2022

SPATIAL HETEROGENEITY OF NITROGEN CYCLING WITHIN AND ACROSS FRESHWATER ECOSYSTEMS

Erin K. Eberhard
Michigan Technological University, ekeberha@mtu.edu

SPATIAL HETEROGENEITY OF NITROGEN CYCLING WITHIN AND ACROSS FRESHWATER ECOSYSTEMS

By

Erin K. Eberhard

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In Biological Sciences

MICHIGAN TECHNOLOGICAL UNIVERSITY

2022

This dissertation has been approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY in Biological Sciences.

Department of Biological Sciences

Dissertation Advisor:	*Amy M. Marcarelli, PhD*
Committee Member:	*Colden V. Baxter, PhD*
Committee Member:	*Evan S. Kane, PhD*
Committee Member:	*Stephen M. Techtmann, PhD*
Department Chair:	*Chandrashekhar Joshi, PhD*

Table of Contents

List of Figures v

List of Tables vii

Author Contribution Statement viii

Acknowledgements ix

Abstract xi

1 Chapter 1: Introduction 1
- 1.1 References 6

2 Chapter 2: Patch dynamics of N_2 fixation and denitrification in streams 11
- 2.1 Abstract 11
- 2.2 Introduction 12
- 2.3 Methods 15
  - 2.3.1 Study Area 15
  - 2.3.2 Study Design 19
    - 2.3.2.1 Mapping and Calculating Patch Area 19
    - 2.3.2.2 Substrate Collection 20
    - 2.3.2.3 N_2 Fixation 21
    - 2.3.2.4 Denitrification 22
    - 2.3.2.5 Microbial sampling and qPCR analysis 23
    - 2.3.2.6 Substrate Analysis 25
    - 2.3.2.7 Environmental Characteristics 25
    - 2.3.2.8 Data Analysis 27
- 2.4 Results 29
- 2.5 Discussion 40
- 2.6 References 47

3 Chapter 3: Heterogeneity in habitat and nutrient availability facilitate the co-occurrence of N_2 fixation and denitrification across wetland – stream – lake interfaces of Lakes Superior and Huron 57
- 3.1 Abstract 57
- 3.2 Introduction 58
- 3.3 Methods 62
  - 3.3.1 Study Area 62
  - 3.3.2 Study Design 67
    - 3.3.2.1 Transect Setup 67
    - 3.3.2.2 Nutrient Limitation 69
    - 3.3.2.3 N Cycling Rate Measurements 69
    - 3.3.2.4 Environmental Characteristics 72

3.3.2.5 Statistical Analysis 73
3.4 Results 77
3.4.1 Environmental Characteristics 77
3.4.2 Nutrient Limitation 79
3.4.3 Process Co-Occurrence and Nutrient Limitation 81
3.4.4 Environmental Factors as Predictors of Process Rates 85
3.5 Discussion 87
3.6 References 93

4 Chapter 4: Diversity of microbial assemblages in streams across ecoregions of the U.S. in relation to N_2 fixation and denitrification 104
4.1 Abstract 104
4.2 Introduction 105
4.3 Methods 107
4.3.1 Study Area 107
4.3.2 Sampling Design 108
4.3.2.1 Chamber Setup 113
4.3.2.2 Environmental Characteristics 113
4.3.2.3 Measurement of N transformation rates 114
4.3.2.4 Microbial Sampling and Illumina Sequencing 118
4.3.2.5 Statistical Analysis 122
4.4 Results 124
4.4.1 Taxonomic Composition and Diversity 124
4.4.2 Environmental Characteristics Across Ecoregions 127
4.4.3 Environmental Characteristics as Drivers of Community Composition 128
4.4.4 % N Fixers and Denitrifiers in Comparison to Average Stream Reach Rates 129
4.5 Discussion 132
4.6 References 136

5 Conclusion 143
5.1 References 147

A Appendix 149
A.1 Chapter 2 149
A.2 Chapter 3 166

List of Figures

Figure 2.1a. Example of habitat maps constructed for South Fork stream reach separated in three parts. 2.1b. Display of marked microhabitats within a riffle geomorphic unit in South Fork 19
Figure 2.2. Log transformed vertical dot plots of N fixation and denitrification rates of Diggie Creek (DC, n = 13), Gratiot River (G, n = 15), Gibson Jack (GJ, n = 18), McGunn (M, n = 11), Pilgrim River (P, n = 12), South Fork Mink Creek (SF, n = 33), and the Upper Portneuf River (UP, n = 12).....30
Figure 2.3. Plot of % contribution vs. % total reach area of all streams, of A. Gratiot River (G, n = 15), B. McGunn (M, n = 11), C. Pilgrim River (P, n = 12), D. Diggie Creek (DC, n = 13), E. Gibson Jack (GJ, n = 18), F. South Fork Mink Creek (SF, n = 33), and G. Upper Portneuf River (UP, n=12)……..32
Figure 2.4. Plot of relative abundances (C_T) of *nifH* and *nirS* for all streams. Diggie Creek (DC, n = 13), Gratiot River (G, n = 15), Gibson Jack (GJ, n = 18), McGunn (M, n = 11), Pilgrim River (P, n = 12), South Fork Mink Creek (SF, n = 33), and the Upper Portneuf River (UP, n = 12)……..33
Figure 2.5. Plot of relative abundances (C_T) of *nifH* and *nirS* versus rates of N_2 fixation and denitrification respectively. Each point represents a sample from one patch in the 7 streams……..34
Figure 2.6. Vertical dot plots of 9 environmental variables across the 7 streams. Diggie Creek (DC, n = 13), Gratiot River (G, n = 15), Gibson Jack (GJ, n = 18), McGunn (M, n = 11), Pilgrim River (P, n = 12), South Fork Mink Creek (SF, n = 33), and the Upper Portneuf River (UP, n = 12)……..36
Figure 2.7. Index values for each stream. Each dot is an index value for a stream where x = the # of patches / total stream reach area and y = stream reach average N flux (denitrification – N fixation). Pg 39……...40
Figure 3.1. Google Earth map of the state of Michigan and Wisconsin and the 5 wetland – stream – lake interfaces sampled in this study marked by a blue pin……63
Figure 3.2. Vertical dot plots of 9 environmental variables across for each transect and study year. Nara 2018 (N, n = 11), Sioux 2018 (Si., n = 9), Mackinac 2018 (M, n = 9), Nara 2019 (N, n = 11), Sioux 2019 (Si., n = 9), Mackinac 2019 (M, n = 7), Saganing 2020 (Sa., n = 9) and Wildfowl 2020 (W, n = 9).....77
Figure 3.3. Nutrient limitation data collected from nutrient diffusing substrates (NDS) for 4 of the 5 transects..……...80
Figure 3.4. Rates of N_2 fixation and denitrification among all transects compared to transect point classification……..……81
Figure 3.5. A heatmap of N_2 fixation (NF) and denitrification (AD) rates across all 5 transects broken down by year, transect number, and substrate type (M = macrophyte or S = sediment).....83
Figure 3.7. Rates of N_2 fixation and denitrification in comparison to nutrient limitation status……..84
Figure 4.1. Map of the continental U.S. broken down by ecoregion……...108
Figure 4.2. Relative abundances of microbial Classes by ecoregion……...124
Figure 4.3. Boxplot of the observed and Shannon alpha diversity for each ecoregion...127
Figure 4.4. Ordination of the principal coordinates analysis and distance based

redundancy analysis of the taxa for each ecoregion……129
Figure 4.5. Plots of the proportion of potential N_2-fixing taxa and stream reach average rates of N_2 fixation (mg $N/m^2/h$)……..131
Figure 4.6. Plots of the proportion of potential denitrifying taxa and stream reach average rates of denitrification (mg $N/m^2/h$)…….131
Figure 5.1. Graphic depiction of heterogeneity of aquatic ecosystems across spatial and temporal scales……143
Figure 1A (part 1 of 3). Rough habitat map for a reach of Diggie Creek in Idaho…....149
Figure 1A (part 2 of 3). Rough habitat map for a reach of Diggie Creek in Idaho…….150
Figure 1A (part 3 of 3). Rough habitat map for a reach of Diggie Creek in Idaho…….151
Figure 2A (part 1 of 3). Rough habitat map for a reach of South Fork in Idaho…...…..152
Figure 2A (part 2 of 3). Rough habitat map for a reach of South Fork in Idaho…….....153
Figure 2A (part 3 of 3): Rough habitat map for a reach of South Fork in Idaho...……..154
Figure 3A (part 1 of 2): Rough habitat map for a reach of the Pilgrim in Michigan…...155
Figure 3A (part 2 of 2): Rough habitat map for a reach of the Pilgrim in Michigan…...156
Figure 4A (part 1 of 2): Rough habitat map for a reach of McGunn in Michigan……..157
Figure 4A (part 2 of 2): Rough habitat map for a reach of McGunn in Michigan.…....158
Figure 5A (part 1 of 2): Rough habitat map for a reach of Gratiot in Michigan…….....159
Figure 5A (part 2 of 2): Rough habitat map for a reach of Gratiot in Michigan…...…..160
Figure 6A (part 1 of 2): Rough habitat map for a reach of Gibson Jack in Idaho……...161
Figure 6A (part 2 of 2): Rough habitat map for a reach of Gibson Jack in Idaho……...162
Figure 7A (part 1 of 2): Rough habitat map for a reach of Upper Portneuf in Idaho…..163
Figure 7A (part 2 of 2): Rough habitat map for a reach of Upper Portneuf in Idaho…..164
Figure 8A. Plot of relative abundances (C_T) of *nifH* and *nirS* across the 7 study streams for the 4 substrate types ordered top to bottom: wood, macrophyte, sediment, and rock……..165
Figure 9A. ArcGIS image of the Nara transect……....168
Figure 10A. ArcGIS image of the Sioux transect……169
Figure 11A. ArcGIS image of the Mackinac transect……..170
Figure 12A. Google earth image of the Saganing transect……..171
Figure 13A. Google earth image of the Wildfowl transect……..172

List of Tables

Table 2.1. Environmental characteristics for the 7 sampling streams in MI and ID arranged in alphabetical order……...17
Table 2.2. Environmental characteristics for the 7 sampling streams in MI and ID arranged in alphabetical order……18
Table 2.3. Predictive modeling results for the response variables N_2 fixation rates and denitrification rates……37
Table 2.4. Predictive modeling results for the response variables *nifH* and *nirS*……..39
Table 3.1. Environmental characteristics for the 5 sampled interfaces in the years 2018, 2019, and 2020……...65
Table 3.2. Environmental characteristics for the 5 sampled interfaces in the years 2018, 2019, and 2020……...66
Table 3.3. The sampling design of each interface is explained including years each transect was sampled, total transect points, total terrestrial points, total aquatic points, number of sites classified as wetland, number of sites classified as wetland to stream transition zones, number of sites classified as stream, number of sites classified as stream to lake transition zones, number of sites classified as lake, and dates sampled……68
Table 3.4. Predictive modeling results for the response variables N_2 fixation rates and denitrification rates……85
Table 4.1a. Physical environmental characteristics of a subset of the 30 study streams in 13 ecoregions by sampling date……...109
Table 4.1b. Physical environmental characteristics of a subset of the 30 study streams in 13 ecoregions by sampling date……...110
Table 4.1c. Nutrient concentrations of a subset of the 30 study streams in 13 ecoregions by sampling date……..111
Table 4.1d. Nutrient concentrations of a subset of the 30 study streams in 13 ecoregions by sampling date……..112
Table 4.2a. The substrate type sampling breakdown of a subset of the 30 streams……119
Table 4.2b. The substrate type sampling breakdown of a subset of the 30 streams……120
Table 1A. The breakdown of deployment of nutrient diffusing substrate (NDS) in all transects sampled in 2019……166
Table 2A. The breakdown of deployment of nutrient diffusing substrate (NDS) in all transects sampled in 2020……167

Author Contribution Statement

Chapter 2: conception and design: Erin Eberhard, Amy Marcarelli, and Colden Baxter; data collection: Erin Eberhard, Amy Marcarelli, Colden Baxter and Stephen Techtmann; analysis and interpretation of results: Erin Eberhard, Amy Marcarelli, Stephen Techtmann; manuscript preparation and revision: Erin Eberhard, Stephen Techtmann, Colden Baxter, and Amy Marcarelli

Chapter 3: conception and design: Erin Eberhard and Amy Marcarelli; data collection: Erin Eberhard; analysis and interpretation of results: Erin Eberhard and Amy Marcarelli; manuscript preparation and revision: Erin Eberhard and Amy Marcarelli

Chapter 4. conception and design: Erin Eberhard and Amy Marcarelli; data collection: Erin Eberhard; analysis and interpretation of results: Erin Eberhard, Stephen Techtmann; manuscript preparation and revision: Erin Eberhard and Amy Marcarelli

Acknowledgements

This has been a roller coaster of a journey and it has taken a village of people to help me get to this point. I could go on forever about how thankful I am and to whom, but I will keep it short for the purposes of this document.

I would first like to thank my advisor, Amy Marcarelli, for her guidance and support over the last 7 years, which has meant the world to me. Thank you for actually caring about my well-being and having a contagious enthusiasm for science that has helped when I've lost my own enthusiasm. I would not be where I am today without you and I can't thank you enough for your mentorship.

I would also like to thank my committee members: Colden Baxter, Stephen Techtmann, and Evan Kane for their guidance and input throughout the years on this dissertation. Thank you, Colden Baxter for welcoming me into your lab in Idaho my first summers of graduate school and for pounding in all those piezometers. Thank you, Stephen Techtmann for being there whenever I had a question and when I needed help troubleshooting the microbial techniques and for being an all-around positive person. Thank you, Evan Kane for your help over the years with sample analysis and troubleshooting wetland sampling techniques.

Thank you to the funding sources for this project: National Science Foundation award DEB 14-51919 to Amy Marcarelli, Michigan Space Grant Consortium Graduate Fellowship under award 80NSSC20M0124 from the National Aeronautics and Space Administration, Michigan Sea Grant Graduate Fellowship under award R/CGLH-11 from the National Oceanic and Atmospheric Administration, Ecosystem Science Center, and the Great Lakes Research Center. Thank you also to the funding support from the Biological Sciences department and the Michigan Tech Graduate School Fall Finishing Fellowship, which was funded by the Nicholas Matwiyoff and Carl Hogberg Graduate Fellowships.

Thank you to Anne Garwood at MI EGLE for being my agency sponsor on the MI Sea Grant Graduate Fellowship and providing guidance on navigating agency jobs. Thank you to Matthew Cooper for helping to sample the Sioux and deploy NDS when I could not physically be there. Thank you to Casey Huckins and members of the Huckarelli lab group for your support over the years with research ideas and presentations.

I would not have all this data without the help of so many people over the years. Thank you to NEON and StreamPulse scientists, staff and students for showing us around field sites across the country. Thank you to the Idaho crew - Jenny Cornell, Adam Eckersell, Jade Ortiz, and James Paris - for your help and support over many sampling summers. Thank you to Claire Allison, Meredith Brehob, Maya Braden, Acacia Copley, Olivia Deans, Andie Norton, Joslyn Pomeroy, Hunter Roose, Renn Schipper, Katie Summers, Molly Warner, and Ryan Van Goethem for field and laboratory assistance. Thank you to Laura Schaerer for your help with microbial lab stuff over the years and for taking the time to show me how to analyze the 16S rRNA sequences and sharing your R code.

Michelle Kelly – Thank you so much for being a friend. Thank you for all of your help in the field, in lab, and with all the modeling in R. Thank you for having the guts to take the lab van to the gate at the Arikaree.

Kevin Nevorski – Your friendship is one of the best things I have gotten from this PhD experience. Thank you for being a confidant over the years and for pushing me outside my comfort zone. There is no one else I'd rather sing along to Disney with and drive cross country with in a tiny van. Thank you for your constant help in the field and lab.

Thank you to Bailey Duxbury and Angela Walczyk for your friendship and support over the years. Thank you to Poppy for being the best giant dog and a source of smiles and cuddles during hard times. Thank you to Rochelle Spencer for the weekly coffee dates, being my running buddy, and being a confidant. Thank you to Taylor Bruns, Scott Loeffler, and Sarah Murphy for listening to me, supporting me from far away, and reminding me I am a person outside of being a graduate student.

I also would not be here without my family. Thank you to my grandparents, Bill and Donna Krist and Beverly and (the late) David Eberhard for believing in me and supporting me. Thank you to my parents, Mark and Kim Eberhard for always being there for me no matter what and believing in me when I don't believe in myself. Thank you to my brothers, Erich, Ean, and Eamon for always being there for me and for all of the belly laughs, especially in times when I hadn't laughed in a while. Thank you to my cat Nala for all of the cuddles and smiles and being the sweetest. Thank you all for reminding me there is life outside of graduate school.

Finally, I would like to thank MTU's Center for Student Mental Health and Well-being. Even in the best of circumstances graduate school can become isolating and stressful and sometimes one needs help to deal with that. I would particularly like to thank Crystal Primaeu McLeod and Lisa Meyer for all of their help over the years. I would not have navigated it as well without your support.

Abstract

Dinitrogen (N_2) fixation and denitrification are two nitrogen (N) cycling processes that despite differences in environmental requirements and constraints, co-occur in aquatic ecosystems. The overall goal of this dissertation was to evaluate how spatial heterogeneity of environmental variables 1) drive hot spots of N_2 fixation, denitrification rates and gene abundances in streams, 2) facilitate co-occurrence of these processes across wetland – stream – lake interfaces, and 3) affect differences in microbial community composition in streams across U.S. ecoregions. We found hot spots of both processes within 7 stream reaches in Michigan and Idaho, but rates of N_2 fixation were not directly related to relative gene abundances of *nifH*, while denitrification rates were related to relative gene abundances of *nirS.* Spatial heterogeneity of organic matter and dissolved oxygen concentrations were important predictors of rates of both processes. In a survey across 5 wetland – stream – lake interfaces of Lakes Superior and Huron, we found that rates of N_2 fixation and denitrification occurred across stream, wetland and shallow lake habitats and that phosphorus (P) availability was important for predicting rates of both processes, while N availability was an important predictor of denitrification and carbon (C) availability was important predictor of N_2 fixation. Finally, in a survey of microbial assemblages from 30 streams across 13 U.S. ecoregions, we found that microbial community composition differed across ecoregions in alpha diversity and relative Class abundances, but little of this variation was explained by environmental variables. Together, these studies show that N_2 fixation and denitrification co-occurred in stream and coastal ecosystems and across spatial scales from stream reaches to ecoregions. However, rates and microbial community composition are not explained fully

by differences in environmental variables on the microhabitat, cross-habitat, or ecoregion scale. N alone was not always an important predictor of the processes despite N being thought of as the best indicator of these processes in the past. Overall, these studies highlight the need to include both N_2 fixation and denitrification measurements in biogeochemical studies for a better understanding of the complexity of N cycling in aquatic ecosystems.

1 Chapter 1: Introduction

Heterogeneity is defined as the variability in a process or pattern over space and time (Palmer and Poff 1997). Spatial heterogeneity is an influential factor in ecological systems that can affect the flux of organisms and materials in an environment (Pickett and Cadenasso 1995). Spatial heterogeneity has been shown to affect the distribution and diversity of organisms in environmental space such as algae and invertebrates in marine intertidal habitats and birds in terrestrial ecosystems. (Paine and Levin 1981, Pickett and Cadenasso 1995). Analyzing the effects of spatial heterogeneity can help explain many complexities within an ecosystem such as species distribution, nutrient concentrations, and biogeochemical fluxes (Pickett and Cadenasso 1995, Pringle et al. 1988, Dent and Grimm 1999).

Aquatic ecosystems exhibit a high degree of spatial complexity that regulates ecosystem processes. The spatial complexity of lakes, streams, and wetlands within a drainage network can affect temperature, mixing of water, and nutrient processing within each system (Jones 2010). Streams are characterized by habitat heterogeneity at multiple, nested scales (Frissell et al. 1986) that, in turn, influences heterogeneity in streamwater chemistry, organisms, and ecosystem processes (e.g., Dent and Grimm 1999, McGuire et al. 2014). The flow of water in streams creates a spiraling transport pattern where nutrients and organic matter are being primarily transported downstream rather than stored (Newbold et al. 1983). In wetlands, however, water movement is slower, with much water located in standing pools. Wetland ecosystems can trap organic matter, remove excess nutrients from runoff (McCarthy et al. 2007, Uzarski et al. 2009), and display physical and chemical gradients of plants and nutrients from the edge to the

interior (Cooper et al. 2012, Mitsch and Gosselink 2015). In near-shore lake areas, nutrients can also be stored in sediments despite the vertical and horizontal mixing of water. Differences in nutrient storage within and nutrient transport across these ecosystems can promote differences in biogeochemical processes within and across ecosystems.

Biogeochemical processing of elements like C and N are spatially and temporally variable within ecosystems, which has consequences for whole-ecosystem fluxes and budgets (McClain et al. 2003). Net N_2 flux in aquatic ecosystems is controlled by N_2 fixation, denitrification, and anammox (Schlesinger and Bernhardt 2020, Zehr and Capone 2021). Anammox is an anerobic process where ammonium is oxidized to N_2 using nitrite and it may account for ~30-50% of N_2 production in oceans (Thamdrup and Dalsgaard 2002). However, in freshwater ecosystems the contribution of anammox to N_2 production is largely unknown (Crowe et al. 2017, Schlesinger and Bernhardt 2020), so studies in freshwater ecosystems have focused on the contributions of N_2 fixation and denitrification as the primary controls of N_2 flux. N_2 fixation is the microbial conversion of atmospheric N_2 gas into an input of biologically available N, while denitrification is the microbial conversion of nitrate (NO_3^-) into N_2 gas, removing N from the ecosystem (Schlesinger and Bernhardt 2020). Denitrifying bacteria are primarily heterotrophic anaerobes, while N_2 fixing bacteria are more diverse encompassing heterotrophic anaerobes in sediments to photosynthetic cyanobacteria (Groffman et al. 2009, Vitousek et al. 2002). Both processes have different N requirements, which is why they have not been extensively studied together in freshwater ecosystems. N_2 fixation tends to be favored when NO_3^- concentrations are low because N_2 fixation has significant energy

costs to the organism (Grimm and Petrone 1997, Kunza and Hall 2013), while denitrification requires higher concentrations of NO_3^- to use as an oxidant (Arango et al. 2007). Despite this difference in N requirements, we have previously found that N_2 fixation and denitrification co-occur in streams (Eberhard et al. 2018), similar to what others have found in coastal areas (Fulweiler and Weiss 2014, Newell et al. 2016). Our previous results suggest that spatial heterogeneity of controlling factors is an important mechanism underlying this co-occurrence, but spatially explicit sampling is required to understand how spatial heterogeneity may facilitate hotspots and whole-ecosystem contributions of these processes.

Spatial heterogeneity of environmental variables has been shown to affect N cycling at multiple scales in aquatic ecosystems. In wetlands, the spatial gradient of nutrient concentrations can create conditions more suitable for different microorganisms to perform processes that are not all favorable under one type of nutrient limitation (Cooper et al. 2016). Variation in substrate, light, temperature, and organic matter can also affect rates of these processes at the reach and sub-reach scale of streams (Holmes et al. 1996, Marcarelli and Wurtsbaugh 2009, Eberhard et al. 2018). Spatial variation in environmental conditions can create hot spots for biogeochemical processes, which may facilitate coexistence of processes that require very different environmental conditions to occur (McClain et al. 2003). These hot spots may also be promoted by variation in microbial community composition that is driven by variation in environmental conditions within and/or among ecosystems (McClain et al. 2003). Through the use of 16S rRNA sequencing and qPCR that targets genes associated with enzymes involved in each process (nitrogenase (*nifH*) with N_2 fixation and nitrite reductase (*nirS, nirK*), nitrate

reductase (*narG*), and nitrous oxide reductase (*nosZ*) with denitrification) studies have shown that variation in environmental factors can be correlated to changes in the microbial community and their resulting biogeochemical potential (Zehr and Capone 2021, Wallenstein et al. 2006). This can also be observed across biomes where variability in microbial assemblages can control biogeochemical processes because organisms that differ in processing capabilities thrive in environments that differ (Fierer et al. 2012).

The goal of this dissertation was to examine how spatial heterogeneity affects the biogeochemical processing of N in aquatic freshwater ecosystems. Specifically, I examined how the variability in environmental conditions and limitation by key nutrients affect rates of N_2 fixation and denitrification as well as composition of microbial assemblages at different spatial scales.

In chapter 2, I measured rates of N_2 fixation and denitrification and relative abundances of genes *nifH* and *nirS* in patches of 7 stream reaches across MI and ID. I hypothesized that spatial variability in environmental variables would facilitate hot spots of process rates and gene abundances. I found that hot spots of N_2 fixation and denitrification occurred across all streams at the patch scale and that there was no direct relationship between N_2 fixation rates and relative abundances of *nifH*, but there was a relationship between denitrification rates and relative abundances of *nirS*. Spatial heterogeneity of organic matter and dissolved oxygen availability were important to predicting rates of both N_2 fixation and denitrification in these streams. These results suggest that spatial heterogeneity in environmental variables is important to the occurrence of N_2 fixation and denitrification in streams and the overall N_2 flux of these ecosystems.

In chapter 3, I evaluated how spatial variability in nutrient limitation and other environmental variables facilitate rates of N_2 fixation and denitrification across 5 wetland - stream - lake interfaces of Lakes Superior and Huron. I hypothesized that there would be spatial variability in nutrient limitation across the interfaces and that the spatial variability of nutrient limitation and other environmental variables would facilitate the occurrence of N_2 fixation and denitrification across the interfaces. My results showed that there was spatial variability in nutrient limitation of attached algae across the interfaces with no limitation, N limitation and N + P limitation detected, but that was not directly related to process rates. Rates of both processes were variable among wetland, stream, nearshore lake, and transition zone habitat types in the interfaces and dissolved P concentrations were important predictors of both processes, while dissolved N concentrations were an important to predictor only to denitrification and C (as organic matter and dissolved organic carbon) was an important predictor to N_2 fixation rates. These results suggest that coastal ecosystems should not be thought of us a simple sink of N, but as more biogeochemically complex because processes both removing and inputting N are occurring at relatively high rates.

In chapter 4, I evaluated the microbial community composition in 30 streams across 13 ecoregions using 16S rRNA Illumina sequencing to test if differences in environmental variables across ecoregions drove differences in microbial assemblages. Using a distance-based redundancy analysis I found that microbial community composition was significantly different across ecoregions, but that differences in environmental variables only explained a small portion of the variance in this relationship with the two most important axes explaining 4.2% and 2.8% of the variability. The

program FAPROTAX was used to assess the functional potential of these communities to perform N_2 fixation and/or denitrification, which were compared to average stream reach rates of N_2 fixation and denitrification. The functional potential of these communities was not directly related to rates of N_2 fixation or denitrification at the stream reach scale. Variability of environmental characteristics within stream reaches may better explain differences in microbial community composition.

1.1 References

Arango, C. P., Tank, J. L., Schaller, J. L., Royer, T. V., Bernot, M. J., & David, M. B. (2007). Benthic organic carbon influences denitrification in streams with high nitrate concentration. *Freshwater Biology*, *52*(7), 1210-1222.

Cooper, M. J., Gyekis, K. F., & Uzarski, D. G. (2012). Edge effects on abiotic conditions, zooplankton, macroinvertebrates, and larval fishes in Great Lakes fringing marshes. *Journal of Great Lakes Research*, *38*(1), 142-151.

Cooper, M. J., Costello, G. M., Francoeur, S. N., & Lamberti, G. A. (2016). Nitrogen limitation of algal biofilms in coastal wetlands of Lakes Michigan and Huron. *Freshwater Science*, *35*(1), 25-40.

Crowe, S. A., Treusch, A. H., Forth, M., Li, J., Magen, C., Canfield, D. E., ... & Katsev, S. (2017). Novel anammox bacteria and nitrogen loss from Lake Superior. *Scientific reports*, *7*(1), 1-7.

Dent, C. L., & Grimm, N. B. (1999). Spatial heterogeneity of stream water nutrient concentrations over successional time. *Ecology*, *80*(7), 2283-2298.

Eberhard, E. K., Marcarelli, A. M., & Baxter, C. V. (2018). Co-occurrence of in-stream nitrogen fixation and denitrification across a nitrogen gradient in a western US watershed. *Biogeochemistry*, *139*(2), 179-195.

Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., ... & Caporaso, J. G. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences*, *109*(52), 21390-21395.

Frissell, C. A., Liss, W. J., Warren, C. E., & Hurley, M. D. (1986). A hierarchical framework for stream habitat classification: viewing streams in a watershed context. *Environmental management*, *10*(2), 199-214.

Fulweiler, R. W., & Heiss, E. M. (2014). (Nearly) A Decade of Directly Measured Sediment N_2 Fluxes: What Can Narragansett Bay Tell Us About the Global Ocean Nitrogen Budget?. *Oceanography*, *27*(1), 184-195.

Grimm, N. B., & Petrone, K. C. (1997). Nitrogen fixation in a desert stream ecosystem. *Biogeochemistry*, *37*(1), 33-61.

Groffman, P. M., Butterbach-Bahl, K., Fulweiler, R. W., Gold, A. J., Morse, J. L., Stander, E. K., ... & Vidon, P. (2009). Challenges to incorporating spatially and temporally explicit phenomena (hotspots and hot moments) in denitrification models. *Biogeochemistry*, *93*(1), 49-77.

Holmes, R. M., Jones, J. B., Fisher, S. G., & Grimm, N. B. (1996). Denitrification in a nitrogen-limited stream ecosystem. *Biogeochemistry*, *33*(2), 125-146.

Jones, N. E. (2010). Incorporating lakes within the river discontinuum: longitudinal

changes in ecological characteristics in stream–lake networks. *Canadian Journal of Fisheries and Aquatic Sciences*, *67*(8), 1350-1362.

Kunza, L. A., & Hall Jr, R. O. (2013). Demographic and mutualistic responses of stream nitrogen fixers to nutrients. *Freshwater Science*, *32*(3), 991-1004.

Marcarelli, A. M., & Wurtsbaugh, W. A. (2009). Nitrogen fixation varies spatially and seasonally in linked stream-lake ecosystems. *Biogeochemistry*, *94*(2), 95-110.

McCarthy, M. J., Gardner, W. S., Lavrentyev, P. J., Moats, K. M., Jochem, F. J., & Klarer, D. M. (2007). Effects of hydrological flow regime on sediment-water interface and water column nitrogen dynamics in a Great Lakes coastal wetland (Old Woman Creek, Lake Erie). *Journal of Great Lakes Research*, *33*(1), 219-231.

McClain, M. E., Boyer, E. W., Dent, C. L., Gergel, S. E., Grimm, N. B., Groffman, P. M., ... & Pinay, G. (2003). Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems*, 301-312.

McGuire, K. J., Torgersen, C. E., Likens, G. E., Buso, D. C., Lowe, W. H., & Bailey, S. W. (2014). Network analysis reveals multiscale controls on streamwater chemistry. *Proceedings of the National Academy of Sciences*, *111*(19), 7030-7035.

Mitsch, W. J., & Gosselink, J. G. (2015). *Wetlands*. John Wiley & Sons.

Newbold, J. D., Elwood, J. W., O'neill, R. V., & Sheldon, A. L. (1983). Phosphorus dynamics in a woodland stream ecosystem: a study of nutrient spiralling. *Ecology*, *64*(5), 1249-1265.

Newell, S. E., McCarthy, M. J., Gardner, W. S., & Fulweiler, R. W. (2016). Sediment

nitrogen fixation: a call for re-evaluating coastal N budgets. *Estuaries and Coasts*, *39*(6), 1626-1638.

Paine, R. T., & Levin, S. A. (1981). Intertidal landscapes: disturbance and the dynamics of pattern. *Ecological monographs*, *51*(2), 145-178.

Palmer, M. A., & Poff, N. L. (1997). The influence of environmental heterogeneity on patterns and processes in streams. *Journal of the North American Benthological Society*, *16*(1), 169-173.

Pickett, S. T., & Cadenasso, M. L. (1995). Landscape ecology: spatial heterogeneity in ecological systems. *Science*, *269*(5222), 331-334.

Pringle, C. M., NAiMAN, R. J., Bretschko, G., Karr, J. R., Oswood, M. W., Webster, J. R., ... & Winterbourn, M. J. (1988). Patch dynamics in lotic systems: the stream as a mosaic. *Journal of the North American benthological society*, *7*(4), 503-524.

Schlesinger, W. H., & Bernhardt, E. S. (2020). Biogeochemistry: an analysis of global change. Academic Press.

Thamdrup, B., & Dalsgaard, T. (2002). Production of N2 through anaerobic ammonium oxidation coupled to nitrate reduction in marine sediments. *Applied and environmental microbiology*, *68*(3), 1312-1318.

Uzarski, D. G., Burton, T. M., Kolar, R. E., & Cooper, M. J. (2009). The ecological impacts of fragmentation and vegetation removal in Lake Huron's coastal wetlands. *Aquatic Ecosystem Health & Management*, *12*(1), 45-62.

Vitousek, P. M., Cassman, K. E. N., Cleveland, C., Crews, T., Field, C. B., Grimm, N. B., ... & Sprent, J. I. (2002). Towards an ecological understanding of biological nitrogen fixation. *Biogeochemistry*, *57*(1), 1-45.

Wallenstein, M. D., Myrold, D. D., Firestone, M., & Voytek, M. (2006). Environmental controls on denitrifying communities and denitrification rates: insights from molecular methods. *Ecological applications*, *16*(6), 2143-2152.

Zehr, J. P., & Capone, D. G. (2021). *Marine Nitrogen Fixation*. Springer Nature.

2 Chapter 2: Patch dynamics of N_2 fixation and denitrification in streams

2.1 Abstract

Stream ecosystems are characterized by high degrees of spatial heterogeneity that can result in patches of microhabitats that vary in their composition and support different groups of organisms. This feature may facilitate the co-occurrence of biogeochemical processes thought to be incompatible due to contrasting environmental constraints, like N_2 fixation and denitrification, through the creation of hot spots. We hypothesized that hot spots of N_2 fixation and denitrification would occur in streams and that variation in environmental variables would facilitate the occurrence of both processes. To test this hypothesis, we measured rates of N_2 fixation and denitrification along with relative abundances of the genes *nifH* and *nirS* in patches determined by channel geomorphic units and substrate type in 4 Idaho and 3 Michigan streams. We found that hot spots of N_2 fixation and denitrification occurred in all stream reaches, and that rates of N_2 fixation were not correlated to *nifH* relative abundances, while *nirS* was positively correlated with denitrification rates. Predictive modeling showed that organic matter and dissolved oxygen concentrations predicted rates of N_2 fixation, denitrification, and *nifH* relative abundance. Phosphorus (P) concentrations and the ratio of dissolved inorganic N to total dissolved P also were important predictors of N_2 fixation rates and the relative abundances of *nifH* and *nirS*, while variables related to N alone were important predictors of relative abundances of *nifH* and *nirS*. However, there was generally more variation in all measured environmental variables among than within streams. These results suggest

that spatial heterogeneity in environmental variables among streams is important to the occurrence of N_2 fixation and denitrification and the overall N_2 flux of these ecosystems.

2.2 Introduction

Patches, or spatially-related areas that control ecosystem structure and function, are created by stream heterogeneity (Pringle et al. 1988). The framework of patch dynamics has inspired the study of hot spots, which are patches that show high reaction rates relative to the surrounding spatial matrix (McClain et al. 2003). Hot spots can have disproportionate contributions to ecosystem nutrient fluxes even if the overall, average conditions do not favor a particular biogeochemical process (McClain et al. 2003, Groffman et al. 2009, Pinay et al. 2015). Hence, hot spots may facilitate the co-occurrence of biogeochemical processes assumed to be mutually exclusive, like N_2 fixation and denitrification.

N_2 fixation and denitrification are important N cycle processes that together control net N_2 flux in ecosystems (Fulweiler and Heiss 2014). N_2 fixation is the microbial conversion of atmospheric N_2 gas into an input of biologically available N, while denitrification is the microbial conversion of nitrate into N_2 gas, removing nitrogen from the ecosystem (Schlesinger and Bernhardt 2020). Both processes have differing environmental constraints. N_2 fixation is favored when nitrate (NO_3^-) concentrations are low because N_2 fixation has significant energy costs to the organism (Grimm and Petrone 1997, Kunza and Hall 2013), while denitrification requires higher concentrations of NO_3^- to use as an oxidant (Arango et al. 2007). Small-scale variation in environmental factors can control or limit rates of N_2 fixation and denitrification. Denitrification rates vary

spatially with organic matter and temperature at the reach and sub-reach scale (Holmes et al. 1996, Groffman et al. 2005, Eberhard et al. 2018). Both N_2 fixation and denitrification rates vary on the microhabitat scale among substrate types, with higher rates of N_2 fixation on rocks and higher rates of denitrification on fine benthic organic matter (Kemp and Dodds 2002, Marcarelli and Wurtsbaugh 2009, Eberhard et al. 2018). These environmental variables create fine-scale differences in process rates that may also lead to hot spots of the processes in stream ecosystems, yet these processes are rarely studied simultaneously due to assumptions that reach-average conditions are most important for determining process rates (Marcarelli et al. 2008, Eberhard et al. 2018).

Variation in environmental conditions can also lead to spatial variability in microbial assemblages, including the microbes responsible for N_2 fixation or denitrification (McClain et al. 2003). Spatial heterogeneity in dissolved inorganic nitrogen (DIN) and NO_3^- concentrations have been shown to affect the distribution of N_2 fixing cyanobacteria in stream reaches (Dent and Grimm 1999, Henry and Fisher 2003). Through the use of 16S rRNA sequencing and qPCR that targets genes associated with enzymes involved in each process (nitrogenase (*nifH*) with N_2 fixation and nitrite reductase (*nirS, nirK*), nitrate reductase (*narG*), and nitrous oxide reductase (*nosZ*) with denitrification) studies have shown that variation in environmental factors can be correlated to changes in the microbial community and their resulting biogeochemical potential (Zehr and Capone 2021, Wallenstein et al. 2006). For example, Wakelin et al. (2008) showed that the spatial proximity of sediment to a point nutrient source from a wastewater treatment plant affected the abundance of *narG* in stream sediments. Ambient nutrient and iron concentrations also can affect the abundance of N_2-fixing microbes,

nifH, and N_2 fixation activity (Santos-Caton 2007, Larson et al. 2018). Gene abundance of *nirK* and *nirS* can be correlated to changes in ash free dry mass content of sediment, ambient N concentrations, and NO_3^- uptake (Knapp et al. 2009, Graham et al. 2010). The abundance of the gene *nosZ* can be correlated to changes in organic matter concentrations and temperature in streams, but not necessarily to denitrification rates (Baxter et al. 2012, Baxter et al. 2013). The presence of a gene in a microbial community does not necessarily mean the process the gene is involved in is actually occurring. However, spatial variability in the presence of these genes could lead to spatial variability in the potential for biogeochemical processes.

The goal of this study was to evaluate how the spatial heterogeneity of environmental variables in patches facilitate hot spots of N_2 fixation, denitrification and microbial gene abundances in stream ecosystems. First, we evaluated whether hot spots of N_2 fixation and denitrification were present in study reaches by comparing process rates measured within patches to reach-average rates. Second, we hypothesized that patches with more light availability and lower dissolved inorganic nitrogen (DIN) concentrations would have higher rates of N_2 fixation and *nifH* gene abundance, while patches with more organic matter, lower hyporheic dissolved oxygen concentrations, and higher DIN concentrations would have higher rates of denitrification and *nirS* gene abundance. Finally, we evaluated how habitat complexity was related to the contributions of N_2 fixation and denitrification at the whole-reach scale. We hypothesized that reaches with more patches would have more balanced fluxes of N_2 at the reach-scale, as they would facilitate both removal of N_2 via fixation and creation of N_2 by denitrification.

2.3 Methods

2.3.1 Study Area

This study was conducted in 3 streams in Michigan and 4 streams in Idaho that had a gradient of nitrate (NO_3^-) concentrations among streams and high variability in substrate cover within streams (Table 2.1). The 3 Michigan streams (Gratiot, Pilgrim, and McGunn) are all located on the Keweenaw Peninsula (237 m.a.s.l) of Michigan's Upper Peninsula, and are tributaries of Lake Superior or the Keweenaw Waterway. The Pilgrim River had a mean daily discharge over a 4-year period of 0.91 m^3/s (USGS 04043016 Pilgrim River at Paradise Road), the Gratiot River had a mean daily discharge over a 1-year period of 0.34 m^3/s (USGS 04040260 Gratiot River at 5 Mile Point Road), and McGunn Creek had a spot measurement discharge of 0.21 m^3/s on August 16th, 2016 (Table 2.2). This region received ~ 107 cm of rain and ~ 389 cm of snow in 2017 (NOAA online weather data). Of the 4 Idaho streams, 3 streams (South Fork Mink Creek, Upper Portneuf, and Gibson Jack) were located in the Portneuf River watershed and 1 stream (Diggie Creek) was located in the Snake River watershed, both located near Pocatello, Idaho. The Portneuf River watershed drains a 3,445 km^2 basin (elevation 1,330 to 2,823 m.a.s.l) and the Snake River watershed drains a 280,000 km^2 basin (elevation 109 to 2,806 m.a.s.l). Both watersheds are located in a semi-arid region that receives approximately 30 cm of rainfall annually and the Portneuf River is dependent on the underlying aquifer and snowmelt runoff from surrounding mountains for water (Minshall and Andrews 1973). In 2017, the Pocatello area received ~178 cm of total snow accumulation (NOAA online weather data). The annual mean discharge of the Portneuf River measured at Pocatello ranged from 3.7 – 9.7 m^3/s over the last ten years (USGS

Water Resources, Station 13075500). South Fork Mink Creek had a spot discharge measurement of 0.01 m^3/s on July 23rd, 2016, Gibson Jack had a measurement of 0.12 m^3/s on July 21st, 2017, and Diggie Creek had a measurement of 0.42 m^3/s on August 8th, 2016 (Table 2.2). In both regions, discharge is highly variable with spring snowmelt. In the Portneuf region, except for Diggie Creek where flows are more stable, discharge begins to rise in late March/early April as temperatures rise and snowmelt occurs in higher elevations, reaching a peak in late May/early June and declining to baseflow discharge in late July (Marcarelli et al. 2010). In the Keweenaw, the annual hydrograph can be highly variable in timing and duration with a general trend of increased and peak discharge occurring somewhere between late March and early May due to snowmelt and declining to baseflow by July, with variation in summer and fall associated with storm events (Meingast et al. 2020).

Table 2.1: Environmental characteristics for the 7 sampling streams in MI and ID arranged in alphabetical order by region. NO_3 (nitrate), NH_4 (ammonium), DIN (dissolved inorganic nitrogen), SRP (soluble reactive phosphorus), TDP (total dissolved phosphorus), DIN:TDP, Canopy Cover, and Patch Depth are all average stream reach values collected in 2016 or 2017 and presented with standard deviation in parentheses.

State	Stream	NO_3 (mg/L)	NH_4 (mg/L)	DIN (mg/L)	SRP (mg/L)	TDP (mg/L)	DIN:TDP	Canopy Cover (%)	Patch Depth (cm)
MI	Gratiot	0.063 (± 0.012)	0.001 (± < 0.001)	0.063 (± 0.012)	0.008 (± 0.004)	0.008 (± 0.004)	8.97 (± 2.45)	43.5 (± 25.1)	30.6 (± 11.6)
	McGunn	0.797 (± 0.128)	0.010 (± 0.002)	0.806 (± 0.127)	0.017 (± 0.003)	0.022 (± 0.003)	37.57 (± 9.74)	82.7 (± 6.56)	13.3 (± 5.09)
	Pilgrim	0.072 (± 0.034)	0.009 (± 0.001)	0.081 (± 0.034)	0.008 (± 0.007)	0.007 (± 0.002)	12.35 (± 5.36)	54.4 (± 22.9)	27.3 (± 9.15)
ID	Diggie Creek	0.703 (± 0.125)	0.004 (± 0.001)	0.707 (± 0.125)	0.001 (± 0.001)	0.011 (± 0.005)	97.51 (± 67.20)	0.30 (± 0.70)	51.1 (± 21.3)
	Gibson Jack	0.246 (± 0.034)	0.004 (± 0.001)	0.249 (± 0.034)	0.011 (± 0.002)	0.019 (± 0.002)	13.28 (± 1.72)	86.8 (± 16.9)	12.7 (± 6.58)
	South Fork	0.051 (± 0.024)	0.014 (± 0.005)	0.065 (± 0.025)	0.009 (± 0.007)	0.124 (± 0.065)	0.723 (± 0.518)	65.2 (± 22.4)	10.3 (± 6.19)
	Upper Portneuf	0.741 (± 0.026)	0.006 (± 0.001)	0.747 (± 0.026)	0.020 (± 0.002)	0.017 (± 0.006)	51.50 (± 14.35)	2.45 (± 3.72)	44.8 (± 11.9)

Table 2.2: Environmental characteristics for the 7 sampling streams in MI and ID arranged in alphabetical order by region. Velocity, DO (dissolved oxygen) Surface and Hyporheic, AFDM (ash free dry mass) for N_2 fixation and denitrification (Denit.) chambers, and Temperature (Temp.) are all average stream reach values collected in 2016 or 2017 and presented with standard deviation in parentheses. Discharge data was collected in previous studies and the total reach area sampled was calculated from habitat maps made of each stream during this study.

State	Stream	Velocity (m^3/s)	DO Surface (mg/L)	DO Hyporheic (mg/L)	AFDM N_2 Fixation (g/m^2)	AFDM Denit. (g/m2)	Temp. (°C)	Discharge (L/s)	Total Reach Area Sampled (m^2)
MI	Gratiot	0.52 (± 0.84)	10.5 (± 0.16)	6.52 (± 2.93)	226 (± 249)	251 (± 260)	15.8 (± 0.50)	240	2784
	McGunn	0.18 (± 0.19)	10.4 (± 0.14)	7.11 (± 2.05)	626 (± 686)	356 (± 299)	15.7 (± 0.82)	212	590
	Pilgrim	0.14 (± 0.16)	9.93 (± 0.31)	5.11 (± 1.51)	275 (± 397)	139 (± 153)	19.7 (± 0.52)	353	657
ID	Diggie Creek	0.22 (± 0.18)	14.9 (± 1.16)	6.24 (± 1.64)	636 (± 489)	746 (± 615)	17.0 (± 1.31)	N/A	2552
	Gibson Jack	0.17 (± 0.18)	10.9 (± 0.13)	7.87 (± 2.26)	586 (± 764)	487 (± 532)	11.7 (± 0.29)	121	214
	South Fork	0.17 (± 0.14)	10.2 (± 0.71)	3.55 (± 3.71)	1132 (± 1280)	1233 (± 1738)	18.6 (± 1.88)	12	544
	Upper Portneuf	0.29 (± 0.22)	13.0 (± 2.33)	5.245 (± 1.64)	979 (± 786)	1232 (± 910)	16.9 (± 1.62)	3255	654

2.3.2 Study Design

2.3.2.1 Mapping and Calculating Patch Area

On the first sampling day at each stream, a rough habitat map was made for an ~50-80 m long reach. Reach length differed among sites based on stream width and the number of pools or riffles it took to reach approximately 20 sampling patches. Grid paper and measuring tape were used to create the patch-level maps of each stream reach (Fig. 2.1, also see Appendix 1 Fig. 1A-7A).

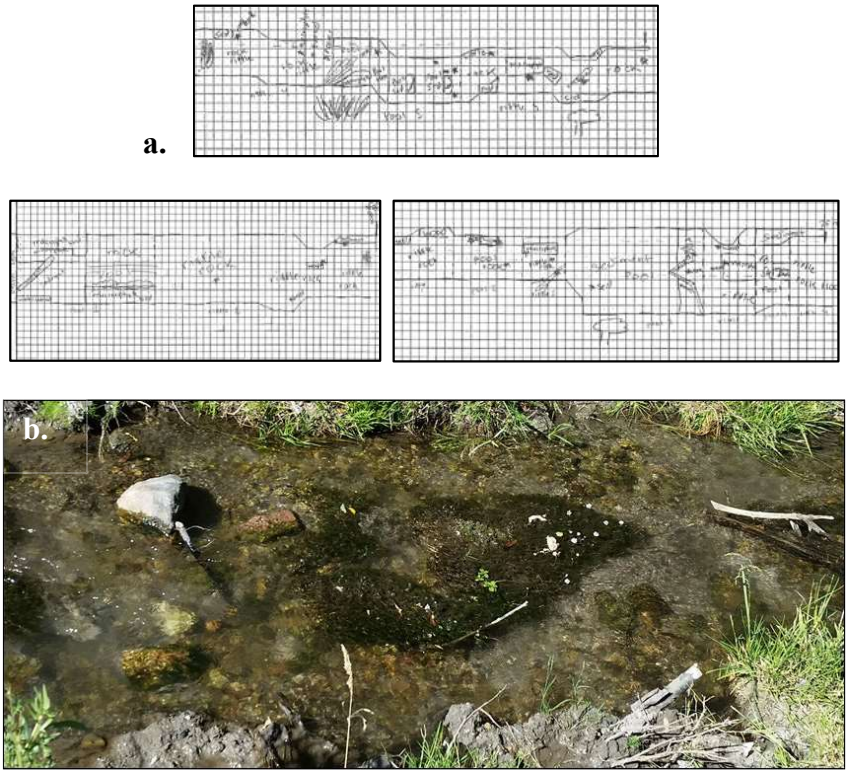

Figure 2.1a. Example of habitat maps constructed for South Fork stream reach separated in three parts. **2.1b.** Display of marked microhabitats within a riffle geomorphic unit in South Fork. From left to right there is a rock, macrophyte and sediment microhabitat.

Visual assessment in each stream reach was used to break down the reach into pools and riffles first. Then, substrate type within each pool and riffle was used to assign

patches. Patch area was later calculated by counting each cell of grid paper occupied by a given patch and then multiplying it by the cell area in m^2.

2.3.2.2 Substrate Collection

N_2 fixation and denitrification rates were measured by acetylene reduction and acetylene block, respectively, mid-day during peak hours of sunlight. Each stream was sampled over two days due to the number of patches in each stream. Both process rates were measured on the same sampling day for each patch. The two sampling days per stream were typically back to back, but in some cases were not due to inclement weather. Chambers used for these techniques varied by substrate type. 2-L polycarbonate food storage containers were used for rock and larger macrophyte substrate (Gettel et al. 2007, Eberhard et al. 2018). The chamber lids were sealed airtight with a Viton o-ring, and were fit with a 13x20 mm septa for sample collection. For sediment, wood, and smaller macrophyte substrate, chambers were made from pint size glass mason jars and lids were similarly fit with an airtight sampling septa.

Rock substrate was collected in each patch by haphazardly sampling rocks from the study area and placing them in the polycarbonate chamber until the bottom was covered. Sediment substrate was collected haphazardly from sediment patches within each stream using a 7 cm diameter suction corer to collect ~200-400 mL of sediment that was then placed into the mason jars. Macrophytes were collected using chamber lids to approximate surface area of macrophyte to sample, then pulling from the root and placing in chambers. Wood was collected by haphazardly sampling wood from wood patches until the bottom of a mason jar was mostly covered.

For each patch there were 1-4 sample chambers and 1-4 blank chambers, with each sample chamber having a paired blank chamber. The number of sample and blank chambers depended on the size of a patch in a stream. Some streams had very large patches of a single substrate type, so we collected replicate samples in these larger patches. For example, one sediment patch in Diggie Creek was 390 m^2 that was represented in the data with 2 replicates, while the largest sediment patch in South Fork was 59 m^2 and had 1 replicate (see Appendix 1A, 2A). The blank chambers were set up to simulate an environment with minimal N_2 fixation or denitrification to control for chamber effects. Materials used for the blanks were selected based on their relative specific heats to mimic the specific heats of incubated substrates to correct for changes in temperature. Rocks found on the shore near the stream were used for blanks for stream rocks, and stream water was used as a blank for sediment, wood, and macrophyte substrates.

2.3.2.3 N_2 Fixation

N_2 fixation rates were measured using acetylene reduction (Capone 1993, Dodds et al. 2017). An acetylene-filled balloon was added to each chamber. Chambers were filled with stream water and sealed underwater, then balloons were popped with a needle through the sampling septum to introduce a 20% acetylene headspace. Chambers were then shaken for approximately 20 seconds to equilibrate the gas dissolved in the water with that in the headspace. Initial gas samples were collected within 5 minutes of sealing the chambers. Chambers were placed in the stream for a 2-hour incubation to maintain ambient stream temperatures. Chambers were shaken again to equilibrate and then final

samples were collected. All gas samples were placed into evacuated 9-mL serum vials and kept in the dark until analyzed. Ethylene concentrations were measured using a SRI 8610C gas chromatograph equipped with a Hayesep T column, He carrier gas, and a flame ionization detector. The column oven was set to 40 °C. To obtain N_2 fixation rates, ethylene concentrations in the chambers were compared to 100 ppm ethylene standards (Matheson Tri Gas). N_2 fixation rates were calculated following Capone (1993) and Dodds et al. (2017), then converted to µg of N assuming a ratio of 3 mols of ethylene produced for every 1 mol of N_2 gas potentially fixed (Capone 1993).

2.3.2.4 Denitrification

Denitrification rates were measured using the acetylene block method (Groffman et al. 2006). Chloramphenicol was used to suppress additional protein synthesis during the incubation in all chambers. We measured nutrient-amended, potential rates because most previous stream studies have used this method and we wanted to be able to compare estimates to these studies, and because this method is quick and easy to run with a large number of replicates to estimate rate variability. Moreover, the acetylene block method also inhibits nitrification, so measuring without amendments of nitrate can underestimate denitrification rates (Dodds et al. 2017). However, the chambers were not sparged with nitrogen or helium to create anoxic conditions. Each chamber received 0.62 g L^{-1} Glucose as a C source and 0.62 g L^{-1} $NaNO_3$ as an N source, plus chloramphenicol (2 g L^{-1}). After the amendment, acetylene was introduced, chambers were incubated, and initial and final gas samples were collected as described previously for N_2 fixation. Nitrous oxide (N_2O) concentrations were measured using a SRI 8610C gas chromatograph equipped with a

Hayesep D column, He carrier gas, and an electron capture detector. The column oven was set to 40 °C. N_2O concentrations in chambers were compared to standard concentrations of 1000 ppm N_2O (Matheson Tri Gas). Denitrification rates were calculated following Dodds et al. (2017).

2.3.2.5 Microbial sampling and qPCR analysis

Biofilm (from rock substrate) and substrate samples (sediment, wood, or macrophyte) were taken from each sample chamber in every stream patch and placed in a sterile 15 mL falcon tube. Rocks were taken out of the chamber and scrubbed, and 12 mL of scrub water was poured into the falcon tubes to collect biofilm samples. 12-mL sediment cores were taken from sediment chambers using a 10 mL syringe and placed into the falcon tubes. Wood chambers were sampled by using a pocketknife to cut off ~ 4 surface shavings from each stick in a chamber and then placed into the falcon tubes with chamber water. Macrophyte was sampled by tearing off a small part of the macrophyte and placing it in a falcon tube with the chamber water. All 15 mL falcon tubes were placed in a mobile -20 °C freezer after collection and in a -10 °C freezer upon return to the lab for storage. DNA from each sample was extracted using the Power-soil DNA Isolation Kit (Qiagen). DNA extracts were then diluted 1:10 in preparation for quantitative polymerase chain (qPCR) analysis to dilute out potential PCR inhibitors and stored in a -20 °C freezer. The 10-fold dilution of the DNA extracts was derived empirically by testing that a 1:10 dilution was still detectable compared to a 1:100 dilution. Analyzed samples for the gene *nifH* came from chambers where rates of N_2 fixation were measured and samples for the gene *nirS* came from chambers where rates

of denitrification were measured. To quantify the gene of interest for N_2 fixation, *nifH,* the primers PolF (5'-TGC GAY CCS AAR GCB GAC TC-3') and PolR (5'-ATS GCC ATC ATY TCR CCG GA-3') were used (Poly et al. 2001). The primers NirS1F (5'-CTT AYT GGC CGG CRC ART-3') and NirS3R (5'-GCC GCC GTC RTG VAG GAA-3') were used in reactions to quantify the gene of interest for denitrification, *nirS* (Braker et al. 1998). For each qPCR reaction a 20 μL mixture was made of 10 μL of PowerTrack SYBR Green Master Mix, 0.5 μL of the forward primer (20 μM), 0.5 μL of the reverse primer (20 μM), 3 μL of 1:10 dilution of sample DNA, and 6 μL of PCR grade water. qPCR was performed on an Applied Biosystems Step ONE plus qPCR machine. The thermal cycling conditions for the *nifH* assay were 10 min at 95°C, then 40 cycles of denaturation at 95°C for 10 s, annealing at 60°C for 30 s, and elongation at 72°C for 30 s (Fan Lu 2013). For the *nirS* assay the thermal cycling conditions began at 95°C for 15 min, then 40 cycles of 94°C for 15 s, 55°C for 30 s, and 72°C for 30 s (Newell et al. 2016). Each assay was performed in a 96-well plate with triplicate samples of a negative control and triplicate environmental samples. Automatic analysis settings were used to determine the Cycle threshold values (C_T) of each sample. C_T values that were originally undetermined were given a C_T value of 40 to indicate in the 40 cycles the concentration of DNA in the sample did not pass the threshold. C_T values were then averaged for each patch and relativized to the median value of the respective target gene. The relativized C_T values were then multiplied by -1 to account for a greater negative difference between the median C_T value and original C_T values meaning more DNA concentration in the samples.

2.3.2.6 Substrate Analysis

To scale process rates by substrate area, all substrate material was collected and analyzed after incubations. Sediments were analyzed for ash free dry mass (AFDM), which provides an estimate of the total organic material present in a sample and is measured as the difference between the mass of the oxidized samples and the initial dry samples. AFDM samples were dried at 50°C, weighed for dry mass and then oxidized in a muffle furnace at 550°C, rewetted, and dried before a final weighing. Surface area and volume of all substrates were also measured to scale process rates for surface area. Surface area for rocks and wood was determined by weighing tracings of the sampled rocks. The weights were then compared to a standard curve to calculate area (Bergey and Getty 2006). Sediment surface area was calculated as the diameter of the corer. Macrophyte surface area was calculated as the diameter of the chamber lid. Rock volume was determined using displacement and sediment volume was determined by multiplying the surface area by average sediment core depth in the jar.

2.3.2.7 Environmental Characteristics

To test the hypothesis that variation in nutrient concentrations within patches would drive differences in N_2 fixation and denitrification process rates as well as gene abundances of *nifH* and *nirS*, we collected ~40 mL water samples from each stream patch. The water was filtered using Millipore 0.45 μm nitrocellulose membrane filters into 60 mL bottles. Samples were frozen until later laboratory analysis. NH_4^+ was analyzed using a fluorometric method (Holmes et al. 1999, Taylor et al. 2007) on a Turner Aquafluor (Turner Designs, Palo Alto California). NO_3^- samples from 2016 were

analyzed via the cadmium reduction method on an auto analyzer by the University of Michigan Biological Station Analytical Lab and in 2017 were analyzed on a SEAL AQ_2 discrete water analyzer using the AQ_2 method EPA-127-A Rev. 9. DIN concentration was then calculated by adding concentrations of NH_4^+ and NO_3^-. Soluble reactive phosphorus (SRP) and total dissolved phosphorus (TDP) samples were analyzed on a Thermo Scientific 10s UV-Vis spectrophotometer in 2016 (and 2017 for TDP) using the ascorbic acid method and molybdenum antimony colorimetric determination methods (APHA 2005). For TDP samples, an ammonium persulfate digestion was used prior to this analysis. The 2017 SRP samples were analyzed on a SEAL AQ_2 discrete water analyzer using the sed AQ_2 method EPA-155-A Rev. 0.

To further test the hypothesis regarding environmental variables as predictors of N_2 fixation and denitrification rates as well as gene abundances of *nifH* and *nirS*, we measured canopy cover (%) using a spherical densiometer in each patch (Lemmon 1956). Water velocity (m/s) in each patch was measured using a Marsh McBirney Flo-mate attached to a wading rod to measure velocity (m/s) at 0.6 × stream depth. Peizometers were installed used in each patch to measure hyporheic dissolved oxygen concentrations (DO mg/L). Peizometers were made of 5/8 in chlorinated polyvinyl chloride pipe (CPVC) with an inner diameter of 7/16 in. Each CPVC peizometer had 30 evenly spaced holes drilled over the bottom 15 cm and was plugged with a stopper at the bottom. The installation process for each piezometer followed the protocol described in Baxter et al. (2003) where a driver mechanism was used to install each piezometer into the streambed. Hyporheic and surface DO were measured with a YSI ProODO probe in ID and MI.

2.3.2.8 Data Analysis

To evaluate whether hot spots of each process were present in streams reaches we examined the patch average process rate to the reach average rate. Individual patch rate (equation a), average stream reach rate (equation b), and overall contribution for rates and area (equations c and d) were calculated as:

(a) Chamber rate (μg / m^2 / h) × Patch area (m^2) = Patch Rate (μg / h)

(b) Sum patch rates (μg / h) ÷ Total reach area (m^2) = Avg Stream Reach Rate (μg / m^2 / h)

(c) Patch rate (μg / h) ÷ Sum patch rates (μg / h) = Overall Rate Contribution (%)

(d) Patch area (m^2) ÷ Total reach area (m^2) = Overall Area Contribution (%)

For patches where there was more than one sample chamber used, an average of the chamber rates from that patch was used in equation (a). Patches with a higher average patch rate than the overall average stream reach rate for a process were considered hot spots. Hot spots were then further evaluated on if they had a high percent overall contribution and low percent total reach area.

To evaluate whether there was a direct relationship between process rates and relative gene abundances (N_2 fixation and *nifH*, denitrification and *nirS*), we used Spearman's

rank correlation in RStudio (R version 4.1.2), as the data did not follow a normal distribution.

We used predictive modeling to evaluate the hypothesis that light, DIN, organic matter, and dissolved oxygen concentration would predict rates of N_2 fixation and denitrification and abundances of the gene *nirS*, and the gene *nifH*. Predictive modeling is a mathematical process that uses known results to create and validate a model that generates predictions accurately (Kuhn and Johnson 2013). We chose predictive modeling because the expected relationships among rates and predictors in this dataset were potentially nonlinear and data collected within streams were spatially autocorrelated. Four separate models were generated with N_2 fixation rates, denitrification rates, relative nifH abundances, and relative nirS abundances as response variables. For all models the predictor variables were substrate type, patch area (m^2), canopy cover (%), depth (cm), velocity (m/s), temperature (°C), AFDM (g/m^2), NH_4^+ (µg/L), NO_3^- (µg/L), SRP (µg/L), TDP (µg/L), dissolved inorganic nitrogen (DIN, µg/L), DIN:TDP, surface dissolved oxygen (mg/L) and hyporheic dissolved oxygen (mg/L) concentrations. All predictor variables included in the models were based on a priori hypotheses and general knowledge of factors that may control these processes in streams. All data were pre-processed by centering, scaling, removing near-zero variables, and imputing missing variables using 5-nearest neighbors. Highly correlated variables were removed at a cutoff value of 85%, which resulted in NO_3^- being removed from all models. Each dataset was split into training and testing sets using stratified random sampling based on stream name, so that each set would have an even distribution of the streams sampled. 80% of the data was placed into a training set and 20% of the data was placed

into a testing set. We used bootstrap resampling methods with 10 resamples for each test model with replacement due to the small size of each dataset (114 observations, 17 variables). For each dataset we then trained a variety of regression-based models including: partial least squares, ridge regression, elastic net/lasso, neural networks, support vector machines, MARS/FDA, K-nearest neighbors, single trees, model trees/rules, bagged trees, random forest, boosted trees, and cubist (summarized in Kuhn and Johnson 2013). Using predictive modeling there is a trade-off between accurate predictability vs. direct interpretation (Kuhn and Johnson 2013), and only the simplest models may be directly interpretable. For each model, the seed was set to 100 and test set performance was evaluated. A best fit model was selected for each of the 4 response variables by finding the model with the lowest root mean square error (RMSE) and a high R^2 value. For each best fit model, we looked at the predictor variables of most importance to evaluate our hypotheses. All modeling was done in RStudio (R version 4.1.2) using the caret package (Kuhn 2019).

Finally, to evaluate the hypothesis that a stream reach with a greater number of patches would have a more balanced N_2 flux at the reach level, we compared the overall average N_2 flux (calculated as reach average denitrification – N_2 fixation) in each individual stream to each stream's number of patches per total stream reach area with a Spearman's rank correlation in RStudio (R version 4.1.2).

2.4 Results

We found patch rates higher than reach average rates for both N_2 fixation and denitrification within all study streams (Fig. 2.2). Stream reach average rates ranged

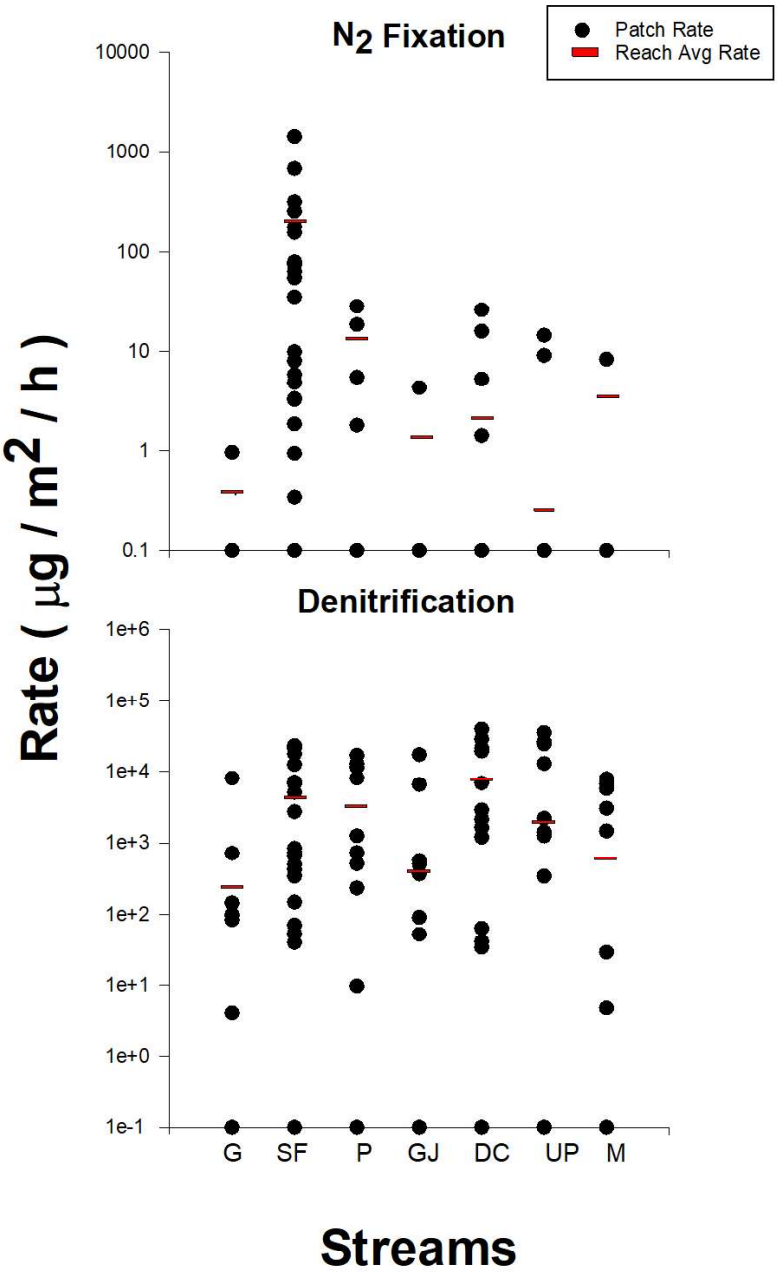

Figure 2.2: Log transformed vertical dot plots of N_2 fixation and denitrification rates of Diggie Creek (DC, n = 13), Gratiot River (G, n = 15), Gibson Jack (GJ, n = 18), McGunn (M, n = 11), Pilgrim River (P, n = 12), South Fork Mink Creek (SF, n = 33), and the Upper Portneuf River (UP, n = 12). Streams are arranged in order of lowest to highest N concentrations. Each dot denotes an average of a single microhabitat measurement of a rate. The red line denotes the reach average rate measurement for that stream. Note the Y-axis for denitrification rates is 100x that of N_2 fixation. Dots at 0.1 µg/m^2/h on the log scale denote average patch rates that were actually 0 µg/m^2/h and there can be multiple 0's for each stream.

from 226 to 7719 μg m^{-2} h^{-1} with a median of 1952 μg m^{-2} h^{-1} for denitrification and 0.25 to 202 μg m^{-2} h^{-1} with a median of 2 μg m^{-2} h^{-1} for N_2 fixation. Average denitrification rates were 100x higher than average N_2 fixation rates at the patch level. We found hot spots of both processes occurring in all study streams when comparing overall percent rate contribution, as well as overall percent area contribution for each individual patch in a stream (Fig. 2.3). For denitrification, Upper Portneuf, McGunn, and Gratiot had 1 patch that contributed 27–44% of the total denitrification for the stream and made up only 1-4% of the total reach area. Gibson Jack, Diggie Creek, and the Pilgrim had 2 patches that contributed 37-55% of the total stream denitrification and made up only 1-15% of the total stream reach area. South Fork had 1 patch that accounted for 52% of the total stream denitrification rate and < 11% of the total stream reach area. All patches that exhibited high contributions to reach-scale denitrification were sediment patches. For N_2 fixation, Diggie Creek, South Fork, and the Upper Portneuf had at least 2 patches that made up < 5-25% of the stream reach area and accounted for 25 - 60% of total N_2 fixation rate for the stream reach. Moreover, Gratiot, McGunn, and Gibson Jack all had one patch that comprised < 43% of the total stream area yet accounted for 100% of the overall N_2 fixation for the stream. The Pilgrim had one patch that accounted for 85% of the overall N_2 fixation rate for the stream and an area that comprised < 40% of the total stream area. All patches that exhibited high contributions to reach-scale N_2 fixation were rock patches.

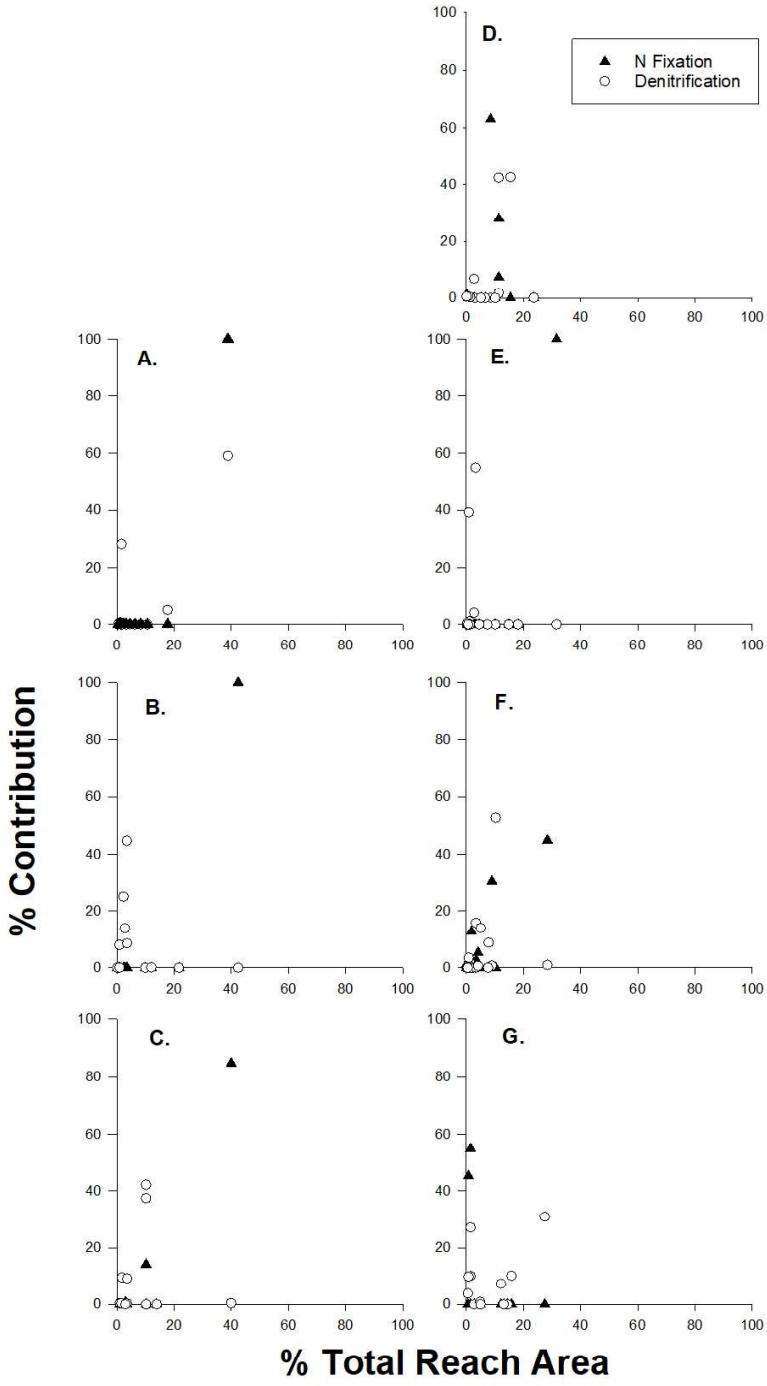

Figure 2.3: Plot of % contribution vs. % total reach area of all streams, of A. Gratiot River (G, n = 15), B. McGunn (M, n = 11), C. Pilgrim River (P, n = 12), D. Diggie Creek (DC, n = 13), E. Gibson Jack (GJ, n = 18), F. South Fork Mink Creek (SF, n = 33), and G. Upper Portneuf River (UP, n = 12). Each shape represents an individual patch from the stream.

When comparing all streams, the average relative abundance of *nifH* was highest in Diggie Creek (relativized $C_T = 3.99$ or 3.99 cycles greater than the median value). Differences of 1 CT value represent a 2-fold difference in concentration. Diggie Creek had a relative abundance of *nifH* that was ~2^8 higher concentration than the stream with the lowest average relative abundance of *nifH*, which was Gratiot (relativized $C_T = -4.95$, Fig. 2.4). For nirS, the highest average relative abundance was in the Pilgrim River (relativized $C_T = 2.91$) and lowest in South Fork (relativized $C_T = -2.94$). Relative abundances for both *nirS* and *nifH* tended to be higher on sediment substrate across streams (Appendix 1 Fig. 8A). When evaluating the relationship between *nifH* relative abundances and N_2 fixation rates, there was no significant correlation using Spearman's correlation coefficient (p-value = 0.99, $\rho < 0.01$). There was a significant relationship between *nirS* relative abundances and denitrification rates (p-value = 0.01), with a Spearman's correlation coefficient (ρ) of 0.25 indicating a positive relationship (Fig. 2.5).

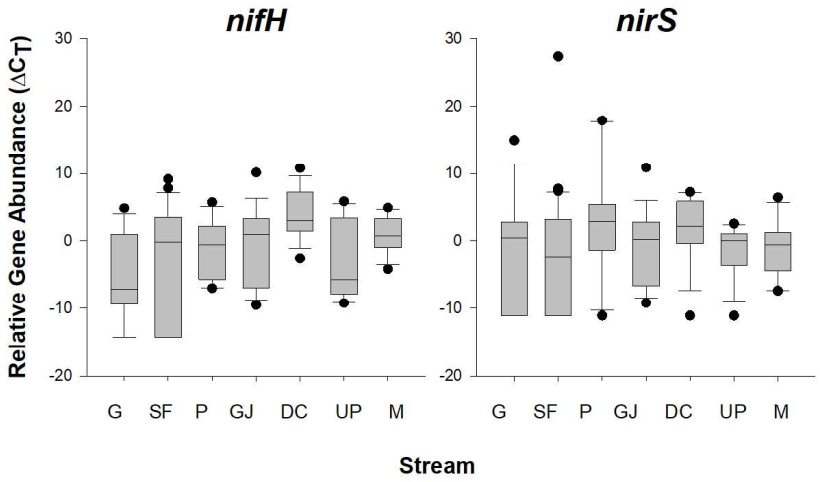

Figure 2.4: Plot of relative abundances (C_T) of *nifH* and *nirS* for all streams. Diggie Creek (DC, n = 13), Gratiot River (G, n = 15), Gibson Jack (GJ, n = 18), McGunn (M, n = 11), Pilgrim River (P, n = 12), South Fork Mink Creek (SF, n = 33), and the Upper Portneuf River (UP, n = 12). Streams are arranged in order of lowest to highest N concentrations. Note that these are relativized values to the median, so negative indicates less abundance than the median value and positive indicates more.

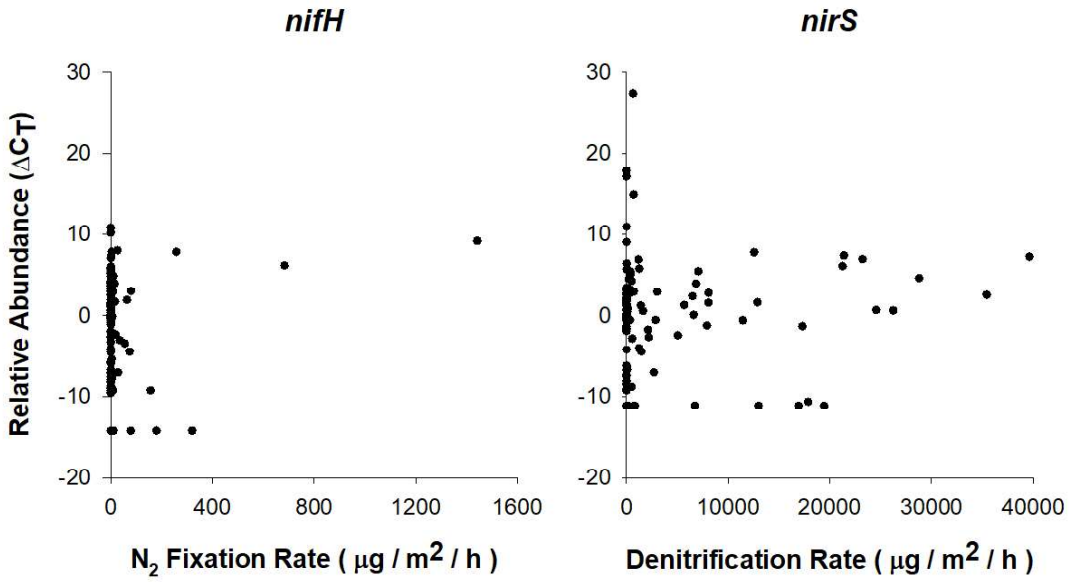

Figure 2.5: Plot of relative abundances (C_T) of *nifH* and *nirS* versus rates of N_2 fixation and denitrification respectively. Each point represents a sample from one patch in the 7 streams. Note the x-axis for denitrification is 8x that for N_2 fixation.

We found variation in environmental characteristics within and among streams. DIN concentrations varied more among streams (standard deviation (s.d.) ± 0.312 mg/L) than within streams (s.d. ± 0.012 – 0.127 mg/L) (Fig. 2.6). For NH_4^+ concentrations, the variation was largest among streams (s.d. ± 0.006 mg/L), as the s.d. within each stream was only ~ ± 0.001 mg/L (Table 2.1, Fig. 2.6). However, South Fork had a s.d. of ± 0.005 mg/L NH_4^+ within stream, which was similar variation to NH_4^+ concentrations among streams. TDP concentrations varied more within streams, particularly South Fork where s.d. was ± 0.065 mg/L compared to a s.d. of ± 0.002 – 0.005 within other streams and a s.d. of ± 0.061 mg/L among streams. DIN:TDP varied the most among streams (s.d. ± 38.5) overall, but Diggie Creek had the highest within-stream s.d. in DIN:TDP at ± 67.2. Canopy cover showed more variability among (s.d. ± 35.3 %) than within streams (s.d. ± 0.7 – 25.2 %). In Gratiot, Pilgrim, and South Fork, in-stream canopy cover varied > ±

20%, which is considerable but still less than among stream variation. (Table 2.1, Fig. 2.6). Depth had more variation within Diggie Creek (s.d. ± 22.3 cm) than among streams (s.d. ± 18.1 cm). However, depth in the rest of the streams varied less within (s.d. ± 5.1 – 11.9 cm) than among streams. Surface dissolved oxygen varied more in Upper Portneuf (s.d. ± 2.33 mg/L) than among streams (s.d. ± 1.83 mg/L). Hyporheic dissolved oxygen was more variable within South Fork (s.d. ± 3.71 mg/L) than among streams (s.d. ± 3.04 mg/L). For AFDM of both N_2 fixation and denitrification chambers there was more variation within South Fork (s.d. ± 1280 and ± 1738 g/m^2, respectively) than among streams (s.d. ± 1113 and ± 892, respectively, Table. 2.2, Fig. 2.6).

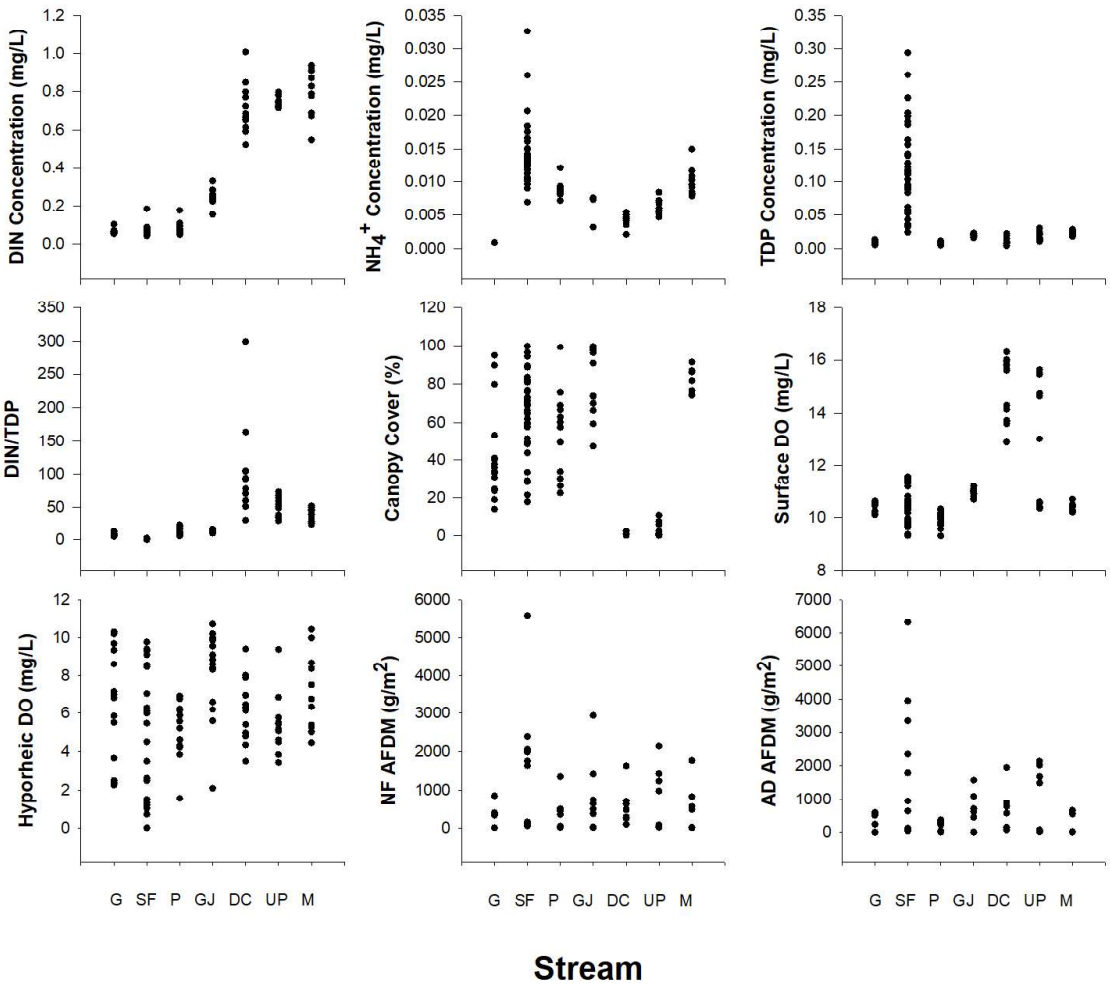

Figure 2.6: Vertical dot plots of 9 environmental variables across the 7 streams. Diggie Creek (DC, n = 13), Gratiot River (G, n = 15), Gibson Jack (GJ, n = 18), McGunn (M, n = 11), Pilgrim River (P, n = 12), South Fork Mink Creek (SF, n = 33), and the Upper Portneuf River (UP, n = 12). Streams are arranged in order of lowest to highest N concentrations. From left to right beginning on the top row the variables are DIN (dissolved inorganic nitrogen) concentration, NH_4^+ (ammonium) concentration, and TDP (total dissolved phosphorus concentration. Row 2 is DIN/TDP, canopy cover, and surface DO (dissolved oxygen). The third row is hyporheic DO, NF AFDM (N_2 fixation chambers ash free dry mass), and AD AFDM (amended denitrification chambers ash free dry mass). Note the Y-axis magnitude is different for all environmental variables.

Predictive modeling for N_2 fixation did not support our hypothesis that patches with more light availability and lower DIN concentrations could predict N_2 fixation rates. With N_2 fixation rate as the response variable, the best fit model was a support vector machine (SVM) with a RMSE of 0.10 and an R^2 of 63% (Table 2.3). SVMs are highly

Table 2.3: Predictive modeling results for the response variables N_2 fixation rates and denitrification rates. Items in bold represent the model of best fit based on the lowest RMSE. RMSE = root mean square error.

Response Variable	Model Type	RMSE	R^2
N_2 Fixation Rates	Partial Least Squares	0.46	0.27
	Ridge Regression	0.61	0.24
	Elastic Net/Lasso	0.33	0.57
	Neural Networks	0.33	0.38
	Support Vector Machines	**0.10**	**0.63**
	MARS/FDA	0.35	0.71
	K-Nearest Neighbor	0.09	0.51
	Single Trees	0.32	0.74
	Model Trees	0.28	0.81
	Bagged Trees	0.46	0.75
	Random Forest	0.26	0.82
	Boosted Trees	0.21	0.85
	Cubist	0.14	0.37
Denitrification Rates	Partial Least Squares	0.77	0.48
	Ridge Regression	0.75	0.50
	Elastic Net/Lasso	0.81	0.41
	Neural Networks	0.78	0.79
	Support Vector Machines	**0.43**	**0.84**
	MARS/FDA	0.78	0.44
	K-Nearest Neighbor	0.57	0.70
	Single Trees	0.72	0.52
	Model Trees	0.60	0.71
	Bagged Trees	0.70	0.55
	Random Forest	0.56	0.73
	Boosted Trees	0.49	0.79
	Cubist	0.64	0.66

flexible models that minimize the effect of outliers on the regression equations (Kuhn and Johnson 2013). SVMs use a kernel function to map complicated data patterns in a more simplistic way (Drake et al. 2006). TDP was the variable of most importance to the SVM

predicting N_2 fixation rates, followed by AFDM, and then DIN:TDP and surface dissolved oxygen. For *nifH* relative abundance, the best fit model was model trees with an RMSE of 0.74 and an R^2 of 26% (Table 2.4). Model trees are a type of regression tree where terminal nodes predict the outcome using a linear model and the model is created with a split variable that is associated with the largest reduction in error (Kuhn and Johnson 2013). The tree growing process continues in this way until there are no further improvements in the error rate or there are not enough samples and then each linear model undergoes simplification, potentially dropping some of the model terms (Kuhn and Johnson 2013). Variables of most importance to the model trees were AFDM, followed by NH_4^+, hyporheic dissolved oxygen, TDP, and then surface dissolved oxygen.

Predictive modeling also provided support for the hypothesis that patches with more organic matter, lower benthic dissolved oxygen, and higher DIN concentrations would have higher rates of denitrification and *nirS* gene abundance. For denitrification rate, the best fit model was a SVM with a RMSE of 0.43 and a R^2 of 84% (Table 2.3). Sediment substrate was the variable of most importance to the SVM model predicting denitrification rates, followed by surface dissolved oxygen, AFDM, and hyporheic dissolved oxygen as the variables of secondary importance. For *nirS* relative abundance, the best model was random forest with an RSME of 0.90 and an R^2 of 13% (Table 2.4). A random forest model combines the output of multiple decision trees made from resampling and bagging into a single result (Kuhn and Johnson 2013). For the random forest model the variable of most importance was depth, then DIN:TDP, followed by patch area, then NH_4^+, and then TDP.

Table 2.4: Predictive modeling results for the response variables nifH and nirS. Items in bold represent the model of best fit based on the lowest RMSE. RMSE = root mean square error.

Response Variable	Model Type	RMSE	R^2
nifH relative abundance	Partial Least Squares	0.80	0.14
	Ridge Regression	0.87	0.08
	Elastic Net/Lasso	0.84	0.05
	Neural Networks	0.95	0.05
	Support Vector Machines	1.14	0.00
	MARS/FDA	0.84	0.08
	K-Nearest Neighbor	0.79	0.15
	Single Trees	0.84	0.14
	Model Trees	**0.74**	**0.26**
	Bagged Trees	0.81	0.18
	Random Forest	0.83	0.11
	Boosted Trees	0.80	0.13
	Cubist	1.15	0.09
nirS relative abundance	Partial Least Squares	0.91	0.03
	Ridge Regression	0.95	0.00
	Elastic Net/Lasso	0.91	0.00
	Neural Networks	0.99	0.00
	Support Vector Machines	0.96	0.02
	MARS/FDA	0.88	0.03
	K-Nearest Neighbor	0.93	0.00
	Single Trees	0.88	0.08
	Model Trees	0.93	0.00
	Bagged Trees	0.92	0.01
	Random Forest	**0.86**	**0.13**
	Boosted Trees	0.90	0.07
	Cubist	0.96	0.00

When evaluating our hypothesis that a stream reach with a greater number of patches would have a more balanced N_2 flux at the reach level, we found that there was no correlation ($p = 0.59$, spearman's correlation coefficient = -0.25, Fig. 2.7). The stream with the greatest number of patches / total reach area did have one of the most balanced

overall average N_2 flux rates at 385 µg/m^2/h. However, one of the streams with the lowest number of patches / total reach area had the lowest average N_2 flux rate at 226 µg/m^2/h, when we hypothesized that it should have a higher N_2 flux. The other stream with the lowest number of patches / total reach area did have the highest average N_2 flux rate of 7716 µg/m^2/h.

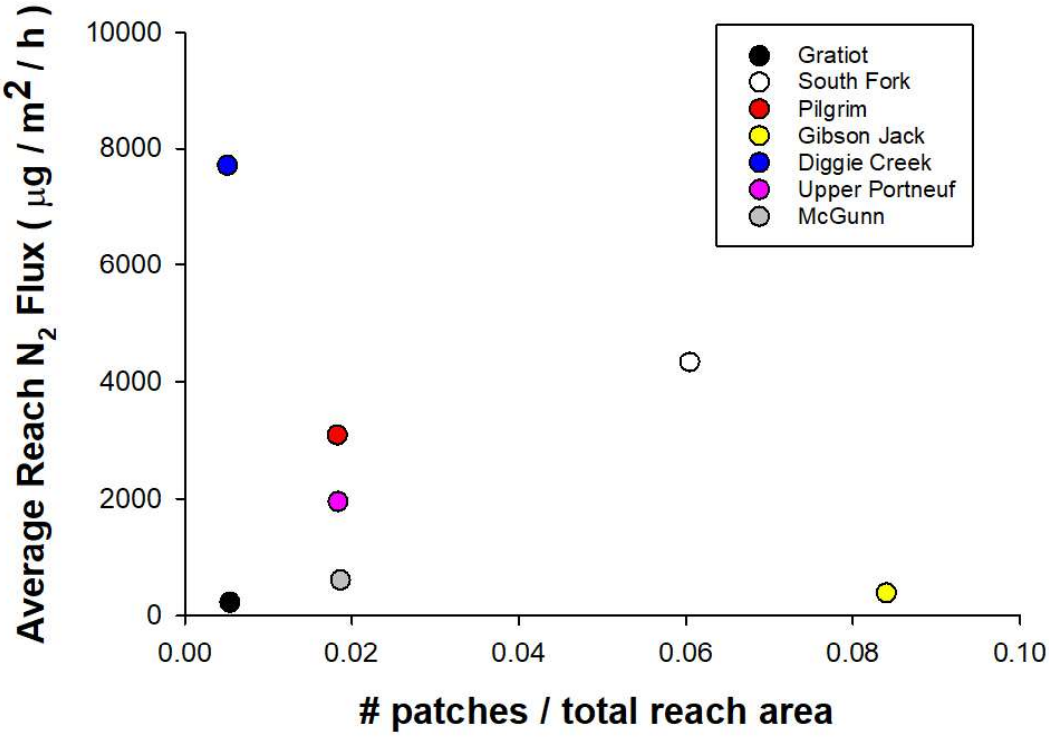

Figure 2.7: Index values for each stream. Each dot is an index value for a stream where x = the # of patches / total stream reach area and y = stream reach average N flux (denitrification – N_2 fixation). Each stream is depicted by a different color.

2.5 Discussion

We found hot spots of both N_2 fixation and denitrification were present in all study streams, with N_2 fixation hot spots occurring exclusively on rocks and denitrification hot spots occurring exclusively on sediment, while the highest relative

gene abundances of *nifH* and *nirS* were observed over all streams on sediment substrate. There was no direct correlation between N_2 fixation rates and relative *nifH* abundances, but there was a positive correlation between denitrification rates and relative *nirS* abundances. When evaluating our second hypothesis that environmental factors like light, DIN, organic matter, and dissolved oxygen concentration would predict rates and gene abundances, we found that organic matter and dissolved oxygen concentrations predicted N_2 fixation rates, denitrification rates, and *nifH* relative abundance, while sediment substrate type only predicted denitrification rates. Phosphorus (P) concentrations and the relation of N to P also were important predictors of N_2 fixation rates and the relative abundances of *nifH* and *nirS*, while variables related to N alone were only important predictors of relative abundances of *nifH* and *nirS*. All of the best-fit predictive models were black-box type, non-linear models with low to no model transparency and low interpretability (Kuhn and Johnson 2013, Visser et al. 2022), so although we can say which environmental variables are important to a model, we cannot report statistics to describe their specific interactions. Together, our findings that hot spots of both N_2 fixation and denitrification occurring across all study streams and variation in relative abundances of *nifH* and *nirS* can be predicted by models informed by variation in environmental characteristics suggests that the spatial heterogeneity of streams is important to the occurrence of both these processes and the overall N_2 flux of these systems.

Hotspots of both N_2 fixation and denitrification were observed in all study streams. The occurrence of these hot spots were substrate-specific for both processes. This is not surprising, as its been shown previously that N_2 fixation rates tend to be

higher on rock substrate and denitrification rates tend to be higher on sediment (Kemp and Dodds 2002, Marcarelli and Wurtsbaugh 2009, Eberhard et al. 2018). Moreover, the variables of importance identified in both the N_2 fixation rates model and denitrification rates model are known constraints on rates of these processes. For N_2 fixation, it has been shown that P concentrations alone and in relation to N (N:P) can be an important limiting factor to microorganisms performing the process (Elwood et al. 1981, Smith 1983, Marcarelli and Wurtsbaugh 2007). Also, oxygen concentrations are known to affect N_2 fixation rates, as the nitrogenase enzyme which performs the process is sensitive to and can be inactivated by high levels of oxygen (Gallon 1981). Organic matter availability could also affect N_2 fixation rates by increasing the availability of trace metals used in the nitrogenase enzyme like molybdenum and iron (Howarth et al. 1988). Like the model for N_2 fixation rates, the model for *nifH* relative abundances included AFDM, dissolved oxygen, and TDP as variables of importance to the model, although this model only explained 26% of the variation in abundances. In oceans, *nifH* has been shown to vary in abundance in oxygen minimum zones (Jayakumar et al. 2012) and in coastal areas can have increased abundance with increases in organic matter concentrations (Li et al. 2021). The importance of organic matter concentrations to *nifH* relative abundance could also suggest the potential for heterotrophic N_2 fixation in stream sediments, which is becoming more widely recognized in ocean and coastal sediments (Bombar et al. 2016, Li et al. 2021). Additionally, the spatial distribution of pore-water dissolved inorganic P has been shown to be positively correlated to *nifH* abundance in shallow coastal sediments (Andersson et al. 2014). However, the model for *nifH* relative abundances also included NH_4^+ as a variable of importance. N_2 fixation rates have been shown to be

depressed by NH_4^+, but also to continue at high rates at concentrations of NH_4^+ of ~200 μM (Knapp 2012). This suggests that N is of some importance to N_2 fixation at the gene level, but that variation in other environmental variables drives rates of N_2 fixation overall in streams.

For denitrification, there were different environmental variables of importance between the models for rates and *nirS* relative abundances, though the *nirS* model explained little variation. Organic matter and dissolved oxygen concentrations were important to the model predicting denitrification rates. Organic matter is important as a source of carbon (C) as an electron donor in the denitrification process (Knowles 1982, Holmes et al. 1996), and dissolved oxygen concentrations can be important in regulating denitrification as it is an anerobic process (Kemp and Dodds 2002). Depth was the most important predictor in the *nirS* model, which could be because depth can be related to the relative amount of dissolved oxygen. Differences in dissolved oxygen concentrations in micro profiles of lake sediment have been shown to affect the abundance of *nirS* (Hong et al. 2019). N and P in the forms of NH_4^+, DIN:TDP, and TDP were also important to the model predicting the relative abundance of *nirS*. Previously, the abundance of *nirS* has been shown to be significantly correlated with SRP concentrations in prairie streams (Graham et al. 2010). In coastal wetlands, NH_4^+ concentrations have been shown to affect the community structure of *nirS*-encoding denitrifiers (Gao et al. 2016). Interestingly, N availability as either NH_4^+ or DIN did not come out as an important factor in either of the models for denitrification or N_2 fixation rates alone, although NH_4^+ was an important factor for both *nirS* and *nifH* relative abundance models. This further highlights findings in previous research in streams that suggests that N alone cannot be used to predict the

occurrence of these processes (Eberhard et al. 2018). Our results demonstrate that heterogeneity in predominantly organic matter and dissolved oxygen concentrations can be used to build accurate predictive models for N_2 fixation and denitrification in streams.

We observed no direct relationship between *nifH* relative gene abundances and N_2 fixation rates, but we did observe a positive correlation of *nirS* relative gene abundances and denitrification rates. However, this correlation only had a Spearman's coefficient value of 0.22. Sequences for both genes are highly conserved among microorganisms (Zehr and Capone 2021, Knowles 1982) and just because the genes are present in the DNA does not mean that they are actively being transcribed. Also, for denitrification, there are four enzymes involved in the complete denitrification process and *nirS* only targets the nitrite reductase or the second step of the process (Knowles 1982). Using a different gene like *nosZ* that targets the nitrous oxide reductase or the last step of the complete denitrification process may give a more accurate representation of the potential denitrifiers in the samples (Kandeler et al. 2006). Moreover, denitrification is a facultative process where differences in environmental characteristics, like increases in the concentration of humic substances, can upregulate the expression of genes related to the process (Dong et al. 2017), so the relative abundances may not be an exact match for true denitrification potential. For N_2 fixation, there are also other genes that are related to the common Mo-Fe nitrogenase enzyme, which are *nifD*, *nifK*, and *nifT* (Cornejo-Castillo and Zehr 2021, Delmont et al. 2021), that by not measuring could possibly underrepresent organisms with the genetic potential for N_2 fixation. Also, there are other alternative nitrogenase enzymes, V (encoded by *vnfHDGK*) and Fe (encoded by *anfHDGK*), which are often thought as backups when Mo is limiting, but have been

shown to be maintained in multiple environments and organisms (Bothe et al. 2010, McRose et al. 2017, Bellenger et al. 2020). Moreover, horizontal gene transfer is known to have occurred in the evolutionary history of *nifH*, which complicates its lineage among microorganisms (Gaby and Buckley 2014). Therefore, for both N_2 fixation and denitrification, the relative abundances of selective genes cannot represent the full functional potential of these processes on a microbial level.

While there is evidence that spatial heterogeneity in environmental variables may affect N_2 fixation and denitrification process rates as well as relative abundances of *nirS* and *nifH* in streams, we did not specifically address how the spatial structure of these variables or patches could be affecting these processes and abundances. We observed no correlation between the relative patchiness (number of patches per total stream reach area) and the balance of the overall stream N_2 flux (denitrification – N_2 fixation rate). This could be because this test did not address the type or structure of patches, which could be very important to the overall N_2 flux. For example, in this study alone hotspots of both N_2 fixation and denitrification occurred exclusively on different substrate types, so if the patchiness of that substrate type increased or decreased in a stream that could have consequences for the overall N_2 flux. Using simulations, it has been shown that sediment heterogeneity can influence spatial patterns of denitrification hotspots in intertidal mixing zones (Heiss et al. 2020). The spatial distribution of anoxic hotspots has also been shown to affect denitrification rates in terrestrial soils (Schlüter et al. 2019). Thus, if a stream becomes more homogenous and loses its heterogeneity then that could affect the occurrence of both N_2 fixation and denitrification, which would alter the overall N_2 flux of the stream.

The overall N_2 flux of the stream could also be altered by the temporal variability in the hotspots of N_2 fixation and denitrification. Although we did not address temporal variability specifically in this study it is worth mentioning that the discussion around the occurrence of hotspots is changing. Recently there has been a reframing of examining hotspots and hot moments as distinct occurrences in biogeochemical cycling. Instead, it has been proposed that hotspots and should be thought of as "ecosystem control points" where there is both an inherent spatial and temporal component (Bernhardt et al. 2017). These ecosystem control points have disproportionate influence on the biogeochemical cycles of an ecosystem that can be altered by location in the landscape, delivery of limiting environmental variables, and the timing/duration of the presence of limiting variables (Bernhardt et al. 2017). Both N_2 fixation and denitrification have been shown to vary day-to-day in the Pilgrim River (maximum daily change of 4,390 μg $N/m^2/hr$ for denitrification and 39 μg $N/m^2/hr$ for N_2 fixation), which was one of our study streams (Nevorski 2021). In contrast, at South Fork we have observed consistently high rates of both N_2 fixation on rocks and denitrification in sediment across multiple sampling events (Eberhard et al. 2018). So, while we did observe hotspots and spatial variation in rates of N_2 fixation and denitrification that may be related to the spatial heterogeneity of limiting environmental variables, we cannot say how this variation may change with time.

In conclusion, we found that hot spots of N_2 fixation and denitrification occur across seven streams in the Midwest and Western United States, and that variability in process rates and relative gene abundances of *nirS* and *nifH* could be predicted by environmental variables on the patch scale. Variation in organic matter and dissolved oxygen concentrations were the most common variables of importance, being present in

models of N_2 fixation, denitrification, and *nifH* relative abundance. N alone was not an important predictor or rates of either process. This variation of N_2 fixation and denitrification rates on the small scale could have an effect on the overall stream reach N_2 flux. Previous studies have shown that emergent properties of small-scale heterogeneity can affect reach-level N uptake (Peipoch et al. 2016). Without consideration of spatial heterogeneity in streams, the importance of both N_2 fixation and denitrification to overall stream N dynamics could be misrepresented or misunderstood.

2.6 References

American Public Health Administration (2005). Standard Methods for examination of water and wastewater. APHA, Washington D.C.

Andersson, B., Sundbäck, K., Hellman, M., Hallin, S., & Alsterberg, C. (2014). Nitrogen fixation in shallow-water sediments: Spatial distribution and controlling factors. *Limnology and Oceanography*, *59*(6), 1932-1944.

Arango, C. P., Tank, J. L., Schaller, J. L., Royer, T. V., Bernot, M. J., & David, M. B. (2007). Benthic organic carbon influences denitrification in streams with high nitrate concentration. *Freshwater Biology*, *52*(7), 1210-1222.

Baxter, C., Hauer, F.R., & Woessner, W.W. (2003). Measuring groundwater–stream water exchange: new techniques for installing minipiezometers and estimating hydraulic conductivity. *Trans Am Fish Soc, 132(3)*, 493-502.

Baxter, A. M., Johnson, L., Edgerton, J., Royer, T., & Leff, L. G. (2012). Structure and function of denitrifying bacterial assemblages in low-order Indiana streams. *Freshwater Science*, *31*(2), 304-317.

Baxter, A. M., Johnson, L., Royer, T., & Leff, L. G. (2013). Spatial differences in denitrification and bacterial community structure of streams: relationships with environmental conditions. *Aquatic Sciences*, *75*(2), 275-284.

Bellenger, J. P., Darnajoux, R., Zhang, X., & Kraepiel, A. M. L. (2020). Biological nitrogen fixation by alternative nitrogenases in terrestrial ecosystems: a review. *Biogeochemistry*, *149*(1), 53-73.

Bergey, E.A., & Getty, G.M. (2006). A review of methods for measuring the surface area of stream substrates. *Hydrobiologia 556*:7-16.

Bernhardt, E. S., Blaszczak, J. R., Ficken, C. D., Fork, M. L., Kaiser, K. E., & Seybold, E. C. (2017). Control points in ecosystems: moving beyond the hot spot hot moment concept. *Ecosystems*, *20*(4), 665-682.

Braker, G., Fesefeldt, A., & Witzel, K. P. (1998). Development of PCR primer systems for amplification of nitrite reductase genes (nirK and nirS) to detect denitrifying bacteria in environmental samples. *Applied and environmental microbiology*, *64*(10), 3769-3775.

Bombar, D., Paerl, R. W., & Riemann, L. (2016). Marine non-cyanobacterial diazotrophs: moving beyond molecular detection. *Trends in microbiology*, *24*(11), 916-927.

Bothe, H., Schmitz, O., Yates, M. G., & Newton, W. E. (2010). Nitrogen fixation and hydrogen metabolism in cyanobacteria. *Microbiology and molecular biology reviews*, *74*(4), 529-551.

Capone, D.G. (1993). Determination of nitrogenase activity in aquatic samples using the

acetylene reduction procedure. In: Kemp P.F., Sherr, B.F., Sherr, E.B., Cole J.J. (ed) Handbook of Methods in Aquatic Microbial Ecology, Lewis Publishers, Boca Raton.

Cornejo-Castillo, F. M., & Zehr, J. P. (2021). Intriguing size distribution of the uncultured and globally widespread marine non-cyanobacterial diazotroph Gamma-A. *The ISME journal*, *15*(1), 124-128.

Delmont, T. O., Pierella Karlusich, J. J., Veseli, I., Fuessel, J., Eren, A. M., Foster, R. A., ... & Pelletier, E. (2021). Heterotrophic bacterial diazotrophs are more abundant than their cyanobacterial counterparts in metagenomes covering most of the sunlit ocean. *The ISME journal*, 1-10.

Dent, C. L., & Grimm, N. B. (1999). Spatial heterogeneity of stream water nutrient concentrations over successional time. *Ecology*, *80*(7), 2283-2298.

Dodds, W. K., Burgin, A. J., Marcarelli, A. M., & Strauss, E. A. (2017). Nitrogen transformations. In *Methods in stream ecology* (pp. 173-196). Academic Press.

Dong, S., Li, M., & Chen, Y. (2017). Inherent humic substance promotes microbial denitrification of landfill leachate via shifting bacterial community, improving enzyme activity and up-regulating gene. *Scientific reports*, *7*(1), 1-10.

Drake, J. M., Randin, C., & Guisan, A. (2006). Modelling ecological niches with support vector machines. *Journal of applied ecology*, *43*(3), 424-432.

Eberhard, E. K., Marcarelli, A. M., & Baxter, C. V. (2018). Co-occurrence of in-stream nitrogen fixation and denitrification across a nitrogen gradient in a western US watershed. *Biogeochemistry*, *139*(2), 179-195.

Elwood, J. W., Newbold, J. D., Trimble, A. F., & Stark, R. W. (1981). The limiting role

of phosphorus in a woodland stream ecosystem: effects of P enrichment on leaf decomposition and primary producers. *Ecology*, *62*(1), 146-158.

Fan, Lu. (2013). Response of diazotrophic microbial community to nitrogen input and glyphosate application in soils cropped to soybean. Ph.D. Dissertation, Auburn University, Auburn.

Fulweiler, R. W., & Heiss, E. M. (2014). (Nearly) A Decade of Directly Measured Sediment N_2 Fluxes:What Can Narragansett Bay Tell Us About the Global Ocean Nitrogen Budget?. *Oceanography*, *27*(1), 184-195.

Gaby, J. C., & Buckley, D. H. (2014). A comprehensive aligned nifH gene database: a multipurpose tool for studies of nitrogen-fixing bacteria. *Database*, *2014*.

Gallon, J. R. (1981). The oxygen sensitivity of nitrogenase: a problem for biochemists and micro-organisms. *Trends in Biochemical Sciences*, *6*, 19-23.

Gao, J., Hou, L., Zheng, Y., Liu, M., Yin, G., Li, X., ... & Sun, X. (2016). nirS-Encoding denitrifier community composition, distribution, and abundance along the coastal wetlands of China. *Applied microbiology and biotechnology*, *100*(19), 8573-8582.

Gettel, G. M., Giblin, A. E., & Howarth, R. W. (2007). The effects of grazing by the snail, Lymnaeaelodes, on benthic N2 fixation and primary production in oligotrophic, arctic lakes. *Limnology and oceanography*, *52*(6), 2398-2409.

Graham, D. W., Trippett, C., Dodds, W. K., O'Brien, J. M., Banner, E. B., Head, I. M., ... & Knapp, C.W. (2010). Correlations between in situ denitrification activity and nir-gene abundances in pristine and impacted prairie streams. *Environmental Pollution*, *158*(10), 3225-3229.

Grimm, N. B., & Petrone, K. C. (1997). Nitrogen fixation in a desert stream

ecosystem. *Biogeochemistry*, *37*(1), 33-61.

Groffman, P. M., Dorsey, A. M., & Mayer, P. M. (2005). N processing within geomorphic structures in urban streams. *Journal of the North American Benthological Society*, *24*(3), 613-625.

Groffman, P. M., Altabet, M. A., Böhlke, J. K., Butterbach-Bahl, K., David, M. B., Firestone, M. K., ... & Voytek, M. A. (2006). Methods for measuring denitrification: diverse approaches to a difficult problem. *Ecological Applications*, *16*(6), 2091-2122.

Groffman, P. M., Butterbach-Bahl, K., Fulweiler, R. W., Gold, A. J., Morse, J. L., Stander, E. K., ... & Vidon, P. (2009). Challenges to incorporating spatially and temporally explicit phenomena (hotspots and hot moments) in denitrification models. *Biogeochemistry*, *93*(1), 49-77.

Heiss, J. W., Michael, H. A., & Koneshloo, M. (2020). Denitrification hotspots in intertidal mixing zones linked to geologic heterogeneity. *Environmental Research Letters*, *15*(8), 084015.

Henry, J. C., & Fisher, S. G. (2003). Spatial segregation of periphyton communities in a desert stream: causes and consequences for N cycling. *Journal of the North American Benthological Society*, *22*(4), 511-527.

Holmes, R. M., Jones, J. B., Fisher, S. G., & Grimm, N. B. (1996). Denitrification in a nitrogen-limited stream ecosystem. *Biogeochemistry*, *33*(2), 125-146.

Holmes, R. M., Aminot, A., Kérouel, R., Hooker, B. A., & Peterson, B. J. (1999). A

simple and precise method for measuring ammonium in marine and freshwater ecosystems. *Canadian Journal of Fisheries and Aquatic Sciences*, *56*(10), 1801-1808.

Hong, P., Wu, X., Shu, Y., Wang, C., Tian, C., Gong, S., ... & Xiao, B. (2019). Denitrification characterization of dissolved oxygen microprofiles in lake surface sediment through analyzing abundance, expression, community composition and enzymatic activities of denitrifier functional genes. *AMB Express*, *9*(1), 1-10.

Howarth, R. W., Marino, R., & Cole, J. J. (1988). Nitrogen fixation in freshwater, estuarine, and marine ecosystems. 2. Biogeochemical controls. *Limnology and Oceanography*, *33*(4), 688-701.

Jayakumar, A., Al-Rshaidat, M. M., Ward, B. B., & Mulholland, M. R. (2012). Diversity, distribution, and expression of diazotroph nifH genes in oxygen-deficient waters of the Arabian Sea. *FEMS microbiology ecology*, *82*(3), 597-606.

Kandeler, E., Deiglmayr, K., Tscherko, D., Bru, D., & Philippot, L. (2006). Abundance of narG, nirS, nirK, and nosZ genes of denitrifying bacteria during primary successions of a glacier foreland. *Applied and environmental microbiology*, *72*(9), 5957-5962.

Kemp, M. J., & Dodds, W. K. (2002). The influence of ammonium, nitrate, and dissolved oxygen concentrations on uptake, nitrification, and denitrification rates associated with prairie stream substrata. *Limnology and Oceanography*, *47*(5), 1380-1393.

Knapp, C. W., Dodds, W. K., Wilson, K. C., O'Brien, J. M., & Graham, D. W. (2009). Spatial heterogeneity of denitrification genes in a highly homogenous urban stream. *Environmental science & technology*, *43*(12), 4273-4279.

Knapp, A. (2012). The sensitivity of marine N2 fixation to dissolved inorganic nitrogen. *Frontiers in microbiology*, *3*, 374.

Knowles R (1982) Denitrification. Microbiological Rev 46:43.

Kunza, L. A., & Hall Jr, R. O. (2013). Demographic and mutualistic responses of stream nitrogen fixers to nutrients. *Freshwater Science*, *32*(3), 991-1004.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26, p. 13). New York: Springer.

Kuhn, M. (2019). The caret package: Classification and regression training. *R Found Stat Comput Vienna, Austria http://topepo. github. io/caret/index. html*.

Larson, C. A., Mirza, B., Rodrigues, J. L. M., & Passy, S. I. (2018). Iron limitation effects on nitrogen-fixing organisms with possible implications for cyanobacterial blooms. *FEMS microbiology ecology*, *94*(5), fiy046.

Lemmon, P. E. (1956). A spherical densiometer for estimating forest overstory density. *Forest science*, *2*(4), 314-320.

Li, S., Twilley, R. R., & Hou, A. (2021). Heterotrophic nitrogen fixation in response to nitrate loading and sediment organic matter in an emerging coastal deltaic floodplain within the Mississippi River Delta plain. *Limnology and Oceanography*, *66*(5), 1961-1978.

Marcarelli, A. M., & Wurtsbaugh, W. A. (2007). Effects of upstream lakes and nutrient limitation on periphytic biomass and nitrogen fixation in oligotrophic, subalpine streams. *Freshwater Biology*, *52*(11), 2211-2225.

Marcarelli, A. M., Baker, M. A., & Wurtsbaugh, W. A. (2008). Is in-stream N2 fixation

an important N source for benthic communities and stream ecosystems?. *Journal of the North American Benthological Society*, *27*(1), 186-211.

Marcarelli, A. M., & Wurtsbaugh, W. A. (2009). Nitrogen fixation varies spatially and seasonally in linked stream-lake ecosystems. *Biogeochemistry*, *94*(2), 95-110.

Marcarelli, A. M., Kirk, R. W. V., & Baxter, C. V. (2010). Predicting effects of hydrologic alteration and climate change on ecosystem metabolism in a western US river. *Ecological Applications*, *20*(8), 2081-2088.

McClain, M. E., Boyer, E. W., Dent, C. L., Gergel, S. E., Grimm, N. B., Groffman, P. M., ... & Pinay, G.(2003). Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems*, 301-312.

McRose, D. L., Zhang, X., Kraepiel, A. M., & Morel, F. M. (2017). Diversity and activity of alternative nitrogenases in sequenced genomes and coastal environments. *Frontiers in microbiology*, *8*, 267.

Meingast, K. M., Kane, E. S., Coble, A. A., Marcarelli, A. M., & Toczydlowski, D. (2020). Climate, snowmelt dynamics and atmospheric deposition interact to control dissolved organic carbon export from a northern forest stream over 26 years. *Environmental Research Letters*, *15*(10), 104034.

Minshall, W.G., & Andrews, D. A. (1973). An ecological investigation of the Portneuf River, Idaho: a semiarid-land stream subjected to pollution. *Freshwater Biology*, *3*(1), 1-30.

Nevorski, K. (2021). *Denitrification and nitrogen fixation commonly co-occur but rates vary throughout the year and in different environments*. Open Access Dissertation, Michigan Technological University.

Newell, S. E., McCarthy, M. J., Gardner, W. S., & Fulweiler, R. W. (2016). Sediment nitrogen fixation: a call for re-evaluating coastal N budgets. *Estuaries and Coasts*, *39*(6), 1626-1638.

Peipoch, M., Gacia, E., Bastias, E., Serra, A., Proia, L., Ribot, M., ... & Martí, E. (2016). Small-scale heterogeneity of microbial N uptake in streams and its implications at the ecosystem level. *Ecology*, *97*(5), 1329-1344.

Pinay, G., Peiffer, S., De Dreuzy, J. R., Krause, S., Hannah, D. M., Fleckenstein, J. H., ... & Hubert-Moy, L. (2015). Upscaling nitrogen removal capacity from local hotspots to low stream orders' drainage basins. *Ecosystems*, *18*(6), 1101-1120.

Poly, F., Monrozier, L. J., & Bally, R. (2001). Improvement in the RFLP procedure for studying the diversity of nifH genes in communities of nitrogen fixers in soil. *Research in microbiology*, *152*(1), 95-103.

Pringle, C. M., Naiman, R. J., Bretschko, G., Karr, J. R., Oswood, M. W., Webster, J. R., ... & Winterbourn, M. J. (1988). Patch dynamics in lotic systems: the stream as a mosaic. *Journal of the North American benthological society*, *7*(4), 503-524.

Santos-Caton, I. R. (2007). *Abundance of nifH genes in urban, agricultural, and pristine prairie streams exposed to different levels of nitrogen loading* (Doctoral dissertation).

Schlesinger, W. H., & Bernhardt, E. S. (2020). Biogeochemistry: an analysis of global change. Academic Press.

Schlüter, S., Zawallich, J., Vogel, H. J., & Dörsch, P. (2019). Physical constraints for respiration inmicrobial hotspots in soil and their importance for denitrification. *Biogeosciences*, *16*(18), 3665-3678.

Smith, V. H. (1983). Low nitrogen to phosphorus ratios favor dominance by blue-green algae in lake phytoplankton. *Science*, *221*(4611), 669-671.

Taylor BW, Keep CF, Hall Jr. RO (2007) Improving the fluorometric ammonium method: matrix effects, background fluorescence, and standard additions. *Journal of the North American benthological society* 26:167-177.

Visser, H., Evers, N., Bontsema, A., Rost, J., de Niet, A., Vethman, P., ... & de Lange, H. J. (2022). What drives the ecological quality of surface waters? A review of 11 predictive modeling tools. *Water Research*, *208*, 117851.

Wakelin, S. A., Colloff, M. J., & Kookana, R. S. (2008). Effect of wastewater treatment plant effluent on microbial function and community structure in the sediment of a freshwater stream with variable seasonal flow. *Applied and environmental microbiology*, *74*(9), 2659-2668.

Wallenstein, M. D., Myrold, D. D., Firestone, M., & Voytek, M. (2006). Environmental controls on denitrifying communities and denitrification rates: insights from molecular methods. *Ecological applications*, *16*(6), 2143-2152.

Zehr, J. P., & Capone, D. G. (2021). *Marine Nitrogen Fixation*. Springer Nature.

3 Chapter 3: Heterogeneity in habitat and nutrient availability facilitate the co-occurrence of N_2 fixation and denitrification across wetland – stream – lake interfaces of Lakes Superior and Huron

3.1 Abstract

Great Lakes coastlines are mosaics of wetland, stream and lake habitats, characterized by a high degree of spatial heterogeneity that may facilitate the co-occurrence of seemingly incompatible biogeochemical processes. We measured nutrient limitation and rates of N_2 fixation and denitrification along transects in 5 wetland-stream-lake interfaces with different nutrient loading in Lakes Superior and Huron. We found that N_2 fixation (0 - 1,950 ug/m^2/h) and denitrification (0 - 16,536 ug/m^2/h) co-occurred in 48% of points measured across all 5 transects and across all habitat types. Denitrification rates were approx.100x higher than N_2 fixation rates where they co-occurred. N_2 fixation occurred on sediment and macrophyte substrate, while denitrification occurred mostly in sediment. Nutrient limitation of biofilms determined using nutrient-diffusing substrates at 31 transect points indicated N limitation at 32%, co-limitation of N and P at 26%, and no nutrient limitation at 42% of the points. Rates of N_2 fixation and denitrification did not differ significantly among points with differing nutrient limitations (Kruskal-Wallis p-value = 0.07 and 0.36 respectively). Predictive models for N_2 fixation and denitrification rates both included P availability as variables of importance. Denitrification models also included N and light availability, while the model for N_2 fixation included macrophyte substrate, temperature, and C availability. Heterogeneity in habitat characteristics and nutrient concentrations facilitate the co-

occurrence of N_2 fixation and denitrification across wetland – stream – lake interfaces. The potential for both processes to co-occur in coastal ecosystems highlights the need to think of these systems as more biogeochemically complex than as a simple sink of N, which is a common conceptualization of coastal wetland ecosystems.

3.2 Introduction

Wetland – stream – lake interfaces are critical locations regulating complex biogeochemical cycling (Hedin 1998, Sierzen et al. 2012, Flint and McDowell 2015). Wetlands and lakes are known to store nutrients in sediments, and wetlands can decrease outflowing concentrations of some nutrients through retention, whereas streams mainly transport material downstream (Knuth and Kelly 2011, Sierzen et al. 2012, Flint and McDowell 2015). These three aquatic habitats, though diverse, are spatially connected through cross-interface processes that alter material form and export magnitude (Kling et al. 2000, Jones 2010, Baker et al. 2016). For example, streams and wetlands are important sources of nutrients and organic matter to lakes, where they are used to support primary and secondary production (Biddanda and Cotner 2002, Dila and Biddanda 2015). Stream inflows to lakes can be hotspots of productivity and biodiversity where large amounts of organic matter and different invertebrates are delivered (Richardson et al. 2021). Lakes in watersheds have been shown to control the hydrology, temperature, and flux of nutrients to outflow streams and wetlands, which can affect metabolic processes within these downstream environments (Goodman et al. 2010, Arp et al. 2012, Epstein et al. 2013). Upstream wetlands can supply dissolved organic carbon to streams, and the presence of embedded lakes in these wetland-stream networks can influence the

hydrological delivery of these nutrients (Lottig et al. 2013). Therefore, differences in environmental variables created by spatial heterogeneity across the wetland – stream – lake interface can have consequences for ecosystem functions like biogeochemical processes within these systems.

One important variable that could facilitate hotspots of biogeochemical processes across wetland – stream – lake interfaces is nutrient limitation. There is abundant evidence that primary producers in the water column of the Great Lakes are primarily limited by phosphorus (P) (Schelske et al. 1987). However, both nitrogen (N) and/or P may limit primary producers in tributary streams and coastal wetlands of the Great Lakes, where nutrient diffusing substrate experiments have shown that increased N concentrations can increase algal standing crops and microbial activity (Allen and Hershey 1996, Wold and Hershey 1999, Cooper et al. 2016). Moreover, these experiments have also shown that in wetlands that are degraded by high N inputs, biofilms can then become P limited (Cooper et al. 2016), which creates conditions more suitable for different microorganisms to perform processes that were not favorable under N limitation. Therefore, spatial gradients in nutrient limitation across wetland – stream – lake interfaces may promote the co-occurrence of different biogeochemical processes – particularly N_2 fixation and denitrification, which have long been thought to be mutually exclusive in freshwater ecosystems. N_2 fixation is the conversion of N_2 gas into an input of biologically available N, while denitrification is the metabolic conversion of nitrate (NO_3^-) into N_2 gas, both of which are microbially-mediated in aquatic ecosystems. Together these two processes control N_2 flux, however different environmental factors favor each process. Traditionally it has been assumed that differences in N concentrations

was the major factor driving the occurrence of these processes, where N_2 fixation is favored when NO_3^- concentrations are low because the process has significant energy costs to the organism, while denitrification requires higher concentrations of NO_3^- to use as an oxidant (Grimm and Petrone 1997, Arango et al. 2007). However, the occurrence of these processes cannot be predicted by N concentrations alone consistently and are related to other environmental variables, like P and carbon (C) availability, across ecosystems (Chapter 2; Eberhard et al. 2018, Marcarelli et al. 2022).

Beyond nutrient limitation, hot spots of N_2 fixation and denitrification may be driven spatially by other environment variables across wetland – stream – lake interfaces. N and C cycling can exhibit spatial patterning with the presence of plants, water depth, organic matter, total N, and soil moisture in wetland and floodplain ecosystems (Bellinger et al. 2014, Orr et al. 2014, Wang et al. 2016). In stream ecosystems, a positive relationship between denitrification rates and organic matter has long been recognized (Holmes et al. 1996, Groffman et al. 2005, Eberhard et al. 2018), however these types of relationships have not been widely studied spatially across wetland – stream – lake interfaces (Larson et al. 2013, 2016). Hydrologic flow paths between these aquatic ecosystems can also be an important environmental variable facilitating hotspots of biogeochemical processing (Sierzen et al. 2012). Streamside wetlands and streams can have hot spots of N inputs from flow paths through terrestrial alder stands (Callahan et al. 2017). Oxbow wetlands can receive stream and storm flow that result in the wetlands being significant sinks of N (Harrison et al. 2014). Examining the spatial heterogeneity of these environmental variables across the wetland – stream – lake interface may better explain biogeochemical process dynamics across these systems.

The contributions of denitrification and N_2 fixation have been overlooked in the Great Lakes region, where wetlands are thought of as sponges of N and P via retention and removal, dissolved N concentrations are high and/or rising in these oligotrophic lake ecosystems, and P limitation of water column primary producers is common (McDonald et al. 2010, Small et al. 2014a). Since primary producers in the water column of the Great Lakes are primarily P-limited, N_2 fixation may not be expected, whereas denitrification may occur in lake and stream sediment if there is sufficient NO_3^- availability and organic matter content (Bellinger et al. 2014, Small et al. 2014b). However, some studies in the Great Lakes have shown that N_2 fixation is important despite differing levels of N concentrations and in periods of N limitation in eutrophic waters of Lake Erie, N_2 fixation rates can exceed NO_3^- and NH_4^+ uptake (Salk et al. 2018, Natwora and Sheik 2021). In a stream tributary of Lake Erie, denitrification has been shown to drive N limitation in downstream wetlands when outflow from the stream to the lake was blocked by a sand barrier (McCarthy et al. 2007). Outside of the Great Lakes region in constructed wetlands, the ratio of N:P can decrease downstream as N is permanently removed through denitrification, creating ideal conditions for N_2 fixing organisms downstream (Scott et al. 2005, 2008). Since wetlands are shallow, they may have warmer temperatures than surrounding streams and lakes that could be more conducive to organisms performing N_2 fixation, as higher temperatures have been shown to stimulate N_2 fixation activity (Marcarelli and Wurtsbaugh 2006, Welter et al. 2015). Indeed, N_2 fixing cyanobacteria and diatoms were observed in a study of N limitation in Great Lakes coastal wetlands (Cooper et al. 2016). Lakes and wetlands also have the potential for N_2 fixation through attached epiphytes on macrophytes (Finke and Seely 1978, Doyle and

Fisher 1994, Scott et al. 2005, Marcarelli and Wurtsbaugh 2009). Quantifying these processes along the full spatial continuum of the wetland – stream – lake interface could change our understanding of the importance of these two processes to the overall N cycle of Great Lakes coastal ecosystems.

The goal of this study was to evaluate how the spatial heterogeneity across a wetland – stream - lake interface controls the net N_2 flux in these ecosystems. We first hypothesized that the spatial heterogeneity of the wetland-stream-lake interface would lead to spatial variability in nutrient limitation. Secondly, we hypothesized that the spatial variability in nutrient limitation would facilitate the co-occurrence of N_2 fixation and denitrification across wetland-stream-lake interfaces, where sites with N or N+P limitation would have higher rates of N_2 fixation and sites with P limitation would have higher rates of denitrification. Finally, we hypothesized that spatial patterns of nutrients, oxygen, organic matter, and temperature would predict the occurrence of these processes. Particularly, denitrification rates would be highest where there is high organic matter and anoxic conditions (e.g., wetland, lake, and stream sediments), while N_2 fixation would occur where there are warm temperatures and low nitrate (e.g., stream microhabitats, shallow water in wetlands, by epiphytes on macrophytes in wetlands and lakes).

3.3 Methods

3.3.1 Study Area

This study was conducted in 5 wetland-stream-lake interfaces in Lakes Superior and Huron, selected to span a gradient of nutrient loading and human impact conditions (Fig. 3.1). The Sioux and Mackinac interfaces were selected as sites where we expected

low levels of human impact, while Nara was selected as a moderate level, and the Saganing and Widlfowl were selected as high impact levels. The Nara interface was in

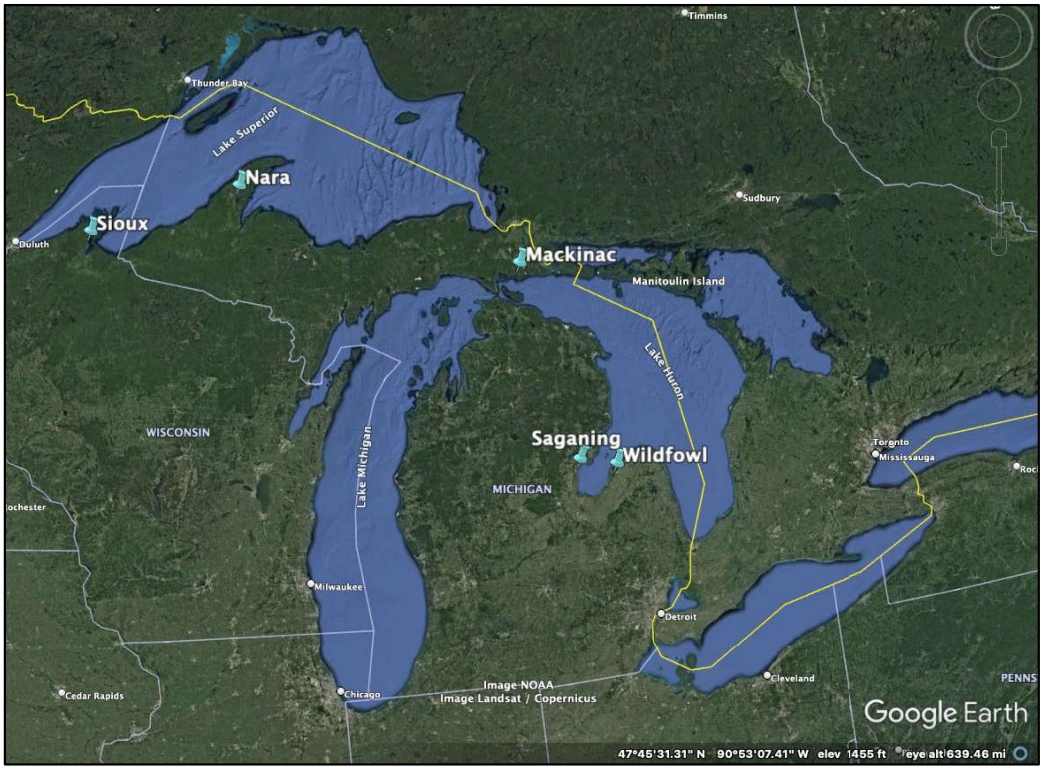

Figure 3.1: Google Earth map of the states of Michigan and Wisconsin and the 5 wetland – stream – lake interfaces sampled in this study marked by a blue pin. The 5 sites were Sioux, Nara, Mackinac, Saganing, and Wildfowl.

the Nara Nature Area in Houghton, MI that encompasses part of the Pilgrim River, a tributary to the Keweenaw Waterway, which flows into Lake Superior. Nara is managed and owned by the city of Houghton, MI. The Sioux interface was along the Sioux River in Washburn, WI which is a tributary to Lake Superior and managed by WI Department of Natural Resources. The Mackinac Bay interface was located near the Les Cheneaux Islands in northern Lake Huron and managed by the Little Traverse Conservancy. Both the Saganing and Wildfowl Bay interfaces drain into Saginaw Bay of Lake Huron. Saganing is managed by the Saginaw Bay Land Conservancy and Wildfowl is managed by the Michigan Department of Natural Resources. All interfaces were categorized as

shrub swamp and emergent marsh cover types in the Great Lakes Coastal Wetland Mapping tool developed by Bourgeau-Chavez et al. (2015). The Wildfowl and Saganing transects were the only transects noted for the presence of the invasive plant *Phragmites australis* at the time of mapping in 2015.

Table 3.1: Environmental characteristics for the 5 sampled interfaces in the years 2018, 2019, and 2020. NO_3 (nitrate), NH_4 (ammonium), DIN (dissolved inorganic nitrogen), SRP (soluble reactive phosphorus), TDP (total dissolved phosphorus), DIN:TDP, DOC (dissolved organic carbon), TDN (total dissolved nitrogen) are all average transect values presented with standard deviation in parentheses.

Year	Transect	NH4 (mg/L)	NO3 (mg/L)	SRP (mg/L)	TDP (mg/L)	DIN (mg/L)	DIN:TDP	DOC (mg/L)	TDN (mg/L)
2018	Nara	0.003 (± 0.003)	0.055 (± 0.040)	0.004 (± 0.001)	0.007 (± 0.001)	0.058 (± 0.040)	9.64 (± 7.05)	6.28 (± 1.89)	0.279 (± 0.099)
	Sioux	0.008 (± 0.005)	0.034 (± 0.052)	0.006 (± 0.003)	0.014 (± 0.004)	0.042 (± 0.049)	3.88 (± 5.76)	5.91 (± 4.82)	0.382 (± 0.368)
	Mackinac	0.006 (± 0.006)	0.022 (± 0.015)	0.005 (± 0.002)	0.003 (± 0.001)	0.028 (± 0.017)	9.08 (± 4.27)	4.70 (± 0.92)	0.182 (± 0.062)
2019	Nara	0.007 (± 0.004)	0.068 (± 0.069)	0.008 (± 0.007)	0.007 (± 0.003)	0.075 (± 0.070)	14.45 (± 15.93)	5.23 (± 2.44)	0.279 (± 0.049)
	Sioux	0.011 (± 0.010)	0.099 (± 0.096)	0.018 (± 0.000)	0.013 (± 0.007)	0.110 (± 0.095)	10.60 (± 13.59)	4.52 (± 3.47)	0.363 (± 0.243)
	Mackinac	0.008 (± 0.010)	0.027 (± 0.035)	0.018 (± 0.000)	0.005 (± 0.002)	0.035 (± 0.037)	7.30 (± 8.55)	5.96 (± 0.86)	0.242 (± 0.04)
2020	Saganing	0.008 (± 0.004)	0.018 (± 0.008)	0.004 (± 0.003)	0.017 (± 0.012)	0.026 (± 0.009)	2.16 (± 1.29)	5.74 (± 0.72)	0.431 (± 0.105)
	Wildfowl	0.038 (± 0.034)	0.223 (± 0.343)	0.010 (± 0.002)	0.016 (± 0.004)	0.260 (± 0.376)	14.69 (± 20.66)	5.85 (± 0.63)	0.584 (± 0.251)

Table 3.2 Environmental characteristics for the 5 sampled interfaces in the years 2018, 2019, and 2020. Depth, Canopy Cover, Temperature (Temp.), and ash freed dry mass (AFDM) for both N_2 fixation (NF) and amended denitrification (AD) chambers are all average transect values presented with standard deviation in parentheses. GPS Coordinates are presented in °N, °W.

Year	Transect	Depth (cm)	Canopy Cover (%)	Temp. (°C)	AFDM NF (g/m^2)	AFDM AD (g/m^2)	GPS Coordinates (°N, °W)
2018	Nara	80.7 (± 54.3)	16.7 (± 19.3)	21.5 (± 4.1)	2263 (± 2930)	2260 (± 2750)	47.10612, -88.51456
	Sioux	51.1 (± 31.4)	1.42 (± 2.08)	20.5 (± 3.3)	2242 (± 2148)	1467 (± 1434)	46.73722, -90.87925
	Mackinac	54.6 (± 51.2)	10.7 (± 23.5)	18.1 (± 3.3)	896 (± 805)	948 (± 751)	46.00581, -84.41543
2019	Nara	73.5 (± 68.9)	0.39 (± 0.73)	11.8 (± 7.0)	1301 (± 739)	1129 (± 612)	47.10612, -88.51456
	Sioux	38.9 (± 16.9)	1.39 (± 4.16)	17.7 (± 1.6)	1311 (± 1480)	723 (± 614)	46.73722,- 90.87925
	Mackinac	74.3 (± 37.8)	1.08 (± 1.40)	14.7 (± 2.5)	2267 (± 642)	1682 (± 1014)	46.00581, -84.41543
2020	Saganing	60.4 (± 22.3)	4.71 (± 6.55)	15.7 (± 2.4)	1720 (± 1247)	1237 (± 889)	43.924037, -83.904138
	Wildfowl	80.7 (± 20.3)	1.56 (± 4.68)	13.8 (± 0.99)	873 (± 557)	732 (± 202)	43.828629,- 83.392688

3.3.2 Study Design

3.3.2.1 Transect Setup

The Nara, Sioux, and Mackinac interfaces were sampled in summers 2018 and 2019, while the Saganing and Wildfowl interfaces were only sampled in summer 2020 (Table 3.3, see also Appendix 2 Fig. 9A-13A). Although we had planned to sample all 5 sites in summer 2020, limitations to field work caused by the COVID-19 pandemic prevented this. Each interface was sampled across 1-3 sampling days in each year due to the number of transect points and the duration of incubations. Sampling days in each interface were typically sequential, but in some cases, there was a day in between sampling due to inclement weather and in one case there was a 3-day intervening period between sampling dates (Table 3.3). On the first day at each interface, a transect of 8-15 points was established that encompassed the wetland-stream-lake interface (Table 3.3). Based on loose classifications of the 83 total transect points among all transects, 37 were wetland, 18 were transition zones from wetland to stream, 15 were stream, 2 were stream to lake transition zones and 11 were lake sites (Table 3.3). The number of wetland sites per transect ranged from 3 -7, the number of stream sites ranged from 0-3, and the number of lake sites ranged from 0 - 4. Wetland to stream transition zones ranged from 1 - 4 sites per transect and the number or stream to lake transition zones were either 0 or 1 in each transect. In the Saganing and Wildfowl transects there were no sites that could be strictly classified as stream or lake because the stream bed itself and nearshore lake areas were too deep to safely deploy chamber incubations given our sampling equipment.

Table 3.3: The sampling design of each interface is explained including years each transect was sampled, total transect points, total terrestrial points, total aquatic points, number of sites classified as wetland, number of sites classified as wetland to stream transition zones, number of sites classified as stream, number of sites classified as stream to lake transition zones, number of sites classified as lake, and dates sampled. Not all transects were sampled every year and some transects varied in the total number of transect points from year to year as well as the classification type of the transect points. Terrestrial points were those with no standing water and aquatic points were those with standing water, stream, or nearshore lake.

Transect	Year Sampled	# of Transect Points	# of Total Terrestrial Points	# of Total Aquatic Points	# of Wetland	# of Wetland-Stream	# of Stream	# of Stream-Lake	# of Lake	Sample Dates
Nara	2018	15	4	11	6	4	3	0	2	18-Jul, 19-Jul
Sioux	2018	9	0	9	4	1	3	0	1	14-Jul
Mackinac	2018	12	3	9	4	1	2	1	4	22-Aug, 23-Aug
Nara	2019	12	1	11	3	4	3	0	2	26-Aug, 30-Aug, 2-Sep
Sioux	2019	9	0	9	4	1	3	0	1	24-Jul, 25-Jul
Mackinac	2019	8	1	7	4	1	1	1	1	17-Jul, 18-Jul
Saganing	2020	9	0	9	7	2	0	0	0	7-Sep, 9-Sep
Wildfowl	2020	9	0	9	5	4	0	0	0	10-Sep, 11-Sep

3.3.2.2 Nutrient Limitation

To test the first hypothesis that the spatial complexity of the wetland-stream-lake interface would lead to spatial variability of nutrient limitation for primary producers, we deployed nutrient diffusing substrates (NDS) at each transect point in an interface (Tank et al. 2017). NDS were constructed using 45 mL plastic containers filled with a 2% (by weight) agar solution amended with 0.8 M N added as $NaNO_3$ (N treatment), 0.05 M P added as NaH_2PO_4 (P treatment), both (N+P treatment), or neither as a control treatment. A 25 mm porous porcelain disc (Leco Corporation, St. Joseph, MI) was placed on top of each hardened NDS for algae to grow on. At transect points, a total of 16 NDS were deployed with 4 control, 4 N, 4 P, and 4 N+P replicates. NDS were deployed at transect points two weeks prior to the first sampling day at a site. After these two weeks, the discs were collected, wrapped in aluminum foil and frozen for later analysis of algal standing crop using chlorophyll-*a*. Laboratory analysis of chlorophyll-*a* followed standard method using a Thermo Scientific 10s UV-Vis spectrophotometer and ethanol extraction (APHA 2005). NDS were only deployed in summers 2019 and 2020. NDS were deployed at all transect points except those that were too deep, had high wave action, or had no standing water (Appendix 2 Tables 1A and 2A).

3.3.2.3 N Cycling Rate Measurements

Chamber incubations were used to measure rates of N_2 fixation and denitrification in all transect points. The chambers used during these incubations varied by substrate type. 2-L polycarbonate food storage containers were used for larger macrophyte substrate (Gettel et al. 2007, Eberhard et al. 2018). The chamber lids were sealed airtight

with a Viton o-ring, and were fit with a 13x20 mm septum for sample collection. For sediment and smaller macrophyte substrate, chambers were made from pint size glass mason jars and lids were similarly fit with an airtight sampling septum. Sediment substrate was collected haphazardly from transect points using a 7 cm diameter suction corer to collect ~200-400 mL of sediment that was then placed into the mason jars. Macrophytes were collected using chamber lids to approximate surface area of macrophyte to sample, then pulling from the root and placing in chambers. For each transect point there were 1-4 sample chambers and 1-4 blank chambers, with each sample chamber having a paired blank chamber. The blank chambers were set up to simulate an environment with minimal N_2 fixation or denitrification to control for chamber effects. Stream water was used as a blank for sediment and macrophyte substrates.

N_2 fixation rates were measured using acetylene reduction (Capone 1993, Dodds et al. 2017). An acetylene-filled balloon was added to each chamber. Chambers were filled with stream water and sealed underwater, then balloons were popped with a needle through the sampling septum to introduce a 20% acetylene headspace. Chambers were shaken for approximately 20 seconds to equilibrate the gas dissolved in the water with that in the headspace, and initial gas samples were collected within 5 minutes of sealing the chambers. Chambers were placed in the stream for a 2-hour incubation to maintain ambient stream temperatures, then shaken again to equilibrate and final samples were collected. All gas samples were placed into evacuated 9-mL serum vials and kept in the dark until analyzed. Ethylene concentrations were measured using a SRI 8610C gas chromatograph equipped with a Hayesep T column, He carrier gas, and a flame

ionization detector. The column oven was set to 40 °C. To obtain N_2 fixation rates, ethylene concentrations in the chambers were compared to 100 ppm ethylene standards (Matheson Tri Gas). N_2 fixation rates were calculated following Capone (1993) and Dodds et al. (2017), then converted to µg of N assuming a ratio of 3 mols of ethylene produced for every 1 mol of N_2 gas potentially fixed (Capone 1993).

Denitrification rates were measured using the acetylene block method (Groffman et al. 2006). Chloramphenicol was used to suppress additional protein synthesis during the incubation in all chambers. We measured nutrient-amended rates because most previous stream studies have used this method and we wanted to be able to compare estimates to these studies, and because this method is quick and easy to run with a large number of replicates to estimate rate variability. Moreover, the acetylene block method also inhibits nitrification, so measuring without amendments of nitrate can underestimate denitrification rates (Dodds et al. 2017). However, the chambers were not sparged with nitrogen or helium to create anoxic conditions. Each chamber received 0.62 g L^{-1} Glucose as a C source and 0.62 g L^{-1} $NaNO_3$ as an N source, plus chloramphenicol (2 g L^{-1}). After the amendment, acetylene was introduced, chambers were incubated, and initial and final gas samples were collected as described previously for N_2 fixation. Nitrous oxide (N_2O) concentrations were measured using a SRI 8610C gas chromatograph equipped with a Hayesep D column, He carrier gas, and an electron capture detector. The column oven was set to 40 °C. N_2O concentrations in chambers were compared to standard concentrations of 1000 ppm N_2O (Matheson Tri Gas). Denitrification rates were calculated following Dodds et al. (2017).

To scale process rates by substrate area, all substrate material was collected and analyzed after incubations. Sediment and macrophyte material were analyzed for ash free dry mass (AFDM), which provides an estimate of the total organic material present in a sample and is measured as the difference between the mass of the oxidized samples and the initial dry samples. AFDM samples were dried at 50 °C, weighed for dry mass and then oxidized in a muffle furnace at 550 °C, rewetted, and dried before a final weighing. Surface area and volume of all substrates were also measured for use in scaling process rates for surface area. Sediment surface area was calculated as the diameter of the corer. Macrophyte surface area was calculated as the diameter of the chamber lid. Sediment volume was determined by multiplying the surface area by average sediment core depth in the jar and macrophyte volume was measured using the displacement method in a graduated cylinder.

3.3.2.4 Environmental Characteristics

To test the third hypothesis that variation in nutrient concentrations would predict the occurrence of N_2 fixation and denitrification process rates, we collected ~40 mL water samples from each transect point. The water was filtered using Millipore 0.45 μm nitrocellulose membrane filters into 60 mL bottles. Samples were frozen until later laboratory analysis for NO_3^-, ammonium (NH_4^+), soluble reactive phosphorus (SRP), total dissolved phosphorus (TDP), dissolved organic carbon (DOC), and total dissolved nitrogen (TDN). NH_4^+ was analyzed using a fluorometric method (Holmes et al. 1999, Taylor et al. 2007) on a Turner Aquafluor (Turner Designs, Palo Alto California). NO_3^- + NO_2 samples were analyzed on a SEAL AQ_2 discrete water analyzer using the AQ_2

method EPA-127-A Rev. 9. DIN concentration was then calculated by adding concentrations of NH_4^+ and $NO_3^- + NO_2$. SRP samples were analyzed on a SEAL AQ_2 discrete water analyzer using the AQ_2 method EPA-155-A Rev. 0. TDP samples were analyzed on a Thermo Scientific 10s UV-Vis spectrophotometer using the ascorbic acid method and molybdenum antimony colorimetric determination methods (APHA 2005). For TDP samples, an ammonium persulfate digestion was used prior to this analysis. DOC and TDN samples were run on a Shimadzu TOC-L_{CPH} analyzer with TNM-L module in the AQUA lab at Michigan Tech.

To further test our hypothesis regarding environmental variables as drivers of process rates we measured depth at each transect point. Canopy cover (%) was also measured at each transect point using a spherical densiometer (Lemmon 1956).

3.3.2.5 Statistical Analysis

To evaluate our first hypothesis that spatial heterogeneity of the wetland-stream-lake interface would lead to spatial variability in nutrient limitation, we used a two-way analysis of variance (ANOVA) with N and P as factors to test whether chlorophyll *a* concentrations were significantly different (p-value ≤ 0.05) among NDS treatments at each transect point (Tank and Dodds 2003, Tank et al. 2017). Single nutrient limitation was indicated if just one of the individual treatments (N or P) indicated a positive response, but the interaction term of the ANOVA was not significant. Co-limitation by N and P was determined when either both individual treatments indicated a positive response, the interaction term of the ANOVA was significant, or if both the interaction term of the ANOVA was significant and one of the individual treatments indicated a

positive response. No significant terms (p-value ≤ 0.05) indicated no nutrient limitation. The ANOVAs were performed in RStudio (R version 4.1.2).

To evaluate the second hypothesis that spatial variability in nutrient limitation would facilitate the co-occurrence of N_2 fixation and denitrification across wetland-stream-lake interfaces we performed Kruskal-Wallis tests in RStudio (R version 4.1.2), as the data did not follow a normal distribution. We ran two separate tests with N_2 fixation or denitrification rates as the response variable and nutrient limitation status as the predictor variable. To evaluate if rates of N_2 fixation and denitrification varied significantly among the different habitat types of the interfaces (wetland, wetland to stream transition zone, stream, stream to lake transition zone, and lake) we also performed Kruskal-Wallis tests.

To evaluate the third hypothesis that spatial patterns of nutrients, light availability, and organic matter would predict the occurrence of these processes we used predictive modeling. Since the dataset for this study was nonlinear and spatially autocorrelated, we chose to use predictive modeling to assess our data. Predictive modeling is a mathematical process that uses known results to create and validate a model that generates predictions accurately (Kuhn and Johnson 2013). Two separate models were generated with N_2 fixation rates and denitrification rates as response variables. For all models, the predictor variables were substrate type, canopy cover (%), depth (cm), temperature (°C), AFDM (g/m^2), NH_4^+ (mg/L), $NO_3^- + NO_2$ (mg/L), SRP (mg/L), TDP (mg/L), dissolved inorganic nitrogen (DIN, mg/L), DIN:TDP, DOC (mg/L), and TDN (mg/L) concentrations. All predictor variables included in the models were

based on a priori hypotheses and general knowledge of biogeochemistry. Nutrient limitation was not used as a predictor variable because this data was only collected in 2019 and 2020 and the models used process rates from 2018, 2019, and 2020 as response variables. 9 total transect points were removed from the data matrix because they were categorized as terrestrial or with no standing water, so they did not have the full suite of predictors collected at those sites. All data were pre-processed by centering, scaling, removing near-zero variables, and imputing missing variables using 5-nearest neighbors. Highly correlated variables were removed at a cutoff value of 85%, which resulted in DIN being removed from all models. The variables NO_3^-+ NO_2 and macrophyte substrate also were indicated as variables with > 85% correlation, but ultimately were kept in the models as N concentrations and substrate type are important variables to N cycling processes. Each dataset was split into training and testing sets using stratified random sampling based on stream name, so that each set would have an even distribution of the streams sampled. 80% of the data was placed into a training set and 20% of the data was placed into a testing set. Replacement was used due to the small size of each dataset (105 observations, 11 variables) and we used bootstrap resampling methods with 10 resamples for each test model. For each dataset we then trained a variety of regression-based models including: partial least squares, ridge regression, elastic net/lasso, neural networks, support vector machines, MARS/FDA, K-nearest neighbors, single trees, model trees/rules, bagged trees, random forest, boosted trees, and cubist (summarized in Kuhn and Johnson 2013). For each model the seed was set to 100 and test set performance was evaluated. A best fit model was selected for each of the 2 response variables by finding

the model with the lowest root mean square error (RMSE) and a high R^2 value. For each best fit model, we then looked at the predictor variables of most importance to that model to evaluate our hypotheses. Using predictive modeling there is a trade-off of accurate predictability and direct interpretation (Kuhn and Johnson 2013). All data analysis for the models was done in RStudio (R version 4.1.2) using the caret package (Kuhn 2019).

3.4 Results

3.4.1 Environmental Characteristics

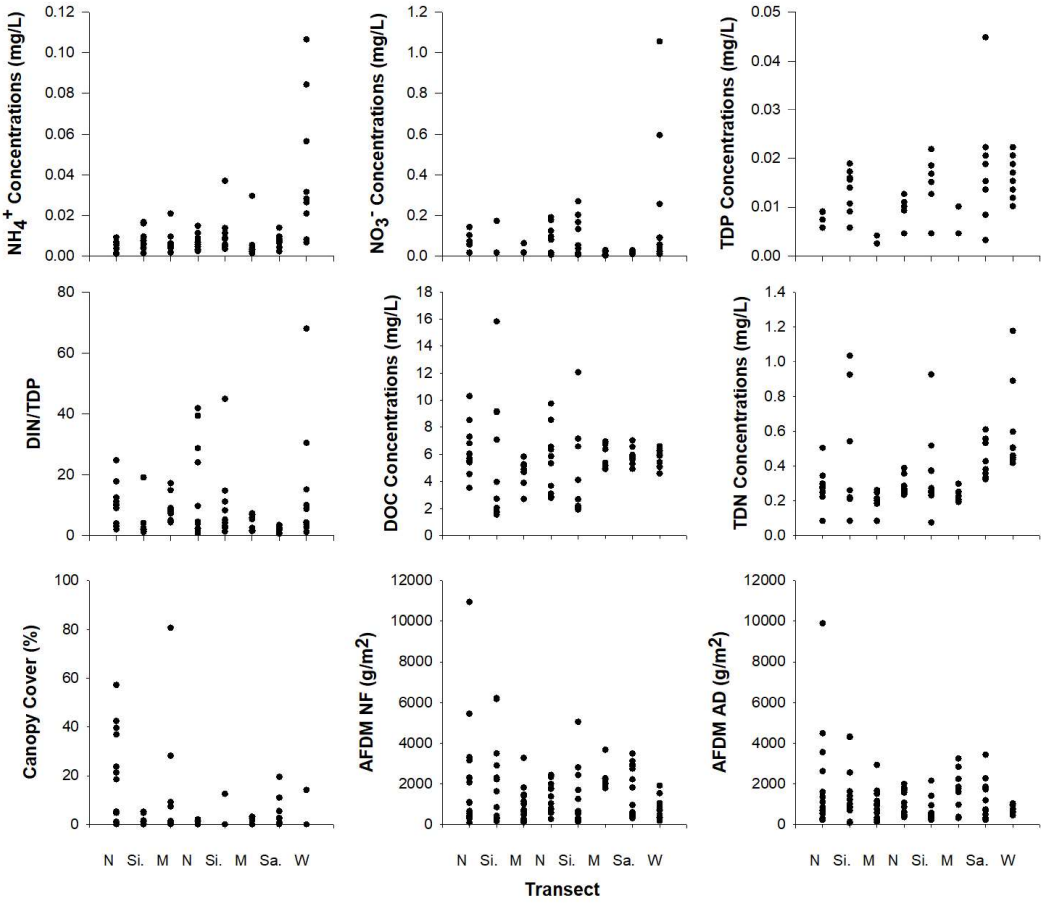

Figure 3.2: Vertical dot plots of 9 environmental variables across for each transect and study year. Nara 2018 (N, n = 11), Sioux 2018 (Si., n = 9), Mackinac 2018 (M, n = 9), Nara 2019 (N, n = 11), Sioux 2019 (Si., n = 9), Mackinac 2019 (M, n = 7), Saganing 2020 (Sa., n = 9) and Wildfowl 2020 (W, n = 9). From left to right beginning on the top row the variables are NH_4^+ (ammonium) concentration, NO_3^- (nitrate) concentration, and TDP (total dissolved phosphorus concentration. Row 2 is DIN/TDP, DOC (dissolved organic carbon) concentration, and TDN (total dissolved nitrogen) concentration. The third row is canopy cover, NF AFDM (N_2 fixation chambers ash free dry mass), and AD AFDM (amended denitrification chambers ash free dry mass). Note the Y-axis magnitude is different for all environmental variables.

We observed variation in environmental characteristics across all the transects. Of the 114 chamber measurements across all transects, 82 were on sediment substrate and 32

were on macrophyte substrate. Mackinac 2018 and Saganing had the highest number of overall macrophyte samples (n = 7, n = 8, respectively), while all other transects had 2 – 4 macrophyte samples. NH_4^+ concentrations were more variable among transects with a s.d. of ± 0.017 mg/L than within the transects (s.d. ± 0.003 – 0.010), except for Wildfowl which had a s.d. of ± 0.034 mg/L (Table 3.1, Fig. 3.2). NO_3^- + NO_2 concentrations were also more variable among transects (s.d. ± 0.143 mg/L) than within transects (s.d. ± 0.008 – 0.096), except for Wildfowl which had a s.d. of ± 0.223 mg/L. SRP concentrations tended to be more variable among transects (s.d. ± 0.006 mg/L) than within transects (s.d ± 0.000 – 0.007). TDP concentrations also tended to be more variable among transects (s.d. ± 0.007 mg/L) than within transects (s.d. ± 0.001 – 0.012, Fig. 3.2). DIN concentrations were the most variable in Wildfowl with a s.d. of ± 0.376 mg/L, but for the rest of the transects the variation within (s.d. ± 0.009 – 0.095 mg/L) was less than the variation among them (s.d. ± 0.157 mg/L). DIN:TDP varied among streams with a s.d of ± 11.9, DOC concentrations varied among streams with a s.d of ± 2.36 mg/L, and TDN concentrations varied among streams with a s.d of ± 0.21 mg/L, but a few transects had higher variability of DIN:TDP, TDN, and DOC within (Table 3.1, Fig. 3.2). Canopy cover was generally more variable among transects than within (s.d. ± 13.6 %), except for Nara 2018 (s.d. ± 19.3 %) and Mackinac 2018 (s.d. ± 23.5 %, Table 3.2, Fig. 3.2). Temperature was more variable among transects with a s.d. of ± 4.7 °C than within transects (s.d. ± 1.0 – 4.1 °C), except for Nara 2019 (Table 3.2, Figure 3.2) AFDM for N_2 fixation and denitrification chambers were more variable among transects (s.d. ± 1640

g/m^2 and ± 1332 g/m^2, respectively) but the transects of Nara and Sioux had more variation in AFDM within (Table 3.2, Fig. 3.2).

3.4.2 Nutrient Limitation

We observed spatial variability in nutrient limitation of biofilms across 4 of the 5 interfaces determined using NDS (Fig. 3.3). Of all 31 transect points measured, N limitation was indicated at 32%, co-limitation of N and P was indicated at 26%, and no nutrient limitation at was indicated at 42% of the points. At the Nara, Wildfowl, and Saganing transects, we observed a range of nutrient limitation responses, with N limitation, P limitation and co-limitation of N and P at different points along the transects. In the Nara transect, 2 sites indicated significant N limitation and 4 indicated significant co-limitation of N and P. In the Wildfowl transect 4 sites were significantly N limited and one was significantly co-limited by N and P. In the Saganing transect 2 sites were significantly co-limited by N and P. In contrast to these 3 interfaces, at the Sioux transect, only N limitation was observed at 4 sites, while 4 sites showed no nutrient limitation. No nutrient limitation data were available from the Mackinac transect because most NDS were lost due to high-water levels and storms.

Transect	Transect Number	N effect	P effect	Interaction NxP
Nara 2019	1	p-value < 0.01		p-value = 0.03
	2	p-value < 0.01		
	3	p-value < 0.01		
	4	p-value = 0.01	p-value = 0.02	
	6	p-value < 0.01	p-value = 0.05	p-value = 0.03
	7			
	11	p-value = 0.02	p-value = 0.04	p-value = 0.03
Sioux 2019	1	p-value = 0.01		
	2			
	3	p-value = 0.04		
	4			
	5			
	6	p-value = 0.03		
	7	p-value < 0.01		
	8			
Saganing 2020	1			p-value = 0.02
	2			
	4			
	5			
	6			
	7			
	8			
	9	p-value < 0.01	p-value = 0.02	
	10			
Wildfowl 2020	1			
	3	p-value = 0.02		
	4	p-value < 0.01		
	5	p-value = 0.03		
	6	p-value < 0.01	p-value < 0.01	p-value = 0.01
	7			
	10	p-value < 0.01		

Figure 3.3: Nutrient limitation data collected from nutrient diffusing substrates (NDS) for 4 of the 5 transects. Type is the classification of each transect point as either wetland, wetland to stream transition zone, stream, stream to lake transition zone, or lake. Transect points with N effect are colored blue, P effect yellow, and N:P effect green. No nutrient limitation is colored gray. N = nitrogen and P = phosphorus. P-values are denoted where significant (p-value ≤ 0.05).

3.4.3 Process Co-Occurrence and Nutrient Limitation

N_2 fixation and denitrification co-occurred across all wetland-stream-lake interfaces. Rates of N_2 fixation ranged from 0 to 1950 µg/m^2/h with a median of 5.49 µg/m^2/h, while denitrification rates ranged from 0 to 16,536 µg/m^2/h with a median of 914 µg/m^2/h. There were no significant differences in rates of N_2 fixation or denitrification across habitat types of the interfaces ($\chi^2 = 2.57$, df = 4, p-value = 0.63 and $\chi^2 = 4.74$, df = 4, p-value = 0.31 respectively, Fig. 3.4). The highest rates of N_2 fixation

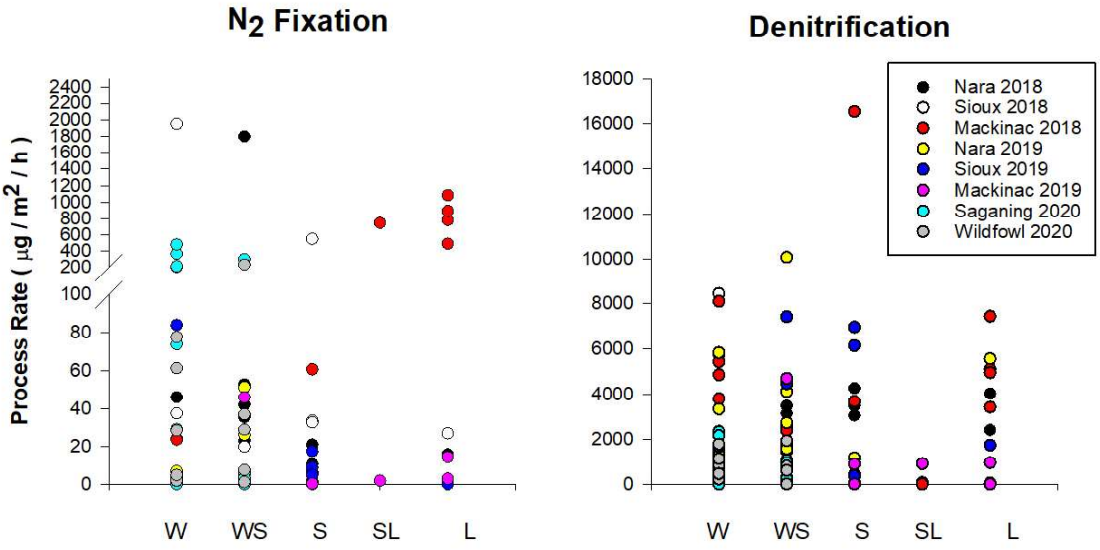

Figure 3.4: Rates of N_2 fixation and denitrification among all transects compared to transect point classification. The classifications were wetland (W, chamber n = 46), wetland to stream transition (WS, n = 27), stream (S, n = 22), stream to lake transition (SL, n = 3), and lake (L, n = 16). Note the Y-axis for denitrification is 6x that of N_2 fixation. Different sites are denoted by different colors.

occurred in wetlands and wetland to stream transition zones, while the highest denitrification rate occurred in a stream site, but high rates of denitrification were observed across all habitat types except the stream to lake transition zones (Fig. 3.4). Across habitat types, N_2 fixation occurred on both sediment and

macrophytes (Fig. 3.5). For N_2 fixation, of the 27 transect points with rates higher than the 75th percentile of all N_2 fixation rates (> 45.1 μg/m^2/h), 14 of those rates occurred on sediment substrate and 13 occurred on macrophyte substrate (Fig. 3.5). Denitrification occurred mostly in sediments, but occasionally macrophytes. For denitrification, of the 29 transect points with rates higher than the 75th percentile (> 3129 μg/m^2/h), 26 of those rates occurred on sediment substrate and 3 occurred on macrophyte substrate (Fig. 3.5).

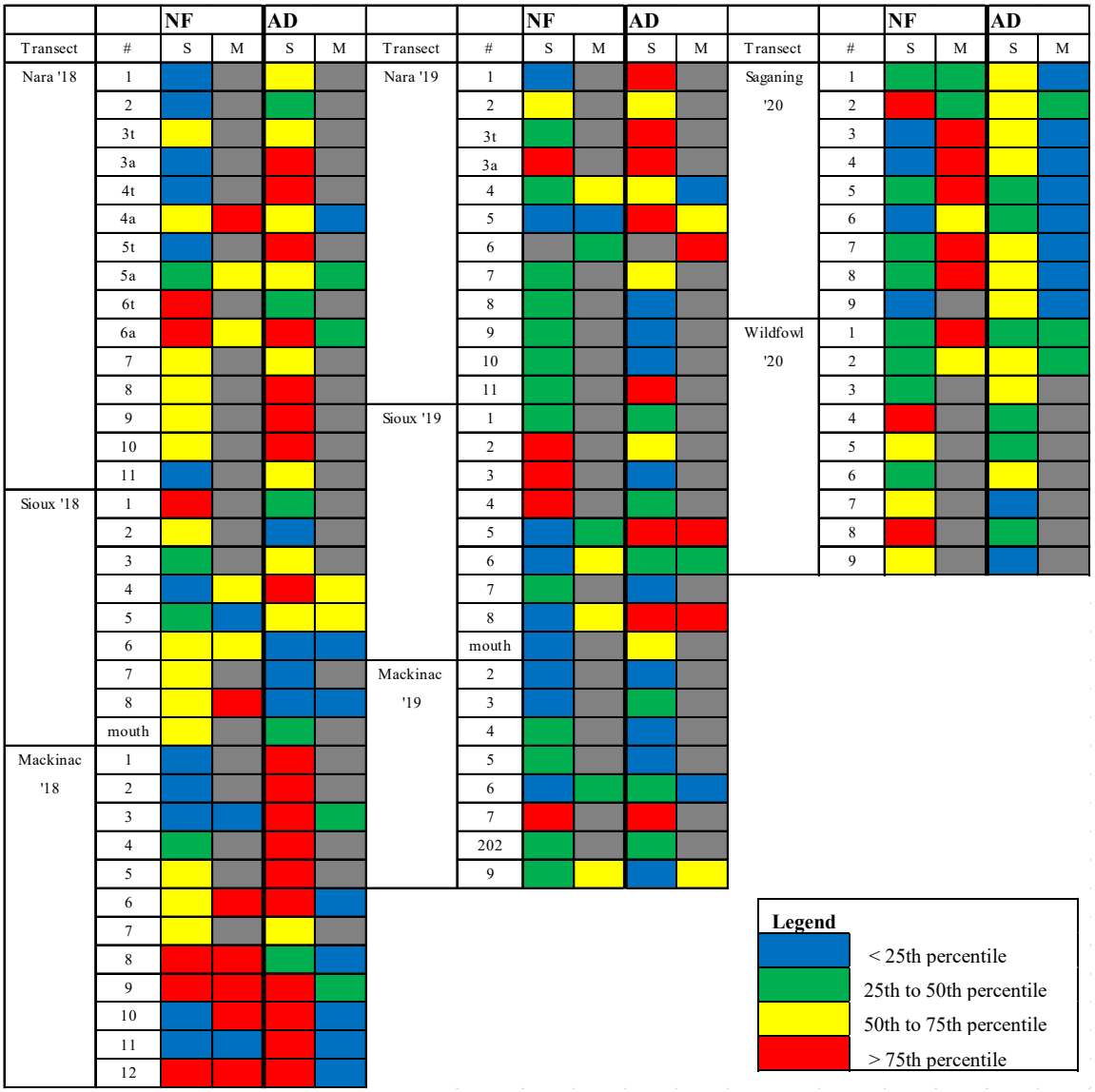

Figure 3.5: A heatmap of N_2 fixation (NF) and denitrification (AD) rates across all 5 transects broken down by year, transect number, and substrate type (M = macrophyte or S = sediment). For transect number t indicates a terrestrial site and a indicates aquatic. If no t or a is present, then a site is just an aquatic site. As transect numbers increase the transect moves from wetland to stream to lake. Rates of both processes are color coded based on quartiles. Blue indicates a rate that is < the 25th percentile (Q1), green indicates a rate that is between the 25th and 50th (Q2) quartile, yellow indicates a rate is between the 50th and 75th (Q3) quartile, and red indicates a rate is above the 75th quartile. Grey indicates no rate was measured for the transect point and substrate combination. For N_2 fixation Q1= 0.02, Q2 = 5.49, and Q3 = 45.1 $\mu g/m^2/h$ and for denitrification Q1 = 0, Q2 = 913.8, and Q3 = 3129.1 $\mu g/m^2/h$.

To test our second hypothesis that spatial variability in nutrient limitation would facilitate co-occurrence of N_2 fixation and denitrification, we found no significant relationship between either N_2 fixation or denitrification rates and nutrient limitation status ($\chi^2 = 5.45$, df = 2, p-value = 0.07 and $\chi^2 = 2.04$, df = 2, p-value = 0.36 respectively, Fig. 3.6). Observationally, the highest rates of N_2 fixation were in sites with no nutrient limitation followed by N limitation, whereas for denitrification the highest rate was in an N limitation site, followed by high rates in no limitation sites (Fig. 3.6).

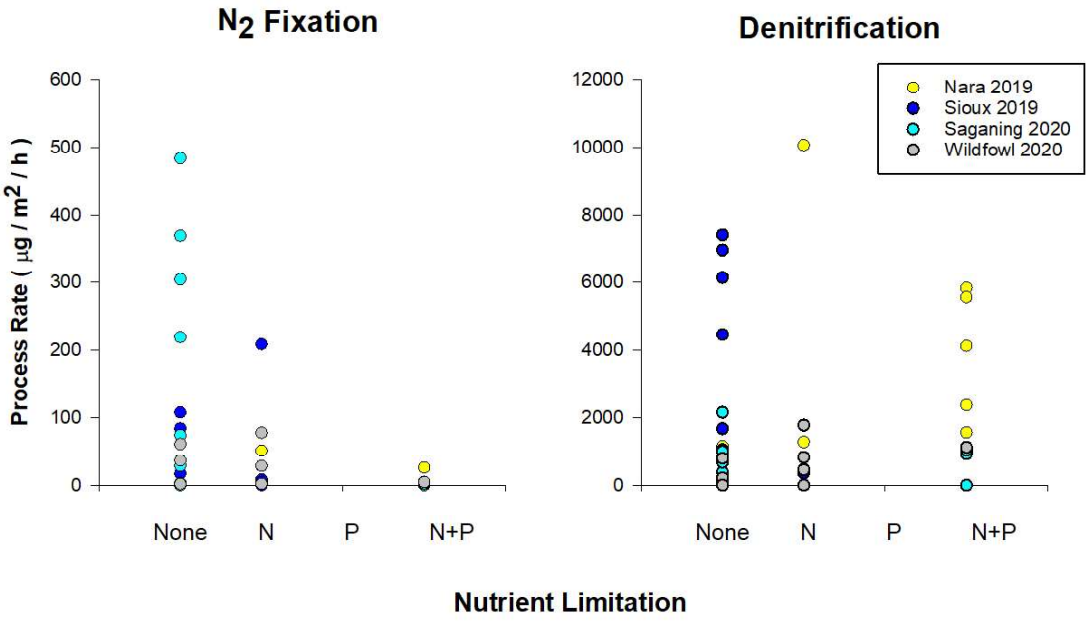

Figure 3.6: Rates of N_2 fixation and denitrification in comparison to nutrient limitation status. Note the Y-axis for denitrification is 20x that of N_2 fixation. N = nitrogen, P = phosphorus, and N+P = co-limitation of nitrogen and phosphorus. Different sites are denoted by different colors.

3.4.4 Environmental Factors as Predictors of Process Rates

Predictive modeling for denitrification did not support our hypothesis that higher organic matter concentrations would facilitate higher rates of denitrification. When testing predictive models with denitrification rates as the response variable, the best fit model was a model tree using bootstrap resampling with a RMSE of 1.45 and a R^2 of 20% (Table 3.4). Model trees are a type of regression tree where terminal nodes predict

Table 3.4: Predictive modeling results for the response variables N_2 fixation rates and denitrification rates. Items in bold represent the model of best fit based on the lowest RMSE. RMSE = root mean square error.

Response Variable	Model Type	RMSE	R^2
N_2 Fixation Rates	Partial Least Squares	0.74	0.01
	Ridge Regression	0.77	0.03
	Elastic Net/Lasso	0.71	0.08
	Neural Networks	0.72	0.01
	Support Vector Machines	0.77	0.01
	MARS/FDA	0.69	0.08
	K-Nearest Neighbor	0.71	0.05
	Single Trees	0.89	0.04
	Model Trees	0.62	0.52
	Bagged Trees	0.76	< 0.01
	Random Forest	0.72	0.04
	Boosted Trees	0.71	0.01
	Cubist	**0.60**	**0.33**
Denitrification Rates	Partial Least Squares	1.55	0.05
	Ridge Regression	1.62	< 0.01
	Elastic Net/Lasso	1.56	0.20
	Neural Networks	1.55	0.06
	Support Vector Machines	1.53	0.08
	MARS/FDA	1.61	< 0.01
	K-Nearest Neighbor	1.57	0.01
	Single Trees	1.66	< 0.01
	Model Trees	**1.45**	**0.20**
	Bagged Trees	1.56	0.06
	Random Forest	1.54	0.06
	Boosted Trees	1.56	0.05
	Cubist	1.51	0.06

the outcome using a linear model and the model is created with a split variable that is associated with the largest reduction in error (Kuhn and Johnson 2013). The tree growing process continues in this way until there are no further improvements in the error rate or there are not enough samples and then each linear model undergoes simplification, potentially dropping some of the model terms (Kuhn and Johnson 2013). Smoothing is also used in model trees to decrease the potential for overfitting (Kuhn and Johnson 2013). The variables of importance to the model tree predicting denitrification rates were NH_4^+, TDN, SRP, canopy cover, and DIN:TDP. NH_4^+ was the variable of most importance, followed by TDN, then SRP, and then canopy cover and DIN:TDP as variables of second, third, and fourth importance.

Predictive modeling for N_2 fixation did support our hypothesis that temperature would be important to predicting N_2 fixation rates, but did not support that NO_3^- concentrations would be important as well. When testing predictive models with N_2 fixation as the response variable, the best fit model was a cubist model using bootstrap resampling, with a RMSE of 0.60 and an R^2 of 33% (Table 3.4). Cubist models are also rule-based, but they are a mixture of several methodologies (Kuhn and Johnson 2013). Cubist models use linear predictions at each node and those are all collected into a single linear model that is smoothed that represents each individual model. Macrophyte substrate was the variable of most importance to the model predicting N_2 fixation rates, followed by temperature, AFDM, and then DOC, and then SRP.

3.5 Discussion

Our results demonstrate that N_2 fixation and denitrification do co-occur across habitats in wetland – stream – lake interfaces and that the occurrence of these processes cannot simply be explained by differences in nutrient limitation or concentrations across the interfaces. When evaluating our first hypothesis that nutrient limitation would vary spatially across the interfaces, we found there was N, N+P, and/or no nutrient limitation at transect points across interfaces. When evaluating our second hypothesis that spatial variation in nutrient limitation would facilitate the co-occurrence of N_2 fixation and denitrification across the interfaces, there was no significant relationship between nutrient limitation and rates of either N_2 fixation or denitrification. N_2 fixation and denitrification did co-occur across the 5 wetland – stream – lake interfaces of Lakes Superior and Huron encompassing a range of N and P concentrations (DIN = 0.026 to 0.260 mg/L, TDP = 0.003 to 0.017 mg/L among sites). Rates of both processes varied spatially within the transects among the different habitat types, but the variation of rates among habitat types was not significant. Predictive models did not support our hypothesis that high rates of denitrification would be related to high organic matter concentrations or that high rates of N_2 fixation would be related to NO_3^- concentrations. Predictive models did indicate that SRP was a variable of importance to both N_2 fixation and denitrification rates. Sources of N were important to denitrification rates and sources of C were important to N_2 fixation rates. Also, macrophyte substrate and temperature were the variables of most importance to the N_2 fixation model. Yet it should be noted that both process models only explained 20-33% of the variation, so although these environmental conditions may be of some importance to process rates across coastal wetland ecosystems, they do not fully predict

the variation we observed. Thus, spatial heterogeneity is important to the occurrence of both of these processes in coastal wetland ecosystems and these ecosystems cannot simply be thought of as nutrient sinks that remove N, which is how wetlands are commonly conceptualized.

The spatial variability of habitat is important to the co-occurrence of N_2 fixation and denitrification in coastal wetland ecosystems. Our results show that N_2 fixation and denitrification do occur across all habitat types within the wetland – stream – lake interface, so any alteration of these habitats could alter N cycling within the interface. Variability of substrate type within the wetland-stream-lake interface helped facilitate this co-occurrence of N_2 fixation and denitrification. This was particularly the case for N_2 fixation, where high rates of the process were observed evenly among sediment and macrophyte substrate, and macrophyte substrate was the variable of most importance to the model predicting N_2 fixation rates. Macrophytes could be important hosts to bound epiphytes that have the potential to fix N (Scott et al. 2005), while sediment could be an important habitat for heterotrophic N_2 fixers or cyanobacteria in sediment-bound microphytobenthos (Scott et al. 2008, Newell et al. 2016). Both of these substrate types across the interfaces facilitated N_2 fixation and without both we would see a great reduction in the total amount of N_2 fixation. Temperature was also important to the model predicting N_2 fixation rates. Temperature has been shown to be a limiting factor of N_2 fixation rates, with increasing temperatures leading to an increase in rates (Marcarelli and Wurtsbaugh 2006, Scott and Marcarelli 2012). In these coastal wetland ecosystems, increased temperature could be linked to the shallower depths of the wetlands where more N_2 fixation activity could possibly occur, though we also observed high rates of N_2

fixation in deeper nearshore areas. N_2 fixation rates have been shown to differ amongst depths in the Great Lakes, which could be due to light availability (Natwora and Sheik 2021). Light availability was also a variable of importance to the model for denitrification rates. Light can affect the amount of oxygen present in a system, which could constrain denitrification rates as it is an anerobic process. High concentrations of oxygen have been shown to negatively impact denitrification rates in sediment of the Great Lakes and across inland marshes of the Great Lakes region (Small et al. 2016, Dybiec et al. 2021).

The spatial variability of primary producer nutrient limitation in these interfaces also highlights the importance of habitat complexity for nutrient dynamics. Overall, we observed a majority of no nutrient limitation across transect points. However, like other studies of nutrient limitation of primary producers in wetlands (Cooper et al. 2016), we did find more N limitation followed by co-limitation of N and P at sites classified as wetland. Previous studies in streams of Lake Superior have shown a predominant co-limitation of N and P (Wold and Hershey 1999), but we observed sites with no limitation, N limitation, and co-limitation of N and P in our study streams. Due to sample recovery we only had nutrient limitation data for one lake site and it was N and P co-limited. Primary producers in the water column of the Great Lakes are primarily limited by phosphorus (P) (Schelske et al. 1987), but studies in other lakes have shown that nutrient limitation can differ between species (Fairchild et al. 1985) and between benthic and planktonic organisms of lakes (Bonilla et al. 2005, Steinman et al. 2016). In a eutrophic lake, benthic algae were found to be co-limited by N and P, while phytoplankton were P-limited (Steinman et al. 2016), which indicates that nutrient limitation of primary producers within lakes is complex. The spatial variation in nutrient limitation of primary

producers measured by NDS had no significant relationship to rates of N_2 fixation or denitrification in our study. This could be because nutrient limitation alone is not the best predictor of process rates.

Although nutrient limitation was not a predictor of N_2 fixation and denitrification rates, the availability of N and/or P were variables of importance included in the predictive models for rates of N_2 fixation and denitrification across the coastal wetland interfaces. Our results suggest N in the forms of NH_4^+ and TDN are important to predicting denitrification rates. NH_4^+ may be of importance to denitrification because when reduced through the process of nitrification NH_4^+ can be a source of NO_3^-, which is used as an oxidant in the denitrification process. Previous studies in Great Lakes sediment have shown that denitrification may be fueled by nitrification, while others have shown that nitrification and denitrification are uncoupled in Lake Superior estuary sediments (Small et al. 2014b, Bellinger et al. 2014). TDN encompasses all the dissolved forms of nitrogen which would include NH_4^+ and NO_3^-, which individually are known constraining variables of denitrification rates. Interestingly, N concentrations were not important to the model for N_2 fixation rates, which is similar to other studies suggesting the relationship between N_2 fixation rates and N concentrations are not always direct (Knapp et al. 2016, Eberhard et al. 2018, Tang et al. 2020). The availability of P in the forms of SRP and/or DIN:TDP were also important to predicting rates of denitrification and N_2 fixation. We found this previously in streams with TDP and DIN:TDP being both important positive predictors to denitrification rates (Eberhard et al. 2018) and suggested that the mechanism could be similar to that proposed in lakes where increased P stimulates algal production and N uptake and when algae die they end up in sediments,

delivering N and organic matter, which increase denitrification rates (Finlay et al. 2013). P availability has been shown to limit N_2 fixation rates in aquatic ecosystems (Elwood et al. 1981, Howarth et al. 1988, Marcarelli and Wurtsbaugh 2007). C availability was also an important predictor to N_2 fixation rates as DOC and AFDM. AFDM is a measure of organic matter, which has been shown to affect N_2 fixation rates by increasing the availability of trace metals used in the nitrogenase enzyme like molybdenum and iron (Howarth et al. 1988). Variability in nutrient concentrations across wetland-stream-lake interfaces may also play a role in facilitating the co-occurrence of N_2 fixation and denitrification in these ecosystems.

Our study did not assess how temporal variability may play a factor in the spatial variability of environmental characteristics and rates of N_2 fixation and denitrification across coastal wetland ecosystems. Recent literature has noted that a fundamental trait of spatial areas of high rates of a biogeochemical process, or hot spots, is that they are temporally dynamic and that these areas should be thought of more of as "control points" that can be turned on or off depending on the timing and magnitude of delivery of limiting factors (Bernhardt et al. 2017). There is evidence of temporal dynamics in our data where in 2018 and 2019 we sampled the same transects. When looking at Table 4, there are different patterns of the process rates of N_2 fixation and denitrification in the same transect points in different years. Moreover, due to rising water levels in the Great Lakes in summer 2019 we were not able to access all the same transect points in Nara or Mackinac as in 2018 and in some cases substrate that was prevalent at a transect point in one year was absent in the other. Rising water levels could affect temperature and light availability, which were important predictors of N_2 fixation and denitrification in these

ecosystems. Water level fluctuations have been shown to have the potential to alter sediment and water nutrient exchange in Great Lakes coastal wetlands in previous studies (Steinman et al. 2012, 2014). Changing substrate presence would also alter the potential for denitrification and particularly N_2 fixation to occur where we saw high rates on both sediment and macrophyte substrate. Previously, rates of both processes have been found to vary day-to-day (maximum daily change of 4,390 μg $N/m^2/hr$ for denitrification and 39 μg $N/m^2/hr$ for N_2 fixation) across seasons in the Pilgrim River, which was part of the Nara transect (Nevorski 2021).

Spatial heterogeneity of environmental variables and habitat in coastal wetland ecosystems facilitate the co-occurrence of N_2 fixation and denitrification in these ecosystems. This means that losses via denitrification must be considered relative to inputs from N_2 fixation to accurately understand the role that wetlands play in nutrient uptake and load mitigation because not as much N will be removed as we may think looking at denitrification rates alone. Plus, the occurrence of both of these processes across coastal wetland ecosystems could affect N dynamics of the larger Great Lakes. For example, recent studies have shown that Lake Superior may be seeded with cyanobacteria through fluvial inputs (Reinl et al. 2020). Therefore, alterations to the stream and/or wetland N dynamics could have an effect on what is being transported to the larger bodies and their biogeochemical cycles. The spatial heterogeneity within coastal wetland ecosystems is key to maintaining this diversity in nutrient cycling. Anything that may reduce physical habitat or biodiversity complexity, such as the invasive wetland plant *Phragmites australis*, will alter the way that wetlands cycle, store, and transport nutrients (Duke et al. 2015, Judd and Francoeur 2019). Therefore, from a

restoration and conservation perspective, it is important to maintain and restore spatial heterogeneity in these ecosystems to preserve their function in complex biogeochemical cycling.

3.6 References

Allen, N. S., & Hershey, A. E. (1996). Seasonal changes in chlorophyll a response to nutrient amendments in a North Shore tributary of Lake Superior. *Journal of the North American Benthological Society*, *15*(2), 170-178.

American Public Health Administration (2005). Standard Methods for examination of water and wastewater. APHA, Washington D.C.

Arango, C. P., Tank, J. L., Schaller, J. L., Royer, T. V., Bernot, M. J., & David, M. B. (2007). Benthic organic carbon influences denitrification in streams with high nitrate concentration. *Freshwater Biology*, *52*(7), 1210-1222.

Arp, C. D., Whitman, M. S., Jones, B. M., Kemnitz, R., Grosse, G., & Urban, F. E. (2012). Drainage network structure and hydrologic behavior of three lake-rich watersheds on the Arctic Coastal Plain, Alaska. *Arctic, Antarctic, and Alpine Research*, *44*(4), 385-398.

Baker, M. A., Arp, C. D., Goodman, K. J., Marcarelli, A. M., & Wurtsbaugh, W. A. (2016). Stream-lake interaction: understanding coupled hydro-ecological systems. In *Stream ecosystems in a changing environment* (pp. 321-348). Academic Press.

Bellinger, B. J., Jicha, T. M., Lehto, L. P., Seifert-Monson, L. R., Bolgrien, D. W., Starry, M. A., ... & Hill, B. H. (2014). Sediment nitrification and denitrification in a Lake Superior estuary. *Journal of Great Lakes Research*, *40*(2), 392-403.

Bernhardt, E. S., Blaszczak, J. R., Ficken, C. D., Fork, M. L., Kaiser, K. E., & Seybold, E. C. (2017). Control points in ecosystems: moving beyond the hot spot hot moment concept. *Ecosystems*, *20*(4), 665-682.

Biddanda, B. A., & Cotner, J. B. (2002). Love handles in aquatic ecosystems: the role of dissolved organic carbon drawdown, resuspended sediments, and terrigenous inputs in the carbon balance of Lake Michigan. *Ecosystems*, *5*(5), 431-445.

Bonilla, S., Villeneuve, V., & Vincent, W. F. (2005). Benthic and planktonic Algal communities in a high arctic lake: pigment structure and contrasting responses to nutrient enrichment. *Journal of Phycology*, *41*(6), 1120-1130.

Bourgeau-Chavez, L., Endres, S., Battaglia, M., Miller, M.E., Banda, E., Laubach, Z., Higman, P., Chow-Fraser, P., Marcaccio, J. (2015). Development of a Bi-National Great Lakes Coastal Wetland and Land Use Map Using Three-Season PALSAR and Landsat Imagery. *Remote Sens.*, *7*, 8655-8682.https://www.mdpi.com/2072-4292/7/7/8655

Callahan, M. K., Whigham, D. F., Rains, M. C., Rains, K. C., King, R. S., Walker, C. M., ... & Baird, S.J. (2017). Nitrogen subsidies from hillslope alder stands to streamside wetlands and headwater streams, Kenai Peninsula, Alaska. *JAWRA Journal of the American Water Resources Association*, *53*(2), 478-492.

Capone, D.G. (1993) Determination of nitrogenase activity in aquatic samples using the acetylene reduction procedure. In: Kemp P.F., Sherr, B.F., Sherr, E.B., Cole J.J. (ed) Handbook of Methods in Aquatic Microbial Ecology, Lewis Publishers, Boca Raton.

Cooper, M. J., Costello, G. M., Francoeur, S. N., & Lamberti, G. A. (2016). Nitrogen

limitation of algal biofilms in coastal wetlands of Lakes Michigan and Huron. *Freshwater Science*, *35*(1), 25-40.

Dila, D. K., & Biddanda, B. A. (2015). From land to lake: contrasting microbial processes across a Great Lakes gradient of organic carbon and inorganic nutrient I nventories. *Journal of Great Lakes Research*, *41*, 75-85.

Dodds, W. K., Burgin, A. J., Marcarelli, A. M., & Strauss, E. A. (2017). Nitrogen transformations. In *Methods in stream ecology* (pp. 173-196). Academic Press.

Doyle, R. D., & Fisher, T. R. (1994). Nitrogen fixation by periphyton and plankton on the Amazon floodplain at Lake Calado. *Biogeochemistry*, *26*(1), 41-66.

Duke, S. T., Francoeur, S. N., & Judd, K. E. (2015). Effects of Phragmites australis invasion on carbon dynamics in a freshwater marsh. *Wetlands*, *35*(2), 311-321.

Dybiec, J. M., Danz, N. P., Learman, D. R., & Uzarski, D. G. (2021). Evaluating relationships between plants, water chemistry, and denitrification potential in palustrine freshwater marshes. *Ecological Indicators*, *131*, 108163.

Eberhard, E. K., Marcarelli, A. M., & Baxter, C. V. (2018). Co-occurrence of in-stream nitrogen fixation and denitrification across a nitrogen gradient in a western US watershed. *Biogeochemistry*, *139*(2), 179-195.

Elwood, J. W., Newbold, J. D., Trimble, A. F., & Stark, R. W. (1981). The limiting role of phosphorus in a woodland stream ecosystem: effects of P enrichment on leaf decomposition and primary producers. *Ecology*, *62*(1), 146-158.

Epstein, D. M., Neilson, B. T., Goodman, K. J., Stevens, D. K., & Wurtsbaugh, W. A. (2013). A modeling approach for assessing the effect of multiple alpine lakes in sequence on nutrient transport. *Aquatic Sciences*, *75*(2), 199-212.

Fairchild, G. W., Lowe, R. L., & Richardson, W. B. (1985). Algal periphyton growth on nutrient-diffusing substrates: an in situ bioassay. *Ecology*, *66*(2), 465-472.

Finlay, J. C., Small, G. E., & Sterner, R. W. (2013). Human influences on nitrogen removal in lakes. *Science*, *342*(6155), 247-250.

Finke, L. R., & Seeley Jr, H. W. (1978). Nitrogen fixation (acetylene reduction) by epiphytes of freshwater macrophytes. *Applied and Environmental Microbiology*, *36*(1), 129-138.

Flint, S. A., & McDowell, W. H. (2015). Effects of headwater wetlands on dissolved nitrogen and dissolved organic carbon concentrations in a suburban New Hampshire watershed. *Freshwater Science*, *34*(2), 456-471.

Gettel, G. M., Giblin, A. E., & Howarth, R. W. (2007). The effects of grazing by the snail, Lymnaeaelodes, on benthic N2 fixation and primary production in oligotrophic, arctic lakes. *Limnology and oceanography*, *52*(6), 2398-2409.

Grimm, N. B., & Petrone, K. C. (1997). Nitrogen fixation in a desert stream ecosystem. *Biogeochemistry*, *37*(1), 33-61.

Goodman, K. J., Baker, M. A., & Wurtsbaugh, W. A. (2010). Mountain lakes increase organic matter decomposition rates in streams. *Journal of the North American Benthological Society*, *29*(2), 521-529.

Groffman, P. M., Dorsey, A. M., & Mayer, P. M. (2005). N processing within geomorphic structures in urban streams. *Journal of the North American Benthological Society*, *24*(3), 613-625.

Groffman, P. M., Altabet, M. A., Böhlke, J. K., Butterbach-Bahl, K., David, M. B.,

Firestone, M. K., ... & Voytek, M. A. (2006). Methods for measuring denitrification: diverse approaches to a difficult problem. *Ecological Applications*, *16*(6), 2091-2122.

Harrison, M. D., Miller, A. J., Groffman, P. M., Mayer, P. M., & Kaushal, S. S. (2014). Hydrologic controls on nitrogen and phosphorous dynamics in relict oxbow wetlands adjacent to an urban restored stream. *JAWRA Journal of the American Water Resources Association*, *50*(6), 1365-1382.

Hedin, L. O., Von Fischer, J. C., Ostrom, N. E., Kennedy, B. P., Brown, M. G., & Robertson, G. P. (1998). Thermodynamic constraints on nitrogen transformations and other biogeochemical processes at soil–stream interfaces. *Ecology*, *79*(2), 684-703.

Holmes, R. M., Jones, J. B., Fisher, S. G., & Grimm, N. B. (1996). Denitrification in a nitrogen-limited stream ecosystem. *Biogeochemistry*, *33*(2), 125-146.

Holmes, R. M., Aminot, A., Kérouel, R., Hooker, B. A., & Peterson, B. J. (1999). A simple and precise method for measuring ammonium in marine and freshwater ecosystems. *Canadian Journal of Fisheries and Aquatic Sciences*, *56*(10), 1801-1808.

Howarth, R. W., Marino, R., & Cole, J. J. (1988). Nitrogen fixation in freshwater, estuarine, and marine ecosystems. 2. Biogeochemical controls. *Limnology and Oceanography*, *33*(4), 688-701.

Jones, N. E. (2010). Incorporating lakes within the river discontinuum: longitudinal changes in ecological characteristics in stream–lake networks. *Canadian Journal of Fisheries and Aquatic Sciences*, *67*(8), 1350-1362.

Judd, K. E., & Francoeur, S. N. (2019). Short-term impacts of Phragmites management on nutrient budgets and plant communities in Great Lakes coastal freshwater marshes. *Wetlands Ecology and Management*, *27*(1), 55-74.

Kling, G. W., Kipphut, G. W., Miller, M. M., & O'Brien, W. J. (2000). Integration of lakes and streams in a landscape perspective: the importance of material processing on spatial patterns and temporal coherence. *Freshwater Biology*, *43*(3), 477-497.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26). New York: Springer.

Kuhn, M. (2019). The caret package: Classification and regression training. *R Found Stat Comput Vienna, Austria http://topepo. github. io/caret/index. html*.

Knapp, A. N., Casciotti, K. L., Berelson, W. M., Prokopenko, M. G., & Capone, D. G. (2016). Low rates of nitrogen fixation in eastern tropical South Pacific surface waters. *Proceedings of the National Academy of Sciences*, *113*(16), 4398-4403.

Knuth, M. L., & Kelly, J. R. (2011). Denitrification rates in a Lake Superior coastal wetland. *Aquatic Ecosystem Health & Management*, *14*(4), 414-421.

Larson, J. H., Trebitz, A. S., Steinman, A. D., Wiley, M. J., Mazur, M. C., Pebbles, V., ... & Seelbach, P. W. (2013). Great Lakes rivermouth ecosystems: scientific synthesis and management implications. *Journal of Great Lakes Research*, *39*(3), 513-524.

Larson, J. H., Frost, P. C., Vallazza, J. M., Nelson, J. C., & Richardson, W. B. (2016). Do rivermouths alter nutrient and seston delivery to the nearshore?. *Freshwater biology*, *61*(11), 1935-1949.

Lemmon, P. E. (1956). A spherical densiometer for estimating forest overstory density. *Forest science*, *2*(4), 314-320.

Lottig, N. R., Buffam, I., & Stanley, E. H. (2013). Comparisons of wetland and drainage lake influences on stream dissolved carbon concentrations and yields in a north temperate lake-rich region. *Aquatic sciences*, *75*(4), 619-630.

Marcarelli, A. M., & Wurtsbaugh, W. A. (2006). Temperature and nutrient supply interact to control nitrogen fixation in oligotrophic streams: an experimental examination. *Limnology and Oceanography*, *51*(5), 2278-2289.

Marcarelli, A. M., & Wurtsbaugh, W. A. (2007). Effects of upstream lakes and nutrient limitation on periphytic biomass and nitrogen fixation in oligotrophic, subalpine streams. *Freshwater Biology*, *52*(11), 2211-2225.

Marcarelli, A. M., & Wurtsbaugh, W. A. (2009). Nitrogen fixation varies spatially and seasonally in linked stream-lake ecosystems. *Biogeochemistry*, *94*(2), 95-110.

Marcarelli, A. M., Fulweiler, R. W., & Scott, J. T. (2022). Nitrogen fixation: A poorly understood process along the freshwater-marine continuum. *Limnology and Oceanography Letters*, *7*(1), 1-10.

McCarthy, M. J., Gardner, W. S., Lavrentyev, P. J., Moats, K. M., Jochem, F. J., & Klarer, D. M. (2007).Effects of hydrological flow regime on sediment-water interface and water column nitrogen dynamics in a Great Lakes coastal wetland (Old Woman Creek, Lake Erie). *Journal of Great Lakes Research*, *33*(1), 219-231.

McDonald, C. P., Urban, N. R., & Casey, C. M. (2010). Modeling historical trends in

Lake Superior total nitrogen concentrations. *Journal of Great Lakes Research*, *36*(4), 715-721.

Natwora, K. E., & Sheik, C. S. (2021). Assessment of nitrogen fixation rates in the Laurentian Great Lakes. *Journal of Great Lakes Research*, *47*(5), 1288-1295.

Nevorski, K. (2021). Denitrification and nitrogen fixation commonly co-occur but rates vary throughout the year and in different environments. Open Access Dissertation, Michigan Technological University.

Newell, S. E., McCarthy, M. J., Gardner, W. S., & Fulweiler, R. W. (2016). Sediment nitrogen fixation: a call for re-evaluating coastal N budgets. *Estuaries and Coasts*, *39*(6), 1626-1638.

Orr, C. H., Predick, K. I., Stanley, E. H., & Rogers, K. L. (2014). Spatial autocorrelation of denitrification in a restored and a natural floodplain. *Wetlands*, *34*(1), 89-100.

Reinl, K. L., Sterner, R. W., Lafrancois, B. M., & Brovold, S. (2020). Fluvial seeding of cyanobacterial blooms in oligotrophic Lake Superior. *Harmful Algae*, *100*, 101941.

Richardson, J. S., Michalski, T., & Becu, M. (2021). Stream inflows to lake deltas: A tributary junction that provides a unique habitat in lakes. *Freshwater Biology*, *66*(11), 2021-2029.

Salk, K. R., Bullerjahn, G. S., McKay, R. M. L., Chaffin, J. D., & Ostrom, N. E. (2018). Nitrogen cycling in Sandusky Bay, Lake Erie: oscillations between strong and weak export and implications for harmful algal blooms. *Biogeosciences*, *15*(9), 2891-2907.

Schelske, C. L., Conley, D. J., Stoermer, E. F., Newberry, T. L., & Campbell, C. D.

(1987). Biogenic silica and phosphorus accumulation in sediments as indices of eutrophication in the Laurentian Great Lakes. In *Paleolimnology IV* (pp. 79-86). Springer, Dordrecht.

Scott, J. T., Doyle, R. D., & Filstrup, C. T. (2005). Periphyton nutrient limitation and nitrogen fixation potential along a wetland nutrient-depletion gradient. *Wetlands*, *25*(2), 439-448.

Scott, J. T., McCarthy, M. J., Gardner, W. S., & Doyle, R. D. (2008). Denitrification, dissimilatory nitrate reduction to ammonium, and nitrogen fixation along a nitrate concentration gradient in a created freshwater wetland. *Biogeochemistry*, *87*(1), 99-111.

Scott, J. T., & Marcarelli, A. M. (2012). Cyanobacteria in freshwater benthic environments. In *Ecology of cyanobacteria II* (pp. 271-289). Springer, Dordrecht.

Sierszen, M. E., Morrice, J. A., Trebitz, A. S., & Hoffman, J. C. (2012). A review of selected ecosystem services provided by coastal wetlands of the Laurentian Great Lakes. *Aquatic Ecosystem Health & Management*, *15*(1), 92-106.

Small, G. E., Sterner, R. W., & Finlay, J. C. (2014a). An ecological network analysis of nitrogen cycling in the Laurentian Great Lakes. *Ecological modelling*, *293*, 150-160.

Small, G. E., Cotner, J. B., Finlay, J. C., Stark, R. A., & Sterner, R. W. (2014b). Nitrogen transformations at the sediment–water interface across redox gradients in the Laurentian Great Lakes. *Hydrobiologia*, *731*(1), 95-108.

Small, G. E., Finlay, J. C., McKay, R. M. L., Rozmarynowycz, M. J., Brovold, S.,

Bullerjahn, G. S., ... & Sterner, R. W. (2016). Large differences in potential denitrification and sediment microbial communities across the Laurentian great lakes. *Biogeochemistry*, *128*(3), 353-368.

Steinman, A. D., Ogdahl, M. E., Weinert, M., Thompson, K., Cooper, M. J., & Uzarski, D. G. (2012). Water level fluctuation and sediment–water nutrient exchange in Great Lakes coastal wetlands. *Journal of Great Lakes Research*, *38*(4), 766-775.

Steinman, A. D., Ogdahl, M. E., Weinert, M., & Uzarski, D. G. (2014). Influence of water-level fluctuation duration and magnitude on sediment–water nutrient exchange in coastal wetlands. *Aquatic ecology*, *48*(2), 143-159.

Steinman, A., Abdimalik, M., Ogdahl, M. E., & Oudsema, M. (2016). Understanding planktonic vs. benthic algal response to manipulation of nutrients and light in a eutrophic lake. *Lake and Reservoir Management*, *32*(4), 402-409.

Tang, W., Cerdán-García, E., Berthelot, H., Polyviou, D., Wang, S., Baylay, A., ... & Cassar, N. (2020). New insights into the distributions of nitrogen fixation and diazotrophs revealed by high-resolution sensing and sampling methods. *The ISME journal*, *14*(10), 2514-2526.

Tank, J. L., & Dodds, W. K. (2003). Nutrient limitation of epilithic and epixylic biofilms in ten NorthAmerican streams. *Freshwater Biology*, *48*(6), 1031-1049.

Tank, J. L., Reisinger, A. J., & Rosi, E. J. (2017). Nutrient limitation and uptake. In *Methods in stream ecology* (pp. 147-171). Academic Press.

Taylor, B. W., Keep, C. F., Hall Jr, R. O., Koch, B. J., Tronstad, L. M., Flecker, A. S., &

Ulseth, A. J. (2007). Improving the fluorometric ammonium method: matrix effects, background fluorescence, and standard additions. *Journal of the North American Benthological Society*, *26*(2), 167-177.

Wang, H., Liao, G., D'Souza, M., Yu, X., Yang, J., Yang, X., & Zheng, T. (2016). Temporal and spatial variations of greenhouse gas fluxes from a tidal mangrove wetland in Southeast China. *Environmental Science and Pollution Research*, *23*(2), 1873-1885.

Welter, J. R., Benstead, J. P., Cross, W. F., Hood, J. M., Huryn, A. D., Johnson, P. W., & Williamson, T. J. (2015). Does N2 fixation amplify the temperature dependence of ecosystem metabolism?. *Ecology*, *96*(3), 603-610.

Wold, A. P., & Hershey, A. E. (1999). Spatial and temporal variability of nutrient limitation in 6 North Shore tributaries to Lake Superior. *Journal of the North American Benthological Society*, *18*(1), 2-14.

4 Chapter 4: Diversity of microbial assemblages in streams across ecoregions of the U.S. in relation to N_2 fixation and denitrification

4.1 Abstract

Microbial assemblages can differ across environmental gradients that define ecoregions, which can lead to differences in the potential for biogeochemical processing. We hypothesized that microbial community composition would differ amongst streams located in different ecoregions with more extreme differences in environmental variables such as temperature, canopy cover, depth, and nutrient concentrations. To evaluate this hypothesis, we collected environmental samples of dominant substrate types from 30 streams across 13 ecoregions in the U.S.A. for 16S rRNA Illumina gene sequencing. Alphaproteobacteria, Gammaproteobacteria, and Cyanobacteria were the Classes of microbes with the highest relative abundance observed across all ecoregions. Observed and Shannon alpha diversity were significantly different among the ecoregions (Kruskal-Wallis p-value = 0.01 and 0.02, respectively). A distance-based redundancy analysis showed that environmental variables did not explain much of the variance in microbial communities among ecoregions (axis 1 = 4.6%, axis 2 = 2.8%), but the difference in microbial communities among ecoregions was significant (PERMANOVA p-value < 0.01, R^2 = 17.6%). When comparing the predicted functional potential of the microbial community composition determined using FAPROTAX vs. stream reach average rates of N_2 fixation and denitrification, we found no significant relationship between process rates and functional potential (Spearman's correlation p-value = 0.77, ρ = -0.09 and p-value = 0.23, ρ = -0.36, respectively). Our results demonstrate that there are significant

differences in microbial community composition among streams across ecoregions, but these differences cannot fully be explained by environmental variables at this scale.

4.2 Introduction

Biogeography is the study of spatial patterns in biodiversity in relation to evolutionary events over space and time. The biogeography of large organisms such as plants and animals has been investigated since the 18th century and more recently, the biogeography of microorganisms has been a focus of increasing study (Martiny et al. 2006). Microorganisms are crucial to ecosystem processes because they decompose organic matter, serve as a food resource, and transform nutrients from one form to another. Microorganisms exhibit more phenotypic and genotypic diversity within their group than among all other organisms (Findlay 2010), so evaluating the biodiversity of microorganisms is key to understanding variation in ecosystem processes across space.

Ecoregions are characterized by groups of organisms that share similar characteristics based on the environments where they exist. Within each ecoregion environmental characteristics differ and thus constraints under which organisms can operate differ, which has driven the evolution of organisms to thrive within specific ecoregions and effected species richness (Hawkins et al. 2003, Kreft and Jetz 2007, Mittelbach and McGill 2019). Variables such as pH, temperature, ion concentration, and organic matter concentration can alter the taxonomy of microbial assemblages within ecoregions (Pagaling et al. 2009, Findlay 2010, Bru et al. 2011, Fierer et al. 2012). Larger-scale factors like latitudinal and elevation gradients, climate, lithology, and

historical contexts can also shape the microbial assemblages within and among ecoregions (LaRouche et al. 2012, Hendershot et al 2017). Variability in microbial assemblages across ecoregions can control biogeochemical processes because organisms that differ in processing capabilities thrive in environments that differ (Fierer et al. 2012).

Biogeochemical cycles such as the N cycle include multiple steps that are accomplished by a variety of microorganisms. For example, N_2 fixing and denitrifying bacteria differ in constraints that affect their establishment within a microbial assemblage. Enzymes in the denitrification pathway are highly conserved in many distantly related species of bacteria (Wellington et al. 2003, Schimel and Gulledge 1998). Furthermore, dominant genera that have genes encoding a pathway for processes like denitrification and nitrification can be highly similar across habitat types, suggesting N functional groups may be less specialized than thought for certain soil characteristics across landscapes (Nelson et al. 2016). However, differences in variables that control enzymes in the denitrification process could produce different levels of activity across ecosystems (Ferguson 1994). The N_2 fixation activity of microbes can be constrained by multiple variables like temperature, light, dissolved N, dissolved iron, and phosphate concentrations (Monteiro et al. 2011). N_2 fixing microbes like photosynthetic cyanobacteria increase in abundance with low N and high P and iron supply in stream microcosms (Larson et al. 2018). Analyzing how the assemblages of microorganisms differ based on environmental constraints will further understanding of the occurrence of N_2 fixation and denitrification in streams across ecoregions.

The goal of this study was to evaluate the taxonomy of microbial communities in streams across ecoregions of the U.S in relation to N_2 fixation and denitrification. We

evaluated taxonomy through 16S rRNA Illumina gene sequencing of microorganisms found on stream substrates. We first hypothesized that microbial assemblages would have different taxonomic compositions in streams from biomes that have extreme differences in environmental variables like temperature, canopy cover, depth, and nutrient concentrations, due to the different conditions for the microorganisms performing these processes. Our second hypothesis was that those streams with higher reach average rates of N_2 fixation or denitrification would have a higher % composition of microorganisms functionally capable of each process respectively.

4.3 Methods

4.3.1 Study Area

Samples were collected for 16S rRNA sequencing in 30 streams across 13 ecoregions during summer between 2017 and 2019 (Fig. 4.1). Most of the streams and rivers were selected from sites that are part of the National Ecological Observatory Network (NEON, Keller et al. 2008, Utz et al. 2013, and Goodman et al. 2015) or the StreamPULSE project (Bernhardt et al. 2018), as these sites are continuously monitored for discharge, temperature, and dissolved O_2. All streams were sampled once, except for New Hope Creek, Arikaree River, Kings Creek, and Pilgrim River, which were sampled multiple times across the study years (Table 4.1a-d). Stream size ranged in average width from 1 to 75 m and average depth from 0.06 to 1.08 m. There was a variety of benthic substrate across all streams, but the most common types across all streams were rock, sediment, and wood (Table 4.2a-b).

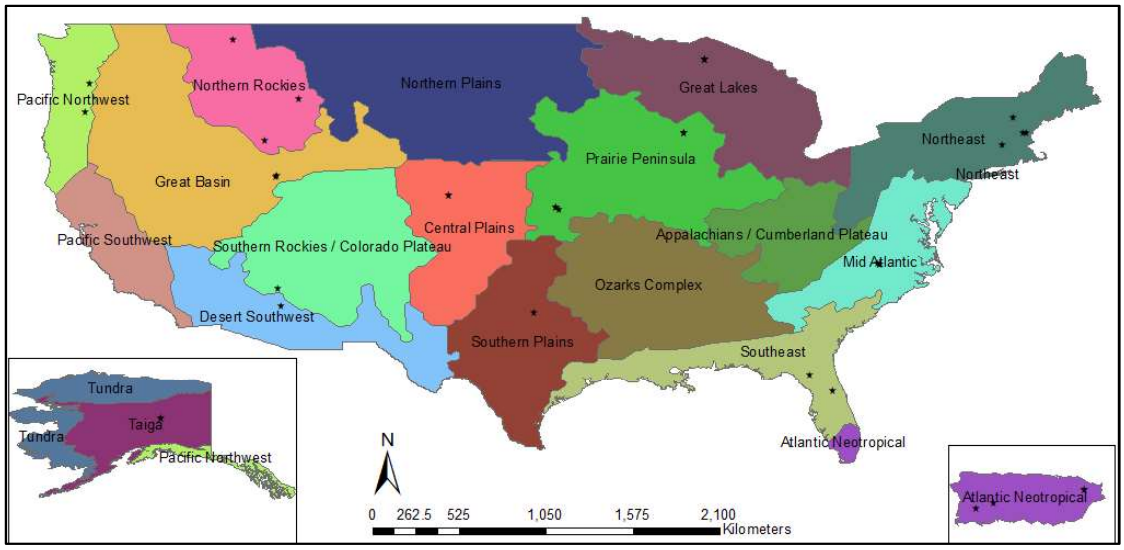

Figure 4.1: Map of the continental U.S. broken down by ecoregion. Each ecoregion is labeled and represented by a different color. Sampling sites are denoted by a black star. This map was created in ArcGIS using the domain layer created by the National Ecological Observatory Network (NEON).

4.3.2 Sampling Design

On a given sampling day, stream reaches were first mapped for habitat characteristics and relative substrate abundance present in the streambed. Samples of each representative substrate in the streambed were then collected and placed into chambers. Chambers then underwent acetylene reduction or acetylene block assays during 2 hour in-stream incubations to measure rates of N_2 fixation and denitrification, respectively. After the incubations, samples of substrate were taken from each chamber and placed in 15 mL falcon tubes to be frozen for 16S rRNA Illumina sequencing.

Table 4.1a: Physical environmental characteristics of a subset of the 30 study streams in 13 ecoregions by sampling date. Lat. Is latitude and Long. is longitude. Sites are arranged west to east by ecoregion.

Site Name	Ecoregion	Sample Date	Watershed Area (km^2)	Average reach width (m)	Average transect depth (cm)	Average canopy cover (%)	Lat. (°N)	Long. (°W)
Caribou Creek	Taiga	9-Jul-19	30.7	2.9	15.3	15.4	65.15306	-147.50246
Poker Creek	Taiga	10-Jul-19	NA	3.9	27.8	18.3	65.1528	-147.4867
McRae Creek	Pacific Northwest	3-Aug-19	5.7	3.7	7.2	89.4	44.2596	-122.16555
Martha Creek	Pacific Northwest	6-Aug-19	7.7	2.9	11.4	43.7	45.79125	-121.93196
South Fork Mink Creek	Northern Rockies	27-Jul-17	28.0	1.9	12.4	39.8	42.7074	-112.4228
Blacktail Deer Creek	Northern Rockies	30-Jul-19	35.0	2.5	18.2	41.2	44.95011	-110.58715
Creston Creek	Northern Rockies	10-Aug-19	93.8	1.9	10.6	0.0	48.1885	-114.1327
Lower Knowton Fork Red Butte Creek	Great Basin	18-Jul-17	4.1	1.4	8.8	87.6	40.8045	-111.7654
1300 E Red Butte Creek	Great Basin	19-Jul-17	26.2	5.3	30.8	60.7	40.7451	-111.8546
Sycamore Creek	Desert Southwest	14-May-19	273.0	2.2	8.7	3.9	33.74906	-111.5069
Wet Beaver Creek	Desert Southwest	16-May-19	294.0	12.9	47.0	18.5	34.6793	-111.695
Arikaree River	Central Plains	11-Jul-17	2890.0	1.8	6.2	2.9	39.75825	-102.4471
Arikaree River	Central Plains	24-May-19	2890.0	4.0	19.2	1.2	39.75825	-102.4471
Pringle Creek	Southern Plains	9-Jun-18	49.1	5.3	18.9	59.2	33.37859	-97.78226
Black Earth Creek	Prairie Peninsula	8-Aug-17	21.1	6.7	69.8	17.2	43.1097	-89.6408
Brewery Creek	Prairie Peninsula	11-Aug-17	28.2	3.9	18.6	81.5	43.125	-89.635
McDiffitt Creek	Prairie Peninsula	12-Jun-18	23.1	3.4	14.1	92.2	38.94428	-96.44197
Kings Creek	Prairie Peninsula	13-Jun-18	13.2	3.2	21.7	85.8	39.10506	-96.60336
Kings Creek	Prairie Peninsula	10-May-19	13.2	6.5	33.5	63.3	39.10506	-96.60336

Table 4.1b: Physical environmental characteristics of a subset of the 30 study streams in 13 ecoregions by sampling date. Lat. is latitude and Long. is longitude. Sites are arranged west to east by ecoregion.

Site Name	Ecoregion	Sample Date	Watershed Area (km^2)	Average reach width (m)	Average transect depth (cm)	Average canopy cover (%)	Lat. (°N)	Long. (°W)
Pilgrim River	Great Lakes	21-Jun-18	63.0	8.6	58.8	9.1	47.1013	-88.5176
Pilgrim River	Great Lakes	17-Jul-18	63.0	8.4	58.0	17.6	47.1013	-88.5176
Pilgrim River	Great Lakes	13-Jun-19	63.0	9.1	59.0	14.4	47.1013	-88.5176
Pilgrim River	Great Lakes	9-Sep-19	63.0	10.0	53.8	13.1	47.1013	-88.5176
Ichetucknee River	Southeast	31-May-18	991.0	25.3	108.4	90.0	29.9526	-82.7863
Alexander Springs	Southeast	2-Jun-18	285.0	75.0	67.2	25.7	29.0812	-81.5662
New Hope Creek	Mid-Atlantic	3-Jun-17	80.8	14.2	24.8	94.2	35.9795	-79.0018
Eno River	Mid-Atlantic	4-Jun-17	171.2	14.9	48.7	79.8	36.0715	-79.0968
New Hope Creek	Mid-Atlantic	27-May-18	80.8	57.3	25.3	90.3	35.9795	-79.0018
Hubbard Brook	Northeast	7-Aug-18	0.4	10.7	15.2	78.0	43.9549	-71.7225
Wednesday Hill Brook	Northeast	9-Aug-18	1.0	1.7	5.9	92.7	43.1222	-71.0049
Dowst Cast Forest	Northeast	10-Aug-18	7.0	6.6	15.2	95.6	43.1344	-71.1838
Hop Brook	Northeast	15-Aug-18	12.8	5.2	27.1	94.5	42.47179	-72.32963
Rio Guillarte	Atlantic Neotropical	21-Feb-19	10.2	4.6	26.3	40.5	18.1741	-66.79851
Rio Cupeyes	Atlantic Neotropical	23-Feb-19	4.9	3.4	8.7	82.4	18.11352	-66.98676
Prieta Stream	Atlantic Neotropical	3-Mar-19	NA	3.1	11.9	87.2	18.324	-65.8151
Sonadora Stream	Atlantic Neotropical	5-Mar-19	2.6	11.8	25.0	81.1	18.3213	-65.8171

Table 4.1c: Nutrient concentrations of a subset of the 30 study streams in 13 ecoregions by sampling date. Sites are arranged west to east by ecoregion. DOC is dissolved organic carbon, TDN is total dissolved nitrogen, NH4 is ammonium, NO3 is nitrate, SRP is soluble reactive phosphorus, TDP is total dissolved phosphorus, TP is total phosphorus, DIN is dissolved inorganic nitrogen, and DIN:TDP is the ratio of DIN to TDP.

Site Name	Ecoregion	Sample Date	DOC (mg / L)	TDN (mg / L)	NH4 (ug / L)	NO3 (ug / L)	SRP (ug / L)	TDP (ug / L)	TP (ug / L)	DIN (ug / L)	DIN:TDP
Caribou Creek	Taiga	9-Jul-19	4.7	0.4	9.5	285.2	7.8	4.6	4.1	294.7	64.1
Poker Creek	Taiga	10-Jul-19	4.5	0.3	4.1	261.7	5.4	4.6	7.4	265.7	57.8
McRae Creek	Pacific Northwest	3-Aug-19	1.0	0.0	1.1	10.7	7.8	4.6	7.4	11.8	2.6
Martha Creek	Pacific Northwest	6-Aug-19	1.7	0.2	4.8	43.4	33.2	28.6	22.4	48.2	1.7
South Fork Mink Creek	Northern Rockies	27-Jul-17	2.1	0.4	18.2	56.1	12.6	42.3	75.3	74.3	1.8
Blacktail Deer Creek	Northern Rockies	30-Jul-19	3.1	0.1	4.6	41.9	5.0	37.4	39.1	46.5	1.2
Creston Creek	Northern Rockies	10-Aug-19	14.3	3.6	18.4	3667.3	10.6	4.6	10.7	3685.7	801.2
Lower Knowton Fork Red Butte Creek	Great Basin	18-Jul-17	0.7	0.1	4.5	54.1	3.1	24.0	58.4	58.6	2.4
1300 E Red Butte Creek	Great Basin	19-Jul-17	2.5	1.1	8.2	919.2	3.1	19.3	33.0	927.4	48.1
Sycamore Creek	Desert Southwest	14-May-19	7.9	0.2	9.2	97.6	36.6	64.0	60.8	106.8	1.7
Wet Beaver Creek	Desert Southwest	16-May-19	4.5	0.1	3.8	37.4	23.9	26.8	24.1	41.2	1.5
Arikaree River	Central Plains	11-Jul-17	6.2	0.3	59.5	51.3	18.0	21.1	109.2	110.8	5.3
Arikaree River	Central Plains	24-May-19	14.9	0.4	5.1	12.6	31.5	12.6	14.1	17.7	1.4
Pringle Creek	Southern Plains	9-Jun-18	7.7	0.2	6.7	5.7	6.4	1.2	17.1	12.4	10.3
Black Earth Creek	Prairie Peninsula	8-Aug-17	5.1	2.5	7.7	2510.0	3.1	12.2	46.6	2517.7	206.4
Brewery Creek	Prairie Peninsula	11-Aug-17	8.1	2.6	40.8	3295.0	3.1	52.8	132.9	3335.8	63.2
McDiffitt Creek	Prairie Peninsula	12-Jun-18	16.9	0.9	84.8	98.7	2.9	23.8	57.0	183.6	7.7
Kings Creek	Prairie Peninsula	13-Jun-18	14.6	0.2	74.7	5.7	8.9	42.3	103.9	80.4	1.9
Kings Creek	Prairie Peninsula	10-May-19	6.9	0.2	1.3	56.6	10.5	10.8	15.7	57.9	5.4

Table 4.1d: Nutrient concentrations of a subset of the 30 study streams in 13 ecoregions by sampling date. Sites are arranged west to east by ecoregion. DOC is dissolved organic carbon, TDN is total dissolved nitrogen, NH4 is ammonium, NO3 is nitrate, SRP is soluble reactive phosphorus, TDP is total dissolved phosphorus, TP is total phosphorus, DIN is dissolved inorganic nitrogen, and DIN:TDP is the ratio of DIN to TDP.

Site Name	Ecoregion	Sample Date	DOC (mg / L)	TDN (mg / L)	NH4 (ug / L)	NO3 (ug / L)	SRP (ug / L)	TDP (ug / L)	TP (ug / L)	DIN (ug / L)	DIN:TDP
Pilgrim River	Great Lakes	21-Jun-18	4.8	0.4	5.4	149.0	12.2	11.5	18.8	154.4	13.4
Pilgrim River	Great Lakes	17-Jul-18	4.0	0.3	NA	115.0	5.2	NA	NA	115.0	NA
Pilgrim River	Great Lakes	13-Jun-19	8.2	0.2	5.7	40.7	50.5	10.8	19.1	46.5	4.3
Pilgrim River	Great Lakes	9-Sep-19	4.6	0.2	6.0	50.0	6.5	10.8	9.1	56.0	5.2
Ichetucknee River	Southeast	31-May-18	5.0	0.6	4.6	790.5	2.9	19.7	24.0	795.1	40.4
Alexander Springs	Southeast	2-Jun-18	11.2	0.4	12.9	28.4	41.1	37.6	39.7	41.3	1.1
New Hope Creek	Mid-Atlantic	3-Jun-17	3.2	0.3	16.9	102.0	12.3	12.6	14.9	118.9	9.4
Eno River	Mid-Atlantic	4-Jun-17	5.2	0.7	25.6	436.5	28.1	19.3	22.9	462.0	23.9
New Hope Creek	Mid-Atlantic	27-May-18	NA	NA	14.7	134.0	9.1	NA	22.3	148.7	NA
Hubbard Brook	Northeast	7-Aug-18	4.1	0.2	1.6	22.1	2.9	1.2	2.7	23.8	19.8
Wednesday Hill Brook	Northeast	9-Aug-18	8.6	0.8	1.6	494.7	9.1	19.7	18.8	496.3	25.2
Dowst Cast Forest	Northeast	10-Aug-18	15.0	0.7	1.6	5.7	2.9	13.6	18.8	7.4	0.5
Hop Brook	Northeast	15-Aug-18	6.2	0.2	3.6	5.7	2.9	9.5	13.6	9.3	1.0
Rio Guillarte	Atlantic Neotropical	21-Feb-19	2.0	0.3	6.4	268.8	12.9	25.0	19.1	275.2	11.0
Rio Cupeyes	Atlantic Neotropical	23-Feb-19	5.9	0.4	2.0	368.5	17.5	4.6	7.4	370.5	80.5
Prieta Stream	Atlantic Neotropical	3-Mar-19	2.4	0.2	12.4	120.6	4.9	4.6	12.4	133.0	28.9
Sonadora Stream	Atlantic Neotropical	5-Mar-19	2.7	0.2	3.4	116.8	4.7	4.6	5.7	120.2	26.1

4.3.2.1 Chamber Setup

Chambers used for the acetylene reduction and acetylene block assays varied by substrate type. 2-L polycarbonate food storage containers were used for rock and larger macrophyte substrate (Gettel et al. 2007, Eberhard et al. 2018). The chamber lids were sealed airtight with a Viton o-ring, and were fit with a 13x20 mm septa for sample collection. For sediment, wood, and smaller macrophyte substrate, chambers were made from pint size glass mason jars and lids were similarly fit with an airtight sampling septa.

Rock substrate was collected by haphazardly sampling rocks from the study area and placing them in the polycarbonate chamber until its bottom was covered. Sediment substrate was collected haphazardly within each stream using a 7 cm diameter suction corer to collect ~200-400 mL of sediment that was then placed into the mason jars. Macrophytes were collected using chamber lids to approximate surface area of macrophyte to sample, then pulling from the root and placing in chambers. Wood was collected by haphazardly sampling wood until the bottom of a mason jar was mostly covered.

4.3.2.2 Environmental Characteristics

To test the hypothesis that variation in environmental characteristics in each biome would drive differences in microbial community composition, we collected 7 ~40 mL water samples from each stream. The water was filtered using Millipore 0.45 μm nitrocellulose membrane filters into 60 mL bottles. Samples were frozen until later laboratory analysis. TDN and DOC samples were acidified to a pH < 2 and sent to Michigan Tech's Laboratory for Environmental Analysis of Forests (LEAF) core facility

which used a Shimadzu 210 TOC-VCSN with a total N module TNM-1 (Shimadzu Scientific Instruments, Columbia, Maryland). NH_4^+ was analyzed using a fluorometric method (Holmes et al. 1999, Taylor et al. 2007) on a Turner Aquafluor (Turner Designs, Palo Alto California). NO_3^- samples were analyzed on a SEAL AQ_2 discrete water analyzer using the AQ_2 method EPA-127-A Rev. 9. DIN concentration was then calculated by adding concentrations of NH_4^+ and NO_3^-. SRP samples were analyzed on a SEAL AQ_2 discrete water analyzer using the sed AQ_2 method EPA-155-A Rev. 0. TDP samples were analyzed on a Thermo Scientific 10s UV-Vis spectrophotometer in using the ascorbic acid method and molybdenum antimony colorimetric determination methods (APHA 2005). For TDP samples, an ammonium persulfate digestion was used prior to this analysis. TP samples used the same method as TDP, but with unfiltered samples.

To further test our hypothesis regarding environmental variables as drivers of differences in microbial community structure we measured depth at 10 points across transects located every 10 m along the sampling reach in every stream and the width of each transect was measured as well. Canopy cover (%) was measured in the middle of each transect using a spherical densiometer (Lemmon 1956).

4.3.2.3 Measurement of N transformation rates

N_2 fixation rates were measured using acetylene reduction (Capone 1993, Dodds et al. 2017). An acetylene-filled balloon was added to each chamber. Chambers were filled with stream water and sealed underwater, then balloons were popped with a needle through the sampling septum to introduce a 20% acetylene headspace. Chambers were then shaken for approximately 20 seconds to equilibrate the gas dissolved in the water

with that in the headspace. Initial gas samples were collected within 5 minutes of sealing the chambers. Chambers were placed in the stream for a 2-hour incubation to maintain ambient stream temperatures. Chambers were shaken again to equilibrate and then final samples were collected. All gas samples were placed into evacuated 9-mL serum vials and kept in the dark until analyzed. Ethylene concentrations were measured using a SRI 8610C gas chromatograph equipped with a Hayesep T column, He carrier gas, and a flame ionization detector. The column oven was set to 40 °C. To obtain N_2 fixation rates, ethylene concentrations in the chambers were compared to 100 ppm ethylene standards (Matheson Tri Gas). N_2 fixation rates were calculated following Capone (1993) and Dodds et al. (2017), then converted to µg of N assuming a ratio of 3 mols of ethylene produced for every 1 mol of N_2 gas potentially fixed (Capone 1993).

Denitrification rates were measured using the acetylene block method (Groffman et al. 2006). Chloramphenicol was used to suppress additional protein synthesis during the incubation in all chambers. We measured nutrient-amended, potential rates because most previous stream studies have used this method and we wanted to be able to compare estimates to these studies, and because this method is quick and easy to run with a large number of replicates to estimate rate variability. Moreover, the acetylene block method also inhibits nitrification, so measuring without amendments of nitrate can underestimate denitrification rates (Dodds et al. 2017). However, the chambers were not sparged with nitrogen or helium to create anoxic conditions in these denitrification rates measurements. Each chamber received 0.62 g L^{-1} Glucose as a C source and 0.62 g L^{-1} $NaNO_3$ as an N source, plus chloramphenicol (2 g L^{-1}). After the amendment, acetylene was introduced, chambers were incubated, and initial and final gas samples were

collected as described previously for N_2 fixation. Nitrous oxide (N_2O) concentrations were measured using a SRI 8610C gas chromatograph equipped with a Hayesep D column, He carrier gas, and an electron capture detector. The column oven was set to 40 °C. N_2O concentrations in chambers were compared to standard concentrations of 1000 ppm N_2O (Matheson Tri Gas). Denitrification rates were calculated following Dodds et al. (2017).

In most cases, in each stream there were 2-4 sample chambers and 2-3 blank chambers per assay and substrate type. The number of sample and blank chambers depended on what substrate type was dominant in the stream study reach. More dominant substrates had more chamber replicates per assay. The blank chambers were set up to simulate an environment with minimal N_2 fixation or denitrification to control for chamber effects. Materials used for the blanks were selected based on their relative specific heats to mimic the specific heats of incubated substrates to correct for changes in temperature. Rocks found on the shore near the stream were used for blanks for stream rocks, and stream water was used as a blank for sediment, wood, algae, and macrophyte substrates. The sample chambers had stream rock, sediment, wood, or macrophyte placed in them as described above.

Surface area and volume of all substrates were measured to scale process rates for surface area. Surface area for rocks and wood was determined by weighing tracings of the sampled rocks. The weights were then compared to a standard curve to calculate area (Bergey and Getty 2006). Sediment surface area was calculated as the diameter of the corer. Macrophyte and algae surface area was calculated as the diameter of the chamber

lid. Rock volume was determined using displacement and sediment volume was determined by multiplying the surface area by average sediment core depth in the jar.

To evaluate how the composition of potential N_2-fixing and denitrifying microorganisms in each stream differed in comparison to overall rates of N_2 fixation and denitrification in streams, the reach average rates for each stream were calculated. Stream reach average rates were only calculated for NEON streams (n = 12 of the 30 total) because only these streams had percent cover of each substrate type on the streambed available as part of their dataset for each year. Percent cover for each substrate type was calculated in each stream using the equation:

$$percentCover_i = \frac{N_i}{N_t} \times 100$$

Where N_i is the number of observed points in a transect that match class type "i" and N_t is the total number of points observed in the transect. This calculation can generate percent cover values >100% if there is vertical stacking of plants (https://data.neonscience.org/data-products/DP1.20072.001). For this analysis, algae and macrophyte were lumped as one substrate in order to match the substrate categories to NEON categories for percent cover.

The percent cover was then used to scale average chamber measurements of each process to whole reach rates by multiplying average chamber rates by the average percent cover of the substrate in the stream, respectively for each substrate present. All of the scaled substrate rates were then added together to get whole stream reach-scaled rates for each stream.

4.3.2.4 Microbial Sampling and Illumina Sequencing

At the end of each sampling day, substrate samples (biofilm, sediment, wood, algae, or macrophyte) were removed from each sample chamber and placed in sterile 15 mL falcon tubes. Rocks were removed from the chambers and scrubbed, and 12 mL of scrub water was poured into the falcon tubes to collect biofilm samples. 8-mL sediment cores were taken from sediment chambers using a 10 mL syringe and placed into the falcon tubes. Wood substrates were sampled by using a pocketknife to cut off ~ 4 surface shavings from each stick in a chamber and placing those into falcon tubes with chamber water. Macrophyte and algae were sampled by tearing off a small part of the macrophyte or algae mat and placing it in a falcon tube with chamber water. All 15 mL falcon tubes were placed in a mobile -20°C freezer after collection and moved to a -10° C freezer upon return to the lab. DNA from each sample was extracted using the Power-soil DNA Isolation Kit (MO Bio) and stored in a -20°C freezer until analysis.

Samples that were selected for use in the 16S rRNA sequencing analysis came from sample chambers that underwent acetylene reduction assays because they had no additions of C, N, or chloramphenicol that may have altered the microbial community. However, samples from these chambers were exposed to acetylene, which has been shown to alter microbial community structure after 7-hour incubations (Fulweiler et al. 2015). The overall samples for 16S rRNA sequencing analysis had 1-4 samples per substrate depending on the dominant substrate of each stream (Table 4.2a-b).

Table 4.2a: The substrate type sampling breakdown of a subset of the 30 streams. The substrate types were rock (which included cobble and pebble), sediment, wood, macrophyte, algae, and coarse particulate organic matter (CPOM). If a "-" is present it means those samples were collected but did not make it through sequencing analysis either as they did not amplify or the sample was contaminated. "NA" means that substrate type was not dominant in the stream and was not collected.

Site Name	Ecoregion	Sample Date	# of Rock Samples	# of Sediment Samples	# of Wood Samples	# of Macrophyte Samples	# of Algae Samples	# of CPOM Samples
Caribou Creek	Taiga	9-Jul-19	-	2	2	NA	NA	NA
Poker Creek	Taiga	10-Jul-19	1	3	2	NA	NA	NA
McRae Creek	Pacific Northwest	3-Aug-19	-	1	1	NA	NA	NA
Martha Creek	Pacific Northwest	6-Aug-19	-	2	2	NA	NA	NA
South Fork Mink Creek	Northern Rockies	27-Jul-17	4	2	1	2	NA	NA
Blacktail Deer Creek	Northern Rockies	30-Jul-19	2	2	2	NA	NA	NA
Creston Creek	Northern Rockies	10-Aug-19	2	1	2	2	NA	NA
Lower Knowton Fork Red Butte Creek	Great Basin	18-Jul-17	1	2	2	NA	NA	NA
1300 E Red Butte Creek	Great Basin	19-Jul-17	-	NA	-	NA	NA	NA
Sycamore Creek	Desert Southwest	14-May-19	1	-	2	NA	2	NA
Wet Beaver Creek	Desert Southwest	16-May-19	3	-	2	NA	NA	NA
Arikaree River	Central Plains	11-Jul-17	NA	7	NA	5	NA	NA
Arikaree River	Central Plains	24-May-19	NA	3	NA	NA	3	NA
Pringle Creek	Southern Plains	9-Jun-18	1	3	1	NA	NA	NA
Black Earth Creek	Prairie Peninsula	8-Aug-17	2	2	2	2	NA	NA
Brewery Creek	Prairie Peninsula	11-Aug-17	NA	3	2	2	NA	NA
McDiffitt Creek	Prairie Peninsula	12-Jun-18	2	3	1	NA	NA	NA
Kings Creek	Prairie Peninsula	13-Jun-18	2	2	2	NA	NA	NA
Kings Creek	Prairie Peninsula	10-May-19	-	3	2	NA	NA	NA

Table 4.2b: The substrate type sampling breakdown of a subset of the 30 streams. The substrate types were rock (which included cobble and pebble), sediment, wood, macrophyte, algae, and coarse particulate organic matter (CPOM). If a "-" is present it means those samples were collected but did not make it through sequencing analysis either as they did not amplify or the sample was contaminated. "NA" means that substrate type was not dominant in the stream and was not collected.

Site Name	Ecoregion	Sample Date	# of Rock Samples	# of Sediment Samples	# of Wood Samples	# of Macrophyte Samples	# of Algae Samples	# of CPOM Samples
Pilgrim River	Great Lakes	21-Jun-18	-	2	1	NA	NA	NA
Pilgrim River	Great Lakes	17-Jul-18	1	2	2	NA	NA	NA
Pilgrim River	Great Lakes	13-Jun-19	1	2	2	NA	NA	NA
Pilgrim River	Great Lakes	9-Sep-19	-	1	-	NA	NA	NA
Ichetucknee River	Southeast	31-May-18	NA	2	-	3	NA	NA
Alexander Springs	Southeast	2-Jun-18	NA	1	2	4	NA	NA
New Hope Creek	Mid-Atlantic	3-Jun-17	3	2	1	NA	NA	NA
Eno River	Mid-Atlantic	4-Jun-17	1	3	2	NA	NA	NA
New Hope Creek	Mid-Atlantic	27-May-18	1	3	2	NA	NA	NA
Hubbard Brook	Northeast	7-Aug-18	-	NA	4	NA	NA	NA
Wednesday Hill Brook	Northeast	9-Aug-18	4	6	4	NA	NA	NA
Dowst Cast Forest	Northeast	10-Aug-18	2	3	5	NA	NA	NA
Hop Brook	Northeast	15-Aug-18	1	1	2	NA	NA	NA
Rio Guillarte	Atlantic Neotropical	21-Feb-19	4	NA	NA	1	1	-
Rio Cupeyes	Atlantic Neotropical	23-Feb-19	4	NA	NA	NA	NA	1
Prieta Stream	Atlantic Neotropical	3-Mar-19	-	1	2	NA	NA	NA
Sonadora Stream	Atlantic Neotropical	5-Mar-19	-	NA	2	NA	NA	NA

In total, 272 DNA samples were prepped for submission to the Michigan State University Genomics Core for 16S-V4 amplicon Illumina sequencing. Samples were first quantified using Qubit dsDNA HS assay kit (Invitrogen) and normalized to approximately the same concentration of 1-10 ng/μL. A test PCR amplification was performed using a mixture of 1 μL DNA, 6.5 μL of 0.5 μM primer mix of the 515f/926r primer pair for the 16S V4-V5 region, and 7.5 μL of 2X Phusion Flash High-Fidelity PCR Master Mix. The PCR cycling conditions were 95° C for 3 minutes, then 30 cycles of denaturation at 95° C for 45 s, annealing at 50° C for 60 s, and elongation at 72° C for 90 s, followed by 72° C for 10 minutes. PCR products were run on a 2% agarose gel to confirm amplification and size. Samples that did not amplify were not sent out for analysis (n = 62). Overall, 206 confirmed DNA samples were sent to the Michigan State University Genomics Core where samples underwent library preparation for 16S-V4 amplicon using the primer pair 515f/806r (Kozich et al. 2013) and then Illumina sequencing using MiSeq v2 Standard 500 cycle (2x250bp paired end). Sequence data were demultiplexed and converted to FastQ format. The sequence data was imported into RStudio (R version 3.6.0) with a sum of 9,746,880 reads and a mean of 47,315 reads. An amplicon sequence variant (ASV) table and taxa table were made using the *dada2 version 1.14.1* package and SILVA version 138 to assign taxonomy (Quast et al. 2012, Yilmaz et al. 2014, Callahan et al. 2016). These tables were combined into one single object using the *phyloseq* package (McMurdie and Holmes 2013). The data was normalized and rarified at 1000 reads with a seed of 81 and a sample size of 2202. After being rarified there were 47,159 taxa and 205 samples in total.

4.3.2.5 Statistical Analysis

To assess our first hypothesis that microbial community taxonomy would differ in biomes that have extreme differences in environmental variables we examined the observed (richness) and Shannon alpha diversity of ASVs in each ecoregion. A Kruskal-Wallis test was performed to assess if there was a difference in the observed alpha diversity among ecoregions and also if there was a difference in the Shannon alpha diversity among ecoregions. A post-hoc Dunn test was performed with Bonferroni correction to assess differences in alpha diversity between pairs of ecoregions. To further assess the composition of the microbial communities across ecoregions, we compared relative abundances at the Class level for each ecoregion.

We performed a distance-based redundancy analysis (db-RDA) or a constrained analysis of principle coordinates to evaluate if environmental variables across each ecoregion could explain the variability in ASVs. The method of db-RDA tries to detect linear relationships on dissimilarities using non-Euclidean distance. The db-RDA takes a dissimilarity matrix and creates a principal coordinates analysis (PCoA) on the matrix. The eigenvalues generated in the PCoA are then put into an RDA that is constrained by a matrix of explanatory variables (Legendre and Anderson 1999). The matrix of explanatory variables for this analysis included the variables watershed area (km^2), average stream reach width (m), average transect depth (cm), average canopy cover (%), DOC (mg/L), TDN (mg/L), NH_4^+ (µg/L), NO_3^- (µg/L), SRP (µg/L), TDP (µg/L), TP (µg/L), DIN (µg/L) and DIN:TDP. Prior to analysis, all samples with NA values for any of the explanatory variables were removed from the matrix because this test does not handle missing data. This resulted in the removal of 23 of the 205 samples, which were

from the 2018 sampling of New Hope Creek and the Pilgrim River and the 2019 sampling of Rio Prieta and Poker Creek (Table 1). The ordination used the Bray-Curtis distance method. A permutational analysis of variance (ANOVA) was then used on the constrained axes of the ordination to evaluate if the db-RDA was significant. To test if the groupings in the ordination were significantly different by ecoregion (beta diversity), a permutational multivariate analysis of variance (PERMANOVA) was performed with Bray-Curtis distance. All analyses were performed using the *phyloseq* and *vegan* (Oksanen et al. 2020) packages in RStudio (R version 4.1.2)

To evaluate if streams with higher reach average rates of N_2 fixation or denitrification would have a higher % composition of microorganisms functionally capable of each process, we used FAPROTAX to map our taxa to metabolic or ecologically relevant functions based on a review of the literature of cultured species. FAPROTAX uses Python script to convert taxa tables to functional tables (Louca et al. 2016). This program is non-exhaustive and is not specific to aquatic habitats. Taxon can be listed as multiple functions if they are known for multiple types of metabolisms and functions can be nested, meaning that if an organism is known to carry out one part of a process like the 4-step process of denitrification then it will also be listed as capable of performing the whole denitrification process. The functional table produced from FAPROTAX was then used to assess the average % of N_2 fixation and denitrification functional groups detected in each stream. Only data for the 12 NEON streams (Arikaree River, Pringle Creek, McDiffet Creek, Kings Creek, Hop Brook, Rio Guillarte, Rio Cupeyes, Sycamore Creek, Caribou Creek, Blacktail Deer Creek, McRae Creek, and Martha Creek) were used because they were the only streams where stream reach average

rates of N_2 fixation and denitrification were able to be calculated. A Spearman's rank correlation was then used in RStudio (R version 4.1.2) to evaluate the differences between average % of N_2 fixation and denitrification functional groups to stream reach average rates of both processes.

4.4 Results

4.4.1 Taxonomic Composition and Diversity

When evaluating our first hypothesis that microbial communities would have different taxonomical compositions in biomes with extreme differences in environmental variables we observed differences in the relative abundances at the Class level of the microbial assemblages across the ecoregions (Fig. 4.2).

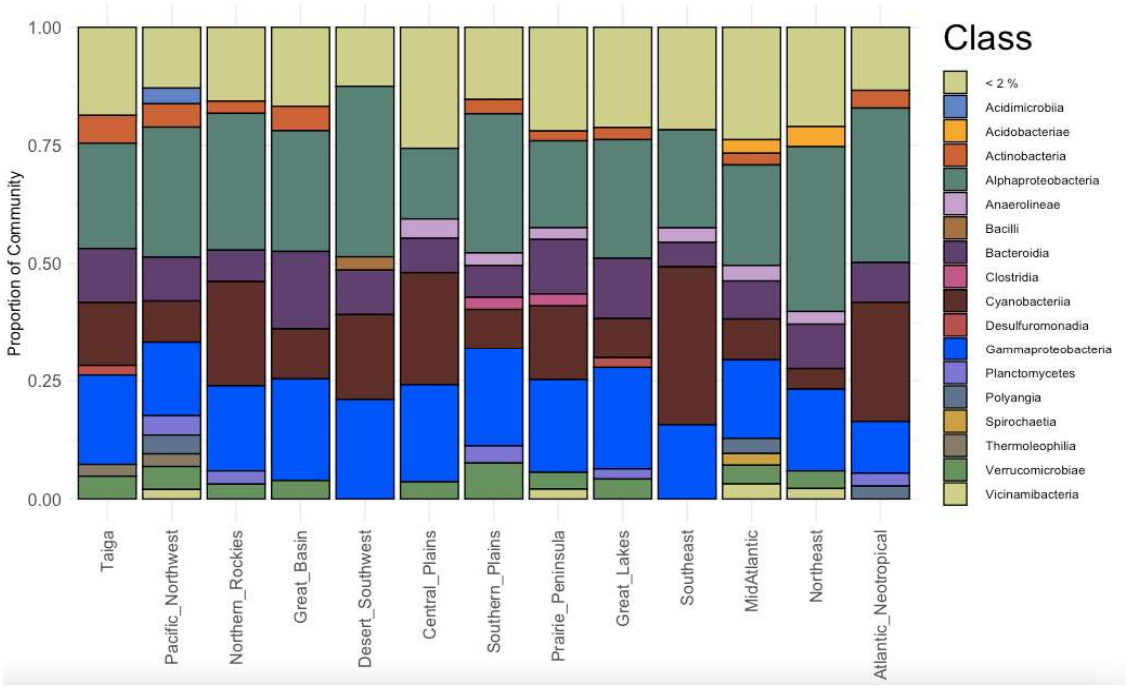

Figure 4.2: Relative abundances of microbial Classes by ecoregion. All Classes that contributed <2% to the overall relative abundance of the microbial community were grouped and labled "<2%". The x-axis is arranged in direction of ecoregion location West to East.

The three most prominent Classes present in all ecoregions were Alphaproteobacteria, Cyanobacteria, and Gammaproteobacteria. Alphaproteobacteria made up 15 - 36% of community compositions and Gammaproteobacteria made up 11-22% of the community compositions. Cyanobacteria made up 4.4 – 33.6% of total community compositions with the lowest percentage in the Northeast and the highest in the Southeast. Bacteroidia were also present in abundance in all ecoregions and made up 5 – 16% of community compositions. Actinobacteria was observed in abundance in all ecoregions except the Central Plains, Desert Southwest, and Southeast. When Actinobacteria were present in an ecoregion they made up 2.1 to 5.9% of the total microbial community composition. Anaerolineae were present in abundance in the Central Plains (4.1%), Mid Atlantic (3.3%), Northeast (2.7%), Prairie Peninsula (2.5%), Southeast (3.1%), and Southern Plains (2.7%). Planctomycetes was present in the Atlantic Neotropical (2.7%), Great Lakes Basin (2.1%), Northern Rockies (2.8%), Pacific Northwest (4.1%), and Southern Plains (3.6%) ecoregions. Polyangia was present in abundance in the Atlantic Neotropical (2.7%), Mid Atlantic (3.2%), and Pacific Northwest (4.0%). Verrucomicrobiae were present in abundance (3.1-7.6%) in all ecoregions, but the Atlantic Neotropical, Desert Southwest, and Southeast. Acidomicrobiia were observed in abundance only in the Pacific Northwest, making up 3.3% of the overall microbial community. Thermoleophilla was of abundance in the Pacific Northwest (2.6%) and Taiga (2.5%) ecoregions. Clostridia were only present in abundance in the Pacific Northwest and Southeast at 2.5% and 2.6% respectively. The Desert Southwest was the only ecoregion with Bacilli in abundance at 2.8%. Desulfuromonadia were only present in abundance in the Great Lakes and Taiga

ecoregions at 2%. Spirochaetia was only of abundance in the Mid Atlantic at 2.5% of the community composition. Acidobacteriae was observed in abundance only in the Mid Atlantic (2.9%) and Northeast (4.2%). Vicinamibacteria were only present in abundance in the Mid Atlantic (3.2%), Northeast (2.2%), Pacific Northwest (2.0%), and Prairie Peninsula (2.1%). Across all ecoregions, Classes that were <2% of the total community relative abundances made up 12.5 to 25.6% of each ecoregions' total microbial community composition; these Classes are not discussed in detail here.

When examining differences in the microbial community composition of the ASVs among the 13 ecoregions, we found observed alpha diversity or richness of taxonomy ranged from 54 to 970 (standard deviation (s.d.) = ± 199) and Shannon alpha diversity ranged from 0.95 to 6.58 (s.d. ± 1.07, Fig. 4.3). The ecoregion with the highest average observed alpha diversity was the Great Lakes (636) and the lowest was the Desert Southwest (636). For Shannon alpha diversity, the ecoregion with the highest average diversity was the Pacific Northwest (5.86) and the lowest was in the Northern Rockies (4.75). There were significant differences among the observed alpha diversity for each ecoregion ($\chi^2 = 27.4$, df = 13, p-value = 0.01), as well as among the Shannon alpha diversity for each ecoregion ($\chi^2 = 24.9$, df = 13, p-value = 0.02) using Kruskal-Wallis tests. However, when assessing post-hoc differences between ecoregions, there were no significant differences (p-values > 0.05) for both the observed richness and Shannon alpha diversity.

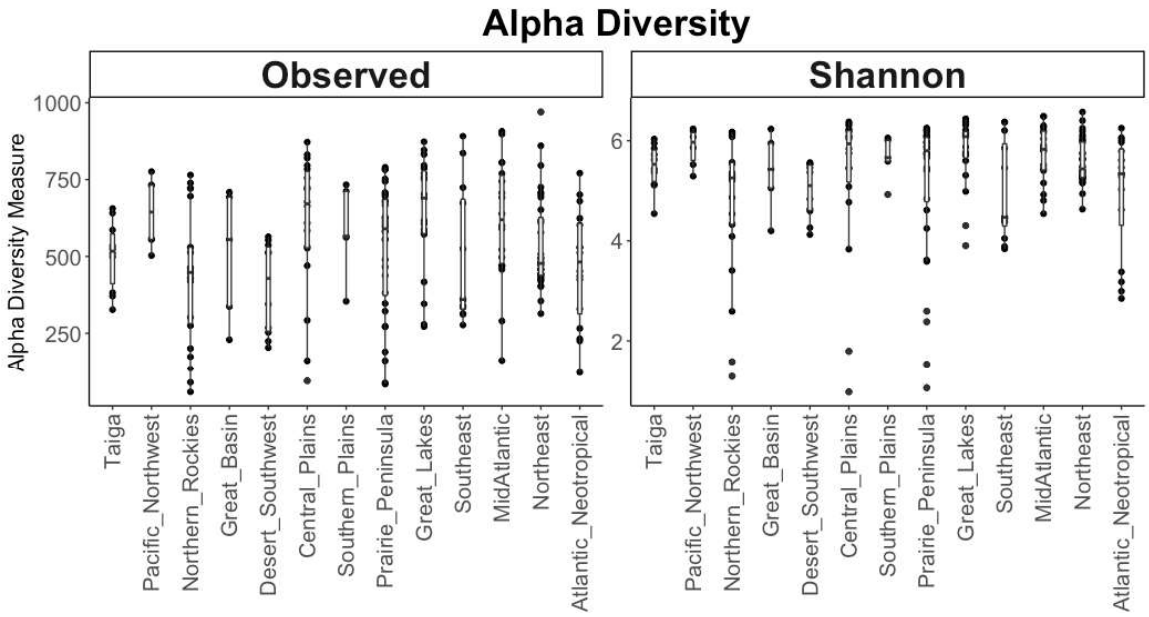

Figure 4.3: Boxplot of the observed richness and Shannon alpha diversity for each ecoregion. The x-axis is arranged in direction of ecoregion location from west to east. The boxplots represent median values with upper and lower hinges corresponding to the 25th and 75th percentiles. The whiskers extend ± 1.5 * interquartile range and all points beyond that are considered outliers.

4.4.2 Environmental Characteristics Across Ecoregions

Across the 12 ecoregions and 30 study streams there was variability in the environmental characteristics (Table 4.1a-d). The watershed area ranged from 0.4 to 2890 km^2. The lowest canopy cover was 0% in Creston Creek in the Northern Rockies and the highest was 95% in Dowst Cast Forest Stream in the Northeast. The lowest DOC concentration of 1.0 mg/L was observed in McRae Creek in the Pacific Northwest, while the highest concentration of 16.9 mg/L was observed in McDiffett Creek in the Prairie Peninsula. NH_4^+ concentrations ranged from 1.1 to 84.8 µg/L, with the 4 highest concentrations occurring in the Central Plains or Prairie Peninsula ecoregions. NO_3^- concentrations ranged from 5.7 to 3667 µg/L, with a median of 98.2 µg/L and the highest

concentrations occurred in streams of the Great Basin, Northern Rockies, and Prairie Peninsula. TDP concentrations ranged from 1.2 to 64.0 µg/L with a median of 13.1 µg/L and the 2 highest concentrations occurred in streams in the Desert Southwest and Northern Rockies. TP concentrations ranged from 2.7 to 132.9 µg/L with a median of 19.1 µg/L and all concentrations > 100 µg/L occurred in the Prairie Peninsula and Central Plains. When looking at the relation of N to P, DIN:TDP ranged from 0.5 to 801.2 µg/L, with a median of 8.6 µg/L. Streams with DIN:TDP > 100 occurred in the Prairie Peninsula and Northern Rockies.

4.4.3 Environmental Characteristics as Drivers of Community Composition

In contrast to the first hypothesis that ecoregions with more extreme environmental characteristics would show the most differences in taxonomy, we found that differences in environmental variables across ecoregions did not explain much of the variability in microbial community structure (Fig. 4.4). The db-RDA was statistically significant (p-value < 0.01) and had 22 constrained axes that explained 27% of the total variation. The two most important axes of the db-RDA explained 4.2% and 2.8% of the overall variability in microbial community structure (Eigenvalues = 3.47 and 2.35). From the db-RDA, the environmental vectors that had the most impact (largest vector magnitude) on the structure of the microbial communities were substrate type (rock, sediment, wood, or macrophyte), NH_4^+, TP, watershed area, and average canopy cover. PERMANOVA analysis showed there were significant groupings of the ASVs by ecoregions (p-value < 0.01), but that these groupings only explained 17.6% of the variance in the data. The Northeast tended to group together in the direction of the

vectors for wood substrate and average canopy cover. The Desert Southwest and Atlantic Neotropical were grouped in the direction of the vectors for macrophyte and rock substrate. The Central Plains and Prairie Peninsula tended to group in the directions of the vectors for NH_4^+, TP, watershed area, and sediment substrate.

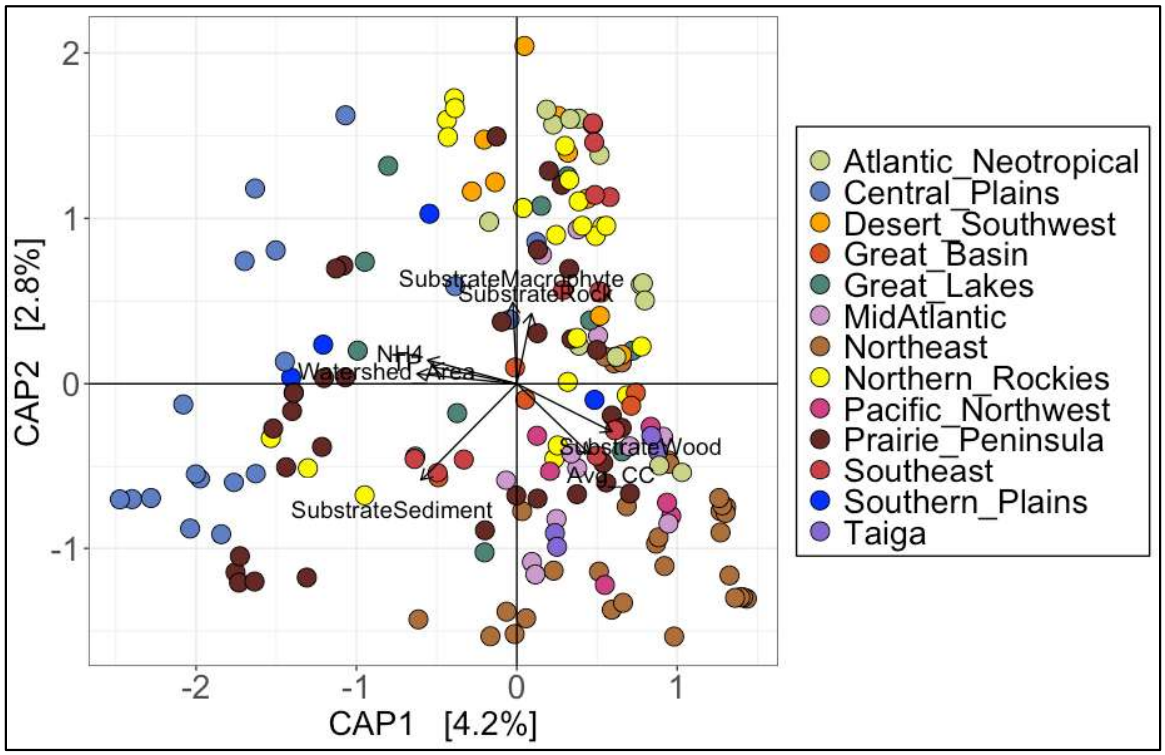

Figure 4.4: Ordination of the principal coordinates analysis and distance based redundancy analysis of the taxa for each ecoregion. Color coding of each dot is based on the ecoregion from which that sample came from. Environmental vectors shown are those had the longest arrows. NH4 is ammonium, TP is total phosphorus**,** and Avg_CC is average canopy over.

4.4.4 % N Fixers and Denitrifiers in Comparison to Average Stream Reach Rates

When evaluating our second hypothesis that streams with higher reach average rates of N_2 fixation and denitrification would also have higher average % composition of

taxa functionally capable of carrying out these processes, we found that the average % of potential N fixers and denitrifiers were less than 10% of the total community functional potential across all ecoregions. For N_2 fixation, the highest average % of potential N fixers was 10.0 % in McRae Creek, while the lowest was 0.03 % in Rio Cupeyes. Stream reach average N_2 fixation rates ranged from 1.93×10^{-5} mg N/m^2/h to 0.35 mg N/m^2/h with a median of 6.7×10^{-4} mg N/m^2/h (Fig. 4.5). For Denitrification, the highest average % of denitrifiers was 0.08 % in McRae Creek and the lowest was 0 % detected in both Rio Guillarte and Rio Cupeyes. Stream reach average rates of denitrification ranged from 0 mg N/m^2/h to 8.75 mg N/m^2/h with a median of 0.45 mg N/m^2/h (Fig. 4.6). There were no significant relationships between the average % of taxa functionally capable of carrying out N_2 fixation or denitrification and rates of these processes (Spearman's correlation p-value = 0.77, $\rho = -0.09$ and p-value = 0.23, $\rho = -0.36$, respectively).

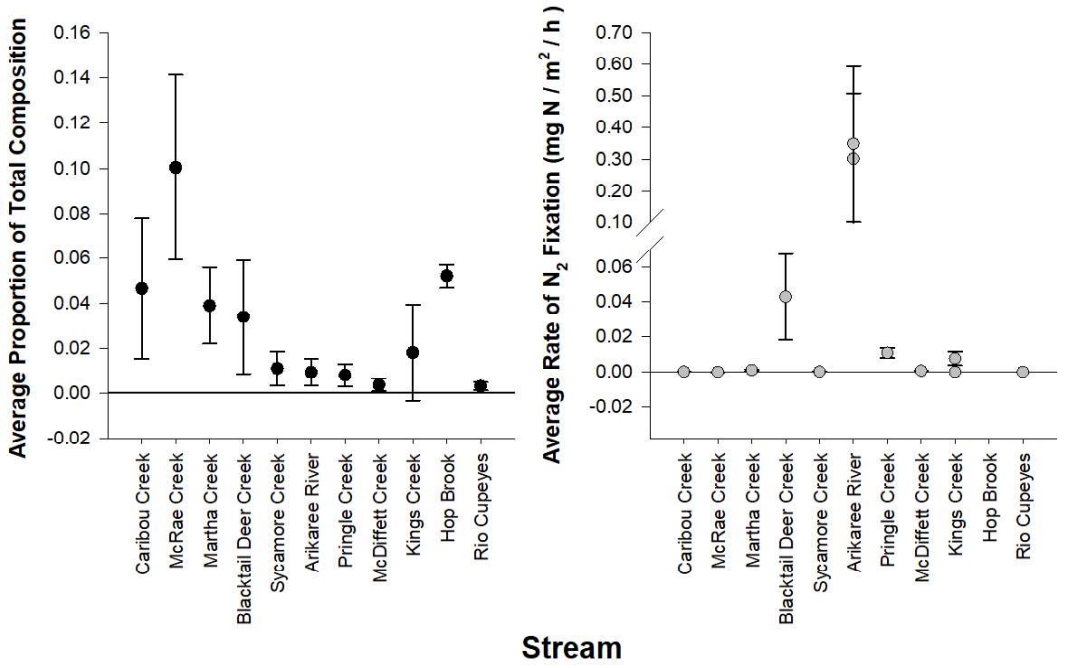

Figure 4.5: Plots of the proportion of potential N_2-fixing taxa and stream reach average rates of N_2 fixation (mg N/m^2/h). The x-axis is arranged in direction of stream location West to East. Error bars are standard deviation.

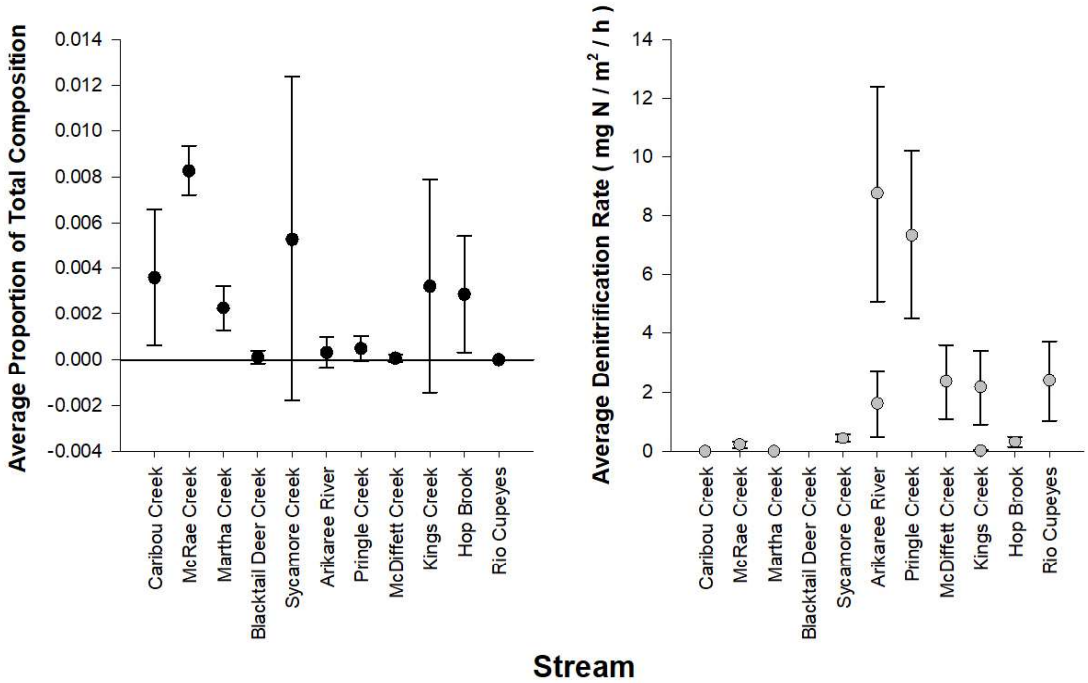

Figure 4.6: Plots of the proportion of potential denitrifying taxa and stream reach average rates of denitrification (mg N/m^2/h). The x-axis is arranged in direction of stream location West to East. Error bars are standard deviation.

4.5 Discussion

Our results showed that there was diversity in the microbial community composition across the 13 ecoregions. The Classes Alphaproteobacteria (15-36%), Cyanobacteria (4.4-33.6%), and Gammaproteobacteria (11-22%) were the most abundant across all ecoregions. We found that ASVs grouped significantly by ecoregions, and that these groupings were mainly related to substrate type, watershed area, canopy cover, NH_4^+, and TP concentrations. However, the ecoregions grouping explained < 20 % of the variation in the ASVs and the axes constrained by environmental variables explained <10 % of the variation. We also found no relationship between % of N_2 fixing and denitrifying bacteria in comparison to stream reach average rates of N_2 fixation and denitrification. These findings highlight that although there is diversity at the ecoregion scale in microbial community structure in streams, much of this diversity cannot be explained by commonly measured stream environmental variables.

When assessing our hypothesis that differences in microbial community composition would be the most apparent across ecoregions with extreme differences in environmental variables, we did find some variability explained by ecoregions grouping. In other large-scale studies of microbial community composition, streambed bacterial community structure has been shown to be more variable across biomes than within using NMDS (Findlay et al. 2008). This has also been observed in other studies where stream pH, quality of fine benthic organic matter, and concentrations of dissolved organic carbon and nitrogen were correlated with relative distance between communities in streams of the Hubbard Brook watershed (Fierer et al. 2007). Bacterial community composition in arctic streams has been shown to be correlated to nutrients, base cations, dissolved

organic carbon, and landscape-scale lithology (Larouche et al. 2012). However, our results indicated that differences in environmental factors at the ecoregion scale did not help explain much of the variability in microbial community structure across ecoregions, which could be because factors at the local scale may be more important environmental drivers of stream microbial communities (Heino et al. 2014). In a floodplain river system, it has been proposed that environmental spatiotemporal heterogeneity determines ecological processes that shape bacterial metacommunities (Huber et al. 2020). Though there are differences in microbial community composition across ecoregions, the differences may be better explained by within stream variation of environmental variables.

Across the 13 ecoregions, both observed richness and Shannon alpha diversity were significantly different. Alpha diversity is a common first step in analysis of microbial communities across environments, but can be biased due to sampling design and the nature of environmental microbial samples. For example, if one environment has more microbial reads than another, then it is more likely to observe a greater number of different taxa in those samples, which can put greater emphasis on the library sizes over the biology (Willis 2019). Measures of diversity can also be inflated by organisms present in a sample that are inactive or dormant and primer bias in the sequencing of amplicons can underrepresent some microbial lineages (Shade 2017). Under-sampling communities is also possible and can skew samples where there are few highly uneven, rich communities (Adams et al. 2013). This has led to the idea that the true diversity of microbial communities is inestimable, but it is still possible to compare relative diversity measures (Hughes et al. 2001). It should be noted that in our study the total number of

samples across ecoregions was not similar, so the relative abundances of each ecoregion are based off of different number of total samples which can introduce error in undercounting members present in each stream's microbial community.

In addition to diversity measures, relative abundances of Classes of microbes can provide further insight into the microbial community composition across ecoregions. The Classes that were most abundant (Alphaproteobacteria, Cyanobacteria, Gammaproteobacteria) were similar across all ecoregions despite these ecosystems having different defining characteristics. Yet, the actual species composition within these classes and their functional abilities could be different across ecoregions, which could have consequences for N cycling across ecoregions. For example, members of the Classes Alphproteobacteria and Gammaprotetobacteria are known to be heterotrophic N_2-fixers in marine sediments (Hamersley et al. 2011, Gier et al. 2016). Even recently, Alphproteobacteria have been found to have the potential to be heterotrophic N_2-fixers in the oxygen minimum zones of oceans (Martinez-Perez et al. 2017). Cyanobacteria are commonly known N_2-fixers in aquatic environments (Scott and Marcarelli 2012). Members of Gammaproteobacteria have been shown to have the potential for nitrification and denitrification in oil sands lakes (Padilla et al. 2017). The potential for denitrification by members of Alphaproteobacteria has also been shown in oxic and suboxic waters (Wyman et al. 2013). Since these three Classes had relatively similar abundance across all sites, the potential for N_2 fixation and denitrification activity is possible across all sites. Yet, just because a species of a certain class has the potential to perform a process like N_2 fixation or denitrification it does not mean that the process is actively being carried out. This can be seen with the analysis of functional groups of our DNA

sequences based off of FAPROTAX in comparison to stream reach rates. The average % of taxa capable of N_2 fixation and denitrification did not correlate with stream reach average rates of each process. For denitrification, some streams had 0 taxa identified as possible denitrifies, but we measured rates of denitrification occurring there. Assessing the composition of microbial communities through DNA sequencing while helpful in determining who is there, cannot be directly related to what is actually occurring in an ecosystem.

In this study, we used 16S rRNA sequencing for all analysis and were able to investigate differences in microbial community structure. However, by sequencing DNA we have just hit on what taxa are present in each stream and ecoregion and not their actual functional capabilities, which would be better evaluated through RNA and transcriptome analysis that can more accurately assess what genes are being expressed. Without knowing what genes are expressed we cannot say for sure that an organism is performing a metabolic process, which is similar to what we found in Chapter 1 where relative gene abundance could not be directly linked to process occurrence. Our results showed that the dominant Classes of microorganisms are similar across ecoregions despite differences in environmental characteristics, although there were differences detected across the streams in overall assemblage composition. Overall, differences in environmental variables among ecoregions did not explain much of the variability in microbial community composition in streams across ecoregions. This could be because microbial communities are not that different at the class level among streams across ecoregions.

4.6 References

Adams, R. I., Amend, A. S., Taylor, J. W., & Bruns, T. D. (2013). A unique signal distorts the perception of species richness and composition in high-throughput sequencing surveys of microbial communities: a case study of fungi in indoor dust. *Microbial ecology*, *66*(4), 735-741.

American Public Health Administration (2005). Standard Methods for examination of water and wastewater. APHA, Washington D.C.

Bergey, E. A., & Getty, G. M. (2006). A review of methods for measuring the surface area of stream substrates. *Hydrobiologia*, *556*(1), 7-16.

Bernhardt, E. S., Heffernan, J. B., Grimm, N. B., Stanley, E. H., Harvey, J. W., Arroita, M., ... & Yackulic, C. B. (2018). The metabolic regimes of flowing waters. *Limnology and Oceanography*, *63*(S1), S99-S118.

Bru, D., Ramette, A., Saby, N. P. A., Dequiedt, S., Ranjard, L., Jolivet, C., ... & Philippot, L. (2011). Determinants of the distribution of nitrogen-cycling microbial communities at the landscape scale. *The ISME journal*, *5*(3), 532-542.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Capone, D.G. (1993). Determination of nitrogenase activity in aquatic samples using the acetylene reduction procedure. In: Kemp P.F., Sherr, B.F., Sherr, E.B., Cole J.J. (ed) Handbook of Methods in Aquatic Microbial Ecology, Lewis Publishers, Boca Raton.

Dodds, W. K., Burgin, A. J., Marcarelli, A. M., & Strauss, E. A. (2017). Nitrogen transformations. In *Methods in stream ecology* (pp. 173-196). Academic Press.

Eberhard, E. K., Marcarelli, A. M., & Baxter, C. V. (2018). Co-occurrence of in-stream nitrogen fixation and denitrification across a nitrogen gradient in a western US watershed. *Biogeochemistry*, *139*(2), 179-195.

Ferguson, S. J. (1994). Denitrification and its control. *Antonie van Leeuwenhoek*, *66*(1), 89-110.

Fierer, N., Morse, J. L., Berthrong, S. T., Bernhardt, E. S., & Jackson, R. B. (2007). Environmental controls on the landscape-scale biogeography of stream bacterial communities. *Ecology*, *88*(9), 2162-2173.

Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., ... & Caporaso, J. G. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences*, *109*(52), 21390-21395.

Findlay, R. H., Yeates, C., Hullar, M. A., Stahl, D. A., & Kaplan, L. A. (2008). Biome-level biogeography of streambed microbiota. *Applied and Environmental Microbiology*, *74*(10), 3014-3021.

Findlay, S. (2010). Stream microbial ecology. *Journal of the North American Benthological Society*, *29*(1), 170-181.

Fulweiler, R. W., Heiss, E. M., Rogener, M. K., Newell, S. E., LeCleir, G. R., Kortebein, S. M., & Wilhelm, S. W. (2015). Examining the impact of acetylene on N-fixation and the active sediment microbial community. *Frontiers in Microbiology*, *6*, 418.

Gettel, G. M., Giblin, A. E., & Howarth, R. W. (2007). The effects of grazing by the

snail, Lymnaeaelodes, on benthic N2 fixation and primary production in oligotrophic, arctic lakes. *Limnology and oceanography*, *52*(6), 2398-2409.

Gier, J., Sommer, S., Löscher, C. R., Dale, A. W., Schmitz, R. A., & Treude, T. (2016). Nitrogen fixation in sediments along a depth transect through the Peruvian oxygen minimum zone. *Biogeosciences*, *13*(14), 4065-4080.

Goodman, K. J., Parker, S. M., Edmonds, J. W., & Zeglin, L. H. (2015). Expanding the scale of aquatic sciences: the role of the National Ecological Observatory Network (NEON). *Freshwater Science*, *34*(1), 377-385.

Groffman, P. M., Altabet, M. A., Böhlke, J. K., Butterbach-Bahl, K., David, M. B., Firestone, M. K., ... & Voytek, M. A. (2006). Methods for measuring denitrification: diverse approaches to a difficult problem. *Ecological Applications*, *16*(6), 2091-2122.

Hamersley, M. R., Turk, K. A., Leinweber, A., Gruber, N., Zehr, J. P., Gunderson, T., & Capone, D. G. (2011). Nitrogen fixation within the water column associated with two hypoxic basins in the Southern California Bight. *Aquatic Microbial Ecology*, *63*(2), 193-205.

Hawkins, B. A., Field, R., Cornell, H. V., Currie, D. J., Guégan, J. F., Kaufman, D. M., ... & Turner, J. R. (2003). Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, *84*(12), 3105-3117.

Heino, J., Tolkkinen, M., Pirttilä, A. M., Aisala, H., & Mykrä, H. (2014). Microbial diversity and community–environment relationships in boreal streams. *Journal of Biogeography*, *41*(12), 2234-2244.

Hendershot, J. N., Read, Q. D., Henning, J. A., Sanders, N. J., & Classen, A. T. (2017).

Consistently inconsistent drivers of microbial diversity and abundance at macroecological scales.

Holmes, R. M., Aminot, A., Kérouel, R., Hooker, B. A., & Peterson, B. J. (1999). A simple and precise method for measuring ammonium in marine and freshwater ecosystems. *Canadian Journal of Fisheries and Aquatic Sciences*, *56*(10), 1801-1808.

Huber, P., Metz, S., Unrein, F., Mayora, G., Sarmento, H., & Devercelli, M. (2020). Environmental heterogeneity determines the ecological processes that govern bacterial metacommunity assembly in a floodplain river system. *The ISME journal*, *14*(12), 2951-2966.

Hughes, J. B., Hellmann, J. J., Ricketts, T. H., & Bohannan, B. J. (2001). Counting the uncountable: statistical approaches to estimating microbial diversity. *Applied and environmental microbiology*, *67*(10), 4399-4406.

Keller, M., Schimel, D. S., Hargrove, W. W., & Hoffman, F. M. (2008). A continental strategy for the National Ecological Observatory Network. *The Ecological Society of America: 282-284*.

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and environmental microbiology*, *79*(17), 5112-5120.

Kreft, H., & Jetz, W. (2007). Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences*, *104*(14), 5925-5930.

Larson, C. A., Mirza, B., Rodrigues, J. L. M., & Passy, S. I. (2018). Iron limitation

effects on nitrogen-fixing organisms with possible implications for cyanobacterial blooms. *FEMS microbiology ecology*, *94*(5), fiy046.

Larouche, J. R., Bowden, W. B., Giordano, R., Flinn, M. B., & Crump, B. C. (2012). Microbial biogeography of arctic streams: exploring influences of lithology and habitat. *Frontiers in Microbiology*, *3*, 309.

Legendre, P., & Anderson, M. J. (1999). Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological monographs*, *69*(1), 1-24.

Lemmon, P. E. (1956). A spherical densiometer for estimating forest overstory density. *Forest science*, *2*(4), 314-320.

Louca, S., Parfrey, L.W., Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Science* 353:1272-1277

Martínez-Pérez, C., Mohr, W., Schwedt, A., Dürschlag, J., Callbeck, C. M., Schunck, H., ... & Kuypers, M. M. (2018). Metabolic versatility of a novel N2-fixing Alphaproteobacterium isolated from a marine oxygen minimum zone. *Environmental Microbiology*, *20*(2), 755-768.

Martiny, J. B. H., Bohannan, B. J., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., ... & Staley, J. T. (2006). Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, *4*(2), 102-112.

McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*, *8*(4), e61217.

Mittelbach, G. G., & McGill, B. J. (2019). *Community ecology*. Oxford University Press.

Monteiro, F. M., Dutkiewicz, S., & Follows, M. J. (2011). Biogeographical controls on the marine nitrogen fixers. *Global Biogeochemical Cycles*, *25*(2).

Nelson, M. B., Martiny, A. C., & Martiny, J. B. (2016). Global biogeography of microbial nitrogen-cycling traits in soil. *Proceedings of the National Academy of Sciences*, *113*(29), 8033-8040.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., ... & Wagner, H. (2020). Community ecology package. *R package version 2.5-7*.

Padilla, C. C., Bertagnolli, A. D., Bristow, L. A., Sarode, N., Glass, J. B., Thamdrup, B., & Stewart, F. J.(2017). Metagenomic binning recovers a transcriptionally active gammaproteobacterium linking methanotrophy to partial denitrification in an anoxic oxygen minimum zone. *Frontiers in Marine Science*, *4*, 23.

Pagaling, E., Wang, H., Venables, M., Wallace, A., Grant, W. D., Cowan, D. A., ... & Heaphy, S. (2009). Microbial biogeography of six salt lakes in Inner Mongolia, China, and a salt lake in Argentina. *Applied and environmental microbiology*, *75*(18), 5750-5760.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research*, *41*(D1), D590-D596.

Schimel, J. P., & Gulledge, J. A. Y. (1998). Microbial community structure and global trace gases. *Global change biology*, *4*(7), 745-758.

Scott, J. T., & Marcarelli, A. M. (2012). Cyanobacteria in freshwater benthic environments. In *Ecology of cyanobacteria II* (pp. 271-289). Springer, Dordrecht.

Shade, A. (2017). Diversity is the question, not the answer. *The ISME journal*, *11*(1), 1-6.

Taylor, B. W., Keep, C. F., Hall Jr, R. O., Koch, B. J., Tronstad, L. M., Flecker, A. S., & Ulseth, A. J. (2007). Improving the fluorometric ammonium method: matrix effects, background fluorescence, and standard additions. *Journal of the North American Benthological Society*, *26*(2), 167-177.

Utz, R. M., Fitzgerald, M. R., Goodman, K. J., Parker, S. M., Powell, H., & Roehm, C. L. (2013). The National Ecological Observatory Network: an observatory poised to expand spatiotemporal scales of inquiry in aquatic and fisheries science. *Fisheries*, *38*(1), 26-35.

Wellington, E. M., Berry, A., & Krsek, M. (2003). Resolving functional diversity in relation to microbial community structure in soil: exploiting genomics and stable isotope probing. *Current opinion in microbiology*, *6*(3), 295-301.

Willis, A. D. (2019). Rarefaction, alpha diversity, and statistics. *Frontiers in microbiology*, *10*, 2407.

Wyman, M., Hodgson, S., & Bird, C. (2013). Denitrifying alphaproteobacteria from the Arabian Sea that express nosZ, the gene encoding nitrous oxide reductase, in oxic and suboxic waters. *Applied and environmental microbiology*, *79*(8), 2670-2681.

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., ... & Glöckner, F. O. (2014). The SILVA and "all-species living tree project (LTP)" taxonomic frameworks. *Nucleic acids research*, *42*(D1), D643-D648.

5 Conclusion

Throughout this dissertation I found that N_2 fixation and denitrification co-occur in freshwater ecosystems at multiple spatial scales including at the patch scale within stream reaches (Chapter 2), cross-ecosystem scale at wetland-stream-lake interfaces (Chapter 3), and in streams across ecoregions of the U.S. (Chapter 4, Fig. 5.1). This finding further supports the findings of my previous research, where I found that N_2 fixation and denitrification co-occurred in stream reaches across a gradient of N concentrations (Eberhard et al. 2018). These findings also support studies in coastal marine and ocean environments, where N_2 fixation and denitrification have been found to co-occur within sediments and the water column (Fulweiler and Heiss 2014, Deutsch et al. 2007).

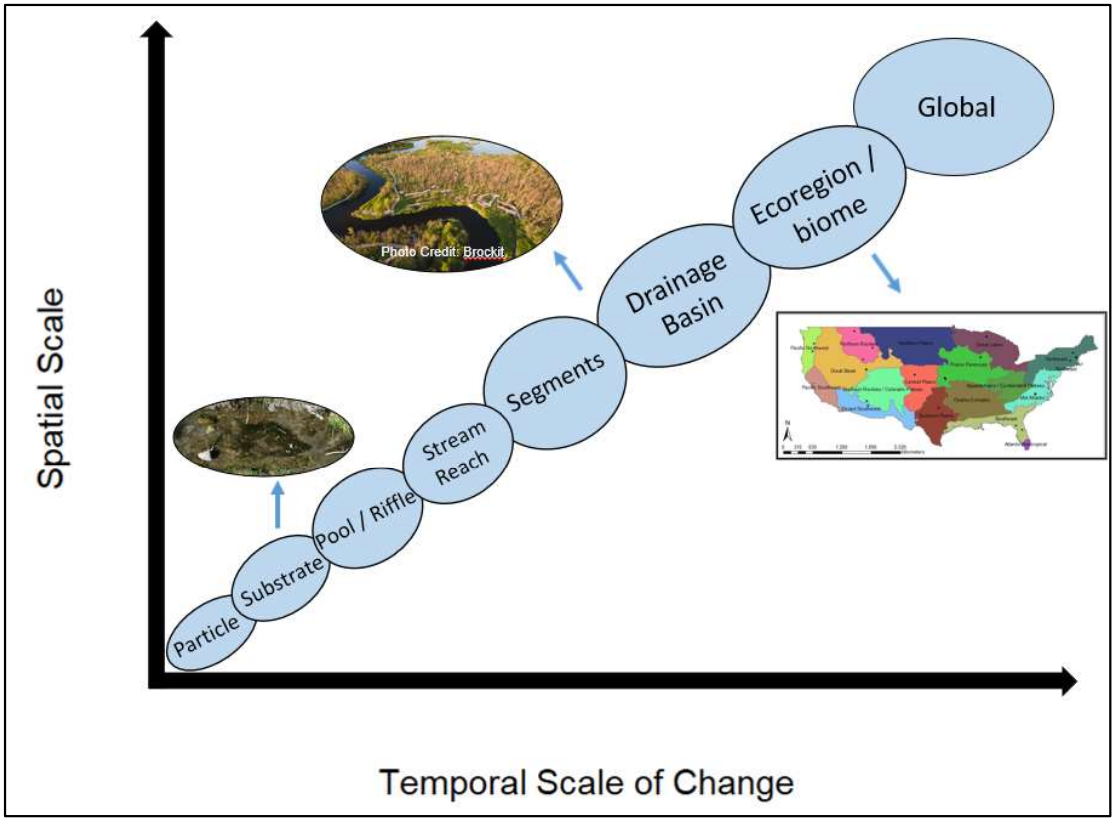

Figure 5.1. Graphic depiction of heterogeneity of aquatic ecosystems across spatial and temporal scales. We specifically looked at the substrate scale (Chapter 2), cross-ecosystem scale (Chapter 3), and across ecoregions (Chapter 4).

The co-existence of N_2 fixation and denitrification at multiple spatial scales was facilitated through variability in substrate. In Chapter 2, at the patch scale within stream reaches we observed high rates of N_2 fixation occurring exclusively on rock substrate and high rates of denitrification occurring exclusively on sediment substrate. Thus, the presence of both substrates facilitated the co-occurrence of both processes in streams. In Chapter 3, high rates of N_2 fixation occurred on both macrophyte and sediment substrate, while high rates of denitrification occurred primarily on sediment substrate. Within multiple transect points, high rates of N_2 fixation and denitrification simultaneously occurred within sediment cores collected inches from one another. This is similar to what others have found, with N_2 fixation and denitrification co-occurring within centimeters of each other in sediment cores from coastal sediments (Newell et al. 2016). While we were able to capture some of this small-scale variability in process rates, these findings from our own study and others suggest that the microorganisms mediating these processes experience variability in environment on a finer scale that we are not able to capture through our commonly used benthic sampling techniques.

Although N_2 fixation and denitrification cooccurs, the magnitude of rates differed and were not always comparable. Within stream reaches (Chapter 2), denitrification rates tended to be 100x higher than N_2 fixation rates, which was also observed at some wetland transect points (Chapter 3). However, at other wetland transect points the magnitude of N_2 fixation and denitrification rates were equal. This could have consequences for understand the N balance in wetland – stream - lake interfaces in the Great Lakes region, where N_2 fixation rates have previously been thought of as negligible. If we base our understanding of N balance off of select point measurements of process rates within these

complex ecosystems then we could miss out on important areas where the two rates are contributing equally to the overall N balance.

Another important finding across all of the research I conducted is that as spatial scale increased, less of the variation in rates of N_2 fixation and denitrification and microbial community composition could be explained by environmental characteristics. At the patch scale, 63% of the variation in N_2 fixation rates and 84% of the variation in denitrification rates were explained by environmental variables, while only 13-26% of variability in microbial gene abundances were explained by environmental variables (Chapter 2). This could be because at the patch scale we were measuring at the scale the organisms performing these processes would be affected on, which would alter process rates. Relative gene abundances may not be fully explained by variance in environmental variables even at this scale because if a gene is present does not mean it is actively being expressed. At the cross-ecosystem scale only 20-33% of the variation in process rates was explained by environmental variables (Chapter 3). At the ecoregion scale only 27% of the variability in microbial community composition could be explained by differences in environmental variables, while 17% of the variation in community composition could be explained by ecoregion (Chapter 4). This decline in explanatory power as spatial scale increased could be because cross-ecosystem and across ecoregions variation in environmental characteristics can only apply a certain amount of selective pressure on which biogeochemical process could occur and which microorganisms are present. Within a system, micro-scale variation in environmental variables could be what is providing the primary control on process rates and microbial assemblages. Temporal variability could account for some of the variance in process rates not explained by

spatial variability, as our measurements are just a snapshot of time. There could be times when conditions are right for a process to occur, or where shifts in microbial community structure occur that could allow for moments in time of high process rates (McClain et al. 2003, Matulich et al. 2015, Bernhardt et al. 2017, Gautam et al. 2021).

Across our studies, N alone was rarely an important predictor for rates of N_2 fixation and denitrification. This is in contradiction with what has been regularly assumed in the past, which is that N concentrations are the best indicator of the occurrence of these processes (Marcarelli et al. 2008, Marcarelli et al. 2022). Instead, environmental variables like P, dissolved oxygen, and organic matter concentrations were frequently identified as important predictors of the occurrence of these processes. Therefore, we cannot continue to make assumptions about the occurrence of these process in aquatic ecosystems based off N concentrations alone. By using N concentrations as the main indicator of process occurrence we could be biasing our understanding and study of both of these processes, and missing much of the spatial variability in rates of N_2 fixation and denitrification within and across aquatic ecosystems. This would severely affect our understanding of overall N dynamics in these ecosystems and underestimate the complexity of N cycling.

Overall, N_2 fixation and denitrification must be studied together to better understand the complexity of N cycling in aquatic ecosystems. Both of these processes are occurring at multiple spatial scales, and in some instances N_2 fixation is occurring at relatively high, ecological significant rates in close proximity to, simultaneously with N removal via denitrification. To better understand the occurrence of these processes we need to better address how heterogeneity, both spatial and temporal, affect process rates of N_2 fixation and denitrification.

5.1 References

Bernhardt, E. S., Blaszczak, J. R., Ficken, C. D., Fork, M. L., Kaiser, K. E., & Seybold, E. C. (2017). Control points in ecosystems: moving beyond the hot spot hot moment concept. *Ecosystems*, *20*(4), 665-682.

Deutsch, C., Sarmiento, J. L., Sigman, D. M., Gruber, N., & Dunne, J. P. (2007). Spatial coupling of nitrogen inputs and losses in the ocean. *Nature*, *445*(7124), 163-167.

Eberhard, E. K., Marcarelli, A. M., & Baxter, C. V. (2018). Co-occurrence of in-stream nitrogen fixation and denitrification across a nitrogen gradient in a western US watershed. *Biogeochemistry*, *139*(2), 179-195.

Fulweiler, R. W., & Heiss, E. M. (2014). (Nearly) A Decade of Directly Measured Sediment N_2 Fluxes: What Can Narragansett Bay Tell Us About the Global Ocean Nitrogen Budget?. *Oceanography*, *27*(1), 184-195.

Gautam, A., Lear, G., & Lewis, G. D. (2021). Analysis of spatial and temporal variations in bacterial community dynamics within stream biofilms. *New Zealand Journal of Marine and Freshwater Research*, *55*(4), 505-523.

Marcarelli, A. M., Baker, M. A., & Wurtsbaugh, W. A. (2008). Is in-stream N2 fixation an important N source for benthic communities and stream ecosystems?. *Journal of the North American Benthological Society*, *27*(1), 186-211.

Marcarelli, A. M., Fulweiler, R. W., & Scott, J. T. (2022). Nitrogen fixation: A poorly understood process along the freshwater-marine continuum. *Limnology and Oceanography Letters*, *7*(1), 1-10.

Matulich, K. L., Weihe, C., Allison, S. D., Amend, A. S., Berlemont, R., Goulden, M. L.,

... & Martiny, J. B. (2015). Temporal variation overshadows the response of leaf litter microbial communities to simulated global change. *The ISME journal*, *9*(11), 2477-2489.

McClain, M. E., Boyer, E. W., Dent, C. L., Gergel, S. E., Grimm, N. B., Groffman, P. M., ... & Pinay, G. (2003). Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems*, 301-312.

Newell, S. E., McCarthy, M. J., Gardner, W. S., & Fulweiler, R. W. (2016). Sediment nitrogen fixation: a call for re-evaluating coastal N budgets. *Estuaries and Coasts*, *39*(6), 1626-1638.

A Appendix

A.1 Chapter 2

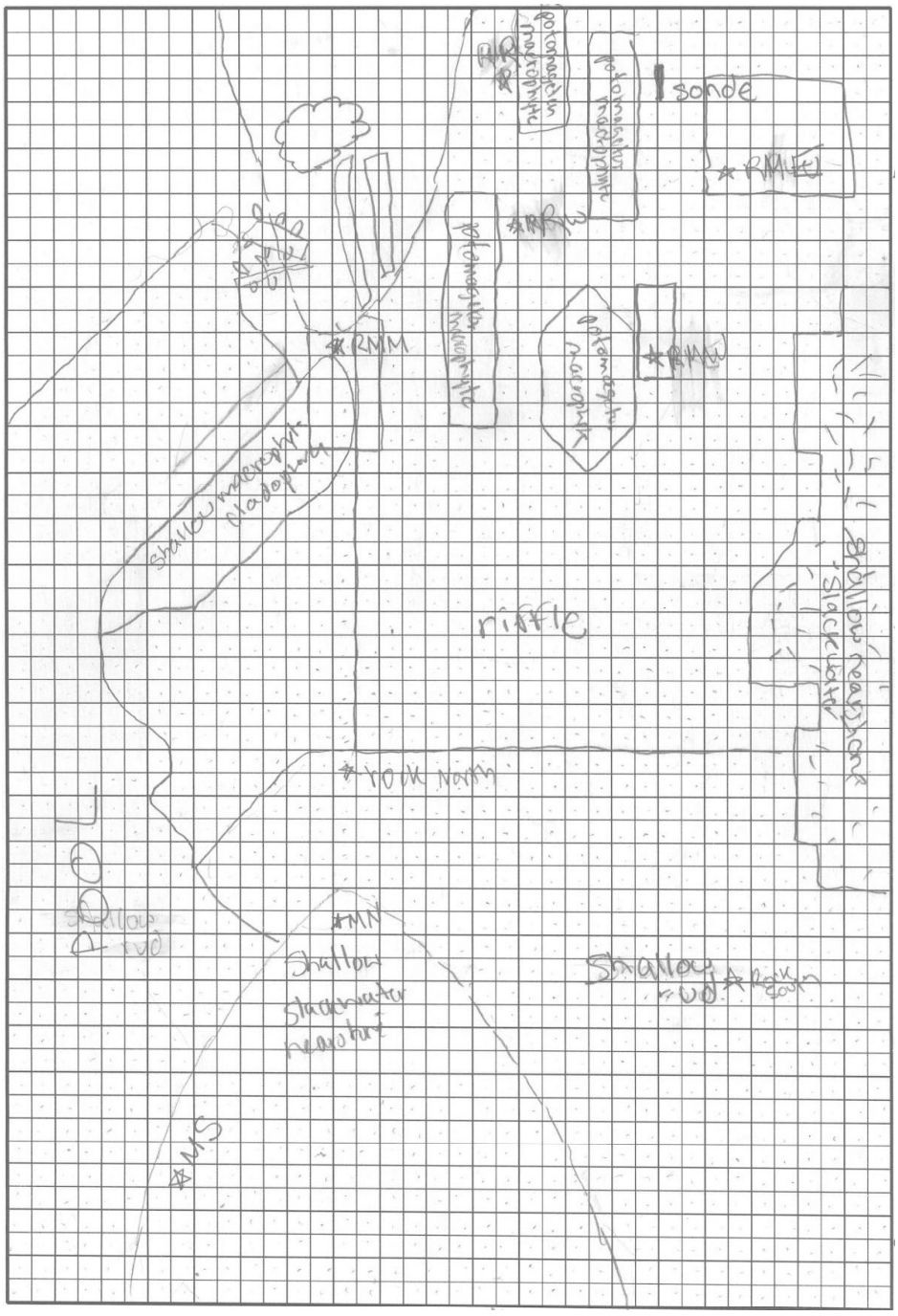

Figure 1A (part 1 of 3): Rough habitat map for a reach of Diggie Creek in Idaho. Each square has 1 x 1 m dimensions. Maps begin upstream to downstream.

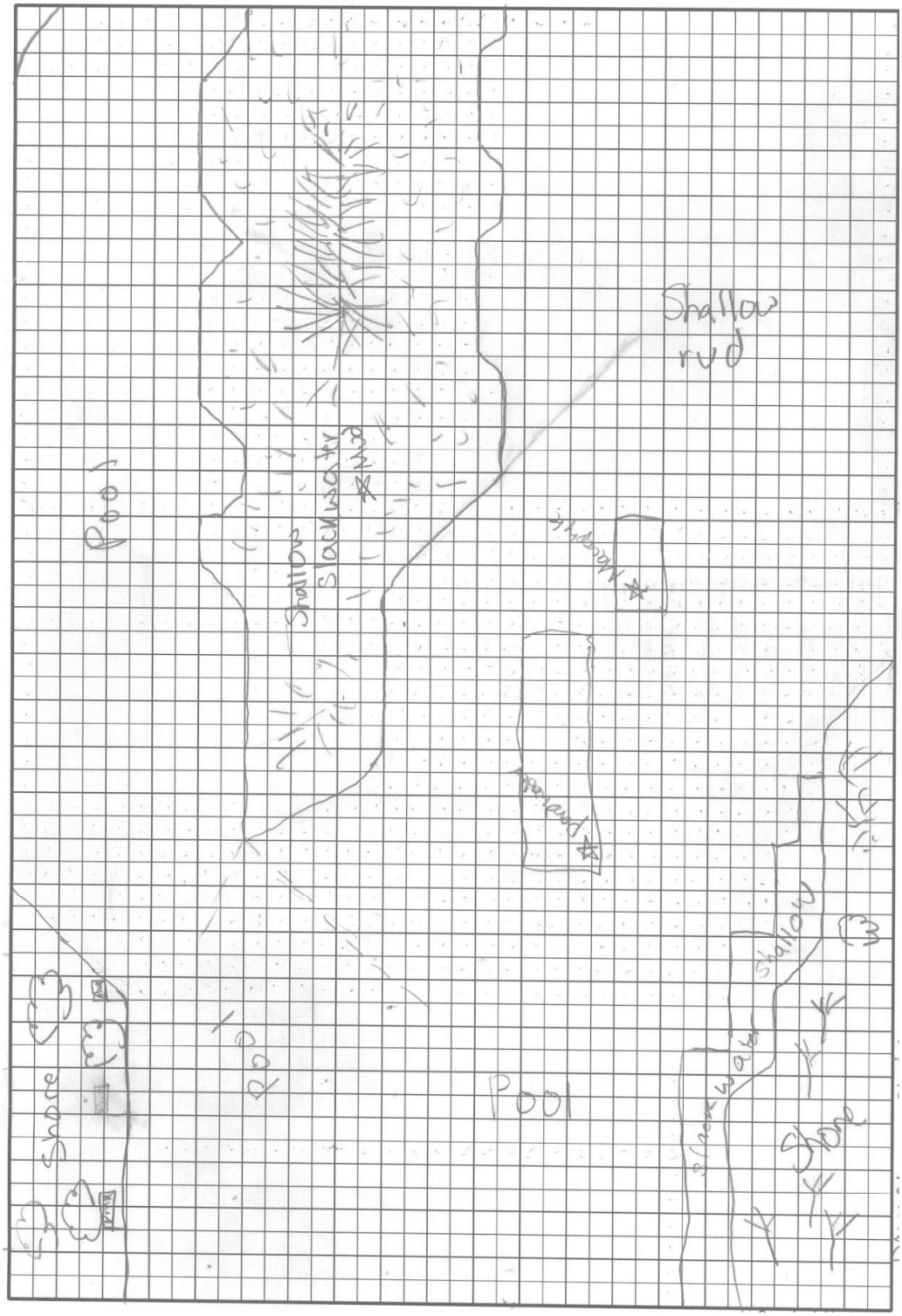

Figure 1A (part 2 of 3): Rough habitat map for a reach of Diggie Creek in Idaho. Each square has 1 x 1 m dimensions. Maps begin upstream to downstream.

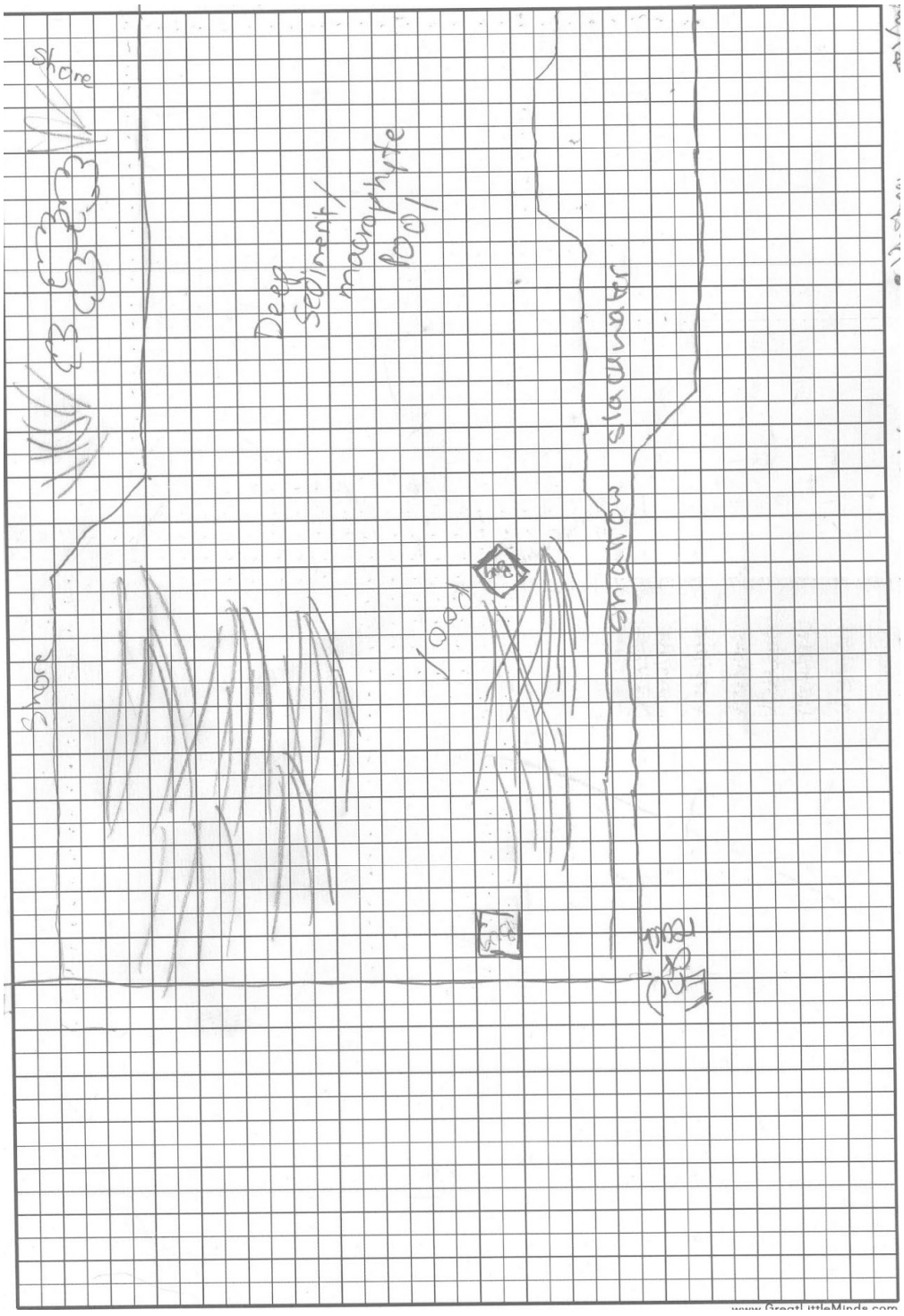

Figure 1A (part 3 of 3): Rough habitat map for a reach of Diggie Creek in Idaho. Each square has 1 x 1 m dimensions. Maps begin upstream to downstream.

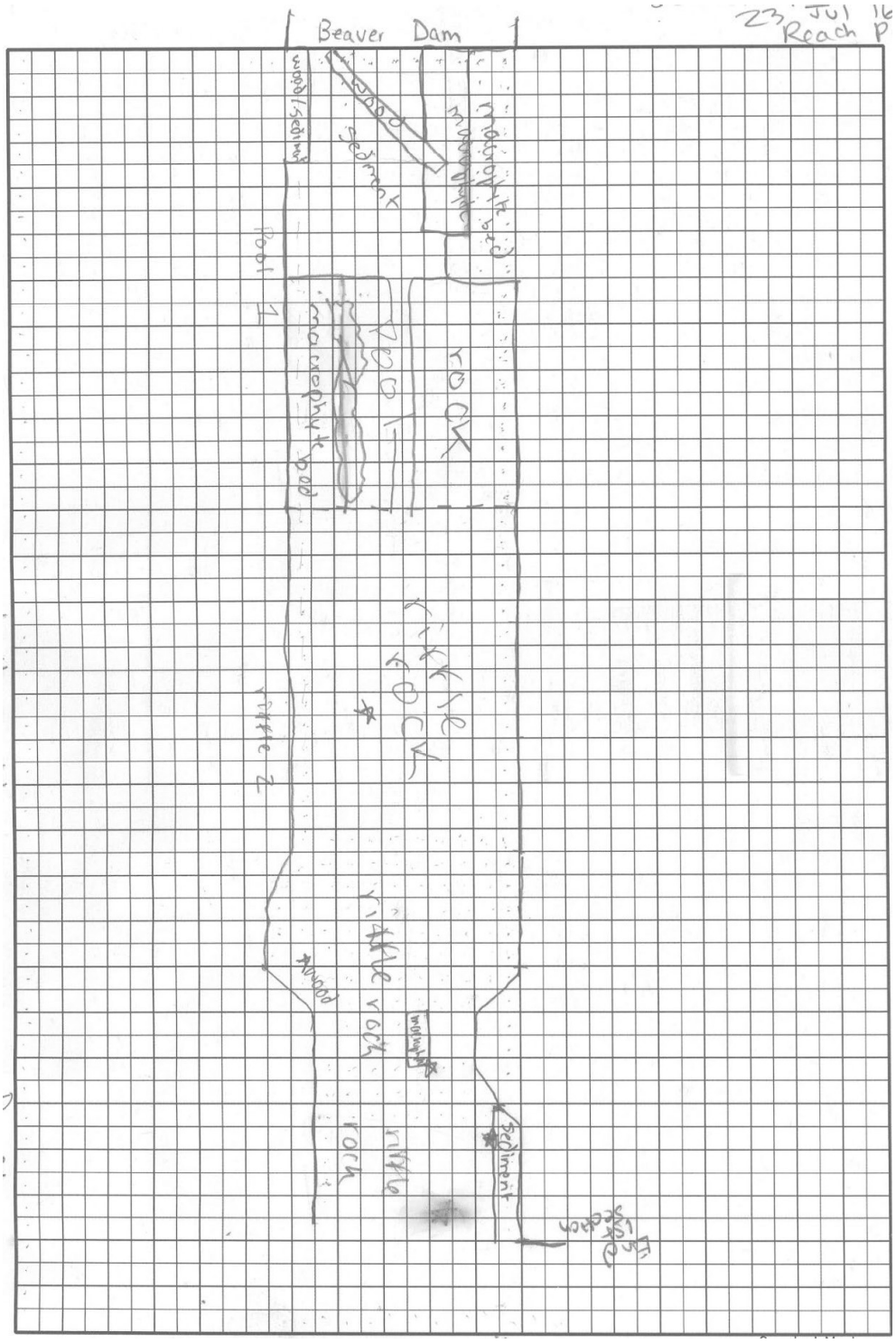

Figure 2A (part 1 of 3): Rough habitat map for a reach of South Fork in Idaho. Each square has dimensions of 1 m horizontal x 0.5 m vertical. Maps begin upstream to downstream.

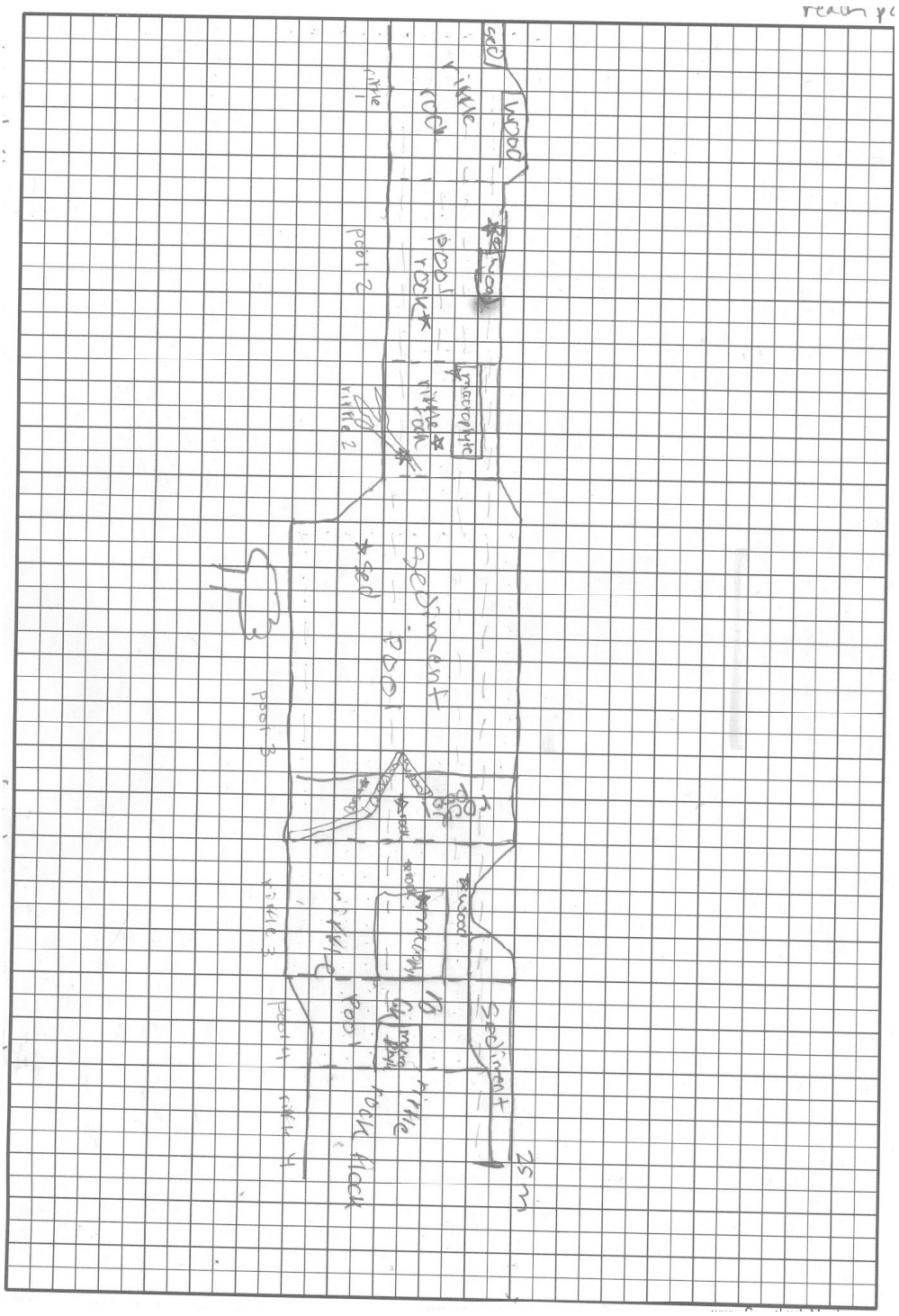

Figure 2A (part 2 of 3): Rough habitat map for a reach of South Fork in Idaho. Each square has dimensions of 1 m horizontal x 0.5 m vertical. Maps begin upstream to downstream.

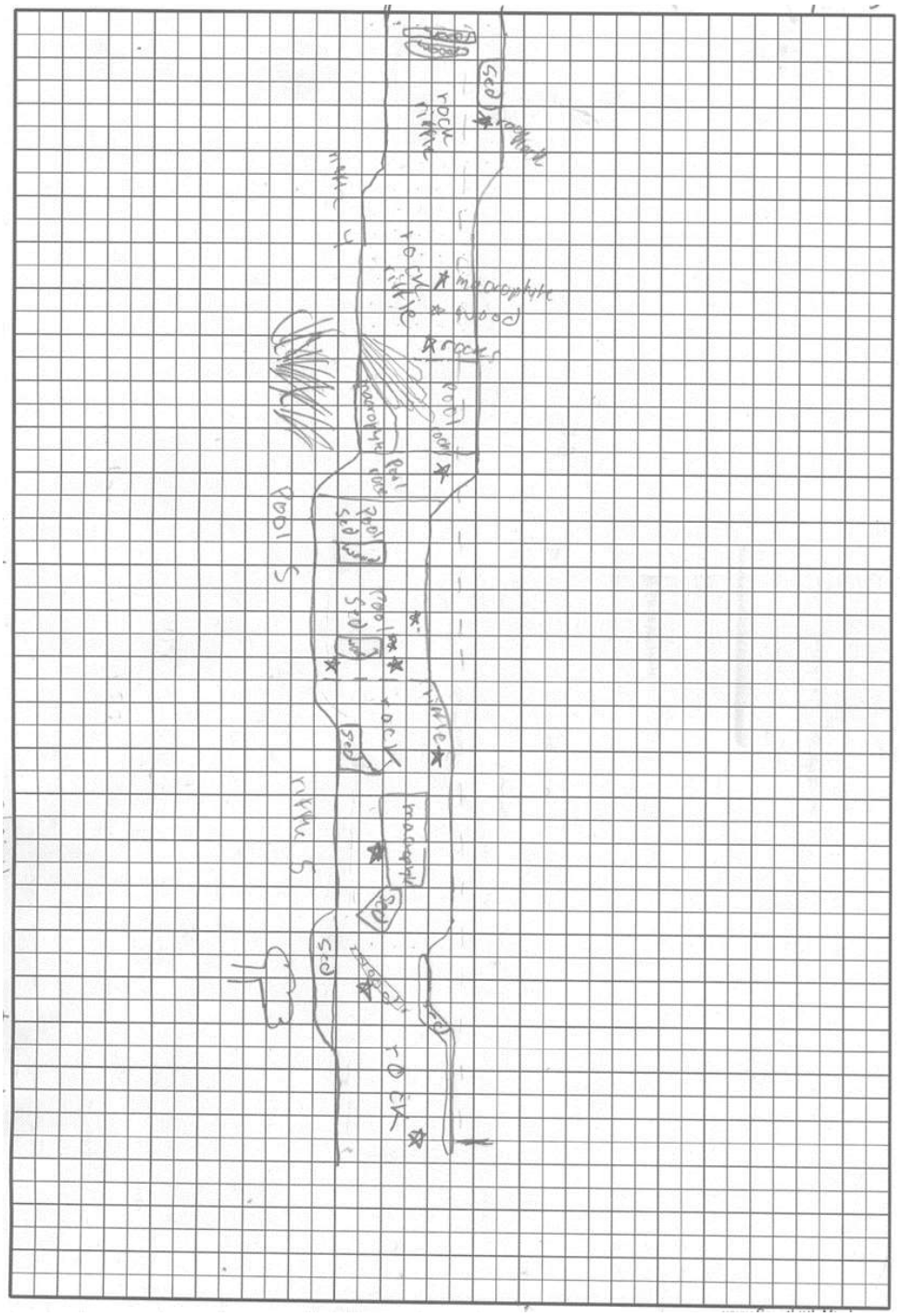

Figure 2A (part 3 of 3): Rough habitat map for a reach of South Fork in Idaho. Each square has dimensions of 1 m horizontal x 0.5 m vertical. Maps begin upstream to downstream.

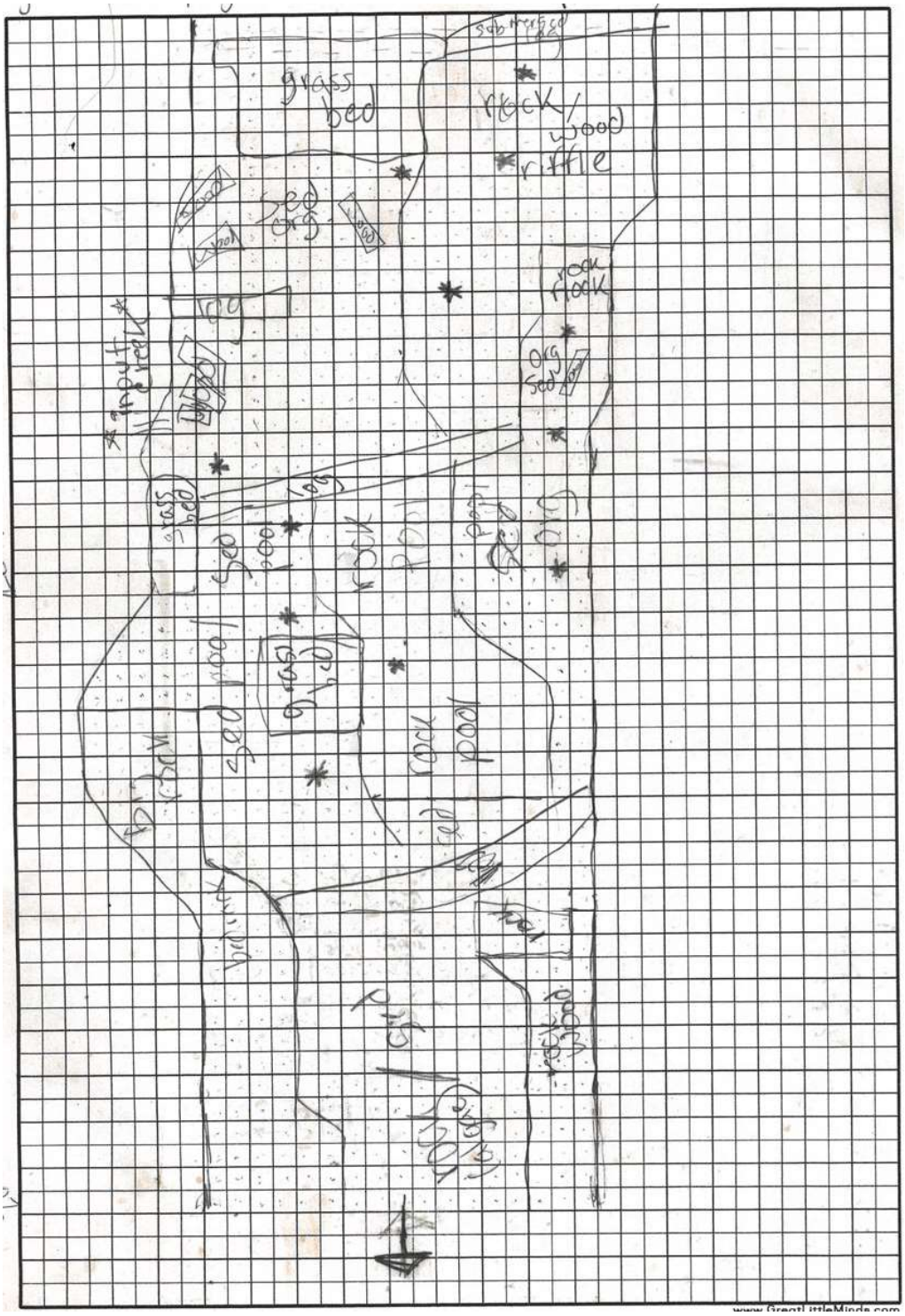

Figure 3A (part 1 of 2): Rough habitat map for a reach of the Pilgrim in Michigan. Each square has dimensions of 1 m horizontal x 1 m vertical. Maps begin upstream to downstream.

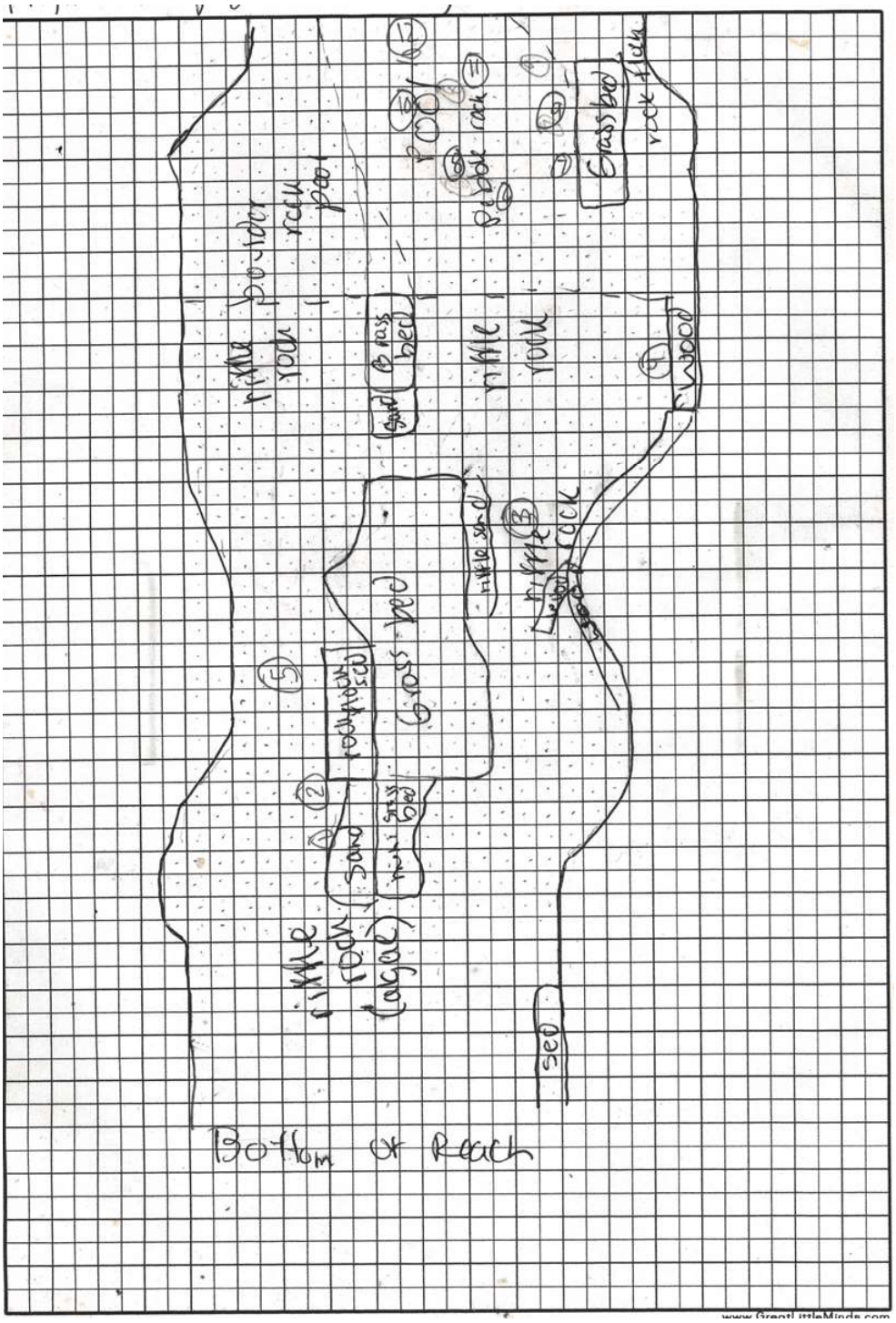

Figure 3A (part 2 of 2): Rough habitat map for a reach of the Pilgrim in Michigan. Each square has dimensions of 1 m horizontal x 1 m vertical. Maps begin upstream to downstream.

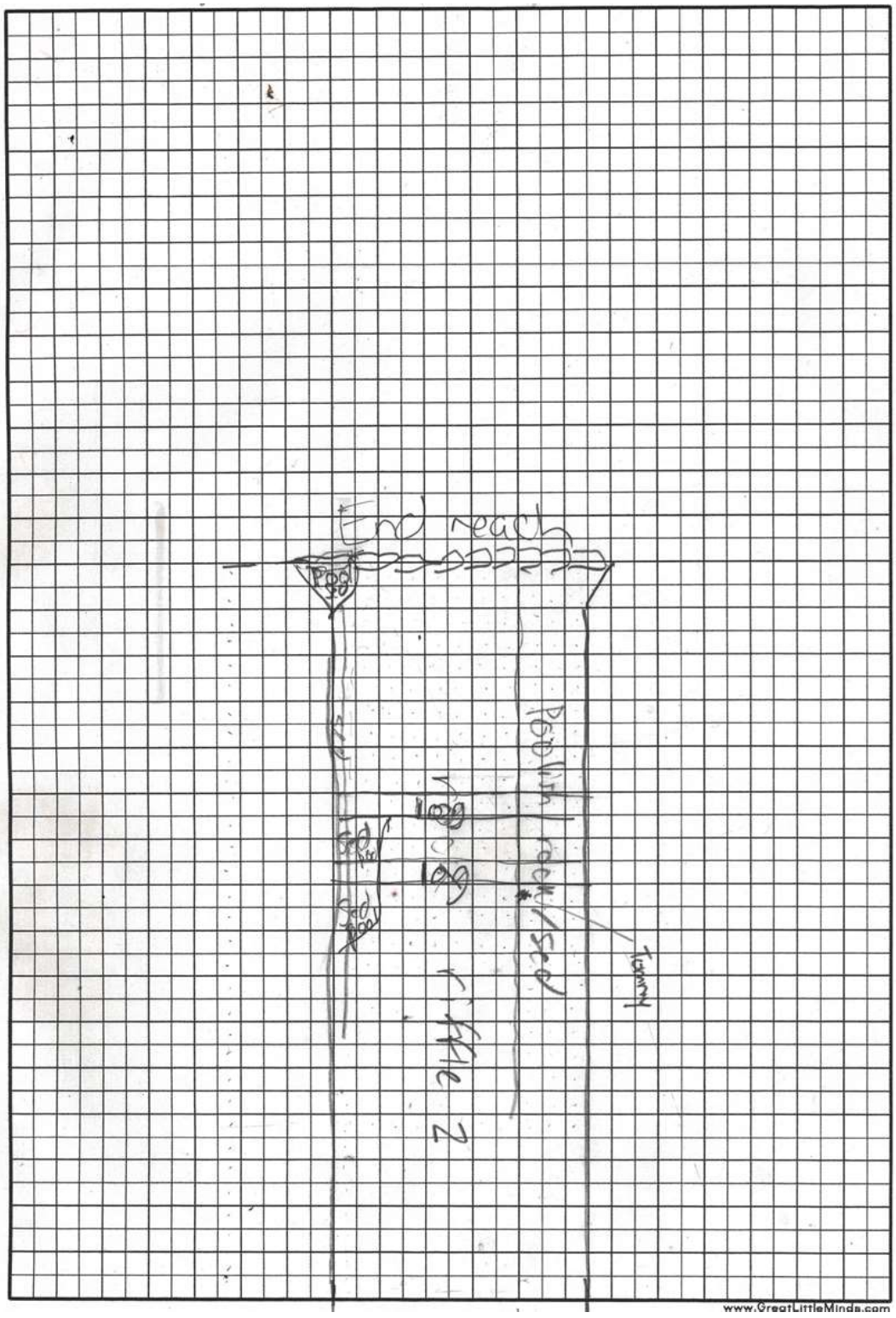

Figure 4A (part 1 of 2): Rough habitat map for a reach of McGunn in Michigan. Each square has dimensions of 1 m horizontal x 1 m vertical. Maps begin upstream to downstream.

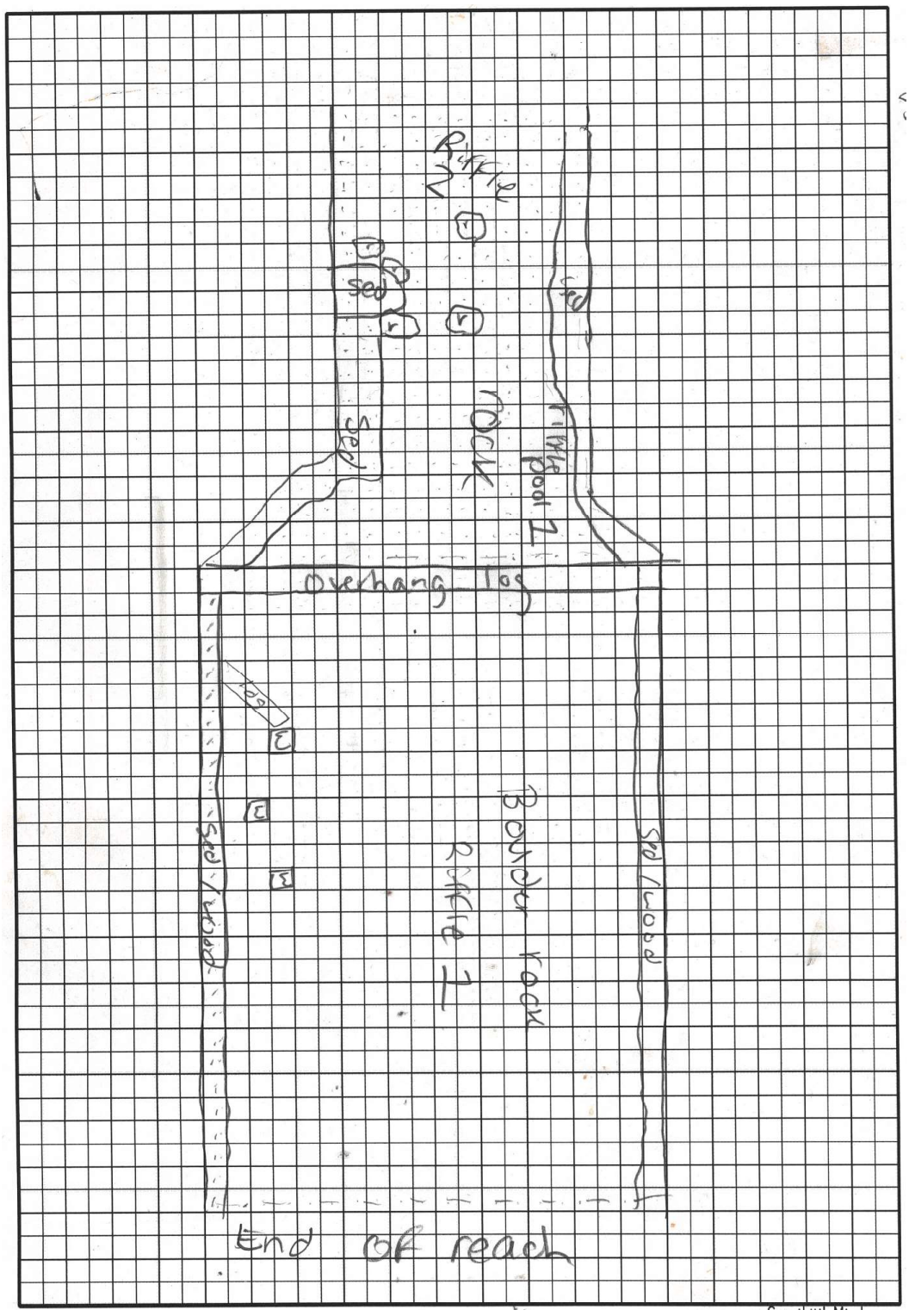

Figure 4A (part 2 of 2): Rough habitat map for a reach of McGunn in Michigan. Each square has dimensions of 1 m horizontal x 1 m vertical. Maps begin upstream to downstream.

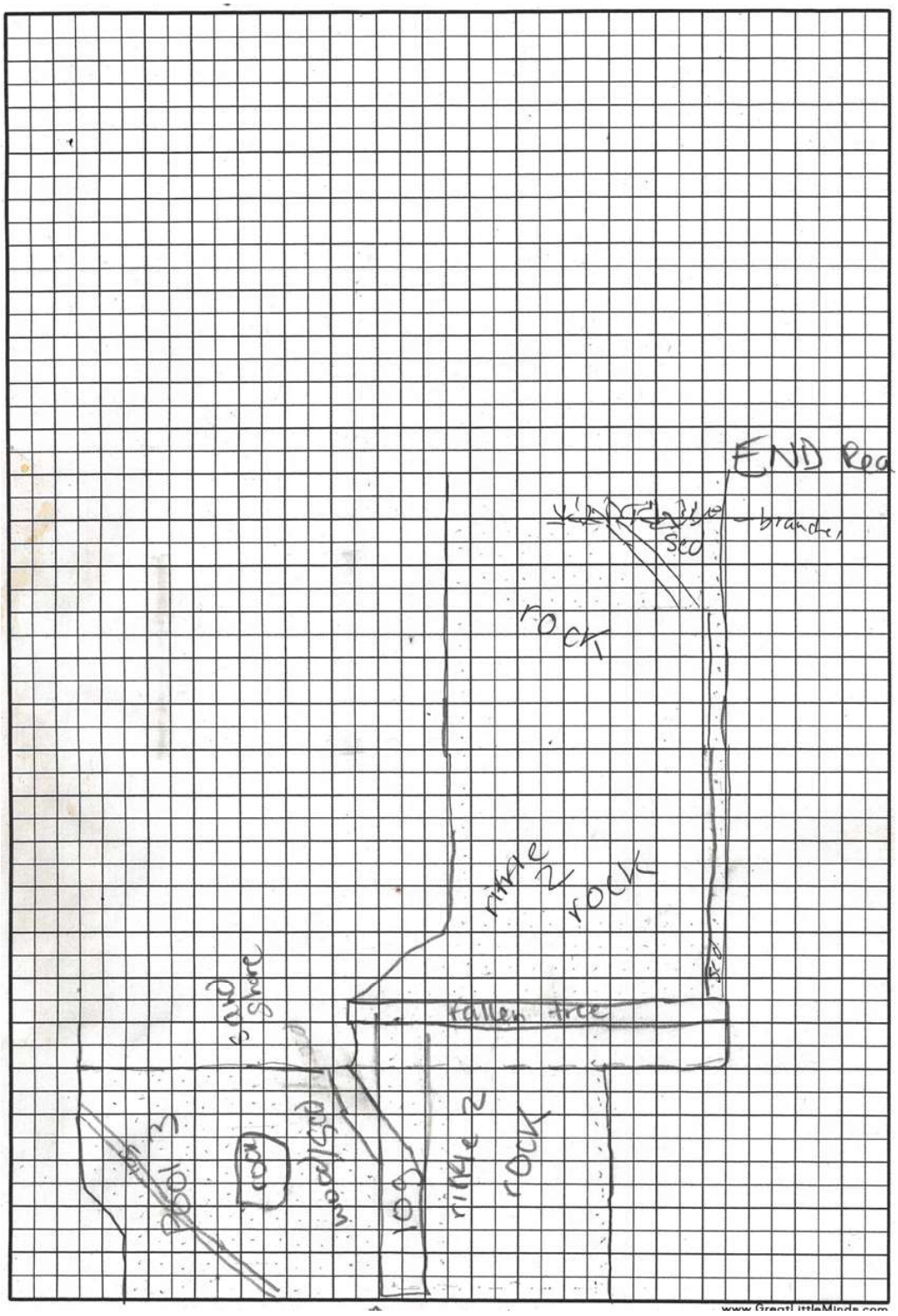

Figure 5A (part 1 of 2): Rough habitat map for a reach of Gratiot in Michigan. Each square has dimensions of 2 m horizontal x 1 m vertical. Maps begin upstream to downstream.

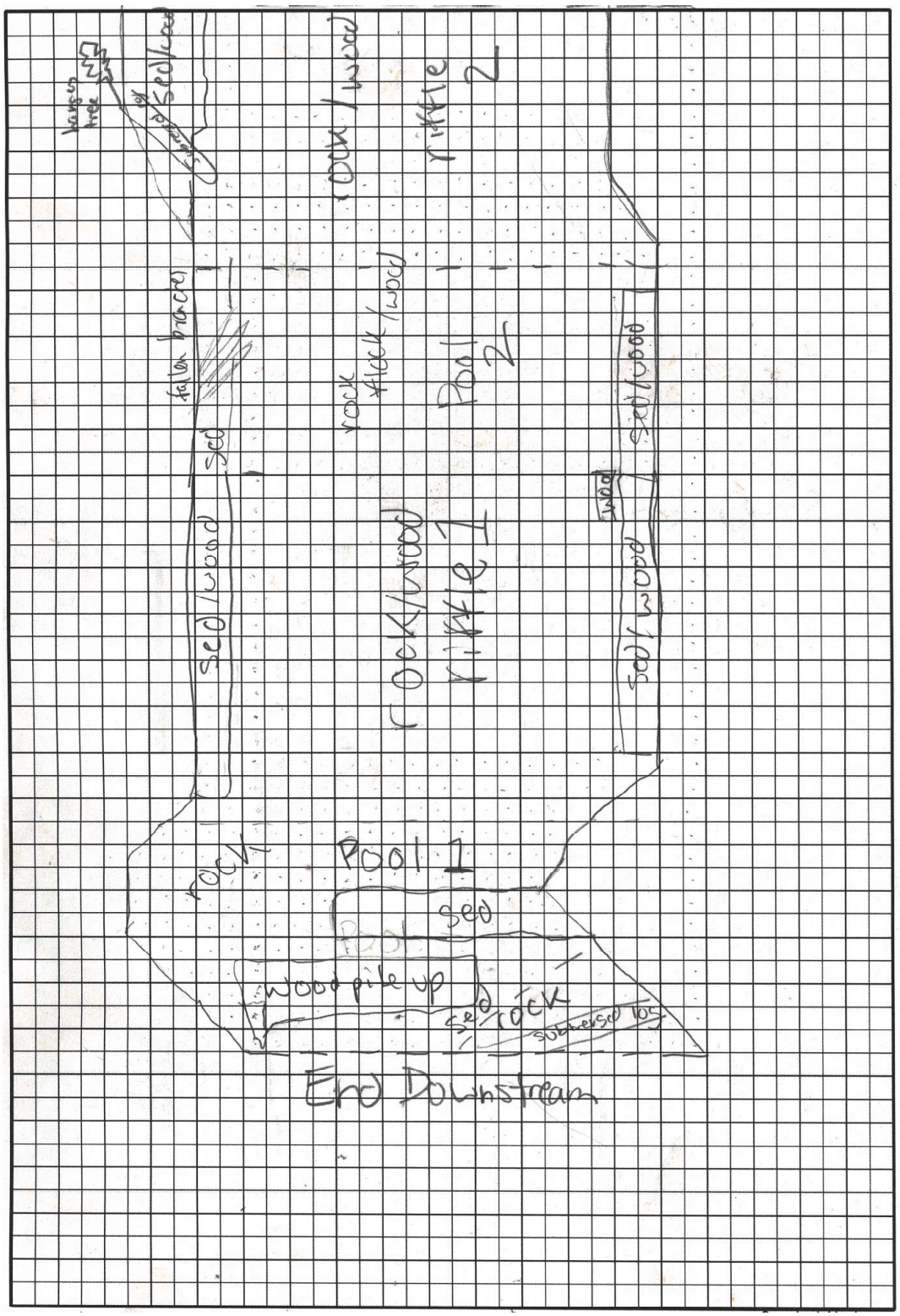

Figure 5A (part 2 of 2): Rough habitat map for a reach of Gratiot in Michigan. Each square has dimensions of 2 m horizontal x 1 m vertical. Maps begin upstream to downstream.

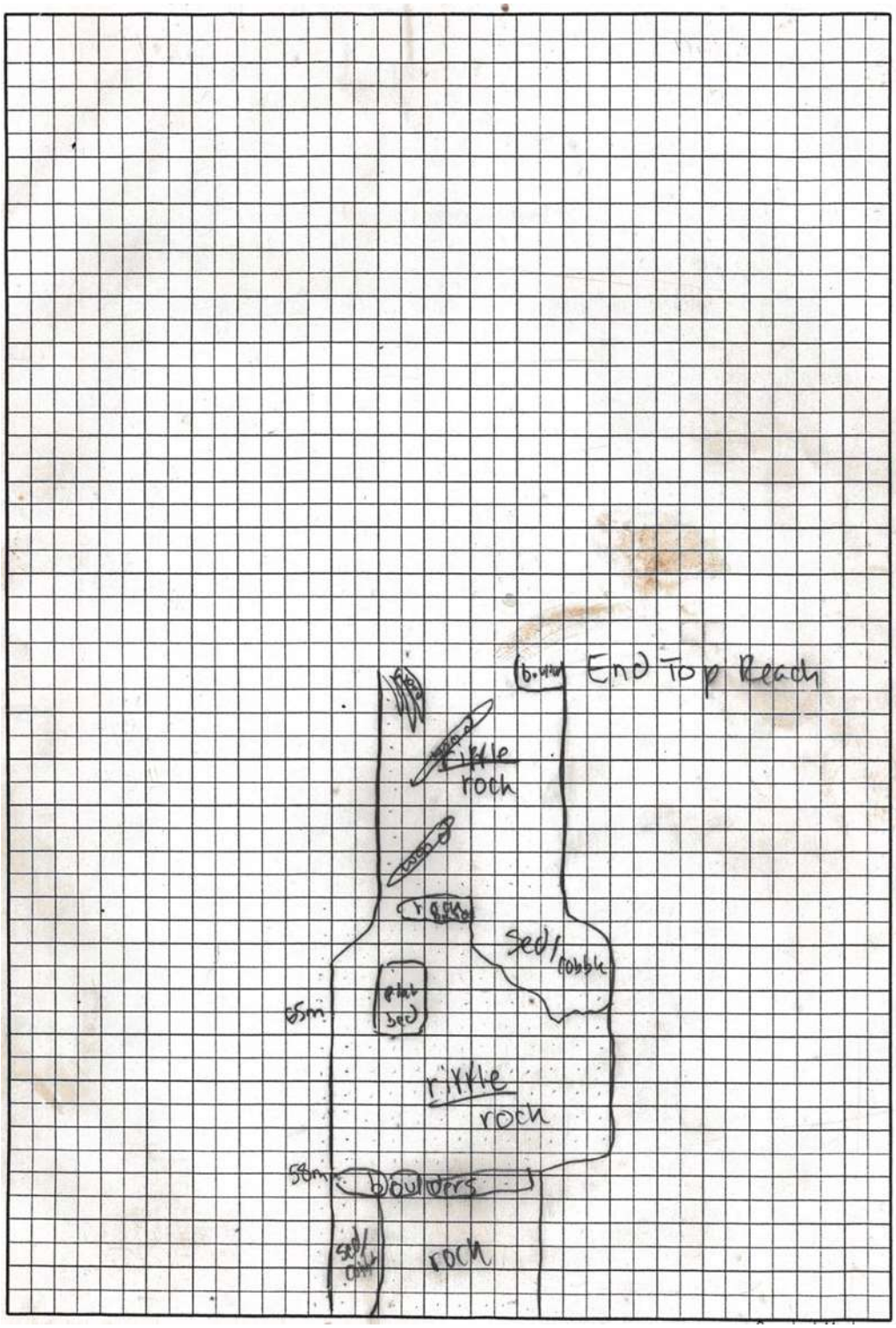

Figure 6A (part 1 of 2): Rough habitat map for a reach of Gibson Jack in Idaho. Each square has dimensions of 0.25 m horizontal x 1 m vertical. Maps begin upstream to downstream.

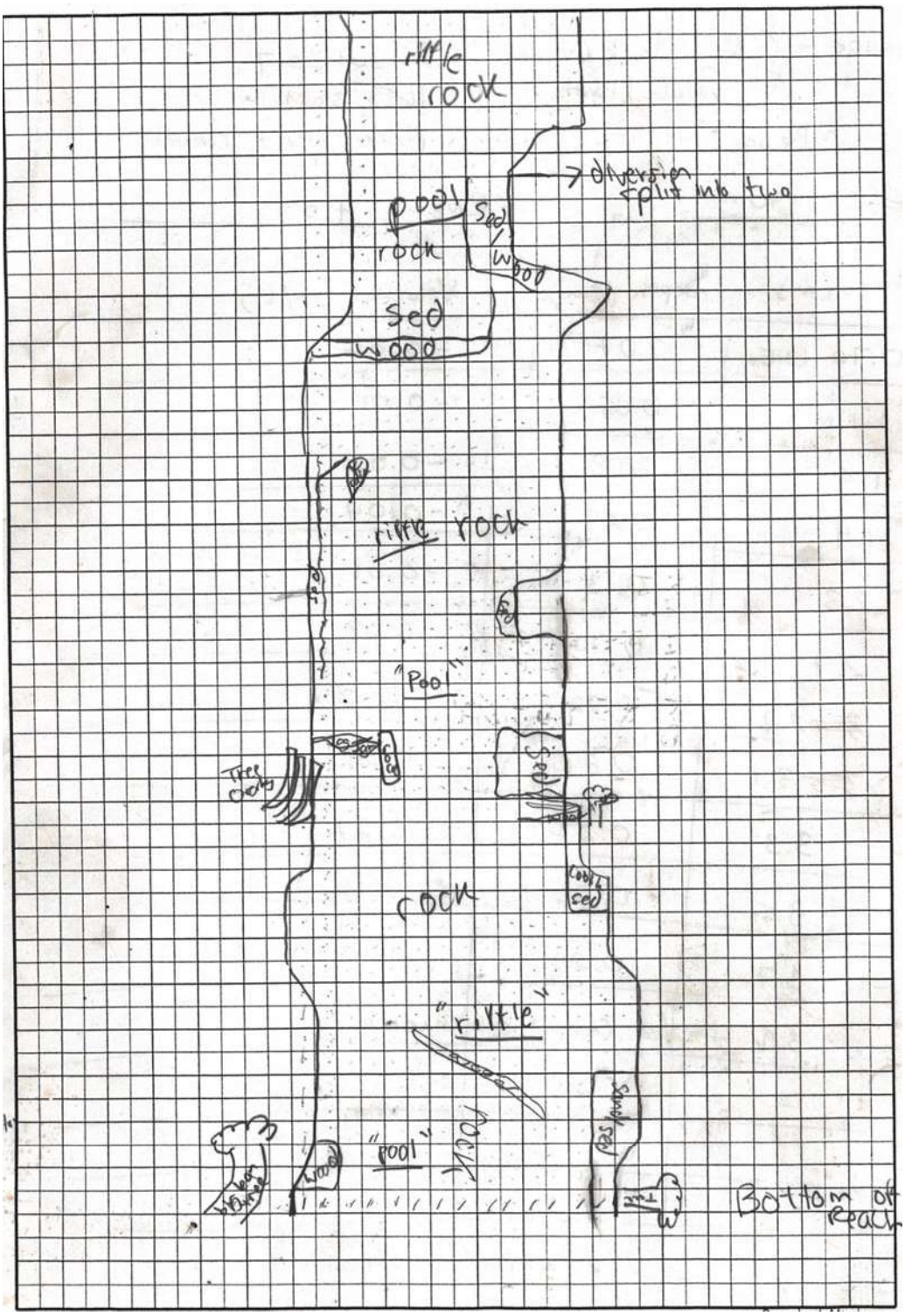

Figure 6A (part 2 of 2): Rough habitat map for a reach of Gibson Jack in Idaho. Each square has dimensions of 0.25 m horizontal x 1 m vertical. Maps begin upstream to downstream.

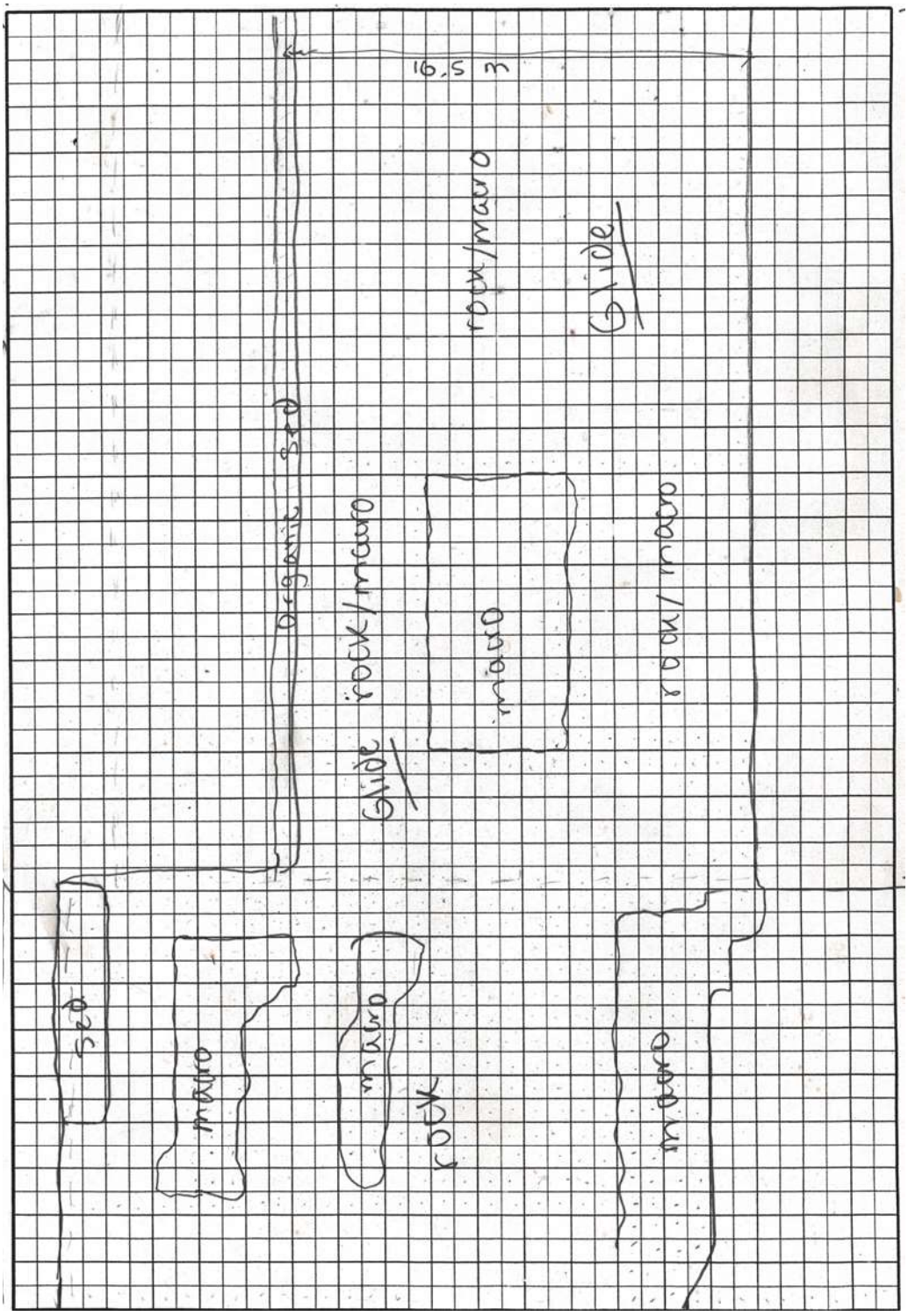

Figure 7A (part 1 of 2): Rough habitat map for a reach of Upper Portneuf in Idaho. Each square has dimensions of 0.5 m horizontal x 0.5 m vertical. Maps begin upstream to downstream.

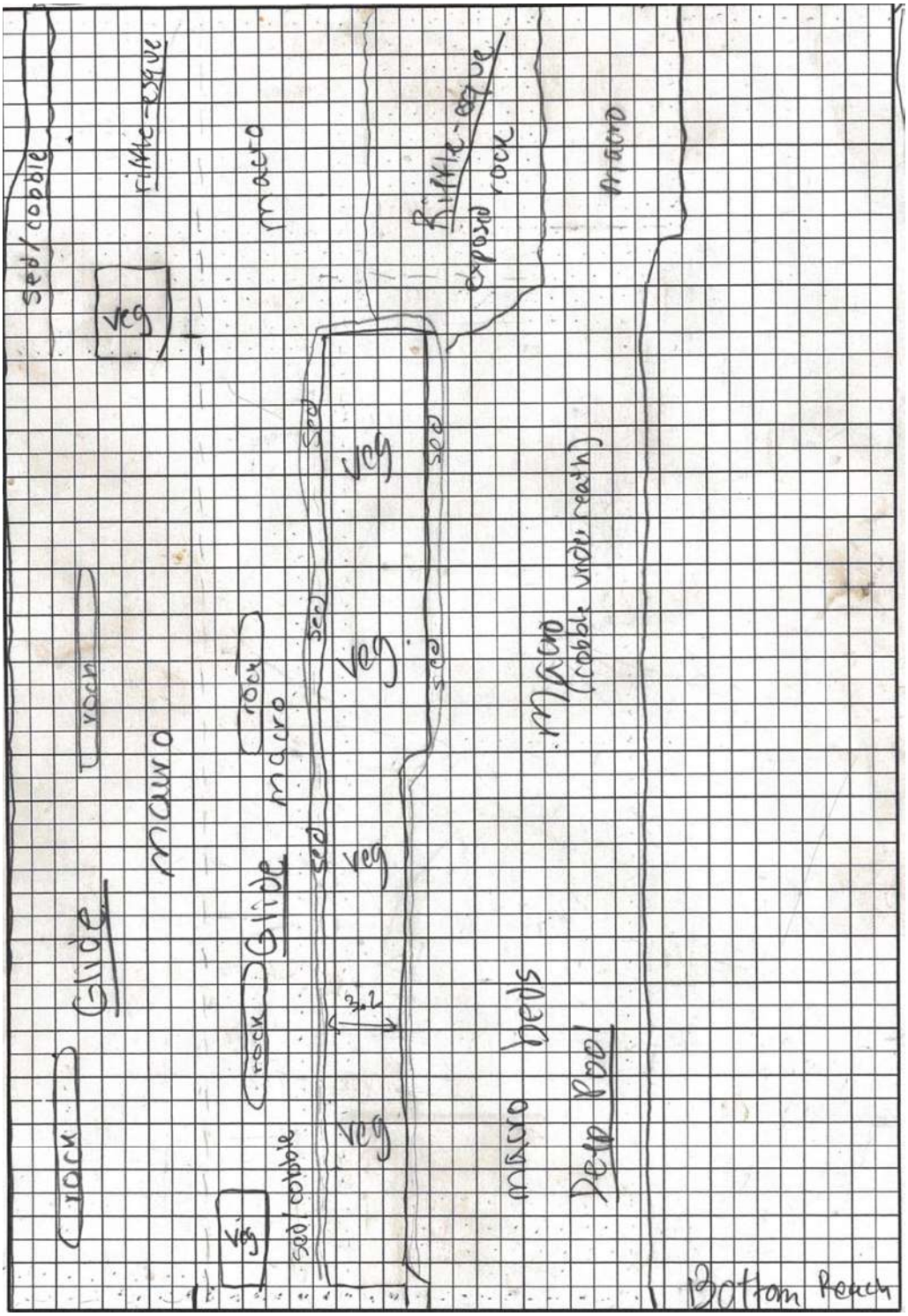

Figure 7A (part 2 of 2): Rough habitat map for a reach of Upper Portneuf in Idaho. Each square has dimensions of 0.5 m horizontal x 0.5 m vertical. Maps begin upstream to downstream.

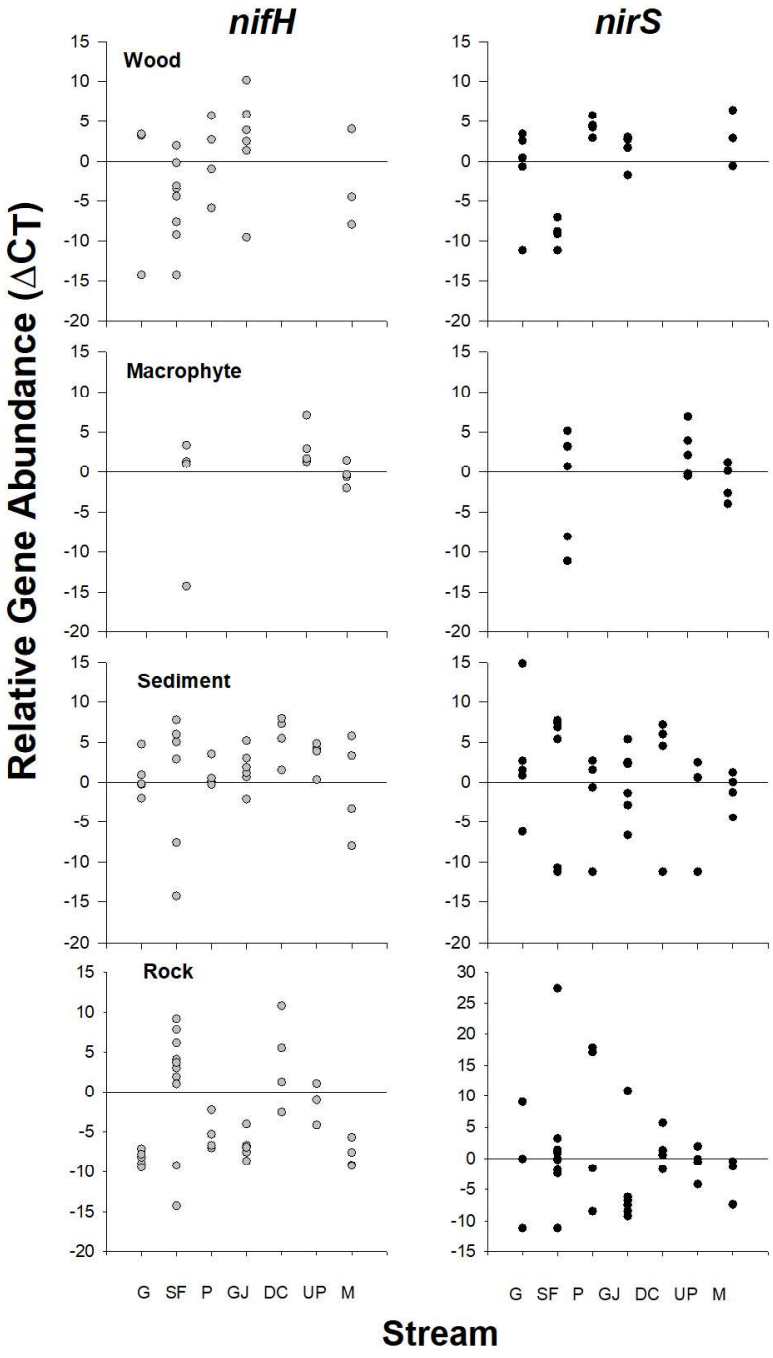

Figure 8A. Plot of relative abundances (C_T) of *nifH* and *nirS* across the 7 study streams for the 4 substrate types ordered top to bottom: wood, macrophyte, sediment, and rock. Streams are arranged from left to right in increasing dissolved inorganic nitrogen concentrations. Diggie Creek (DC), Gratiot River (G), Gibson Jack (GJ), McGunn (M), Pilgrim River (P), South Fork Mink Creek (SF), and the Upper Portneuf River (UP). Note that these are relativized values to the median, so negative indicates less abundance than the median value and positive indicates more. All Y-axes are the same except for *nirS* on rock substrate.

A.2 Chapter 3

Table 1A. The breakdown of deployment of nutrient diffusing substrate (NDS) in all transects sampled in 2019. Transect points are broken down by habitat type. Y= yes and N = no. The recovered NDS are broken down into the number of replicates of each treatment that were recovered (N, P, N+P, and control), with the highest number of possible replicates being 4. If NDS was deployed at a site but not recovered or not deployed at all, then a "-" is used in the number of replicates.

Site	Transect #	Type	NDS Deployed (Y/N)	NDS Recovered (Y/N)	N Replicates Recovered	P Replicates Reovered	N+P Replicates Recovered	Control Replicates Recovered
Nara	1	Wetland	Y	Y	4	4	4	4
	2	Wetland	Y	Y	4	4	4	4
	3	Wetland-Stream	Y	Y	4	4	4	4
	4	Wetland-Stream	Y	Y	4	4	4	4
	5	Wetland-Stream	Y	N	0	0	0	0
	6	Wetland-Stream	Y	Y	4	4	4	4
	7	Stream	Y	Y	4	3	4	4
	8	Stream	N	N	-	-	-	-
	9	Stream	N	N	-	-	-	-
	10	Lake	N	N	-	-	-	-
	11	Lake	Y	Y	3	4	3	3
Sioux	1	Wetland	Y	Y	4	4	4	4
	2	Wetland	Y	Y	4	4	4	4
	3	Wetland	Y	Y	4	4	4	4
	4	Wetland	Y	Y	4	4	4	4
	5	Wetland-Stream	Y	Y	4	4	4	4
	6	Stream	Y	Y	4	4	4	4
	7	Stream	Y	Y	4	4	3	3
	8	Stream	Y	Y	4	4	4	4
	mouth	Lake	N	N	-	-	-	-
Mackinac	1	Wetland	Y	Y	4	1	3	2
	2	Wetland	Y	Y	2	1	0	0
	3	Wetland	Y	N	0	0	0	0
	4	Wetland	Y	Y	1	1	3	2
	5	Wetland	Y	Y	3	4	1	3
	6	Stream	Y	Y	2	1	2	2
	7	Wetland-Stream	N	N	-	-	-	-
	202	Stream-Lake	Y	Y	3	0	1	0
	9	Lake	N	N	-	-	-	-

Table 2A. The breakdown of deployment of nutrient diffusing substrate (NDS) in all transects sampled in 2020. Transect points are broken down by habitat type. Y= yes and N = no. The recovered NDS are broken down into the number of replicates of each treatment that were recovered (N, P, N+P, and control), with the highest number of possible replicates being 4. If NDS was deployed at a site but not recovered or not deployed at all, then a "-" is used in the number of replicates.

Site	Transect #	Type	NDS Deployed (Y/N)	NDS Recovered (Y/N)	N Replicates Recovered	P Replicates Reovered	N+P Replicates Recovered	Control Replicates Recovered
Saganing	1	Wetland	Y	Y	4	4	4	3
	2	Wetland-Stream	Y	Y	4	4	3	2
	3	Wetland	Y	N	0	0	0	0
	4	Wetland	Y	Y	4	4	4	3
	5	Wetland	Y	Y	3	4	4	4
	6	Wetland	Y	Y	4	4	4	3
	7	Wetland	Y	Y	3	4	3	2
	8	Wetland	Y	Y	4	4	2	4
	9	Wetland-Stream	Y	Y	4	4	4	4
	10	Wetland-Stream	Y	Y	4	4	4	4
Wildfowl	1	Wetland	Y	Y	4	4	3	3
	2	Wetland-Stream	Y	N	0	0	0	0
	3	Wetland	Y	Y	4	4	4	4
	4	Wetland	Y	Y	3	4	4	4
	5	Wetland	Y	Y	4	4	4	4
	6	Wetland	Y	Y	4	4	4	4
	7	Wetland-Stream	Y	Y	4	3	4	4
	8	Wetland-Stream	Y	N	0	0	0	0
	9	Wetland-Stream	Y	N	0	0	0	0
	10	Wetland-Stream	Y	Y	4	4	4	4

Figure 9A. ArcGIS image of the Nara transect. The pink dots denote transect points. From left to right the transect points start at 1 and end at 11.

Figure 10A. ArcGIS image of the Sioux transect. The pink dots denote transect points. From top to bottom the transect points start at 1 go to 8. The mouth transect point is in the nearshore lake area.

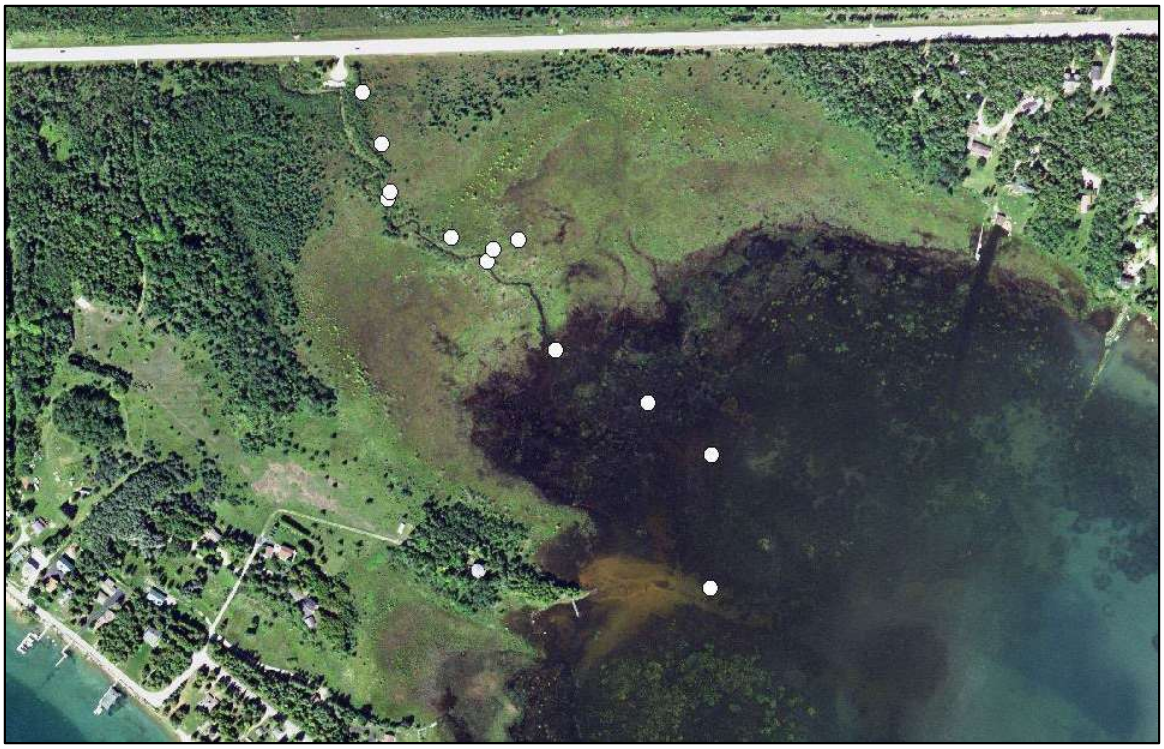

Figure 11A. ArcGIS image of the Mackinac transect. The white dots denote transect points. From top to bottom the transect points start at 1 go to 12.

Figure 12A. Google earth image of the Saganing transect. The green push pins denote transect points. From top to bottom the transect points start at 1 go to 10.

Figure 13A. Google earth image of the Wildfowl transect. The green push pins denote transect points. From right to left the transect points start at 1 go to 10.