences were retained (mean = 2,255 ± 5,498 reads per sample, range = 16 to 41,096). Due to lack of water accessibility or other physical restraints in the field, not every site was sampled 12 times. Wetlands with shorter hydroperiods and smaller area were at times, not as comprehensively sampled as larger wetlands with longer hydroperiods. The mean number of reads in the negative control samples (across types) was 865 (± 3,124). The median number of reads for detections in negative controls (median = 15) was used as a threshold for detections in water samples associated with our project. Individual detections based on fewer than 15 reads were removed from the community matrix prior to further analysis.

Alpha Diversity Measures and Generalized Linear Modeling

Measures of diversity were calculated for each of 50 sites across Michigan, Indiana, and Ohio were calculated (Table 2.4). Across all sites, we detected 27 herpetofauna species. Species richness ranged widely across the 50 sites, but averaged 9.92 (± 2.83) herpetofauna species per site. The mean Shannon diversity index value was 1.14 (± 0.38). Our Porter County site (IN) had the lowest species richness (3) and the following two sites had the highest values of species richness: the Livingston County site (MI) and Delaware County site (IN) (15). For Shannon diversity values, the site with the lowest value was the St. Clair County site (MI; 0.33) and the site with the highest value was in Putnam County (OH; 1.93). Overall, there was a trend of increased diversity in the two southern states (Indiana and Ohio) in comparison to sites in

Michigan. Ohio had a mean Shannon diversity of 1.19 (± 0.42) across all sites sampled, and Indiana had a mean Shannon diversity of 1.18 (± 0.34). In comparison, Michigan had a mean Shannon diversity of 1.05 (± 0.38).

Table 2.4. Measures of alpha diversity for 50 sites in Michigan, Ohio, and Indiana (Shannon diversity index and species richness).

Site (County)	Shannon diversity	Species Richness
Williams	0.633	8
Defiance	0.764	7
Wayne, Ashland	1.259	14
Delaware (OH)	0.483	6
Marion	1.599	12
Medina	1.306	7
Allen	1.300	13
Allen	1.114	10
Putnam	1.925	10
Hancock	1.517	8
Champaign	1.600	12
Clinton	1.403	12
Ingham	0.864	7
Cass	1.268	11
Livingston	1.613	15
Paulding	1.380	7
Wabash	1.365	11
Marion	1.721	12
Mecosta	1.090	14
Wayne	0.805	11
Macomb	0.975	11
Porter	0.816	3
Pulaski	0.456	8
Starke	0.833	5
Jay	1.467	12
Newton	1.264	10
Porter	1.296	7
Lucas	1.696	11

Table 2.4 (cont'd)

Mecosta	1.466	9
Delaware (IN)	1.186	15
Erie	0.789	6
Steuben	1.797	14
St. Joseph	0.841	7
Tippecanoe	0.890	9
Bay	1.089	9
Saginaw	1.325	11
St Clair	0.329	6
Pulaski	1.049	11
Kosciusko	1.103	13
Marshall	1.320	9
Franklin	1.165	5
Portage	0.993	9
Kalamazoo	0.466	13
St. Joseph	1.283	13
Oakland	0.902	10
Summit	0.649	9
Guernsey	1.350	11
Allegan	1.426	9
Montcalm	1.323	11
Barry	0.551	13

For the thirteen models analyzed, the best-fit model related species richness to the number of wetland types sampled at a site (AICc value of 248.50, Table 2.5). This model carried 25% of the cumulative AICc weight. The estimated coefficient for the number of wetland type was -0.93, but the coefficient was not significantly different from zero ($p = 0.10$, t = -1.68). The second best-fitting model was the null (intercept) model with an AICc value of 249.07, carrying 19% of the total weight. For the generalized linear models used to assess Shannon diversity, the best-fitting model was the null (intercept) model (AICc = 48.40; Table 2.6). The second best-fitting model related Shannon diversity to the percentage of the landscape (within 1km) that was categorized as "developed" (i.e., areas with greater than 20% impervious cover, NLCD classifications 22, 23, and 24; AICc = 49.10). The estimated coefficient for the developed land cover coefficient was -0.05, but this coefficient was not significantly different from zero ($p = 0.84$ t = -0.20).

Table 2.5. AIC table for 13 generalized linear models to assess relationships between species richness and environmental characteristics for 50 sites in Michigan, Ohio, and Indiana.

Model	AICc	Delta AICc	Model Likelihood	AICc Weight	Log-likelihood	Cumulative Weight
Number of wetland types sampled	248.50	0.00	1.00	0.25	-120.99	0.25
Intercept only	249.07	0.57	0.75	0.19	-122.41	0.43
Mean annual temperature	250.07	1.57	0.46	0.11	-121.77	0.55
Latitude	250.68	2.19	0.33	0.08	-122.08	0.63
% Forest (within 1km) (NLCD)	250.79	2.30	0.32	0.08	-122.14	0.71
% Wetland (within 1km) (NLCD)	251.00	2.51	0.29	0.07	-122.24	0.78
% Developed (within 1km) (NLCD)	251.01	2.52	0.28	0.07	-122.25	0.85
Canopy cover	251.26	2.77	0.25	0.06	-122.37	0.91
Mean annual precipitation (PRISM)	251.29	2.79	0.25	0.06	-122.38	0.97
Mean temperature, mean precipitation, latitude	254.32	5.83	0.05	0.01	-121.48	0.98
Habitat (canopy cover, forest, wetland, # wetland types sampled)	254.65	6.16	0.05	0.01	-120.35	1.00
Landcover (canopy cover, forest, wetland, developed, # wetland types sampled)	257.53	9.03	0.01	0.00	-121.79	1.00
Global	258.60	10.11	0.01	0.00	-119.55	1.00

Table 2.6. AIC table for 13 generalized linear models to assess relationships between the Shannon diversity index and environmental characteristics for 50 sites in Michigan, Ohio, and Indiana.

Model	AICc	Delta AICc	Model Likelihood	AICc Weight	Log-likelihood	Cumulative Weight
Intercept only	48.40	0.00	1.00	0.21	-22.07	0.21
% Developed (within 1km) (NLCD)	49.10	0.70	0.71	0.15	-21.29	0.37
Latitude	49.63	1.23	0.54	0.12	-21.55	0.48
Mean annual precipitation (PRISM)	50.35	1.94	0.38	0.08	-21.91	0.56
% Wetland (within 1km) (NLCD)	50.42	2.02	0.37	0.08	-21.95	0.64
Mean annual temperature	50.46	2.06	0.36	0.08	-21.97	0.72
% Forest (within 1km) (NLCD)	50.56	2.16	0.34	0.07	-22.02	0.79
Canopy cover	50.61	2.21	0.33	0.07	-22.04	0.86
Number of wetland types sampled	50.62	2.22	0.33	0.07	-22.05	0.93
Mean temperature, mean precipitation, latitude	52.04	3.64	0.16	0.03	-20.34	0.97
Habitat (canopy cover, forest, wetland, # wetland types sampled)	53.51	5.11	0.08	0.02	-19.78	0.98
Landcover (canopy cover, forest, wetland, developed, # wetland types sampled)	53.58	5.17	0.08	0.02	-19.81	1.00
Global	60.94	12.54	0.00	0.00	-20.71	1.00

Occupancy Models

Three species from the list of 22 focal species developed for this project were detected in samples across the study region and one additional species had possible genus level detections. Due to this, eDNA occurrences were only available to fit occupancy models for four species: *Ambystoma texanum*, *Emydoidea blandingii*, *Thamnophis unclassified*, and *Rana palustris*. A total of 17, 10, 16, and 17 models (Table 2.7) were evaluated and relationships between occupancy of a species (respectively) (See Table 2.1 in bold) and environmental variables (Table 2.8) were determined.

Table 2.7. Occupancy models and WAIC values for three species of conservation concern.

A. texanum

Model	WAIC
ψ(% urban), θ(·), p(·)	56.76
ψ(# wetland type), θ(·), p(·)	57.42
ψ(MAT + MAP), θ(·), p(·)	58.06
ψ(% forest), θ(·), p(·)	58.60
ψ(MAT + MAP), θ(season), p(·)	59.12
ψ(% forest), θ(season), p(·)	59.32
ψ(% wetland), θ(season), p(·)	59.62
ψ(·), θ(·), p(·)	59.78
ψ(# wetland type), θ(season), p(·)	59.82
ψ(% urban), θ(season), p(·)	60.02
ψ(% wetland), θ(·), p(·)	60.12
ψ(% wetland + % developed + % forest + canopy cover), θ(season), p(·)	60.80
ψ(·), θ(season + predators), p(·)	61.11
ψ(% wetland + % developed + % forest + canopy cover), θ(·), p(·)	61.41
ψ(·), θ(predators), p(·)	62.82
ψ(season), θ(season + predators), p(·)	65.19
ψ(% wetland + % developed + % forest + canopy cover + # wetlands + MAP + MAT), θ(% wetland + % developed + % forest + canopy cover + # wetlands + MAP + MAT + predators), p(·)	65.56

E. blandingii

Model	WAIC
ψ(% urban), θ(·), p(·)	4.52
ψ(·), θ(·), p(·)	4.57
ψ(% wetland), θ(·), p(·)	4.83
ψ(% forest), θ(·), p(·)	4.85
ψ(% forest), θ(season), p(·)	5.13
ψ(# wetland type), θ(·), p(·)	5.36
ψ(MAT + MAP), θ(·), p(·)	5.69
ψ(# wetland type), θ(season), p(·)	5.69
ψ(% urban), θ(season), p(·)	5.85
ψ(MAT + MAP), θ(season), p(·)	6.28

R. palustris

Model	WAIC
ψ(·), θ(·), p(·)	12.09
ψ(% wetland), θ(·), p(·)	12.39
ψ(# wetland type), θ(·), p(·)	12.63
ψ(·), θ(predators), p(·)	12.66
ψ(MAT + MAP), θ(·), p(·)	12.69
ψ(% urban), θ(season), p(·)	12.94
ψ(% urban), θ(·), p(·)	13.08

Table 2.7 (cont'd)

ψ(# wetland type), θ(season), p(·)	13.69
ψ(MAT + MAP), θ(season), p(·)	13.76
ψ(% forest), θ(season), p(·)	13.77
ψ(% forest), θ(·), p(·)	13.77
ψ(% wetland + % developed + % forest + canopy cover), θ(·), p(·)	13.98
ψ(season), θ(season + predators), p(·)	14.16
ψ(% wetland + % developed + % forest + canopy cover), θ(season), p(·)	14.56
ψ(% wetland), θ(season), p(·)	14.62
ψ(·), θ(season + predators), p(·)	15.94

***Thamnophis* unclassified**

Model	**WAIC**
ψ(# wetland type), θ(·), p(·)	75.41
ψ(MAT + MAP), θ(·), p(·)	76.31
ψ(% urban), θ(season), p(·)	76.59
ψ(·), θ(·), p(·)	76.75
ψ(% forest), θ(·), p(·)	76.78
ψ(% wetland + % developed + % forest + canopy cover), θ(·), p(·)	76.82
ψ(% urban), θ(·), p(·)	76.82
ψ(% wetland), θ(·), p(·)	78.66
ψ(MAT + MAP), θ(season), p(·)	79.64
ψ(% forest), θ(season), p(·)	80.40
ψ(% wetland), θ(season), p(·)	81.72
ψ(% wetland + % developed + % forest + canopy cover), θ(season), p(·)	81.73
ψ(# wetland type), θ(season), p(·)	82.23
ψ(·), θ(predators), p(·)	83.90
ψ(season), θ(season + predators), p(·)	84.25
ψ(·), θ(season + predators), p(·)	84.41
ψ(% wetland + % developed + % forest + canopy cover + # wetlands + MAP + MAT), θ(% wetland + % developed + % forest + canopy cover + # wetlands + MAP + MAT + predators), *p*(·)	84.62

Table 2.8. Environmental variables included in generalized linear models (GLMs) to assess relationships between species richness (alpha diversity) and environmental characteristics, and variables used in occupancy models (OMs) to determine relationships between site characteristics and species occupancy.

Environmental variable	**Source**	**Model used for (GLM/OM)**
Canopy cover	NLCD	GLM/OM
% Wetland	NLCD	GLM/OM
% Forest	NLCD	GLM/OM
% Developed	NLCD	GLM/OM
Latitude	Collected project data	GLM/OM
Mean annual temperature	PRISM	GLM/OM
Mean annual precipitation	PRISM	GLM/OM
Number of wetland type sampled	Collected project data	GLM/OM
Season (Early summer/late summer)	Collected project data	OM
Fish predator presence	Collected project sequence data	OM

A limited number of detections for *E. blandingii* prevented us from evaluating more than 10 of the candidate models. For *A. texanum* and *E. blandingii*, the best model related site occupancy to the percentage of the landscape (within 1km) that was categorized as "developed", with no covariates for sample-level occupancy (WAIC = 56.76 and 4.52, respectively; Table 2.7). For *A. texanum*, the estimated coefficient for percent developed landcover in relation to site occupancy was was 0.02 (95 % HPD interval = -0.74-1.44). For *E. blandingii*, the estimated coefficient was -0.21 (95 % HPD interval = -1.84-1.54). For *Thamnophis* unclassified, the best-fit occupancy model related site occupancy to the number of wetlands sampled at a site (a surrogate measure of habitat heterogeneity) (WAIC = 75.41, coefficient = 0.04, 95 % HPD interval = -1.15-0.81). Meanwhile, for *Rana palustris*, the intercept-only model was the best fit to the data (WAIC = 12.09). The estimated detection probabilities for all four species were $p = 0.07$ (*A. texanum*), $p = 0.20$ (*E. blandingii*), $p = 0.11$ (*R. palustris*). and $p = 0.07$ (*Thamnophis* unclassified).

DISCUSSION

Alpha Diversity Measures and Generalized Linear Modeling

Results provide insight into the conservation status of both imperiled herpetofauna and wetlands alike, while also highlighting the utility of community metabarcoding tools in diverse wetland systems of the Great Lakes region. Linear modeling results suggest no deviation from the intercept only models, identifying no significant patterns between Shannon diversity, species richness, and environmental variables. In our occupancy models, though developed land cover was the best fit model for *A. texanum*, HPD intervals include zero, suggesting a lack of significance. The occupancy model findings for *E. blandingii* were also insignificant. For *Thamnophis* unclassified, the best fitting model (number of wetland types sampled) was also competitive with the intercept only model, and HPD intervals for the coefficient included zero,

again indicating a lack of a significant effect of number of wetland types sampled on site occupancy.

Based on model analysis, two independent variables in our best-fit models were identified as potential important contributing factors describing species richness and Shannon diversity using generalized linear model functions. Both the number of wetland types sampled at a site and the percentage of developed land cover within 1km of a site contributed to models that were assessed as best-fitting. However, in both cases, coefficients associated with model predictors were not significantly different from zero, indicating a lack of certainty about the direction of effects of wetland diversity and land cover on diversity of the herpetofauna communities inhabiting the sampled sites.

Species Distribution Modeling

When considering the stringent methods used to filter our environmental DNA sequence counts, we took a conservative approach to ensure we did not report false positives based on potential contamination at the various levels of DNA processing including sampling, extraction, and DNA amplification. While we have conservatively reported our findings, there is a potential for false negatives to occur for any species detected below the 15 sequence read threshold. Species detections were removed based on this threshold less often for species of conservation concern than for more abundant herpetofauna, but there were filtered detections of *A. texanum, Pantherophis spiloides, R. palustris,* and *Thamnophis unclassified.*

Occupancy Modeling

Previous literature (Peterman et al., 2013) suggested that the use of occupancy models in conjunction with SDMs can build upon complementary strengths to direct sampling efforts for target species. We detected three species of conservation concern in our survey for the purpose

of occupancy modeling. Our choice of marker (16S) and amplification bias may have affected detection probabilities for the species we did detect, and potentially had an effect on which species were not detected from our conservation concern list. In addition to the three species, we detected an unclassified genus group, *Thamnophis* unclassified. When developing our taxonomic database combined with our operational taxonomic unit data, we identified that we require more information on intra-specific variation over our study region to ensure we identify and assign species properly. This finding may explain why we saw an unclassified group as large as the genus *Thamnophis* appear in our detection data. For the purpose of the occupancy models, we used Thamnophis unclassified as a classification bin for other Thamnophis species that occur on our conservation concern list such as the Northern ribbon snake and the Plains garter snake. Additionally, all species used for our occupancy modeling had low estimated detection probabilities. This indicated that an increase in samples collected per wetland, per site, and/or additional PCR replicates could have strengthened our estimates of site occupancy across the three species of conservation concern.

Based on our results, two variables represented the best-fit models in relation to the occupancy of a species at a given site including the number of wetlands sampled at the site (which we used as a surrogate measure of the diversity of wetlands in a geographic area) and the level of development within 1km of the site. When targeting the number of wetlands to sample at a site, our team attempted to capture the most diversity possible. If three types of wetland types were present across the geographic area of a site, we sampled three separate wetland types (with two water samples collected from each). If only one was present, we still aimed to capture the most diversity by sampling multiple wetlands of the same classification. When considering an entire wetland herpetofauna community, the life history traits and biological requirements must

be considered. Certain species such as anurans and urodelans use highly ephemeral wetlands for breeding purposes due to decrease in larval and egg predation by fishes (Holbrook & Dorn, 2016; Skelly, 1997), while other species like reptiles may facultatively use other wetland types for water access, sunning, and other daily activities. Additionally, as wetlands change over the course of a season, wetland obligate species may shift patterns of habitat use among different wetland types present in an area. Our data showed that *Thamnophis* unclassified occupancy was slightly positively related to the number of wetlands sampled. Meanwhile, occupancy of both *A. texanum* and *E. blandingii* were best explained by the model using percentage of developed land cover, though HPD intervals for model coefficients included zero (Figures 2.3 and 2.4). As urbanization and agricultural needs increase, wetlands are degraded and fragmented (Antonio et al., 2022) potentially reducing the effective delivery of their ecosystem services to humans and wildlife alike. Fyson & Blouin-Demers (2021) found that areas of development were less likely to host Blanding's turtles using a combination of eDNA and visual survey data. While not previously noted for *A. texanum*, the suggestion that some amphibian species (including frogs and salamanders) may experience a positive relationship with urban land cover aligns with previous literature that suggests some species persist, and even thrive in developed landscapes (Browne et al., 2009; Wilk et al., 2020). Because the best-fit models were close in AICc value to our intercept-only models for *A. texanum*, *E. blandingii*, and *Thamnophis* unclassified, and HPD intervals for model coefficients include zero, findings from this study were inconclusive for the factors that may drive patterns of site occupancy for these species.

Figure 2.3. Site level detection for the Smallmouth salamander (*A. texanum*) in relation to the percentage of developed land within 1km of the site sampled.

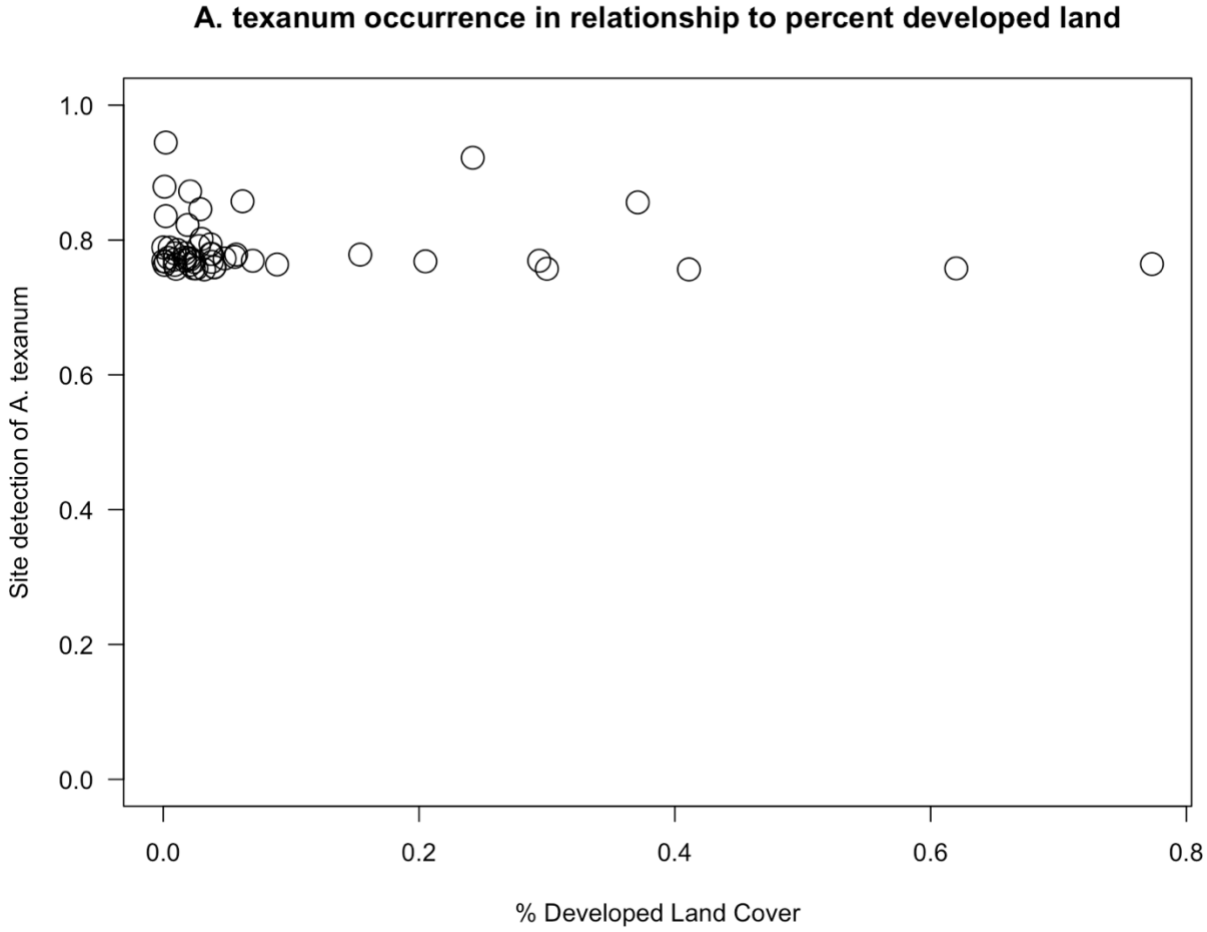
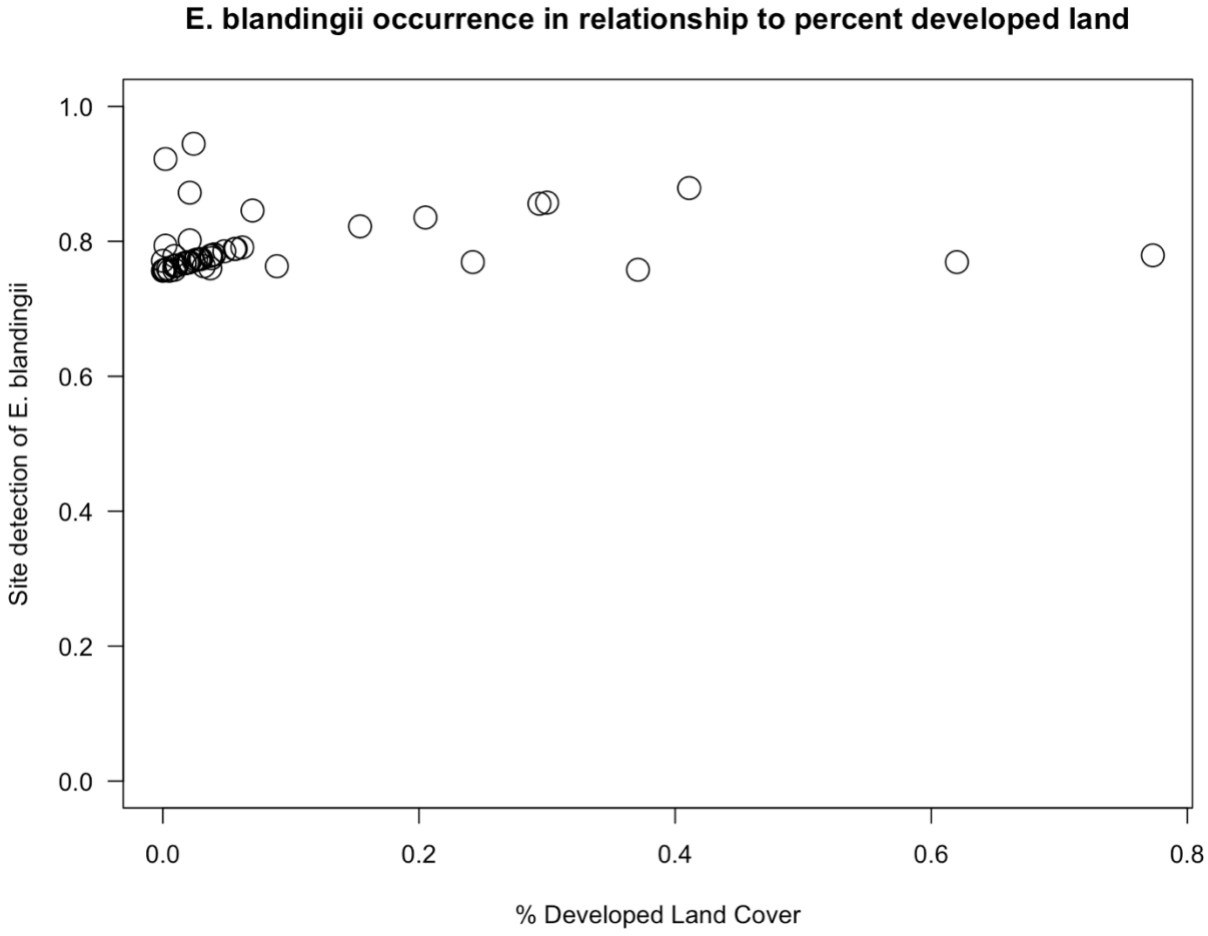


Figure 2.4. Site level detection for Blanding's turtle (*E. blandingii*) in relation to the percentage of developed land within 1km of the site sampled.



CONCLUSIONS AND REFLECTIONS

This study found no significant patterns when analyzing species richness and occupancy for threatened and endangered herpetofauna in Great Lakes wetland systems. After assessing collected data and the modeling results, our group had a few critical takeaways of how to improve and further answer queries surrounding habitat characteristics associated with wetland herpetofauna biodiversity and individual species occupancy. The first would be the utilization of ground truthing, or comparisons between previously collected data using traditional sampling methods and our collected eDNA data. The use of additional data in conjunction with our own may provide insight as to why we did or did not detect certain species at sampling sites across our study region. Second, additional interrogations into the breadth and validity of data used for species distribution models would be useful to create more accurate predictions. This could include sourcing data from additional or alternative databases and may involve further filtering to ensure an accurate set of occurrence data is used. In addition to the occurrence records, a focus on habitat characteristics versus climactic variables may contribute to more accurate species richness predictions. Important habitat characteristics could include soil composition and wetland specific information. The third takeaway was that additional sampling is needed for our occupancy modeling analyses. Whether this is more robust physical sampling in the field, or additional PCR replicates in the laboratory, more replication could provide us with more confidence in both our detection probabilities and the patterns associated with individual species presence. The fourth takeaway, which is still feasible for future research, is pursuing a smaller scale focus on species and taxonomic group habitat use. While this project was expansive both geographically and technically with metabarcoding applications, there is room for additional

work to explore finer-scale patterns to provide insight on habitat characteristic requirements for herpetofauna of conservation concern in the Great Lakes.

LITERATURE CITED

Akaike, H. 1974. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. doi:10.1109/TAC.1974.1100705

Antonio, G., Sandoval Herazo, L.C., Luis, J., López Méndez, M.C., & Arturo, E. 2022. Factors Affecting Wetland Loss: A Review. *Land, 11*(3), 434. doi:10.3390/land11030434

Binckley, C.A., & Resetarits, J.W. (2003). Functional equivalence of non-lethal effects: Generalized fish avoidance determines distribution of gray treefrog, Hyla chrysoscelis, larvae. *Oikos, 102*(3), 623-629. doi:10.1034/j.1600-0706.2003.12483.x

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Yu, D.W., de Bruyn, M. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution*, 29, 358–367. doi:10.1016/j.tree.2014.04.003

Browne, C.L., Paszkowski, C.A., Foote, A.L., Moenting, A. & Boss, S.M. 2009. The relationship of amphibian abundance to habitat features across spatial scales in the Boreal Plains, *Écoscience*, 16:2, 209-223, doi:10.2980/16-2-3220

Dahl, T.E. 2000. Status and trends of wetlands in the conterminous United States 1986 to 1997. U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C. 82 pp

Davidson, N.A. 2014. How much wetland has the world lost? Long-term and recent trends in global wetland area. *Marine and Freshwater Research*, 65, 934–941 http://dx.doi.org/10.1071/MF14173

Deagle, B.E., Kirkwood, R., & Jarman, S.N. 2009. Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular Ecology, 18*(9), 2022– 2038.

Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., de Vere, N., Pfrender, M.E., Bernatchez, L., 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, 26, 5872–5895. doi:10.1111/mec.14350

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

Dertien, J.S., Self, S., Ross, B.E., Barrett, K., & Baldwin, R.F. 2020. The relationship between biodiversity and wetland cover varies across regions of the conterminous United States. *PLOS ONE, 15*(5), e0232052. doi:10.1371/journal.pone.0232052

Dewitz, J. 2021. National Land Cover Database (NLCD) 2019 Products [Data set]. U.S.

Geological Survey. doi:10.5066/P9KZCM54

Dickie, I.A., Boyer, S., Buckley, H.L., Duncan, R.P., Gardner, P.P., Hogg, I.D., Holdaway, R.J., Lear, G., Makiola, A., Morales, S.E., Powell, J.R., Weaver, L. 2018. Towards robust and repeatable sampling methods in eDNA-based studies. *Molecular Ecology Resources*, *18*(2), 940–952. doi:10.1111/1755-0998.12907

Dorazio, R. M., & Erickson, R. A. 2018. Ednaoccupancy: An r package for multiscale occupancy modelling of environmental DNA data. *Molecular Ecology Resources*, *18*(2), 368-380. doi:10.1111/1755-0998.12735

Evans, N.T., Li, Y., Renshaw, M.A., Olds, B.P., Deiner, K., Turner, C.R., Jerde, C.L., Lodge, D.M., Lamberti, G.A., Pfrender, M.E. 2017. Fish community assessment with eDNA metabarcoding: Effects of sampling design and bioinformatic filtering. *Canadian Journal of Fisheries and Aquatic Sciences*, 74, 1362–1374. doi:10.1139/cjfas-2016-0306

Franklin, J., 2013. Species distribution models in conservation biogeography: Developments and challenges. *Diversity and Distributions*, 19, 1217–1223. doi:10.1111/ddi.12125

Fyson, V. K. & Blouin-Demers, G.. 2021. Effects of landscape composition on wetland occupancy by Blanding's Turtles (*Emydoidea blandingii*) as determined by environmental DNA and visual surveys. *Canadian Journal of Zoology*. *99*(8): 672-680. doi:10.1139/cjz-2021-0004

Gibbons, J.W., Scott, D.E., Ryan, T.J., Buhlman, K.A., Tuberville, T.D., Metts, B.S., Green, J.L., Mills, T., Leiden, Y., Poppy, S., Winne, C.T. 2000. The global decline of reptiles, déjà vu amphibians. *Bioscience,* 50, 653. doi:10.1641/00063568(2000)050[0653:tgdord]2.0.co;2

Gibbons, J.W. 2003. Terrestrial habitat: A vital component for herpetofauna of isolated wetlands. *Wetlands*, 23, 630–635. doi:10.1672/0277-5212(2003)023[0630:THAVCF]2.0.CO;2

Guisan, A., Thuiller, W., 2005. Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, 8, 993–1009. doi:10.1111/j.1461-0248.2005.00792.x

Hecnar, S.J. 2004. Great Lakes wetlands as amphibian habitats: A review. *Aquatic Ecosystem Health & Management,* 7(2): 289–303. doi:10.1080/14634980490461542

Holbrook, J.D. & Dorn, N.J. 2016. Fish reduce anuran abundance and decrease herpetofaunal species richness in wetlands. *Freshwater Biology,* 61, 100–109

Hurvich, C.M. and Tsai, C. 1989. Regression and time series model selection in small samples. *Biometrika*, *76*(2), 297-307.

Kingsford, R.T., Basset, A., & Jackson, L. 2016. Wetlands: Conservation's poor cousins.

Aquatic Conservation: Marine and Freshwater Ecosystems, *26*(5), 892-916. doi:10.1002/aqc.2709

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

Mackenzie, D.I., Royle, J.A., 2005. Designing occupancy studies: General advice and allocating survey effort. *Journal of Applied Ecology*, 42, 1105–1114. doi:10.1111/j.1365-2664.2005.01098.x

Marsh, D.M., Cosentino, B.J., Jones, K.S., Apodaca, J.J., Beard, K.H., Bell, J.M., Bozarth, C., Carper, D., Charbonnier, J.F., Dantas, A., Forys, E.A., Foster, M., General, J., Genet, K.S., Hanneken, M., Hess, K.R., Hill, S.A., Iqbal, F., Karraker, N.E., . . . Vonesh, J.R. 2017. Effects of roads and land use on frog distributions across spatial scales and regions in the Eastern and Central United States. *Diversity and Distributions*, *23*(2), 158-170. doi:10.1111/ddi.12516

Martinuzzi, S., Withey, J.C., Pidgeon, A.M., Plantinga, A.J., McKerrow, A.J., Williams, S.G., Helmers, D.P., & Radeloff, V.C. 2015. Future land-use scenarios and the loss of wildlife habitats in the southeastern United States. *Ecological Applications*, *25*(1), 160-171. doi:10.1890/13-2078.1

McKee, A.M., Calhoun, D.L., Barichivich, W.J., Spear, S.F., Goldberg, C.S., Glenn, T.C. 2015. Assessment of environmental DNA for detecting presence of imperiled aquatic amphibian species in isolated wetlands. *Journal of Fish and Wildlife Management*, *6*(2): 498–510.

Muha, T.P., Rolla, M., & Tricarico, E. 2017. Using environmental DNA to improve species distribution models for freshwater invaders. *Frontiers in Ecology and Evolution*, *5*, 312785. doi:10.3389/fevo.2017.00158

Nordstrom, B., Mitchell, N., Byrne, M., & Jarman, S. 2022. A review of applications of environmental DNA for reptile conservation and management. *Ecology and Evolution*, *12*(6), e8995. doi:10.1002/ece3.8995

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., and Wagner, H. 2020. vegan: Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Peterman, W.E., Crawford, J.A., Kuhns, A.R., 2013. Using species distribution and occupancy modeling to guide survey efforts and assess species status. *Journal for Nature Conservation*, 21, 114–121. doi:10.1016/j.jnc.2012.11.005

Phillips, S.J., Anderson, R.P., & Schapire, R.E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3-4), 231-259. doi:10.1016/j.ecolmodel.2005.03.026

Phillips, S.J., Dudik, M., & Schapire, R.E. 2023. Maxent software for modeling species niches and distributions (Version 3.4.1). Available from url: http://biodiversityinformatics.amnh.org/open_source/maxent/. Accessed on June 6 2023.

Piaggio, A.J., Engeman, R.M., Hopken, M.W., Humphrey, J.S., Keacher, K.L., Bruce, W.E., Avery, M.L., 2014. Detecting an elusive invasive species: A diagnostic PCR to detect Burmese python in Florida waters and an assessment of persistence of environmental DNA. *Molecular Ecology Resources*, 14, 374–380. doi:10.1111/1755-0998.12180

Prince, A.M., Andrus, L., 1992. PCR: How to kill unwanted DNA. *Biotechniques*, 12, 358–360.

Pukk, L., Kanefsky, J., Heathman, A.L., Weise, E.M., Nathan, L.R., Herbst, S.J., Robinson, J.D. 2021. eDNA metabarcoding in lakes to quantify influences of landscape features and human activity on aquatic invasive species prevalence and fish community diversity. *Diversity and Distributions*, 27, 2016–2031. doi:10.1111/ddi.13370

Quesnelle, P.E., Fahrig, L., & Lindsay, K.E. 2013. Effects of habitat loss, habitat configuration and matrix composition on declining wetland species. *Biological Conservation*, 160, 200-208. doi:10.1016/j.biocon.2013.01.020

Saenz-Agudelo, P., Delrieu-Trottin, E., DiBattista, J. D., Martínez-Rincon, D., Morales-González, S., Pontigo, F., Ramírez, P., Silva, A., Soto, M., & Correa, C. 2022. Monitoring vertebrate biodiversity of a protected coastal wetland using eDNA metabarcoding. *Environmental DNA*, *4*(1), 77-92. doi:10.1002/edn3.200

Sard, N., Scribner, K.T., Robinson, J.D., Herbst, S., Kanefsky, J., 2019. Comparison of fish detections, community diversity and relative abundance using environmental DNA metabarcoding and traditional gears. *Environmental DNA*, 1– 31. doi:10.1111/j.1365-2958.2003.03935.x

Sayers, E.W., Cavanaugh, M., Clark, K., Ostell, J., & Pruitt, K.D. 2020. GenBank. *Nucleic Acids Research*, *48*(D1), D84-D86. doi:10.1093/nar/gkz956

Schloss, P.D. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, 75:7537–7541.

Sierszen, M., Morrice, J., Trebitz, A., Hoffman, J. 2012. A review of selected ecosystem services provided by coastal wetlands of the Laurentian Great Lakes. *Aquatic Ecosystem Health & Management*, 15, 92-106. 10.1080/14634988.2011.624970.

Skelly, D.K. 1997. Tadpole communities: pond permanence and predation are powerful forces

shaping the structure of tadpole communities. *American Scientist*, 85:36–45.

Smith, A.B. enmSdm: tools for modeling niches and distributions of species. R package version 0.9.4. 2022.

Smith, A.B., Murphy, S.J., Henderson, D., and Erickson, K.D. 2023. Including imprecisely georeferenced specimens improves accuracy of species distribution models and estimates of niche breadth. *Global Ecology and Biogeography*. In press.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., Willerslev, E. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045-2050. doi:10.1111/j.1365-294X.2012.05470.x

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

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G. H., Geniez, P., Pont, D., Argillier, C., Baudoin, M., . . . Dejean, T. 2016. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, *25*(4), 929-942. doi:10.1111/mec.13428

Watanabe, S. 2013. A widely applicable Bayesian information criterion. *Journal of Machine Learning*, 14, 867–897.

Wikston, M., Breton, B., Vilaça, S.T., Bennett, A.M., Kyle, C.J., Beresford, D.V., Lesbarrères, D., Wilson, C.C., Green, D.M., Fortin, M., & Murray, D.L. 2023. Comparative efficacy of eDNA and conventional methods for monitoring wetland anuran communities. *Frontiers in Ecology and Evolution*, *11*, 1179158. doi:10.3389/fevo.2023.1179158

Wilk, A.J., Donlon, K.C. & Peterman, W.E. 2020. Effects of habitat fragment size and isolation on the density and genetics of urban red-backed salamanders (*Plethodon cinereus*). *Urban Ecosystems*, 23, 761–773. doi:10.1007/s11252-020-00958-8

APPENDIX

Table 1. Wetland classifications for site sampling.

Wetland type	**Abbreviation**	**Definition**
Vernal pool	VP	Classified and sourced by the National Wetlands Inventory
Marsh	M	Classified and sourced by the National Wetlands Inventory
Forested shrub	FS	Classified and sourced by the National Wetlands Inventory
Fen	F	Ground water sourced lowlands with small pools of water scattered throughout the plain
Pond	P	A medium sized natural or man-made pool larger than 5 feet across and has a long or permanent hydroperiod
Creek/drain	C	A small flowing body of water, often found adjacent to roads or fields
Bog	B	Poorly drained area rich with vegetation, often identified by bog-specific plant species (Mosses and sundews)
Other	O	Water access points not able to be classified into a traditional wetland categorization system

Figure 1. Trace plots of convergence for occupancy model (Model 5) for *A. texanum.*

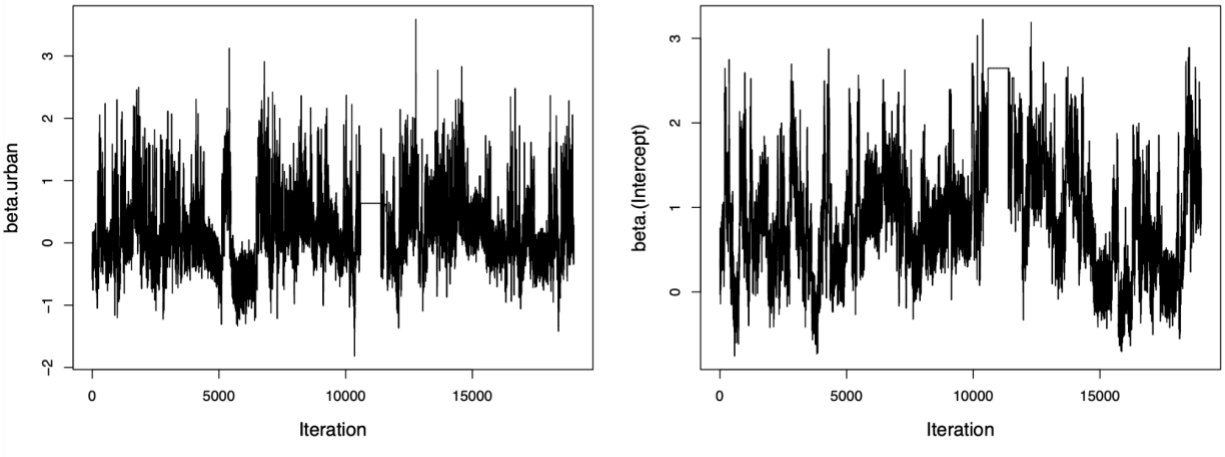


Figure 2. Trace plots of convergence for occupancy model (Model 5) for *E. blandingii*.

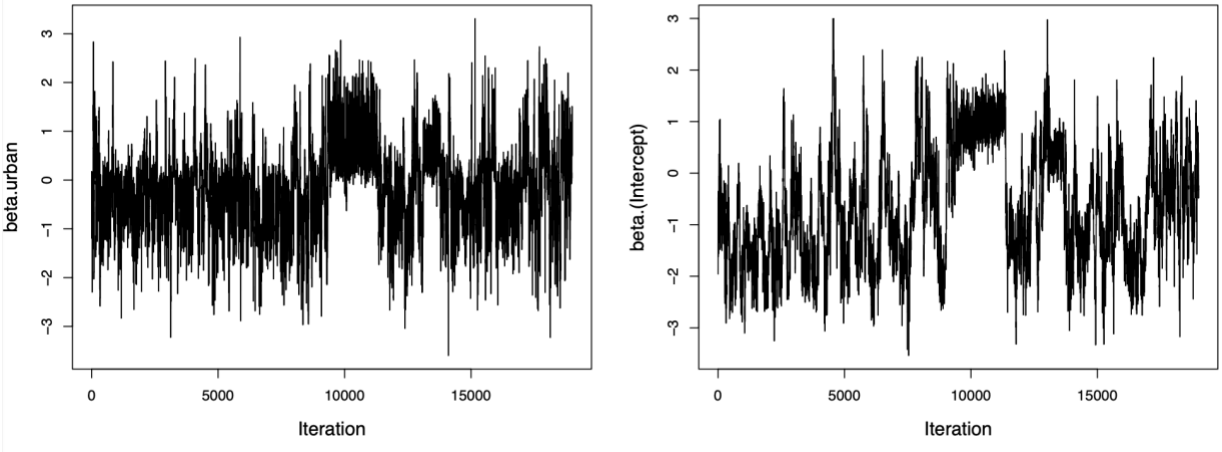


Figure 3. Trace plots of convergence for occupancy model (Model 1) for *R. palustris*.

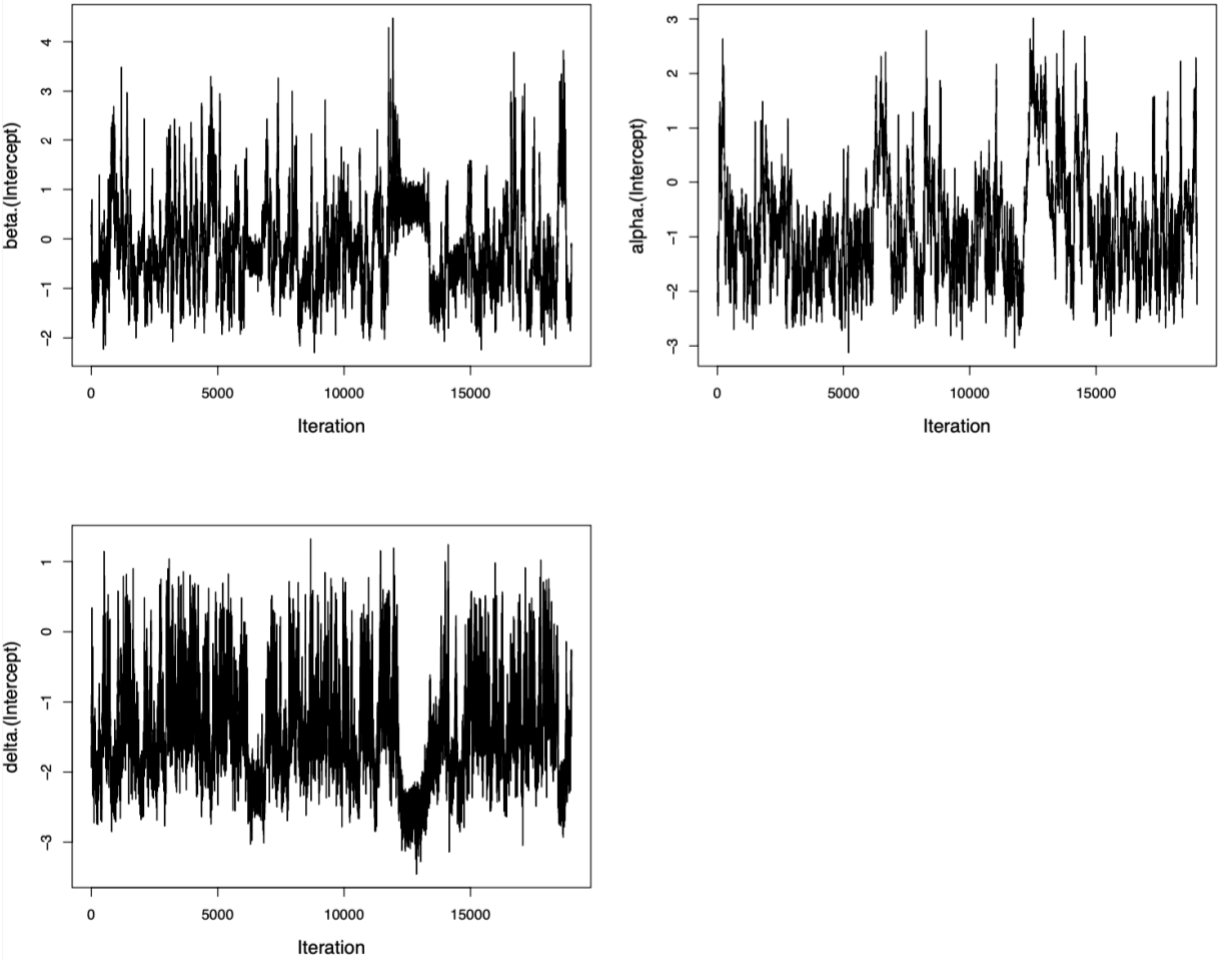
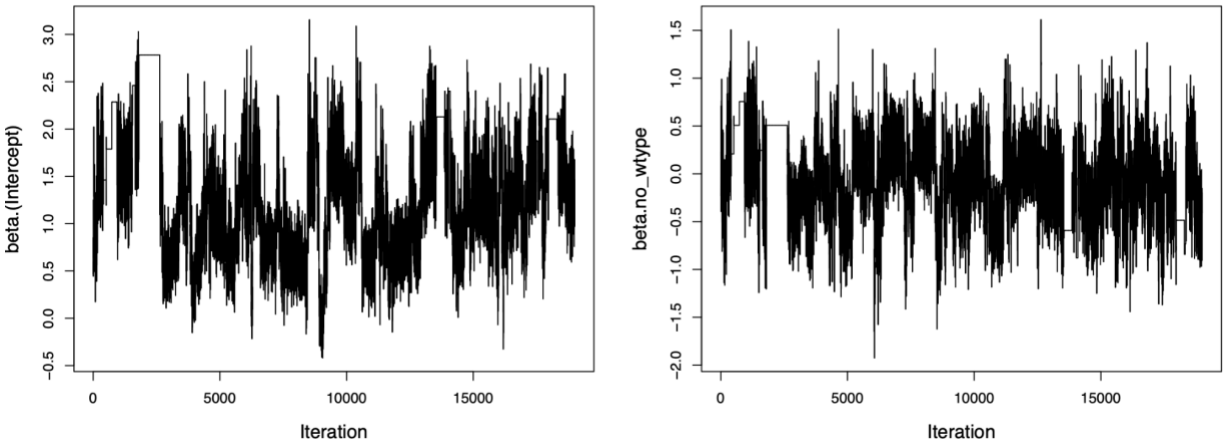



Figure 4. Trace plots of convergence for occupancy model (Model 3) for *T. unclassified.*