UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

LOOKING OUT FOR THE LITTLE GUYS: HOW MUSSELS FACILITATE MICROBES AND SCIENTISTS FACILITATE STUDENTS

A DISSERTATION

SUBMITTED TO THE GRADAUTE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

EDWARD DOUGLAS HIGGINS
Norman, Oklahoma
2022

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/336581

LOOKING OUT FOR THE LITTLE GUYS: HOW MUSSELS FACILITATE MICROBES AND SCIENTISTS FACILITATE STUDENTS

A DISSERTATION APPROVED FOR THE

DEPARTMENT OF BIOLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Caryn Vaughn, Chair

Dr. Carla Atkinson

Dr. Cameron Siler

Dr. Lara Souza

Dr. Krithi Sankaranaryanan

© Copyright by EDWARD DOUGLAS HIGGINS 2022

All Rights Reserved.

ACKNOWLEDGEMENTS

I would like to thank the Department of Biology, the Oklahoma Biological Survey, and the University of Oklahoma for the opportunity to pursue my Ph.D. The faculty, staff, and students here have been tremendously supportive, and I have neither the space, nor the memory to list everyone here.

Most importantly, I would like to thank my advisor, Dr. Caryn Vaughn. I have undergone enormous growth as a scientist and as a person under her mentorship. My interests are varied and sometimes scattered, and Caryn has had the patience, wisdom, and more patience to successfully guide me through these past five years. Without her encouragement and support, I would have stuck to a much narrower path through graduate school, and I hope to be as good of a mentor as she is someday. Thank you doesn't begin to cover it. I am also grateful to my advisory committee, each of whom provided support and guidance during my Ph.D.; Dr. Carla Atkinson, Dr. Cam Siler, Dr. Lara Souza, and Dr. Krithi Sankaranarayanan. The breadth and depth of expertise from my committee members was vital to the development and execution of these projects and I am extremely grateful.

Although he wasn't on my committee, Dr. Thomas Parr certainly counts as an honorary member. Thomas seems to have an unceasing amount of background knowledge on the topics that interest me most and he was always happy to talk through them. It was an honor any time I was able to contribute something he didn't know. I'm fortunate to have him as a mentor and I'm better for it.

Thank you to Dr. Jeff Kelly and Dr. Sara Mata and not just for your work on the NRT program (for which I am indebted to the whole team!). Because of the NRT, I have been able to accomplish way more than I thought I'd be able to do in grad school. Jeff and Sara also went

above and beyond when they showed that they believe in me and supported my crazy idea to conduct a research project alongside a workshop I was working on. Along those lines, thank you Dr. Hayley Lanier, Dr. Katie Marske, and Dr. Lara Stein for all of your time and help with that same research and workshop!

Jonathan Lopez approached me out of the blue in the auditorium of my first official training at OU to cheerily introduce himself as a fellow Vaughn Lab Ph.D. student starting at the same time. He would continue to show that same level of extroverted camaraderie throughout our time at OU and I am glad that he did. He very rudely finished his degree ahead of me and I miss him already. I would then meet Traci Dubose and Beth Tweedy in the Vaughn Lab as well. The three of us, plus the rest of our Dungeons and Dragons crew (Elyse Freitas, Sean Dubose, Rachel Hartnett, Brian Mannel, and Ben Wilder), would go on to make a lot of memories that played no small part in getting me through my degree. Adventure!

I am grateful for McKenna Burns, a resilient, kind-hearted, and valued friend and collaborator. Her hard work, along with that of Paula Cimprich, Mehrnaz Akhami, Meelyn Pandit, and Liz Besozzi, has ensured that the workshop in my third chapter has continued for years after the first one. It's an incredible program that has helped hundreds of students over the years and I hope it continues.

Thank you to my friends that have shown me support from afar—Chris Jaeger, Brent Knauff, Mike Weber, and Friendan Pitts have all been my buddies for what feels like forever. Caroline Shaw is a constant in my life and has helped me through a lot of my struggles. Her friendship means the world to me. Trevor, I wish you were around to see me succeed. I miss you.

I am especially grateful for Michelle Busch. She has become one of my closest friends over the past few years. Throughout our time at OU, Michelle was there for me. From our first

time hanging out getting free pizza, to talking through research problems last week, Michelle has always been a compassionate friend and advocate. I owe a lot of my mannerisms to her and I spending so much time together. Thanks dude, you're the best.

Ben Wilder, my partner, left a perfectly decent job and moved with me to Oklahoma. It was a silly choice. It is also a choice for which I will never be able to repay him (but I'll sure try anyway). He has provided me stability, sanity, and support. He bought me Goldfish for my general exam. He is the reason I am who I am today. I could write several more paragraphs dedicated to thanking him specifically, but I won't. Thank you, Ben.

Finally, this dissertation is dedicated to my older sister, Sam Higgins, my father, Edward Higgins, and my mother, Terry Higgins. Sam has been a role model for me, whether she wanted the attention or not, since I was young and I don't see that changing any time soon. My dad has worked extremely hard to make sure I know I am supported. It's difficult to take risks without a safety net—I'm glad I won't have to. My mother has contributed an incredible amount to my success. I cherish my talks with her more than she knows. She is, hands down, the strongest person I know, and I owe a lot of the resilience I've been able to muster to her. I love my family dearly and I owe them so much gratitude.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv

TABLE OF CONTENTS vii

LIST OF TABLES ix

LIST OF FIGURES xi

ABSTRACT xiv

CHAPTER ONE 1

- Abstract 2
- Introduction 3
- Methods 5
- Results 10
- Discussion 12
- Acknowledgements 18
- Literature cited 19
- Tables 30
- Figure legends 37
- Figures 38

CHAPTER TWO 42

- Abstract 43
- 1 | INTRODUCTION 45
- 2 | MATERIALS AND METHODS 48
- 3 | RESULTS 54
- 4 | DISCUSSION 57
- ACKNOWLEDGEMENTS 64
- LITERATURE CITED 65
- Tables 75
- Figure legends 77
- Figures 79

CHAPTER THREE 85

- Abstract 86
- Introduction 87
- Methods 91
- Results 95
- Discussion 97
- Acknowledgements 103
- Literature cited 104
- Figure legends 112

Tables..... 113
Figures..... 114

LIST OF TABLES

Chapter One

Table 1. Pairwise PERMANOVA results comparing microbial communities among microhabitats with Holm adjusted p values. 30

Table 2. Axis loadings for PCA comparing sediment, shell, and biodeposit microbial communities. Higher r values correlated with axis 1 suggest taxa associated with sediment and lower values with mussel shells. Lower r values correlated with axis 2 will correlate with taxa differentiating mussel biodeposits from shell and sediment communities. 31

Table 3. PERMANOVA outputs for sediment, shell, and biodeposit models. Statistically significant results in bold. Higher R2 values indicate more variance is explained by a given variable. 32

Table 4. Axis loadings for PCA sediment microbial communities. Higher r values correlated with axis 1 suggest taxa associated with subsurface sediments while lower values suggest taxa associated with surface sediments. Lower r values correlated with axis 2 may suggest values associated with the site K1. 33

Table 5. Axis loadings for PCA biodeposit microbial communities. Higher r values correlated with axis 1 suggest taxa associated with the site K3 while lower values suggest taxa associated with K2 and K1. 34

Table 6. Axis loadings for PCA comparing shell microbial communities. Higher r values correlated with axis 1 suggest taxa distinguishing the sites K1 and K2 from K3 while lower

indicates the opposite. Higher r values correlated with axis two suggest taxa distinguishing sites K2 from K3 and K1. 35

Table 7. Wolman pebble counts, substratum heterogeneity (D60/D10), flow measurements, and canopy cover collected in 2016, from Vaughn et al. 2021. 36

Chapter Two

Table 1. Pairwise comparisons of PERMANOVA models testing for effects of mussels and nutrient on sediment microbial community structure between nutrient treatments on day 7. "N:M" indicates the effect of the interaction between nutrient and mussel treatments. Significant values are in bold. Significant effects of single variables were not evaluated if the interaction effect was significant. 75

Table 2. ANOVA results for ecoenzyme activity and ecoenzyme activity ratios. Significant results are in bold. 76

Chapter Three

Table 1. Course titles and student enrollment. All courses had a cap of 25 students, however, introductory courses had two sections. GS indicates a graduate student instructor or assistant instructor, F indicates a faculty member of either OU or MU, and SL indicates a science librarian. 113

LIST OF FIGURES

Chapter One

Figure 1. Map of the Kiamichi River drainage in Southeastern Oklahoma..... 38

Figure 2. Comparisons of microbial community (A) richness and (B) evenness among sampled microhabitats. 39

Figure 3. Principal coordinates analysis including samples from all microhabitats. Ellipses are drawn around centroids with a 50% confidence interval..... 40

Figure 4. Principal coordinates analysis of (A) mussel biodeposit microbial communities, (B) mussel shell microbial communities, and (C) sediment microbial communities. Communities are colored based on the site from which they were sampled. Ellipses are drawn around centroids with a 50% confidence interval. Vectors were added post hoc and are based on envfit analysis 41

Chapter Two

Figure 1. (A) Sediment and mussels from the Kiamichi River were placed in enclosures within each flow-through mesocosm. (B) Experimental design used in study. Colors and letters correspond to nutrient treatments (N – nitrogen, P – phosphorus, NP – both nitrogen and phosphorus, ctrl – no nutrient control) and line type represent mussel treatment (solid – mussel treatment, dashed (non-mussel).. 79

Figure 2. Boxplots of alpha diversity metrics of sediment microbial communities over sampling dates. (A) Shannon diversity. (B) Inverse Simpson's Evenness Index..... 80

Figure 3. PCA of sediment microbial communities collected on the first day of the experiment before the introduction of mussels and the final day of the experiment for (A) nutrient control

treatments, (B) nitrogen treatments, (C) nitrogen and phosphorus treatments, and (D) phosphorus treatments. Lines are drawn from the centroid of the samples collected on the first day to the centroid of samples on the final day. Ellipses are drawn at a 50% confidence interval around centroids. Lighter shades and dashed lines indicate non-mussel treatments. Samples collected on the first day of the experiment are all considered non-mussel treatments as mussels had not been introduced. All panels are based on the same PCA but with all ellipses other than the nutrient in question masked for clarity. 81

Figure 4. Boxplots of alpha diversity metrics for samples collected on the final day of the experiment by nutrient and mussel treatment. (A) Shannon diversity. (B) Inverse Simpson's Evenness Index. 82

Figure 5. PCA of sediment microbial communities collected on the final day of the experiment for (A) control treatments, (B) nitrogen treatments, (C) nitrogen and phosphorus treatments, and (D) phosphorus treatments. Ellipses are drawn at a 50% confidence interval around centroids. Lighter shades and dashed lines indicate non-mussel treatments. All panels are based on the same PCA, but with all ellipses other than the nutrient in question masked for clarity. 83

Figure 6. Boxplots of ecoenzyme activity within nutrient and mussel treatments on the final day of the experiment. Enzyme activities shown are (A) C:N enzyme activity ratio, (B) C:P enzyme activity ratio, (C) N:P enzyme activity ratio, (D) BG enzyme activity (ln[nmol · hr-1· mg AFDM-1]), (E) sum of NAG + LAP enzyme activity (ln[nmol · hr-1· mg AFDM-1]) and, (F) AP enzyme activity (ln[nmol · hr-1· mg AFDM-1]) 84

Chapter Three

Figure 1. Participant demographics based on survey results 114

Figure 2. Box plots of student survey responses. (A) Student comfort when approaching a problem requiring coding knowledge. Higher values indicate more self-confidence. (B) Student confidence in their ability to use coding software or coding skills. Higher values indicate more self-confidence. (C) Student comparison of their coding abilities compared to their peers' abilities. A value of 2 or lower indicates the student thinks they are less skilled than their peers, 3 indicates they are equally as skilled, and 4 or higher indicates they believe they are more skilled. (D) Student self-assessment of the overlap of their own identity and that of a STEM professional. Lower values indicate they do not view themselves as a STEM professional, a value of four indicates they view themselves partially as a STEM professional, and higher values indicate much more self-concept as a STEM professional..... 115

ABSTRACT

Animals can have large impacts on how ecosystems function, from influencing population dynamics of other plant or animal species, to modifying hydrogeological flow dynamics. One ecosystem function that has received widespread attention is the biogeochemical transformation of key nutrients required for primary production. Freshwater mussels are benthic species that in dense communities, act as biogeochemical hotspots with landscape-level impacts. Mussels can cycle nutrients through their own metabolism, but the observed changes in ecosystem-scale nutrient transformations are also largely influenced by microbial metabolism. My dissertation examined the complex interactions between freshwater mussels and environmental microbial communities and how these interactions shape nutrient dynamics. Then, I investigated how scientists can effectively communicate their research in ways that promote participation in the broader scientific community.

Chapter one explored microbial communities associated with mussels in discrete, but proximate microhabitats in a southern US river. Mussel microbiomes (shell and biodeposits), were less diverse than those in surface and subsurface sediments. Mussel abundance was a significant predictor of sediment microbial community composition. Mussel species richness and distance between sample sites were not significant predictors of microbial community composition. These data suggest that mussels and local habitat conditions that change dynamically along streams, such as discharge, water turnover, and canopy cover, work in tandem to influence environmental microbial community assemblages at discreet rather than landscape scales. Further, mussel burrowing activity and mussel shells may provide interactions between microbial communities critical to nutrient cycling in these systems

Chapter two investigated how mussels can influence microbial community structure and function in the sediment under different nutrient regimes. I transplanted freshwater mussels and natural river sediment to flow-through mesocosms with different nutrient amendment treatments and monitored changes in microbial community composition over one week. I compared these microbial communities to activity measurements of ecoenzymes known to correlate to microbial function and nutrient availability. Mussels always changed sediment microbial community composition, but the final microbial community composition was also dependent on ambient nutrient concentrations. Further, mussels homogenized the stoichiometric ratios of ecoenzyme activities, indicating a consistent function of sediment microbes associated with freshwater mussels. My results suggest that mussels may promote functional redundancy in sediment microbial communities and highlight the importance of animals in controlling biogeochemical transformations under changing nutrient conditions.

Chapter three studied how STEM intervention programs can effectively improve the self-efficacy and self-concept of younger generations of scientists. I surveyed students who participated in a week-long, virtual workshop with a focus on computing in the biological sciences. The workshop had daily seminars and a career mixer to introduce students to scientists in academic, government, and commercial careers. Courses were most frequently taught by graduate students to implement a peer-to-peer mentoring strategy. The workshop was successful in improving student self-efficacy but had more modest success in improving student self-concept. Results were not different among demographic groups, indicating this STEM program did not mirror detrimental impacts on minority students seen systemically across STEM.

Taken together, this dissertation demonstrates that investigations into interactions between animal and microbial communities can improve our understanding of how ecosystems function.

Further, it demonstrates that effective outreach and education of scientific knowledge is crucial to the continuity of science as a broader ecosystem.

CHAPTER ONE

Mussels and local conditions interact to influence microbial communities in mussel beds

Keywords:

freshwater, mussel, microbiome, nutrient cycling, sulfur, ecosystem function

Published in *Frontiers in Microbiology*

Higgins, E., Parr, T. B., & Vaughn, C. C. (2022). Mussels and Local Conditions Interact to Influence Microbial Communities in Mussel Beds. Frontiers in Microbiology, 12. doi: 10.3389/fmicb.2021.790554

Abstract

Microbiomes are increasingly recognized as widespread regulators of function from individual organism to ecosystem scales. However, the manner in which animals influence the structure and function of environmental microbiomes has received considerably less attention. Using a comparative field study, we investigated the relationship between freshwater mussel microbiomes and environmental microbiomes. We used two focal species of unionid mussels, *Amblema plicata* and *Actinonaias ligamentina*, with distinct behavioral and physiological characteristics. Mussel microbiomes, those of the shell and biodeposits, were less diverse than both surface and subsurface sediment microbiomes. Mussel abundance was a significant predictor of sediment microbial community composition, but mussel species richness was not. Our data suggest that local habitat conditions that change dynamically along streams, such as discharge, water turnover, and canopy cover, work in tandem to influence environmental microbial community assemblages at discreet rather than landscape scales. Further, mussel burrowing activity and mussel shells may provide habitat for microbial communities critical to nutrient cycling in these systems.

Introduction

Key ecosystem processes are carried out by both microbes and animals, but microbial communities are particularly important to evaluate in tandem with animal influences on ecosystem function as microbiome data combined with environmental data improve our understanding of ecosystem processes (Graham et al., 2016). Further, effects of animals on microbial communities are important and underexplored (Skelton et al., 2017; Fitzpatrick et al., 2018; Thoemmes and Cove, 2020). As more systems are investigated, the implications of animal-microbial interactions and their impact on ecosystem function become more apparent. For example, marine birds translocate nutrients from the ocean to islands, where those nutrients increase organic matter decomposition rates by soil bacteria (Fukami et al., 2006); earthworms affect the function, but not community composition, of methanotrophic bacteria in landfills (Héry et al., 2008); and marsupial burrowing activity causes successional shifts in microbial community composition and increases nitrogen availability in soils (Eldridge et al., 2015). It is particularly important to understand baseline interactions between animals and microbes in the wake of climate and land use change. In this context, streams are a good study system because they are globally threatened by pollution and climate change (Jury and Vaux, 2005).

Within these systems, freshwater mussels (bivalve mollusks in the order Unionida) are large (~10 to 100 mm adult shell length), long-lived (~10 to over 100 yrs) benthic animals that perform important ecosystem functions in streams, such as filtering the water and recycling and storing nutrients (Vaughn and Hoellein 2018). Freshwater mussels are globally imperiled and as mussel communities shift and populations decline (Spooner and Vaughn, 2008; Vaughn et al., 2015), evaluating mussel-microbiome interactions is critical to predicting changes in ecosystem function. Mussels often occur as dense (10 ~ 100 mussels/m^2), multispecies aggregations called

mussel beds and can comprise a significant portion of benthic biomass (Vaughn and Spooner, 2006). These communities can have large impacts on both biotic and abiotic factors in streams (Vaughn and Hoellein, 2018).

Of importance to stream microbial function, filter-feeding mussels burrow in the sediment and transform as well as transport organic matter from the water column into the sediment via excretion and biodeposition of feces and pseudofeces (rejected particles encapsulated in mucus and expelled before ingestion). We know mussel beds can significantly influence nutrient cycling in the sediment on an ecosystem-wide scale (Hoellein et al., 2017; Nickerson et al., 2019) indicating interactions with sediment microbiomes. Aquatic sediment is a unique environment in which the interface between an oxygenated surface and an anoxic subsurface microhabitat is relatively shallow and mussel burrowing activity can directly influence the microhabitats of both layers, often introducing oxic microniches into anoxic habitats (Brune et al., 2000). Levels of oxygenation can affect microbial community composition and function (Sinsabaugh et al., 2009) and benthic organisms can couple microbially driven biogeochemical processes (nitrification-denitrification, elemental sulfur cycling, etc; Nickerson et al., 2019). However, incorporating drivers of benthic microbial community structure and diversity into riverine ecosystem function requires further research (Zeglin, 2015).

Here we consider the mussel microbiome to be comprised of the microbial communities on mussel shells and in their biodeposits. How these communities interact with the sediment microbial communities, and potential differences between mussel species in how this occurs, is a key research gap. Host physiology and diet are known to impact hosted microbiomes in a diversity of organisms (Turnbaugh and Gordon, 2009; Phillips et al., 2012; Ye et al., 2014; Pierce et al., 2016) and freshwater mussels have species-specific physiological and behavioral

traits (Haag, 2012). Interspecific differences may influence the mussel microbiome and therefore interactions with the sediment microbiome. For example, mussel species investigated in southeastern Oklahoma U.S. are either thermally sensitive (e.g. *Actinonaias ligamentina)* or thermally tolerant (e.g. *Amblema plicata),* and thermally sensitive species excrete nutrients at higher rates at warm temperatures, with different stoichiometric ratios when stressed (Spooner and Vaughn, 2008). Additionally, Allen and Vaughn (2009) found that mussels exhibit species-specific differences in burrowing activity with thermally sensitive species demonstrating higher activity.

Here we asked, how similar are the freshwater mussel microbiome and the sediment microbiome, and how do these relationships change with mussel abundance, species composition, and environmental conditions? We addressed these questions with a field study comparing benthic microbiomes in three mussel beds in a small river in the southern US focusing on two dominant mussel species. We sampled microbial communities from four microhabitats: the surface layer of sediment, sediment from 6 – 10 cm below the surface, mussel shells, and mussel biodeposits. Microbes were identified using 16S rRNA analysis. We predicted that environmental conditions among microhabitats would be sufficiently distinct to host unique assemblages of microbes. We expected that differences in mussel species' behavior and nutrient excretion would produce species-specific host-associated microbial community composition. We also expected that microbial community structure within the sediment would reflect mussel-associated changes in biogeochemical cycling.

Methods

Study area and focal mussel species – We conducted our study in the Kiamichi River, Oklahoma, a well-studied stream in the southcentral U.S. known for its high freshwater mussel

biodiversity (Matthews et al., 2005). Mussel assemblages in this river are typically dominated by two species, *Actinonaias ligamentina* and *Amblema plicata*, that make up ~70% of mussel biomass in this region but differ morphologically, behaviorally, and physiologically (Vaughn, 2010; Hopper et al., 2019). *Amblema plicata* has a ridged shell and tends to be sedentary, while *A. ligamentina* is an active burrower with a smooth shell (Allen and Vaughn, 2009). The two species also differ in their thermal preferences, which influences filtration rates as well as nutrient excretion rates and stoichiometry (Spooner and Vaughn, 2008; Trentman et al., 2018).

Field study – In July 2018, we collected microbial samples from mussel beds in the Kiamichi River. We selected three sites (Fig. 1) with previously documented abundant, diverse mussel assemblages and data on mussel roles in nutrient recycling and storage (Atkinson et al., 2013; Atkinson and Vaughn, 2015; Hopper et al., 2018). Sites varied in abiotic characteristics that may influence environmental microbiomes such as flow, substrate, and shade. Thus, we characterized sites by measuring flow (with a Hach LDO meter), sediment particle sizes (with Wolman pebble counts), and shading (using a densiometer) in summers 2015 – 16 as part of a larger study (Hopper et al., 2018; Vaughn et al., 2021).

We sampled four individuals of both *A. plicata* and *A. ligamentina* from mussel beds at three sites along the river. For each mussel bed, we first conducted tactile searches on the sediment surface to locate mussels. Searches were conducted from downstream to upstream to minimize disturbance of the sediment. We then placed a 0.25 m^2 quadrat around locations that contained at least one individual of each species. Although quadrats could contain multiple individuals of each species, they all contained both focal species, and we only sampled one individual of each species per quadrat. We had a total of 12 sampling locations and 24 total mussels (12 *A. plicata* and 12 *A. ligamentina*) across all sites. We used a custom 4.9 cm

diameter, clear acrylic sediment corer to collect one sediment core from each quadrat, for a total of 12 sediment cores. To prevent cross-contamination between samples, we used 90% ethanol to rinse the interior of our corer between samples. Mussels have been shown to affect bacterial growth and metabolism at depths of 6 cm or greater below the surface (McCall et al., 1986; Black and Just, 2018) so we extruded the sediment from the corer with a rubber stopper and took subsamples from the surface layer (n = 12) and 6-10 cm below (n = 12) using an ethanol rinsed, flame sterilized spatula.

We removed mussel individuals from the sediment and used sterile razor blades to collect a single biofilm sample from the shell of each mussel. Blades were rinsed with 90% ethanol, wiped with a sterile kimwipe, and flame sterilized for one minute between samples (Horton et al., 2019; Miller-ter Kuile et al., 2021). Five shell samples did not successfully sequence (three from *A. ligamentina* and two from *A. plicata*) resulting in a final n of 19. We then gently scrubbed mussel shells using sterile nylon mesh to remove the remaining biofilm and left them in containers with 1L of filtered river water for 4 to 6 hours to allow time for mussels to biodeposit sufficient material for collection. Biodeposits were collected using an ethanol rinsed, flame-sterilized spatula. Similarly, not all biodeposits samples sequenced successfully (2 from each species) for a final n of 20. After mussels were removed from the sediment, quadrats were excavated to a depth of 15 cm and any additional mussels were identified to species and counted (Vaughn et al., 1997). While storage at -80° C is considered optimal for microbial community samples, short term cold storage demonstrates little change in fecal and soil microbiome community structure (Rubin et al., 2013; Choo et al., 2015), and so all microbiome samples were placed in sterile cryovials, stored on ice in coolers on the shaded riverbank for no more than five hours, and then placed in liquid nitrogen until transfer to a -20° C freezer within four days.

Amplicon library construction and sequencing – All samples were thawed, spun at 10,000 x gravity for two minutes, and water was removed via pipette. DNA was extracted using DNeasy PowerSoil® kits (Qiagen, Hilden, Germany). We amplified the v4 region of the 16s rRNA gene using primers and PCR protocols from (Kozich et al., 2013). We purified post PCR samples with Ampure XP beads (BeckmanCoulter, Indianapolis, IL, U.S.) at 1x concentration, quantified with a Qubit Fluorometer (ThermoFisher Scientific, Waltham, Massachusetts, U.S.), diluted with lab grade water to 4nM equimolar concentrations, and pooled. Library preparation was performed at the Sam Noble Oklahoma Museum of Natural History and library sequencing was performed at the University of Oklahoma Consolidated Core Lab using 2 × 250 bp paired-end sequencing on an Illumina MiSeq.

Bioinformatics and data analyses – Sequencing reads were merged and filtered and using the program '*AdapterRemoval'* (Lindgreen, 2012). We performed closed reference OTU picking using '*uParse'* (Edgar, 2013) at 97% sequence similarity and assigned taxonomy with the SILVA reference database (v.32, Quast et al., 2013). After filtering out read abundances less than 0.1% of the average sequencing depth, we quantified richness and evenness of our samples with the number of unique OTUs and the Berger-Parker Dominance Index respectively, using Quantitative Insights into Microbial Ecology (QIIME; Berger and Parker, 1970; Bolyen et al., 2019). We used Kruskal-Wallis tests to determine statistical differences in richness and evenness using the base R software (R Core Team, 2020). Significant results were further examined using Holm adjusted pairwise-Wilcoxon Rank Tests between microhabitats (Wright, 1992).

To quantify differences in beta diversity we calculated an Aitchison distance matrix (Euclidean distance of centered log-ratio transformed OTU counts) in R using the Compositions and Vegan packages (van den Boogaart and Tolosana-Delgado, 2008; Filzmoser et al., 2010;

Gloor et al., 2017; Quinn et al., 2018; Oksanen et al., 2020). We used PERMANOVA to determine differences in bacterial community structure and permdisp to determine differences in dispersion among all microhabitats using Vegan's '*adonis'* and '*betadisper'* functions respectively. We conducted post-hoc, pairwise PERMANOVA to evaluate differences among mussel and sediment microhabitats using the '*pairwise.adonis*' test from the pairwiseAdonis package (Wright, 1992; Martinez, 2020). We then tested the effects of site, sediment layer, mussel abundance and mussel richness on sediment microbial community structure as well as the effects of site and mussel species on shell and biodeposit microbial community structure. For both models, we used the 'adonis2' function in Vegan for which the relative importance of each term is indicated by an R^2 value (McArdle and Anderson, 2001; Oksanen et al., 2020). Differences in dispersion for these models were tested with Holm adjusted '*betadipser'* calculations. Environmental variables measured at each site (Ratio of D60 to D10 Wolman pebble counts, average discharge, average canopy cover, and average turnover) were correlated with bacterial communities using the '*envfit'* program in vegan.

To visualize differences in community structure, we performed principal coordinates analyses (PCA) using Aitchison distance matrices (Gloor et al., 2017). We were interested in microbial community patterns in each microhabitat, so in addition to our entire dataset we generated individual PCAs for sediment, shell, and biodeposit communities. Then to examine taxa contributing to differences in microbial community structure, we calculated axis loadings of each PCA by calculating Pearson rank-sum correlations between axis scores and CLR transformed abundances using R (Comrey and Lee, 1992; Quinn et al., 2018). Loadings with absolute r values ≥ 0.70 were considered sufficiently correlated to evaluate (Comrey and Lee, 1992; Curry and Patten, 2014; Weingarten et al., 2019). We interpreted correlations on both the

first and second PC axis. With this method, significant correlations with an r value above 0.70 are interpreted such that higher r values indicated a given taxa had higher abundances as PC values increase and r values below -0.70 are interpreted as taxa demonstrating higher abundances as PC values decrease. Further, we were interested in functional differences between microbial communities and so we only interpreted taxa identified to family as higher classifications tended to encompass taxa with broad metabolic and niche preferences.

Results

Microhabitats displayed significant differences in microbial richness ($\chi^2 = 24.65$, $p<0.001$) and evenness ($\chi^2 = 40.23$, $p<0.001$; Fig. 2). Post-hoc Wilcoxon rank-sum tests showed that these differences were likely driven by mussel biodeposit and shell samples that were 44% less rich ($W = 117.5$, $p < 0.001$) and 8% less even ($W = 811$, $p < 0.001$) than sediment samples. Pairwise comparisons further showed that mussel shell and mussel biodeposit microbial communities were not significantly different in richness ($p = 0.31$), but biodeposit samples were significantly less even ($p < 0.001$). The top layer of sediment had no significant differences in richness from lower layers ($p = 0.32$) or evenness ($p = 0.31$). Bacterial community structure ($F = 13.71$, $p = 0.001$) and dispersion ($F = 3.74$, $p = 0.016$) were significantly distinct among the four microhabitats (Fig. 3). Pairwise PERMANOVA demonstrated that all microhabitats were significantly different from each other (Table 1). Axis loading calculations resulted in 15 taxa with identified genera and 28 unique families significantly correlated with either the first or second axis (Table 2).

Within quadrats, mussel abundance ranged from 3 to 13 mussels while richness ranged from 1 to 6 species. The PERMANOVA model testing the effects of site, sediment layer, mussel richness, and mussel abundance on sediment community structure was a statistically significant

fit to these data ($F = 2.52$, $p < 0.001$). Of these variables, the strongest driver of sediment microbial community structure was sediment layer, followed by site, and then mussel abundance (Table 3). Mussel richness was not a significant predictor of sediment community structure ($F = 1.144$, $p = 0.11$). Sediment layer ($F = 0.030$, $p = 0.86$), mussel abundance ($F = 1.65$, $p = 0.58$), and mussel richness ($F = 1.69$, $p = 0.58$) did not show significant differences in dispersion and sites were only marginally significantly different in dispersion ($F = 4.76$, $p = 0.07$; Fig. 4A). Sediment axis loadings resulted in 7 identified genera and 17 unique families (Table 4).

Overall, the mussel biodeposit ($F = 3.01$, $p = 0.001$) and shell ($F = 2.87$, $p = 0.001$) models were statistically significant fits to these data. Mussel biodeposit microbial community structure seems to be driven by site, but mussel species was not significant (Table 3; Fig 4B). Differences in dispersion were significant based on site ($F = 5.75$, $p = 0.0248$) but not mussel species ($F = 0.201$, $p = 0.659$). Similar to biodeposits, shell microbial communities seem driven by site, but not species (Table 3; Fig 4C). There were no significant differences in dispersion by either site ($F = 2.16$, $p = 0.296$) or mussel species ($F = 0.059$, $p = 0.811$). Axis loadings for shell communities resulted in 8 identified genera and 18 unique families while biodeposit communities resulted in 4 identified genera and 16 unique families (Tables 5 and 6).

The K2 mussel bed had larger substrates and pebble sizes that were evenly distributed indicated by a low D60/10. In comparison, the substrates of K1 and K3 were smaller but less evenly distributed. Discharge measurements suggest that the mussel bed at K3 typically had longer water turnover times (Table 7).

Discussion

Our investigation revealed that the microbial communities hosted by freshwater mussels are distinct from those of the surrounding sediment. These microhabitats are in constant contact; mussels deposit feces and pseudofeces directly into the sediment and we collected biofilm from shells that were exposed to surface and subsurface sediment. Yet, sediment communities demonstrated higher alpha diversity than those that were mussel-associated, and we also found low overlap in microbial community composition among animal-associated and environmental microhabitats. Our data indicate that both environmental conditions specific to locations along the river and animal activity shape these microbial communities. Interactions relevant to critical ecological function between these microhabitats can be inferred by examining the ecology of taxa that distinguish these distinct communities.

Site characteristics supersede mussel species identity, abundance, and richness as a driver of microbial community composition within mussel beds – While Weingarten et al., (2019) found that microbial communities *retained* by freshwater mussels were influenced by species as well as site, our study found that microbes on the shell and in material passed through the gut, did not differ between our focal species. These results are complementary. Much of the phytoplankton, detritus, and bacteria filtered by mussels survives gut passage alive and undamaged (Vaughn et al., 2008) and so it is possible that which taxa are retained in the gut may be influenced by mussel species, but taxa that pass through the gut are not. Mussel biodeposits, regardless of species, may reflect the same background food sources and differences in biodeposit community composition may be minimal when occupying the same site. The algae and bacteria able to colonize shells may similarly be site specific and this signal may overwhelm any differences in community on the basis of different shell characteristics between species.

In contrast to species, site was a strong predictor of microbial community composition in every model tested. K1 and K3 sediment communities are more similar to each other than to K2 (Fig 4C) and mussel shell and biodeposit communities at K2 and K1 are more similar to each other than to K3 (Fig. 4B and 4A). These patterns are not entirely expected based on scale and hydrology. If increasing spatial scale were to predict our microbial community assemblages as it can in soil (Averill et al., 2021), then we would expect sites closer together to show greater overlap in community composition yet our most distant sites (K1 and K3) are more similar to each other than to K2 (Fig. 1). Additionally, K1 is located upstream of a tributary impoundment, while K2 and K3 are downstream. Lack of releases from this impoundment during recent severe drought years has led to patchy drying of the lower river and increased water temperatures in shallow areas which has led to mussel declines and changes in mussel community composition (Atkinson et al., 2014; Vaughn et al., 2015). Based on changes in flow regimes and mussel communities, these results are similarly unexpected.

Local characteristics at the stream reach level may offer insight into differences in these microbial communities. Sediment depth and particle size are both significant predictors of microbial community structure in streams (Sliva and Williams, 2005). K2's relatively larger but more evenly distributed sediment particle sizes may drive distinct microbial communities from K1 and K3 and we see this reflected in our *envfit* results which indicate higher values as community compositions grow more distinct from those of K2 (Figure 4, Table 6). Similarly, differences in canopy cover govern shading and thus influence photosynthetic organisms on shells. K3 has the highest canopy cover and contains shell microbial communities that are most distinct from K2 and K1. K3 also has the greatest water turnover time which may partially explain its distinction in biodeposit community composition. Slower turnover in the water

column will increase the duration of seston delivery which impacts what seston mussels filter (Byllaardt and Ackerman, 2014; Mistry and Ackerman, 2018), and therefore egest as biodeposits, as well as impact which bacterial taxa can colonize shells.

Mussel abundance also significantly affected sediment microbial communities. This result is supported by the findings of Black et al. (2017). They investigated relationships between sediment microbial communities and the presence or absence of mussels in the upper Mississippi River and found that sediment below mussels hosted distinct microbial communities. However, our results may underestimate the impact of mussel beds on sediment community structure. Our sampling resulted in a range of 12 to 56 mussels per m^2 and unionid mussel abundances in this system can reach up to 100 mussels per m^2. Investigating mussel impacts on sediment communities when present at higher densities may demonstrate greater significance.

Interactions between the distinct microbial communities found within benthic microhabitats may be a driver of ecosystem function – Our data suggests that as a system, mussel shells, biodeposits, and the surrounding sediment contain microbial communities that work synergistically across microhabitats to cycle sulfur and nitrogen in aquatic environments. Interactions between mussels and the surrounding sediment is particularly relevant to ongoing investigations of the impact of freshwater mussels on ecosystem function (Vaughn and Hoellein, 2018). Studies on mussel-driven changes in nutrient cycling often focus either on the nutrients mussels cycle themselves (Atkinson and Vaughn, 2015; Trentman et al., 2018) or on ecosystem processes carried out by sediment microbial heterotrophs (Black and Just, 2018; Nickerson et al., 2019).

Across microhabitats, there are microbes that are important to sulfur cycling. Typically in anaerobic subsurface sediments, sulfate reducing bacteria (SRB) use sulfate as a terminal

electron acceptor instead of oxygen, resulting in sulfide compounds (Hansen, 1994). We found multiple SRB taxa within the family Thermodesulfovibrionaceae and the genus *Desulfococcus* that distinguish sediment from other microhabitats (Table 2; Galushko and Kuever, 2019; Umezawa et al., 2021). Then, much of the sulfide formed by SRB in sediments is oxidized back to sulfate by sulfur-oxidizing bacteria (SOB; Jørgensen and Nelson, 2004). Among the bacteria that distinguish shell bacterial communities from other microhabitats, we find SOB in the genera *Novosphingobium* and *Rhodobacter* (Table 2; Imhoff et al., 1984; Imhoff, 2015; Haosagul et al., 2020). We also see the family Comamonadaceae associated more strongly with mussel biodeposits than other microhabitats and this family has also been shown to oxidize sulfur compounds (Zhang et al., 2017). While we detect evidence of both SRB and SOB within the sediment, these SOB distinguish mussel communities from those of the sediment and their differential abundances may be facilitated by mussel activity. It is not unusual in marine sediments for sulfur bacteria to form symbiotic relationships with bivalves that depend on their primary production (Vaughn and Hoellein, 2018). If mussel activity expands the oxic-anoxic interface and delivers SOB to anoxic regions to respire sulfide produced by SRB, this may be an additional and critical mechanism by which mussels influence nutrient dynamics and primary production in freshwater environments. The presence of both SOB and SRB in these environments may prove especially interesting considering the interactions between sulfide produced by SRB and nitrogen removal in streams.

One of the primary mechanisms by which excess nitrogen (N) is removed as N_2 gas from aquatic ecosystems rather than assimilated into microbial and algal biomass, is through dissimilatory N respiration by microbes in the sediment. However, under laboratory conditions, sulfide has been demonstrated to inhibit enzyme pathways required for both nitrification (Joye

and Hollibaugh, 1995) and denitrification (Brunet and Garcia-Gil, 1996; Burgin and Hamilton, 2007). Yet, mussels have been shown to increase the potential for nitrogen removal in sediments proven to have relatively high amounts of sulfur deposition (Newton et al., 2013; Nickerson et al., 2019). When adjusting the ratio of carbon and nitrogen in a bioreactor, dos Santos et al. (2021) found that microbial communities with significantly higher N removal demonstrated increased abundances of six families of bacteria: Saprospiraceae, Chitinophagaceae, Xanthomonadaceae, Comamonadaceae, Bacillaceae, and Planctomycetaceae. We found five of these six families associated with different microhabitats across mussel beds (Tables 2-6) indicating that it may not simply be mussel driven changes in sediment communities that alter nutrient cycling in these systems, but rather interactions between mussel-associated and environmental microbes.

Nitrification and sulfide oxidation occur in oxic sediments while denitrification and sulfate reduction occur in anoxic environments. Mussels often traverse the boundary between oxic and anoxic layers of the sediment and so it is not surprising that we find microbes that encompass these diverse modes of respiration associated with them in the sediment. However, these modes of respiration are complementary and we find key players in these cycles across both mussels and the sediment. These patterns suggest that the role of burrowing bivalves in facilitating interactions between these microbes may be underestimated. Further, delivery of both S contaminants and nitrates to anaerobic sediments can enhance both sulfide oxidation and denitrification (Cardoso et al., 1996) and sulfur cycling by microbes has been shown to account for a large portion of nitrate removal in streams, lakes, and wetlands (Burgin and Hamilton, 2008). Mussels may provide a niche for microbes with a wide array of respiratory functions that

in summation serve to remove contaminants of ecological concern (Turner and Rabalais, 2003; Burgin and Hamilton, 2007).

Conclusions – In our exploratory study of the microbiomes of mussels (shell and biodeposits) and the environment they inhabitat (sediment) we found that mussel microbiomes were less diverse than those of the sediment, mussel abundance was a significant predictor of sediment microbial community composition, and local habitat influenced microbial assemblage composition more than site spatial location along the river. Our findings indicate that rather than a continuous shift in beta diversity along the river, microbial communities further away from each other are more similar than communities next to each other. In our system, a regional pool of bacterial taxa may thus be filtered by site-specific environmental conditions along the river continuum. Further, the presence of macro-organisms may be an additional mechanism that shapes the microbiome of benthic communities. Although synergistic communities of microbes are likely to persist in the sediment, interactions may be limited. Animal-facilitated interactions between freshwater microhabitats have implications for the removal of environmentally impactful metabolites such as nitrates and sulfides. We suggest more thorough testing of the impact of mussels and other burrowing organisms on microbial community diversity and function. Mussel aggregations may provide a niche for microbial communities that undoubtedly play a role in the complex cycling of multiple nutrients critical to primary production. Alterations to nitrogen and sulfide removal become especially important when considering anthropogenic inputs into freshwater systems.

Acknowledgements

We thank T. DuBose, N. Ferreira-Rodríguez, J. Hartwell, G. Hopper, J. Lopez and B. Tweedy for assistance with field work. We thank C. Siler for use of molecular facilities and K. Sankaranarayanan for bioinformatic pipeline and aid in analyses. We thank C. Atkinson, K. Sankaranarayanan, C. Siler, and L. Souza for their input to study design and comments on the manuscript. This paper is part of a dissertation at the University of Oklahoma and is a contribution to the program of the Oklahoma Biological Survey.

Literature cited

Allen, D. C., and Vaughn, C. C. (2009). Burrowing behavior of freshwater mussels in experimentally manipulated communities. *Journal of the North American Benthological Society* 28, 93–100. doi:10.1899/07-170.1.

Atkinson, C. L., Julian, J. P., and Vaughn, C. C. (2014). Species and function lost: Role of drought in structuring stream communities. *Biological Conservation* 176, 30–38. doi:10.1016/j.biocon.2014.04.029.

Atkinson, C. L., and Vaughn, C. C. (2015). Biogeochemical hotspots: Temporal and spatial scaling of the impact of freshwater mussels on ecosystem function. *Freshwater Biology* 60, 563–574. doi:10.1111/fwb.12498.

Atkinson, C. L., Vaughn, C. C., Forshay, K. J., and Cooper, J. T. (2013). Aggregated filter-feeding consumers alter nutrient limitation: Consequences for ecosystem and community dynamics. *Ecology* 94, 1359–1369. doi:10.1890/12-1531.1.

Averill, C., Werbin, Z. R., Atherton, K. F., Bhatnagar, J. M., and Dietze, M. C. (2021). Soil microbiome predictability increases with spatial and taxonomic scale. *Nat Ecol Evol* 5, 747–756. doi:10.1038/s41559-021-01445-9.

Berger, W. H., and Parker, F. L. (1970). Diversity of Planktonic Foraminifera in Deep-Sea Sediments. *Science* 168, 1345–1347. doi:10.1126/science.168.3937.1345.

Black, E. M., and Just, C. L. (2018). The Genomic Potentials of NOB and Comammox Nitrospira in River Sediment Are Impacted by Native Freshwater Mussels. *Frontiers in Microbiology* 9, 2061. doi:10.3389/fmicb.2018.02061.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37, 852–857. doi:10.1038/s41587-019-0209-9.

Brune, A., Frenzel, P., and Cypionka, H. (2000). Life at the oxic-anoxic interface: Microbial activities and adaptations. *FEMS Microbiology Reviews* 24, 691–710. doi:10.1016/S0168-6445(00)00054-1.

Brunet, R. C., and Garcia-Gil, L. J. (1996). Sulfide-induced dissimilatory nitrate reduction to ammonia in anaerobic freshwater sediments. *FEMS Microbiology Ecology* 21, 131–138. doi:10.1016/0168-6496(96)00051-7.

Burgin, A. J., and Hamilton, S. K. (2007). Have we overemphasized the role of denitrification in aquatic ecosystems? A review of nitrate removal pathways. *Frontiers in Ecology and the Environment* 5, 89–96. doi:10.1890/1540-9295(2007)5[89:HWOTRO]2.0.CO;2.

Burgin, A. J., and Hamilton, S. K. (2008). NO3−-Driven SO42− Production in Freshwater Ecosystems: Implications for N and S Cycling. *Ecosystems* 11, 908–922. doi:10.1007/s10021-008-9169-5.

Byllaardt, J. V., and Ackerman, J. D. (2014). Hydrodynamic habitat influences suspension feeding by unionid mussels in freshwater ecosystems. *Freshwater Biology* 59, 1187–1196. doi:10.1111/fwb.12339.

Cardoso, R. B., Sierra-Alvarez, R., Rowlette, P., Razo Flores, E., Go'mez, J., and Field, J. A. (1996). Sulfide Oxidation Under Chemolithoautotrophic Denitrifying Conditions. *Journal of Anatomy* 189 (Pt 3, 503–505. doi:10.1002/bit.

Choo, J. M., Leong, L. E., and Rogers, G. B. (2015). Sample storage conditions significantly influence faecal microbiome profiles. *Sci Rep* 5, 16350. doi:10.1038/srep16350.

Comrey, A. L., and Lee, H. B. (1992). *A first course in factor analysis, 2nd ed.* Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Curry, C. M., and Patten, M. A. (2014). Current and Historical Extent of Phenotypic Variation in the Tufted and Black-crested Titmouse (Paridae) Hybrid Zone in the Southern Great Plains. *The American Midland Naturalist* 171, 271–300. doi:10.1674/0003-0031-171.2.271.

dos Santos, C. E. D., Costa, R. B., Rabelo, C. A. B. S., Ferraz Júnior, A. D. N., Persinoti, G. F., Pozzi, E., et al. (2021). Hacking biofilm developed in a structured-bed reactor (SBRRIA) with integrated processes of nitrogen and organic matter removal. *Bioprocess Biosyst Eng* 44, 1841–1851. doi:10.1007/s00449-021-02564-0.

Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* 10, 996–998. doi:10.1038/nmeth.2604.

Eldridge, D. J., Woodhouse, J. N., Curlevski, N. J. A., Hayward, M., Brown, M. V., and Neilan, B. A. (2015). Soil-foraging animals alter the composition and co-occurrence of microbial communities in a desert shrubland. *ISME Journal* 9, 2671–2681. doi:10.1038/ismej.2015.70.

Filzmoser, P., Hron, K., and Reimann, C. (2010). The bivariate statistical analysis of environmental (compositional) data. *Science of The Total Environment* 408, 4230–4238. doi:10.1016/j.scitotenv.2010.05.011.

Fitzpatrick, C. R., Copeland, J., Wang, P. W., Guttman, D. S., Kotanen, P. M., and Johnson, M. T. J. (2018). Assembly and ecological function of the root microbiome across angiosperm plant species. *Proc Natl Acad Sci USA* 115, E1157–E1165. doi:10.1073/pnas.1717617115.

Fukami, T., Wardle, D. A., Bellingham, P. J., Mulder, C. P. H., Towns, D. R., Yeates, G. W., et al. (2006). Above- and below-ground impacts of introduced predators in seabird-dominated island ecosystems. *Ecology Letters* 9, 1299–1307. doi:10.1111/j.1461-0248.2006.00983.x.

Galushko, A., and Kuever, J. (2019). "Desulfococcus," in *Bergey's Manual of Systematics of Archaea and Bacteria* (American Cancer Society), 1–5. doi:10.1002/9781118960608.gbm01016.pub2.

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., and Egozcue, J. J. (2017). Microbiome datasets are compositional: And this is not optional. *Frontiers in Microbiology* 8, 1–6. doi:10.3389/fmicb.2017.02224.

Graham, E. B., Knelman, J. E., Schindlbacher, A., Siciliano, S., Breulmann, M., Yannarell, A., et al. (2016). Microbes as engines of ecosystem function: When does community structure enhance predictions of ecosystem processes? *Frontiers in Microbiology* 7, 1–10. doi:10.3389/fmicb.2016.00214.

Haag, W. R. (2012). *North American freshwater mussels: natural history, ecology, and conservation*. Available at: https://www.srs.fs.usda.gov/pubs/49874 [Accessed October 14, 2019].

Hansen, T. A. (1994). Metabolism of sulfate-reducing prokaryotes. *Antonie van Leeuwenhoek* 66, 165–185. doi:10.1007/BF00871638.

Haosagul, S., Prommeenate, P., Hobbs, G., and Pisutpaisal, N. (2020). Sulfur-oxidizing bacteria in full-scale biogas cleanup system of ethanol industry. *Renewable Energy* 150, 965–972. doi:10.1016/j.renene.2019.11.140.

Héry, M., Singer, A. C., Kumaresan, D., Bodrossy, L., Stralis-Pavese, N., Prosser, J. I., et al. (2008). Effect of earthworms on the community structure of active methanotrophic bacteria in a landfill cover soil. *ISME Journal* 2, 92–104. doi:10.1038/ismej.2007.66.

Hoellein, T. J., Zarnoch, C. B., Bruesewitz, D. A., and DeMartini, J. (2017). Contributions of freshwater mussels (Unionidae) to nutrient cycling in an urban river: filtration, recycling, storage, and removal. *Biogeochemistry* 135, 307–324. doi:10.1007/s10533-017-0376-z.

Hopper, G. W., DuBose, T. P., Gido, K. B., and Vaughn, C. C. (2019). Freshwater mussels alter fish distributions through habitat modifications at fine spatial scales. *Freshwater Science*, 000–000. doi:10.1086/705666.

Hopper, G. W., Gido, K. B., Vaughn, C. C., Parr, T. B., Popejoy, T. G., Atkinson, C. L., et al. (2018). Biomass distribution of fishes and mussels mediates spatial and temporal heterogeneity in nutrient cycling in streams. *Oecologia* 188, 1133–1144. doi:10.1007/s00442-018-4277-1.

Horton, D. J., Theis, K. R., Uzarski, D. G., and Learman, D. R. (2019). Microbial community structure and microbial networks correspond to nutrient gradients within coastal wetlands of the Laurentian Great Lakes. *FEMS Microbiology Ecology* 95. doi:10.1093/femsec/fiz033.

Imhoff, J. F. (2015). Rhodobacter. *Bergey's Manual of Systematics of Archaea and Bacteria*, 1–12. doi:10.1002/9781118960608.gbm00862.

Imhoff, J. F., TRuPER, H. G., and Pfennig, N. (1984). Rearrangement of the Species and Genera of the Phototrophic "Purple Nonsulfur Bacteria." *International Journal of Systematic Bacteriology* 34, 340–343. doi:10.1099/00207713-34-3-340.

Jørgensen, B. B., and Nelson, D. C. (2004). Sulfide oxidation in marine sediments: Geochemistry meets microbiology. *Sulfur Biogeochemistry – Past and Present*, 63–81.

Joye, S. B., and Hollibaugh, J. T. (1995). Influence of sulfide inhibition of nitrification on nitrogen regeneration in sediments. *Science* 270, 623–625. doi:10.1126/science.270.5236.623.

Jury, W. A., and Vaux, H. (2005). The role of science in solving the world's emerging water problems. *Proceedings of the National Academy of Sciences of the United States of America* 102, 15715–15720. doi:10.1073/pnas.0506467102.

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., and 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, 5112–5120. doi:10.1128/aem.01043-13.

Lindgreen, S. (2012). AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Research Notes* 5, 337. doi:10.1186/1756-0500-5-337.

Martinez, A. (2020). *pairwiseAdonis: Pairwise multilevel comparison using adonis*. Available at: https://github.com/pmartinezarbizu/pairwiseAdonis.

Matthews, W. J., Vaughn, C. C., Gido, K. B., and Marsh-Matthews, E. (2005). "Southern Plains Rivers," in *North American Rivers*, eds. A. Benke and C. Cushings (Elsevier), 283–325.

McArdle, B. H., and Anderson, M. J. (2001). Fitting Multivariate Models to Community Data: A Comment on Distance-Based Redundancy Analysis. *Ecology* 82, 290–297. doi:10.2307/2680104.

McCall, P. L., Matisoff, G., and Tevesz, M. J. S. (1986). The effects of a unionid bivalve on the physical, chemical, and microbial properties of cohesive sediments from Lake Erie. *American Journal of Science* 286, 127–159. doi:10.2475/ajs.286.2.127.

Miller-ter Kuile, A., Apigo, A., and Young, H. S. (2021). Effects of consumer surface sterilization on diet DNA metabarcoding data of terrestrial invertebrates in natural environments and feeding trials. *Ecology and Evolution* 11, 12025–12034. doi:10.1002/ece3.7968.

Mistry, R., and Ackerman, J. D. (2018). Flow, Flux, and Feeding in Freshwater Mussels. *Water Resources Research* 54, 7619–7630. doi:10.1029/2018WR023112.

Newton, T. J., Vaughn, C. C., Spooner, D. E., Nichols, S. J., Michael, T., and Arts, M. T. (2013). Profiles of Biochemical Tracers in Unionid Mussels Across a Broad Geographical Range. 32, 497–507. doi:10.2983/035.032.0229.

Nickerson, Z. L., Mortazavi, B., and Atkinson, C. L. (2019). Using functional traits to assess the influence of burrowing bivalves on nitrogen-removal in streams. *Biogeochemistry* 146, 125–143. doi:10.1007/s10533-019-00612-2.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2020). *vegan: Community Ecology Package*. Available at: https://CRAN.R-project.org/package=vegan.

Phillips, C. D., Phelan, G., Dowd, S. E., McDonough, M. M., Ferguson, A. W., Delton Hanson, J., et al. (2012). Microbiome analysis among bats describes influences of host phylogeny, life history, physiology and geography. *Molecular Ecology* 21, 2617–2627. doi:10.1111/j.1365-294X.2012.05568.x.

Pierce, M. L., Ward, J. E., Holohan, B. A., Zhao, X., and Hicks, R. E. (2016). The influence of site and season on the gut and pallial fluid microbial communities of the eastern oyster, Crassostrea virginica (Bivalvia, Ostreidae): community-level physiological profiling and genetic structure. *Hydrobiologia* 765, 97–113. doi:10.1007/s10750-015-2405-z.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41, D590–D596. doi:10.1093/nar/gks1219.

Quinn, T. P., Erb, I., Richardson, M. F., and Crowley, T. M. (2018). Understanding sequencing data as compositions: an outlook and review. *Bioinformatics* 34, 2870–2878. doi:10.1093/bioinformatics/bty175.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing Available at: https://www.R-project.org/.

Rubin, B. E. R., Gibbons, S. M., Kennedy, S., Hampton-Marcell, J., Owens, S., and Gilbert, J. A. (2013). Investigating the Impact of Storage Conditions on Microbial Community Composition in Soil Samples. *PLOS ONE* 8, e70460. doi:10.1371/journal.pone.0070460.

Sinsabaugh, R. L., Hill, B. H., and Follstad Shah, J. J. (2009). Ecoenzymatic stoichiometry of microbial organic nutrient acquisition in soil and sediment. *Nature* 462, 795–798. doi:10.1038/nature08632.

Skelton, J., Geyer, K. M., Lennon, J. T., Creed, R. P., and Brown, B. L. (2017). Multi-scale ecological filters shape the crayfish microbiome. *Symbiosis* 72, 159–170. doi:10.1007/s13199-016-0469-9.

Sliva, L., and Williams, D. D. (2005). Exploration of riffle-scale interactions between abiotic variables and microbial assemblages in the hyporheic zone. *Can. J. Fish. Aquat. Sci.* 62, 276–290. doi:10.1139/f04-190.

Spooner, D. E., and Vaughn, C. C. (2008). A trait-based approach to species' roles in stream ecosystems: Climate change, community structure, and material cycling. *Oecologia* 158, 307–317. doi:10.1007/s00442-008-1132-9.

Thoemmes, M. S., and Cove, M. V. (2020). Bacterial communities in the natural and supplemental nests of an endangered ecosystem engineer. *Ecosphere* 11, e03239. doi:10.1002/ecs2.3239.

Trentman, M. T., Atkinson, C. L., and Brant, J. D. (2018). Native freshwater mussel effects on nitrogen cycling: impacts of nutrient limitation and biomass dependency. *Freshwater Science* 37, 000–000. doi:10.1086/697293.

Turnbaugh, P. J., and Gordon, J. I. (2009). The core gut microbiome, energy balance and obesity. *Journal of Physiology* 587, 4153–4158. doi:10.1113/jphysiol.2009.174136.

Turner, R. E., and Rabalais, N. N. (2003). Linking Landscape and Water Quality in the Mississippi River Basin for 200 Years. *BioScience* 53, 563–572. doi:10.1641/0006-3568(2003)053[0563:LLAWQI]2.0.CO;2.

Umezawa, K., Kojima, H., Kato, Y., and Fukui, M. (2021). Dissulfurispira thermophila gen. nov., sp. nov., a thermophilic chemolithoautotroph growing by sulfur disproportionation, and proposal of novel taxa in the phylum Nitrospirota to reclassify the genus Thermodesulfovibrio. *Systematic and Applied Microbiology* 44, 126184. doi:10.1016/j.syapm.2021.126184.

van den Boogaart, K. G., and Tolosana-Delgado, R. (2008). "compositions": A unified R package to analyze compositional data. *Computers & Geosciences* 34, 320–338. doi:10.1016/j.cageo.2006.11.017.

Vaughn, C. C. (2010). Biodiversity Losses and Ecosystem Function in Freshwaters: Emerging Conclusions and Research Directions. *BioScience* 60, 25–35. doi:10.1525/bio.2010.60.1.7.

Vaughn, C. C., Atkinson, C. L. C. L., and Julian, J. P. J. P. (2015). Drought-induced changes in flow regimes lead to long-term losses in mussel-provided ecosystem services. *Ecology and Evolution* 5, 1291–1305. doi:10.1002/ece3.1442.

Vaughn, C. C., Gido, K. B., and Parr, T. B. (2021). Overlapping fish and mussel hotspots data archived in Open Science Framework. Available at: DOI 10.17605/OSF.IO/FWJH9.

Vaughn, C. C., and Hoellein, T. J. (2018). Bivalve Impacts in Freshwater and Marine Ecosystems. *Annual Review of Ecology, Evolution, and Systematics* 49, 183–208. doi:10.1146/annurev-ecolsys-110617-062703.

Vaughn, C. C., Nichols, S. J., and Spooner, D. E. (2008). Community and foodweb ecology of freshwater mussels. *Journal of the North American Benthological Society* 27, 409–423. doi:10.1899/07-058.1.

Vaughn, C. C., and Spooner, D. E. (2006). Unionid mussels influence macroinvertebrate assemblage structure in streams. *Journal of the North American Benthological Society* 25, 691–700. doi:10.1899/0887-3593(2006)25[691:UMIMAS]2.0.CO;2.

Vaughn, C. C., Taylor, C. M., & Eberhard, K. J. (1997). "A comparison of the effectiveness of timed searches vs. quadrat sampling in mussel surveys," in *Conservation and Management of Freshwater Mussels II: Initiatives for the Future*, 157–162.

Weingarten, E. A., Atkinson, C. L., and Jackson, C. R. (2019). The gut microbiome of freshwater Unionidae mussels is determined by host species and is selectively retained from filtered seston. *PLOS ONE* 14, e0224796. doi:10.1371/journal.pone.0224796.

Wright, S. P. (1992). Adjusted P-Values for Simultaneous Inference. *Biometrics* 48, 1005–1013. doi:10.2307/2532694.

Ye, L., Amberg, J., Chapman, D., Gaikowski, M., and Liu, W. T. (2014). Fish gut microbiota analysis differentiates physiology and behavior of invasive Asian carp and indigenous American fish. *ISME Journal* 8, 541–551. doi:10.1038/ismej.2013.181.

Zeglin, L. H. (2015). Stream microbial diversity in response to environmental changes: Review and synthesis of existing research. *Frontiers in Microbiology* 6, 1–15. doi:10.3389/fmicb.2015.00454.

Zhang, Y., Wang, X., Zhen, Y., Mi, T., He, H., and Yu, Z. (2017). Microbial Diversity and Community Structure of Sulfate-Reducing and Sulfur-Oxidizing Bacteria in Sediment Cores from the East China Sea. *Frontiers in Microbiology* 8, 2133. doi:10.3389/fmicb.2017.02133.

Tables

Table 1. Pairwise PERMANOVA results comparing microbial communities among microhabitats with Holm adjusted p values.

Comparison	Df	Sums Of Squares	F Model	R2	Adjusted p value
Mussel Biodeposit vs Mussel Shell	1	9027.225	17.38	0.32	**0.006**
Mussel Biodeposit vs Surface Sediment	1	6428.088	13.58	0.31	**0.006**
Mussel Biodeposit vs Subsurface Sediment	1	6710.488	13.93	0.32	**0.006**
Mussel Shell vs Surface Sediment	1	7352.131	14.43	0.33	**0.006**
Mussel Shell vs Subsurface Sediment	1	7736.168	14.90	0.35	**0.006**
Surface Sediment vs Subsurface Sediment	1	1649.57	3.64	0.15	**0.006**

Table 2. Axis loadings for PCA comparing sediment, shell, and biodeposit microbial communities. Higher r values correlated with axis 1 suggest taxa associated with sediment and lower values with mussel shells. Lower r values correlated with axis 2 will correlate with taxa differentiating mussel biodeposits from shell and sediment communities.

Family	Genus	Axis 1 r	Axis 2 r
Rhodobacteraceae	Rhodobacter	**-0.82**	0.06
Xanthomonadaceae		**-0.76**	0.34
Xenococcaceae		**-0.76**	0.42
Rhodobacteraceae		**-0.75**	0.34
Sphingomonadaceae	Novosphingobium	**-0.73**	0.14
Gemmataceae		**-0.72**	0.50
Synechococcaceae	Synechococcus	-0.49	**-0.71**
Synechococcaceae	Synechococcus	-0.45	**-0.77**
Verrucomicrobiaceae	Prosthecobacter	-0.41	**-0.77**
Synechococcaceae	Synechococcus	-0.41	**-0.80**
Synechococcaceae	Synechococcus	-0.39	**-0.81**
Chthoniobacteraceae	CandidatusXiphinematobacter	-0.37	**-0.84**
Planctomycetaceae	Planctomyces	-0.32	**-0.87**
Thermaceae	Meiothermus	-0.31	**0.71**
Pirellulaceae		-0.30	**-0.86**
Gemmatimonadaceae	Gemmatimonas	-0.29	**-0.72**
Chitinophagaceae	Sediminibacterium	-0.27	**-0.80**
Comamonadaceae		-0.25	**-0.72**
Chitinophagaceae		-0.23	**-0.88**
Pirellulaceae		-0.22	**-0.82**
Rhodocyclaceae		-0.19	**-0.86**
Fusobacteriaceae		-0.14	**-0.77**
Acetobacteraceae		-0.14	**-0.78**
Synechococcaceae	Synechococcus	-0.14	**-0.91**
Synechococcaceae	Synechococcus	-0.09	**-0.81**
Acetobacteraceae		-0.07	**-0.81**
Armatimonadaceae		-0.03	**-0.85**
Chitinophagaceae		-0.01	**-0.82**
Isosphaeraceae		0.01	**-0.73**
Bryobacteraceae		0.01	**-0.77**
Pirellulaceae		0.02	**-0.84**
Synechococcaceae	Synechococcus	0.06	**-0.83**
Caulobacteraceae	Phenylobacterium	0.06	**-0.74**
Desulfobacteraceae	Desulfococcus	**0.70**	-0.02
Thermodesulfovibrionaceae	HB118	**0.70**	0.09
Syntrophobacteraceae	Syntrophobacter	**0.71**	0.12
Myxococcaceae	Anaeromyxobacter	**0.71**	-0.02
Myxococcaceae	Anaeromyxobacter	**0.72**	0.03
Myxococcaceae	Anaeromyxobacter	**0.73**	0.07
Thermodesulfovibrionaceae	GOUTA19	**0.74**	-0.10
Desulfarculaceae		**0.76**	0.10
Thermodesulfovibrionaceae		**0.76**	0.17
Thermodesulfovibrionaceae	GOUTA19	**0.78**	0.08
Myxococcaceae	Anaeromyxobacter	**0.79**	0.19
Syntrophaceae		**0.89**	-0.03

Table 3. PERMANOVA outputs for sediment, shell, and biodeposit models. Statistically significant results in bold. Higher R2 values indicate more variance is explained by a given variable.

Model	Variable	R^2	F	p value
Sediment	Sediment Layer	0.15	4.49	**0.001**
	Site	0.15	2.23	**0.002**
	Mussel Abundance	0.05	1.6	**0.047**
	Mussel Richness	0.05	1.14	0.11
Shell	Site	0.31	3.7	**0.001**
	Mussel Species	0.05	1.22	0.19
Biodeposit	Site	0.31	3.9	**0.001**
	Mussel Species	0.047	1.19	0.22

Table 4. Axis loadings for PCA sediment microbial communities. Higher r values correlated with axis 1 suggest taxa associated with subsurface sediments while lower values suggest taxa associated with surface sediments. Lower r values correlated with axis 2 may suggest values associated with the site K1.

Family	Genus	Axis 1 r	Axis 2 r
Methanomassiliicoccaceae		**0.90**	0.31
Thermodesulfovibrionaceae	GOUTA19	**0.77**	-0.30
Thermodesulfovibrionaceae	LCP	**0.76**	-0.01
Xenococcaceae		**0.74**	-0.10
Syntrophaceae		**0.72**	-0.19
Isosphaeraceae		**0.72**	-0.32
Comamonadaceae		**-0.87**	0.08
Rhodocyclaceae		**-0.84**	0.07
Rhodocyclaceae	Dechloromonas	**-0.83**	0.13
Synechococcaceae	Synechococcus	**-0.79**	0.02
Desulfuromonadaceae		**-0.78**	-0.10
Geobacteraceae	Geobacter	**-0.78**	-0.27
Cytophagaceae		**-0.76**	0.23
Alcaligenaceae		**-0.76**	-0.23
Comamonadaceae		**-0.74**	-0.16
Chitinophagaceae		**-0.74**	0.01
Rhodocyclaceae		**-0.74**	0.05
Rhodobacteraceae	Rhodobacter	**-0.72**	-0.17
Cytophagaceae		**-0.72**	0.19
Sphingomonadaceae	Novosphingobium	**-0.70**	-0.40
Comamonadaceae		**-0.70**	-0.05
Myxococcaceae	Anaeromyxobacter	-0.07	**-0.79**
Crenotrichaceae	Crenothrix	-0.17	**-0.79**
Comamonadaceae		-0.08	**-0.70**

Table 5. Axis loadings for PCA biodeposit microbial communities. Higher r values correlated with axis 1 suggest taxa associated with the site K3 while lower values suggest taxa associated with K2 and K1.

Family	Genus	Axis 1 r	Axis 2 r
Synechococcaceae	Synechococcus	**-0.96**	-0.24
Synechococcaceae	Synechococcus	**-0.89**	-0.24
Synechococcaceae	Synechococcus	**-0.86**	-0.21
Fusobacteriaceae		**-0.85**	-0.21
Rhodospirillaceae		**-0.82**	-0.40
Synechococcaceae	Synechococcus	**-0.82**	-0.34
Armatimonadaceae		**-0.77**	-0.41
Enterobacteriaceae		**-0.74**	-0.13
Synechococcaceae	Synechococcus	**-0.73**	-0.41
Chitinophagaceae		**-0.73**	-0.29
Acetobacteraceae		**-0.73**	-0.38
Sinobacteraceae		**-0.71**	0.28
Pirellulaceae		-0.34	**-0.80**
auto67_4W		-0.25	**0.73**
Pirellulaceae		-0.19	**-0.78**
Syntrophaceae		-0.16	**0.84**
Desulfobacteraceae	Desulfococcus	-0.10	**0.73**
Caldilineaceae		-0.04	**-0.70**
Sinobacteraceae	Steroidobacter	0.16	**-0.76**
Thermodesulfovibrionaceae		0.24	**0.86**
Desulfobacteraceae	Desulfococcus	**0.72**	-0.08
Chthoniobacteraceae	CandidatusXiphinematobacter	**0.87**	-0.05
Nostocaceae	Dolichospermum	**0.89**	-0.21
Chthoniobacteraceae	CandidatusXiphinematobacter	**0.94**	-0.15

Table 6. Axis loadings for PCA comparing shell microbial communities. Higher r values correlated with axis 1 suggest taxa distinguishing the sites K1 and K2 from K3 while lower indicates the opposite. Higher r values correlated with axis two suggest taxa distinguishing sites K2 from K3 and K1.

Family	Genus	Axis 1 r	Axis 2 r
Sphingomonadaceae		**-0.83**	-0.19
Thermaceae	Meiothermus	**-0.81**	0.00
Kouleothrixaceae	Kouleothrix	**-0.77**	0.20
Chitinophagaceae		**-0.76**	0.18
Planctomycetaceae	Planctomyces	**-0.76**	-0.19
Sphingomonadaceae	Kaistobacter	**-0.74**	0.28
Hyphomicrobiaceae		**-0.72**	0.01
Syntrophobacteraceae		-0.43	**0.73**
Chitinophagaceae		-0.26	**0.73**
Sinobacteraceae	Steroidobacter	-0.25	**0.76**
Syntrophaceae		-0.08	**-0.70**
Hyphomicrobiaceae	Hyphomicrobium	**0.70**	-0.15
Rivulariaceae	Calothrix	**0.71**	-0.18
Synechococcaceae	Synechococcus	**0.72**	0.34
Synechococcaceae	Synechococcus	**0.73**	0.24
Synechococcaceae	Synechococcus	**0.73**	0.34
Clostridiaceae		**0.74**	0.03
Sinobacteraceae		**0.75**	0.13
Synechococcaceae	Synechococcus	**0.76**	0.25
Nostocaceae	Dolichospermum	**0.76**	0.00
Hyphomicrobiaceae	Hyphomicrobium	**0.76**	0.08
Caldilineaceae	Caldilinea	**0.77**	-0.31
Comamonadaceae		**0.77**	0.38
Chthoniobacteraceae	CandidatusXiphinematobacter	**0.77**	0.08
Acetobacteraceae		**0.78**	-0.10
Synechococcaceae	Paulinella	**0.79**	0.35
Hyphomicrobiaceae		**0.79**	-0.02
Nostocaceae		**0.81**	-0.18
Caldilineaceae		**0.85**	-0.17
Acetobacteraceae		**0.86**	-0.07
Sinobacteraceae		**0.88**	0.25
Oscillochloridaceae	Oscillochloris	**0.97**	0.00

Table 7. Wolman pebble counts, substratum heterogeneity (D60/D10), flow measurements, and canopy cover collected in 2016, from Vaughn et al. 2021.

	K1	K2	K3
D10	0.13	22.1	0.18
D50	34.17	86.9	22.78
D60	51.25	115.33	28.09
D90	112	225	58
D60/D10	861.54	5.22	156.06
Average Discharge (m^3/s)	0.25	0.33	0.34
Average Turnover (s)	58.62	15.47	105.69
Average Canopy Cover (%)	19.83	16.37	27.36

Figure legends

Figure 1. Map of the Kiamichi River drainage in Southeastern Oklahoma.

Figure 2. Principal coordinates analysis including samples from all microhabitats. Ellipses are drawn around centroids with a 50% confidence interval.

Figure 3. Principal coordinates analysis of (A) mussel biodeposit microbial communities, (B) mussel shell microbial communities, and (C) sediment microbial communities. Communities are colored based on the site from which they were sampled. Ellipses are drawn around centroids with a 50% confidence interval. Vectors were added *post hoc* and are based on *envfit* analysis.

Figure 4. Principal coordinates analysis of (A) mussel biodeposit microbial communities, (B) mussel shell microbial communities, and (C) sediment microbial communities. Communities are colored based on the site from which they were sampled. Ellipses are drawn around centroids with a 50% confidence interval. Vectors were added post hoc and are based on envfit analysis.

Figures

Figure 1.

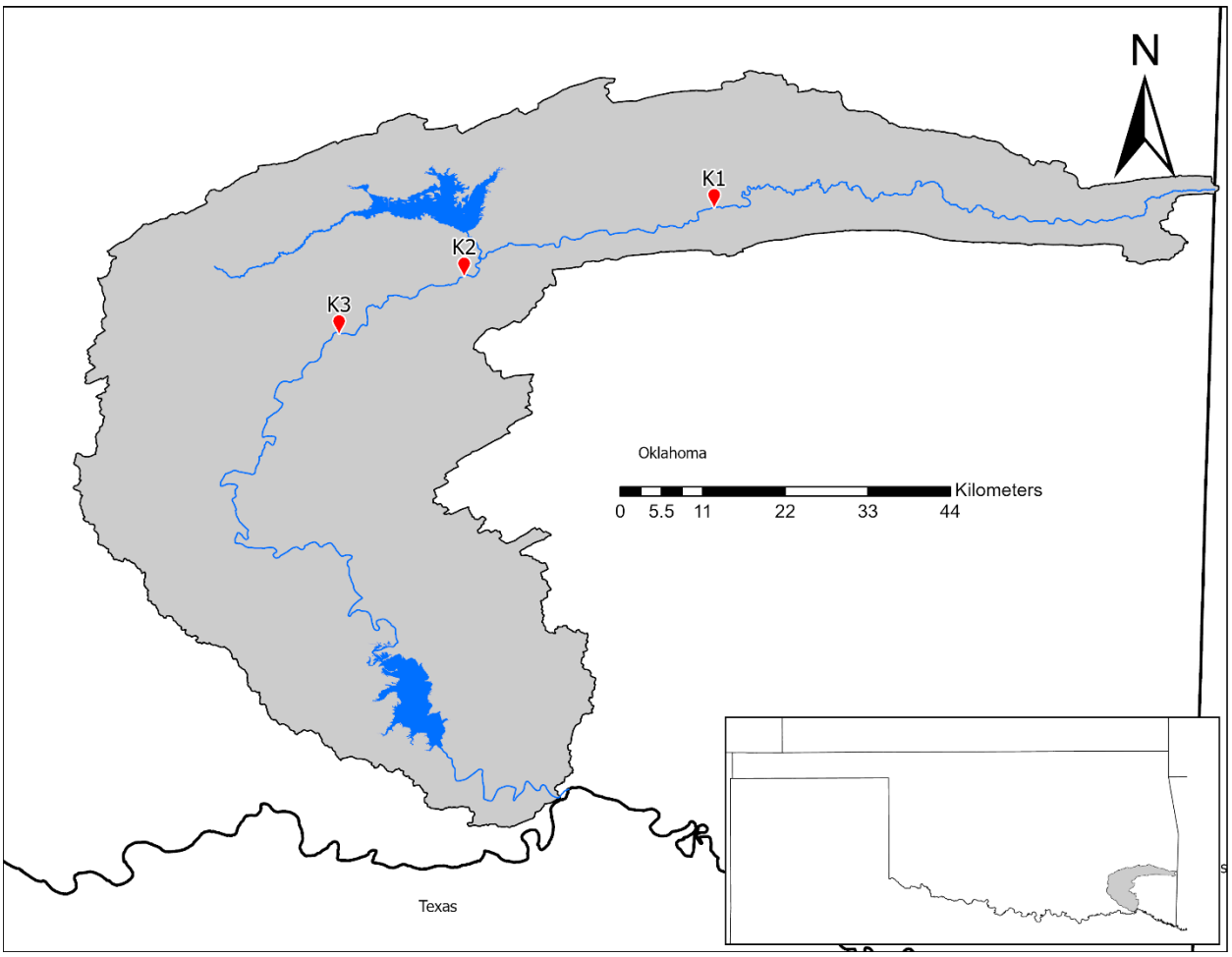


Figure 2.

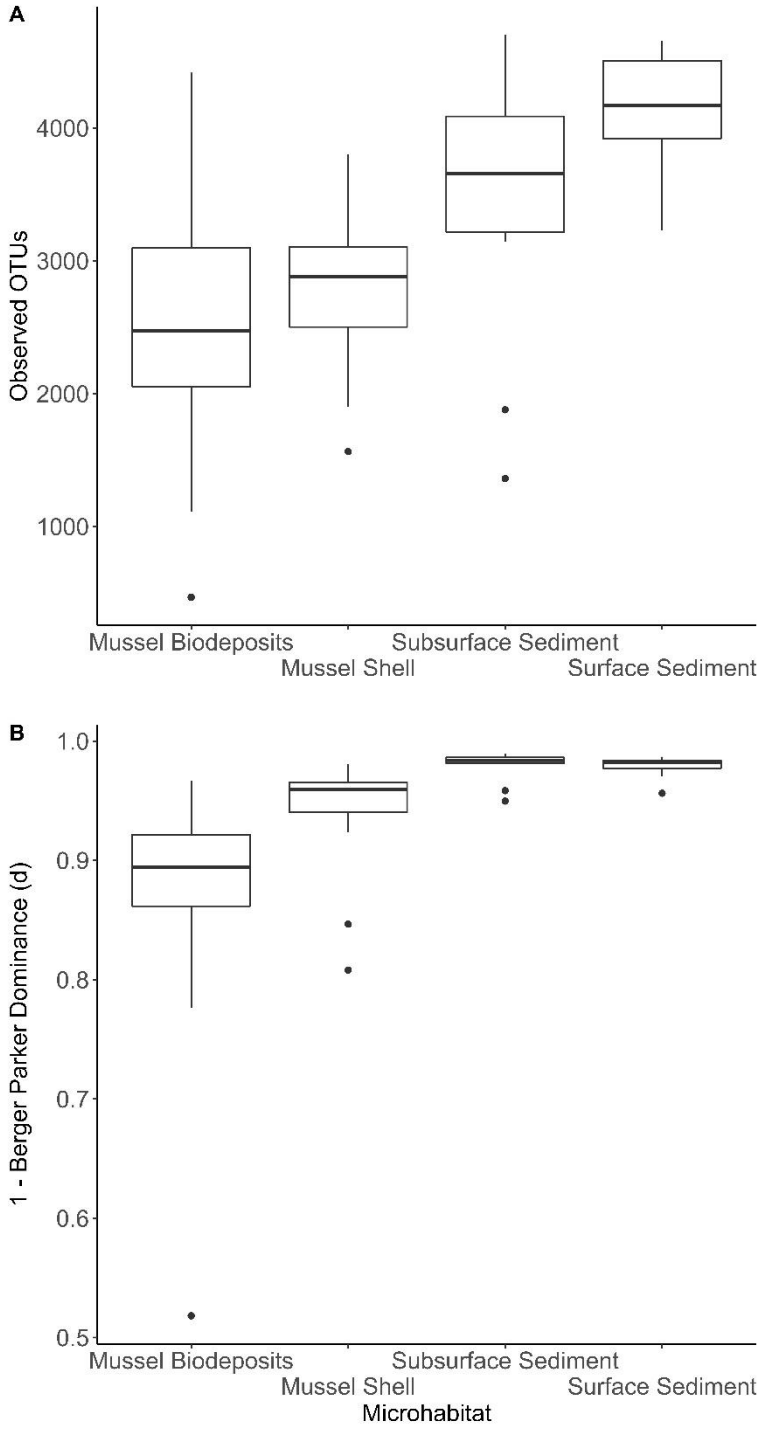

A
Observed OTUs
4000
3000
2000
1000
Mussel Biodeposits
Mussel Shell
Subsurface Sediment
Surface Sediment
B
1 - Berger Parker Dominance (d)
1.0
0.9
0.8
0.7
0.6
0.5
Mussel Biodeposits
Mussel Shell
Subsurface Sediment
Surface Sediment
Microhabitat

Figure 3.

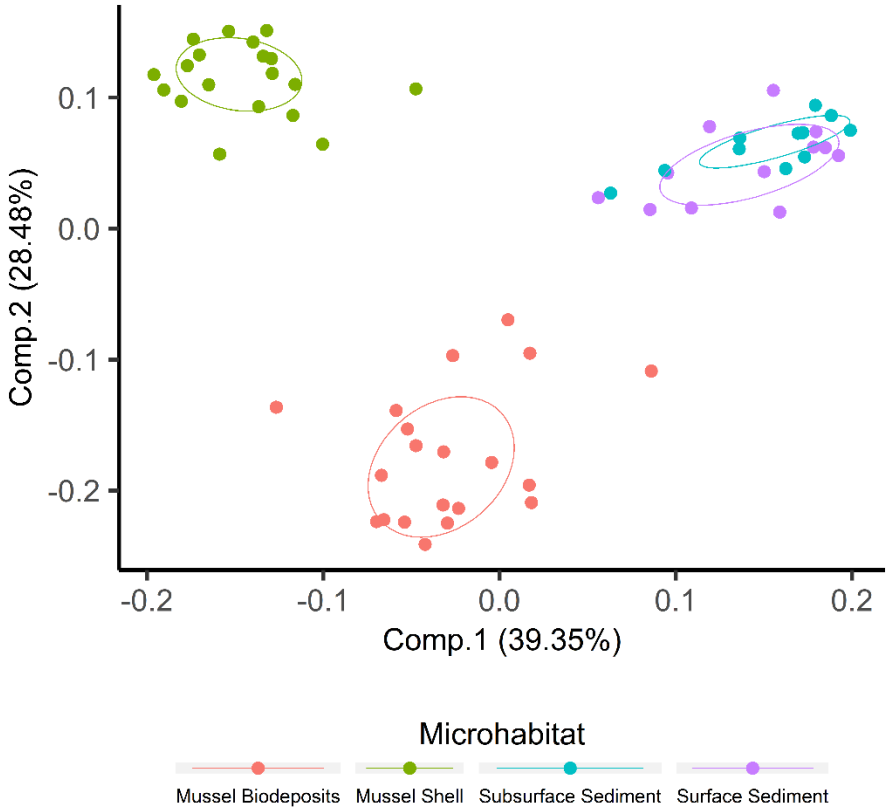
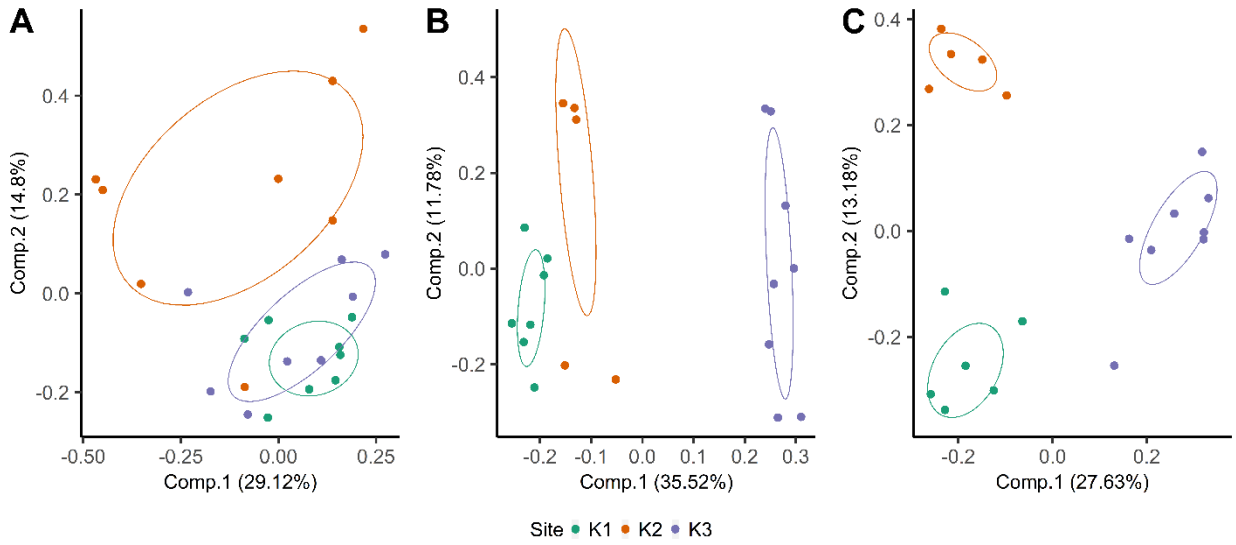

0.1
0.0
-0.1
-0.2
Comp.2 (28.48%)
-0.2
-0.1
0.0
0.1
0.2
Comp.1 (39.35%)
Microhabitat
Mussel Biodeposits
Mussel Shell
Subsurface Sediment
Surface Sediment

Figure 4.


A
B
C
Comp.2 (14.8%)
Comp.1 (29.12%)
Comp.2 (11.78%)
Comp.1 (35.52%)
Comp.2 (13.18%)
Comp.1 (27.63%)
Site K1 K2 K3

CHAPTER TWO

Freshwater mussels promote functional redundancy in sediment microbial communities under different nutrient regimes

Keywords:

Rivers, nitrogen, phosphorus, ecoenzyme, biogeochemical transformations

Formatted for submission to *Functional Ecology*

Authors: Edward Higgins, Thomas B. Parr, Caryn C. Vaughn

Abstract

1. Animals can have large impacts on how ecosystems function, from influencing population dynamics of other plant or animal species, to modifying landscape scale hydrogeological flow dynamics. One ecosystem function that has received widespread attention is the biogeochemical transformation of key nutrients required for primary production.
2. Freshwater mussels are benthic species with ecologically significant effects on nutrient cycling. Dense, high biomass mussel communities act as biogeochemical hotspots with landscape-level impacts. Here we investigated if mussel influences on biogeochemical cycling in stream sediment are accompanied by changes in sediment microbial community composition and functions and if these relationships change under different nutrient regimes.
3. We transplanted freshwater mussels and natural river sediment to flow-through mesocosms and monitored changes in microbial community composition over one week. On the final day we measured the activity of ecoenzymes known to correlate to microbial function and nutrient availability.
4. We predicted that the observed mussel effects on ecosystem function are caused by modified microbial communities. We hypothesized that if changes in sediment ecoenzymatic function are driven by mussel-derived nutrient amendments, we should see muted changes in microbial community assemblage or function when a given nutrient is not limiting. However, if microbial community and function are influenced by other mussel functions, then we should see uniform changes regardless of nutrient availability.

5. Our results indicate that mussels always changed sediment microbial community composition but the way communities changed was dependent on ambient nutrient concentrations. Further, mussels homogenized the stoichiometric ratios of ecoenzyme activities, indicating a consistent function of sediment microbes associated with freshwater mussels. Our results suggest that mussels may promote functional redundancy in sediment microbial communities and highlight the importance of animals in controlling biogeochemical transformations under changing nutrient conditions.

1 | INTRODUCTION

Although the term varies – keystone species, flagship species, ecosystem engineer, etc. – some organisms have disproportionate effects on ecosystem-scale processes and interspecies interactions. To date, terrestrial mammals have received a great deal of attention and as a result, there are plentiful, well-characterized examples (Coggan et al., 2018). Examples include beavers that alter river flow and change understory plant diversity (Brazier et al., 2021); deer that can reduce recruitment of hardwood saplings and influence population dynamics of forest-breeding birds (Baiser et al., 2008; Rushing et al., 2020; Shelton et al., 2014); and elephants that increase water availability by digging and provide understory habitat by uprooting trees (Coverdale et al., 2016; Haynes, 2012). However, ecosystem engineers are not limited to terrestrial systems and mammals. Intertidal crabs in saltmarshes and estuaries burrow and impact sediment chemistry, enhance soil drainage and aeration, influence juvenile fish behavior, and even modify rodent foraging (Spivak, 2010). In freshwater ecosystems, grazing fish can uproot macrophytes and increase turbidity, and crayfishes can change sediment patch dynamics and provide drought refugia for other species (Emery-Butcher et al., 2020).

Across ecosystems, a growing body of work demonstrates that animal physiology and behavior can shape the ecosystem level biogeochemical cycles of key nutrients (Atkinson et al., 2017; Schmitz et al., 2018; Vanni, 2002). While biogeochemical cycling is primarily carried out by microbial heterotrophs, animals can modify nutrient cycling both directly and indirectly. For example, terrestrial herbivores can alter nutrient dynamics in soil directly, through excretion or egestion, or indirectly through consumptive effects in which they alter plant-microbe interactions (Buchkowski & Schmitz, 2022; Sitters et al., 2017). Invasive marine worms can influence nitrogen fluxes in the sediment by aerating sediments by burrowing which impacts respiratory

pathways available to microbial heterotrophs (Tait et al., 2020). However, these interactions can come with caveats in which context dependency dictates the dominant microbial process. Bonaglia et al. (2014) found that while high meiofauna densities can increase denitrification in marine sediment, in the presence of bivalves, dissimilatory nitrate reduction to ammonium and methane efflux are significantly enhanced instead.

Bivalves in particular have many documented impacts on ecosystems (e.g., biofiltration, nutrient storage, habitat modification, stimulation of primary and secondary production, etc.), partially due to the economic value of marine species (Vaughn & Hoellein, 2018). Freshwater mussels (order Unionida) are bivalve ecosystem engineers that have strong, documented effects on nutrient cycling in streams (Vaughn & Hoellein, 2018). Unionid mussels are long-lived (life spans range from 6 to >100 years) sedentary filter-feeders that often occur in patchily distributed, multispecies aggregations in rivers called mussel beds (Haag, 2012). Dense, high biomass mussel beds act as biogeochemical hotspots with ecosystem-wide impacts (Atkinson et al., 2018; Atkinson & Vaughn, 2015). Mussels exhibit species-specific nutrient stoichiometry in both body tissue and excretion, and mussel communities can directly influence biogeochemical cycling of water column carbon and nutrients—both nutrient storage and nutrient remineralization—through their own physiology (Atkinson et al., 2020; Atkinson & Vaughn, 2015; Hopper et al., 2021; Parr et al., 2019). Research has also shown that mussels can have indirect impacts on the cycling of key nutrients by altering microbially mediated biogeochemical processes in the sediment, increasing nitrogen removal, likely through bioturbation or by alleviating nutrient limitation (Atkinson & Forshay, 2022; Trentman et al., 2018).

Recent work has also shown that the observed mussel influences on sediment microbial function may also be reflected in changes in the sediment microbial communities themselves.

The presence of unionid mussels has been linked to changes in sediment microbial community structure and metagenomic nitrogen cycling potential (Black et al., 2017; Black & Just, 2018). These mussel effects on sediment microbial communities appear to be context dependent in which local environmental conditions interact with mussels to shape microbial communities, as opposed to landscape scale factors or mussels alone (Higgins et al., 2022). However, while there is evidence for mussel-driven changes in sediment microbial function and sediment microbial community structure, it is not known if these changes happen concurrently, if mussel impacts on community structure or function are comparable across environmental gradients, and what aspect of mussel behavior or physiology drives these changes.

Here we asked, when mussels change sediment microbial function, is that due to altered microbial community structure or does the community persist, but alter the metabolic pathways by which they cycle nutrients? We hypothesized that if the observed mussel effects on ecosystem function are caused by modified microbial communities, we should see differences in sediment microbial community composition when mussels are present compared to when they are absent. Further, is the primary mechanism for altering microbially mediated nutrient cycling alleviating nutrient limitation or via mussel behavior e.g., bioturbation? If these changes in sediment function are driven by mussel excretion or egestion alleviating nutrient limitation, we should see muted changes in microbial community assemblage or function when a given nutrient is not limiting, but if they are driven by mussel bioturbation or other mussel activity, then we should see uniform changes regardless of nutrient availability. To address these questions we conducted a mesocosm experiment in which we manipulated mussel abundance and nutrient concentration over a seven day period and measured the response of the sediment microbial assemblage and

the activity of ecoenzymes commonly used to measure microbial nutrient assimilation and cycling (Sinsabaugh et al., 2009).

2 | MATERIALS AND METHODS

2.1 | Study organism and system

Mussels and sediment were collected from the Kiamichi River, a fifth order river in southeastern Oklahoma US. This relatively pristine river is known for its high aquatic diversity, including ~30 species of freshwater mussels (Matthews et al., 2005). Ecosystem effects of mussels in this river have been well-studied, especially their effects on nutrient cycling (Atkinson & Vaughn, 2015; Atkinson et al., 2018). For this study we used one freshwater mussel species, *Actinonaias ligamentina.* This species has high biomass in the river with documented strong effects on nutrient cycling as well as primary and secondary production (Spooner & Vaughn, 2006; Vaughn et al., 2007; Vaughn, 2010).

2.2 | Experimental design

We conducted the experiment in outdoor, flow-through mesocosms at the University of Oklahoma Aquatic Research Facility (ARF; 35°10'58.5"N 97°26'49.8"W). These mesocosms have been used successfully in other recent experiments with freshwater mussels (Parr et al., 2019; Vaughn et al., 2022). A small-scale pilot study in 2019 demonstrated that by the end of one week in a mesocosm, natural sediment microbial communities had significantly decreased in alpha diversity but this change subsequently plateaued. Thus, to give our treatments enough time to produce observable results, but not so much time that mesocosm effects obscured treatment effects, we ran the experiment for 7 days. We used 30 flow-through mesocosms which consisted of a small inflow (1.33 m long x 0.45 m wide, 0.2 m deep), a large circular pool filled with gravel (1.8 m diameter and 0.35 m deep), and a short outflow (0.71 m long x 0.45 m wide x 0.2

m deep; Parr et al., 2019). In each pool we buried one 33.02 cm x 33.02 cm x 27.94 cm enclosure to a depth of 20 cm (Fig. 1).

We had eight treatments; a mussel and a non-mussel treatment crossed with four nutrient amendments. Nutrient treatments were amendments of nitrogen (N), phosphorus (P), nitrogen and phosphorus (NP), and a no nutrient control. Each treatment was replicated four times except for the control which had three replicates. The mussel treatments contained four *A. ligamentina* individuals within the enclosures resulting in a density of ~36 individuals$\cdot m^{-2}$, a mussel density found in large mussel beds in the Kiamichi River (Vaughn & Spooner, 2006). Nutrient amendments were based on the background concentration of N and P in the groundwater supplied to the mesocosms (Parr et al., 2019). Groundwater nutrient concentrations were relatively stable and in a previous experiment, the average background concentration of SRP was ~27.2 $\mu g \cdot L^{-1}$ while the average concentration of N was ~1273 $\mu g \cdot L^{-1}$ as nitrate with much less N present as ammonium (~8.2 $\mu g \cdot L^{-1}$). The inflow rate of water to mesocosms was roughly 1 L/min resulting in a delivery of ~27.2 $\mu g\ P \cdot min^{-1}$ and ~8.2 $\mu g\ NH_4^+ \cdot min^{-1}$ to mesocosms. To our P and NP mesocosms, we targeted an amendment rate of monosodium phosphate (NaH_2PO_4) in solution at 54.4 $\mu g\ SRP \cdot min^{-1}$ to raise the influx of P three-fold. To mimic *A. ligamentina* excretion stoichiometric ratios of N:P, we calculated our N amendments to reach 20N:1P with most N as NH_4^+ (Spooner & Vaughn, 2008). Thus, we added ammonium chloride (NH_4Cl) in solution at 155.8 $\mu g \cdot min^{-1}$ to our N treatments to target an influx of N to 168 $\mu g\ N \cdot min^{-1}$ (20 times the influx of P in our P treatment). For our +NP treatments we then added NH_4Cl at 483.8 $\mu g\ N \cdot min^{-1}$ to adjust for the three-fold increase in P delivery. Nutrients were administered using IV drip bags manually adjusted to add 500 mL of solution over 24 hours. IV bags drip rates were measured every morning and readjusted as needed.

The day before the experiment, we began amending nutrients to the mesocosms. We also collected fresh river sediment from the Kiamichi River the day before the experiment as the river is a 3-hour drive from the research facility. We collected sediment with a 90% ethanol-rinsed, flame sterilized trowel into 90% ethanol-rinsed plastic buckets. Sediment was immediately transferred on shore to coolers with a layer of ice separated by 90% ethanol rinsed plastic liners. Sediment was transported back to the ARF within 5 hours and added to enclosures in random order to fill up to ~13cm depth. Mussels were collected in June and held in a recirculating tank until placement within enclosures after the first sampling effort. The experiment was initiated on July 31st, 2020 and ran until August 7th, 2020.

To ensure mussels were actively feeding and thus excreting and egesting, a commercial algal diet was added to all mesocosms on day two and five (algal density ≈ 7 million cells/mL; Instant Algae®: Shellfish Diet 1800, Reed Mariculture, Campbell, California (Blakeslee et al., 2013; Galbraith et al., 2015)). Commercial algae stocks were used rather than lab cultured algae to avoid the addition of the supplemental nutrients needed to grow algae in the laboratory. Instead, commercial algae feed comes as a concentrated stock suspended in clean water with known quantities of algal cells to ensure uniformity of distribution.

2.3 | Sampling

Sediment samples for microbial community composition analyses were collected on days 0, 3, and 7. Sediment was removed from a random location within the enclosure with a 70% ethanol sterilized, 4 cm diameter PVC sediment corer and an ethanol sterilized putty knife. Sediment particle size has been shown to select for different communities of microbes (Sliva & Williams, 2005) and so to get a more complete sampling of sediment microbial communities, sediment was sieved and two samples were taken as separate microhabitats: one of particle sizes

above 495 µm and one between 61 µm and 495 µm. Samples were stored in sterile cryovials and immediately placed in liquid nitrogen. On days 0 and 7, an additional sample was taken from each mesocosm between 61 µm and 495 µm particle sizes for ecoenzyme activity analysis. Ecoenzyme samples were stored on ice until they were transported to a refrigerator. Pandemic induced delays and restrictions at our institution did not allow for day 0 samples to be analyzed within one week of collection; thus these were not used. Day 7 ecoenzyme samples were analyzed within 8 hours of collection.

Algae grew on mussel shells and in the mesocosms during the experiment. To quantify algae on mussel shells, on day 7 after the final samples were collected, mussels were removed from mesocosms, and three independent observers ranked the algal growth on mussel shells from 1-5. To train observers, 20 mussels from a holding tank separate from the experiment were selected and sorted into scores of 1-5 based on consensus before ranking. To quantify algae in the mesocosms, on day 7 a photo was taken from the same angle above each mesocosm to quantify the amount of visible filamentous algae and periphyton that may not have been sampled from the water column following Parr et al. (2019). From these photos, four independent observers ranked the algal growth on a scale of 1-10. Observers were provided with examples of mesocosms with scores of 1, 5, and 10 ahead of scoring to ensure consistency (Parr et al., 2019).

2.4 | Sample processing

Microbiome library preparation and sequencing followed Higgins et al. (2022). In brief, we extracted DNA using DNeasy PowerSoil® kits (Qiagen, Hilden, Germany), amplified the v4 region of the 16s rRNA gene (Kozich et al., 2013), and sequenced with 2 × 250 bp paired-end sequencing on an Illumina MiSeq.

We quantified the activity of ecoenzymes often linked to important biogeochemical transformations of C, N, and P (Sinsabaugh et al., 2009; Sinsabaugh & Follstad Shah, 2012); β-Glucosidase (BG), β-N-acetylglucosaminidase (NAG), leucine-aminopeptidase (LAP), and phosphatase (AP). To quantify activity, we measured the change in fluorescence over time of 1-2 g of sediment samples exposed to 5 nM solutions of 4-methylumbelliferone (BG, NAG, and AP) or 7-amino-4-methyl coumarin (LAP) linked substrates using the methods and calculations outlined by Findlay & Parr (2017).

2.5 | Bioinformatics and data analysis

16S sequencing data were analyzed in QIIME2 (Bolyen et al., 2019). Sequences were denoised into ASVs with DADA2 (Callahan et al., 2016) and classified using the Silva 138 Reference Database (Quast et al., 2013). After removing samples that unsuccessfully sequenced, we had 157 remaining samples. One sample had an ASV count 5 standard deviations from the mean and was removed as an outlier for a final N of 156 samples total and 59 for the final day of the experiment. We calculated Shannon Diversity and Simpson's Index as our alpha diversity metrics on the resulting taxa table. To determine statistical differences in alpha diversity we used Kruskal-Wallis tests in the base R software (R Core Team, 2020). Significant results were further examined using Holm adjusted pairwise-Wilcoxon Rank Tests between microhabitats (Wright, 1992).

To quantify differences in beta diversity we first filtered ASVs with less than 0.1% of the average sequencing depth (Higgins et al., 2022). To account for the compositional nature of sequencing data, we calculated an Aitchison distance matrix in R using the Compositions and Vegan packages (Gloor et al., 2017; Oksanen et al., 2020; Quinn et al., 2018; van den Boogaart & Tolosana-Delgado, 2008). To visualize differences in community structure, we performed

principal coordinates analyses (PCA) using our distance matrices (Gloor et al., 2017). We used PERMANOVA tests, with samples nested within the mesocosm from which they were sampled, to determine differences in bacterial community structure among treatments with Vegan's '*adonis'* package. We conducted post-hoc, pairwise PERMANOVA to evaluate differences among mussel and sediment microhabitats using the '*pairwise.adonis*' test from the pairwiseAdonis package (Martinez, 2020; Wright, 1992). We included particle size in our PERMANOVA to test if mussels had a distinct effect on different particle sizes. We used permdisp to determine differences in dispersion using the '*betadisper'* functions. Because '*betadisper'* can only test singular variables at a time, we tested the effect of each variable in our PERMANOVA models individually and reported Holm adjusted p-values (Chen et al., 2017; Holm, 1979). For interaction effects in our permdisp tests, we concatenated the interacting variables into one dummy variable (Penha et al., 2017). To investigate the impacts of our experimental variables without mesocosm effects, and to better compare microbial community data to ecoenzyme data, we separated samples collected on day 7 and reanalyzed them separately. For these analyses, we removed ASVs that were not present in any sample from the resulting filtered taxa table.

2.6 | Ecoenzyme data analysis

Ecoenzyme fluorescence measurements were converted to activity measurements using the methods described in (Findlay & Parr, 2017). The stoichiometry of extracellular enzymes provides insight into the relative availability of limiting nutrients (Sinsabaugh & Follstad Shah, 2012). The ratios of β-glucosidase activity to that of phosphatase ($C:P_{EEA}$) and to the sum of β-N-acetylglucosaminidase and leucine-aminopeptidase activities ($C:N_{EEA}$) provides insight into C:P and C:N ratios limiting microbes, with similar implications for the ratio of nitrogen and

phosphorus enzymes ($N:P_{EEA}$). We expected that exudates from algae growing on mussels and in the water could possibly influence ecoenzyme activity in the sediment and so we tested if incorporating our average mesocosm algal score and mussel shell algal score measurements as random effects into mixed linear models produced singular fits, indicating the models were over-fit (Barr et al., 2013; Kuznetsova et al., 2017). If either variable produced a singular fit, they were not used. For both linear models and mixed linear models we used Type II ANOVA for unequal design to test the effect of our treatments on ecoenzyme activity in the sediment (Langsrud, 2003).

3 | RESULTS

3.1 | Treatment effects over time on microbial alpha and beta diversity

Both alpha and beta diversity of microbial communities changed with treatment and time. There was a statistically significant 9% reduction in Shannon diversity between the three sampling dates ($\chi^2 = 45.75$, $p < 0.0001$; Fig 2A) and pairwise comparisons show significant differences between the first day of the experiment and both the third (t ratio = 5.31, $p < 0.0001$) and final day (t ratio = 6.24, $p < 0.0001$), but not between the third and final day (t ratio = 0.46, $p = 0.643$). There was also a 49% reduction in the inverse Simpson's evenness index among the three sampling dates ($\chi^2 = 19.20$, $p < 0.0001$; Fig 2B) although the only statistically significant pairwise comparison was between the first and final sampling date (t ratio = 4.38, $p < 0.0001$; Fig. 2B). Our PERMANOVA model testing the effects of the experiment date and the interactions between mussel and nutrient treatments was an overall good fit to our data ($F = 3.05$, $p = 0.001$). Both the interaction between mussel and nutrient treatments ($F = 1.32$, $p = 0.001$) and the sampling date ($F = 8.67$, $p < 0.001$) were significant predictors of microbial community

beta diversity (Fig. 3). There was a significant increase in dispersion among communities over time ($F = 24.21$, $p < 0.001$).

3.2 | Treatment effects on sediment microbial community structure

By the seventh day of the experiment, although no Shannon diversity measurements were significantly different among treatments, nutrient treatments had a significant effect on microbial community evenness ($\chi^2 = 12.64$, $p = 0.006$; Fig. 4A). Post hoc tests demonstrate that this difference was driven by N (t ratio = 3.29, $p = 0.02$) and NP (t ratio = 2.82, $p = 0.05$) treatments with slightly reduced evenness than control treatments (Fig. 4B). Our model evaluating the effect of the interaction between nutrients and mussels as well as the interaction between mussels and particle size on microbial community structure was a significant fit to the data ($F = 1.45$, $p = 0.001$). PERMANOVA showed mussel and nutrient treatment interactions resulted in significantly distinct microbial communities ($F = 1.45$, $p = 0.001$; Fig. 5), but mussels did not interact differently with different sediment size classes to affect microbial communities ($F = 0.73$, $p = 0.65$). Posthoc comparisons show that in comparison to controls, nutrient and mussel treatment interactions were always a significant predictor of microbial community structure but for comparisons between nutrient treatments, there were no interaction effects and instead mussels were always significant predictors (Table 1). There were no differences in dispersion among nutrient treatments alone ($F = 0.88$, $p = 0.46$), mussel treatments alone, ($F = 0.48$, $p = 0.49$) or our nutrient-mussel variable ($F = 1.41$, $p = 0.22$).

3.3 | Treatment effects on sediment microbial community function

The average mesocosms algal score produced a singular fit and were removed from mixed models. Linear models demonstrated that mussels had a significant effect on $C:N_{EEA}$ and $C:P_{EEA}$ but not on the $N:P_{EEA}$ (Fig. 6A-C, Table 2). Mixed linear models controlling for algal

growth on mussel shells showed no significant effect of treatments or interactions on the activity of P acquiring enzymes or the summed activity N-acquiring enzymes, but there was a significant effect of nutrient treatment on the activity of carbon acquiring enzymes (Fig. 6D-F; Table 2).

4 | DISCUSSION

4.1 | Mussels influenced sediment microbial communities under strong environmental selection

We predicted that we would see differences in sediment microbial community composition when mussels were present compared to when they were absent. Mussels did impact microbial communities, despite the fact that these communities also changed when removed from their natural river environment. We observed significant decreases in both richness and evenness of the sediment microbial communities that stabilized by day three and sediment microbial community composition was distinct at each sampling date. Thus, while sampling date was a strong predictor of microbial community structure, mussel-driven changes to microbial communities were strong enough to persist despite these background changes.

By the end of the experiment, mussels influenced sediment microbial community composition, regardless of nutrient treatment, but not alpha diversity. Interestingly, previous work has shown that the addition of increasing concentrations of simple labile carbon and nutrients decreases bacterial richness and evenness (Van Horn et al., 2011). What we know of the effects of labile carbon addition on microbial community structure may be due to the preponderance of studies using monomers (e.g. acetate, glucose, etc.) as opposed to complex natural forms of labile carbon. While still highly labile, the diversity of molecules in the mussel biodeposits may be critical for maintaining community richness. Further, while other studies have found that mussels influence microbial community beta diversity (Black et al., 2017; Lukwambe et al., 2018), our data suggest that mussels not only influence microbial community structure, they do so under many nutrient regimes.

4.2| Mussel impacts on microbial community structure were context dependent based on concentration, not identity, of nutrients

We hypothesized that if mussel N or P amendments were the primary mechanism of microbial community change, there should be at least one nutrient treatment in which the effect of mussels on microbial community composition was not significant. Instead, mussel treatments were always significant. However, patterns in centroid location (representative of differences in community composition) and the dispersion of sample points (representative of the breadth of possible community composition) suggest that mussels changed microbial community composition differently in high and low nutrient environments.

By the final day of the experiment, there was a large turnover in microbial community composition in nutrient control tanks based on the presence of mussels alone. Interestingly, the sediment microbial communities with mussels in nutrient-control treatments were more similar to communities *without* mussels but with nutrients than to mussel-nutrient treatments (Fig. 5). This suggests that the sediment microbial communities in this study were sensitive to changes in the nutrient environment, but whether those nutrients came from animals or the environment, the resulting microbial communities were similar. This is in contrast to other studies that have found that in natural systems, microbial taxonomic diversity is resilient to external inputs of nutrients (Allison & Martiny, 2008; Bowen et al., 2009, 2011; Graves et al., 2016). However, by transporting natural sediment microbial communities to a nutrient poor environment, the amendment of any animal or ambient nutrient may have played a role in structuring microbial communities.

Although not significant, smaller microbial community dispersion seems to be retained between mussel and non-mussel treatments in nutrient control mesocosms, but mussels appear to

increase microbial community dispersion when nutrients were added. When biologically available nutrients are limited, it is possible that mussel-derived nutrients and energy shift communities towards a state more similar to when those nutrients are elevated in the environment. However, when elevated nutrients are already present in the water column, some other mechanism of mussel behavior increases beta diversity and expands the available niches to allow for a wider variety of possible community compositions. It is worth noting that with elevated P, the centroid of mussel treatment microbial communities was located within the 50% confidence interval of the non-mussel treatment suggesting a smaller change in community composition. It is possible that a larger sample size would highlight more nuanced differences in the impacts of mussels among different nutrient regimes.

Taken together, mussels always impacted sediment microbial community beta diversity, but the nature of that change was context dependent. Mussels in a low-nutrient environment shifted microbial communities to a community composition similar to those with higher nutrients but no mussels, but mussels in a high-nutrient environment seemed to increase the breadth of possible community compositions. Other investigations into ecosystem engineers demonstrate that not all engineers have strong effects on soil microbial communities and those that do, can have similar, beneficial effects across disturbance gradients (Berga et al., 2015; Eldridge et al., 2015, 2016). Identifying animals that play a role in mediating microbial communities across environmental gradients is important when considering the impacts of animal conservation on ecosystem function.

4.3 | Mussel impacts on sediment microbial function were not context dependent

Our data suggest that mussels play a significant role in homogenizing the balance of microbial carbon and nutrient assimilatory function. In all treatments the presence of mussels

either increased or decreased $C:N_{EEA}$ and $C:P_{EEA}$ to a consistent ratio across treatments. Elevated carbon to nutrient ratios can indicate either more investment in C acquiring enzymes, in which mussels supplied more labile C, or less investment in N or P acquiring enzymes, in which mussels delivered more bioavailable N or P (Sinsabaugh et al., 2009). β-Glucosidase is an assimilatory enzyme that has been shown to be released in the presence of high quality, phytoplankton exudates (Sieczko et al., 2015). Given this relationship, elevated activity of β-Glucosidase may be a result of mussel delivery of lysed phytoplankton and water-column algae through biodeposits. Spatial ecosystem ecology is a theory often framed with the idea of animal movement transporting important energy or nutrients between ecosystems (Schmitz et al., 2018), but in this instance, animals that simply exist and function at critical boundary layers may serve similar, important roles. Further, in the mesocosms without mussels but with nutrients, there was much more variability in the response of $C:N_{EEA}$ and $C:P_{EEA}$. This may suggest the important role that mussels and other burrowing animals play in stabilizing some ecosystem functions.

We hypothesized our $N:P_{EEA}$ results would reflect a stronger abiotic (sorption) control as compared to the hypothesized biotic controls. We expected that adding P would result in higher $N:P_{EEA}$ as phosphatase has been demonstrated to have a consistent, inverse relationship with bioavailable P (Sinsabaugh & Follstad Shah, 2012) but we instead saw depressed $N:P_{EEA}$ in non-mussel P mesocosms in comparison to non-mussel nutrient control mesocosms. Phosphorus availability is dictated more strongly by physical than biological processes and elevated levels of water column phosphorus and sedimentation of organic matter can increase adsorption to sediments in aquatic ecosystems (Allan & Castillo, 2007; Boström et al., 1988; Yang et al., 2019). If the combination of elevated water column P and autochthonous carbon (from commercial algae added to each tank) increased adsorption to surface sediments, there may have

been less biologically available P to the subsurface microbial heterotrophs and altered ambient nutrients enough to shift heterotrophic communities to become P-limited. This is further supported by the depressed activity of β-Glucosidase in phosphorus tanks, in which P-limitation may have limited microbial growth over time, reducing the need for carbon assimilation. If mussels delivered bio-available P past surface sediments, this may have shifted the system to a co-limitation of N and P, explaining the observed increase in $N:P_{EEA}$ in mussel tanks.

While we acknowledge that $N:P_{EEA}$ among treatments did not produce significant results, we suspect that the combination of an unexpected result in phosphorus mesocosms and small sample sizes may have played a role in the lack of significance of the effects of our treatments. Mussel treatments demonstrated a consistent $N:P_{EEA}$ across treatments, while control and P non-mussel treatments produced stratified $N:P_{EEA}$. This opposite, inverse result may have obscured the impacts of both mussels and nutrients in our models. If the interaction effect involved too many comparisons given our small sample size, this may explain the lack of observed significance despite a consistent pattern across treatments. Future studies should investigate this interaction as our sampling effort seems to have revealed an interesting pattern.

4.4 | Mussels contribute to sediment microbial community functional redundancy

We hypothesized that based on the mechanism by which mussels may influence microbial communities, there should be either a predictable pattern across nutrient treatments or the same result in each treatment. However, neither of these scenarios was supported by our community or ecoenzyme data. Despite the abundance of studies investigating microbial diversity in natural systems, conflicting results can arise with regards to microbial diversity and function. Recent research has shown that microbial diversity is tightly linked to chemodiversity in dissolved organic matter in soil and estuaries (Chen et al., 2021; Li et al., 2018), yet other

studies have pointed to microbial community resilience to change with excess nutrients (Bowen et al., 2009, 2011). Our study contributes to the growing body of work demonstrating that microbial taxonomic diversity is not a strong predictor of microbial function. Graves et al. (2016) found that bacterial communities can amplify a single functional gene across a wide range of functionally redundant taxa, resulting in taxonomically similar yet functionally distinct bacterial communities. Here, we found the opposite, where mussel presence resulted in taxonomically distinct microbial communities that were functionally similar. If the primary mechanism of mussel influence was providing a diversity of nutrients and energy, instead of providing key nutrients like N and P, this may better explain our results.

Mussels excrete forms of dissolved organic matter distinct from ambient sources (Hopper et al., 2021). The complexity of resources and abundance of carbon provided by mussels may ultimately promote functional redundancy in sediment communities and cause a stabilization of microbial decomposition and ecoenzymatic stoichiometry despite ambient nutrients. Some bacteria can specialize in degrading specific substrates and provide increased nutrients for the community as a whole (Baty et al., 2000). If opportunistic, functionally redundant specialists stochastically filled available niches provided by mussel-derived nutrients, this in turn may be responsible for the observed increase in beta diversity associated with mussels across nutrient regimes. Once both mussel-derived and specialist-enhanced resources (or genes), entered into the community, they may have been amplified and homogenized function regardless of community composition. The presence of mussels and other animals with similar ecosystem impacts may become increasingly important to maintaining ecosystem function as anthropogenic inputs to environments change over time.

This study has some implicit limitations that should be addressed in future studies. There are a broad number of possible ecoenzymes that could be investigated. Further, ecoenzymes serve primarily assimilatory functions and do not give insights into respiratory functions that also play a large role in ecosystem nutrient dynamics (Sinsabaugh & Follstad Shah, 2012). Finally, sample size and mesocosm conditions may have obscured some findings. Our results indicate that the factor with the greatest impact on microbial community structure was time from translocation. The large change in environmental factors between the Kiamichi River (where we collected the sediment) and the mesocosms, such as interstitial flow, temperature, and water nutrient content, all likely contributed to observed changes in microbial community structure (Febria et al., 2010, 2012; Tian et al., 2022). During collection, digging and transporting sediments into buckets to transport to shore likely aerated the sediment. Oxygen plays a crucial role in microbial community structure and depending on the taxa, ranges from an extremely efficient electron acceptor to an active detriment to microbial function (Fenchel & Finlay, 2008). As the experiment progressed, communities may have adapted to slowly stratifying oxygenation in the sediment. Future studies should consider investigating the role of bivalves in influencing microbial nutrient decomposition and nutrient transformation across natural systems with known differences in nutrient regimes.

However, this study points to the important role that mussels that interact with the sediment hyporheic zone—or other animals that act at boundaries between functionally discrete environments—may play in moderating ecosystem function. Our data support the hypothesis that mussels impact both sediment microbial community structure and function simultaneously, but the mechanism by which they affect both may not be strictly through nutrient limitation or bioturbation. It may be that mussels deliver and transform water column nutrients to lower levels

of the sediment and homogenize microbial metabolism. In the context of anthropogenic inputs into freshwater ecosystems, benthic ecosystem engineers may provide a buffer to microbial communities through the promotion of functional redundancy. Many nutrient transformations occur in aquatic sediment and terrestrial soils, and animals that interact in transformative ways with either warrant further investigation.

ACKNOWLEDGEMENTS

We thank M. Busch and A. Cooper for assistance with field work. We thank M. Busch, A. Cooper, and B. Wilder for assistance with carrying out the experiment. We thank C. Siler for use of molecular facilities and D. Ning for their bioinformatic pipeline. We thank C. Atkinson, K. Sankaranarayanan, C. Siler, and L. Souza for their input to study design and comments on the manuscript. Funding was provided by the University of Oklahoma Research Council and the Department of Biology through an Adams Summer Scholarship. This paper is part of a dissertation at the University of Oklahoma and is a contribution to the program of the Oklahoma Biological Survey.

LITERATURE CITED

Allan, J. D., & Castillo, M. M. (2007). Stream Ecology. Springer Netherlands. https://doi.org/10.1007/978-1-4020-5583-6

Allison, S. D., & Martiny, J. B. H. (2008). Resistance, resilience, and redundancy in microbial communities. Proceedings of the National Academy of Sciences, 105(supplement_1), 11512–11519. https://doi.org/10.1073/pnas.0801925105

Atkinson, C. L., Capps, K. A., Rugenski, A. T., & Vanni, M. J. (2017). Consumer-driven nutrient dynamics in freshwater ecosystems: From individuals to ecosystems. Biological Reviews, 92(4), 2003–2023. https://doi.org/10.1111/brv.12318

Atkinson, C. L., & Forshay, K. J. (2022). Community patch dynamics governs direct and indirect nutrient recycling by aggregated animals across spatial scales. Functional Ecology, 36(3), 595–606. https://doi.org/10.1111/1365-2435.13982

Atkinson, C. L., Sansom, B. J., Vaughn, C. C., & Forshay, K. J. (2018). Consumer Aggregations Drive Nutrient Dynamics and Ecosystem Metabolism in Nutrient-Limited Systems. Ecosystems, 21(3), 521–535. https://doi.org/10.1007/s10021-017-0166-4

Atkinson, C. L., van Ee, B. C., & Pfeiffer, J. M. (2020). Evolutionary history drives aspects of stoichiometric niche variation and functional effects within a guild. Ecology, 101(9), 1–12. https://doi.org/10.1002/ecy.3100

Atkinson, C. L., & Vaughn, C. C. (2015). Biogeochemical hotspots: Temporal and spatial scaling of the impact of freshwater mussels on ecosystem function. Freshwater Biology, 60(3), 563–574. https://doi.org/10.1111/fwb.12498

Baiser, B., Lockwood, J. L., La Puma, D., & Aronson, M. F. J. (2008). A perfect storm: Two ecosystem engineers interact to degrade deciduous forests of New Jersey. Biological

Invasions, 10(6), 785–795. https://doi.org/10.1007/s10530-008-9247-9

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Berga, M., Östman, Ö., Lindström, E. S., & Langenheder, S. (2015). Combined effects of zooplankton grazing and dispersal on the diversity and assembly mechanisms of bacterial metacommunities. Environmental Microbiology, 17(7), 2275–2287. https://doi.org/10.1111/1462-2920.12688

Black, E. M., Chimenti, M. S., & Just, C. L. (2017). Effect of freshwater mussels on the vertical distribution of anaerobic ammonia oxidizers and other nitrogen-transforming microorganisms in upper Mississippi river sediment. PeerJ, 5, e3536. https://doi.org/10.7717/peerj.3536

Black, E. M., & Just, C. L. (2018). The Genomic Potentials of NOB and Comammox Nitrospira in River Sediment Are Impacted by Native Freshwater Mussels. Frontiers in Microbiology, 9(September), 2061. https://doi.org/10.3389/fmicb.2018.02061

Blakeslee, C. J., Galbraith, H. S., Robertson, L. S., & St. John White, B. (2013). The effects of salinity exposure on multiple life stages of a common freshwater mussel, *Elliptio complanata*. Environmental Toxicology and Chemistry, 32(12), 2849–2854. https://doi.org/10.1002/etc.2381

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology, 37(8), 852–857.

https://doi.org/10.1038/s41587-019-0209-9

Bonaglia, S., Nascimento, F. J. A., Bartoli, M., Klawonn, I., & Brüchert, V. (2014). Meiofauna increases bacterial denitrification in marine sediments. Nature Communications, 5(1), 5133. https://doi.org/10.1038/ncomms6133

Boström, B., Andersen, J. M., Fleischer, S., & Jansson, M. (1988). Exchange of Phosphorus Across the Sediment-Water Interface. In G. Persson & M. Jansson (Eds.), Phosphorus in Freshwater Ecosystems (pp. 229–244). Springer Netherlands. https://doi.org/10.1007/978-94-009-3109-1_14

Bowen, J. L., Crump, B. C., Deegan, L. A., & Hobbie, J. E. (2009). Salt marsh sediment bacteria: Their distribution and response to external nutrient inputs. The ISME Journal, 3(8), 924–934. https://doi.org/10.1038/ismej.2009.44

Bowen, J. L., Ward, B. B., Morrison, H. G., Hobbie, J. E., Valiela, I., Deegan, L. A., & Sogin, M. L. (2011). Microbial community composition in sediments resists perturbation by nutrient enrichment. The ISME Journal, 5(9), 1540–1548. https://doi.org/10.1038/ismej.2011.22

Brazier, R. E., Puttock, A., Graham, H. A., Auster, R. E., Davies, K. H., & Brown, C. M. L. (2021). Beaver: Nature's ecosystem engineers. WIREs Water, 8(1), e1494. https://doi.org/10.1002/wat2.1494

Buchkowski, R. W., & Schmitz, O. J. (2022). Weak interactions between strong interactors in an old-field ecosystem: Control of nitrogen cycling by coupled herbivores and detritivores. Functional Ecology, 36(1), 133–147. https://doi.org/10.1111/1365-2435.13932

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & 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

Chen, S.-Y., Feng, Z., & Yi, X. (2017). A general introduction to adjustment for multiple comparisons. Journal of Thoracic Disease, 9(6), 1725–1729. https://doi.org/10.21037/jtd.2017.05.34

Chen, Q., Chen, F., Gonsior, M., Li, Y., Wang, Y., He, C., Cai, R., Xu, J., Wang, Y., Xu, D., Sun, J., Zhang, T., Shi, Q., Jiao, N., & Zheng, Q. (2021). Correspondence between DOM molecules and microbial community in a subtropical coastal estuary on a spatiotemporal scale. Environment International, 154, 106558. https://doi.org/10.1016/j.envint.2021.106558

Coggan, N. V., Hayward, M. W., & Gibb, H. (2018). A global database and "state of the field" review of research into ecosystem engineering by land animals. Journal of Animal Ecology, 87(4), 974–994. https://doi.org/10.1111/1365-2656.12819

Coverdale, T. C., Kartzinel, T. R., Grabowski, K. L., Shriver, R. K., Hassan, A. A., Goheen, J. R., Palmer, T. M., & Pringle, R. M. (2016). Elephants in the understory: Opposing direct and indirect effects of consumption and ecosystem engineering by megaherbivores. Ecology, 97(11), 3219–3230. https://doi.org/10.1002/ecy.1557

Eldridge, D. J., Delgado-Baquerizo, M., Woodhouse, J. N., & Neilan, B. A. (2016). Mammalian engineers drive soil microbial communities and ecosystem functions across a disturbance gradient. Journal of Animal Ecology, 85(6), 1636–1646. https://doi.org/10.1111/1365-2656.12574

Eldridge, D. J., Woodhouse, J. N., Curlevski, N. J. A., Hayward, M., Brown, M. V., & Neilan, B. A. (2015). Soil-foraging animals alter the composition and co-occurrence of microbial communities in a desert shrubland. ISME Journal, 9(12), 2671–2681. https://doi.org/10.1038/ismej.2015.70

Emery-Butcher, H. E., Beatty, S. J., & Robson, B. J. (2020). The impacts of invasive ecosystem

engineers in freshwaters: A review. Freshwater Biology, 65(5), 999–1015. https://doi.org/10.1111/fwb.13479

Findlay, S. E. G., & Parr, T. B. (2017). Dissolved Organic Matter. Methods in Stream Ecology, 21–36. https://doi.org/10.1016/B978-0-12-813047-6.00002-4

Galbraith, H. S., Blakeslee, C. J., & Lellis, W. A. (2015). Behavioral responses of freshwater mussels to experimental dewatering. Freshwater Science, 34(1), 42–52. https://doi.org/10.1086/679446

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., & Egozcue, J. J. (2017). Microbiome datasets are compositional: And this is not optional. Frontiers in Microbiology, 8(NOV), 1–6. https://doi.org/10.3389/fmicb.2017.02224

Graves, C. J., Makrides, E. J., Schmidt, V. T., Giblin, A. E., Cardon, Z. G., & Rand, D. M. (2016). Functional Responses of Salt Marsh Microbial Communities to Long-Term Nutrient Enrichment. Applied and Environmental Microbiology, 82(9), 2862–2871. https://doi.org/10.1128/AEM.03990-15

Haag, W. R. (2012). North American freshwater mussels: Natural history, ecology, and conservation. Cambridge University Press.

Haynes, G. (2012). Elephants (and extinct relatives) as earth-movers and ecosystem engineers. Geomorphology, 157–158, 99–107. https://doi.org/10.1016/j.geomorph.2011.04.045

Higgins, E., Parr, T. B., & Vaughn, C. C. (2022). Mussels and Local Conditions Interact to Influence Microbial Communities in Mussel Beds. Frontiers in Microbiology, 12. https://www.frontiersin.org/article/10.3389/fmicb.2021.790554

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics, 6(2), 65–70.

Hopper, G. W., Chen, S., Sánchez González, I., Bucholz, J. R., Lu, Y., & Atkinson, C. L. (2021). Aggregated filter-feeders govern the flux and stoichiometry of locally available energy and nutrients in rivers. Functional Ecology, 35(5), 1183–1195. https://doi.org/10.1111/1365-2435.13778

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. https://doi.org/10.1128/AEM.01043-13

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software, 82, 1–26. https://doi.org/10.18637/jss.v082.i13

Langsrud, Ø. (2003). ANOVA for unbalanced data: Use Type II instead of Type III sums of squares. Satistics and Computing, 13, 163–167. https://doi.org/10.1023/A:1023260610025

Martinez, A. (2020). pairwiseAdonis: Pairwise multilevel comparison using adonis (R package version 0.4). https://github.com/pmartinezarbizu/pairwiseAdonis

Matthews, W. J., Vaughn, C. C., Gido, K. B., & Marsh-Matthews, E. (2005). Southern Plains Rivers. In A. Benke & C. Cushings (Eds.), North American Rivers (pp. 283–325). Elsevier.

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

Parr, T. B., Vaughn, C. C., & Gido, K. B. (2019). Animal effects on dissolved organic carbon bioavailability in an algal controlled ecosystem. Freshwater Biology, 65, 1298-1310. https://doi.org/10.1111/fwb.13438

Penha, J., Landeiro, V. L., Ortega, J. C. G., & Mateus, L. (2017). Interchange between flooding and drying, and spatial connectivity control the fish metacommunity structure in lakes of the Pantanal wetland. Hydrobiologia, 797(1), 115–126. https://doi.org/10.1007/s10750-017-3164-9

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Research, 41(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

Quinn, T. P., Erb, I., Richardson, M. F., & Crowley, T. M. (2018). Understanding sequencing data as compositions: An outlook and review. Bioinformatics, 34(16), 2870–2878. https://doi.org/10.1093/bioinformatics/bty175

R Core Team. (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/

Rushing, C. S., Rohrbaugh, R. W., Fiss, C. J., Rosenberry, C. S., Rodewald, A. D., & Larkin, J. L. (2020). Long-term variation in white-tailed deer abundance shapes landscape-scale population dynamics of forest-breeding birds. Forest Ecology and Management, 456, 117629. https://doi.org/10.1016/j.foreco.2019.117629

Schmitz, O. J., Wilmers, C. C., Leroux, S. J., Doughty, C. E., Atwood, T. B., Galetti, M., Davies, A. B., & Goetz, S. J. (2018). Animals and the zoogeochemistry of the carbon cycle. Science, 362(6419). https://doi.org/10.1126/science.aar3213

Shelton, A. L., Henning, J. A., Schultz, P., & Clay, K. (2014). Effects of abundant white-tailed deer on vegetation, animals, mycorrhizal fungi, and soils. Forest Ecology and Management, 320, 39–49. https://doi.org/10.1016/j.foreco.2014.02.026

Sieczko, A., Maschek, M., & Peduzzi, P. (2015). Algal extracellular release in river-floodplain dissolved organic matter: Response of extracellular enzymatic activity during a post-flood period. Frontiers in Microbiology, 6. https://www.frontiersin.org/articles/10.3389/fmicb.2015.00080

Sinsabaugh, R. L., & Follstad Shah, J. J. (2012). Ecoenzymatic Stoichiometry and Ecological Theory. Annual Review of Ecology, Evolution, and Systematics, 43(1), 313–343. https://doi.org/10.1146/annurev-ecolsys-071112-124414

Sinsabaugh, R. L., Hill, B. H., & Follstad Shah, J. J. (2009). Ecoenzymatic stoichiometry of microbial organic nutrient acquisition in soil and sediment. Nature, 462(7274), 795–798. https://doi.org/10.1038/nature08632

Sitters, J., Bakker, E. S., Veldhuis, M. P., Veen, G. F., Olde Venterink, H., & Vanni, M. J. (2017). The Stoichiometry of Nutrient Release by Terrestrial Herbivores and Its Ecosystem Consequences. Frontiers in Earth Science, 5. https://www.frontiersin.org/articles/10.3389/feart.2017.00032

Sliva, L., & Williams, D. D. (2005). Exploration of riffle-scale interactions between abiotic variables and microbial assemblages in the hyporheic zone. Canadian Journal of Fisheries and Aquatic Sciences, 62(2), 276–290. https://doi.org/10.1139/f04-190

Spivak, E. D. (2010). The crab *Neohelice* (*Chasmagnathus*) *granulata*: An emergent animal model from emergent countries. Helgoland Marine Research, 64(3), 149–154. https://doi.org/10.1007/s10152-010-0198-z

Spooner, D. E., & Vaughn, C. C. (2006). Context-dependent effects of freshwater mussels on stream benthic communities. Freshwater Biology, 51(6), 1016–1024. https://doi.org/10.1111/j.1365-2427.2006.01547.x

Spooner, D. E., & Vaughn, C. C. (2008). A trait-based approach to species' roles in stream ecosystems: Climate change, community structure, and material cycling. Oecologia, 158(2), 307–317. https://doi.org/10.1007/s00442-008-1132-9

Tait, L. W., Lohrer, A. M., Townsend, M., Atalah, J., Floerl, O., & Inglis, G. J. (2020). Invasive ecosystem engineers threaten benthic nitrogen cycling by altering native infaunal and biofouling communities. Scientific Reports, 10(1), 1581. https://doi.org/10.1038/s41598-020-58557-8

Trentman, M. T., Atkinson, C. L., & Brant, J. D. (2018). Native freshwater mussel effects on nitrogen cycling: Impacts of nutrient limitation and biomass dependency. Freshwater Science, 37(2), 276–286. https://doi.org/10.1086/697293

van den Boogaart, K. G., & Tolosana-Delgado, R. (2008). "compositions": A unified R package to analyze compositional data. Computers & Geosciences, 34(4), 320–338. https://doi.org/10.1016/j.cageo.2006.11.017

Vanni, M. J. (2002). Nutrient cycling by animals in freshwater ecosystems. Annual Review of Ecology and Systematics, 33, 341–370. https://doi.org/10.1146/annurev.ecolsys.33.010802.150519

Vaughn, C. C. (2010). Biodiversity Losses and Ecosystem Function in Freshwaters: Emerging Conclusions and Research Directions. BioScience, 60(1), 25–35. https://doi.org/10.1525/bio.2010.60.1.7

Vaughn, C. C., & Hoellein, T. J. (2018). Bivalve Impacts in Freshwater and Marine Ecosystems. Annual Review of Ecology, Evolution, and Systematics, 49(1), 183–208. https://doi.org/10.1146/annurev-ecolsys-110617-062703

Vaughn, C. C., Parr, T. B., DuBose, T. P., Gates, K. K., Hopper, G. W., & Gido, K. B. (2022).

Do mobile consumers homogenize the distribution of resources in stream food webs? A test with overlapping fish and mussel aggregations. Freshwater Biology, 67(4), 684–694. https://doi.org/10.1111/fwb.13872

Vaughn, C. C., & Spooner, D. E. (2006). Unionid mussels influence macroinvertebrate assemblage structure in streams. Journal of the North American Benthological Society, 25(3), 691–700. https://doi.org/10.1899/0887-3593(2006)25[691:UMIMAS]2.0.CO;2

Vaughn, C. C., Spooner, D. E., & Galbraith, H. S. (2007). Context-Dependent Species Identity Effects Within a Functional Group of Filter-Feeding Bivalves. Ecology, 88(7), 1654–1662. https://doi.org/10.1890/06-0471.1

Wright, S. P. (1992). Adjusted P-Values for Simultaneous Inference. Biometrics, 48(4), 1005–1013. https://doi.org/10.2307/2532694

Yang, X., Chen, X., & Yang, X. (2019). Effect of organic matter on phosphorus adsorption and desorption in a black soil from Northeast China. Soil and Tillage Research, 187, 85–91. https://doi.org/10.1016/j.still.2018.11.016

Tables

Table 1. Pairwise comparisons of PERMANOVA models testing for effects of mussels and nutrient on sediment microbial community structure between nutrient treatments on day 7. "N:M" indicates the effect of the interaction between nutrient and mussel treatments. Significant values are in bold. Significant effects of single variables were not evaluated if the interaction effect was significant.

	Control			N			P		
		F	Pr(>F)						
N	Nutrient	1.5194	0.011						
	Mussel	1.3879	0.023						
	N:M	**2.0786**	**0.001**						
					F	Pr(>F)			
P	Nutrient	1.3546	0.038	Nutrient	1.1859	0.131			
	Mussel	1.6207	0.006	**Mussel**	**1.5576**	**0.009**			
	N:M	**1.6961**	**0.007**	N:M	1.0875	0.245			
							F	Pr(>F)	
NP	Nutrient	1.2275	0.145	**Nutrient**	**1.4196**	**0.030**	Nutrient	0.9283	0.576
	Mussel	1.3330	0.067 .	**Mussel**	**1.5467**	**0.008**	**Mussel**	**1.4796**	**0.025**
	N:M	**1.8984**	**0.002**	N:M	1.0215	0.415	N:M	1.0110	0.446

Table 2. ANOVA results for ecoenzyme activity and ecoenzyme activity ratios. Significant results are in bold.

	BG Activity		NAG+LAP Activity		Phos Activity		BG/[NAG+LAP] Activity		BG/Phos Activity		[NAG+LAP]/Phos Activity	
	X^2	*p*	X^2	*p*	X^2	*p*	F	*p*	F	*p*	F	*p*
Nutrient	7.900	**0.048**	6.980	0.072	4.192	0.242	1.158	0.348	1.345	0.285	0.884	0.465
Mussel	2.600	0.107	0.037	0.850	0.037	0.847	6.305	**0.020**	6.421	**0.019**	0.265	0.612
Nutrient:Mussel	0.144	0.986	2.323	0.501	3.465	0.325	0.897	0.458	1.283	0.305	1.183	0.339

Figure legends

Figure 1. (A) Sediment and mussels from the Kiamichi River were placed in enclosures within each flow-through mesocosm. (B) Experimental design used in study. Colors and letters correspond to nutrient treatments (N – nitrogen, P – phosphorus, NP – both nitrogen and phosphorus, ctrl – no nutrient control) and line type represent mussel treatment (solid – mussel treatment, dashed (non-mussel).

Figure 2. Boxplots of alpha diversity metrics of sediment microbial communities over sampling dates. (A) Shannon diversity. (B) Inverse Simpson's Evenness Index.

Figure 3. PCA of sediment microbial communities collected on the first day of the experiment before the introduction of mussels and the final day of the experiment for (A) nutrient control treatments, (B) nitrogen treatments, (C) nitrogen and phosphorus treatments, and (D) phosphorus treatments. Lines are drawn from the centroid of the samples collected on the first day to the centroid of samples on the final day. Ellipses are drawn at a 50% confidence interval around centroids. Lighter shades and dashed lines indicate non-mussel treatments. Samples collected on the first day of the experiment are all considered non-mussel treatments as mussels had not been introduced. All panels are based on the same PCA but with all ellipses other than the nutrient in question masked for clarity.

Figure 4. Boxplots of alpha diversity metrics for samples collected on the final day of the experiment by nutrient and mussel treatment. (A) Shannon diversity. (B) Inverse Simpson's Evenness Index.

Figure 5. PCA of sediment microbial communities collected on the final day of the experiment for (A) control treatments, (B) nitrogen treatments, (C) nitrogen and phosphorus treatments, and (D) phosphorus treatments. Ellipses are drawn at a 50% confidence interval around centroids.

Lighter shades and dashed lines indicate non-mussel treatments. All panels are based on the same PCA, but with all ellipses other than the nutrient in question masked for clarity.

Figure 6. Boxplots of ecoenzyme activity within nutrient and mussel treatments on the final day of the experiment. Enzyme activities shown are (A) C:N enzyme activity ratio, (B) C:P enzyme activity ratio, (C) N:P enzyme activity ratio, (D) BG enzyme activity (ln[nmol · hr-1· mg AFDM-1]), (E) sum of NAG + LAP enzyme activity (ln[nmol · hr-1· mg AFDM-1]) and, (F) AP enzyme activity (ln[nmol · hr-1· mg AFDM-1]).

Figures

Figure 1.

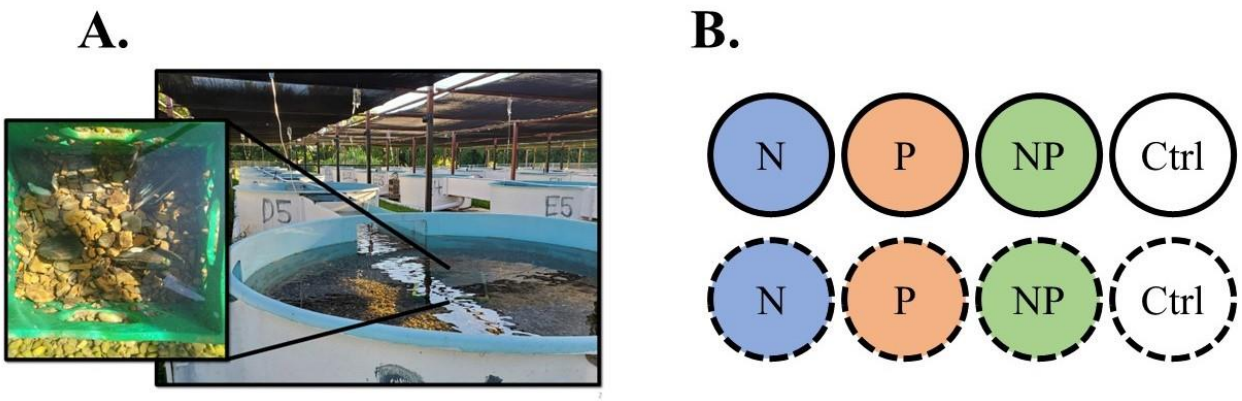


Figure 2.

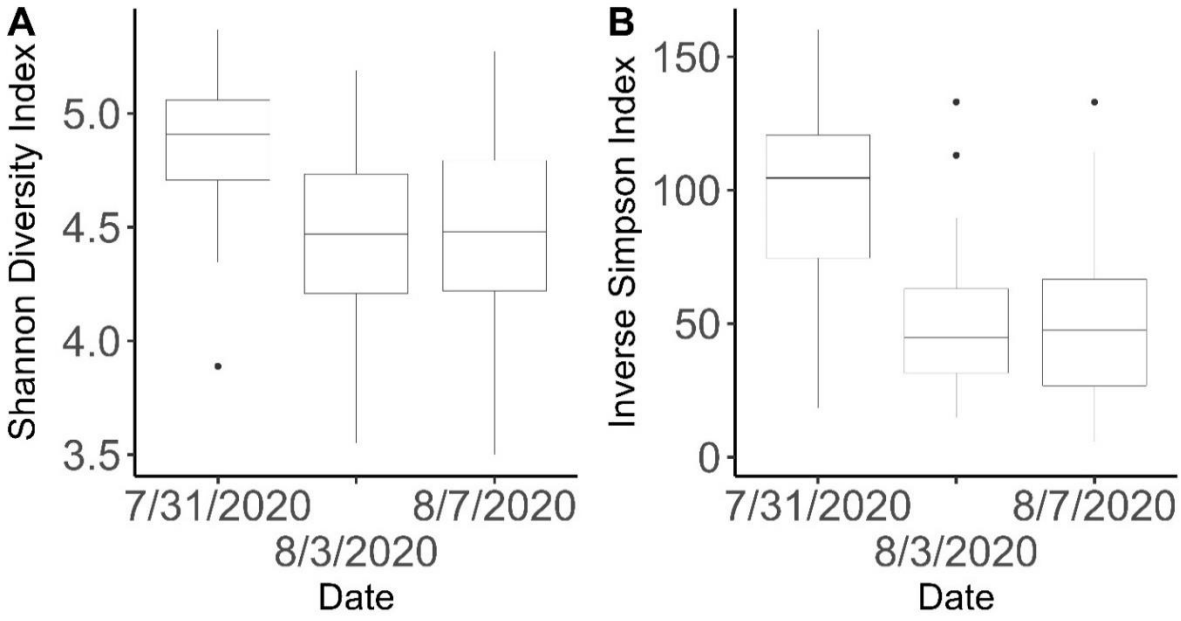

A
Shannon Diversity Index
5.0
4.5
4.0
3.5
7/31/2020
8/3/2020
8/7/2020
Date
B
Inverse Simpson Index
150
100
50
0
7/31/2020
8/3/2020
8/7/2020
Date

Figure 3.

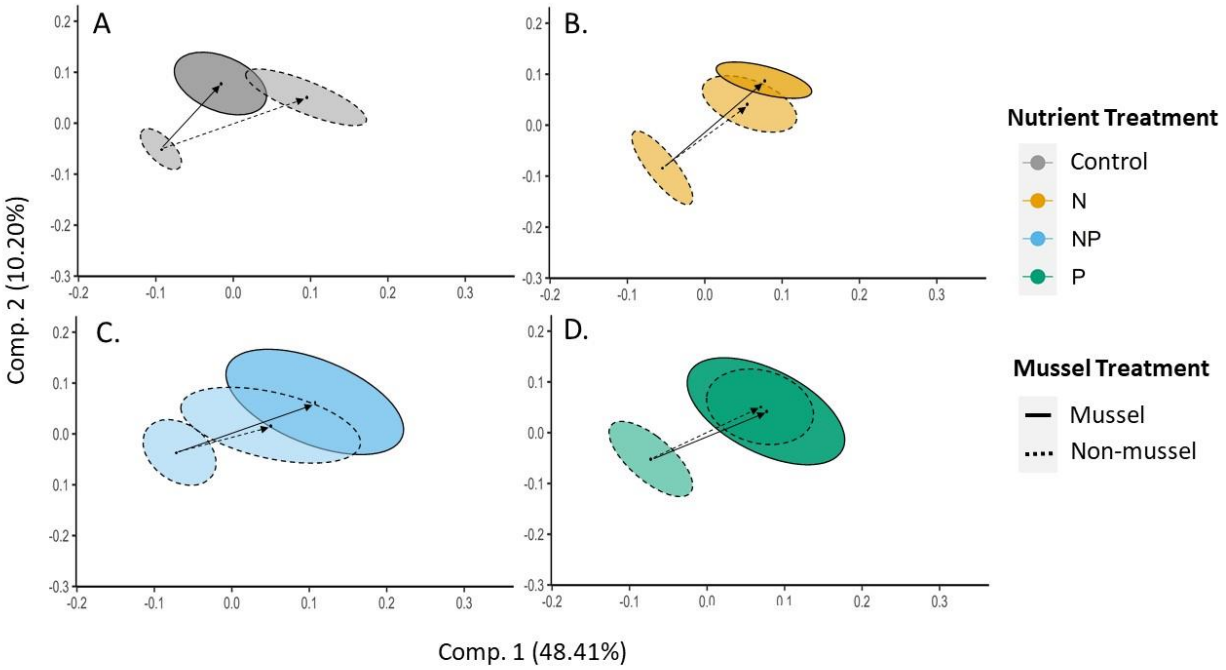

A
B.
C.
D.
Comp. 2 (10.20%)
Comp. 1 (48.41%)
Nutrient Treatment
Control
N
NP
P
Mussel Treatment
Mussel
Non-mussel

Figure 4.

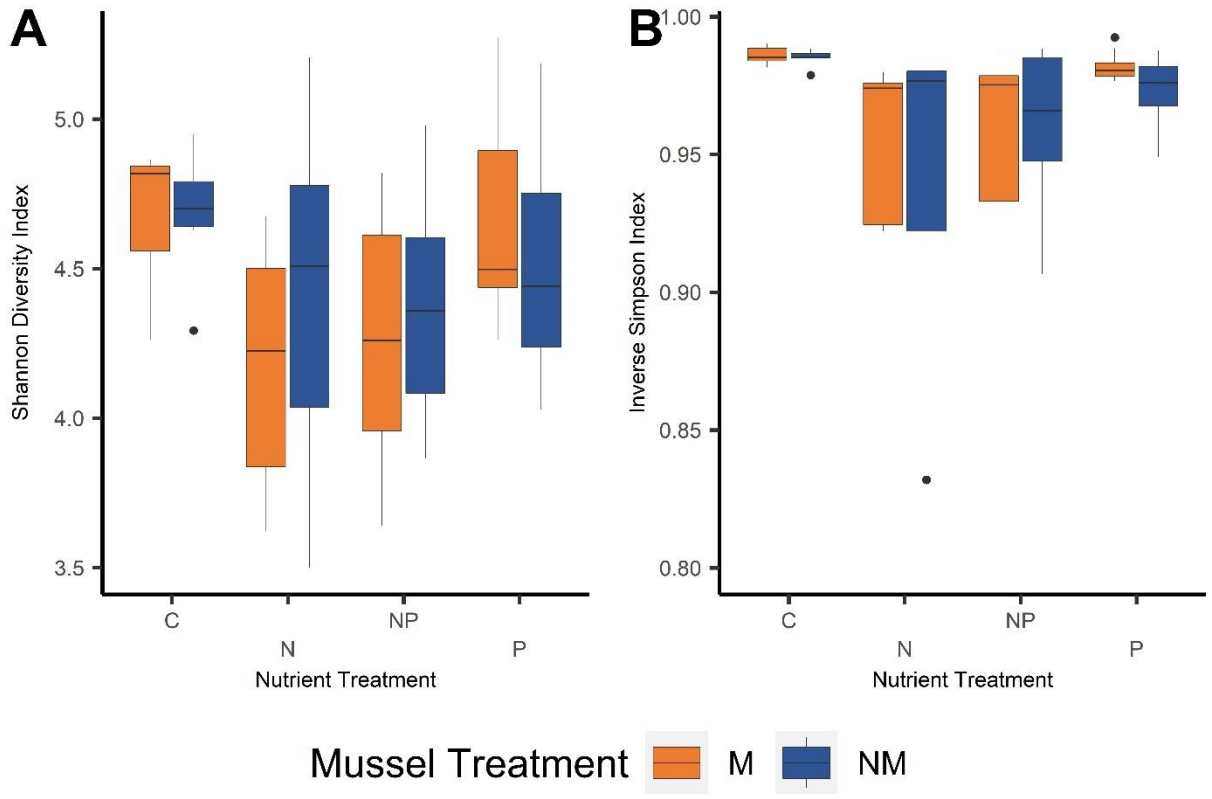

A
Shannon Diversity Index
5.0
4.5
4.0
3.5
C
N
NP
P
Nutrient Treatment
B
Inverse Simpson Index
1.00
0.95
0.90
0.85
0.80
C
N
NP
P
Nutrient Treatment
Mussel Treatment
M
NM

Figure 5.

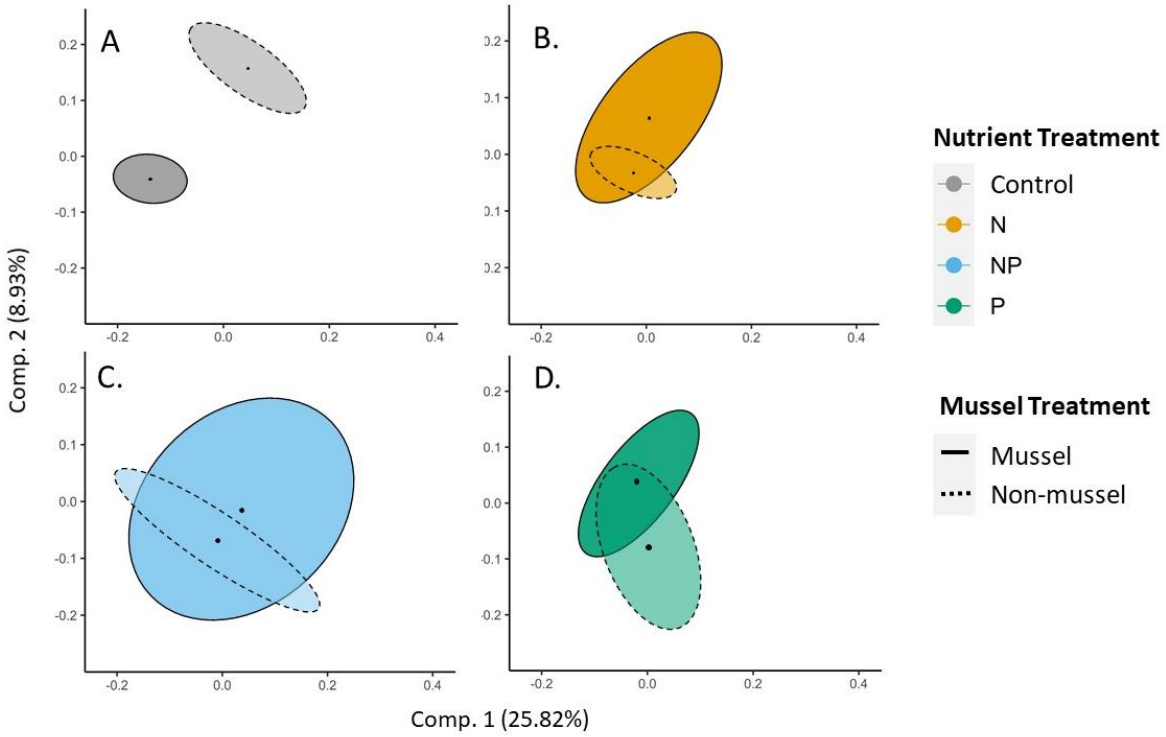
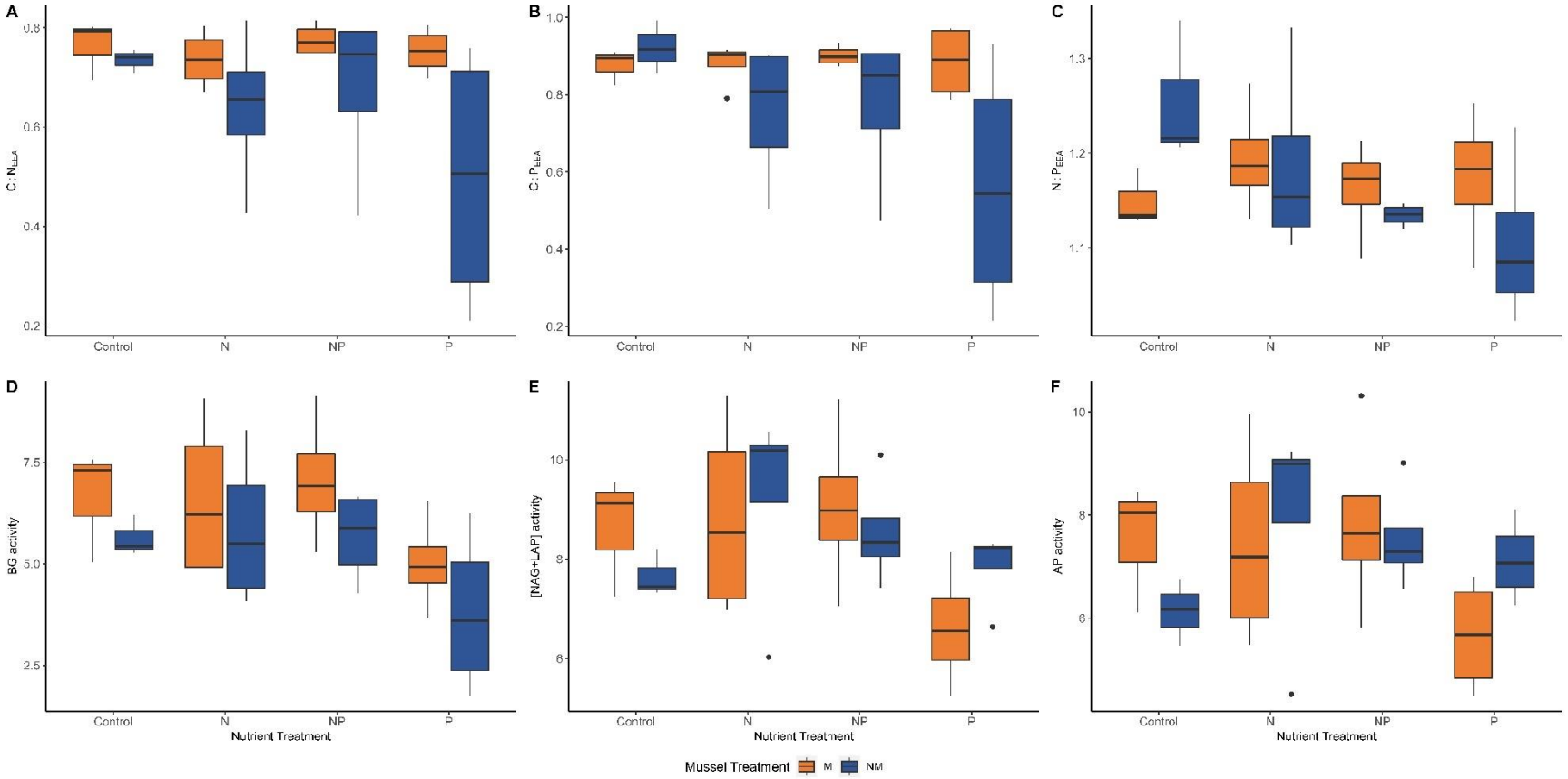

A
B.
C.
D.
Comp. 2 (8.93%)
Comp. 1 (25.82%)
Nutrient Treatment
Control
N
NP
P
Mussel Treatment
Mussel
Non-mussel

Figure 6.


A
B
C
D
E
F
C : N_EEA
C : P_EEA
N : P_EEA
BG activity
[NAG+LAP] activity
AP activity
Control
N
NP
P
Nutrient Treatment
Mussel Treatment
M
NM

CHAPTER THREE

Efficacy of “Computer Science in Modern Biology,” a Peer-to-Peer Virtual Workshop to Improve Student Confidence

Keywords:

biology; DEI; coding; interdisciplinary; STEM education; STEM intervention

Formatted for submission to the *Journal of Science Education and Technology*

Authors: Edward Douglas Higgins, McKenna Burns, Paula Cimprich, Mehrnaz Afkhami, Meelyn Pandit, Hayley Lanier, Katharine Marske, Laura Stein, Sara Mata

Abstract

Early STEM intervention programs can be effective at improving student performance and retention in STEM. However, because these programs vary in their implementation, evaluation of their structure and effectiveness in improving student outcomes is imperative for continued success. In 2020, a cross-university collaboration designed a workshop designed to improve participant's self-efficacy and self-concept—both of which have been shown to improve student persistence in STEM, especially for underrepresented students. The week-long, virtual workshop offered 14 courses in R and Python with a focus on computing in the biological sciences. Courses were most frequently taught by graduate students to implement a peer-to-peer mentoring strategy for undergraduate, high school, and community college students. The workshop was successful in improving student self-efficacy, even three months after the workshop, but had more modest success in improving student self-concept. These results were not different among demographic groups, indicating this STEM program did not mirror detrimental impacts on minority students seen systemically across STEM. Peer-to-peer workshops can be effective methods for improving student confidence in unfamiliar topics, even in virtual environments. However, more robust methods for developing a sense of community and self-concept may be necessary in short-term distance-learning.

Introduction

Policy makers maintain an expectation for the United States to be a world leader in science, technology, engineering, and math (STEM), and recognize the need for adequate STEM training nationwide (The President's Council of Advisors on Science and Technology, 2020). Yet, the U.S. is falling behind other countries in multiple metrics including reduced global share of research and development performance (National Science Board, 2022), as well as student numeracy and problem solving (National Academies of Sciences & Medicine, 2016). While the surplus or dearth of qualified professionals is inconsistent among STEM careers, there remains a consistent pattern across academia, government, and industry: labor shortages in professionals with coding and software experience (Xue & Larson, 2015).

The need for STEM professionals is also reflected in the growth of computer science jobs, where opportunities are predicted to expand at three times the average of other occupations between 2019 and 2029 (Zilberman & Ice, 2021). Even before entering the workforce, many K–12 students use computer skills daily, and many future careers, both STEM and non-STEM, will include some component of computer science (National Science Board, 2020b). In addition, the COVID-19 pandemic has provided insight into the importance of computer skills for continued education, and STEM careers have been shown to be more resilient to the pandemic than non-STEM jobs (National Science Board, 2021). Yet, academia has been unable to meet the STEM workforce need for college graduates with adequate computer science skills (Xue & Larson, 2015).

Disparities in Computer Science – A major factor underlying the dearth in STEM talent is the loss of women and underrepresented minority (URM) students through the academic process, despite strong initial interest (Estrada et al., 2011; Williams et al., 2019) and this

problem has persisted for decades (Sciences et al., 2007). While the STEM industry in the United States has made progress over time in increasing participation of underrepresented groups, this progress is nuanced, and its success limited; ultimately, over the past few decades there has been an increase in the total number of women and other historically URM participating in STEM, but these increases are not keeping pace with the overall growth of the industry and this is especially noticeable in Computer Science (National Science Board, 2020b, 2022). Unfortunately, systemic barriers in K–12 education often prevent equal access to enrolling and succeeding in computer science among Black, Hispanic, and female students (Wang & Hejazi Moghadam, 2017). Although roughly 51 percent of high schools in the U.S. offer foundational computer courses, these courses are much less common in urban, rural, and economically disadvantaged populations and Black, Hispanic, poor, and rural students are less likely to attend schools that offer them (Hendrickson et al., 2021; Upadhyaya et al., 2020).

This inequity then persists into post-secondary education. One study that compared persistence rates among students from different racial backgrounds from all public universities in Ohio between 1998 and 2002 found that by the third year of college, 63 percent of White STEM students remained in their field of study compared to only 48 percent of Black STEM students (Price, 2010). Similarly, using data from the National Center of Education Statistics, Riegle-Crumb et al. (2019) demonstrated that STEM was the only field in which Black and Hispanic students were significantly more likely than their White peers to switch educational fields, even when controlling for social class. This trend is also prevalent between women and men: only 18 percent of all undergraduate computer and information sciences degrees are awarded to women, and Black and Latina women comprise less than 4 percent of bachelor's degrees in engineering and computer science (National Science Board, 2020a). Broadly, there is a dearth of studies on

LGBTQ experiences in STEM, but evidence shows that sexual minorities experience greater career limitations, harassment, and professional devaluation and are more likely to leave STEM than their non-LGBTQ peers (Cech & Pham, 2017; J. Freeman, 2018; J. B. Freeman, 2020).

Many STEM intervention programs have been developed with the goal of addressing these inequalities, but quantitative evaluations of their efficacy are vital to their continued funding and operation (Scott, 2013). To date, many programs focus on recruitment of underrepresented groups as well as the development of support networks, but the longevity of STEM intervention programs depends on institutional commitment (Pon-Barry et al., 2017; Rincon & George-Jackson, 2016). Transient funding opportunities combined with the necessity of long-term efforts to improve equity therefore requires the development and promotion of sustainable educational interventions with evidence of effectiveness in addressing factors pertinent to participation and retention of students in STEM.

Near-Peer Education and Psychological Factors for Retention – Two well documented, crucial predictors of student retention in STEM fields—especially URM retention—are self-efficacy, the confidence a student has to achieve their goals, and self-concept, how one perceives oneself (Dyer-Barr, 2014; Eccles & Wigfield, 2002; Lazarides et al., 2020; Museus et al., 2011). Increasing student exposure to role models in science and engineering can contribute to higher self-efficacy of underrepresented students pursuing STEM degrees (Amelink & Creamer, 2010; Karunanayake & Nauta, 2004). Additionally, encouragement from both instructors as well as peers are important factors relating to retention of women in science courses (Denner et al., 2014; Dingel, 2006). Along these lines, “near-peer” or “peer-to-peer” education is a pedagogical technique in which students are mentored by students near in age or education status which has been demonstrated to benefit both students and their peer mentors in STEM (Jin & Xu, 2019;

Tenenbaum et al., 2014; Trujillo et al., 2015). Implementing educational interventions using near-peer learning may be a fruitful avenue to promote both self-efficacy and self-concept in underrepresented STEM students (Hall et al., 2013; A. J. Nelson et al., 2013).

Students with early positive computational experiences are more likely to pursue a degree in a computationally intensive field in college, to the extent that simply having positive interactions with computing can predict a students' intentions to major in computer science (Barker et al., 2009; McInerney et al., 2006). Historically, women have had fewer positive experiences in STEM courses which can contribute to feeling inadequately prepared for pursuing computationally intensive major (Patall et al., 2018). Providing effective introductory courses can recruit more students to STEM programs; in contrast, ineffective pedagogy of introductory level courses can deter students from pursuing STEM fields (Cohoon et al., 2013; Mervis, 2010; Watkins & Mazur, 2022). To provide more supportive pathways to STEM degrees, and ultimately, into the STEM workforce, it is crucial to develop pedagogically effective coding and computer science interventions for high school and pre-collegiate students.

Here, we describe the results of a workshop aimed at providing foundational training in computationally intensive fields within biology, and examine its ability to improve self-efficacy and self-concept. Students self-enrolled in courses which focused on computer skills essential in a wide variety of biological data science. Most courses were developed and taught by graduate students, and coding skills were focused primarily on teaching basic coding within R and Python, two widely used languages in biology. Students applied newly acquired skills by analysing ecological and evolutionary datasets, often produced through the instructor's current research. The final day of the workshop was a 'career mixer' in which students interacted with a variety of professionals in biological professions. This study examines the efficacy of this educational

intervention by evaluating student's perceptions of their own capabilities before and their progress after participating.

Methods

Workshop design – The "Computer Science in Modern Biology Workshop" (CompSciBio Workshop) was first developed by researchers at Miami University (MU) in 2019. In 2020, due to the COVID-19 pandemic, the workshop was moved to an online format hosted over the Zoom video conferencing platform. The online format allowed for a collaboration between organizers at Miami University (MU) and the University of Oklahoma's (OU) STEM Inclusion Council, a graduate student organization comprised entirely of students from underrepresented backgrounds majoring in STEM related fields. Participation was free and fourteen courses were taught to a total of 100 students, each with one instructor and two assistant instructors (Table 1). Instructors were the primary lecturers while assistant instructors oversaw attendance lists, troubleshot software errors, and answered student questions using Zoom break-out rooms. Instructors were recruited from both universities, with instructor positions offered first to graduate students and then to tenure-track faculty, postdoctoral researchers, and OU science librarians. Four instructors were faculty, nine were graduate students (both masters and doctoral), and one was an OU science librarian. Two assistant instructors were faculty, two were OU science librarians, and the remaining were graduate students (Table 1).

In developing this workshop, the primary aims outlined to instructors before course and curricula design were: (1) to make coding more approachable, (2) to provide insight into the computational and social skills needed by computational biologists, and (3) to help develop a sense of belonging for students from underrepresented groups in STEM. Each instructor met with workshop organizers and were asked to address key issues in their course design:

preventing “zoom fatigue,” how to use assistant instructors to ensure all students needs were addressed when instructors were busy, to share personal stories about their own learning process and, if they were comfortable, their experiences with underrepresentation in STEM. Courses were required to use only free software, and instructors designed their courses on a topic of their choice. The only courses for which there were prescribed topics were the Introduction to R and Introduction to Python course series.

The target audience for these courses were underrepresented minority (URM) students in high school, undergraduate, and graduate school; however, the workshop was open to all students in their junior year of high school or above. Advertising began in April of 2020 and continued through June. The workshop was advertised to five high schools with websites indicating high enrollment of URM students (four confirmed they distributed materials); two community, or regional colleges (one confirmed they distributed materials); and both OU and MU undergraduate and graduate students (both confirmed). Due to the pandemic, in-person recruitment was not possible and so for schools other than OU and MU, workshop organizers requested advertising materials be sent to science teachers to distribute to their students. Other advertising included social media announcements through Twitter and Facebook, word-of-mouth, academic department announcements, and fliers posted at OU and MU.

The workshop was held over five days in the second week of August 2020 and students could take a maximum of four courses throughout the workshop. Courses were designated as either introductory or advanced and brief descriptions of each course were posted on the workshop website. Students self-enrolled in June through the workshop website, with introductory courses recommended to students with limited or no experience in coding. For advanced courses, instructors provided guidelines of suggested prerequisite skills for students to

assess their preparation before selecting their courses. Course enrollment was capped at 25 students. Each course was five total hours and split into two, 2.5-hour course periods across two days. Although students had to use their own computers, students who elected to provide their address during the enrollment process were mailed personalized digital and hard copies of course materials—flash drives and binders— based on their selected courses two weeks prior to the workshop. All binders included printed course materials, blank note sheets, information on installing software, and links to online coding courses.

Morning and afternoon course periods were separated by voluntary lunch hour sessions comprised of a thirty-minute break followed by a thirty-minute seminar. Lunch hours were optional for both participants and graduate student instructors, however there were graduate student organizers at every lunch hour. Therefore, in addition to formal class periods, students could interact informally with graduate students and discuss issues people have faced in STEM programs.

On the final day of the workshop, a three and one-half hour "mixer" between students and 10 professionals from various fields of biology provided students with insights into different career paths. Professionals were invited individually by workshop organizers and included biologists from state wildlife agencies, federal science policy analysts, research faculty at domestic and foreign universities, the United States Environmental Protection Agency, biotechnology scientists, Fulbright Fellows, and natural history museum management. Using Zoom break-out rooms, small groups (between two and three) of students spent fifteen-minute rotations with one to two professionals. Each group was accompanied by a graduate student organizer to facilitate student questions about work skills, career choices, duties, and professional development.

Survey design – To assess the workshop, three anonymous surveys approved by the institutional review board (IRB #12352; Appendix 1-3) were conducted to evaluate student confidence before and after the workshop in their ability to solve problems using coding as well as student feelings of adequacy using the tools and resources provided by the workshop. The surveys were hosted on the online platform Qualtrics and distributed through email. The first survey was given one week prior to the workshop. The second survey was distributed immediately after the conclusion of the workshop and was available to complete for two weeks. The third and final survey was distributed three months following the workshop and available for two weeks. At the end of the third survey there was a link to an optional entry into a monetary raffle. Students could win one of seven digital Visa gift cards of which there was one valued at $100.00, two at $50.00, and four at $25.00 gift cards. Raffle winners were emailed their gift card.

Participants first answered three questions to generate an individualized, anonymous three-character code used to link participant responses through the multiple surveys. Only the first and third survey contained demographic questions, and these anonymous codes were used to determine demographic information of students who took the second survey. Four survey questions examined effectiveness of the workshop and confidence in computer skills using the 5-point Likert Scale (Likert, 1932; Norman, 2010). These questions assessed confidence in downloading software, approaching coding problems, solving coding problems, and how the participants thought their skills compared to those of their peers. To assess student self-concept, students selected a Venn diagram they felt represented the overlap between themselves and a STEM professional (McDonald et al., 2019), answered one question from the Clance Imposter Phenomenon Scale (Chrisman et al., 1995; Clance, 1985), and one question addressing feelings

of being overwhelmed in their studies (Kolligian Jr. & Sternberg, 1991).

Statistical analyses – Differences in student responses between each survey for each question were determined using parametric (Norman, 2010) mixed linear models in R (R Core Team 2020), employing the packages lme4_1.1-26 (Bates et al., 2015) and lmerTest_3.1-3 (Kuznetsova et al., 2017). To examine differential effects of the workshop on student confidence among demographic groups, each model calculated interactions between the time surveyed (pre-workshop, post-workshop, or three months after) and gender identity, sexual identity, and race and ethnicity on student responses to each question. Individualized participant codes were used as a random effect to account for lack of independence among the samples (Zuur et al., 2007). We calculated post-hoc estimated marginal means tests with Tukey adjusted p-values for significant fixed effects with the R package emmeans_1.6.3 (Lenth, 2021).

Results

Sample Characteristics – A total of 50 students responded to the first survey, 40 to the second survey, and 47 to the final survey. Attendance documents indicate a total of 100 students showed up to at least one course and a total of 100 individualized, anonymous three-character code were generated across the three surveys indicating that each student took at least one survey, but few took all three (Fig. 1). Survey respondents ranged from 18–41 years old with a median age of 21. No students under the age of 18 enrolled. 70 percent of respondents identified as heterosexual, 4 percent chose not to answer, and 26 percent reported as a member of the LGBTQ+ community including 17 percent bisexual, 1 percent gay, 2 percent pansexual, 4 percent queer, and 2 percent questioning. 70 percent of respondents identified as female, 28 percent as male, one respondent identified as gender queer, and one chose not to answer. 68 percent of respondents identified as White, 2 percent chose not to answer, and 30 percent

identified as non-White, including 2 percent multiracial, 9 percent Hispanic or Latino, 17 percent Asian or Asian American, and 1 percent American Indian or Alaska Native.

Survey sample sizes were not large enough to account for all races and ethnicities polled, and so survey responses were binned into either White or non-White. Only one student identified as “Genderqueer or gender non-conforming” and was removed from the analyses regarding gender. Like race and ethnicity, there was not a large enough sample size to address all sexual identities, and we therefore binned responses into “heterosexual” and “LGBTQ+.”

Student responses – Four questions measured students' self-reported confidence in solving or approaching computational problems. Students did not report a statistically significant increase in confidence downloading and installing software after attending the workshop ($F = 0.535$, $p = 0.5894$). However, the average confidence level of students began relatively high at a rank of 4.2 out of 5. While student responses increased immediately after the workshop, the average response returns to baseline three months after the workshop. When asked how confident students were in approaching a project that requires coding or computer skills, student confidence increased significantly ($F = 8.75$, $p = 0.0005$; Fig 2A). This degree of confidence was retained after 3 months in relation to the second survey results as there was no difference between the second and third survey (t ratio = 1.026, $p = 0.564$). Student's overall confidence in their coding ability or use of programming/analysis and visualization software was significantly higher after the workshop ($F = 26.32$, $p < 0.0001$; Fig 2B) and was similarly retained between the second and third survey (t ratio = -0.403, $p = 0.915$). When asked to compare their coding ability to that of their peers, students showed a significant increase in confidence ($F = 12.25$, $p < 0.0001$; Fig 2C) which was retained over time (t ratio = 0.121, $p = 0.992$). The average response increased from “somewhat worse” than their peers to “about the same” as their peers. There were

no significant interactions between any self-efficacy question and gender, race, or sexual preference.

Three questions were posed to assess student's self-concept. When asked to select a Venn diagram illustration to best describe themselves as a STEM professional, for which higher values indicate students considered themselves more of a STEM professional, there was a significant difference in student identity between the three surveys ($F = 7.122$, $p = 0.002$; Fig 2D). Post-hoc tests showed this difference was likely driven by a significant increase in self-identification as a STEM professional between the first and final survey (t ratio = -3.098, $p = 0.008$). The average score throughout the three surveys was a four—the intermediate image between zero identification and full identification as a STEM professional—but there was an increase in the upper quartile of responses after the workshop (Fig 2D). However, students saw no significant change in feeling over their head or beyond their capabilities in their schoolwork after the workshop ($F = 2.06$, $p = 0.13$). Students were also asked if they were worried people important to them would find out they are not as capable as they seem, and students were relatively concerned their inabilities would be discovered, with the average rank starting and remaining at a four out of five throughout the duration of the surveys ($F = 0.41$, $p = 0.67$). For these three questions there were no significant interactions with gender, race, or sexual identity.

Discussion

Coding workshops increase student coding confidence – This study points to the success of the Computer Science in Modern Biology workshop's ability to promote self-efficacy in coding within biological sciences of high school, undergraduate, and graduate students seemingly without continuing the trend of discouraging URM students from continuing in STEM.

However, student demographics indicate only moderate success in the recruitment of URM students. With 70 percent White, 17 percent Asian students, and 9 percent Hispanic or Latino, the workshop had racial and ethnic participation that partially reflects that of the United States STEM industry with a bachelor's degree (66 percent White, 16.3 percent Asian, and 7.7 percent Hispanic or Latino; National Science Board, 2022), but notably had zero Black students. This is an important gap to be addressed in future workshops, as retention of Black and Hispanic students continue to lag behind in STEM participation (National Science Board, 2022). Interestingly, at 70 percent participation of women, this workshop focusing on the interdisciplinary field of computational biology more closely reflects participation of women in biology but not computer science (Kahn & Ginther, 2017). In contrast, based on previous research on LGBTQ+ participation in STEM careers, which found 2.8 percent of survey respondents identified as LGBTQ+ (Cech & Pham, 2017), this workshop had considerably higher than expected participation from LGBTQ+ individuals (30 percent). Taken together, workshop recruitment was no more successful than previous efforts in STEM based on race, ethnicity and gender, but more successful based on sexual identity. However, studies on sexual minorities in STEM are scarce and participation of sexual minorities in STEM is not often interrogated. The recruitment of diverse sexual identities may indicate a success but may also reflect an increase in identification as LGBTQ+ among younger generations (Gates, 2017). Yet overall, given the lack of statistical interactions among demographic groups, the workshop still had beneficial impacts on self-efficacy among both minority and non-minority students.

Factors such as race and gender diversity representation in the instructors and speakers may have ensured an improvement of the perception of diverse students regarding their role in computer science and biology. All courses had at least one female instructor or assistant

instructor and twelve out of fourteen courses had members of a graduate student group comprised of underrepresented students in biology as instructors or assistant instructors. This may be an underestimation of the instructor diversity as demographic information on volunteers was not collected. Representation among instructors can improve URM student self-efficacy and self-concept in STEM programs (D. J. Nelson & Rogers, 2003; Shin et al., 2016). Course sizes were limited to 25 students and thus students were able to have more direct interactions with diverse graduate students nearer to their education level than traditional instructors. Further, lunch hours encouraged students to engage in conversations with graduate students involved in the workshop where discussions commonly included hardships overcame by older students and advice on furthering STEM careers. Student comments regarding the lunch hours in the second survey were very favourable, with multiple students reporting that they enjoyed interacting in an informal environment with graduate students and other peers as well as most students explicitly reporting they had no negative comments. The combination of both vocational and psychosocial interactions with near-peers are key aspects of mentorship and can lead to improved student outcomes in STEM performance and retention (Crisp & Cruz, 2009; Zaniewski & Reinholz, 2016). This workshop contributes to the literature supporting the idea that representative instructors and plentiful opportunities for near-peer teaching can contribute to improving outcomes for students involved in SIPs.

Changes in self-concept are comparatively smaller – How a student considers their abilities in comparison to their peers can be considered a measurement of both self-efficacy and self-concept, and students saw favourable changes in this regard. However, other measurements of self-concept saw smaller changes after the workshop.

There was a significant increase in self-concept when considering how much students

consider themselves a STEM professional. However, this change was not ubiquitous among all students. The median age of students was 21, so it is possible younger students felt they were not considered professionals as they were still completing their schooling regardless of skills and interest in the field. Further, while students saw no significant changes in consistently feeling overwhelmed by work, this workshop was offered in late summer while most students were on summer vacation and ended just before the school semester started. However, there were also no significant change in overwhelmed feelings once students had returned to school, and it is possible the workshop's proximity to the school year was able to help mitigate an increase of said feelings. Conversely, it is concerning that student worries of being as seen as incapable were consistently high throughout the duration of the study. It is important to consider that anxiety and mental health concerns increased greatly during the COVID-19 pandemic, especially among adolescents and young adults (Courtney et al., 2020; Hawes et al., 2021). A virtual workshop in the midst of virtual schooling may not have been able to create the sense of community necessary to aid student feelings of self-concept. Similar workshops held in person could potentially better address these needs, resulting in fewer anxious feelings of being seen as incapable.

Taken in conjunction with the survey's self-efficacy questions, it appears that the educational aspects of this workshop were sufficient to assuage student's feelings of unpreparedness, but student's self-concept was at best modestly affected. This is not entirely surprising: while self-efficacy can predict student intention to pursue scientific careers, self-concept is a more evasive yet critical motivation to persistence (Estrada et al., 2011). Lunch seminars and the career mixer were voluntary, and attendance ranged from 25 to 40 students. This modest attendance indicates that most students may have viewed the workshop as primarily

an educational tool rather than a community building opportunity, especially in a virtual setting. Students that engage in collaborative and empowering activities, as opposed to the competitive nature of coursework, can see higher levels of self-concept, and so improving attendance of these aspects in future workshops may yield more improvements in self-concept among students (Hurtado et al., 2010). Additionally, students that seek out educational opportunities, and therefore encounter new programs such as this workshop, may have declined to participate in lunch sessions to focus on coursework to learn new skills and catch up with their peers whom they view as more fluent in their field, or simply take a break from their computer screens. However, once higher self-efficacy is achieved, students can reject stigma associated with being an underrepresented student in science and hold higher degrees of self-concept in STEM (Hurtado et al., 2010). It is possible that STEM intervention programs must first increase self-efficacy before students can internalize a sense of community and self-concept.

Future Directions – Overall, students reported higher levels of confidence regarding both their coding skills and when comparing themselves to their peers. Participation of graduate students and the intimacy of smaller courses may have served to improve student's retention of tools they need to feel more comfortable in computational biology. Student reviews also demonstrated that graduate student participation was especially important and appreciated in the more casual settings of the workshop. There were no statistical interactions between these increases in confidence and race, gender, or sexual identity suggesting that these increases were equally successful for URM and non-URM students. Student increases in self-concept were modest in comparison to confidence, and this may be improved by increasing participation in community building aspects of the workshop, especially interactions with their peers and near-peers. Recruitment efforts were limited by the COVID-19 pandemic, and this may be linked with

the middling success in recruiting students from underrepresented backgrounds. In-person recruitment and building relationships with the same schools, teachers, and communities over time will likely demonstrate more success in future iterations. Given insights from this study, improving the success of future workshops will require reaching a more targeted audience who may further benefit from educational intervention rather than those who self-select for educational opportunities.

Acknowledgements

Funding for the 2020 Computer Science in Modern Biology workshop was provided by the M.I.A.M.I. Women Giving Circle, and the Office of the Vice President of Research and Partnerships and the Department of Biology at the University of Oklahoma. Support for the collection and analysis of data from the workshop was supported by the University of Oklahoma Office of the Vice President of Research and Partnerships. We thank the University of Oklahoma STEM Inclusion Council for their assistance in coordinating the workshop and all graduate student and faculty instructors of both MU and OU for their time and expertise in educating our students. We further gratefully acknowledge J. Kelly for help in grant proposal development and developing IRB surveys and C. Curry for help in data analysis.

Literature cited

Amelink, C. T., & Creamer, E. G. (2010). Gender Differences in Elements of the Undergraduate Experience that Influence Satisfaction with the Engineering Major and the Intent to Pursue Engineering as a Career. *Journal of Engineering Education*, *99*(1), 81–92. https://doi.org/10.1002/j.2168-9830.2010.tb01044.x

Barker, L. J., McDowell, C., & Kalahar, K. (2009). Exploring Factors that Influence Computer Science Introductory Course Students to Persist in the Major. *Association for Computing Machinery*, *41*(1), 5. https://doi-org.ezproxy.lib.ou.edu/10.1145/1539024.1508923

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Cech, E. A., & Pham, M. V. (2017). Queer in STEM organizations: Workplace disadvantages for LGBT employees in STEM related federal agencies. *Social Sciences*, *6*(1), 12. https://doi.org/10.3390/socsci6010012

Chrisman, S. M., Pieper, W. A., Clance, P. R., Holland, C. L., & Glickauf-Hughes, C. (1995). Validation of the Clance Imposter Phenomenon Scale. *Journal of Personality Assessment*, *65*(3), 456–467. https://doi.org/10.1207/s15327752jpa6503_6

Clance, P. (1985). *The impostor phenomenon: Overcoming the fear that haunts your success*. Peachtree Publishers.

Cohoon, J. P., Cohoon, J. M., & Soffa, M. L. (2013). Educating Diverse Computing Students at the University of Virginia. *Computer*, *46*(3), 52–55. https://doi.org/10.1109/MC.2013.39

Courtney, D., Watson, P., Battaglia, M., Mulsant, B. H., & Szatmari, P. (2020). COVID-19 Impacts on Child and Youth Anxiety and Depression: Challenges and Opportunities. *The*

Canadian Journal of Psychiatry, *65*(10), 688–691. https://doi.org/10.1177/0706743720935646

Crisp, G., & Cruz, I. (2009). Mentoring College Students: A Critical Review of the Literature Between 1990 and 2007. *Research in Higher Education*, *50*(6), 525–545. https://doi.org/10.1007/s11162-009-9130-2

Denner, J., Werner, L., O'Connor, L., & Glassman, J. (2014). Community College Men and Women: A Test of Three Widely Held Beliefs About Who Pursues Computer Science. *Community College Review*, *42*(4), 342–362. https://doi.org/10.1177/0091552114535624

Dingel, M. J. (2006). Gendered experiences in the science classroom. *Removing Barriers: Women in Academic Science, Technology, Engineering, and Mathematics*, 161–176.

Dyer-Barr, R. (2014). Research to practice: Identifying best practices for STEM intervention programs for URMs. *Quality Approaches in Higher Education*, *5*(1), 19–25.

Eccles, J. S., & Wigfield, A. (2002). Motivational Beliefs, Values, and Goals. *Annual Review of Psychology*, *53*(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Estrada, M., Woodcock, A., Hernandez, P. R., & Schultz, P. W. (2011). Toward a Model of Social Influence that Explains Minority Student Integration into the Scientific Community. *Journal of Educational Psychology*, *103*(1), 206–222. https://doi.org/10.1037/a0020743

Freeman, J. (2018). LGBTQ scientists are still left out comments. *Nature*, *559*(7712), 27–28. https://doi.org/10.1038/d41586-018-05587-y

Freeman, J. B. (2020). Measuring and Resolving LGBTQ Disparities in STEM. *Policy Insights from the Behavioral and Brain Sciences*, *7*(2), 141–148. https://doi.org/10.1177/2372732220943232

Gates, G. J. (2017). LGBT Data Collection Amid Social and Demographic Shifts of the US LGBT Community. *American Journal of Public Health*, *107*(8), 1220–1222. https://doi.org/10.2105/AJPH.2017.303927

Hall, S., Lewis, M., Border, S., & Powell, M. (2013). Near-peer teaching in clinical neuroanatomy. *The Clinical Teacher*, *10*(4), 230–235. https://doi.org/10.1111/tct.12001

Hawes, M. T., Szenczy, A. K., Klein, D. N., Hajcak, G., & Nelson, B. D. (2021). Increases in depression and anxiety symptoms in adolescents and young adults during the COVID-19 pandemic. *Psychological Medicine*, 1–9. https://doi.org/10.1017/S0033291720005358

Hendrickson, K., Gauthier, L., Osorio Glennon, M., Menocal Harrigan, A., Weissman, H., Fletcher, C., Dunton, S., Baskin, J., & Mak, J. (2021). *2021 State of Computer Science Education* (p. 116). Code.org Advocacy Coalition, Computer Science Teachers Assiocation, and Expanding Computing Education Pathways. https://advocacy.code.org/stateofcs

Hurtado, S., Newman, C. B., Tran, M. C., & Chang, M. J. (2010). Improving the rate of success for underrepresented racial minorities in STEM fields: Insights from a national project. *New Directions for Institutional Research*, *2010*(148), 5–15. https://doi.org/10.1002/ir.357

Jin, W., & Xu, X. (2019). Near-peer Led Workshops on Game Development for Broadening Participation and Diversity in Computing. *Proceedings of the 2019 ACM Southeast Conference*, 39–45. https://doi.org/10.1145/3299815.3314430

Kahn, S., & Ginther, D. (2017). *Women and STEM* (No. w23525; p. w23525). National Bureau of Economic Research. https://doi.org/10.3386/w23525

Karunanayake, D., & Nauta, M. M. (2004). The Relationship Between Race and Students' Identified Career Role Models and Perceived Role Model Influence. *The Career Development Quarterly*, *52*(3), 225–234. https://doi.org/10.1002/j.2161-0045.2004.tb00644.x

Kolligian Jr., J., & Sternberg, R. J. (1991). Perceived Fraudulence in Young Adults: Is There an "Imposter Syndrome"? *Journal of Personality Assessment*, *56*(2), 308–326. https://doi.org/10.1207/s15327752jpa5602_10

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Lazarides, R., Dicke, A.-L., Rubach, C., & Eccles, J. S. (2020). Profiles of motivational beliefs in math: Exploring their development, relations to student-perceived classroom characteristics, and impact on future career aspirations and choices. *Journal of Educational Psychology*, *112*(1), 70–92. https://doi.org/10.1037/edu0000368

Lenth, R. V. (2021). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. https://CRAN.R-project.org/package=emmeans

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*.

McDonald, M. M., Zeigler-Hill, V., Vrabel, J. K., & Escobar, M. (2019). A Single-Item Measure for Assessing STEM Identity. *Frontiers in Education*, *4*(July), 1–15. https://doi.org/10.3389/feduc.2019.00078

McInerney, C. R., DiDonato, N. C., Giagnacova, R., & O'Donnell, A. M. (2006). Students' Choice of Information Technology Majors and Careers: A Qualitative Study. *Information Technology, Learning, and Performance Journal*, *24*(2), 35–53.

Mervis, J. (2010). Better Intro Courses Seen as Key to Reducing Attrition of STEM Majors. *Science*, *330*(6002), 306–306. https://doi.org/10.1126/science.330.6002.306

Museus, S. D., Palmer, R. T., Davis, R. J., & Maramba, D. C. (2011). *Racial and Ethnic Minority Students' Success in STEM Education*. *36*(6), 1–140. http://dx.doi.org/10.1002/aehe.3606

National Academies of Sciences, Engineering & Medicine. (2016). *Developing a National STEM Workforce Strategy: A Workshop Summary* (J. Alper, Ed.). The National Academies Press. https://doi.org/10.17226/21900

National Science Board, N. S. F. (2020a). *Science and Engineering Indicators 2020* (NSB-2020-1). https://ncses.nsf.gov/pubs/nsb20201/

National Science Board, N. S. F. (2020b). *Vision 2030* (NSB-2020-15; p. 24). https://www.nsf.gov/nsb/publications/2020/nsb202015.pdf

National Science Board, N. S. F. (2021). *The STEM Labor Force of Today: Scientists, Engineers and Skilled Technical Workers* (NSB-2021-2; Science and Engineering Indicators 2022). National Science Board. https://ncses.nsf.gov/pubs/nsb20212

National Science Board, N. S. F. (2022). *Science and Engineering Indicators 2022: The State of U.S. Science and Engineering* (NSB-2022-1). https://ncses.nsf.gov/pubs/nsb20221

Nelson, A. J., Nelson, S. V., Linn, A. M. J., Raw, L. E., Kildea, H. B., & Tonkin, A. L. (2013). Tomorrow's educators ... today? Implementing near-peer teaching for medical students. *Medical Teacher*, *35*(2), 156–159. https://doi.org/10.3109/0142159X.2012.737961

Nelson, D. J., & Rogers, D. C. (2003). *A national analysis of diversity in science and engineering faculties at research universities*. National Organization of Women.

https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.180.1551&rep=rep1&type=pdf

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, *15*(5), 625–632.

Patall, E. A., Steingut, R. R., Freeman, J. L., Pituch, K. A., & Vasquez, A. C. (2018). Gender disparities in students' motivational experiences in high school science classrooms. *Science Education*, *102*(5), 951–977. https://doi.org/10.1002/sce.21461

Pon-Barry, H., Packard, B. W.-L., & St. John, A. (2017). Expanding capacity and promoting inclusion in introductory computer science: A focus on near-peer mentor preparation and code review. *Computer Science Education*, *27*(1), 54–77. https://doi.org/10.1080/08993408.2017.1333270

Price, J. (2010). The effect of instructor race and gender on student persistence in STEM fields. *Economics of Education Review*, *29*(6), 901–910. https://doi.org/10.1016/j.econedurev.2010.07.009

Riegle-Crumb, C., King, B., & Irizarry, Y. (2019). Does STEM Stand Out? Examining Racial/Ethnic Gaps in Persistence Across Postsecondary Fields. *Educational Researcher*, *48*(3), 133–144. https://doi.org/10.3102/0013189X19831006

Rincon, B. E., & George-Jackson, C. E. (2016). STEM intervention programs: Funding practices and challenges. *Studies in Higher Education*, *41*(3), 429–444. https://doi.org/10.1080/03075079.2014.927845

Sciences, N. A. of, Engineering, N. A. of, & Medicine, I. of. (2007). *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. The National Academies Press. https://doi.org/10.17226/11463

Science, Technology, Engineering, and Mathematics Education: Governmentwide Strategy Needed to Better Manage Overlapping Programs, House of Representatives (2013) (testimony of George A. Scott). https://www.gao.gov/assets/gao-13-529t.pdf

Shin, J. E. L., Levy, S. R., & London, B. (2016). Effects of role model exposure on STEM and non-STEM student engagement. *Journal of Applied Social Psychology*, *46*(7), 410–427. https://doi.org/10.1111/jasp.12371

Tenenbaum, L. S., Anderson, M. K., Jett, M., & Yourick, D. L. (2014). An Innovative Near-Peer Mentoring Model for Undergraduate and Secondary Students: STEM Focus. *Innovative Higher Education*, *39*(5), 375–385. https://doi.org/10.1007/s10755-014-9286-3

The President's Council of Advisors on Science and Technology. (2020). *Recommendations for Strengthening American Leadership in Industries of the Future*. The President's Council of Advisors on Science and Technology. https://science.osti.gov/-/media/_/pdf/about/pcast/202006/PCAST_June_2020_Report.pdf?la=en&hash=019A4F17C79FDEE5005C51D3D6CAC81FB31E3ABC

Trujillo, G., Aguinaldo, P. G., Anderson, C., Bustamante, J., Gelsinger, D. R., Pastor, M. J., Wright, J., Márquez-Magaña, L., & Riggs, B. (2015). Near-peer STEM Mentoring Offers Unexpected Benefits for Mentors from Traditionally Underrepresented Backgrounds. *Perspectives on Undergraduate Research and Mentoring : PURM*, *4*(1), http://blogs.elon.edu/purm/files/2015/11/Riggs.GT-et-al-PURM-4.1.pdf.

Upadhyaya, B., McGill, M. M., & Decker, A. (2020). A Longitudinal Analysis of K-12 Computing Education Research in the United States: Implications and Recommendations for Change. In *Proceedings of the 51st ACM Technical Symposium on Computer Science*

Education (pp. 605–611). Association for Computing Machinery. http://doi.org/10.1145/3328778.3366809

Wang, J., & Hejazi Moghadam, S. (2017). Diversity Barriers in K-12 Computer Science Education: Structural and Social. *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, 615–620. https://doi.org/10.1145/3017680.3017734

Watkins, J., & Mazur, E. (2022). *Retaining Students in Science, Technology, Engineering, and Mathematics (STEM) Majors*. 7.

Williams, M. J., George-Jones, J., & Hebl, M. (2019). The face of STEM: Racial phenotypic stereotypicality predicts STEM persistence by—And ability attributions about—Students of color. *Journal of Personality and Social Psychology*, *116*(3), 416.

Xue, Y., & Larson, R. C. (2015). STEM crisis or STEM surplus? Yes and yes. *Monthly Labor Review*, *2015*, 10.21916/mlr.2015.14.

Zaniewski, A. M., & Reinholz, D. (2016). Increasing STEM success: A near-peer mentoring program in the physical sciences. *International Journal of STEM Education*, *3*(1), 14. https://doi.org/10.1186/s40594-016-0043-2

Zilberman, A., & Ice, L. (2021). *Why computer occupations are behind strong STEM employment growth in the 2019–29 decade* (No. 1; Beyond the Numbers: Employment & Unemployment). U.S. Bureau of Labor Statistics. https://www.bls.gov/opub/btn/volume-10/why-computer-occupations-are-behind-strong-stem-employment-growth.htm

Zuur, A. F., Ieno, E. N., Smith, G. M., & others. (2007). *Analysing ecological data* (Vol. 680). Springer.

Figure legends

Figure 1. Participant demographics based on survey results.

Figure 2. Box plots of student survey responses. (A) Student comfort when approaching a problem requiring coding knowledge. Higher values indicate more self-confidence. (B) Student confidence in their ability to use coding software or coding skills. Higher values indicate more self-confidence. (C) Student comparison of their coding abilities compared to their peers' abilities. A value of 2 or lower indicates the student thinks they are less skilled than their peers, 3 indicates they are equally as skilled, and 4 or higher indicates they believe they are more skilled. (D) Student self-assessment of the overlap of their own identity and that of a STEM professional. Lower values indicate they do not view themselves as a STEM professional, a value of four indicates they view themselves partially as a STEM professional, and higher values indicate much more self-concept as a STEM professional.

Tables

Table 1. Course titles and student enrollment. All courses had a cap of 25 students, however, introductory courses had two sections. GS indicates a graduate student instructor or assistant instructor, F indicates a faculty member of either OU or MU, and SL indicates a science librarian.

Course	Enrollment	Instructor	Assistant Instructor
Introduction to R, Section One	25	GS	GS,GS
Introduction to R, Section Two	25	GS	GS,GS
Introduction to R II, Section One	15	F	GS,GS
Introduction to R II, Section Two	15	GS	GS,GS
Introduction to Python, Section One	25	GS	F,GS
Introduction to Python, Section Two	25	SL	GS,GS
Introduction to Python II, Section One	22	GS	SL,GS
Introduction to Python II, Section Two	22	GS	GS,GS
Data Management	30	F	SL,GS
Evo History	26	GS	GS,GS
Intro to GIS	25	GS	GS,GS
Advanced GIS	15	GS	GS,GS
Data Visualization	21	GS	GS,GS
Niche Modeling	20	F	GS,GS
Niche Modeling	15	F	GS,GS
Loops and Functions in R	17	GS	F,GS
Introduction to Super Computing	7	F	GS,GS
Conservation Biology	6	GS	GS,F

Figures

Figure 1.

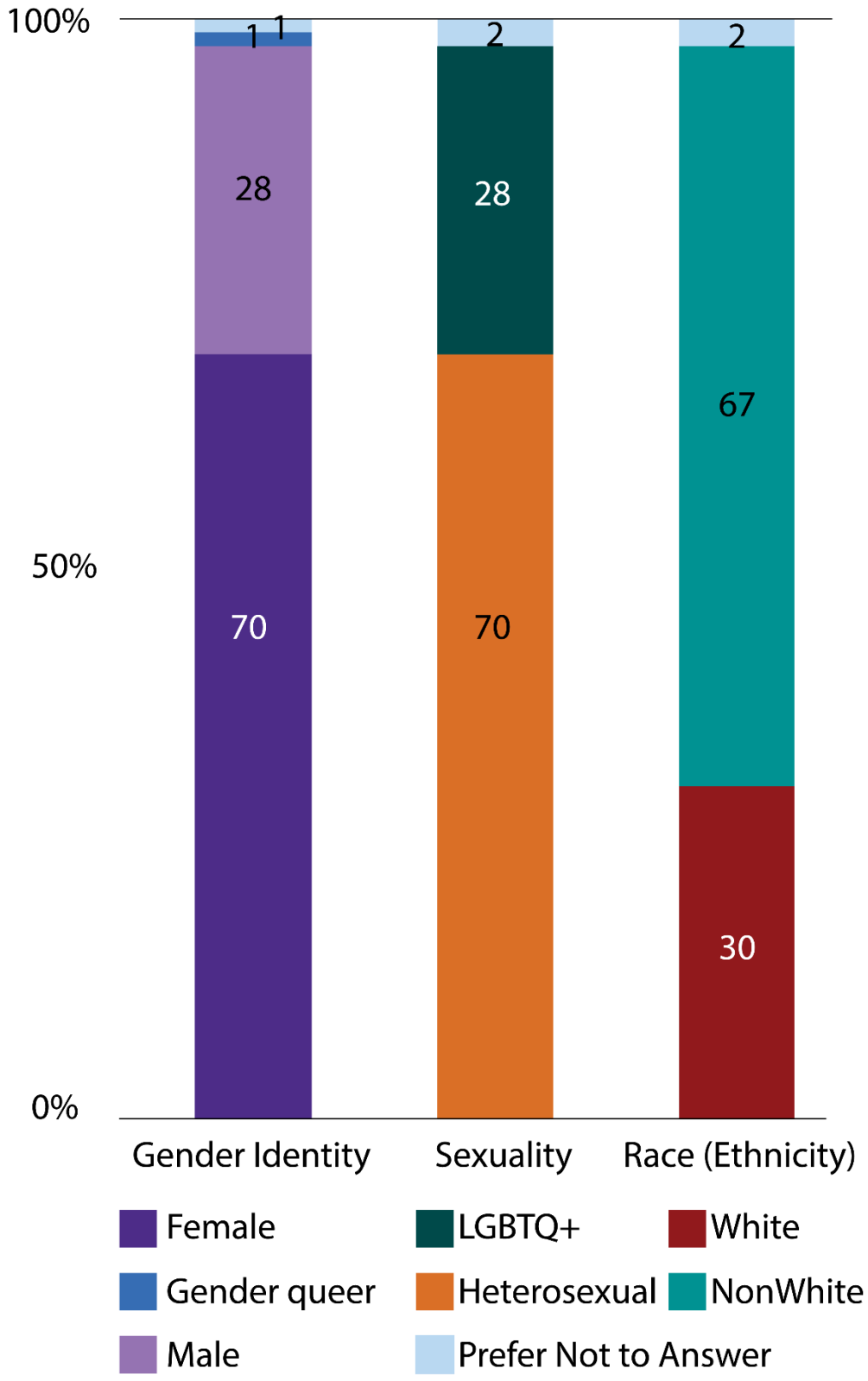
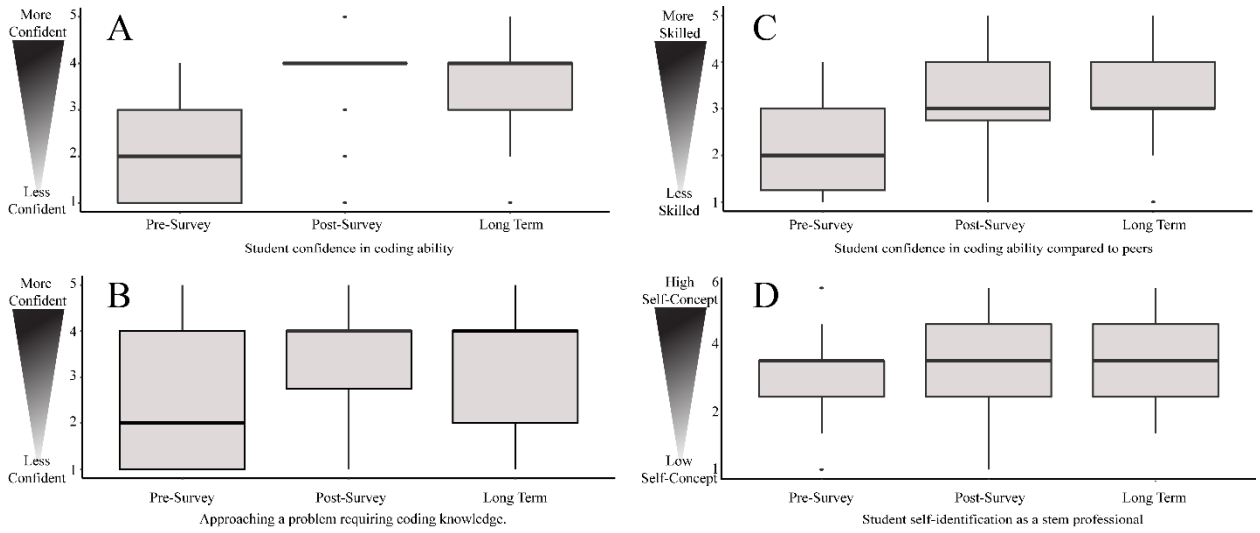


Figure 2.


A
More Confident
Less Confident
5
4
3
2
1
Pre-Survey
Post-Survey
Long Term
Student confidence in coding ability
B
More Confident
Less Confident
Pre-Survey
Post-Survey
Long Term
Approaching a problem requiring coding knowledge.
C
More Skilled
Less Skilled
Pre-Survey
Post-Survey
Long Term
Student confidence in coding ability compared to peers
D
High Self-Concept
Low Self-Concept
6
4
2
1
Pre-Survey
Post-Survey
Long Term
Student self-identification as a stem professional

Supplemental Material

Appendix 1. Survey one distributed to students before workshop.

Appendix 2. Survey two distributed to students immediately after workshop.

Appendix 3. Survey three distributed to students three months following the workshop.

Appendix 1. Survey one distributed to students before workshop.

Journal of Science Education and Technology

Efficacy of "Computer Science in Modern Biology," a Peer-to-Peer Virtual Workshop to Improve Student Confidence

PreSurvey Effectiveness of Computing in Modern Biology Workshop for Broadening Participation in STEM

Start of Block: Default Question Block

Q1

You are invited to participate in research about **your participation in the STEM workshop.** If you agree to participate, you will **complete this online survey.** There are **no risks or benefits**. There **will be no compensation for participation.** Your participation is voluntary and your responses will be **confidential** After removing all identifiers, we might share your data with other researchers or use it in future research without obtaining additional consent from you. Even if you choose to participate now, you may stop participating at any time and for any reason. Data are collected via an online survey system that has its own privacy and security policies for keeping your information confidential. No assurance can be made as to their use of the data you provide. If you have questions about this research, please contact: **Sara Mata, Ph.D. email: sara.mata@ou.edu** You can also contact the University of Oklahoma – Norman Campus Institutional Review Board at 405-325-8110 or irb@ou.edu with questions, concerns or complaints about your rights as a research participant, or if you don't want to talk to the researcher. *Please print this document for your records. By providing information to the researcher(s), I am agreeing to participate in this* research.

Are you 18 years of age or older?

- ○ No (2)
- ○ Yes (1)

Skip To: Q32 If You are invited to participate in research about your participation in the STEM workshop. If yo... = Yes

Skip To: End of Survey If You are invited to participate in research about your participation in the STEM workshop. If yo... = No

Q32 We are not asking you for any personally identifying information, however we're asking three questions that will create an anonymous key to compare before and after surveys.

What is the first letter of the name of your favorite teacher?

A-Z

Q33 What is the day of the month that you were born?

1-31

Q34 What is the second to last digit in your phone number?

1-10

Q2 How comfortable are you downloading and installing software?

- ○ Very comfortable (1)
- ○ Somewhat comfortable (2)
- ○ Neither comfortable nor uncomfortable (3)
- ○ Somewhat uncomfortable (4)
- ○ Very uncomfortable (5)

Q3 How confident are you in your coding ability or use of any programming/analysis and visualization software? If you have any coding experience, please respond based on the coding language you are most comfortable with?

- ○ Very comfortable (1)
- ○ Somewhat comfortable (2)
- ○ Neither comfortable nor uncomfortable (3)
- ○ Somewhat uncomfortable (4)
- ○ Very uncomfortable (5)



Q4 How many hours per week on average do you spend coding or using any type of programming/analysis and visualization software? For example, Excel, R, GIS, or python. Give your answer in numerical format (0, 5, 15, and so on)

_____

Q5 How comfortable do you feel finding or approaching a coding/programming/analysis and visualization project?

- ○ Very comfortable (1)
- ○ Somewhat comfortable (2)
- ○ Neither comfortable nor uncomfortable (3)
- ○ Somewhat uncomfortable (4)
- ○ Very uncomfortable (5)

Q6 How does your coding ability compare to that of your peers?

- Much better (1)
- Somewhat better (2)
- About the same (3)
- Somewhat worse (4)
- Much worse (5)

Q7 Do you aspire to a career within biology or other science field?

- Yes (1)
- Maybe (2)
- No (3)

Display This Question:

If Do you aspire to a career within biology or other science field? = Yes

Or Do you aspire to a career within biology or other science field? = Maybe

Q8 Please tell us your top 3 careers in biology or other science field that you aspire to (even if you're unsure)

_____

_____

_____

_____

_____

Q9 How important do you think coding or related computer software understanding is for your planned career goal?

- Extremely important (1)
- Very important (2)
- Moderately important (3)
- Slightly important (4)
- Not at all important (5)

Q29 I often feel that I am “in over my head” or beyond my capabilities in my school work or study

- ○ Strongly agree (11)
- ○ Somewhat agree (12)
- ○ Neither agree nor disagree (13)
- ○ Somewhat disagree (14)
- ○ Strongly disagree (15)

Q30 I’m afraid people important to me may find out that I’m not as capable as they think I am.

- ○ Strongly agree (11)
- ○ Somewhat agree (12)
- ○ Neither agree nor disagree (13)
- ○ Somewhat disagree (14)
- ○ Strongly disagree (15)

Q10 Select the picture that best describes the current overlap of the image you have of yourself and your image of what a STEM professional is

- Image:stemoverlap1 (1)
- Image:stemoverlap2 (2)
- Image:stemoverlap3 (3)
- Image:stemoverlap4 (4)
- Image:stemoverlap5 (5)
- Image:stemoverlap6 (6)

Q14 What are your goals for this workshop?

_____

_____

_____

_____

_____

End of Block: Default Question Block

Start of Block: Block 1

*

Q16 What is your age (numerical answer e.g. 18, 21, etc.)

_____

Q17 With which gender do you identify?

- ○ Female (1)
- ○ Male (2)
- ○ Non-binary (3)
- ○ Genderqueer or gender nonconforming (4)
- ○ Transgender (5)
- ○ An identity not listed, self-identify (6) _____
- ○ I prefer not to answer (7)

Q18 With which race/ethnicity do you identify? (select all that apply)

- ☐ American Indian or Alaska Native (1)
- ☐ Asian American or Asian (2)
- ☐ Black or African American (3)
- ☐ Native Hawaiian or Pacific Islander (4)
- ☐ White (5)
- ☐ Middle Eastern (6)
- ☐ Multiracial (7)
- ☐ Hispanic or Latino (8)
- ☐ I prefer not to answer (9)
- ☐ We realize that the racial/ethnic category you selected encompasses many different nationalities. If you are interested in sharing more, please describe your nationality (i.e., Armenian, Puerto Rican, Vietnamese): (10)

_____

Q19 With which sexual orientation do you identify? (Select all that apply)

- ○ Asexual (1)
- ○ Bisexual (2)
- ○ Gay (3)
- ○ Heterosexual (4)
- ○ Lesbian (5)
- ○ Pansexual (6)
- ○ Queer (7)
- ○ Questioning or unsure (8)
- ○ An identity not listed, self identify: (9) _____
- ○ I prefer not to answer (10)

Q20 What is the highest level of education you have completed so far?

- ○ Middle school (1)
- ○ High school graduate (2)
- ○ Some college (3)
- ○ 2 year degree (4)
- ○ 4 year degree (5)
- ○ Masters degree (6)

Q22 What is the highest level of education completed by a parent or guardian in your household?:

- ○ Middle school (1)
- ○ High school graduate (2)
- ○ Some college (3)
- ○ 2 year degree (4)
- ○ 4 year degree (5)
- ○ Masters degree (6)
- ○ PhD (7)
- ○ I don't know (8)

Q21 Which social class group do you, as an individual, identify with?

- ○ Poor (1)
- ○ Working Class (2)
- ○ Middle Class (3)
- ○ Upper Class (4)
- ○ I prefer not to answer (5)

Q23 What was your total household income during the past 12 months?

- ○ Less than $25,000 (1)
- ○ $25,000 - $34,999 (2)
- ○ $35,000 - $49,999 (3)
- ○ $50,000 - $74,999 (4)
- ○ $75,000 - $99,999 (5)
- ○ $100,000 - $149,999 (6)
- ○ $150,000 or more (7)
- ○ I don't know (8)
- ○ I prefer not to answer (9)

Q24 Did you receive a Federal Pell Grant as part of your financial aid package?

- ○ Yes (1)
- ○ No (2)
- ○ I don't know (3)

Q31 Where do you call home? (City, State)

_____

Q25 Please select your home living environment:

- Rural (1)
- Suburban (2)
- Urban (3)
- I don't know (4)

Q28 Growing up, was there at least one laptop or desktop computer in your home for you to use? If yes, how many?

- No, we didn't have a laptop or desktop computer in the house (1)
- 1 (2)
- 2 (3)
- 3+ (4)

Q26 Thank you for completing this survey! You are awesome!

End of Block: Block 1

Appendix 2. Survey two distributed to students immediately after workshop.

Journal of Science Education and Technology

Efficacy of "Computer Science in Modern Biology," a Peer-to-Peer Virtual Workshop to Improve Student Confidence

Post: Effectiveness of Computing in Modern Biology Workshop for Broadening Participation in STEM

Start of Block: Default Question Block

Q1 **Consent to Participate in Research at the University of Oklahoma**[OU-NC IRB Number: 12352 Approval Date: July 29, 2020] You are invited to participate in research about **your participation in the STEM workshop.** If you agree to participate, you will **complete this online survey.** There are **no risks or benefits**. There **will be no compensation for participation.** Your participation is voluntary and your responses will be **confidential** After removing all identifiers, we might share your data with other researchers or use it in future research without obtaining additional consent from you. Even if you choose to participate now, you may stop participating at any time and for any reason. Data are collected via an online survey system that has its own privacy and security policies for keeping your information confidential. No assurance can be made as to their use of the data you provide.If you have questions about this research, please contact:**Sara Mata, Ph.D. email: sara.mata@ou.edu** You can also contact the University of Oklahoma – Norman Campus Institutional Review Board at 405-325-8110 or irb@ou.edu with questions, concerns or complaints about your rights as a research participant, or if you don't want to talk to the researcher. *Please print this document for*

your records. By providing information to the researcher(s), I am agreeing to participate in this research. Are you 18 years of age or older?

- ○ Yes (1)
- ○ No (2)

Skip To: End of Survey If Consent to Participate in Research at the University of Oklahoma [OU-NC IRB Number: 12352 ... = No

Skip To: Q43 If Consent to Participate in Research at the University of Oklahoma [OU-NC IRB Number: 12352 ... = Yes

Q43 We are not asking you for any personally identifying information, however we're asking three questions that will create an anonymous key to compare before and after surveys.

What is the first letter of the name of your favorite teacher?

▼ A (1) ... Z (33)

Q45 What is the day of the month that you were born?

▼ 1 (1) ... 31 (31)

Q47 What is the second to last digit in your phone number?

1-10

Q28 Were you satisfied with the workshop?

- Strongly agree (11)
- Somewhat agree (12)
- Neither agree nor disagree (13)
- Somewhat disagree (14)
- Strongly disagree (15)

Q34 Were you successful in achieving the goals you set before the workshop?

- I achieved all of my goals (11)
- I achieved some of my goals (12)
- I achieved none of my goals (13)

Q29 What DID you like about the courses you took or that were offered?

_____

_____

_____

_____

_____

Q31 What DIDN'T you like about the courses you took or that were offered?

_____

_____

_____

_____

_____

Q32 What DID you like about the lunch seminars?

_____

_____

_____

_____

_____

Q33 What DIDN'T you like about the lunch seminars?

_____

_____

_____

_____

_____

Q35 How likely are you to contact someone you interacted with at the workshop in the future?

- Definitely I'll will contact someone (1)
- Maybe I'll contact someone (2)
- I will not contact someone (3)

Q36 I would recommend this workshop to others

- ○ Strongly agree (11)
- ○ Somewhat agree (12)
- ○ Neither agree nor disagree (13)
- ○ Somewhat disagree (14)
- ○ Strongly disagree (15)

Q38 What other topics from guest speakers would you like to see?

_____

_____

_____

_____

_____

Q37 How did you find out about this workshop?

- ○ A peer (4)
- ○ An general email from my High School or University (5)
- ○ A teacher or faculty advisory (6)
- ○ Not listed, Other: (7) _____

Q2 How comfortable now are you downloading and installing software?

- ○ Very comfortable (1)
- ○ Somewhat comfortable (2)
- ○ Neither comfortable nor uncomfortable (3)
- ○ Somewhat uncomfortable (4)
- ○ Very uncomfortable (5)

Q3 How confident are you in your coding ability or use of any programming/analysis and visualization software? If you have any coding experience, please respond based on the coding language you are most comfortable with?

- ○ Very comfortable (1)
- ○ Somewhat comfortable (2)
- ○ Neither comfortable nor uncomfortable (3)
- ○ Somewhat uncomfortable (4)
- ○ Very uncomfortable (5)



Q4 How many hours per week on average do you spend coding or using any type of programming/analysis and visualization software? For example, Excel, R, GIS, or python. Give your answer in numerical format (for example: 2)

_____

Q5 How comfortable do you feel finding or approaching a coding/programming/analysis and visualization project?

- ○ Very comfortable (1)
- ○ Somewhat comfortable (2)
- ○ Neither comfortable nor uncomfortable (3)
- ○ Somewhat uncomfortable (4)
- ○ Very uncomfortable (5)

Q6 How does your coding ability compare to that of your peers?

- ○ Much better (1)
- ○ Somewhat better (2)
- ○ About the same (3)
- ○ Somewhat worse (4)
- ○ Much worse (5)

Q7 Do you aspire to a career within biology or other science field?

- ○ Yes (1)
- ○ Maybe (2)
- ○ No (3)

Display This Question:

If Do you aspire to a career within biology or other science field? = Yes

Or Do you aspire to a career within biology or other science field? = Maybe

Q8 Please tell us your top 3 careers in biology or other science field that you aspire to (even if you're unsure)

_____

_____

_____

_____

_____

Q9 How important do you think coding or related computer software understanding is for your planned career goal?

- ○ Extremely important (1)
- ○ Very important (2)
- ○ Moderately important (3)
- ○ Slightly important (4)
- ○ Not at all important (5)

Q40 I often feel that I am “in over my head” or beyond my capabilities in my school work or study

- ○ Strongly agree (11)
- ○ Somewhat agree (12)
- ○ Neither agree nor disagree (13)
- ○ Somewhat disagree (14)
- ○ Strongly disagree (15)

Q42 I’m afraid people important to me may find out that I’m not as capable as they think I am.

- ○ Strongly agree (11)
- ○ Somewhat agree (12)
- ○ Neither agree nor disagree (13)
- ○ Somewhat disagree (14)
- ○ Strongly disagree (15)

Q10 Select the picture that best describes the current overlap of the image you have of yourself and your image of what a STEM professional is

- ○ Image:stemoverlap1 (1)
- ○ Image:stemoverlap2 (2)
- ○ Image:stemoverlap3 (3)
- ○ Image:stemoverlap4 (4)
- ○ Image:stemoverlap5 (5)
- ○ Image:stemoverlap6 (6)

End of Block: Default Question Block

Appendix 3. Survey three distributed to students three months following the workshop.

Journal of Science Education and Technology

Efficacy of "Computer Science in Modern Biology," a Peer-to-Peer Virtual Workshop to Improve Student Confidence

Long-Term Effectiveness of Computing in Modern Biology Workshop for Broadening Participation in STEM

Start of Block: Default Question Block

Q1 **Consent to Participate in Research at the University of Oklahoma**[OU-NC IRB Number: 12352 Approval Date: July 29th, 2020] You are invited to participate in research about **your participation in the STEM workshop.** If you agree to participate, you will **complete this online survey.** There are **no risks or benefits**. There **will be no compensation for participation.** Your participation is voluntary and your responses will be **confidential** After removing all identifiers, we might share your data with other researchers or use it in future research without obtaining additional consent from you. Even if you choose to participate now, you may stop participating at any time and for any reason. Data are collected via an online survey system that has its own privacy and security policies for keeping your information confidential. No assurance can be made as to their use of the data you provide.If you have questions about this research, please contact:**Sara Mata, Ph.D. email: sara.mata@ou.edu** You can also contact the University of Oklahoma – Norman Campus Institutional Review Board at 405-325-8110 or irb@ou.edu with questions, concerns or complaints about your rights as a research participant, or if you don't want to talk to the researcher. *Please print this document for*

your records. By providing information to the researcher(s), I am agreeing to participate in this research. Are you 18 years of age or older?

- ○ Yes (1)
- ○ No (2)

Skip To: Q36 If Consent to Participate in Research at the University of Oklahoma [OU-NC IRB Number: 12352 ... = Yes

Skip To: End of Survey If Consent to Participate in Research at the University of Oklahoma [OU-NC IRB Number: 12352 ... = No

Q36 We are not asking you for any personally identifying information, however we're asking three questions that will create an anonymous key to compare before and after surveys.

What is the first letter of the name of your favorite teacher?

▼ A (1) ... Z (33)

Q38 What is the day of the month on which you were born?

▼ 1 (1) ... 31 (31)

Q40 What is the second to last digit in your phone number?

▼ 1 (1) ... 0 (10)

Q61 Select the Lunch seminars you attended during the workshop (select all that apply)

- ☐ The Role of Computing in Biology (1)
- ☐ Careers in Biology (2)
- ☐ Imposter's Syndrome: What Is it and How to Cope? (3)
- ☐ Women in Science (5)
- ☐ None attended (7)

Q62 Did you attend the Biology Professionals mixer?

- ○ Yes (4)
- ○ No (5)

Q31 Have you followed-up or connected with anyone you met at the workshop?

- ○ Yes (1)
- ○ No (4)

Display This Question:

If Have you followed-up or connected with anyone you met at the workshop? = Yes

Q32 Please select anyone you have gotten in touch with

- ☐ Teaching Assistant (1)
- ☐ Instructor (2)
- ☐ Guest Lecturer (3)
- ☐ Mixer Contact (4)
- ☐ Peer (5)

Q33 Have you used any of the skills or resources you were given at the workshop?

- ○ Yes I've used resources from the workshop (1)
- ○ I haven't used any resources from the workshop since attending (2)

Display This Question:

If Have you used any of the skills or resources you were given at the workshop? = Yes I've used resources from the workshop

Q34 Please list the workshop resources you have used since attending:

_____

Q2 How comfortable are you downloading and installing software?

- Very comfortable (1)
- Somewhat comfortable (2)
- Neither comfortable nor uncomfortable (3)
- Somewhat uncomfortable (4)
- Very uncomfortable (5)

Q3 How confident are you in your coding ability or use of any programming/analysis and visualization software? If you have any coding experience, please respond based on the coding language you are most comfortable with?

- ○ Very comfortable (1)
- ○ Somewhat comfortable (2)
- ○ Neither comfortable nor uncomfortable (3)
- ○ Somewhat uncomfortable (4)
- ○ Very uncomfortable (5)



Q4 How many hours per week on average do you spend coding or using any type of programming/analysis and visualization software? For example, Excel, R, GIS, or python. Give your answer in numerical format (for example, 3)

_____

Q5 How comfortable do you feel finding or approaching a coding/programming/analysis and visualization project?

- ○ Very comfortable (1)
- ○ Somewhat comfortable (2)
- ○ Neither comfortable nor uncomfortable (3)
- ○ Somewhat uncomfortable (4)
- ○ Very uncomfortable (5)

Q6 How does your coding ability compare to that of your peers?

- ○ Much better (1)
- ○ Somewhat better (2)
- ○ About the same (3)
- ○ Somewhat worse (4)
- ○ Much worse (5)

Q7 Do you aspire to a career within biology or other science field?

- ○ Yes (1)
- ○ Maybe (2)
- ○ No (3)

Display This Question:

If Do you aspire to a career within biology or other science field? = Yes

Or Do you aspire to a career within biology or other science field? = Maybe

Q8 Please tell us your top 3 careers in biology or other science field that you aspire to (even if you're unsure)

_____

_____

_____

_____

_____

Q9 How important do you think coding or related computer software understanding is for your planned career goal?

- ○ Extremely important (1)
- ○ Very important (2)
- ○ Moderately important (3)
- ○ Slightly important (4)
- ○ Not at all important (5)

Q29 I often feel that I am “in over my head” or beyond my capabilities in my school work or study

- ○ Strongly agree (11)
- ○ Somewhat agree (12)
- ○ Neither agree nor disagree (13)
- ○ Somewhat disagree (14)
- ○ Strongly disagree (15)

Q30 I’m afraid people important to me may find out that I’m not as capable as they think I am.

- ○ Strongly agree (11)
- ○ Somewhat agree (12)
- ○ Neither agree nor disagree (13)
- ○ Somewhat disagree (14)
- ○ Strongly disagree (15)

Q10 Select the picture that best describes the current overlap of the image you have of yourself and your image of what a STEM professional is

- ○ Image:stemoverlap1 (1)
- ○ Image:stemoverlap2 (2)
- ○ Image:stemoverlap3 (3)
- ○ Image:stemoverlap4 (4)
- ○ Image:stemoverlap5 (5)
- ○ Image:stemoverlap6 (6)

---

*

Q40 What is your age (numerical answer e.g. 18, 21, etc.)

_____

Q42 With which gender do you identify?

- ○ Female (1)
- ○ Male (2)
- ○ Non-binary (3)
- ○ Genderqueer or gender nonconforming (4)
- ○ Transgender (5)
- ○ An identity not listed, self-identify (6) _____
- ○ I prefer not to answer (7)

Q44 With which race/ethnicity do you identify? (select all that apply)

- ☐ American Indian or Alaska Native (1)
- ☐ Asian American or Asian (2)
- ☐ Black or African American (3)
- ☐ Native Hawaiian or Pacific Islander (4)
- ☐ White (5)
- ☐ Middle Eastern (6)
- ☐ Multiracial (7)
- ☐ Hispanic or Latino (8)
- ☐ I prefer not to answer (9)
- ☐ We realize that the racial/ethnic category you selected encompasses many different nationalities. If you are interested in sharing more, please describe your nationality (i.e., Armenian, Puerto Rican, Vietnamese): (10)

_____

Q46 With which sexual orientation do you identify? (Select all that apply)

- ○ Asexual (1)
- ○ Bisexual (2)
- ○ Gay (3)
- ○ Heterosexual (4)
- ○ Lesbian (5)
- ○ Pansexual (6)
- ○ Queer (7)
- ○ Questioning or unsure (8)
- ○ An identity not listed, self identify: (9) _____
- ○ I prefer not to answer (10)

Q48 What is the highest level of education you have completed so far?

- ○ Middle school (1)
- ○ High school graduate (2)
- ○ Some college (3)
- ○ 2 year degree (4)
- ○ 4 year degree (5)
- ○ Masters degree (6)

Q50 What is the highest level of education completed by a parent or guardian in your household?:

- Middle school (1)
- High school graduate (2)
- Some college (3)
- 2 year degree (4)
- 4 year degree (5)
- Masters degree (6)
- PhD (7)
- I don't know (8)

Q52 Which social class group do you, as an individual, identify with?

- Poor (1)
- Working Class (2)
- Middle Class (3)
- Upper Class (4)
- I prefer not to answer (5)

Q54 What was your total household income during the past 12 months?

- Less than $25,000 (1)
- $25,000 - $34,999 (2)
- $35,000 - $49,999 (3)
- $50,000 - $74,999 (4)
- $75,000 - $99,999 (5)
- $100,000 - $149,999 (6)
- $150,000 or more (7)
- I don't know (8)
- I prefer not to answer (9)

Q56 Did you receive a Federal Pell Grant as part of your financial aid package?

- Yes (1)
- No (2)
- I don't know (3)

Q58 Please select your home living environment:

- Rural (1)
- Suburban (2)
- Urban (3)
- I don't know (4)

Q60 Growing up, was there at least one laptop or desktop computer in your home for you to use? If yes, how many?

- No, we didn't have a laptop or desktop computer in the house (1)
- 1 (2)
- 2 (3)
- 3+ (4)

End of Block: Default Question Block

Start of Block: Block 1

Q51 Thank you for completing the final survey for the Computer Science in Modern Biology Workshop study. Click on the following link to enter your name into the raffle: https://ousurvey.qualtrics.com/jfe/form/SV_9Yvf8dnfEKGJsax

End of Block: Block 1