Spring 2018

Assessing the impacts of native freshwater mussels on nitrogen cycling microbial communities using metagenomics

Ellen Marie Black
University of Iowa

ASSESSING THE IMPACTS OF NATIVE FRESHWATER MUSSELS ON NITROGEN CYCLING MICROBIAL COMMUNITIES USING METAGENOMICS

by

Ellen Marie Black

A thesis submitted in partial fulfillment
of the requirements for the Doctor of Philosophy
degree in Civil and Environmental Engineering in the
Graduate College of
The University of Iowa

May 2018

Thesis Supervisor: Assistant Professor Craig L. Just

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Ellen Marie Black

has been approved by the Examining Committee for
the thesis requirement for the Doctor of Philosophy degree
in Civil and Environmental Engineering at the May 2018 graduation.

Thesis Committee:

Craig L. Just, Thesis Supervisor

Gregory H. LeFevre

Jerald L. Schnoor

Michael S. Chimenti

Timothy E. Mattes

To my parents, who whole-heartedly supported my endeavors and encouraged my scientific curiosity at a young age.

"You cannot get through a single day without having an impact on the world around you. What you do makes a difference, and you have to decide what kind of difference you want to make."

-Jane Goodall

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my PhD mentor, Dr. Craig Just, who encouraged me to pursue unchartered research territory and inspired me to perform meaningful science. Craig first piqued my interest in the nitrogen cycle and water quality when I enrolled in his course, "Introduction to Sustainability", as an undergraduate student. Craig's passion for social and environmental quality was contagious and inspired me to attend graduate school, and eventually to discover a career where I could "save the world".

I'm thankful for Dr. Timothy Mattes and his research group for the countless guidance while I learned environmental microbiology research techniques. Xikun Liu and Yi Liang were essential in helping me troubleshoot qPCR and other experimental methods. Thank you to my additional committee members, Dr. Jerald Schnoor and Dr. Gregory LeFevre, for playing a critical role in the development of my thesis and helping me develop as a scientist, engineer, and scholar. I was supported emotionally and professionally by countless EES faculty and students during graduate school. I am honored to consider many of you friends and colleagues.

Next, I am extremely thankful for my collaborator, Dr. Michael Chimenti, for providing excellent computational support and for guiding my development of bioinformatics pipelines. Dr. Chimenti's bioinformatics work was essential for our amplicon and shotgun sequencing studies and paved the way for my independent bioinformatics research. My PhD thesis would not have been possible without you.

I received numerous sources of funding during my time in graduate school which enabled me to produce this thesis. I would like to thank the University of Iowa Water

Sustainability Initiative for funding the amplicon sequencing project, the University of Iowa Graduate College for supporting me with the Ballard and Seashore Dissertation Fellowship, and the Department of Civil and Environmental Engineering and NSF EPSCoR for supporting my stipend and research. I am grateful for the graduate teaching and research assistantships which for supporting me financially for four years and enabled me to devote my time to research and education.

Finally, I owe many thanks to my parents. Thank you for instilling in me a love for knowledge, a respect for nature, and a solid work ethic. Thank you for your endless love and support.

ABSTRACT

The Upper Mississippi River (UMR) basin contributes over 50,000 metric tons of nitrogen (N) to the Gulf of Mexico each year, resulting in a "dead zone" inhospitable to aquatic life. Land-applied N (fertilizer) in the corn-belt is attributed with a majority of the N-load reaching the Gulf and is difficult to treat as run-off is considered a non-point source of pollution (i.e. not from a pipe). One solution to this "grand challenge" of intercepting N pollution is utilizing filter-feeding organisms native to the UMR. Freshwater mussel (order Unionidae) assemblages collectively filter over 14 billion gallons of water, remove tons of biomass from overlying water, and sequester tons of N each day. Our previous research showed mussel excretions increased the sediment porewater concentrations of ammonium by 160%, and indirectly increased nitrate and nitrite by 40%, presumably from microbial degradation of ammonium. In response, the goal of this research was to characterize how mussels influenced microbial communities (microbiome) to determine the fate of N in UMR sediment.

First, we used qPCR and non-targeted amplicon sequencing within sediment layers to identify the N-cycling microbiome and characterized microbial community changes attributable to freshwater mussels. qPCR identified that anaerobic ammonium oxidizing (anammox) bacteria were increased by a factor of 2.2 at 3 cm below the water-sediment interface when mussels were present. Amplicon sequencing of sediment at depths relevant to mussel burrowing (3 and 5 cm) showed that mussel presence reduced microbial species richness and diversity and indicated that sediment below mussels harbored distinct microbial communities. Furthermore, mussels increased the abundance of ammonia oxidizing bacteria (family *Nitrosomonadaceae*), nitrite oxidizing bacteria

(genus *Nitrospira*), but decreased the abundance of ammonia oxidizing archaea (genus *Candidatus* Nitrososphaera), and microorganisms which couple denitrification with methane oxidation. These findings suggested that mussels may enhance microbial niches at the interface of oxic and anoxic conditions, presumably through excretion of N and burrowing activity.

In response, we performed metagenomic shotgun sequencing to identify which genes of the microbiome were most impacted by mussels. We hypothesized that genes responsible for ammonia and nitrite oxidation would be greater in the sediment with mussel assemblages. We found the largest abundance of N-cycling genes were responsible for nitrate reduction and nitrite oxidation, which is corroborated by the high concentration of nitrates in UMR water. Linear discriminant analysis statistical analyses showed nitrification genes were most impacted by mussels, and this presented an opposing effect on genes responsible for producing nitrous oxide, a potent greenhouse gas. Further investigation showed an increased abundance of a novel organism capable of completely oxidizing ammonia to nitrate (*Candidatus* Nitrospira inopinata) and coexisted with metabolically flexible *Nitrospira*, likely enhancing both carbon and N-cycling.

We demonstrated that native mussels harbor a unique niche for N-cycling microorganisms with large metabolic potentials to degrade mussel excretion products. Our findings suggest the ecosystem services of mussels extend beyond water filtration, and includes enhanced biogeochemical cycling of carbon, N, and reduces the potential for a potent microbially-produced greenhouse gas. Ultimately, this research could be used to advocate for mussel habitat restoration in the UMR to lessen the impacts of non-point pollution.

PUBLIC ABSTRACT

Native freshwater mussels of the Upper Mississippi River filter over 14 billion gallons of water, transfer tons of nitrogen from the overlying water into sediment, but few studies address how this "ecosystem engineer" impacts the surrounding microbial communities. In response, this research utilized metagenomic sequencing and bioinformatics techniques to determine how mussels influence sediment microbiomes with a focus on nitrogen-cycling microorganisms. Our results showed that mussels create a distinct sediment microbiome and increased the abundance nitrifiers, namely ammonia oxidizing bacteria and nitrite oxidizing bacteria. Regardless of mussel presence, microbiomes were equipped with large denitrification capacities which is likely a response to high nitrate concentrations in the river. Linear discriminant statistical analyses showed nitrification genes were most impacted by mussels, and this presented an opposing effect on genes responsible for producing nitrous oxide, a potent greenhouse gas. Further investigation showed an increased abundance of a novel organism capable of completely oxidizing ammonia to nitrate (known as comammox) and coexisted with metabolically flexible microorganisms likely enhancing both carbon and nitrogen cycling. We demonstrated that native mussels harbor a unique niche for nitrogen cycling microorganisms capable of degrading mussel excretion products. Our findings suggest the ecosystem services of mussels extend beyond water filtration, and includes enhanced microbial cycling of nitrogen and carbon, and reduces the potential for a potent microbially-produced greenhouse gas. Ultimately, this research could be used to advocate for mussel habitat restoration in the UMR to lessen the impacts of non-point pollution.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Introduction to nitrogen cycling

Nitrogen (N) is an essential element for life; it is required for nucleotide and protein synthesis, makes up 78% of the Earth's atmosphere, and is commonly a limiting nutrient in the environment [1]. Human activity greatly increased the supply of inorganic nitrogen since the Haber-Bosch process was invented in 1913 to meet the demand for nitrogen fertilizer to feed a growing population through ammonia-based fertilizer [1]. In 2014 alone, 144 million metric tons (Tg) of ammonia were produced world-wide for industrial and agricultural purposes [2]. A portion of ammonia fertilizer applied to land is nitrified to nitrate by microorganisms and becomes mobile in water. It was estimated that about one-third of applied nitrogen fertilizer is not harvested, where it affects water quality downstream [3]. Total flux of reactive nitrogen into global river systems in the 1990's was estimated at 118.1 Tg-N/yr and is predicted to reach 149.8 Tg-N/yr by 2050 [1]. The introduction of nitrogen into water bodies is a concern since it exerts an oxygen demand by microorganisms, known as nitrogenous oxygen demand.

Consequences of large N inputs, commonly referred to as eutrophication, includes enormous primary productivity of bacteria and algae. Once these organisms die, fall to bottom waters and decompose, they fuel microbial respiration which rapidly consumes oxygen. High primary productivity can result in water bodies containing dissolved oxygen (DO) concentrations less than 2.0 mg/L, often referred to as hypoxia, and results in unhospitable life for benthic organisms [4]. The Upper Mississippi River (UMR) basin is the dominant source of N that reaches the Gulf of Mexico [5] and resulted in a 22,720

km^2 "dead zone" in the Gulf of Mexico in 2017 [6]. A trend of increasing N loads reaching the Gulf stresses the need to reduce non-point N pollution to downstream waterways.

1.2 Introduction to freshwater mussels

North American native freshwater mussels, Unionoida, are suspension feeding bivalves and are vital ecosystem engineers that filter 14,000 million gallons of water and non-point pollution per day from the UMR [7], making their collective filtration capacity twice that of the 10 largest wastewater treatment plants world-wide [8]. A dense Unionid population of 8 mussels/m^2 can sequester 2 g carbon (C) each day and 200 mg nitrogen each day per m^2 of river sediment [9]. Freshwater mussel feces exerts a "bottom-up" nutrient control by filtering phytoplankton from the water column and depositing feces and pseudofeces (biodeposition) containing organic N and C into underlying sediment for consumption by benthic organisms [10].

Freshwater mussels have a large muscular foot that serves as an anchor and allows the animal to move vertically and horizontally through sediment substrates (**Figure 1.1**). Mussels enhance the transfer of nutrient-rich water and oxygen into sediment through burrowing activity (bioturbation), which creates a niche for N-cycling sediment microbial communities [11]. For example, researchers commonly measure a flux of nitrate (NO_3^-) from sediment into overlying water when mussels are present [12] and this is assumed to be a byproduct of nitrification occurring in the sediment [13]. Since it is now economically feasible to sequence millions of DNA copies in one day, we can accurately infer the N-cycling capabilities of a single organism and its entire microbial community by analyzing the genetic units responsible for N metabolism [14].

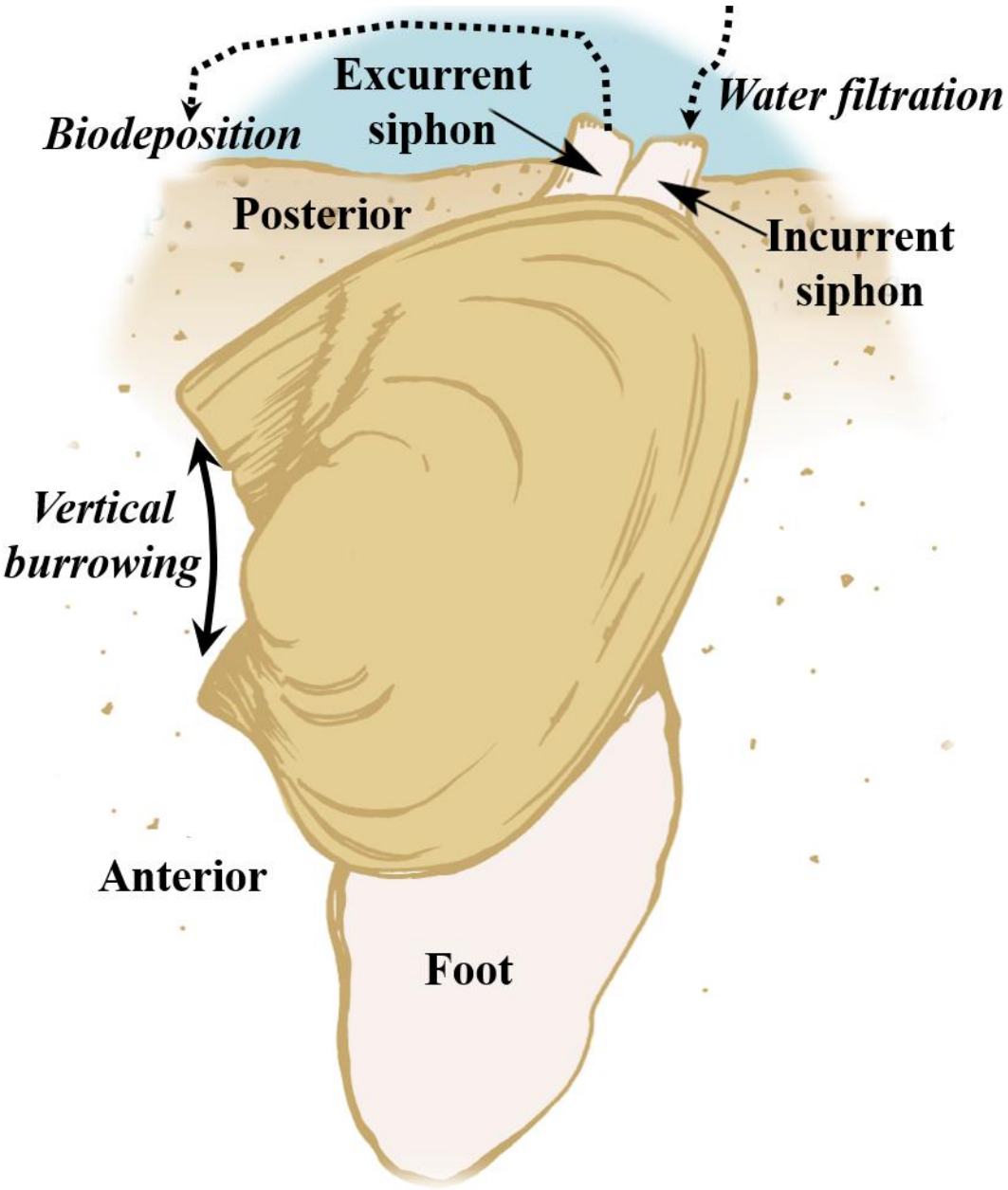


Figure 1.1 Native freshwater mussels use a muscular foot to burrow and move through sediment substrates. The posterior end of mussels lies near the sediment-water interface and allows the animal to filter phytoplankton from overlying water and deposit wastes as the sediment surface. (Adapted from Bruenderman, et al., 2002 [15].)

1.3 Introduction to N cycling microbiology

1.3.1 "Conventional" N cycling microorganisms

The scientific consensus once suggested that the microbial N cycle was controlled by two groups of microorganisms: nitrifiers and denitrifiers. The former group of organisms readily oxidize NH_4^+ to nitrite (NO_2^-) and nitrate (NO_3^-) by expressing the functional genes, ammonia monooxygenase (*amo*) and nitrite oxidoreductase (*nxr*), respectively [16]. In response, nitrates can be reduced sequentially to nitrogen gas by heterotrophic denitrifiers in anoxic conditions using a set of functional genes including nitrite reductase (*nir*), nitrate reductase (*nar* and *nap*), nitric oxide reductase (*nor*), and nitrous oxide reductase (*nos*) [17]. Additionally, organisms capable of dissimilatory nitrate reduction to ammonium (DNRA) present a competing reaction to denitrification, but do not significantly contribute to N flux in freshwater sediments, especially when NO_3^- is present in excess [18].

1.3.2 Anaerobic ammonium oxidation

In contrast to conventional N-cycling, a novel group of microorganisms are capable of anaerobically oxidizing ammonium (anammox) to nitrogen gas by coupling NO_2^- reduction and forming a reactive intermediate, hydrazine [19, 20]. Anammox bacteria are a major sink for N in aquatic environments and are readily competitive in anoxic micro-niches containing NH_4^+, NO_2^-, and NO_3^- [21]. Furthermore, two of the five genera of anammox bacteria (**Table 1.1**) can switch metabolism when NH_4^+ is limited in the environment and instead reduces NO_3^- or NO_2^- while utilizing organic acids as a source of electrons [22]. Since the discovery of anammox bacteria in the 1990's, researchers have increasingly identified anammox-induced N turnover as exceedingly

influential in freshwater and marine environments [20, 23]. Recent studies have identified a large presence and diversity of anammox in freshwater environments but few have investigated the biotic influences on these important organisms and the resulting implications for freshwater N cycling [24-28].

1.3.3 Complete ammonia oxidation

Most recently, genomic sequencing has confirmed the thermodynamic predictions of a "do-it-all nitrifier" in the *Nitrospira* genus [29]. These candidate *Nitrospira* species are genetically capable of completely oxidizing NH_3 (comammox) to NO_3^- by expressing both *amo* and *nxr* gene clusters [30, 31]. Successful isolation of the comammox species *Candidatus* Nitrospira inopinata revealed the organism encodes a unique Amo enzyme with the largest affinity for NH_3 measured to date [32]. In a similar example, *Nitrospira moscoviensis* was shown to be genetically capable of cleaving urea into NH_3 and carbon dioxide, and reciprocally feeding the NH_3 to urease-lacking nitrifiers and receiving NO_2^- in return [33]. Furthermore, *N. moscoviensis* and other nitrifying *Nitrospira* are metabolically flexible and competitive in fluctuating environmental conditions and is capable of utilizing cyanate as an energy source [34], aerobically oxidizing hydrogen [35], or even coupling formate or hydrogen oxidation to NO_3^- reduction [36]. Studying the metabolic potential of these unique microbes enables scientists to hypothesize about how the environment and microbes impact one another.

1.3.4 N-cycling microbial ecology

Environmental factors such as N, oxygen, and carbon (C), select for the N-cycling microbes with the most advantageous metabolic characteristics, in a process termed "niche partitioning" [37]. For example, biogeochemical conditions where carbon-nitrogen

(C/N) ratios are less than two enables anammox bacteria to predominate (85% of microbiome) in anoxic conditions (dissolved oxygen less than 0.05 mg/L) and are more favorable than denitrification reactions [38]. Additionally, resistance to low oxygen concentrations creates vertical partitioning between the nitrifying organisms, aerobic ammonia oxidizing bacteria (AOB) and archaea (AOA). AOA are equipped with high-oxygen affinity enzymes and predominate in deeper surface sediments, whereas AOB are highly competitive in aerobic sediments near the sediment-water interface [39].

The vertical stratification of microbial niches can be depicted in the form of chemical gradient profiles (**Figure 1.2**). For example, the consumption of mussel-derived NH_4^+ should correspond with the consumption of oxygen and the production of oxidized-N (NO_2^- and NO_3^-), but NH_4^+ would begin to accumulate at lower depths as nitrifiers become inhibited by anoxia. The co-occurrence of low NH_4^+ consumption and low, but not void, oxygen availability, results in a niche where NO_2^- and NO_3^- cannot be reduced by denitrifiers. This zone where NO_2^- accumulates corresponds with a steep decline in oxygen availability and the environment transitions from oxic to anoxic ("oxic-anoxic interface"). This oxic-anoxic interface is unique because it represents a niche for N-cycling microorganisms to compete for substrates, but also results in coexistence among N-oxidizers and N-reducers by being proximally close to the microbes producing their substrates. In response to potential impact mussels may have on this ecological partitioning, this research aimed to quantify the abundance of AOB, AOA, NOB, denitrifiers, comammox, and anammox bacteria (**Table 1.1**) because they are highly competitive in sediments at the interface of oxic and anoxic conditions (**Figure 1.2**) [40].

Table 1.1 Overview of N cycle microbiology, colloquial terms, and taxonomic examples.

Process	Acronyms	Generalized N Chemistry	Taxonomic Examples
Nitrogen fixation	Diazotroph	$N_{2\ (g)} \rightarrow NH_3$	Cyanobacteria, *Chlorobiaceae* (Green sulfur bacteria), *Azotobacter, Rhizobium, Spirillum, Frankia*
Aerobic ammonia oxidation	AOB AOA	$NH_4^+ \rightarrow NO_2^-$	*Nitrosomonas, Nitrosococcus, Nitrosospira, Nitrosopumilus, Nitrososphaera*
Nitrite oxidation	NOB	$NO_2^- \rightarrow NO_3^-$	*Nitrobacter, Nitrospina, Nitrococcus, Nitrospira*
Denitrification	Denitrifiers	$NO_3^- \rightarrow NO_2^- \rightarrow NO \rightarrow N_2O \rightarrow N_{2\ (g)}$	Species in over 50 genera including *Pseudomonas, Alkaligenes, Bacillus*
Anaerobic ammonium oxidation	Anammox	$NH_4^+ + NO_2^- \rightarrow N_{2\ (g)}$	"Candidate" genera (Planctomycetes phylum): *Anammoxoglobus, Kuenenia, Brocadia, Jettenia, Scalindua*
Complete ammonia oxidation	Comammox	$NH_4^+ \rightarrow NO_2^- \rightarrow NO_3^-$	"Candidate" species (*Nitrospira* genus): *Ca*. N. inopinata, *Ca*. N. nitrificans, *Ca*. N. nitrosa

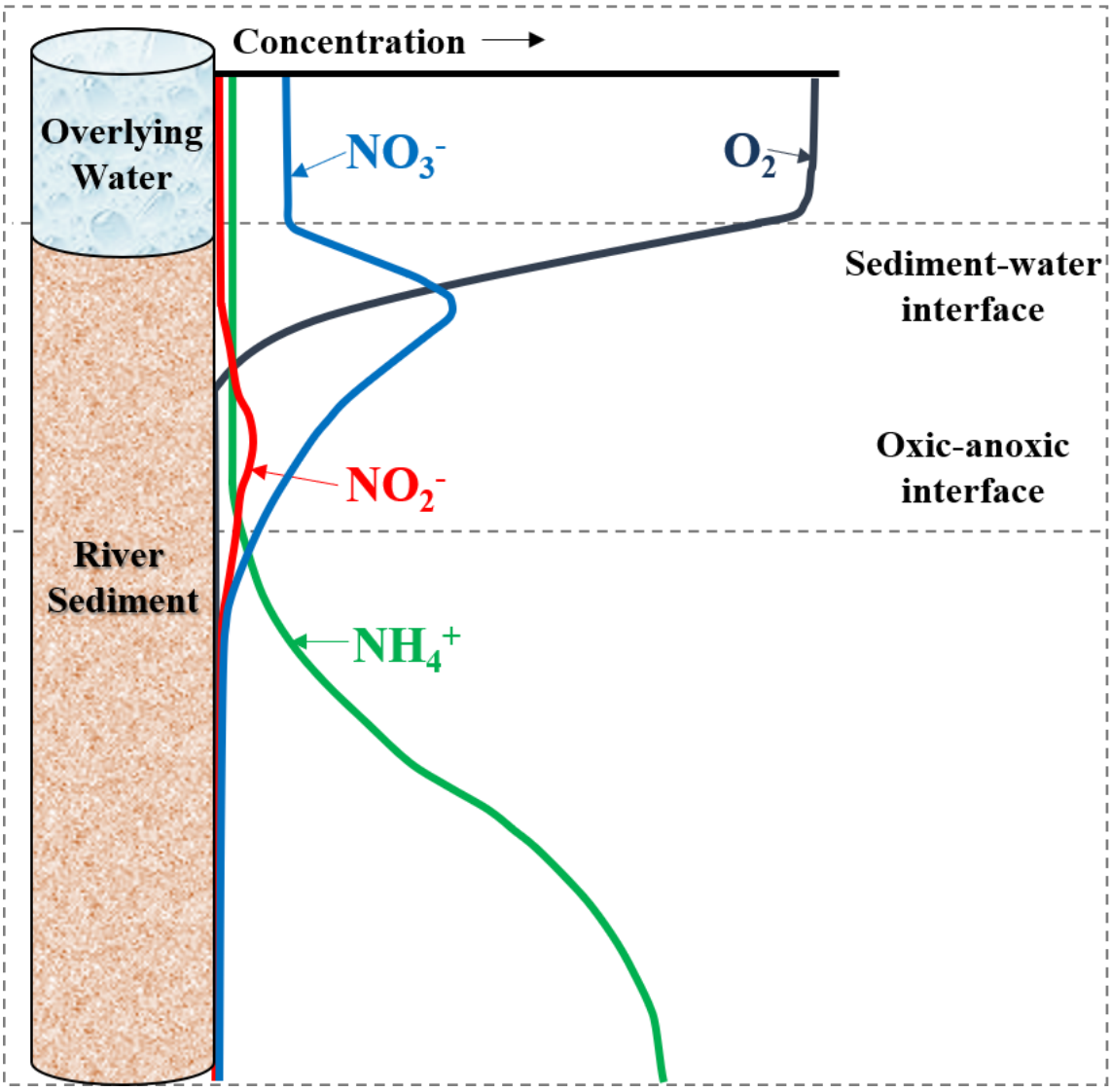


Figure 1.2 Vertical partitioning of inorganic N with sediment depth and an oxygen gradient. (Adapted from Ulloa, O. et al., 2012 [41].)

1.4 Microbial ecology research techniques

High-throughput sequencing allows researchers to identify all the genes present in a complex microbial community (i.e. metagenomics) [42]. The research in this dissertation used two variants of this technology to survey the UMR sediment microbiome (**Figure 1.3**). First, we sequenced 16S rRNA genes to identify the microbial community composition (Chapter 2). The genetic codes isolated from 16S rRNA are compared to a reference database to add a taxonomic identification to each group of similar sequences, referred to as "operational taxonomic units (OTUs)". As a result, we can identify the abundance of each OTU in relative to the whole microbial community and use ecological diversity metrices to classify the number of OTUs and community diversity. Although 16S rRNA sequencing accurately identifies taxonomy, researchers can only infer functions of taxons with known metabolisms.

Metabolic functions of a microbial community are determined through metagenomic shotgun sequencing; gene names and enzymatic functions are inferred when comparing the sequences to reference databases, such as KEGG or SEED (**Figure 1.3**). This research utilized metagenomics to find the relative abundances of protein-encoding genes and the metabolic pathways of sediment microbiomes (Chapters 3 and 4). Applying results from 16S rRNA sequencing allowed us to form hypothesis about how mussels impacted the microbiome's function. Ultimately, combining 16S rRNA sequencing and metagenomic sequencing enables us to specify which microbes are present and which metabolic functions the microbes are capable of performing.

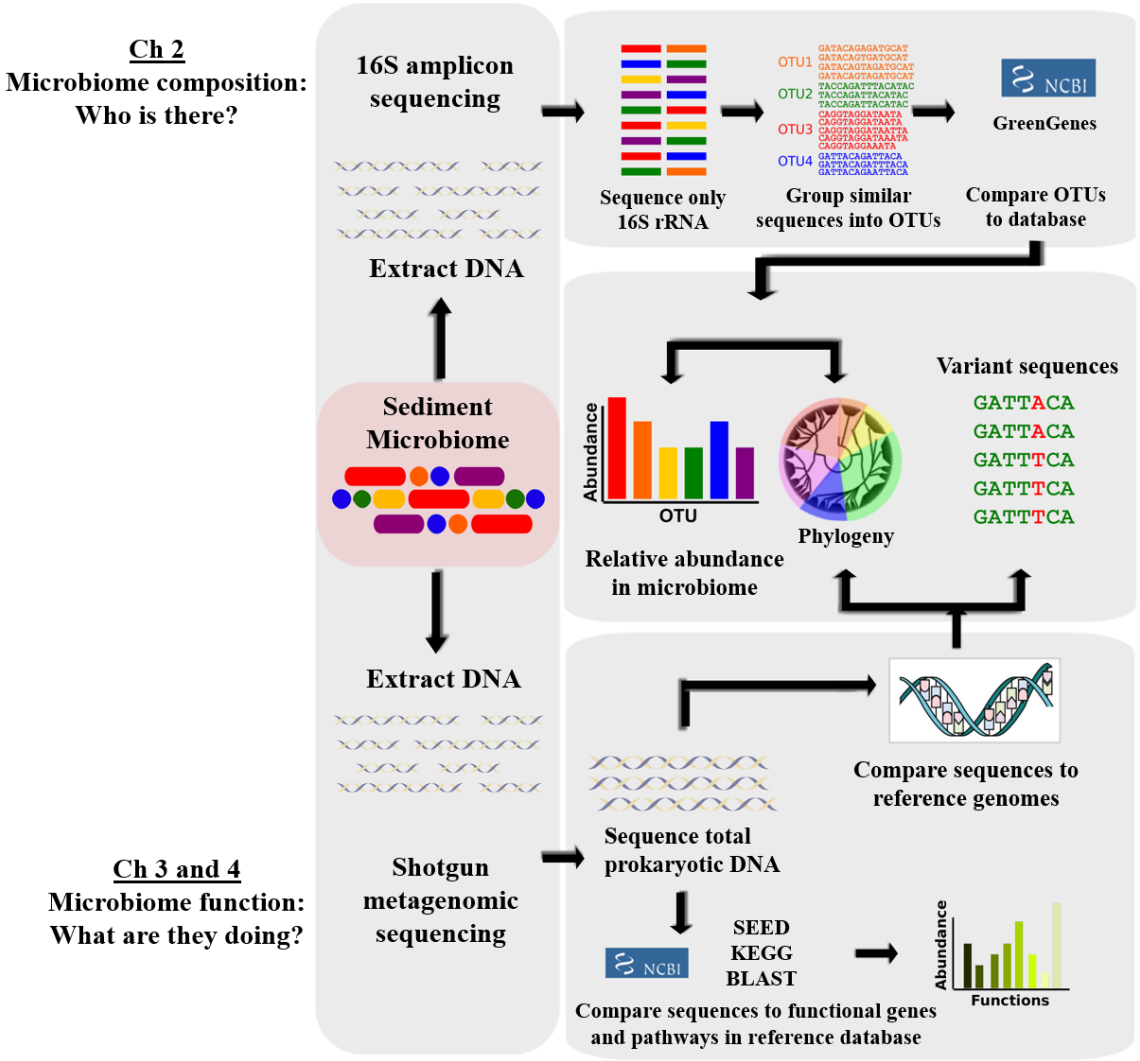


Figure 1.3 Molecular biology and bioinformatics techniques used to identify structure and function of microbial communities. (Adapted from Morgan, X.C. et al., 2012 [43].)

1.5 Thesis overview

Chapter 1 introduces the current knowledge of the N cycle, native freshwater mussel biology, and addresses relevant research techniques used in these fields. Chapter 2 is previously published research (https://doi.org/10.7717/peerj.3536) using qPCR and non-targeted amplicon sequencing to identify how anammox abundance and microbial communities were impacted by the presence of mussels in the UMR. The QIIME bioinformatics pipeline for 16S amplicon sequencing data and differential abundance analysis was performed by our collaborator, Dr. Michael Chimenti. Ellen Black performed sediment extraction and DNA isolation, qPCR, and analyzed statistical analyses for the N-cycling microbial community.

Chapters 3 and 4 utilized metagenomic sequencing to identify how mussel presence impacted the N-cycling potential of UMR sediment microbiomes and are currently being reviewed in two high-profile microbiology journals. Chapter 3 is an analysis of N-cycling functional genes and which N metabolic processes were most impacted by native freshwater mussels. Dr. Michael Chimenti was responsible for executing the metagenomic profiling of microbial pathways and gene family abundances. Ellen Black performed downstream analyses including manipulating HUMAnN2 output tables, LDA statistical analyses, and visualizing data.

Chapter 4 explores the metabolic features most responsible for differences in the N-cycling microbiome. In comparison to the gene-centric approach of Chapter 3, this manuscript focuses on the genomic features of NOB and commammox *Nitrospira*, as these organisms were shown to be greatly impacted by mussels in the preceding research (Chapter 2 and Chapter 3). Ellen Black performed taxonomic and functional microbiome

analyses, statistical analyses, genome reconstruction, and consulted with Dr. Chimenti for the initial pipeline development.

The thesis is concluded with Chapter 5, describing the engineering significance of this research, including preliminary work modeling the flux of N in a sediment microbiome with mussels. Finally, Chapter 5 summarizes future research directions and provides final conclusions drawn by this research.

1.6 Study site description

The mussel habitat study site chosen for this research (**Figure 1.4**) was informed by surveys detailing the historical mussel populations and long-term environmental research performed by the nearby Lucille A. Carver Mississippi River Environmental Research Station (LACMRERS) [44, 45]. Furthermore, the UMR basin is attributed as an N-rich agro-ecosystem [46, 47], potentially making the effects of mussels on a variety of N-cycling bacteria and archaea more pronounced than in any other freshwater environment. The pocketbook mussel (*Lampsilis cardium*) and threeridge mussel (*Amblema plicata*) comprise up to 38% and 56% of the mussel biomass in the UMR, respectively, and the mussel habitat in Navigation Pool 16 has consistent densities of 1.56 *Lampsilis cardium* per m^2 and 7.18 *Amblema plicata* per m^2, on average [45].

Figure 1.4 Study site in Navigation Pool 16 of the UMR depicting the mussel habitat (green) and upstream sediment (red) without mussels.

1.7 Objectives and hypotheses

1.7.1 Chapter 2: Mussels impact the vertical distribution of N-cycling microorganisms and the microbiome composition

Objective 1: Informed by previous research showing native freshwater mussels increased porewater concentrations of NH_4^+, NO_2^-, and NO_3^-, our first objective was to determine if the presence of mussels impacted the composition of the surrounding sediment microbial communities. Since the previously mentioned N-species co-occur in oxic-anoxic interfaces, another objective was to quantify the vertical distribution of anaerobic NH_4^+ oxidizing organisms as an indication of this mussel-influenced niche.

Hypothesis 1: The presence of freshwater mussels increases the abundance of NH_4^+ oxidizing microorganisms and NO_2^- oxidizing microorganisms.

1.7.2 Chapter 3: Mussels increase the abundance of nitrification genes

Objective 2: Since previous work found mussels increase porewater nitrification products, NO_2^- and NO_3^-, the second objective was to compare the abundance of N-cycling functional genes from sediment with and without mussels.

Hypothesis 2: The presence of mussels corresponds to an increased relative abundance of genes-encoding nitrification reactions, namely ammonia monooxygenase (Amo) and nitrite oxidoreductase (Nxr).

1.7.3 Chapter 4: Mussels influence the genomic potentials of NOB and comammox *Nitrospira*

Objective 3: The third objective was to determine if the abundance of functional genes in NOB *Nitrospira* and comammox-*Nitrospira* were increased with mussels. This

finding would provide further evidence that mussels increase the potential for nitrification, but also would suggest which metabolic functions were responsible for enhanced *Nitrospira* with mussels.

Hypothesis 3: The presence of mussels corresponds with an enhanced genomic potential for nitrification metabolism in *Nitrospira* species.

CHAPTER 2: EFFECT OF FRESHWATER MUSSELS ON THE VERTICAL DISTRIBUTION OF ANAEROBIC AMMONIA OXIDIZERS AND OTHER NITROGEN-TRANSFORMING MICROORGANISMS IN UPPER MISSISSIPPI RIVER SEDIMENT[1]

2.1 Abstract

Targeted qPCR and non-targeted amplicon sequencing of 16S rRNA genes within sediment layers identified the anaerobic ammonium oxidation (anammox) niche and characterized microbial community changes attributable to freshwater mussels. Anammox bacteria abundance did not differ from a normal distribution (Shapiro-Wilk normality test, W-statistic=0.954, p=0.773) between 1-15 cm depth and were increased by a factor of 2.2 ($p<0.001$) at 3 cm below the water-sediment interface when mussels were present. Amplicon sequencing of sediment at depths relevant to mussel burrowing (3 and 5 cm) showed that mussel presence reduced observed species richness (p=0.005), Chao1 diversity (p=0.005), and Shannon diversity ($p<0.001$), with more pronounced decreases at 5 cm depth. A non-metric, multidimensional scaling model showed that intersample microbial species diversity varied as a function of mussel presence, indicating that sediment below mussels harbored distinct microbial communities. Mussel presence corresponded with a 4-fold decrease in a majority of operational taxonomic units (OTUs) classified in the phyla Gemmatimonadetes, Actinobacteria, Acidobacteria, Plantomycetes, Chloroflexi, Firmicutes, Crenarcheota, and Verrucomicrobia. 38 OTUs in

[1] A version of this chapter has been published: 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. http://doi.org/10.7717/peerj.3536

the phylum Nitrospirae were differentially abundant ($p<0.001$) with mussels, resulting in an overall increase from 25% to 35%.

Nitrogen (N)-cycle OTUs significantly impacted by mussels belonged to anammmox genus *Candidatus* Brocadia, ammonium oxidizing bacteria family Nitrosomonadaceae, ammonium oxidizing archaea genus *Candidatus* Nitrososphaera, nitrite oxidizing bacteria in genus *Nitrospira*, and nitrate- and nitrite-dependent anaerobic methane oxidizing organisms in the archaeal family "ANME-2D" and bacterial phylum "NC10", respectively. Nitrosomonadaceae (0.9-fold ($p<0.001$)) increased with mussels, while NC10 (2.1-fold ($p<0.001$)), ANME-2D (1.8-fold ($p<0.001$)), and *Candidatus* Nitrososphaera (1.5-fold ($p<0.001$)) decreased with mussels. Co-occurrence of 2-fold increases in *Candidatus* Brocadia and *Nitrospira* in shallow sediments suggests that mussels may enhance microbial niches at the interface of oxic-anoxic conditions, presumably through biodeposition and burrowing. Furthermore, it is likely that the niches of *Candidatus* Nitrososphaera and N-DAMO were suppressed by mussel biodeposition and sediment aeration, as these phylotypes require low ammonium concentrations and anoxic conditions, respectively. As far as we know, this is the first study to characterize freshwater mussel impacts on microbial diversity and the vertical distribution of N-cycle microorganisms in upper Mississippi river sediment. These findings advance our understanding of ecosystem services provided by mussels and their impact on aquatic biogeochemical N-cycling.

2.2 Introduction

Native freshwater mussels (Order Unionida) are ecosystem engineers that significantly alter benthic habitats through biodeposition of feces and pseudofeces, rich in

ammonium (NH_4^+) and organic carbon (C), into sediment [48]. The estimated mussel filtration capacity in a 480 km, Upper Mississippi River (UMR) segment, as a percentage of river discharge, is up to 1.4% at high flows, up to 4.4% at moderate flows and up to 12.2% during low flows [45]. The mussels in this river segment collectively filter over 14 billion gallons of water, remove tons of biomass from the overlying water, and deposit tons of reduced C and nitrogen (N) at the water-sediment interface each day [10]. The pocketbook mussel (*Lampsilis cardium*) and threeridge mussel (*Amblema plicata*) comprise up to 38% and 56% of the mussel biomass in the UMR, respectively [7]. A habitat near Buffalo, Iowa, in UMR Pool 16, had mean densities of 1.56 *L. cardium*-m^{-2} and 7.18 *A. plicata*-m^{-2} that correlated with fine sediment diameters (d_{50}=0.300 ± 0.121 mm) which were presumably influenced by mussel burrowing [49]. Mussels live primarily buried in sediment, with their posterior end often flush with the sediment surface [50], or slightly below the surface in soft sediments [51-53]. This positions adult freshwater mussels 6-10 cm into the sediment with tendencies toward more shallow burrowing during the spring and summer [54]. Extensive observations in the UMR concluded that *A. plicata* were often found with portions of their shell above the water-sediment interface, while *L. cardium* burrow a few cm into the sediment during the summer [55]. Additionally, *A. plicata* often burrowed up to 2.5 cm vertically [53] in response to stressors while *L. cardium* moved more horizontally when stressed [55]. Two common stressors, that happen to be created by the mussels themselves, are low dissolved oxygen (DO) and elevated ammonia (NH_3) and NH_4^+ [50, 56]. We hypothesize that this frequent vertical and horizontal movement by mussels, many times as an indirect

and/or direct response to their own waste production, has a significant impact on porewater chemistry and microbiology in UMR sediments.

The evidence for freshwater mussel impacts on aquatic chemistry is compelling, especially for nutrients. A dense mussel population can sequester 2 g C day^{-1} m^{-2}, 200 mg N day^{-1} m^{-2}, and 50 mg phosphorus day^{-1} m^{-2} from river water into sediment [9]. During the summer months, biodeposition-derived N from mussels was roughly 67% NH_4^+, 28% amino acids, and 5% urea [57]. Mussel biodeposition accounted for up to 40% of total N demand in freshwaters and up to 74% of N in the food web, but was sometimes dampened [58] in high nutrient environments [58]. Our previous work showed mussel burrowing and biodeposition, just below the water-sediment interface, increased porewater NH_4^+, nitrate (NO_3^-), nitrite (NO_2^-), and total organic C concentrations by 160%, 38%, 40%, and 26%, respectively [13, 59]. But, the experimental design of our previous work limited our ability to assess the effects of mussels on the broad microbial community that was transforming N simultaneously and, quite likely, synergistically.

The UMR is an N-rich agro-ecosystem [46, 47, 60-62] shown to foster high microbial N transformations [63, 64] potentially making the effects of mussels on a variety of N-transforming bacteria and archaea more pronounced than in any other freshwater environment. The first step in transforming biologically active N is nitrification by aerobic ammonium oxidizing bacteria (AOB) (**Figure 2.1**, yellow arrows), such as the genera *Nitrosomonas* and *Nitrosospira* in the Nitrosomonadaceae family [65-68], and aerobic ammonium oxidizing archaea (AOA) in multiple candidate genera. AOB and AOA are metabolically diverse [69] and serve a functionally important role of catalyzing the rate limiting step of nitrification [70] in various freshwater niches.

For example, *Candidatus* Nitrososphaera (phylum Thaumarchaetoa), a group of thermophilic AOA [71], and AOB species in genera *Nitrosospira* and *Nitrosococcus* [72] can use urea as an alternative source of NH_4^+ [73]. AOA often outnumber AOB [74] due to their ability to grow at NH_4^+ concentrations below 10 nM [70], compared to 10 µM for some AOB species [75]. In the second step of nitrification, nitrite oxidizing bacteria (NOB), such as *Nitrospira* (phylum Nitrospirae), and *Nitrobacter*, *Nitrococcus*, and *Nitrospina* (phylum Proteobacteria) [68], aerobically oxidize NO_2^- to NO_3^- (**Figure 2.1**, yellow arrows).

Nitrospira are the most abundant and diverse group of NOB and dominate numerous habitats, ranging from freshwater sediment to engineered wastewater treatment plants [33, 76, 77]. Furthermore, NOB species *Nitrospira moscoviensis* and *Nitrospira lenta* can derive NH_4^+ from urea hydrolysis, provide NH_4^+ to AOB, and subsequently may oxidize NO_2^- from AOB in a process deemed "reciprocal feeding of nitrifiers" [33, 76]. *Candidatus* Nitrospira inopinata, can also use urea as an alternative NH_4^+ source [76] and has all the genes necessary for complete ammonia oxidation (comammox) to NO_3^- [30] (**Figure 2.1**, yellow curved arrow). Denitrifiers complete the conventional N-cycle by sequentially reducing NO_3^- to nitric oxide (NO), nitrous oxide (N_2O), and nitrogen gas in anoxic and high C environments (**Figure 2.1**, blue arrows) [65].

A specialized group of bacteria in the phylum Planctomycetes, anaerobic ammonium oxidizing (anammox) bacteria, oxidize NH_4^+, utilize NO_2^- as a terminal electron acceptor, and produce N_2 gas [19, 20] (**Figure 2.1**, gray arrows). Anammox bacteria thrive at the interface of oxic-anoxic conditions due to dependence on NO_2^- production by AOB or AOA [78]. Anammox and NOB compete for NO_2^- in low

substrate environments, and this is especially true for *Nitrospira* NOB, which share a homologous form of the key enzyme catalyzing NO_2^- oxidation with anammox [77]. In another example, *N. moscoviensis* can adapt to a range of oxygen concentrations by coupling formate oxidation and NO_3^- reduction [33]. Recently, an N-cycling enrichment culture revealed comammox bacteria co-occurring with anammox bacteria in the genus *Candidatus* Brocadia, presumably enhanced by the ability of commamox organisms to oxidize NH_4^+ in low oxygen conditions (<3.1μM) [30]. *Nitrospira* species, *N. moscoviensis* and *Candidatus* Nitrospira inopinata in particular, are examples of NOB which harbor a unique ability to assist or compete with anammox for N-substrate in a variety of niches [77].

Shallow sediments also pose a competitive niche for anammox bacteria because of high NH_4^+ fluxes into oxic sediment and NO_2^- limitations from denitrification [78, 79]. Another addition to the suite of known N-transformations includes prokaryotic coupling of anaerobic oxidation of methane with denitrification [80]. In nitrite- and nitrate-dependent anaerobic methane (CH_4) oxidation (N-DAMO) [78, 80, 81], NO_3^- reduction to NO_2^- and NO_2^- reduction to N_2 are coupled with CH_4 oxidation to CO_2 (**Figure 2.1**, gray line and curved arrows). Nitrate-DAMO biochemical processes have been linked to family "ANME-2D" [82, 83], while nitrite-DAMO was discovered for *Candidatus* Methylomirabilis oxyfera [84] in phylum "NC10" [85-87], and both are widespread in anoxic freshwater sediments [82, 85, 88, 89].

Mollusks have been shown to influence the diversity of microbial communities and abundance of N-transforming microorganisms. For example, metagenomic profiling revealed a marine California mussel (*Mytilus californianus*) shell provided a niche for N-

and C-transforming microorganism populations [90], and a restored oyster reef enhanced nitrification and denitrification rates greater than 10-fold [91]. Furthermore, an experimental microcosm study reported enhanced prokaryotic metabolic activity and diversity following a biodeposition rate of 10 g m^{-2} d^{-1} of mussel feces and pseudofeces [92]. Additionally, clusters of the zebra mussel (*Dreissena polymorpha*) in a lake increased heterotrophic bacteria density, activity, and diversity [93]. Since the impact of native freshwater mussels on prokaryotic diversity and abundance in the UMR is largely unknown, this study utilized targeted and non-targeted sequencing of the 16S rRNA gene to determine how N-transforming microorganisms and microbial community structure differs in sediments with mussels compared to sediments without mussels.

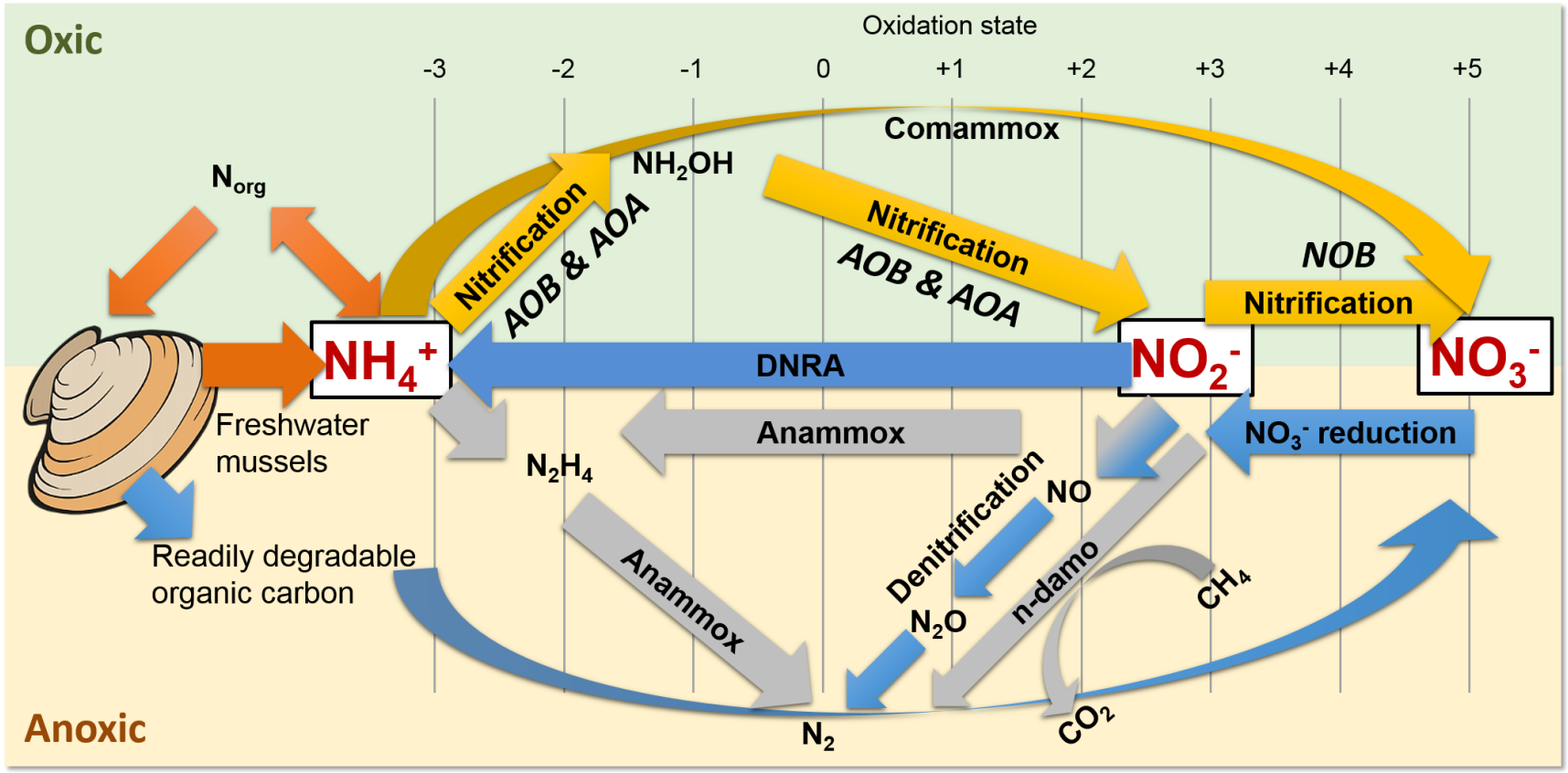


Figure 2.1 Freshwater mussels deposit feces and pseudofeces containing nitrogen and carbon at the water-sediment interface (i.e. oxic-anoxic transition). NH_4^+ resulting from mussel biodeposits may be oxidized via nitrification and comammox in oxic conditions (yellow arrows), and/or by anammox and n-damo near the oxic-anoxic interface (gray arrows). Oxidized nitrogen species (NO_2^- and NO_3^-) may be reduced by dissimilatory nitrate reduction to ammonium (DNRA) and denitrification pathways (blue).

2.3 Materials and Methods

UMR sediments were collected within a dense, well-characterized mussel assemblage in the Buffalo Habitat of UMR Pool 16 [49] (41.452804, -90.763299), and from a slightly up-river location with no mussels (41.451540, -90.753275) using a 3-inch diameter, hammer-driven, acrylic tube (Batch 1 samples) or a 2-inch diameter, post-driver sediment sampler with a polypropylene liner (Multi-Stage Sediment Sampler, Art's Manufacturing and Supply, Inc., American Falls, ID, USA; Batch 2 samples). Batch 1 sediment was used to identify the vertical distribution of anammox bacteria below freshwater mussels. For Batch 1, the acrylic tube for each core (n=3 with-mussels) was penetrated at 1, 3, 5, 7, 11, and 15 cm sediment depths with a 3/8th-inch diameter, ethanol flame-sterilized drill bit to enable sediment collection. In comparison, Batch 2 sediment was used to characterize anammox abundance, microbial diversity, and community structure in shallow sediments below mussels. For Batch 2, the polypropylene liner for each sediment core (n=5 with-mussels, n=5 no-mussels) was penetrated at depths of 3 cm and 5 cm. Sediment was sampled for DNA isolation (in quadruplicate) for a combined sample size of n=20 for 3 cm depth with-mussels, n=20 for 5 cm depth with-mussels, n=20 for 3 cm depth without mussels, and n=20 for 5 cm depth without mussels. Genomic DNA was isolated from 0.25 g of each sediment sample (PowerSoil® DNA Isolation Kit, MoBio Laboratories, Inc., Carlsbad, CA, USA) and stored at -20°C. Batch 2 Genomic DNA was used for anammox-targeted qPCR (n=20 for each treatment) and 16S rRNA gene amplicon sequencing (n=10 for each treatment).

2.3.1 Anammox 16S rRNA gene quantification

Microbial culture from a sidestream deammonification process (Hampton Roads Sanitation District, Virginia Beach, VA) served as a source of anammox genetic material for qPCR standard curve construction. PCR products (primers A483f (5'-GTCRGGAGTTADGAAATG-3') and A684r (5'-ACCAGAAGTTCCACTCTC-3') [16]) of the anammox 16S rRNA gene was purified with Qiaquick PCR purification Kit (Qiagen Inc., Valencia, CA, USA), and cloned into the pCR 2.1-TOPO® vector using the TOPO® TA cloning Kit (Invitrogen Corp., Carlsbad, CA, USA). Clones were Sanger sequenced at the University of Iowa Institute of Human Genetics with M13F (5'-TGTAAAACGACGGCCAGT-3') and M13R (5'-CAGGAAACAGCTATGAC-3') primers to ensure anammox 16S rRNA PCR products were inserted into the vector. Nucleotide sequences were aligned using the Standard Nucleotide Basic Local Alignment Search Tool [94] (GenBank Accession: KU047953) and classified as *Candidatus* Brocadiales (of the Planctomycetes phylum) with a 95% confidence threshold using RDP Naïve Bayesian rRNA Classifier Version 2.10 [95]. Plasmid DNA concentration was quantified with Qubit® Fluorometer 1.0 (Thermo Fisher Scientific, Inc.), serially diluted, and used to construct qPCR calibration curves.

The anammox 16S rRNA gene from batches 1 and 2 was quantified [96] with qPCR using QuantStudio™ 7 Flex Real-Time PCR System (Thermo Fisher Scientific, Inc., Waltham, MA, USA) with primers A483f and A684r [16] and analyzed with QuantStudio™ Real-Time PCR Software (Thermo Fisher Scientific, Inc.). The threshold cycle (C_t) curves were satisfactory (slope=-3.374, Y-int=36.702, R^2=0.998, and amplification efficiency=97.99%), and PCR product dissociation curves revealed singe

peaks centered at a melting temperature of 83°C. The statistical significance of 16S rRNA gene copies was determined via a one-way, repeated measures analysis of variance (ANOVA) (SigmaPlot 13.0, Systat Software, Inc., Chicago, IL, USA) between the 4 treatment groups (n=20) following a passed normality test (p=0.826, Shapiro-Wilk) and an equal variance test (p=0.073, Brown-Forsythe). Pairwise multiple comparison procedures were completed via the Holm-Sidak method with a significance level of 0.050 and a power of 0.990.

2.3.2 Non-targeted amplicon sequencing of the 16S rRNA gene

Batch 2 genomic DNA (20 µL, 1-50 ng/µL) was analyzed by the Argonne National Laboratory, Environmental Sample Preparation and Sequencing Facility (ESPSF) utilizing the Earth Microbiome Project protocol (http://www.earthmicrobiome.org/emp-standard-protocols/16s/). All samples were analyzed together in one batch. The v4 region of prokaryotic 16S rRNA gene (515F-806R) was amplified using the following conditions: 3 minutes at 94°C, 35 cycles of 94°C for 45 seconds, 50°C for 60 seconds, and 72°C for 90 seconds, followed by 10 minutes at 72°C [97]. The PCR mixture consisted of 13.0 µL PCR grade water, 10.0 µL 5 PRIME HotMasterMix (Quanta Biosciences, Beverly, MA), 1.0 µL genomic DNA, and 0.5 µL forward and reverse primers (10 µM). 16S rRNA gene amplicon libraries were sequenced by ESPSF using Illumina MiSeq paired end reads (2x151 bp) [97] and uploaded to MG-RAST (ID's: 4705672.3-4705709.3) and NCBI (BioProject ID PRJNA374585).

Determining the operational taxonomic units ("OTUs") in each sample from the raw 16S rRNA gene amplicon reads was accomplished using the default Quantitative

Insights into Microbial Ecology (QIIME) open-reference pipeline [98]. Briefly, the QIIME open-reference pipeline takes paired-end reads as input, which are then joined, demultiplexed, filtered, and clustered into OTUs with uclust [99]. Representative sequences from each cluster were aligned [100] to GreenGenes 13.5 reference database [101] with a 97% similarity threshold. RDP classifier [95] was used for taxonomy assignment, PyNAST [100] was used for multiple sequence alignment. Phylogenetic trees were constructed using FastTree2.1.3 with default settings [102]. The OTU table from QIIME open reference picking ('otu_table_mc2_w_tax_no_pynast_failures_json.biom' in the standard QIIME workflow) was imported into R using the phyloseq package [103] for downstream analysis, along with the corresponding phylogenetic tree ('rep_set.tre') and a metadata mapping file. These datasets were merged to create a single 'physeq' object representing the experiment. Alpha-diversity was calculated on the unfiltered OTU abundance data using the Observed species, Chao1 [104], and Shannon [105] metrics. Beta-diversity was calculated using a matrix of bray-curtis [106] intersample distances and ordination plots calculated with non-metric multidimensional scaling (NMDS). Differential abundance analysis was carried out using the DESeq2 [107] R package with default settings (test type was "Wald," fit type was "parametric"). Translating physeq objects into a compatible DESeq2 object was performed with the "phyloseq_to_deseq2" function. The complete data analysis R script can be downloaded from the public github repository:

https://github.com/mchimenti/black_chimenti_just_phyloseq/blob/master/phyloseq.r.

Analysis at the OTU level provided a fine scale resolution for significant differences in microbial ecology between mussel and no mussel treatments. To put these

results into a biological context, the genus-level OTU file was used to compare relative abundances for N-cycle phylotypes. These groups include AOA genus *Candidatus* Nitrososphaera, nitrate-damo family "ANME-2D", NOB genus *Nitrospira*, anammox genus *Candidatus* Brocadia, AOB family Nitrosomonadaceae, and nitrite-DAMO phylum "NC10". Relative abundance counts for each N-cycle group was tested for statistical significance between treatments, using metadata groups "3 cm with-mussels" (n=10), "5 cm with-mussels" (n=10), "3 cm no-mussels" (n=10), and "5 cm no-mussels" (n=10). 1-way ANOVA's of each N-cycle group was performed using the Kruskal-Wallis test ($p<0.05$) with Dunn's multiple correction test ($Padj<0.05$) (GraphPad Prism 7.0, La Jolla, CA). Similarly, multiple comparisons were made between all N-cycle phylotype groups and their respective treatments (n=10); significant differences between relative abundances were tested using the Kruskal-Wallis test ($P<0.0001$) and Dunn's multiple comparison test ($Padj<0.05$).

2.4 Results

2.4.1 Anammox-targeted 16S rRNA gene quantification

The targeted 16S rRNA gene data from Batch 1 (n=3, with-mussels) indicated an anammox bacterial gene copy maximum (~$3x10^5$ copies g^{-1} sediment) between 3 cm and 7 cm sediment depth in the presence of mussels (**Figure 2.2A**). The Batch 1 data did not diverge from a normal distribution between 1 cm and 15 cm (Shapiro-Wilk normality test, W-statistic=0.954, p=0.773). Only one sediment core went beyond 7 cm leaving anammox bacterial gene copy data at 11 cm and 15 cm without replicates. The Batch 2 data (n=20 for 3 cm with-mussels, n=20 for 5 cm with-mussels, n=20 for 3 cm no-mussels, n=20 for 5 cm no-mussels) showed that anammox bacteria experienced a 2.2-

fold increase ($p<0.001$) at 3 cm with-mussels compared to the no-mussels control (**Figure 2.2B**). The anammox gene copies measured at 5 cm were statistically indistinguishable between the with-mussels and no-mussels treatments.

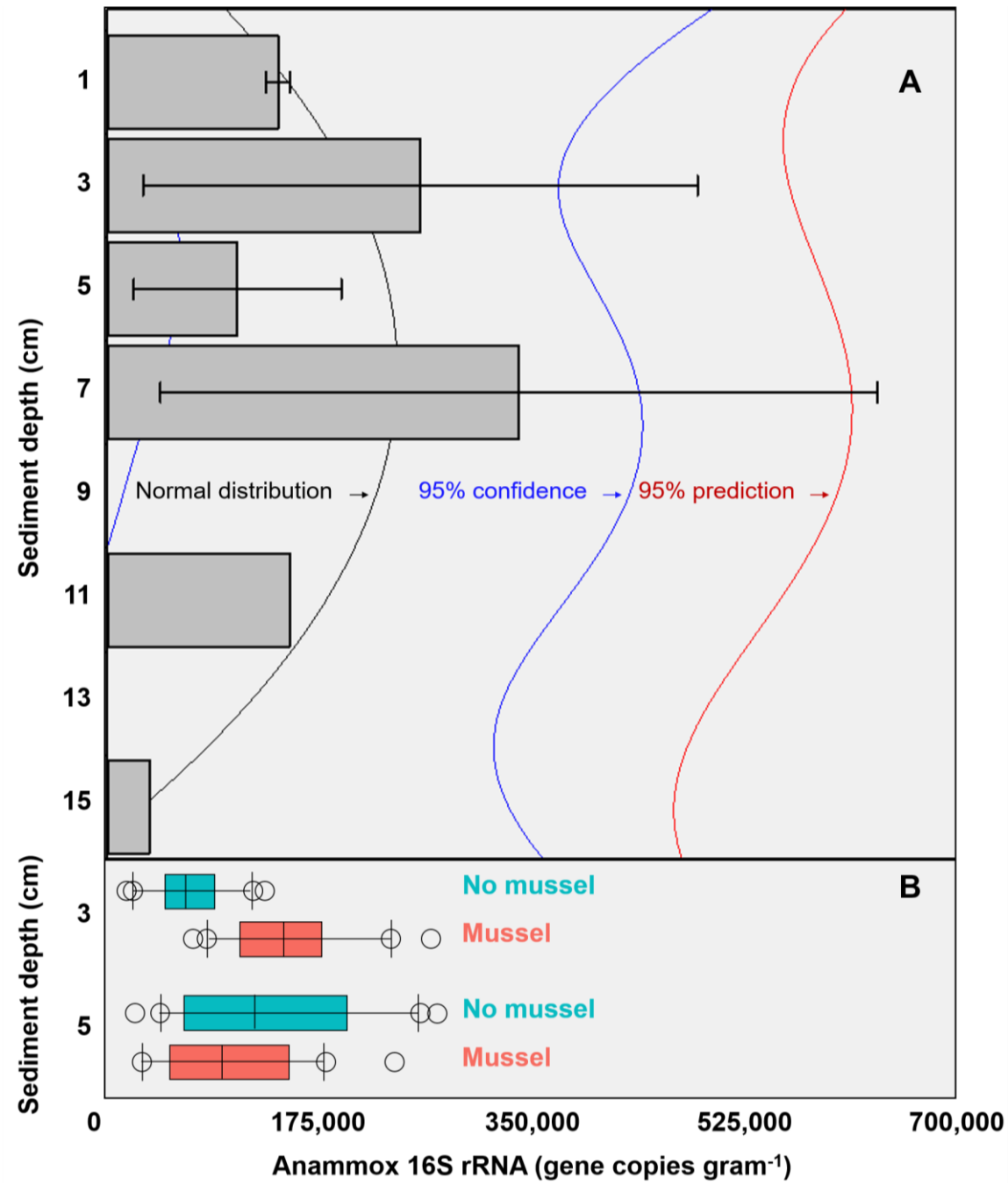


Figure 2.2 A) The mean anammox 16S rRNA gene copies (per gram of sediment) in the presence of mussels were normally distributed (Shapiro-Wilk normality test, W-statistic=0.954, p=0.773) with depth (Batch 1 data). Error bars represent 1 standard deviation from the mean. **B)** Mussels (salmon-colored data) significantly increased the anammox 16S rRNA gene copies at 3 cm depth ($p<0.001$; Batch 2 data). The anammox gene copies were statistically indistinguishable with mussels as compared to the no mussels (turquoise-colored data) sediments at 5 cm (Batch 2 data). The outer most open circles in Figure 2B represent data outliers, box boundaries represent the 25th and 75th percentile, the line within the box is the median, and error bars indicate 10th and 90th percentiles.

2.4.2 Non-targeted sequencing of the 16S rRNA gene

Summing across all samples, a total of 2,103,661 amplicon sequences were analyzed and about 76,000 unique OTUs were reported by QIIME. Of the unique OTUs, 18,777 had 10 or more reads and 3,916 OTUs had counts exceeding 100 reads. Mussel bed samples had read counts of 45,290 (±15,271) at 3 cm sediment depth and 52,451 (±7,044) at 5 cm sediment depth, while no-mussel samples had 48,920 (±7,517) read counts at 3 cm depth and 63,706 (±25,379) at 5 cm sediment depth (read depths depicted in **Figure 2.3**). The top phyla in mussel bed sediments were Proteobacteria (40.7%), Nitrospirae (35.2%), Chloroflexi (5.9 %), Euryarchaeota (5.0%), Chlorobi (4.2%), and Bacteroidetes (2.3%). Proteobacteria decreased by about 6% with mussels while Nitrospirae increased by 10% with mussels. The most abundant taxonomic families in the Nitrospirae phylum were *Thermodesulfovibrionaceae* (55%), "FW" (33%), and *Nitrospiraceae* (13%), and were 5% less, 3% and 2 % greater than in no-mussel samples, respectively. With mussels, Proteobacteria taxonomic classes consisted of the following proportions: 68% Deltaproteobacteria (8% less than without-mussels), 16% Gammaproteobacteria, and 15% Betaproteobacteria. A majority of these Deltaproteobacteria OTUs were from "BPC076", Desulfarculales, and Syntrophobacterales taxanomic orders, while orders Burkholderiales and Xanthomonodales made up a majority of Betaproteobacteria and Gammaproteobacteria taxons.

Species richness was analyzed using three common measures: Observed species, Chao1 and Shannon indices (n=20 with-mussels and n=20 without mussels). Together, the three measures indicated a decrease in microbial community richness and evenness in

the presence of mussels as compared to sediments without mussels (**Figure 2.4**). The observed decrease in alpha-diversity reached significance for each of the three measures tested ($p=0.0054$ or lower). A similar result was obtained when calculating alpha-diversity measures in samples exclusively from 3 cm ($n=10$) or exclusively from 5 cm ($n=10$) depths in the presence and absence of mussels. However, the decrease in richness was more pronounced at 5 cm (**Figure 2.5**) than at 3 cm depth (**Figure 2.6**).

To compare intersample diversity in species abundances and community composition ("beta diversity"), we employed NMDS scaling to accurately visualize, in 2D space, the higher-order community structure between with-mussels and no-mussels samples (**Figure 2.7**). The NMDS model produced an excellent representation of the bray-curtis distances for all samples (convergence in 20 iterations, stress ~ 0.06; shepard plot shown in **Figure 2.8**). The beta diversity clearly differentiated as a function of mussel presence, but not sediment depth (**Figure 2.7**). Taken together, these data show that mussel presence had a pronounced influence on the microbial community evenness, richness, and composition within the sediment.

Differential abundances in OTUs did not reach significance for metadata values of sediment depth or comparisons between sediment cores. On the other hand, there were numerous differences in OTU abundances when comparing sediment with mussels and without mussels. We performed a differential abundance estimation with the DESeq2 R package using mussel presence status ($n=20$ with-mussels, $n=20$ no-mussels) as our covariate. 734 OTUs (or 0.94% of the 77,288 OTUs tested) reached significance with a false discovery rate of 0.01. The vast majority of OTUs belonging to the phyla Gemmatimonadetes, Actinobacteria, Acidobacteria, Plantomycetes, Chloroflexi,

Firmicutes, Crenarcheota, and Verrucomicrobia decreased by at least 4-fold in the presence of mussels. In contrast, Proteobacteria showed a marked decrease in order Alphaproteobacteria, while showing mixed increasing and decreasing OTUs among Beta-, Delta-, and Gammaproteobacteria. Phylum Nitrospirae also had 38 OTUs which were differentially abundant with p-adj<0.001. OTUs assigned to the GreenGenes taxonomic family of "0319-6A21" were the most abundant among those OTUs increasing without mussels, while families *Thermodesulfovibrionaceae* and "*FW*" were most abundant among those OTUs increasing with mussels.

Many of the Nitrospirae taxons that increased without mussels did so from a smaller average abundance (17 average counts for *Nitrospira* and up to 126 average counts for *Thermodesulfovibrionaceae*) relative to those that were increased with mussels (209 average counts for *Nitrospira* and up to 581 average counts for *Thermodesulfovibrionaceae*). This explains the 10% increase in Nitrospirae abundance when summing across all samples with mussels. **Figure 2.9** shows the Log2FC categorized by phyla for OTUs with p-adj<0.0001 (to enhance visual clarity). Significant differences within the Nitrospirae phylum were represented by increases of genus "*HB118*" in family *Thermodesulfovibrionaceae* (2.0Log2FC from a mean count of 52, p<0.001) and unclassified *Nitrospira* species (0.8Log2FC from an average count of 209, p<0.001) with mussels. No-mussel treatments showed increases in genus "*LCP-6*" from family *Thermodesulfovibrionaceae* (3.6Log2FC from an average count of 126, p<0.001) and unclassified *Nitrospira* species (2.1Log2FC from an average count of 17, p<0.001).

Despite seemingly even representation of phylum Thaumarchaeota between treatments, unclassified species from *Candidatus* Nitrososphaera were enhanced from an

average abundance of 126 (1.73Log2FC, $p<0.001$) without mussels, and AOA species, *Candidatus* Nitrososphaera gargensis increased from an average count of 16 (2.85Log2FC, $p<0.001$) without mussels. One OTU classified in the anammox genus, *Candidatus* Brocadia, increased from an average count of 17 (3.72Log2FC, $p<0.001$) without mussels, while another OTU classified as an unknown *Candidatus* Brocadia species increased from a mean count of 16 (1.2Log2FC, $p=0.001$) with-mussels. Furthermore, OTUs belonging to the AOB family *Nitrosomonadaceae* increased from an average abundance of 6 (1.9Log2FC, $p<0.001$) with mussels. Without mussels, taxonomic groups capable of nitrite-DAMO, phylum "NC10", increased from average abundances up to 130 (4.4 Log2FC, $p<0.001$), and nitrate-DAMO family "ANME-2D" increased from average abundances up to 59 (3.4 Log2FC, $p<0.001$). A summary of Log2FC values for OTUs relevant to N-transformations are listed for no-mussel (**Table 2.1**) and with mussel treatments (**Table 2.2**).

N-cycle phylotypes were examined for statistically significant relative abundances between treatments of mussel presence and sediment depth (**Table 2.3**). *Candidatus* Nitrososphaera experienced a 2.6-fold decrease ($p=0.047$) with mussels at 5 cm sediment depth. ANME-2D was 3 times greater ($p=0.049$) at 5 cm sediment depth without mussels, compared to 3 cm sediment depth without mussels. Within the mussel bed, *Nitrospira* were 1.7 times greater ($p=0.0497$) at 3 cm depth, and experienced a 1.9-fold increase ($p=0.025$) with mussels at 3 cm sediment depth versus control. *Candidatus* Brocadia was 3 times greater ($p=0.013$) at 5 cm depth without mussels versus 3 cm without mussels, and the 3 cm sediment showed a 2-fold increase ($p=0.002$) with mussels versus control.

Nitrosomonadaceae was 2.7 times greater ($p=0.015$) at 3 cm with mussels versus 5 cm depth with mussels.

Relative abundances of N-cycle phylotypes were compared within each treatment (**Figure 2.10, Boxes A, B, D, E**) and between treatments (**Figure 2.10, Boxes C, F, G-I**). Within 3 cm sediment samples with mussels (**Figure 2.10, Box A**), *Nitrospira* was statistically greater in abundance than *Candidatus* Brocadia, and ANME-2D was less abundant than *Nitrospira*. Sediment without mussels at 3 cm depth (**Figure 2.10, Box B**) contained statistically greater abundances of *Candidatus* Nitrososphaera than *Candidatus* Brocadia, and greater *Nitrospira* abundances compared to *Candidatus* Brocadia, *Nitrosomonadaceae*, and ANME-2D.

Relative abundance comparisons between mussel and no-mussel treatments at 3 cm depth (**Figure 2.10, Box C**) showed that *Candidatus* Nitrososphaera was reduced in the mussel treatment, while *Nitrospira* and *Candidatus* Brocadia were enhanced with mussels. Within mussel sediment samples at 5 cm depth, *Nitrospira* was more abundant than *Candidatus* Brocadia, ANME-2D, and Nitrosomonadaceae. (**Figure 2.10, Box D**). On the other hand, *Candidatus* Nitrososphaera and *Nitrospira* were both more abundant than Nitrosomonadaceae without mussels at 5 cm sediment depth (**Figure 2.10, Box E**). Comparing microbial communities at 5 cm depth between mussel and no-mussel treatments (**Figure 2.10, Box F**) revealed that *Candidatus* Nitrososphaera was less abundant with mussels versus the no-mussel population. *Nitrospira* and Nitrosomonadaceae phylotypes were more prominent with mussels in shallow sediment depths (**Figure 2.10, Box G**).

Overall, *Nitrospira* made up larger proportions of microbial communities with and without mussels compared to many N-cycle organisms, especially *Candidatus* Brocadia and Nitrosomonadaceae (**Figure 2.10, Box C, F, and G-I**). Without mussels at 3 cm sediment depth, *Candidatus* Brocadia made up a smaller proportion of the N-cycling microbial community, especially when compared to *Candidatus* Nitrososphaera, ANME-2D, NC10, and *Nitrospira* in deeper sediments (**Figure 2.10, Box H**).

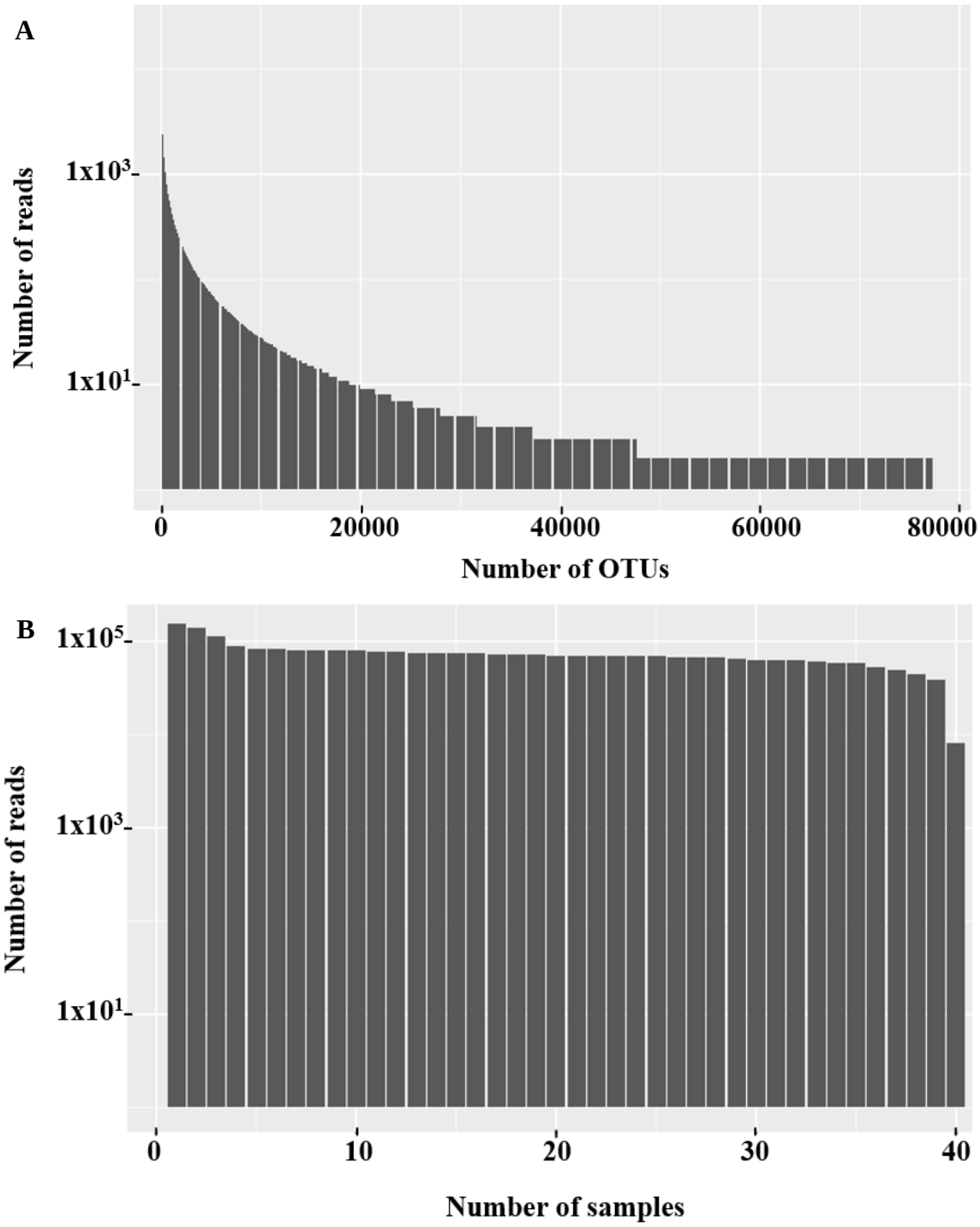


Figure 2.3 Number of reads (y-axis; log-scale) by rank-ordered OTU (**A**) and by sample (**B**) in this experiment. There were ~55,000 OTUs with fewer than 10 reads and the sample read depths were relatively consistent.

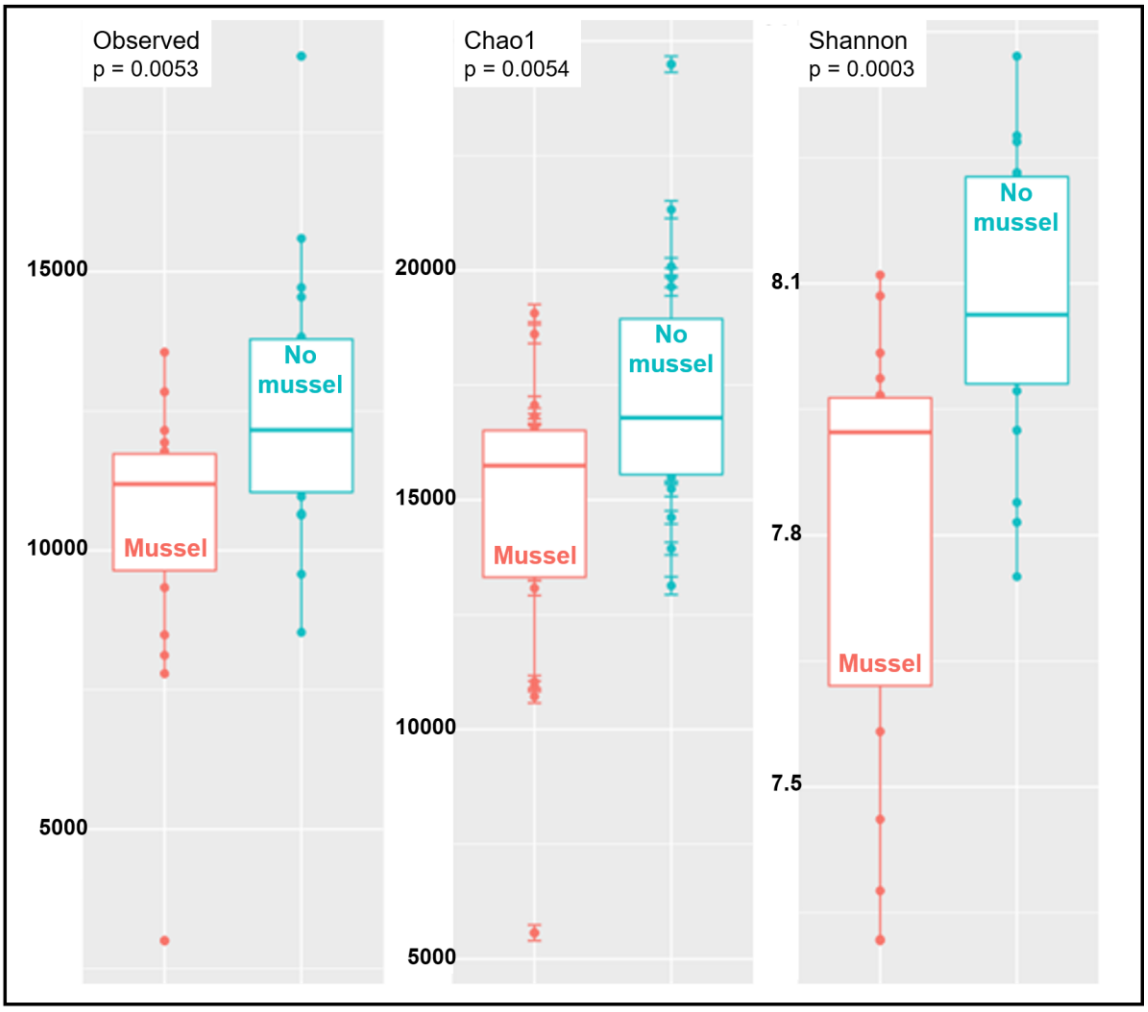


Figure 2.4 Sediments with mussels have lower observed species richness (p=0.005), Chao1 diversity (p=0.005), and Shannon (p=0.0003) diversity than no-mussel sediments.

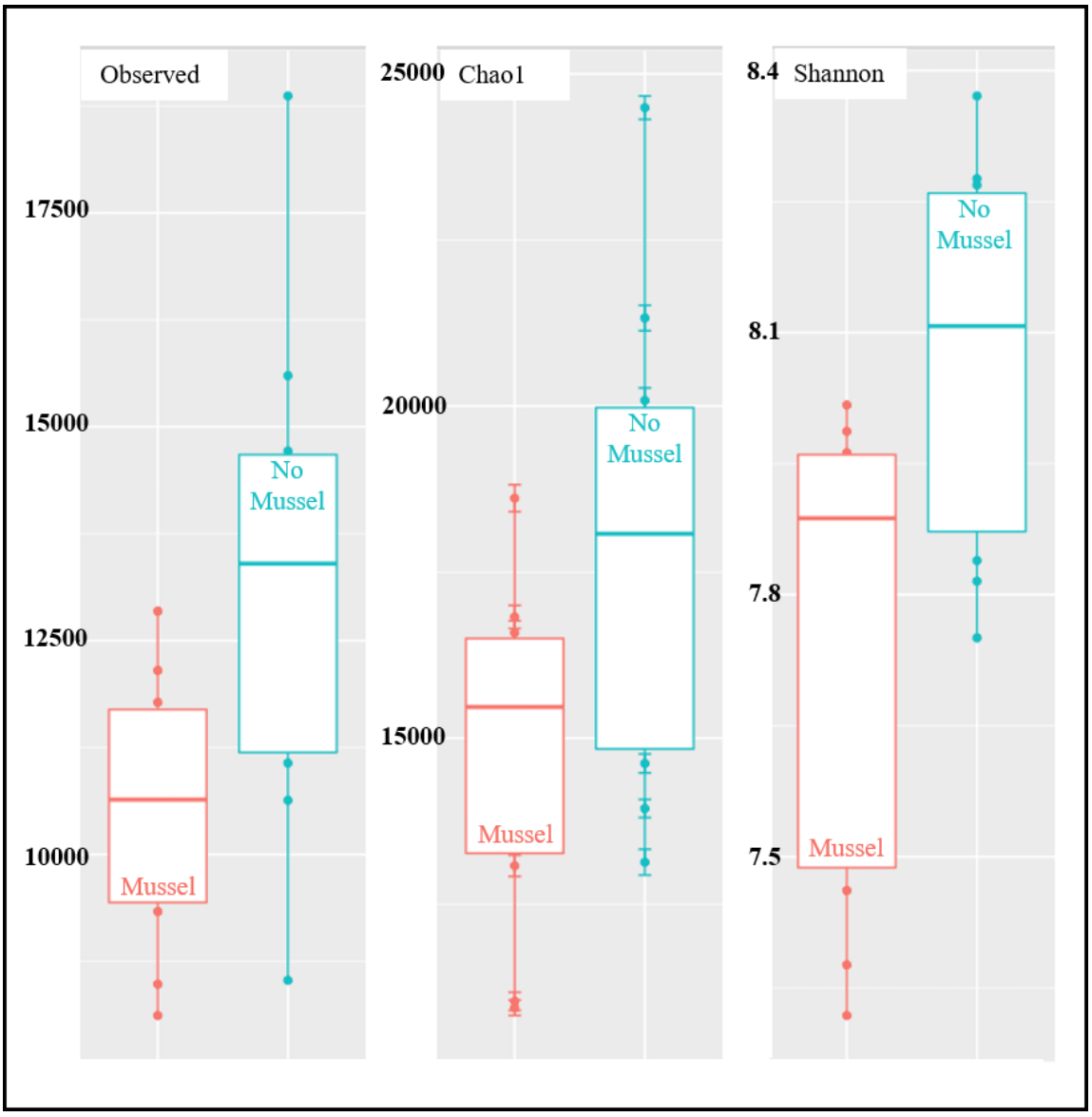


Figure 2.5 Alpha diversity at 5 cm depth in with-mussels and no-mussels samples.

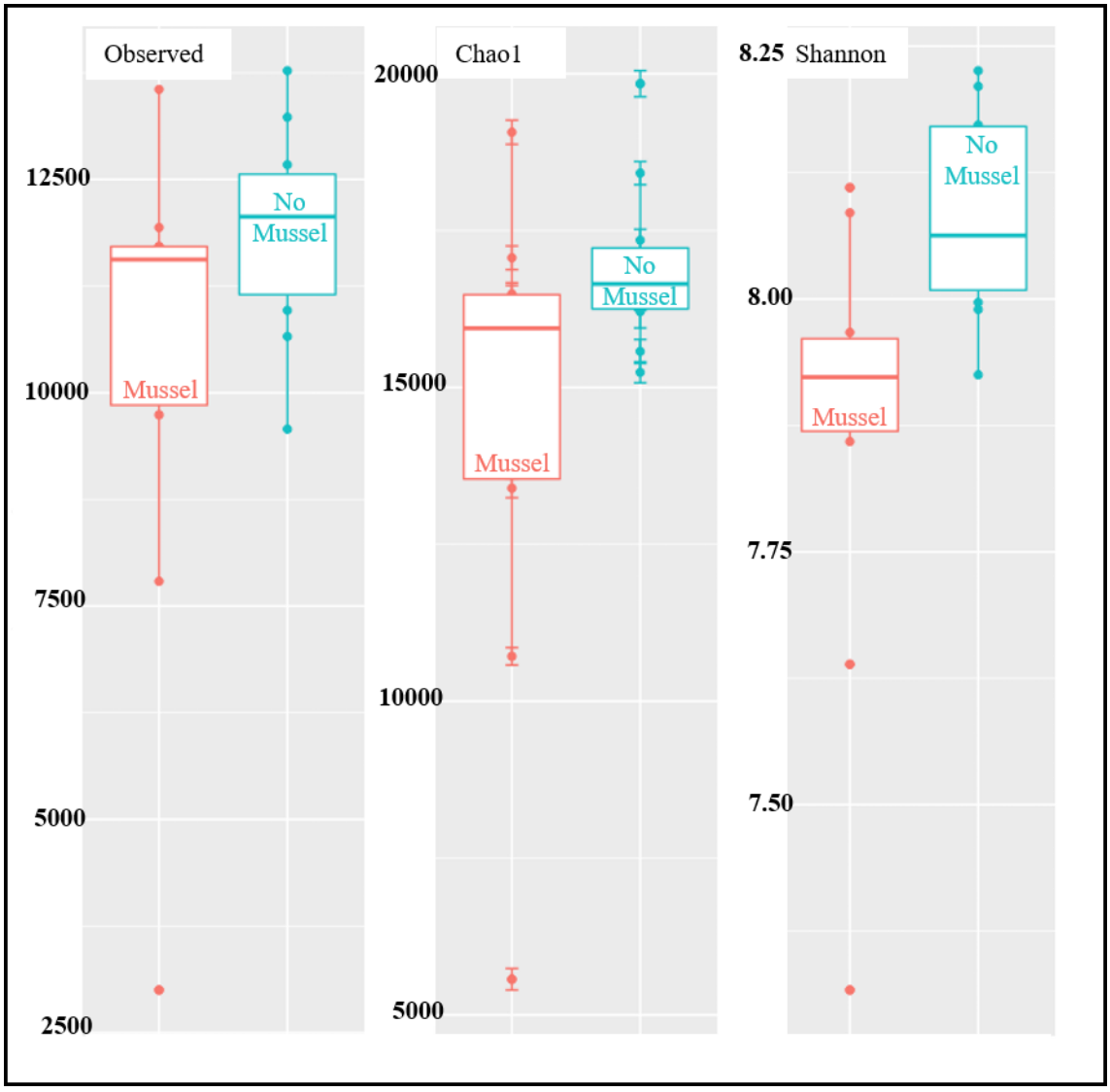


Figure 2.6 Alpha diversity at 3 cm depth in with-mussels and no-mussels samples.

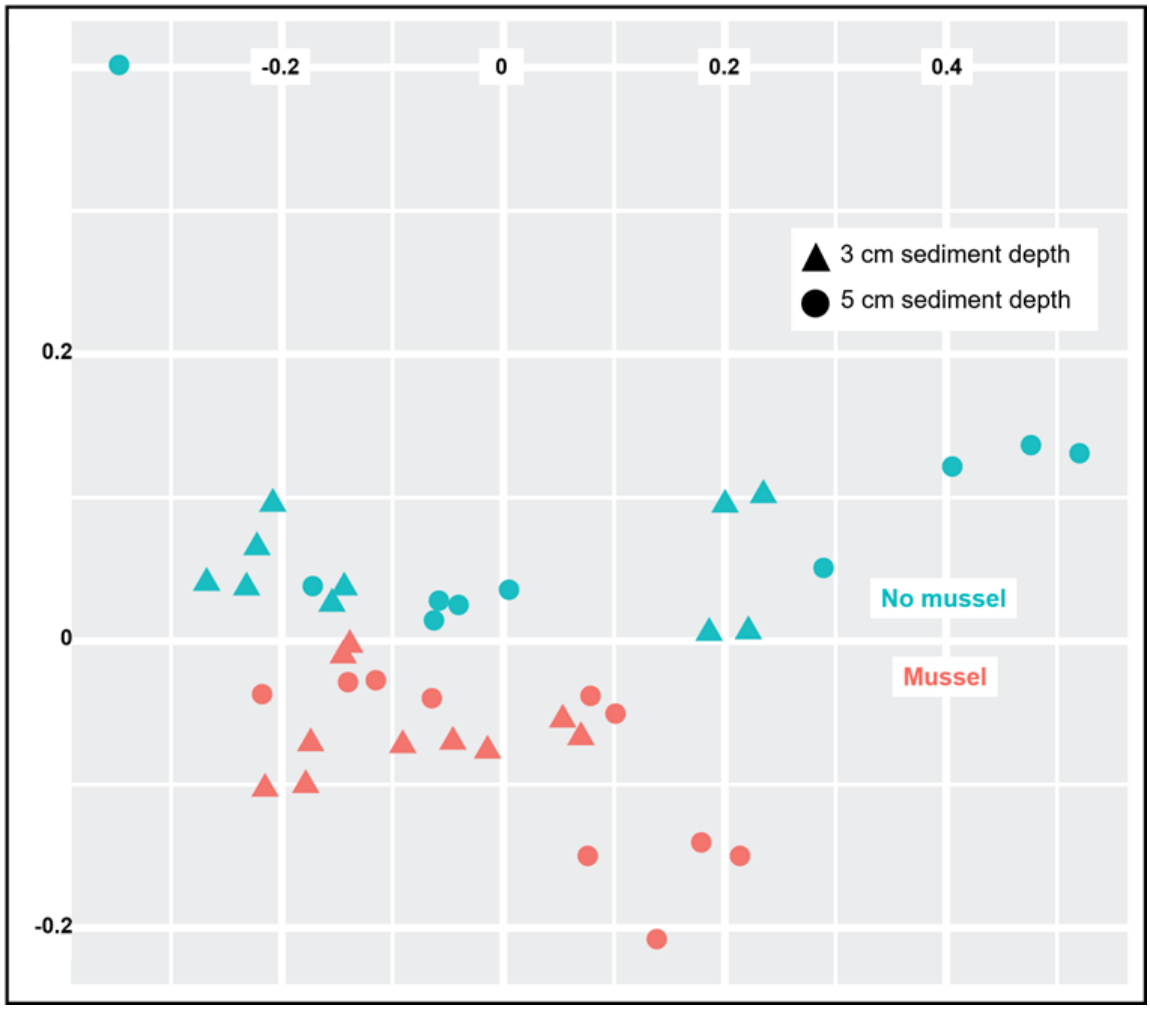


Figure 2.7 NMDS analysis using bray-curtis distances revealed sample clustering as a function of mussel presence, but not sediment depth.

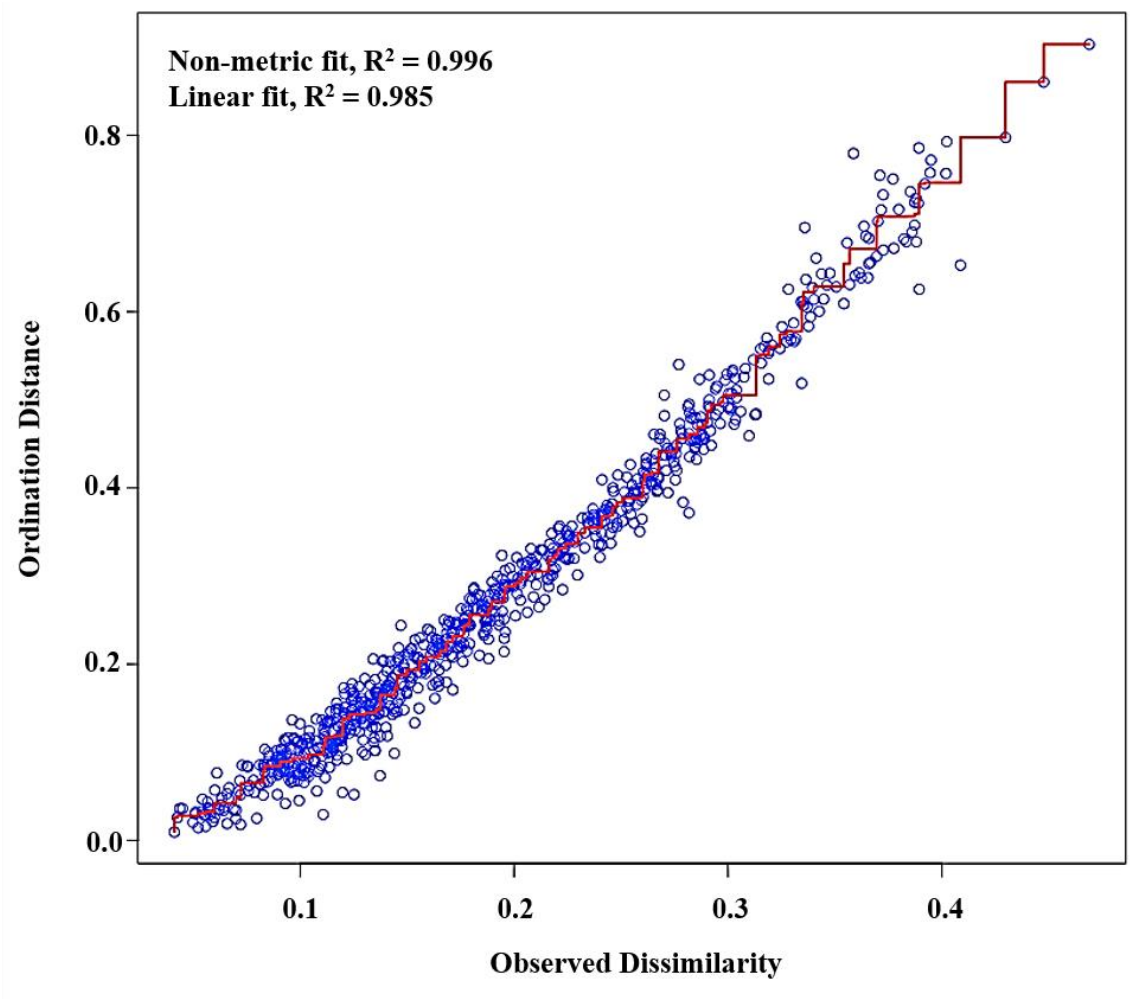


Figure 2.8 Shepard diagram from NMDS modeling of the sample bray-curtis distance matrix. Both the non-metric and linear fits indicate an excellent correlation to the dataset; stress converged at 0.06.

Table 2.1 Taxonomic classification and phylotype of OTUs with a statistically significant effect size (Log2FC) in the no-mussel treatment. OTUs are listed in order of descending effect size. Species is designated in parentheses when applicable and unclassified taxonomic levels are represented by a dash.

	Effect size (log2FC)	Average OTU count	P-adj	Phylum	Class	Order	Family	Genus (species)	Phylotype
No Mussels	4.40	14.42	$2.14e^{-08}$	NC10	12-24	Methylomirabiliales	*Methylomirabiliaceae*	*Candidatus* Methylomirabilis	N-DAMO
	4.08	6.19	$1.41e^{-04}$	NC10	12-24	JH-WHS47	-	-	N-DAMO
	3.88	2.35	$1.88e^{-03}$	NC10	12-24	JH-WHS47	-	-	N-DAMO
	3.80	7.13	$4.00e^{-04}$	NC10	12-24	Methylomirabiliales	*Methylomirabiliaceae*	*Candidatus* Methylomirabilis	N-DAMO
	3.72	7.04	$9.11e^{-05}$	Planctomycetes	Planctomycetia	*Candidatus* Brocadiales	*Candidatus* Brocadiaceae	*Candidatus* Brocadia	anammox
	3.59	21.01	$6.80e^{-07}$	NC10	wb1-A12	-	-	-	N-DAMO
	3.41	50.35	$2.38e^{-06}$	Euryarchaeota	Methanomicrobia	Methanosarcinales	ANME-2D	-	N-DAMO
	3.32	3.50	$1.49e^{-03}$	NC10	wb1-A12	-	-	-	N-DAMO
	2.95	26.12	$1.27e^{-04}$	NC10	12-24	JH-WHS47	-	-	N-DAMO
	2.93	33.33	$4.92e^{-06}$	Euryarchaeota	Methanomicrobia	Methanosarcinales	ANME-2D	-	N-DAMO
	2.85	16.14	$3.43e^{-06}$	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	*Nitrososphaeraceae*	*Candidatus* Nitrososphaera (gargensis)	AOA
	2.82	59.13	$1.12e^{-06}$	Euryarchaeota	Methanomicrobia	Methanosarcinales	ANME-2D	-	N-DAMO
	2.67	25.49	$1.43e^{-05}$	NC10	12-24	JH-WHS47	-	-	N-DAMO
	2.08	129.78	$1.98e^{-04}$	NC10	12-24	JH-WHS47	-	-	N-DAMO
	2.05	17.35	$1.51e^{-04}$	Nitrospirae	Nitrospira	Nitrospirales	*Nitrospiraceae*	*Nitrospira*	NOB
	1.73	126.03	$1.73e^{-06}$	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	*Nitrososphaeraceae*	*Candidatus* Nitrososphaera (SCA1170)	AOA
	1.66	21.87	$1.39e^{-04}$	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	*Nitrososphaeraceae*	*Candidatus* Nitrososphaera	AOA
	1.59	12.25	$1.69e^{-03}$	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	Nitrososphaeraceae	*Candidatus* Nitrososphaera	AOA
	1.58	4.50	$6.64e^{-04}$	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	Nitrososphaeraceae	*Candidatus* Nitrososphaera (SCA1170)	AOA

Table 2.2 Taxonomic classification and phylotype of OTUs with a statistically significant effect size (Log2FC) for sediments with mussels. OTUs are listed in order of descending effect size. Unclassified taxonomic levels are represented by a dash.

	Effect size (log2FC)	Average OTU count	P-adj	Phylum	Class	Order	Family	Genus	Phylotype
With Mussels	1.88	5.57	$2.26e^{-04}$	Proteobacteria	Betaproteobacteria	Nitrosomonadales	*Nitrosomonadaceae*	-	AOB
	1.21	15.57	$1.02e^{-03}$	Planctomycetes	Planctomycetia	Candidatus Brocadiales	*Candidatus* Brocadiaceae	*Candidatus* Brocadia	anammox
	0.80	209.27	$1.31e^{-05}$	Nitrospirae	Nitrospira	Nitrospirales	*Nitrospiraceae*	*Nitrospira*	NOB

Table 2.3 Percent relative abundance of N-cycle organisms for mussel and depth treatments.

Taxonomic Classification	N-Cycle Phylotype	Mean Percent Relative Abundance			
		3 cm with-mussels	3 cm no-mussels	5 cm with-mussels	5 cm no-mussels
Candidatus Nitrososphaera	AOA	0.26	0.44	0.22	0.58
ANME-2D	nitrate-DAMO	0.12	0.21	0.11	0.63
NC10	nitrite-DAMO	0.0039	0.02	0.0035	0.08
Nitrospira	NOB/comammox	1.92	1	1.11	0.85
Candidatus Brocadia	anammox	0.1	0.05	0.07	0.15
Nitrosomonadaceae	AOB	0.27	0.13	0.1	0.08

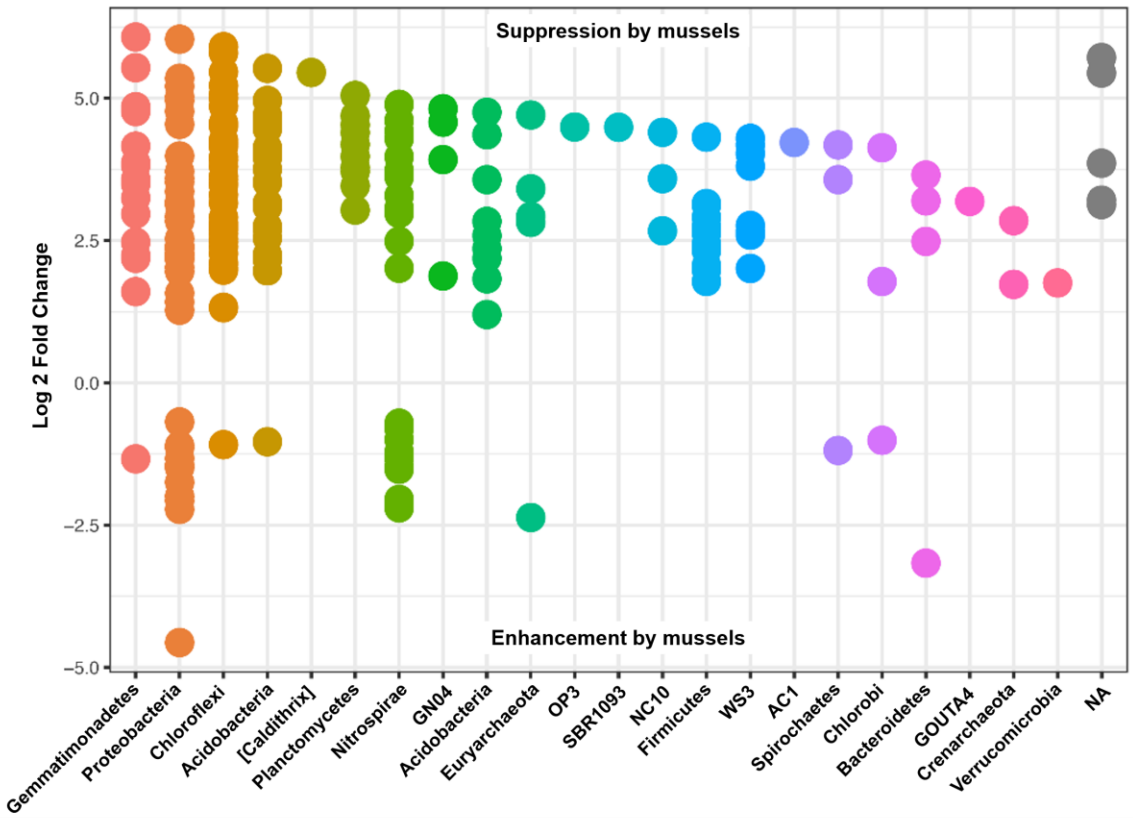


Figure 2.9 Results from a DESeq2 differential abundance analysis expressed as Log2FC comparison of with-mussels and no-mussels samples. Negative Log2FC represent phyla enhanced in the mussel bed and each point represents an individual OTU. To enhance clarity, only those OTUs with p-adj < 0.0001 are shown.

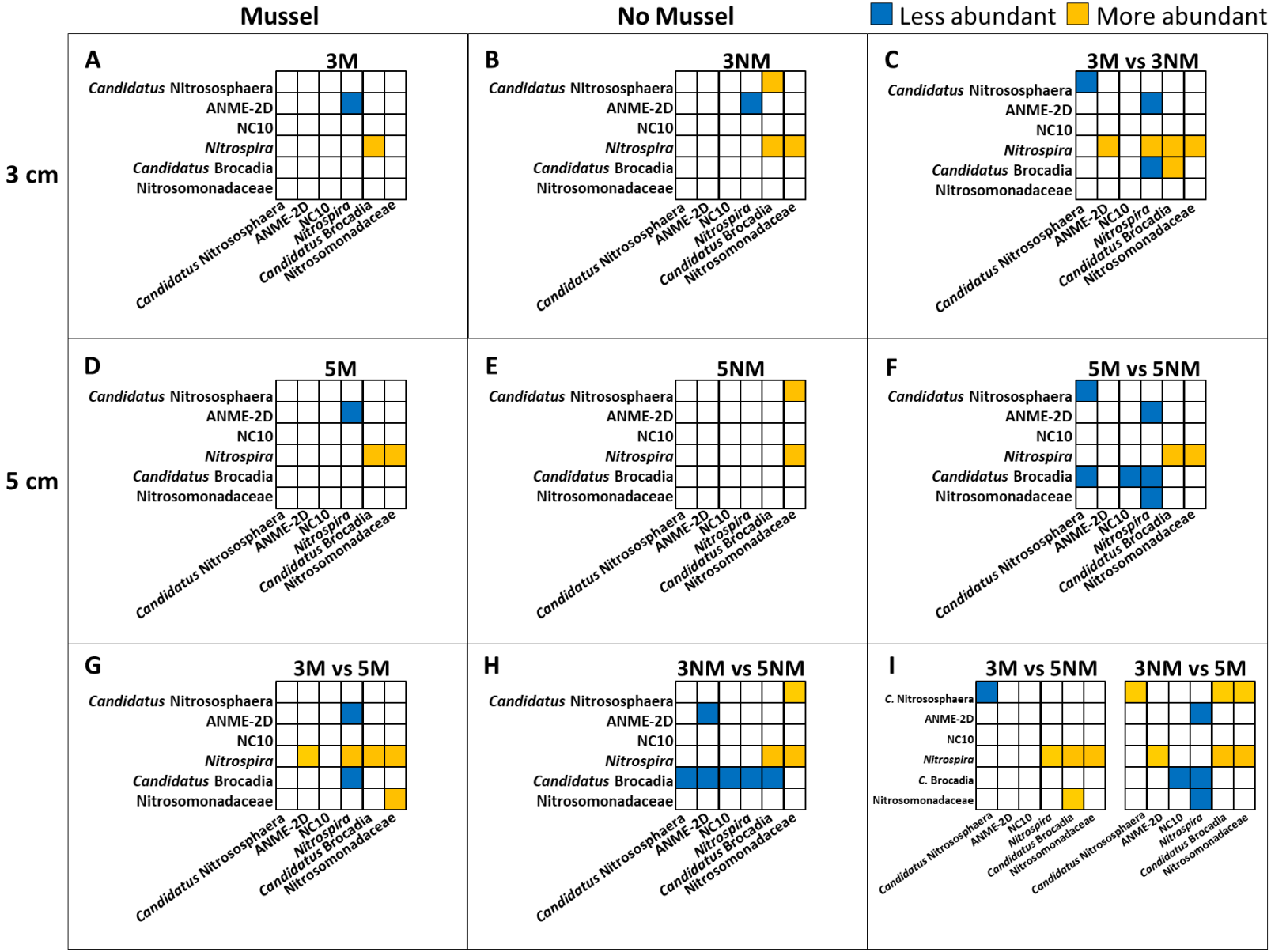

Mussel
No Mussel
Less abundant
More abundant
3 cm
5 cm
A
3M
B
3NM
C
3M vs 3NM
D
5M
E
5NM
F
5M vs 5NM
G
3M vs 5M
H
3NM vs 5NM
I
3M vs 5NM
3NM vs 5M
Candidatus Nitrososphaera
ANME-2D
NC10
Nitrospira
Candidatus Brocadia
Nitrosomonadaceae
C. Nitrososphaera
C. Brocadia

Figure 2.10 Statistically significant differences in N-cycle phylotype relative abundances (Padj<0.05). All boxes show y-axes compared to the "baseline" x-axes, with no boxes representing comparisons not meeting significance. **Boxes A-B, D-E**: Comparisons within treatment conditions of mussel presence and depth. **Box C**: Differentially abundant phylotypes between 3 cm mussel and 3 cm no mussel treatments. **Box F**: Relative abundance comparisons between 5 cm mussel and 5 cm no mussel treatments. **Box G**: Differential N-cycle phylotype abundances between 3 cm mussel and 5 cm mussel samples. **Box H**: Comparisons between 3 cm no mussel and 5 cm no mussel treatments. **Box I**: Abundance comparisons of 3 cm Mussel versus 5 cm no mussel, and 3 no mussel versus 5 mussel samples.

2.5 Discussion

2.5.1 Sediment microbiome

Numerous studies have found Proteobacteria to be the most abundant phylum in freshwater sediments [108-112], sediments with mollusks [113, 114], and also mollusk microbiomes [115-118]. Although our results showed Proteobacteria were the most abundant phylum, we observed a decrease in Proteobacteria by 6% and an increase in Nitrospirae by 10% in the presence of mussels. Families *Thermodesulfovibrionaceae* and "*FW*" accounted for many of the Nitrospirae OTUs that increased with mussels and helps explain decreases in species richness for mussel bed sediment.

Sediments contain the most phylogenetically diverse microbial communities [119] and structure and diversity of soil microbial communities is often determined by soil biogeochemistry [120], further supporting the impact mussels have on biogeochemical cycling. In support of our hypothesis, our data indicated that mussel presence in the UMR had a pronounced influence on the microbial community evenness, richness, and composition within the sediment. The observed changes in sediment microbial community structure and diversity showed mussels created a niche for specific microorganisms and may be attributable to the diverse chemical composition of mussel biodeposits, mixing of sediment from mussel burrowing, or the microbes living on mussels. Our findings of distinct microbial communities in mussel bed sediment are corroborated by a study of the California mussel [121] where taxonomic richness increased and taxa evenness increased following the removal of mussels from a rocky shore habitat.

In contrast to our results of decreased microbial diversity with freshwater mussels, research has shown invasive zebra mussels (*Dreissena polymorpha*) increased bacterial community diversity and richness [114], and metabolic diversity and activity in freshwater sediments [93]. Increased microbial diversity and activity has been attributed to the variety of C and N components in feces and pseudofeces, and also selects for the dominant microbial species [92, 93]. An experiment combining estuarine bivalve species (*N. virens, M. arenaria, and M. balthica*) implicated mussel-induced changes in O_2, NH_4^+ and NO_3^- fluxes for the alteration of microbial community composition [122]. On the other hand, investigation of microbiota in Thick-shelled River Mussel (*Unio crassus*) beds did not find any difference in microorganism diversity, abundance, and composition [123]. This may be explained by the drastic differences in the study site, with high mussel densities (23-433 mussels/m^2) and control plots containing low microbial diversities with mean species richness of 48 OTUs/sample with high evenness [123]. The contrasting findings of microbial community diversity and composition indicate that mussel density and/or mollusk species may produce different responses by microorganism communities.

Additionally, alterations in sediment microbial community structure may arise from exposure to the mussel shell, tissues, or fecal microbiome. Mussel tissue and fecal material has been shown to contain less diverse microbiomes than the surrounding water and sediment for the zebra mussel [116], tropical oyster (*Crassostrea rhizophorae*) [118], and marine mussel, *Mytilus californianus* [116, 121]. Some studies have attributed immediate increased sediment microbial activity to the mussel intestinal microbiome [124]. Furthermore, mollusk biodeposition rates and biodeposit chemical compositions are highly dependent on mollusk species [125, 126], and food availability [11, 56, 127],

so it makes sense that our results differ from studies with dissimilar mollusk species, densities, and study location.

Changes in mussel bed sediment microbial communities was also likely enhanced by mussel burrowing, because diffusion of substrates across the water-sediment interface is a relatively slow process [128] and is increased by mollusk burrowing [11], which ultimately affects microbial communities. For example, the burrow of shrimp species *Upogebia deltaura* and *Callianassa subterranean* contained distinct bacterial communities and a 3-fold increase in taxon richness [129], and the estuarine bivalve, *C. fluminea*, stimulated microbial diversity via bioturbation [130]. It is likely that UMR mussel bed sediments also experience the benefits from bioturbation, such as sediment mixing [131] and aeration [11]. Furthermore, bioturbation has been linked to increased NH_4^+ concentrations which alters the N-transforming microbial community [132], with greatest effects on bacteria growth found at 4 to 6 cm depth below the water-sediment interface [133].

2.5.2 N-cycle microbial community

Our research revealed an increase in anammox bacteria abundance 3 cm below the water-sediment interface when mussels were present, shown for the anammox community using anammox-targeted qPCR (2.2-fold increase) and for *Candidatus* Brocadia using non-targeted 16S rRNA gene amplicon sequencing (2-fold increase). The significance of agreement between these techniques is finding that increases in the genus *Candidatus* Brocadia are representative for the anammox phylotype as a whole. *Candidatus* Brocadia may also make up a majority of the anammox community in UMR sediment, as amplicon sequencing did not detect anammox bacteria belonging to other

genera. We are confident in these conclusions, as *Candidatus* Brocadia is often the dominant anammox genus in freshwater sediments [26, 134, 135]. One study showed that feeding of NH_4^+, NO_2^-, NO_3^-, and acetate led to an 80% enrichment of *Candidatus* Brocadia fulgida, signifying that B. *fulgida* could outcompete anammox species in genera *Candidatus* Anammoxoglobus and *Candidatus* Kuenenia, species *Candidatus* Brocadia anammoxidans, and even denitrifiers when acetate is present [136]. This indicates that *Candidatus* Brocadia has a distinct ecological niche and can utilize intermediates from anaerobic degradation of organic C to reduce NO_3^- [136]. Therefore, it is possible that a portion of our observed increases in *Candidatus* Brocadia with mussels was attributable to C biodeposition in the UMR.

Our research also revealed a vertical distribution of anammox bacteria with higher abundances near the sediment surface, which reflects the vertical distribution found in an agricultural field [137], oxygen minimum zone [138], flooded paddy fields [137, 139], and an urban wetland [140]. A vertical anammox distribution has been shown to coincide with NH_4^+ presence and NO_2^- production [27, 140-142] and anammox "hotspots" occur in zones of low, but not entirely absent, O_2 availability [143]. Anammox abundance in freshwater sediment can range between $7x10^4$ -8 $x10^6$ gene copies g^{-1} sediment [134], or between 10^6-10^7 gene copies g^{-1} sediment in peak NO_2^- microniches at the oxic-anoxic interface [140, 144, 145]. Studies have shown anammox bacteria increase 1.5 to 2-fold within their niche [144, 145], similar to our findings of a 2.2-fold increase in anammox bacteria 3 cm below the water-sediment interface with mussels.

Co-occurrence of aerobic NH_4^+ oxidation and anammox niches are likely due to linked NO_2^- oxidation and reduction, respectively [142]. Interestingly, we found that

mussels also enhanced taxa from the AOB family Nitrosomonadaceae and the OTUs made up a greater proportion of mussel bed sediment populations near the water-sediment interface. To this point, the pacific oyster (*C. gigas*) was found to increase porewater NH_4^+ and elevate the concentration of NH_4^+ oxidizing microorganisms [146]. Furthermore, our previous research [13, 56] showed elevated NH_4^+ and NO_2^- in porewater of a similar depth below mussels. It makes sense that these groups of N-transforming bacteria co-occur where their substrate microniches overlap, and is likely enhanced by mussels periodically aerating the sediment [132]. Intermittent aeration has shown to enrich microbial cultures in AOB and anammox bacteria in engineered partial nitritation-anammox processes [147, 148], and similar to our findings, enriches the anammox genus *Candidatus* Brocadia [147].

On the other hand, we saw a decrease in *Candidatus* Nitrososphaera (AOA) with mussels at 3 cm (1.7-fold) and 5 cm (2.6-fold) sediment depths. It makes sense that mussels suppress abundance of AOA since these organisms typically dominate sediment niches with low NH_4^+ concentrations [70, 149]. Furthermore, a group of OTUs suppressed by mussels were classified at the species level as *Candidatus* Nitrososphaera gargensis, which are partially inhibited by NH_4^+ concentrations (3.08 mM) much lower than AOB [71, 149-151]. Furthermore, nitrifier niche partitioning studies using agricultural soil showed that AOB increased in abundance and activity following the addition of urine-derived N, while AOA remained unchanged [149, 152, 153]. Therefore, it is possible that mussel biodeposits and an increased flux of agriculturally-fed water into sediment by mussel burrowing enhanced porewater NH_4^+ composition such that Nitrosomonadaceae out competed *Candidatus* Nitrososphaera. Our results agree with

Chen et al.(2017), who found bioturbated sediment corresponded with a greater diversity of AOB and lower diversity of AOA microbial communities [132]. On the other hand, our results of decreased abundance of *Candidatus* Nitrososphaera co-occurring with an increase in *Nitrospira* is in contrast to previous findings that these organisms may exhibit similar niche partitioning [151]. For example, some species in *Candidatus* Nitrososphaera can adjust their metabolism for low oxygen availability [154] and *Nitrospira* species are adapted to low oxygen concentrations [154-156] and microoxic environments [157]. Alternatively, our detected increased abundance of *Nitrospira* may include species with a variety of environmental niches.

Some *Nitrospira* species have shown to occupy a niche at oxic-anoxic interfaces, in opposition to NOB with higher O_2 tolerances such as those in genus *Nitrobacter* [157]. This supports our mussel-attributed increases in relative *Nitrospira* abundances (1.9-fold) at 3 cm sediment depths. Although we saw two different *Nitrospira* OTUs suppressed and enhanced by mussels, the mussel-enhanced OTUs had a larger mean abundance by about 12%. Different NOB OTUs enhanced with and without mussels further suggests that mussel bed sediments harbor specific NOB strains sensitive to microoxic niches. Despite OTU variability, we can conclude that mussels enhance the *Nitrospira* phylotype, especially near the water-sediment interface where *Nitrospira* were 1.7-times greater than deeper mussel bed depths.

On the other hand, we did not expect to see an increase in both NOB and anammox phylotypes due to competition of NO_2^- as a substrate. The co-occurrence of *Nitrospira* and anammox bacteria may be explained by the metabolic versatility of *Nitrospira* species, especially if mussel-derived urea provided an additional source of

NH_3 and NO_2^- via reciprocal feeding between ammonia oxidizers and *Nitrospira*. Furthermore, these phylotypes have been shown to coexist in an oxygen minimum zone, where anammox bacteria obtained a majority of NO_2^- from NO_3^- reducers [158]. The similar effect size of mussels on *Nitrospira* (1.9-fold) and *Candidatus* Brocadia (2-fold) at 3 cm depth suggests that mussels may exert similar influences on the niches of these phylotypes. It is possible that these anammox and *Nitrospira* phylotypes were functionally linked in shallow mussel bed sediment, which has been shown for microoxic niches [30]. Furthermore, it is possible that the *Nitrospira* co-occurring with *Candidatus* Brocadia were *Nitrospira* species with the genetic potential for comammox, as a fluorescence in-situ hybridization study confirmed the extensive aggregation of the 2 phylotypes in hypoxic conditions (<3.1 μM O_2) [30]. Despite *Nitrospira* comammox being identified in numerous aquatic environments [76, 159, 160], we cannot conclusively identify comammox without sequencing the ammonia monooxygenase gene [30, 159].

In contrast to studies which found significant N-reduction on both a marine mussel (*Mytilus californianus*) [90] and a freshwater mussel (*Limnoperna fortunei*) [161], our results showed that mussels suppressed N-DAMO OTUs in phylum "NC10" (2.1-fold) and family "ANME-2d" (1.8-fold). One study determined NO_3^--DAMO was responsible for NO_3^- reduction and anammox for NO_2^- reductions in a bioreactor supplied with NH_4^+, NO_2^-, NO_3^-, CH_4, and anoxic conditions, thus concluding anammox outcompeted NO_2^--DAMO [162]. These findings make sense, because anammox bacteria have a higher affinity for NO_2^- [87], anammox outperform N-DAMO in bioturbated sediments with higher NH_4^+ and lower NO_2^- and NO_3^- [132], and anammox and N-

DAMO communities have a competitive relationship in burrowed mangrove sediment [132]. Furthermore, NC10 bacteria in a peatland were most prevalent at depths with porewater CH_4 concentrations near 300 µM, where NO_3^- consumption exceeds production, and in completely anoxic conditions [163]. According to the literature, it makes sense that we found UMR mussels enhanced *Candidatus* Brocadia and suppressed NO_2^- reducing-NC10. Perhaps N-DAMO organisms did not have a favorable niche in mussel bed sediment because biodeposition products created an excess of NH_4^+ in sediment porewater [164], or burrowing activity increased oxygen concentrations and made methane oxidation unfavorable [165]. Our finding that no-mussel sediment contained 3 times more ANME-2D in deeper, and presumably anoxic sediment, further suggests that mussels broaden the oxic-anoxic interface niche [87, 132]. However, we cannot extrapolate these findings to all denitrifying organisms, since denitrifying species are sporadically distributed among various taxonomic lineages, and are difficult to identify solely with16S rRNA amplicon sequencing [166].

Although we observed greater relative abundances of *Nitrospira* than *Candidatus* Brocadia in a majority of treatments, both phylotypes increased by a factor of 2 with mussels at 3 cm depth. No-mussel samples contained a significantly smaller proportion of *Candidatus* Brocadia in shallow sediments compared to almost all N-transformers found in the deeper control sediments. Our phylotype-level analyses revealed similarities with the OTU-level differential abundance comparisons. For example, phylotype comparisons showed ANME-2D was less abundant than *Nitrospira* in 3 cm sediments with mussels, and *Candidatus* Nitrososphaera was more abundant than *Candidatus* Brocadia in 3 cm sediment samples without mussels. These results relate to DESeq2 OTU comparisons

which found *Candidatus* Brocadia and *Nitrospira* enhanced with mussels while ANME-2d and *Candidatus* Nitrososphaera were suppressed with mussels.

Extending our focus beyond N-cycling organisms, we demonstrated that mussels promoted a large effect size for OTUs classified as *Thermodesulfovibrionaceae* (Nitrospirales order). In contrast to *Nitrospira*, the Nitrospirales genus *Thermodesulfovibrio* contains multiple sulfate reducing species [167, 168] and can outcompete other anaerobic organisms when sulfate is present [169]. These findings are corroborated by discoveries of significantly greater C and sulfate concentrations from mussel biodeposits and 63% greater sulfate reduction in sediments with mussels [170]. Biodeposition products often lead to increasingly anoxic sediment and greater activity of anoxic microorganisms [91, 170], presumably due to consumption of excretion products by oxygen-consuming microorganisms [92]. Interestingly, Fdz-Polanco et al. (2001) observed simultaneous N and sulfate removal in an anaerobic fluidized-bed reactor and proposed simultaneous anammox and sulfate reduction [171]. Coupled biological sulfate reduction and anammox reactions are metabolically feasible [172, 173] and have been of interest in the recent history [174-177], therefore warranting further research. Therefore, we showed that *Thermodesulfovibrionaceae* are significantly increased in the presence of mussels which may affect sulfate reduction [178] in tandem with anammox reactions in UMR sediments.

As a whole, mussels do have an impact on microbial niches and lower the overall community diversity. Mussel-influenced changes in microbiological diversity may have larger ecosystem implications, such as macrobiota richness and diversity [179, 180]. Native freshwater mussels are capable of increasing macrobiota diversity as a result of

being keystone species [181] and ecosystem engineers [182, 183]. Mussel biogeochemical hotspots can lead to a bottom-up trophic cascade by enhancing N substrates normally limiting primary productivity, ultimately leading to increased richness [184] and biodiversity [185] of higher trophic levels.

2.6 Conclusion

As far as we know, this is the first study to characterize freshwater mussel effects on microbial community diversity, composition, and the vertical distribution of N-cycle microorganisms in the UMR. qPCR of the anammox-specific 16S rRNA gene revealed an increase in anammox bacteria abundance 3 cm below the water-sediment interface when mussels were present, and confirmed anammox bacteria were normally distributed with depth. Non-targeted 16S rRNA gene amplicon sequencing revealed mussel presence suppressed AOA (*Candidatus* Nitrososphaera) and that the families *Thermodesulfovibrionaceae* and "*FW*" (Nitrospirales order) were overrepresented among the enhanced OTUs with-mussels. Mussel bed sediment contained microbial communities with 10% greater Nitrospirae and 6% fewer OTUs belonging to the phylum Proteobacteria, which ultimately had a pronounced influence on microbial community evenness, richness, and composition. This was indicated by lower observed species richness, Chao1 diversity, Shannon diversity, and clustering of mussel samples in an NMDS analysis. We have shown that native freshwater mussels affect niche differentiation of N-cycle microorganisms, as evidenced by increased abundances of AOB family *Nitrosomonadaceae*, anammox genus *Candidatus* Brocadia, and NOB genus *Nitrospira*, while exhibiting a decrease in AOA genus *Candidatus* Nitrososphaera, and n-damo organisms in the phylum NC10 and family ANME-2d. Co-occurring 2-fold

increases in *Candidatus* Brocadia and *Nitrospira* in shallow sediment suggests mussels may enhance microbial niches at the interface of oxic-anoxic conditions, presumably through biodeposition and burrowing. Ultimately, this study demonstrates the large impact mussels have on biogeochemical N-cycling and ecosystem services in freshwater agroecosystems.

CHAPTER 3: METAGENOMIC ANALYSIS OF NITROGEN CYCLING GENES IN UPPER MISSISSIPPI RIVER SEDIMENT WITH MUSSEL ASSEMBLAGES

3.1 Abstract

We investigated the abundance of nitrogen cycling genes from freshwater sediment of the upper Mississippi river (UMR), a watershed largely impacted by nutrient runoff from agricultural land use. We hypothesized that genomic potential for ammonia and nitrite oxidation would be greater in the sediment with mussel assemblages, presumably from mussel biodeposition products, namely ammonia and organic carbon. The relative abundance of nitrogen metabolic pathways showed the UMR microbial communities had the largest genomic potential for nitrogen fixation, urea catabolism, nitrate metabolism, and nitrate assimilation, listed in decreasing abundance. Furthermore, NarGHI and NxrAB proteins represented the greatest functional potential for microbial communities, both with and without mussels. Using linear discriminant analysis (LDA), we found nitrification genes were the most important biomarkers for nitrogen cycling genomic potential when mussels were present, and this presented an opposing effect on nitric oxide reduction. The genes most responsible for larger genomic potential of nitrification were *amoA* associated with comammox *Nitrospira* and *nxr* homologs associated with *Nitrospira*. On the other hand, the most distinctive biomarkers of microbial communities without mussels were *norB* and *nrfA*. Ultimately this research demonstrates the impact of native mollusks on microbial nitrogen cycling in an aquatic agroecosystem.

3.2 Introduction

The estimated, healthy "planetary boundary" for land applied nitrogen (N) is 35 teragrams (Tg)- N yr^{-1} [1, 186, 187]. Human activity has pushed the planet well beyond this boundary, to 150 Tg-N yr^{-1} [188], resulting in excessive aquatic eutrophication and harmful algal blooms [189]. Excess land-applied urea and ammonia (NH_3) is biologically oxidized to nitrate (NO_3^-), which has a high runoff potential [190], causing negative ecosystem impacts, degraded water quality, and biogeochemical cycling imbalances [191]. The well-studied Upper Mississippi River (UMR) basin [192, 193], that contributes 50,000 metric tons of N to the Gulf of Mexico annually [194], is a global "epicenter" of excessive N-transfer from land to water. In the UMR, microbially-driven N-biogeochemistry [195] is symbiotically linked to freshwater mussel N-cycling. Billions of native freshwater mussels live in assemblages in the UMR, filter billions of gallons of water, and remove tons of biomass from overlying water daily [45]. Previous studies have shown that mussels increase the concentration of N in sediment porewater via bioactivity including burrowing (bioturbation) and excretion of feces and pseudofeces (biodeposition) [13, 59]. Intermittent sediment aeration and elevated nutrient concentrations create a niche ripe for removal of N at the interface of oxic and anoxic conditions, via nitrification, denitrification, and anaerobic ammonium (NH_4^+) oxidizing (anammox) processes, making mussels and microbial activity a functional biological unit for N-cycling in aquatic systems.

The main functions of N-cycling microorganisms in aquatic sediments include N-fixation by benthic organisms, and elemental transformations such as nitrification and denitrification, which oxidizes and reduces inorganic N, respectively [196]. Biological N-

fixation by benthic prokaryotes commonly produces 0.4-1.6 to 76 g-N m^{-2} yr^{-1}, and reaches 76 g-N m^{-2} yr^{-1} in dense microbial mats [197]. Microorganisms fix N with the nitrogenase (Nif) enzyme complex and are responsible for catalyzing half of the bioavailable N on Earth [198]. Bioavailable N (ammonia) can be assimilated into biomass for growth or used as an energy source in nitrifying organisms. The NH_3 monooxygenase (Amo) enzyme catalyzes the oxidation of NH_3 into hydroxylamine (NH_2OH), which may be oxidized to nitrite (NO_2^-) using hydroxylamine oxidoreductase (Hao), and complete nitrification occurs when NO_2^- is oxidized to NO_3^- in organisms containing a nitrite oxidoreductase enzyme (Nxr). Numerous microorganisms are capable of partial nitrification, the oxidation of NH_3 to NO_2^- or NO_2^- to NO_3^-, while only the *Nitrospira* lineage II contains the genetic potential to completely oxidize NH_3 to NO_3^- [199, 200]. Another metabolic pathway for NH_3 oxidation is present in anammox bacteria (Planctomycetes phylum) which oxidize NH_4^+, reduce NO_2^-, and produce a hydrazine intermediate with the hydrazine synthase enzyme (Hzs) to ultimately produce N_2 gas [141]. Competition for N resources arises from the metabolic pathways of dissimilatory NO_3^- reduction to NH_3 (DNRA), stepwise NO_3^- reduction to NOx (NO_2^-, nitric oxide, nitrous oxide), or complete denitrification to N_2.

N-cycling ecosystem services are impacted in agroecosystems due to the increased availability of reactive N [201]. For example, NH_3 oxidizing pathways are enhanced by greater NH_3 concentrations, and the subsequently oxidized-N also enhances nitrate reduction pathways. However, more research is needed to accurately quantify services of biogeochemical cycling in agroecosystems [202], especially in aquatic

systems where macrobiota significantly enhance the transfer of N from overlying water to sediment.

In a previous study, we showed that sediment underlying a native freshwater mussel assemblage harbored microbial communities with lower species richness and evenness as compared to mussel-free sediment [203]. Additionally, mussels had a distinct and significant effect on the vertical distribution of multiple N-cycling microorganisms, including NO_2^- oxidizing bacteria (NOB) in the genus *Nitrospira*, aerobic NH_3 oxidizing bacteria (AOB) in family *Nitrosomonadaceae*, and anammox bacteria from candidate genus *Brocadia*. These N-cycling taxons were increased most drastically at 3 cm depth below the water-sediment interface, a depth which is relevant to mussel burrowing, and suggested the presence of an oxic-anoxic interface niche for N-cycling microorganisms 3 cm below the water-sediment interface. The N-cycling microbial community was most similar between the shallow (3 cm) mussel sediment and deeper (5 cm) no-mussel sediment. Therefore, these two sample groups were chosen for follow-up metagenomic sequencing to assess how mussel presence impacted N-cycling gene abundances in N-cycling communities of similar composition.

This study aimed to determine if mussels increased the abundance of N-cycling genes, especially genes responsible for NH_3 oxidation, NO_2^- oxidation, and would clarify previous findings of increased AOB and NOB taxons with mussels. Our hypothesis was that N-cycling microbial communities of the previously determined oxic-anoxic interface niche will contain greater metabolic potentials for urea degradation, NH_3 oxidation, and NO_2^- oxidation reactions in the presence of mussels. These results would indicate which N metabolic pathways are most impacted by mussel assemblages in the UMR.

3.3 Materials and Methods

3.3.1 Sediment collection and DNA isolation

Sediment cores were removed from our study sites in the Upper Mississippi River Pool 16 using a 2-in diameter, post-driver with a polypropylene liner (Multi-State Sediment Sampler, Art's Manufacturing and Supply, Inc.; American Falls, ID, USA) and sediment samples were removed using an ethanol flame-sterilized 3/8-in diameter drill bit at sediment depths of 3 and 5 cm. Sediment (0.25g) was removed in quadruplicate (n=4, 3 cm depth with mussels; n=4, 5 cm depth without mussels) and stored in sterile bead beating tubes overnight at -20°C. Genomic DNA was isolated (PowerSoil DNA Isolation Kit; MoBio Laboratories, Inc., Carlsbad, CA, USA), assessed for total DNA quality and quantity (NanoDrop 1000; Thermo Fisher Scientific, Waltham, MA), and stored at -20°C prior to sequencing. These samples correspond to representative sequences without-mussels and with mussels (**Table A3.1**) from a previous 16S rRNA amplicon study of N-cycling taxonomic profiling [203].

3.3.2 Metagenomic shotgun sequencing

For each sample, 120 ng of genomic DNA in 60 µL of 10mM Tris-HCl, pH 8.0 buffer, was placed into 1.5 mL RNase-/DNase-free, low binding microcentrifuge tubes. Library creation steps were performed by the University of Iowa Institute for Human Genetics, Genomics Division, and included DNA shearing using the Covaris Adaptive Focused Acoustics™ process (Covaris E220 Focused-ultrasonicator; Covaris, Inc., Woburn, MA), and DNA fragment purification and end polishing (KAPA Hyper prep kits; Kapa Biosystems, Inc., Wilmington, MA) prior to ligation to indexed adaptors. The library size distribution was validated using the Agilent 2100 Bioanalyzer Instrument

(Agilent Technologies, Santa Clara, CA), and quantified using the q-PCR KAPA library amplification module following manufacturer instructions (Kapa Biosystems, Inc.). The indexed libraries were normalized, pooled, and clustered on a flow cell using the cBOT Cluster Generation System (Illumina, Inc., San Diego, CA) and sequenced on the Illumina HiSeq 4000 System (Illumina, Inc.) in high output mode (1 lane, 2x150bp). Metagenomic reads and sequence statistics are accessible at the MG-RAST server, European Nucleotide Archive, and the NCBI Sequence Read Archive (**Table A3.1**).

3.3.3 Bioinformatics pipeline

Using the HUMAnN2 [204] standard workflow, paired-end reads were imported into MetaPhlAn2 [14] and mapped against functionally annotated genomes from the ChocoPhlAn pangenome database (NCBI RefSeq Release 80) using Bowtie2 algorithm with default settings [205]. Unmapped reads were subjected to rapid translated search using the Diamond [206] algorithm against the Universal Protein Reference Database for 90% similarity (UniRef90 [207]) within HUMAnN2 using default values (e-value threshold=1.0, prescreen threshold=0.01, and identity threshold=50.0%). Hits to protein families and organism-specific gene hits were compared to the 2016 *MetaCyc* pathway collection [208] using HUMAnN2 core algorithms. The output of this pipeline included tables of gene family and pathway abundances in units of reads per kilobase (RPK), and pathway coverages. Gene and pathway abundance tables were normalized for sample sequencing depth in copies per million (CPM), labeled with "mussel" and "no-mussel" metadata categories, and stratified by lowest common ancestor (LCA) classification using the scripts "humann2_renorm_table" and "humann2_infer_taxonomy", as provided with the HUMAnN2 package. The CPM-normalized gene families were regrouped into Gene

Ontology (GO) and KEGG Orthology (KO) terms via the "humann2_regroup_table" script and mapping files "uniref90_go" and "uniref90_ko". Of the original protein clusters, 68.6% and 33.0% were successfully regrouped into GO and KO features for further processing, respectively.

Protein clusters that mapped onto KEGG orthologous groups within the Nitrogen metabolism module and urease functions were considered for this study (**Table 3.1**). Parent categories only contained functional genes unique to the given module to obtain discrete categories. As a result, NarGHI and NapAB were specific to dissimilatory nitrate reduction and were not included in denitrification or DNRA pathways. The full functional "tree" used for LDA can be found at https://metacyc.org/group?id=biocyc13-27028-3726046429-frozen. The complete list of KO and GO relative abundance profiles for our metagenomes can be found at the following links:

https://metacyc.org/group?id=biocyc13-27028-3725982783

https://metacyc.org/group?id=biocyc13-27028-3725983191.

Table 3.1 KO protein clusters and KEGG Modules used to categorize functional genes in this study.

KEGG modules	KO and protein name			
nitrogen fixation	nitrogenase		nitrogenase delta subunit	
	NifDKH		AnfG	
	K002586, K002588, K002591		K00531	
assimilatory nitrate reduction	ferredoxin-nitrate reductase		ferredoxin-nitrite reductase	
	NarB		NirA	
	K00367		K00366	
dissimilatory nitrate reduction	nitrate reductase		periplasmic nitrate reductase	
	NarGHI		NapAB	
	K00370, K00371, K00374		K02567-K02568	
denitrification	nitrite reductase (NO-forming)	nitrite reductase (NO-forming)/ hydroxylamine reductase	nitric oxide reductase	nitrous oxide reductase
	NirK	NirS	NorBC	NosZ
	K00368	K15864	K04561	K00376
nitrification	nitrite oxidoreductase	hydroxylamine oxidoreductase	ammonia monooxygenase	
	NxrAB	Hao	AmoCAB	
	K00370, K00371, K00374	K10535	K10944-K10946	
anammox	nitrite reductase (NO-forming)	nitrite reductase (NO-forming)/ hydroxylamine reductase	hydrazine hydrolase	hydrazine synthase
	NirK	NirS	Hdh	Hzs
	K00368	K15864	K20935	K20932-K20934
urease	urease subunit alpha	urease subunit beta	urease subunit gamma	
	UreC	UreB	UreA	
	K01428	K01429	K01430	
DNRA	nitrite reductase (NADH)		nitrite reductase (cytochrome c-552)	
	NirBD		NrfAH	
	K00362-K00363		K03385	

3.3.4 LDA effect size

N-cycling gene families were placed into parent categories of N-cycle function defined within the KEGG "Nitrogen metabolism" module (i.e. nitrification, denitrification, anammox, N-fixation, assimilation), and further specified into functional sub-categories (i.e. ammonia oxidation, nitrite oxidation, etc.), and gene families, as defined by KEGG ontologies (i.e. *amoA*, *nxrA*). Relative abundances (CPM) were assessed for linear discriminant analysis (LDA) effect size (LEfSe), a method to determine the consistent metagenomic features responsible for differences between microbial communities [209]. All samples were labeled by class (n=4 with mussels, n=4 without mussels) and features were compared for differential distribution using the non-parametric factorial Kruskal-Wallis sum-rank test (alpha= 0.05). Features deemed differentially abundant were compared for effect size using the pairwise Wilcoxon rank-sum test (alpha= 0.05), and input into a LDA model which ranked features according to effect size, with a LDA score of 2.0 chosen as a cutoff for inclusion as a significant feature. The LEfSe program ranked genes by effect size, with the highest ranking given to those with biological consistency, meaning differential abundance scores held true for higher order categories of gene and pathway abundances. LEfSe biomarker results were graphically displayed with the "Plot Cladogram" command.

3.3.5 Phylogenetic tree construction

All N-cycle genes identified as differentially abundant were labeled with species of origin from the protein cluster's mapping to NCBI taxonomy ID. The comammox genome from *Candidatus* Nitrospira inopinata was not included in the Chocophlan pangenomes at the time of this study. As a result, we used multiple sequence alignments

of protein sequences to determine if differentially abundant nitrification functional genes originated from the comammox *Nitrospira* lineage. Multiple sequence alignments of AmoA proteins were performed in MEGA7.0.20 [210] using reference sequences from IMG (**Table A3.3**), using the MUSCLE algorithm [211] with default options (Gap penalties: open= -2.9, gap extend= 0, hydrophobicity multiplier= 1.2), the neighbor joining method of clustering (8 iterations, γ=24) [211], and trimmed for quality in Jalview [212]. Phylogeny was reconstructed using 100 bootstrap replications [213] of Maximum Likelihood method based on the Poisson model for amino acid substitutions, assuming gamma distributed evolution rates with 5 discrete categories, and 80% site coverage cutoff for partial deletions. Trees were constructed with the Subtree-Pruning-Regrafting (SPR) maximum likelihood heuristic method and the initial tree was inferred by the Neighbor-Join and BioNJ algorithms [210].

3.4 Results

3.4.1 Nitrogen cycling gene abundances

The goal of this research was to determine which N-cycling processes dominated microbial communities in UMR sediment, and which genes were most characteristic of sediment with mussel assemblages (**Figure 3.1**). Nitrogen compound metabolic processes (GO:0006807) represented an average relative abundance of 66.0 CPM (±3.1 CPM) in microbial communities with mussels and 63.3 CPM (±3.2 CPM) without mussels, and was more abundant than other biological processes such as “aerobic respiration” and “one-carbon metabolic processes”.

According to GO annotations (**Figure 3.2**), microbial communities had the greatest potential for N-fixation even though *nifHDK* was not the most abundant N-

cycling gene family considered in the study. This may be explained by the fact that the GO parent category includes all genes involved in N-fixation and not solely *nifHDK*. Urea catabolism and NO_3^- metabolic processes were similarly abundant and were composed of the most variable gene families. Gene families specific to the denitrification pathway (*nirK*, *nirS*, *norBC*, and *nosZ*) were less variable than other N-cycling gene families, both within biological replicates and between mussel and no-mussel treatments. Nitrification biomarkers were moderately abundant, with NH_3 oxidation representing a majority of the genetic potential. The average abundance of *amoCAB* with mussels was 1.6 (±0.6 CPM), the no-mussel treatment had an average count of 1.4 (±0.3 CPM), and both treatments had an average abundance less than 1 CPM for *hao*. Both the mussel and no-mussel metagenomes had non-detectable abundances of anammox biomarkers, *hzs* and *hdh*. The NO_2^-/NO_3^- transforming gene families (*narGH/nxrAB*) represented the largest N-cycling genomic potential but were also quite variable in the treatment with mussels. Ultimately, these microbial communities show large genomic potentials to transform NO_2^-/NO_3^- rather than removing N through denitrification or anammox processes.

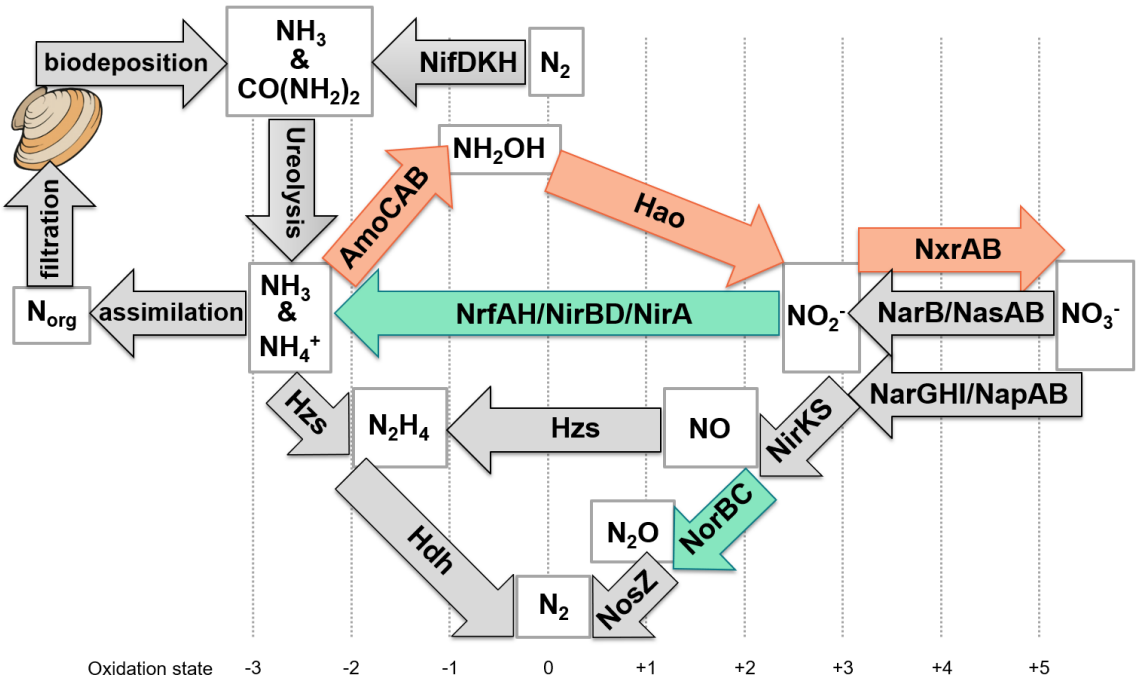


Figure 3.1 Proposed flow of nitrogen in the UMR freshwater mussel bed. N is added to the sediment by microbial N-fixation and can be detected by presence of the functional gene, nitrogenase (NifDKH). N may also be added through mussel biodeposition of NH3 and urea ($CO(NH_2)_2$), which may be hydrolyzed by urease enzymes. Bioavailable N could be assimilated into microbial biomass or utilized in redox reactions. Assimilated N (Norg) is recycled in the aquatic system through bivalve filtration processes. Redox transformations of N include microbial nitrification and is quantified by the functional genes, ammonia monooxygenase (AmoCAB), hydroxylamine dehydrogenase (Hao), and nitrite oxidoreductase (NxrAB). Complete removal of N is possible with anammox biochemical processes (Hzs, Hdh) or denitrification by sequentially reducing NO_3^- to N_2 with reductase enzymes for NO_3^- (NarGHI/NapAB), NO_2^- (NirKS), NO (NorBC), and N_2O (NosZ). Lastly, N may be temporarily sequestered via assimilatory NO_3^- and NO_2^- reduction (NasAB/NarB and NirA, respectively) and dissimilatory reduction to NH_3 (DNRA) (NrfAH). Colored arrows represent the biomarker genes found to be differentially abundant in this research without mussels (green) and with mussels (orange).

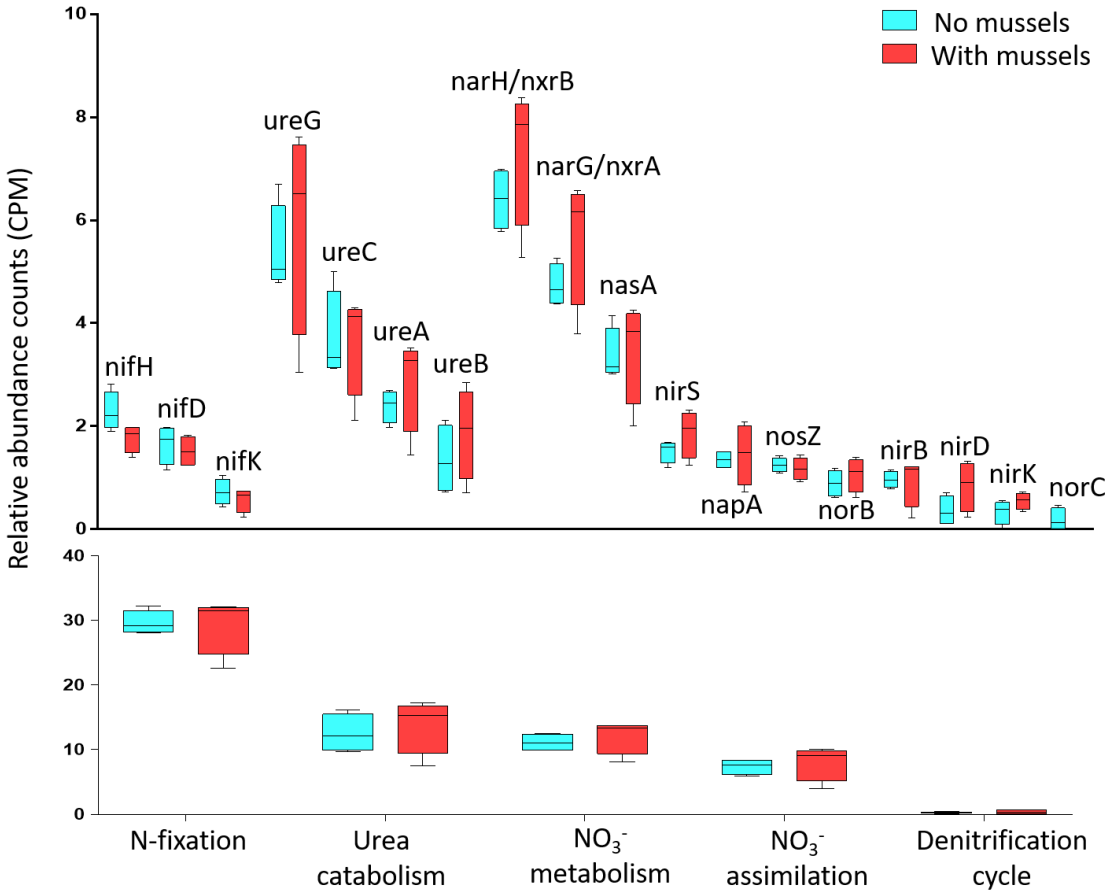


Figure 3.2 Relative abundance counts of the most abundant N-cycling pathways (bottom) grouped according to the GO database, and the corresponding orthologous groups (above) according to KOs within each pathway.

3.4.2 N-cycling biomarker discovery

In support of our hypothesis, nitrification functional genes were the most influential biomarker for the mussel metagenomes (**Table 3.2**), with NO_2^- oxidization and *amoA* gene families responsible for most of the differential abundance (**Figure 3.3**). For NO_2^- oxidation, the large increases in *nxrB* and *nxrC* gene abundances were attributed to protein clusters derived from the genome of *Nitrospira* (**Figure 3.5A, Table A3.2**) as were increases in *nxrA2* and *nxrA1*. The most differentially abundant *amoA* were similar to protein clusters aligning with comammox *amoA*, as shown through phylogenetic analysis (**Figure 3.4**, **Table A3.2**). In further support of our hypothesis, 2 *ureC* genes and 1 *Nitrosomonas hao* gene were differentially abundant with mussels (**Figure 3.5A, Table A3.2**). Although nitrification was the strongest biomarker for microbial communities with mussels, some genes in the denitrification pathway were more abundant with mussels (**Figure 3.5B and C**). These included 1 nosZ protein cluster and 4 *norB* gene families.

Furthermore, N-cycling genes that increased with mussels originated from taxons known for elemental cycling, such sulfur and methane transformation. For example, differentially abundant of protein clusters used in dissimilatory nitrate reduction, narG and narH (**Table A3.4**), were associated with the methane oxidizing *Methylobacter* and methanotrophic *Methylosarcina*, respectively. In one other example, one dissimilatory nitrite reducing cluster (NirB; **Figure 3.5D, Table A3.5**) originated from the methanotrophic genus, *Methyloglobulus*. N-fixation biomarker genes (*nifD*) (**Table A3.6**) were associated with a filamentous sulfur oxidizing genus, *Beggiatoa*, and mesophilic purple sulfur bacterial family, *Chromatiaceae*.

In comparison, NO reduction was revealed as the most evident biomarker (**Figure 3.3**) for UMR sediment microbial communities without mussels, with the abundance of the *norB* gene most responsible for this distinction (**Table 3.3**). Interestingly, the *nrfA* gene family was increased without mussels (**Table 3.3**), despite both DNRA and NO reduction pathways requiring NO_2^- as substrate. Although the nrfA protein was characterized as a biomarker of no-mussel microbial communities, only 2 *nrfA* genes families (**Figure 3.5D**) were differentially abundant. The higher ranking of NO reductase as a no-mussel biomarker was explained by 5 differentially abundant *norB* genes (**Figure 3.5C**), two of which originated from *Ochrobactrum* and *Zoogloea* genera. Numerous other denitrification genes (**Table A3.7**) were more abundant without mussels, including 13 different *nosZ* gene families (**Figure 3.5B**), 6 *narG*, and 4 *nirS* (**Figure 3.5D**), which provides evidence that the no-mussel microbial communities had a larger genetic potential for complete denitrification. In summation, a similar number of narGH orthologs were differentially abundant for both treatments while the samples without mussels contained a greater number of gene clusters responsible for NO reduction, N_2O reduction, and N fixation.

Table 3.2 Biomarker N-cycling pathways, functional roles, and gene families deemed differentially abundant with mussels. Genes encoding amoA, nxrB, and nxrC were statistically greater with mussels, and the higher order classifications of nitrite oxidation and nitrification were also statistically significant.

	N-cycling function or functional gene	LDA effect size	P-value
Level 1: N-cycle pathway	Nitrification	4.38	0.021
Level 2: N-cycle function	Nitrite oxidation	3.98	0.021
Level 3: Functional genes	nxrB	3.74	0.043
	nxrC	3.34	0.018
	amoA	3.73	0.021

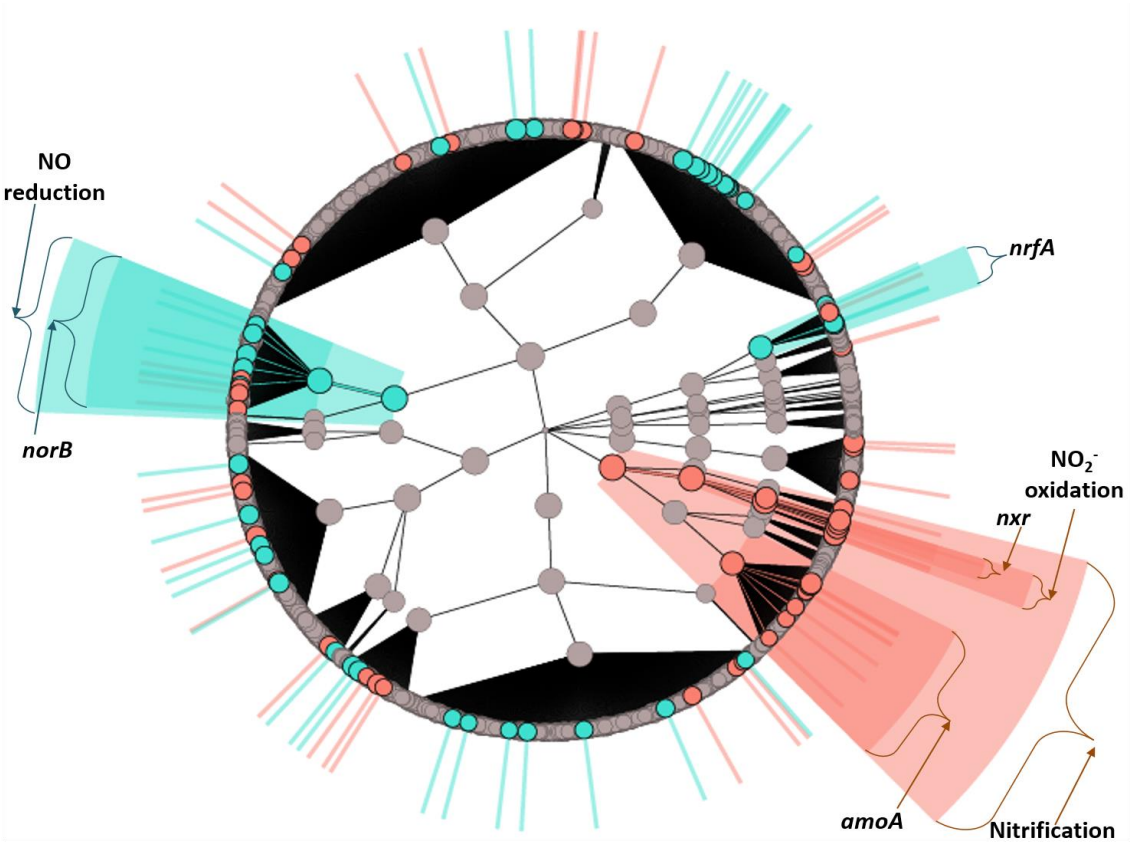


Figure 3.3 Cladogram produced by GraPhlAn depicting N-cycling functional genes present in the metagenomic samples, with the outermost circles representing abundances of individual UniRef90 protein clusters. Genes were placed in functional categories based on KO groups (i.e. nxr), enzymatic reactions (i.e. nitrite oxidation), and N-cycling pathways (i.e. nitrification) as described in the methods. Gene families and functional categories are labeled with colored circles if they were differentially abundant in the treatment with mussels (orange bars) and without mussel (green bars) and are shown with radial extensions beyond the cladogram. Circle sizes represent relative counts (CPM) in each category. Circles near the center represent N-cycling pathways (defined in **Table 3.1**), and categories become more specific as circles are farther from the cladogram center. Genes encoding Nxr and AmoA were the most differentially abundant features with mussels and resulted in a differentially abundant nitrification pathway. No-mussel samples were distinguished by increased NO reduction genes and were increased in abundances of norB and nrfA orthologs.

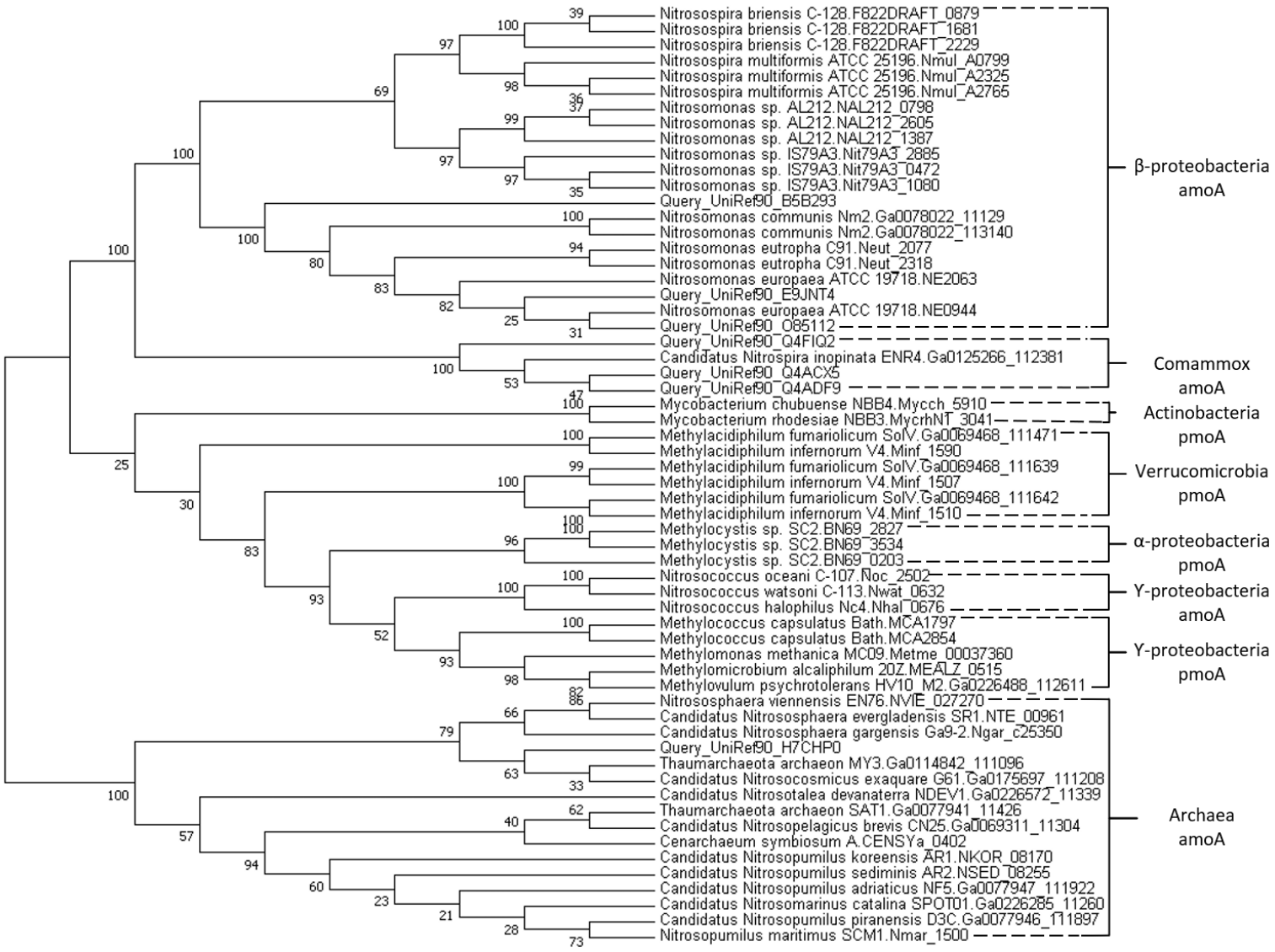
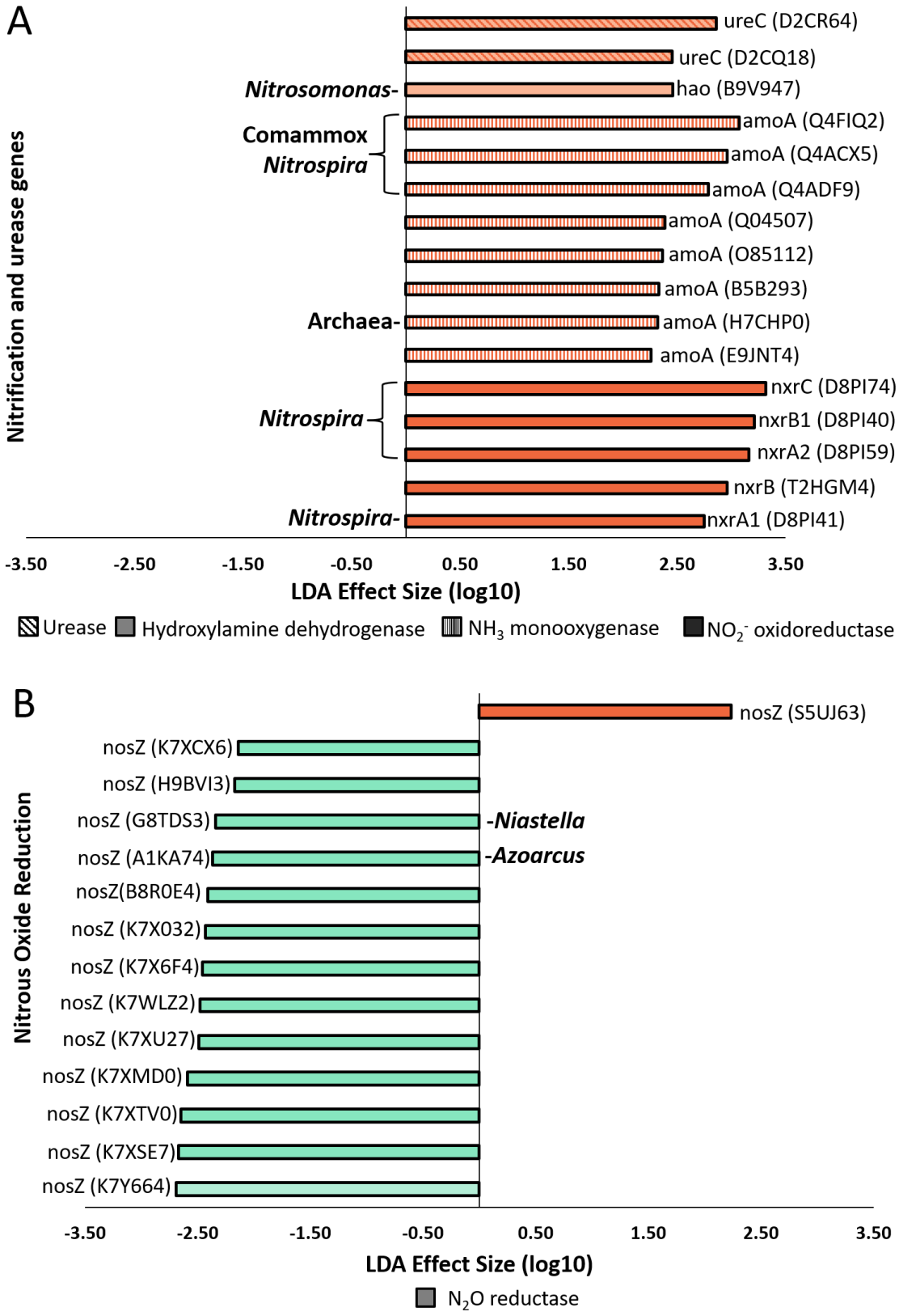
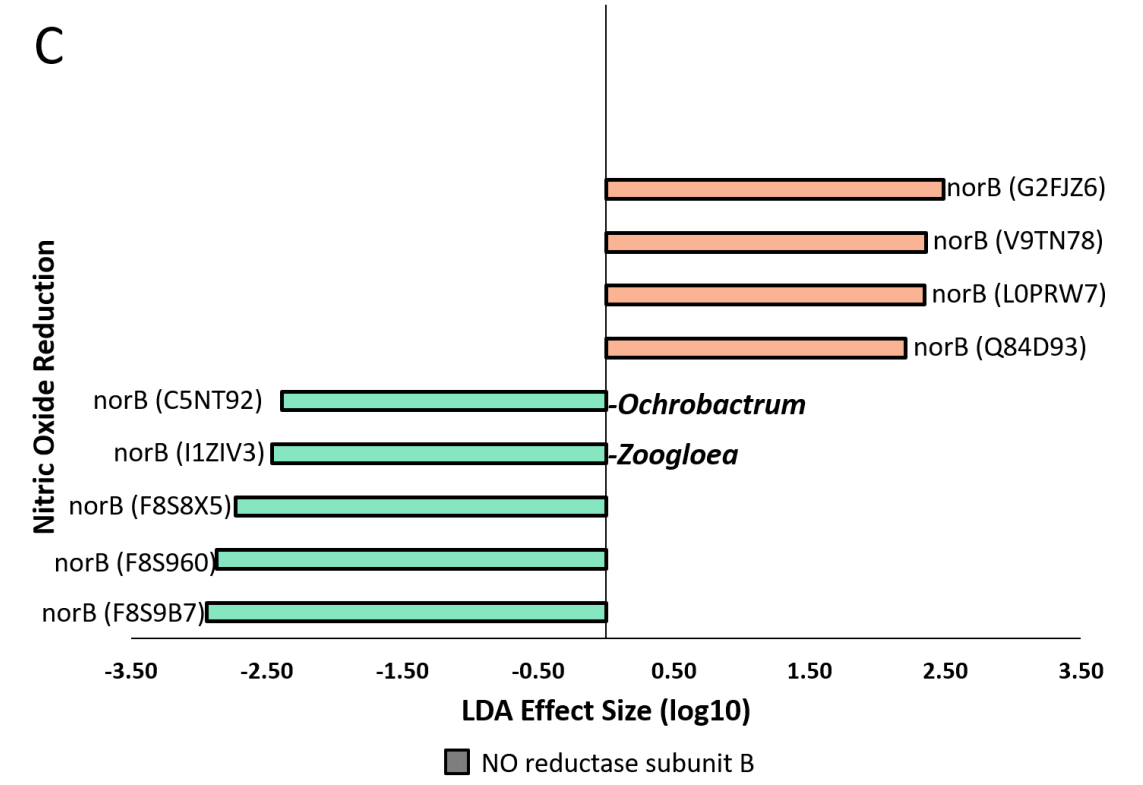

Nitrosospira briensis C-128.F822DRAFT_0879
Nitrosospira briensis C-128.F822DRAFT_1681
Nitrosospira briensis C-128.F822DRAFT_2229
Nitrosospira multiformis ATCC 25196.Nmul_A0799
Nitrosospira multiformis ATCC 25196.Nmul_A2325
Nitrosospira multiformis ATCC 25196.Nmul_A2765
Nitrosomonas sp. AL212.NAL212_0798
Nitrosomonas sp. AL212.NAL212_2605
Nitrosomonas sp. AL212.NAL212_1387
Nitrosomonas sp. IS79A3.Nit79A3_2885
Nitrosomonas sp. IS79A3.Nit79A3_0472
Nitrosomonas sp. IS79A3.Nit79A3_1080
Query_UniRef90_B5B293
Nitrosomonas communis Nm2.Ga0078022_11129
Nitrosomonas communis Nm2.Ga0078022_113140
Nitrosomonas eutropha C91.Neut_2077
Nitrosomonas eutropha C91.Neut_2318
Nitrosomonas europaea ATCC 19718.NE2063
Query_UniRef90_E9JNT4
Nitrosomonas europaea ATCC 19718.NE0944
Query_UniRef90_O85112
β-proteobacteria amoA
Query_UniRef90_Q4FIQ2
Candidatus Nitrospira inopinata ENR4.Ga0125266_112381
Query_UniRef90_Q4ACX5
Query_UniRef90_Q4ADF9
Comammox amoA
Mycobacterium chubuense NBB4.Mycch_5910
Mycobacterium rhodesiae NBB3.MycrhNT_3041
Actinobacteria pmoA
Methylacidiphilum fumariolicum SolV.Ga0069468_111471
Methylacidiphilum infernorum V4.Minf_1590
Methylacidiphilum fumariolicum SolV.Ga0069468_111639
Methylacidiphilum infernorum V4.Minf_1507
Methylacidiphilum fumariolicum SolV.Ga0069468_111642
Methylacidiphilum infernorum V4.Minf_1510
Verrucomicrobia pmoA
Methylocystis sp. SC2.BN69_2827
Methylocystis sp. SC2.BN69_3534
Methylocystis sp. SC2.BN69_0203
α-proteobacteria pmoA
Nitrosococcus oceani C-107.Noc_2502
Nitrosococcus watsoni C-113.Nwat_0632
Nitrosococcus halophilus Nc4.Nhal_0676
Y-proteobacteria amoA
Methylococcus capsulatus Bath.MCA1797
Methylococcus capsulatus Bath.MCA2854
Methylomonas methanica MC09.Metme_00037360
Methylomicrobium alcaliphilum 20Z.MEALZ_0515
Methylovulum psychrotolerans HV10_M2.Ga0226488_112611
Y-proteobacteria pmoA
Nitrososphaera viennensis EN76.NVIE_027270
Candidatus Nitrososphaera evergladensis SR1.NTE_00961
Candidatus Nitrososphaera gargensis Ga9-2.Ngar_c25350
Query_UniRef90_H7CHP0
Thaumarchaeota archaeon MY3.Ga0114842_111096
Candidatus Nitrosocosmicus exaquare G61.Ga0175697_111208
Candidatus Nitrosotalea devanaterra NDEV1.Ga0226572_11339
Thaumarchaeota archaeon SAT1.Ga0077941_11426
Candidatus Nitrosopelagicus brevis CN25.Ga0069311_11304
Cenarchaeum symbiosum A.CENSYa_0402
Candidatus Nitrosopumilus koreensis AR1.NKOR_08170
Candidatus Nitrosopumilus sediminis AR2.NSED_08255
Candidatus Nitrosopumilus adriaticus NF5.Ga0077947_111922
Candidatus Nitrosomarinus catalina SPOT01.Ga0226285_11260
Candidatus Nitrosopumilus piranensis D3C.Ga0077946_111897
Nitrosopumilus maritimus SCM1.Nmar_1500
Archaea amoA

Figure 3.4 Consensus tree representing the evolutionary relationship of amoA and pmoA amino acid sequences following 100 replications. The phylogenetic tree depicts groupings of amoA proteins belonging to comammox *Nitrospira*, Archaea, β-, γ-, and α-proteobacteria lineages, and various pmoA. AmoA queries are named by their UniRef90 accession, and amoA sequences obtained from reviewed genomes (IMG) are designated with the genome name and locus tag. There was a total of 149 parameters in the final dataset, and archaeal amoA was used as the outgroup.

Table 3.3 Biomarker N-cycling pathways, functional roles, and gene families with no mussels. The abundance of genes encoding norB and the higher order functional category of nitric oxide reduction were statistically greater in the no-mussel treatment, but the denitrification pathway was not significantly different. nrfA was statistically more abundant, but the higher order categories of nitrite reduction in the DNRA pathway were not statistically significant.

No-Mussel sediment	N-cycling function or functional gene	LDA effect size	P-value
Level 1: N-cycle pathway	Denitrification	-	-
Level 2: N-cycle function	Nitric oxide reduction	4.03	0.021
Level 3: Functional gene	*norB*	3.88	0.021
Level 1: N-cycle pathway	DNRA	-	-
Level 2: N-cycle function	Nitrite reduction	-	-
Level 3: Functional gene	*nrfA*	3.87	0.043


A
Nitrification and urease genes
ureC (D2CR64)
ureC (D2CQ18)
Nitrosomonas-
hao (B9V947)
Comammox
Nitrospira
amoA (Q4FIQ2)
amoA (Q4ACX5)
amoA (Q4ADF9)
amoA (Q04507)
amoA (O85112)
amoA (B5B293)
Archaea-
amoA (H7CHP0)
amoA (E9JNT4)
Nitrospira
nxrC (D8PI74)
nxrB1 (D8PI40)
nxrA2 (D8PI59)
nxrB (T2HGM4)
Nitrospira-
nxrA1 (D8PI41)
-3.50
-2.50
-1.50
-0.50
0.50
1.50
2.50
3.50
LDA Effect Size (log10)
Urease
Hydroxylamine dehydrogenase
NH3 monooxygenase
NO2- oxidoreductase
B
Nitrous Oxide Reduction
nosZ (S5UJ63)
nosZ (K7XCX6)
nosZ (H9BVI3)
nosZ (G8TDS3)
-Niastella
nosZ (A1KA74)
-Azoarcus
nosZ(B8R0E4)
nosZ (K7X032)
nosZ (K7X6F4)
nosZ (K7WLZ2)
nosZ (K7XU27)
nosZ (K7XMD0)
nosZ (K7XTV0)
nosZ (K7XSE7)
nosZ (K7Y664)
-3.50
-2.50
-1.50
-0.50
0.50
1.50
2.50
3.50
LDA Effect Size (log10)
N2O reductase


C
Nitric Oxide Reduction
norB (G2FJZ6)
norB (V9TN78)
norB (L0PRW7)
norB (Q84D93)
norB (C5NT92)
-Ochrobactrum
norB (I1ZIV3)
-Zoogloea
norB (F8S8X5)
norB (F8S960)
norB (F8S9B7)
-3.50
-2.50
-1.50
-0.50
0.50
1.50
2.50
3.50
LDA Effect Size (log10)
NO reductase subunit B

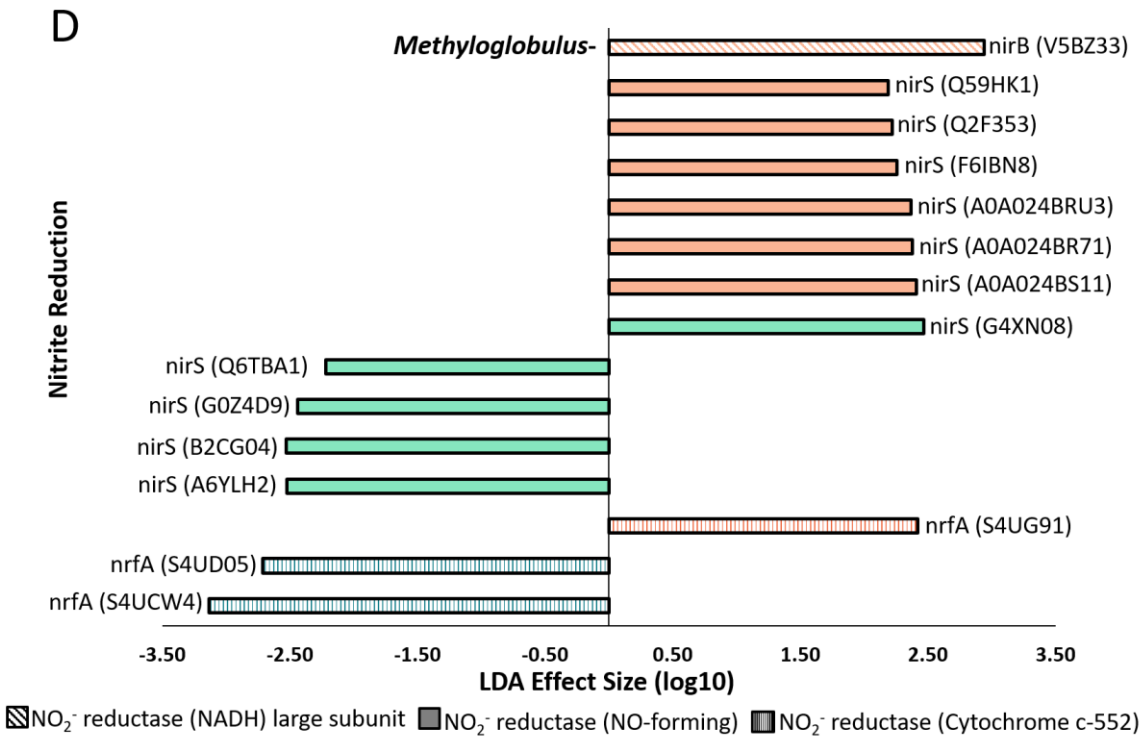

D
Nitrite Reduction
Methyloglobulus-
nirB (V5BZ33)
nirS (Q59HK1)
nirS (Q2F353)
nirS (F6IBN8)
nirS (A0A024BRU3)
nirS (A0A024BR71)
nirS (A0A024BS11)
nirS (G4XN08)
nirS (Q6TBA1)
nirS (G0Z4D9)
nirS (B2CG04)
nirS (A6YLH2)
nrfA (S4UG91)
nrfA (S4UD05)
nrfA (S4UCW4)
-3.50
-2.50
-1.50
-0.50
0.50
1.50
2.50
3.50
LDA Effect Size (log10)
NO2- reductase (NADH) large subunit
NO2- reductase (NO-forming)
NO2- reductase (Cytochrome c-552)

Figure 3.5 Linear discriminant analysis scores for N-cycle protein clusters. Each bar represents the effect size of mussels on a specific gene, with negative LDA scores representing no-mussel samples, and positive scores corresponding with mussel samples. Genes with an LDA score less than 2 are not depicted. Genes are labeled with taxonomic origin according to mapping of LDA to respective UniRef90 IDs and taxonomic origin was not included for those without a LCA defined at the genus level. **A)** N-transforming gene clusters with significant LDA scores included genes encoding urease, hydroxylamine oxidoreductase, ammonia monooxygenase, and nitrite oxidoreductase. Abundances of *amoA* and *nxr* genes were most responsible for the 4.3 effect size of the nitrification pathway with mussels. **B)** Genes encoding nitrous oxide reduction experienced effect sizes up to 2.7 without mussels. **C)** Genes encoding nitric oxide reductases were shown to be a biomarker of no-mussel metagenomes and had effect sizes up to 2.9. **D)** Dissimilatory nitrite reductase genes in the DNRA pathway (NirB, NrfA) had larger effect sizes than the NO-forming protein clusters (NirS) belonging to the denitrification pathway.

3.5 Discussion

3.5.1 Denitrification gene families dominated N-cycling in UMR sediment

As a whole, UMR sediment microbial communities contained the greatest genetic potential for NO_2^-/NO_3^- redox by way of *narGH* and *nxrAB*. This may be explained by reliably high NO_3^- loads found in the UMR agroecosystem [214], with concentrations measured near 14-18 mg/L [46, 215] throughout the UMR watershed. Furthermore, our findings are consistent with previous studies which showed associations between high N in the UMR and denitrification rates [216]. These results show consistency with aquatic sediments outside the UMR, where the genomic potential for denitrification dominated oligotrophic sediments and showed minor contributions of DNRA and anammox in driving N-cycling [40]. Ultimately this suggests that sediment microbial communities in the UMR have the genetic capability to mitigate high NO_3^- loadings, and warrants further research into biostimulation as technique to reduce non-point N concentrations.

3.5.2 Nitrification biomarkers in sediments with mussels

In confirmation of our hypothesis, the UMR mussel bed sediment contained microbial communities with increased genetic potential for NH_3 and NO_2^- oxidation, as well as a differentially abundant *hao* gene family (**Figure 3.6**). The LEfSe biomarker analysis revealed that nitrification pathways were the most definitive biomarker of N-cycling microbial communities with mussels and was largely due to increased genetic potential for NO_2^-oxidation. Furthermore, we identified nitrification biomarkers belonged to the genera *Nitrospira* and *Nitrosomonas* and also matches our previous findings from 16S rRNA amplicon sequencing.

It is possible that mussels had the most impact on nitrification genes because their biodeposition products increase porewater NH_3 concentrations [13] and enhance the flux of N between water and sediment [217]. Other studies have found significantly greater AOB *amoA* genes corresponding with a higher NH_3 load [218] and aerophilic conditions [219]. The most distinct nitrification genes were most closely related to NOB *Nitrospira*, and comammox *Nitrospira*. It is not surprising that *Nitrospira* species dominated the nitrification biomarkers due to their metabolic diversity [33, 220], domination within freshwater sediments [221], and increased abundance in sediments with mussels [203, 222], and greater abundance from invertebrate bioturbation activities [223]. Finding *amoA* biomarkers from comammox *Nitrospira* clades suggests that the presence of mussels may enhance the genetic potential for complete nitrification.

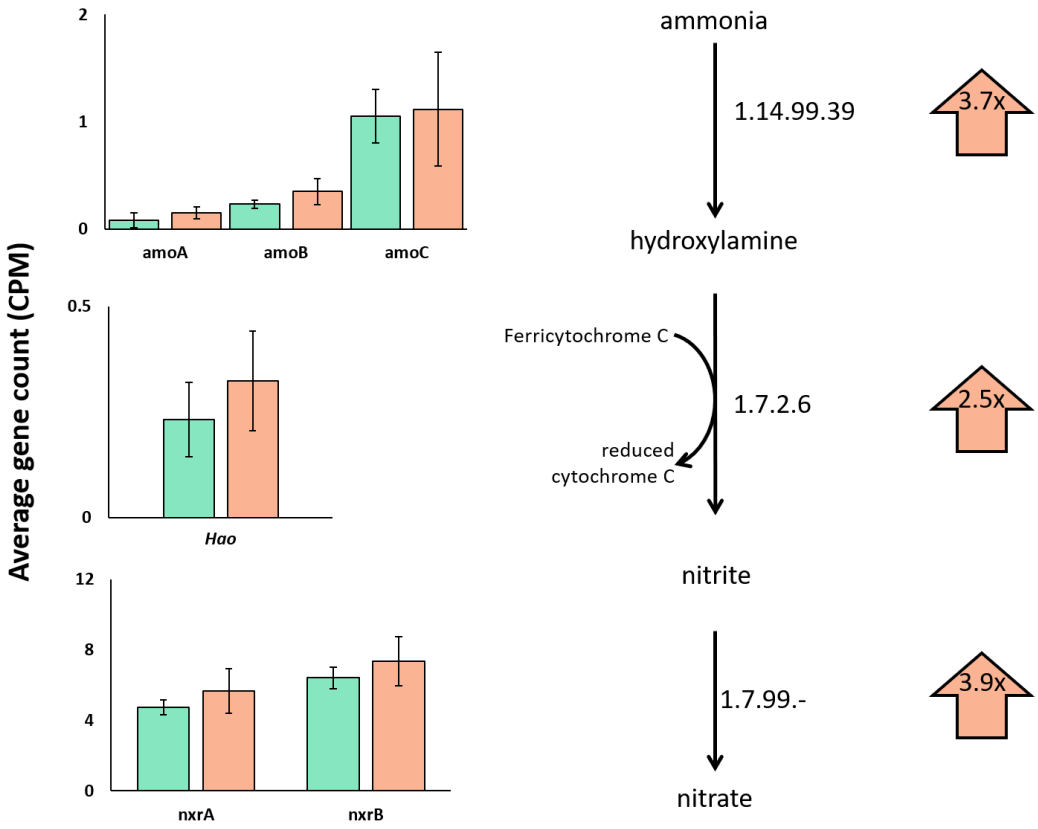


Figure 3.6 Representation of the KEGG nitrification module (M00528) and KEGG orthologous groups composing each reaction (EC numbers) that were increased with mussels. Average counts (CPM) of KOs are depicted in bar graph form (No mussels=green colored bars; With mussels=orange colored bars), with error bars representing standard deviations, and large orange-colored arrows representing the LDA effect size with mussels, as determined by LefSe. The genomic potential for ammonia oxidation with ammonia monooxygenase (*amoA*) was a biomarker with mussels (top), one hydroxylamine hydrogenase protein cluster was increased with mussels (middle), and nitrite oxidation was the most definitive biomarker of the N-cycling communities with mussels (bottom). Overall, the nitrification pathway had an LDA effect size of 4.4 in the treatment with mussels.

3.5.3 Implications of freshwater mussels on N-cycling

Microbial communities without mussel influences had greater metabolic potential for NO reduction and contained high ranking biomarker genes *norB* and *nrfA*. It makes sense that mussels suppressed the genomic potential for NO_2^- and NO reduction since these processes are in opposition to NO_2^- oxidation. Furthermore, our results match a study which found higher abundances of *Nitrospira* near the water-sediment interface of NH_3-enriched, mixed and homogenized sediment, at the expense of DNRA [224]. The suppression of DNRA by mussels would be an important ecosystem service because DNRA recycles bioavailable N and promotes a positive feedback of eutrophication [225]. *NrfA* abundance has been positively correlated to sediment C:N concentrations [226], so it is possible that mussel assemblages lowered this biogeochemical ratio from biodeposition products, and resulted in a suppressed DNRA pathway by microbial communities.

Our main findings of decreased *norB* abundance may be explained by mussel bioturbation activity and aeration of the sediment. One study showed that microaerophilic conditions affect denitrification rates, and decreased *norB* transcripts when O_2 concentrations exceeded 200 nM [227]. Results of decreased genomic potential for NO reduction suggests that mussels could indirectly decrease the production of N_2O, a potent greenhouse gas [228, 229], in UMR sediments. This is an important finding, as studies have noted that denitrification in the UMR is a major source of atmospheric N_2O [230], and N_2O emissions from upper Midwest agroecosystem were primarily from soil [228]. Turner et al. [230] also projected that a doubling in aquatic N concentrations would result

in a 40% increase in N_2O emissions from denitrification in the UMR, and illustrates that mussels may provide a buffering capacity towards future N_2O emissions.

3.6 Conclusion

Metagenomic sequencing of UMR sediments revealed a large genomic potential for nitrate metabolism and minor abundance of genes for anaerobic NH_3 oxidation and DNRA pathways. The presence of a well-established freshwater mussel assemblage in this agroecosystem resulted in significantly increased nitrification potential at the expense of DNRA and NO reduction to N_2O. In support of these findings, *amoA* and *nxr* genes were the most predominant biomarkers of mussel bed, and the most defining genes were associated with comammox *Nitrospira* and NOB *Nitrospira*, respectively. Additionally, our results provide evidence that mussels may offer a buffer against N_2O production by suppressing *norB* and prevent a positive feedback for eutrophication via reducing the abundance of *nrfA* genes. Overall, this research demonstrated the genomic potential of N-cycling microbial communities were impacted by freshwater mussels in a high nutrient agroecosystem.

CHAPTER 4: THE GENOMIC POTENTIALS OF NITRITE OXIDIZING BACTERIA AND COMPLETE AMMONIA OXIDIZING NITROSPIRA IN RIVER SEDIMENT ARE IMPACTED BY NATIVE FRESHWATER MUSSELS

4.1 Abstract

Freshwater mussel assemblages of the Upper Mississippi River sequester tons of ammonia- and urea-based biodeposits each day and aerate sediment through burrowing activities, thus creating a unique niche for nitrogen (N) cycling microorganisms. This study explored how mussels impact the abundance of N-cycling species with an emphasis on *Candidatus* Nitrospira inopinata, the first microorganism known to completely oxidize ammonia (comammox) to nitrate. This study used metagenomic shotgun sequencing of genomic DNA to compare nitrogen cycling species in sediment under a well-established mussel assemblage and in nearby sediment without mussels. Metagenomic reads were aligned to the prokaryotic RefSeq non-redundant protein database using BLASTx, taxonomic binning was performed using the weighted lowest common ancestor algorithm, and protein-coding genes were categorized by metabolic function using the SEED subsystem. Linear discriminant analysis (LDA) effect sizes were used to determine which metagenomes and metabolic features explained the most differences between the mussel habitat sediment and sediment without mussels. Of the N-cycling species deemed differentially abundant, *Nitrospira moscoviensis* and *Ca*. Nitrospira inopinata were responsible for creating a distinctive N-cycling microbiome in the mussel habitat sediment. Further investigation revealed that comammox *Nitrospira* had a large metabolic potential to degrade mussel biodeposits, as evidenced the top ten percent of

protein-coding genes including the cytochrome c-type biogenesis protein required for hydroxylamine oxidation, ammonia monooxygenase, and urea decomposition SEED subsystems. Biomarker analysis of these two *Nitrospira* taxons suggested that *N. moscoviensis* was most impacted by diverse carbon metabolic processes while *Ca*. Nitrospira inopinata was most distinguished by multidrug efflux proteins (acrB), NiFe hydrogenase (hypF) used in hydrogen oxidation and sulfur reduction coupled reactions, and a heme chaperone (ccmE). Furthermore, our research suggests that comammox and NOB *Nitrospira* likely coexisted by utilizing mixotrophic metabolisms. For example, *Ca*. Nitrospira inopinata had the largest potentials for ammonia oxidation, nitrite reduction with nirK, and hydrogen oxidation, while NOB *Nitrospira* had the greatest potential for nitrite oxidation, and nitrate reduction possibly coupled with formate oxidation. Overall, our results suggest that this mussel habitat sediment harbors a niche for NOB and comammox *Nitrospira*, and ultimately impacts N-cycling in backwaters of the Upper Mississippi River.

4.2 Introduction

Water quality of the Upper Mississippi River (UMR) has been documented for decades [193], yet the UMR basin contributes over 50,000 metric tons of bioactive nitrogen (N) to the Gulf of Mexico each year [194]. Research has shown that microbial communities are impacted by the addition of bioactive N [231], and subsequently alter N-biogeochemical cycling in the UMR through nitrification and denitrification processes [232, 233]. Enhancing the vertical exchange between overlying water and groundwater (i.e. water-sediment interface) of UMR backwater channels has been proposed to significantly enhance N removal [234, 235], particularly because biotic removal of N

reaches a maximum efficiency of 40% as N loads increase in large streams [233] and denitrification rates plateau as nitrate (NO_3-N) reaches 5 mg/L in backwater channels [195]. Taken together, these findings emphasize the large N-cycling potential of benthic organisms, by enhancing the flux of nutrients into sediment for microbial transformations [11, 184].

Freshwater mussels (order Unionidae) native to the UMR live in assemblages of 3 to 5 mussels m^{-2}, collectively filter billions of gallons of water, and remove tons of N-containing biomass from overlying water each day [45]. In addition to the ecosystem services of water filtration and enhancing nutrient exchange rates across the water-sediment interface [58], mussel excretion of feces and pseudofeces (biodeposition products) sequesters ammonia (NH_3) and carbon (C) into sediment porewater [13, 45, 59]. As a result, mussel assemblages are attributed with creating "hotspots" of N and C in surrounding sediment [236], and create a microbial niche ripe for nitrification at the interface of oxic and anoxic conditions [203].

Nitrifying organisms capable of mixotrophy may pose an advantage in a mussel-influenced habitat, owing to the adaptation of switching metabolic functions when conditions change from oxic to anoxic [237, 238]. It was previously thought that nitrite (NO_2^-) oxidizing bacteria (NOB) were restricted to oxic environments where ammonia (NH_3) oxidizing bacteria (AOB) produce NO_2^-, but recent genomic analyses have expanded the known metabolic functions of conventional NOB *Nitrospira*. For example, the NO_2^- oxidizing species *Nitrospira moscoviensis* is genetically capable of cyanate degradation [34], aerobic hydrogen oxidation [35], and formate oxidation coupled with nitrate (NO_3) reduction [33]. Nitrification was further expanded after discovering that

Nitrospira moscoviensis can produce NH_3 and CO_2 by way of urea hydrolysis, and can reciprocally feed NH_3 to urease-lacking AOB and receive NO_2^- in return [33]. Furthermore, the N-cycle was transformed after the discovery of a single organism capable of complete NH_3 oxidation (comammox) [200] and confirmation of genes required for complete nitrification encoded by *Candidatus* Nitrospira inopinata [199], and potentially even sulfur reduction [239].

In a previous study, we showed that sediment of a well-established mussel habitat in UMR backwaters contained an enhanced niche for *Nitrospirae* in addition to a greater abundance of microorganisms indicative of an oxic-anoxic niche, like anaerobic ammonium oxidizers (anammox). This presumed oxic-anoxic niche was detected closer to the water-sediment interface in the mussel habitat, since the relative abundance of anammox bacteria peaked at shallow (3 cm) sediment depths with mussels, but were more abundant in deeper (5 cm) sediments in the no-mussel treatment [203]. Furthermore, the 16S rRNA amplicon survey showed fewer differences among N-cycling phylotypes in shallow sediment with mussels and deeper sediment without mussels (i.e. intrasample differences), and the fewest differences when comparing the shallow mussel sediment against the deeper no-mussel sediment (inter-sample differences). In response, this study used the deeper no-mussel sediment as the most stringent baseline to assess mussel influences on the N-cycling community with mussels. We employed metagenomic shotgun sequencing of total DNA corresponding with the aforementioned oxic-anoxic niche sediment components, with the goal of identifying the N-cycling species most impacted by mussels. We hypothesized that sediment from the mussel

habitat would contain an increased abundance of nitrifying taxons, presumably due to an enhanced genomic potential for ammonia oxidation.

4.3 Materials and Methods

4.3.1 Sediment collection and DNA isolation

Our study sites were located in the backwaters of the Upper Mississippi River navigation pool 16, where a consistently populated mussel assemblage has been studied for decades [44, 240-242]. Sediment cores were obtained from the mussel habitat (41.452804, -90.763299) and upstream sediment (41.451540, -90.753275) lacking mussels [203]; both sites had similar hydraulics and sediment composition [240], and will be considered as treatments "with-mussels" and with "no-mussels" according to previous studies [243, 244]. Cores were removed from each site using a 2-in diameter, post-driver with a polypropylene liner (Multi-State Sediment Sampler, Art's Manufacturing and Supply, Inc.; American Falls, ID, USA), and an ethanol flame-sterilized 3/8-in diameter drill bit was used to penetrate the cores at depths of 3 and 5 cm. For each core, samples (0.25g sediment) were removed in quadruplicate (n=4, 3 cm depth with mussels; n=4, 5 cm depth without mussels) and stored in sterile bead beating tubes overnight at -20^{o}C. Genomic DNA was isolated (PowerSoil DNA Isolation Kit; MoBio Laboratories, Inc., Carlsbad, CA, USA) and stored at -20^{o}C. Following verification of DNA quality and quantity (NanoDrop 1000; Thermo Fisher Scientific, Waltham, MA), genomic DNA was sequenced at the University of Iowa Institute for Human Genetics (IIHG). As mentioned earlier, selection of these samples were informed by evidence suggesting oxic-anoxic interface niches were located at 5 cm sediment depth without mussels and 3 cm depth beneath mussels [203]. Samples chosen for this experiment correspond to the following

16S rRNA amplicon sequencing data at MG-RAST: without-mussels (mgm4705698.3, mgm4705704.3, mgm4705686.3, mgm4705697.3) and with-mussels (mgm4705708.3, mgm4705672.3, mgm4705699.3, mgm4705680.3).

4.3.2 Metagenomic shotgun sequencing

For each sample, 120 ng of genomic DNA in 60 µL of 10mM Tris-HCl, pH 8.0 buffer, was placed into 1.5 mL RNase-/DNase-free, low binding microcentrifuge tubes. Library creation steps were performed by the IIHG Genomics Division and included DNA shearing using the Covaris Adaptive Focused Acoustics™ process (Covaris E220 Focused-ultrasonicator; Covaris, Inc., Woburn, MA), and DNA fragment purification and end polishing (KAPA Hyper prep kits; Kapa Biosystems, Inc., Wilmington, MA) prior to ligation to indexed adaptors. The library size distribution was validated using the Agilent 2100 Bioanalyzer Instrument (Agilent Technologies, Santa Clara, CA), and quantified using the q-PCR KAPA library amplification module following manufacturer instructions (Kapa Biosystems, Inc.). The indexed libraries were normalized, pooled, and clustered on a flow cell using the cBOT Cluster Generation System (Illumina, Inc., San Diego, CA) and sequenced on the Illumina HiSeq 4000 System (Illumina, Inc.) in high output mode (1 lane, 2x150bp). FASTQ files are accessible at ENA (Study Accession: PRJEB23134) and NCBI repositories (BioProject ID: PRJNA414922), and MG-RAST contains QA/QC and analyses of metagenomes (MG-RAST project mgp21252) (**Table A3.1**).

4.3.3 Bioinformatics pipeline

FastQC [245] was used for quality control and revealed an average sequence abundance of 42,606,252 ± 2,365,531 sequences, sequence lengths of 151 bp, and 60% (± 0.84%) GC content for no-mussel samples, and 41,402,435 ± 3,444,191 sequences,

with sequence lengths of 151 bp, and 60% (±1.31%) GC content for samples with mussels. For taxonomic and functional binning of reads, we employed the streamlined DIAMOND [206] and MEGAN [246] pipeline specialized for microbiome shotgun sequencing analyses [247]. First, RefSeq [248] non-redundant (nr) archaeal and bacterial protein sequences (Release80) were concatenated to construct a database for BLASTx alignments in DIAMOND using the "make.db" command, and pairwise alignments were performed using the default BLASTx settings (BLOSUM62 matrix, γ= 0.267, K=0.041, Penalties=11/1). The aligned reads from DIAMOND were imported into MEGAN using *daa-meganizer* [247], keeping only the top 100 matches per read. The weighted lowest common ancestor (LCA) algorithm was used for taxonomic binning using default settings in MEGAN6 [247] (min score=50.0, max expected=0.01, top percent= 10.0%, min support percent=0.05, min support=1, 80% coverage for weighted LCA algorithm) and classified according to NCBI taxonomy IDs (Nov 2016 release).

Furthermore, aligned reads were assigned functional roles using accession mapping files for SEED subsystems [249] (SEED May 2015 annotation). The resulting files contained all reads, alignments, taxonomic and functional classifications, and were normalized for sampling read depth (normalized to 12,726,950 reads per sample) and assigned metadata categories, "no-mussel" and "with-mussel". Reads aligned to N-cycling genomes were extracted from MEGAN for differential abundance analysis using LefSe [209]. The most differentially abundant genomes belonged to *Candidatus* Nitrospira inopinata (GCF_001458695.1) and *Nitrospira moscoviensis* (GCF_001273775.1), and were further assembled and annotated in Unipro UGENE [250] and depicted using DNAPlotter [251]. The UGENE workflow included mapping reads to

indexed reference genomes using BWA MEM [252] with default settings, followed by filtering and sorting the BAM files using SAMtools [253], and a final quality control step using FastQC [245].

4.3.4 LDA effect size

A linear discriminant analysis (LDA) method was used to assess which genomes and genomic features were most discriminative of the freshwater mussel habitat. N-cycling taxonomies of interest were chosen based on previous research [254] and included AOB and NOB phylotypes with the prefix "nitro", N-reducing phylotypes (**Table 1.1**) designated by "denitrificans" or "nitroreducens", and anammox candidate genera, *Brocadia* and *Jettenia*. The relative abundance of reads aligned to these N-cycling taxons were assessed for LDA effect size (LEfSe) [209] to determine which N-cycling taxonomic features were most responsible for differences in the mussel habitat microbiomes. All samples were labeled by class ("mussel" and "no-mussel") and features were compared for differential distribution using the non-parametric factorial Kruskal-Wallis sum-rank test (alpha= 0.05) and normalized to a total read count of 1 M. Features deemed differentially abundant were compared for significant effect size using the pairwise Wilcoxon rank-sum test (alpha= 0.05; "all against all"), and ranked features according to greatest effect size. A minimum LDA score of 2.0 was chosen as a cutoff for significant features to limit analysis to the most distinctive metagenomic traits. *Candidatus* Nitrospira inopinata and *Nitrospira moscoviensis* were shown to be the most distinctive genomes with mussels, and follow-up biomarker tests were performed for SEED assignments to address which metabolic functions may be responsible for this differentiation.

4.4 Results

4.4.1 N-cycling taxonomic composition

The DIAMOND/MEGAN pipeline revealed metagenomic reads assigned to N-cycling organisms were slightly more abundant with mussels (157,275 ± 17,503 reads) than without mussels (136,884 ± 20,982 reads), and the mussel habitat contained more reads belonging to N-cycling bacterial lineages (**Figure 4.1**) with *Nitrospirae* representing the most bacteria. In both treatments, *Candidatus* Methanoperdens nitroreducens represented the largest number of archaeal metagenomic reads but was differentially abundant between treatments. NO_2^- oxidizing organisms represented the largest N-cycling group, with *Nitrospira moscoviensis*, *Nitrospira defluvii*, and *Candidatus* Nitrospira inopinata comprising an average of 22%, 15%, and 11% of N-cycling metagenomic reads in the mussel habitat, respectively. *Steroidobacter denitrificans* was also a highly abundant component of the N-cycling community (12-14%) but was not differentially abundant between treatments.

LDA effect size analysis of the metagenomic reads assigned to N-cycle taxons further emphasized the major differences in the mussel habitat (**Figure 4.1**). Of the taxons considered, bacterial lineages were most impacted by mussels (LDA=4.27, P=0.043), and the most differentially abundant species were *Nitrospira moscoviensis* (LDA= 3.80, P=0.021) and *Candidatus* Nitrospira inopinata (LDA=3.63, P=0.021). One other group of nitrifying taxons were differentially abundant with mussels and belonged to the *Nitrosococcus* genus (LDA=2.20, P=0.021). Multiple denitrifying taxons were greater with mussels, including *Methylomonas denitrificans*, *Denitrobacterium detoxificans*, *Competibacter denitrificans*, and a Gammaproteobacterial sulfur oxidizing

symbiont, *Thiohalorhabdus denitrificans*. However, these denitrifying species were lower in abundance than *Nitrospira*, and thus were ranked lower as biomarker species. Protein functional assignments of *Ca*. Nitrospira inopinata and *N. moscoviensis* were also assessed for distinctive features between the mussel and no-mussel treatments, with the goal of discovering niche differentiating functions responsible for the enhanced abundance of NOB and comammox *Nitrospira* genomes [255].

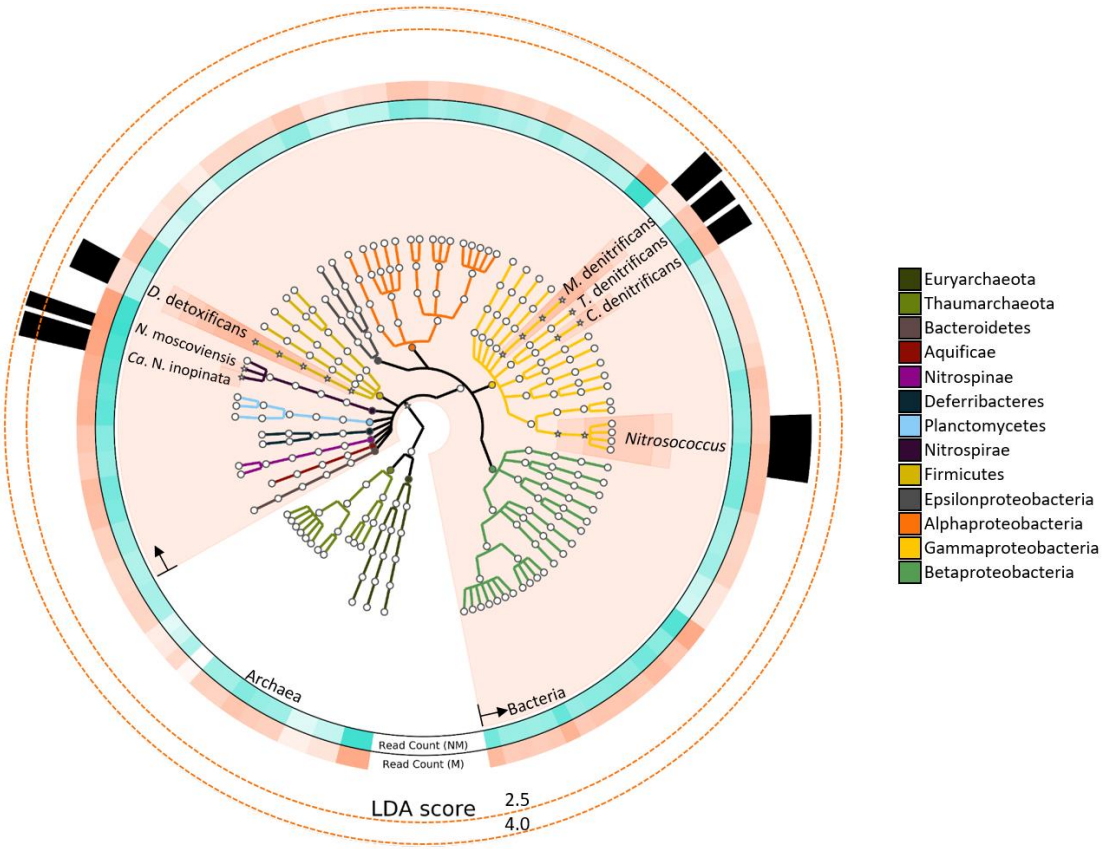


Figure 4.1 Nitrogen-cycling taxonomies assessed for linear discriminant analysis (LDA) effect size are colored based on phylum, as specified in the legend. Taxons shown with a "*" icon and salmon-colored background had statistically significant LDA scores (LDA>2, P<0.05). Rings surrounding the phylogenetic tree depict the relative abundance of reads assigned to the respective species. Ring color intensity represents relative read count with mussels (salmon-colored; "M") and with no mussels (turquoise-colored; "NM). The opacity of "Read count" ring-segments corresponds to the greatest taxonomic abundance. Several N-cycling taxons were differentially abundant with mussels and LDA effect sizes are represented by the height of black bars in the outer-most ring.

4.4.2 *Nitrospira moscoviensis* genomic potential

A total of 435,151 SEED protein functions were assigned to the genome of *N. moscoviensis* (normalized to 36,562 per sample) and was dominated by 5 SEED categories: carbohydrates, cofactors, vitamins, prosthetic groups, pigments, amino acids and derivatives, protein metabolism, and DNA metabolism (**Table A4.1**). The 25% most abundant SEED subsystems included those indicative of metabolic activity and growth, such as peptidoglycan and cytoskeleton biosynthesis, respiration and carbon fixation, DNA replication and repair (**Table A4.1**). Highly abundant protein functions unique to *N. moscoviensis* in the no-mussel treatment included motility and chemotaxis protein groups, biotin synthesis, and thiamin metabolism ("5-FCL-like protein"), while the mussel habitat treatment was dominated by folate and cysteine biosynthesis, carbon cycling ("alpha carboxysome"), and the DNA regulatory proteins YebC and proteasomes (**Table A4.1**).

Of the N-metabolism functional assignments (**Table 4.1**), formate hydrogenases were the most abundant N-cycling category for both treatments, and numerous enzymatic functions were relatively more abundant in the mussel habitat treatment. These included an NH_4^+ permease, NO reductase proteins (NorD and NorQ), formate dehydrogenase subunits (beta and chain D), NO_2^-/NO_3^- transporters and sensors, periplasmic nitrate reductases (NapG and NapF), and urease proteins (ureA and ureF) (**Table 4.1**).

Table 4.1 *N. moscoviensis* N-cycling protein functions from the mussel habitat in relative abundance (RPKM) and as a proportion of SEED enzymatic function. SEED counts are shown as a percentage of the total classified protein function in the mussel habitat.

SEED subsystems and protein functions	Gene abundance (RPKM)	Proportion of total function
Ammonia assimilation		
Nitrogen regulatory protein P-II Gln	1409.04	0.34%
Nitrogen assimilation regulatory protein Ntr	395.687	
Ammonium transporter amtB	300.22	
Glutamate-ammonia-ligase glnEb	3267.89	
Denitrification		
Ferredoxin-type protein NapG (periplasmic nitrate reductase)	801.88	0.13%
Copper-containing nitrite reductase nirK (NO-forming)	ND	
Formate hydrogenase		
Formate hydrogenlyase, membrane subunit HyfB	95.37	0.71%
putative Formate hydrogenlyase, membrane subunit HyfC	ND	
putative formate hydrogenlyase, membrane subunit HyfE	214.91	
Formate hydrogenlyase, membrane subunit HyfF	158.95	
Formate hydrogenlyase, large subunit hyfG	ND	
putative formate hydrogenlyase, small subunit hyfI	94.81	
Formate dehydrogenase, alpha subunit fdsA	17.577	
Formate dehydrogenase, beta subunit fdsB	60.937	
Formate dehydrogenase, gamma subunit fdsG	ND	
formate hydrogenlyase transcriptional activator fhlA	122.83	
formate transporter focA	ND	
Nitrate and nitrite ammonification		
Nitrate transporter narK	80.79	0.15%
Nitrate ABC transporter nrt	333.27	
Nitrite oxidoreductase, alpha subunit nxrA	1097.03	
Nitrite oxidoreductase, beta subunit nxrB	ND	
Nitrite oxidoreductase, membrane subunit nxrC	1695.47	
Nitrite reductase (NADH) small subunit nirD	474.06	
Nitrogen fixation		
Nitrogenase (molybdenum-iron)-specific transcriptional regulator NifA	63.33	0.15%
Nitrogenase (iron-iron) transcriptional regulator	ND	
Urea Degradation		
Urease alpha subunit UreC	139.91	0.10%
Urease gamma subunit UreA	159.583	
Urea ABC transporter, urea binding protein UrtA	36.72	

Table 4.1 cont.

SEED subsystems and protein functions	Gene abundance (RPKM)
Urease accessory protein UreD	ND
Urease accessory protein UreF	70.38
Urease accessory protein UreG	143.91
Urease beta subunit UreB	ND

4.4.3 *Nitrospira moscoviensis* biomarkers

Genetic code processing was a major function of *N*. moscoviensis in the mussel habitat, as evidenced by biomarker proteins used in DNA, RNA, and protein metabolism (LDA up to 3.32). Two of the most distinct features were YebC-like DNA-binding regulatory proteins, an ATP-dependent DNA helicase protein (PcrA), and ribonuclease H III (**Table A4.2**, **Figure 4.2**), in addition to other proteins like DNA polymerase III and LSU ribosome. Unclassified hypothetical proteins (FIG039061) related to heme utilization was a highly ranked SEED subsystem (LDA=3.33, P=0.043) and likely was due to a large abundance of a modular heme utilizing protein (NITMOv2_0147), with other iron-based biomarkers including Cytochrome C553 and an Fe-S cluster regulator (IscR). Other differentially abundant features were related to carbon cycling SEED subsystems (LDA=2.99, P=0.021) and included ribulose phosphate-3 epimerase used for carbon fixation, and two glycogen synthesis enzymes, 4-alpha-glucanotransferase (malQ), and1,4-alpha glucan (glycogen) branching enzyme (GH-13 type) (glgB).

Contrastingly, the *N*. *moscoviensis* protein functions without mussels were largely marked by biomarkers of stress response. The glutathione-regulated K^+ efflux system used in potassium metabolic processes (LDA= 3.28, P=0.043) were the most definitive SEED subsystems of the no-mussel *N*. *moscoviensis* genome. Furthermore, the glutathione-regulated K^+ efflux system protein family (KefB) is activated in the presence of methylglyoxal [256], and the metabolism of methylglyoxal was also a biomarker of the no-mussel treatment (LDA=2.45, P=0.043). These results may suggest a stress response biomarker, as glutathione-regulated potassium efflux systems are often utilized to counteract electrophilic compounds during stress [257, 258]. Additionally, superoxide

dismutase was another highly ranked functional biomarker and suggests an enhanced stress from reactive oxygen. Other stress biomarkers included the carbon starvation stress SEED subsystem (LDA=2.85, P=0.021) and an enhanced abundance of carbon starvation protein A (cstA) (**Table A4.2**). Other biomarkers suggest a metabolic ability to respond to diverse carbon compounds, including gluconolactonase of the Entner Duodoroff pathway (LDA=2.99, P=0.021), acetoacetate metabolism with an enhanced regulatory protein, AtoC, and an acetyl-CoA biosynthesis enzyme, pyruvate decarboxylase (**Table A4.2**). Differential features also included metabolism of complex carbon sources, such as lactose (LDA=2.40, P=0.021) via 2-oxoglutarate decarboxylase, mannose by way of mannose-1-P guanylyltransferase, maltose and maltodextrin degradation via alpha amylase, and N-acetylglucosamine (LDA=2.96, P=0.043) with beta-hexosaminidase enzymes.

N metabolic genes were lowly abundant for *N. moscoviensis* genomes in both treatments, but a NO_3^-/NO_2^- sensor protein was a biomarker of the mussel habitat (LDA=2.61, P=0.014) and suggests an enhanced ability to respond to NO_2^- and NO_3^- in the environment. In comparison, the no-mussel treatment had an enhance genetic ability to uptake and store N by way of a cyanate ABC transporter (LDA=2.40, P=0.043) and asparagine synthetase (asnB) (**Table A4.2**), respectively.

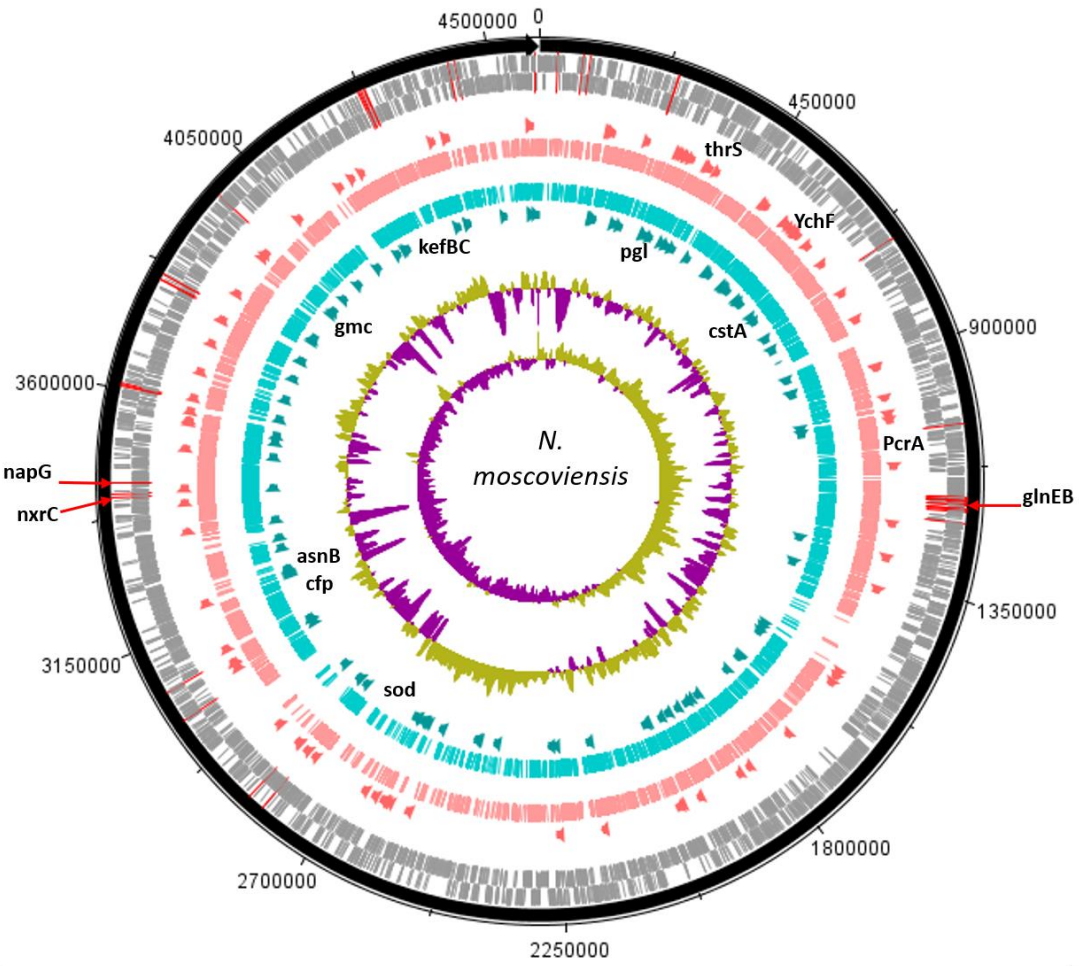


Figure 4.2 The assembled *N. moscoviensis* genome is depicted with tick mark intervals of 225 kBp, and tracks are composed of the following components, starting with the outermost rings: Forward strand coding regions (gray), reverse strand coding regions (gray), the top 25% most abundant enzymes with mussels (dark orange), genomic read coverage from the mussel habitat (salmon colored), genomic coverage without mussels (turquoise colored), the top 25% most abundant enzymes with no mussels (dark green), GC coverage, and GC skew. The N-cycling genes with greatest abundance, GlnEB (NITMOv2_1289, NITMOv2_1290), nitrite oxidoreductase C (NITMOv2_3624), and periplasmic nitrate reductase (NITMOv2_3626) are marked with red lines over the outer rings. Gene names are shown for those with the largest LDA effect size (LDA>3.0). For the mussel habitat, these include the YebC-like proteins, YchF and thrS, heme utilization protein, and DNA helicase (PcrA). Without mussels, these include the CPA2 protein families (kefBC in N. Nitrospira), 6-phosphogluconolactonase (pgl), a Fe-Mn superoxide dismutase (sod), Asparagine synthetase (asnB), glucose methanol choline oxidoreductase (gmc), pyruvate decarboxylase (cfp), and carbon starvation protein (cstA).

4.4.4 Comammox *Nitrospira* genomic potential

A total of 163,253 SEED functionalities were assigned to the genome of *Candidatus* Nitrospira inopinata (normalized to 13,278 per sample) and the top 25% most prominent SEED subsystems did not differ substantially between treatments (**Table A4.3**, **Figure 4.3**). The "restriction-modification system" represented the largest functional category for the no-mussel treatment, while the thiamin metabolism ("5-FCL-like protein"), and stress response ("commensurate regulon activation") SEED assignments were the largest for the genome with mussels. The metabolic function of "urea decomposition" was more abundant without mussels, while "biogenesis of c-type cytochromes" was greater in the mussel treatment. The "ammonia monooxygenase" function was the 4th most abundant SEED subsystem for both mussel and no-mussel treatments. Both treatments shared numerous abundant SEED enzyme assignments (**Table A4.3**), including a resistance-nodulation-cell division (RND) efflux system inner membrane transporter, amoC, alcohol dehydrogenase, and glycogen utilizing enzymes (**Table A4.3**). Numerous N-cycling functions of *Ca*. Nitrospira inopinata were highly abundant in the mussel habitat (**Table 4.2**), including NO- and N_2O-forming enzymes, Amo, and NO_2^-/NO_3^- transforming enzymes (Nxr, NapG, and Nas). Although the total abundance of Amo was slightly larger in the mussel treatment, this trend was most evident for amoA and amoB. Urea transporters (UrtB and UrtC) and urease proteins (UreD, UreG, UreA, and UreF) were relatively more abundant in the mussel treatment, though the "urea decomposition" subsystem did not follow this overall trend.

Finally, N-cycling genes shared by *N. moscoviensis* and *Ca*. Nitrospira inopinata were assessed for differential abundance to reveal if N-cycling genes could explain how

both species were distinct to the mussel habitat, despite competing for similar N substrates. The protein functions found in both species include NH_4^+ transporters and permeases, copper-containing nitrite reductase (nirK), a "NnrS protein involved in response to NO", NO reductase, periplasmic NO_3^- reductase, NO_2^-/NO_3^- sensor and response regulator proteins, nxr, urea transporters, and urease enzymes. A mean rank multiple comparison analysis (Kruskal-Wallis test with Dunn's multiple test correction) revealed that the sum of urea transporters (Urt) (Mean rank difference=53.8, P=0.002) and copper-containing nitrite reductases (NO-forming) (nirK) (Mean rank difference=54.0, P=0.002) were more abundant from the *Ca*. Nitrospira inopinata genome (mussel habitat) compared to *N. moscoviensis.*

Table 4.2 *Ca*. Nitrospira inopinata N-cycling protein functions from the mussel habitat as relative abundance (RPKM) and as a proportion of SEED enzymatic function. SEED counts are shown as a percentage of the total classified protein function in the mussel habitat.

SEED subsystems and protein functions	Average read abundance (RPKM)	Relative proportion of protein function
Ammonia monooxygenase		
Ammonia monooxygenase A-subunit amoA	2955.70	3.79%
Ammonia monooxygenase B-subunit amoB	1405.72	
Ammonia monooxygenase C-subunit amoC	37945.15	
Urea decomposition		
Urea ABC transporter, ATPase protein UrtD	1634.33	3.52%
Urea ABC transporter, ATPase protein UrtE	1459.22	
Urea ABC transporter, permease protein UrtB	1713.41	
Urea ABC transporter, permease protein UrtC	735.02	
Urea ABC transporter, urea binding protein urtA	1443.89	
Urea carboxylase-related amino acid permease UctT	201.53	
Urease accessory protein UreD	ND	
Urease accessory protein UreF	151.66	
Urease accessory protein UreG	303.31	
Urease alpha subunit UreC	724.79	
Urease beta subunit UreB	221.21	
Urease gamma subunit UreA	ND	
Denitrification		
Copper-containing nitrite reductase nirK	10156.73	0.50%
Nitric oxide reductase protein NorQ	127.68	
Nitrate and nitrite ammonification		
Hydroxylamine dehydrogenase (hao)	3472.95	0.28%
Putative haoB	107.52	
Nitrite/Nitrate oxidoreductase, alpha subunit (nxrA)	ND	
Nitrite/Nitrate oxidoreductase, beta subunit (nxrB)	ND	
putative Nitrite/Nitrate oxidoreductase, membrane subunit (nxrC)	522.64	
Ammonia assimilation		
Ammonium transporter Rh50	87.70	0.17%
Glutamate-ammonia-ligase (glnEb)	271.58	

4.4.5 Comammox biomarkers

The most discerning differences in the genome of *Candidatus* Nitrospira inopinata with mussels included a gene encoding the RND efflux system inner membrane transporter (acrB) (**Table A4.4**, **Figure 4.3**) as part of the subclass "FOL commensurate regulon activation" (LDA=3.71, P=0.043). Numerous other stress response and virulence pathways were enhanced in the mussel treatment (**Table A4.4**), such as the RND efflux system membrane fusion protein (TtgA), a membrane protease with a mechanism for aminoglycoside resistance (HflK), an integral inner membrane protein type IV secretion complex (virB), and "death on curing protein" (doc) as part of the biomarker pathway of YdcE and YdcD toxin programmed cell death (LDA=2.46, P=0.014). Results also indicated a larger metal tolerance in the mussel treatment, namely copper tolerance (LDA=2.26, P=0.047) using a periplasmic divalent cation tolerance protein (cutA).

Other biomarkers in the mussel habitat had indirect links to N metabolic through urea cycling and cytochrome biosynthesis. A NiFe hydrogenase (HypF) assembly protein which regulates the sulfur-reducing hydrogenase gene set (hydBGDA and hybD) was increased, and was recently suggested as giving comammox *Nitrospira* the potential function of hydrogen oxidation coupled to sulfur reduction in anaerobic conditions [239]. Cytochrome c-type biogenesis protein (CcmF) was increased with mussels, and is a heme chaperone required for biogenesis of cytochrome c-type proteins located immediately downstream of hydroxylamine dehydrogenase (hao) (**Figure 4.3**). Another biomarker with mussels, carbamoyl phosphate synthetase (carB), would effectively add urea-derived NH_3 into central metabolism.

SEED subsystems of *Candidatus* Nitrospira inopinata most indicative of the no-mussel treatment were potassium homeostasis (LDA=3.53, P=0.043), alpha carboxysome (LDA=3.04, P=0.043), and DNA metabolic CRISPR function (LDA=3.04, P=0.043). The only significantly different enzyme in the "potassium homeostasis" subsystem was a potassium transporting ATPase (KdpC), and is a catalytic chaperone for high-affinity potassium uptake [259]. Greater abundance of rubrerythrin, a stress response protein used to combat oxidative stress (LDA=2.45, P=0.021) [260], was greater in the no-mussel treatment, and may be linked to greater abundances of recombination and repair protein (RecO) and excinuclease ABC genes in DNA repair pathways (LDA=2.38, P=0.021). N-cycling genes were not found to be a significant feature differentiating comammox *Nitrospira* genomes in the no-mussel treatment.

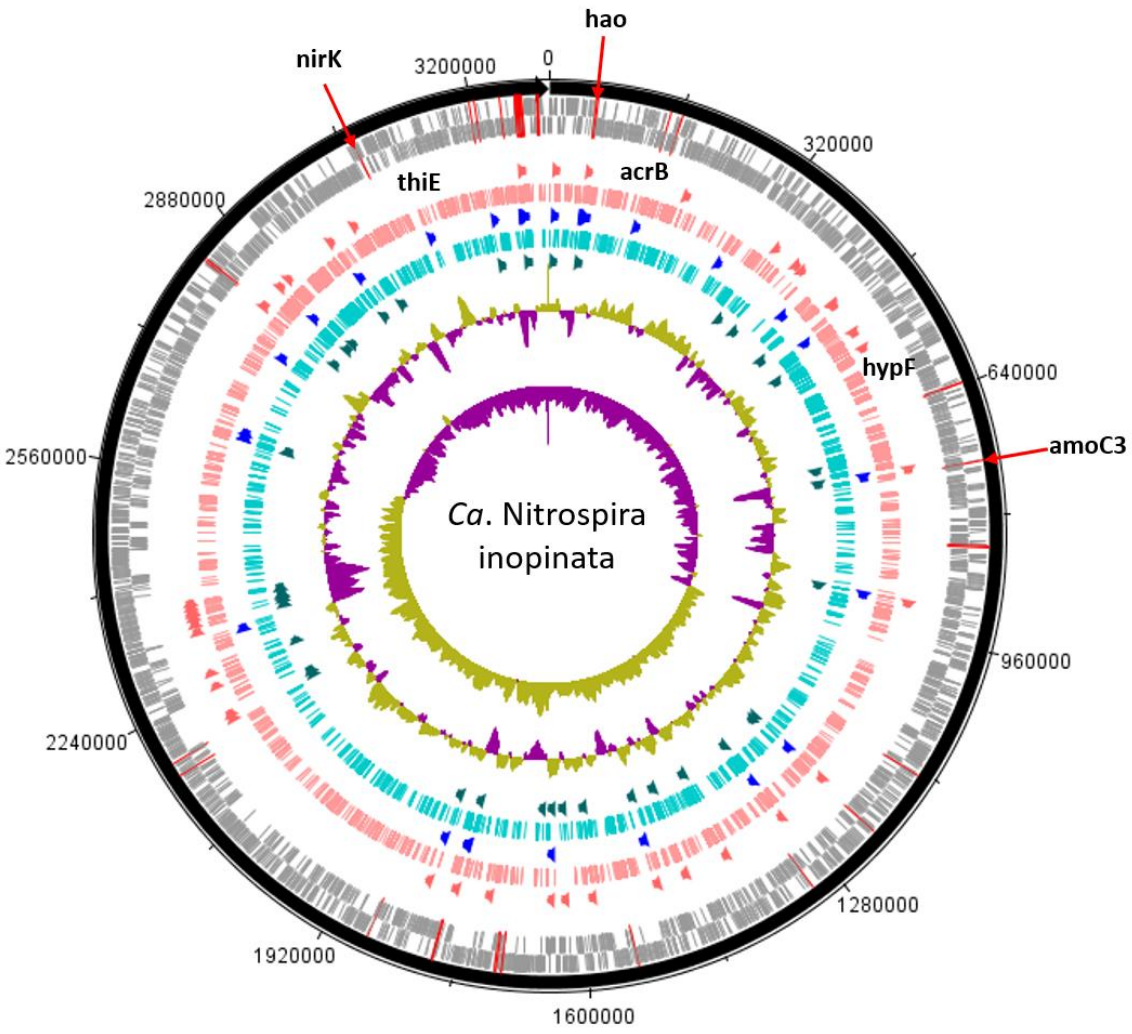


Figure 4.3 *Ca*. Nitrospira inopinata assembly shown with tick mark intervals of 160 kBp. Tracks within the genome are composed of the following, starting with the outermost rings: Forward strand coding regions, reverse strand coding regions, the top 25% most abundant enzymes with mussels (dark orange), genomic coverage with mussels (salmon colored), top 25% most abundant SEED subsystems in both treatments (royal blue), genomic coverage without mussels (turquoise colored), the top 25% most abundant enzymes with no mussels (dark green), GC coverage, and GC skew. N-cycling genes are designated with red lines across the CDS rings and the most abundant genes (amoC3 (NITINOP_0766), nirK (NITINOP_3146), hao (NITINOP_0065) are designated with text. For simplicity, genes with LDA>3 are named (acrB, hypF, and thiE), and therefore limits to the mussel habitat treatment (text shown above salmon-colored ring).

4.5 Discussion

In support of our research goal, we provided metagenomic evidence of niche partitioning features to clarify how *Nitrospira* were greater in UMR sediment with mussels, and agrees with previous research [203]. We found that *Nitrospira moscoviensis* and *Candidatus* Nitrospira inopinata were the most differentiating N-cycling taxons in this mussel habitat niche, and these organisms likely co-existed because of metabolic flexibility beyond the conventional N-cycle.

4.5.1 NOB Nitrospira were most impacted by carbon cycling

The most definitive biomarker of the *N. moscoviensis* genome recovered from the mussel habitat was YebC (LDA=3.49) and suggests NOB-like *Nitrospira* were capable of enhanced genetic diversity. For example, the YebC protein regulates genetic recombination and resolution of Holliday Junctions [261], and enzymes classified within the YebC subsystem, a GTP-binding and nucleic acid-binding protein (YchF) and Threonyl-tRNA synthetase (ThrRS), function as translational control factors [262, 263]. These results suggest that NOB-like *Nitrospira* of the mussel habitat had the potential to synthesize proteins and respond to their environment, perhaps in response to nutrient supply [264]. NOB *Nitrospira* have high substrate affinities for N and O_2, and thus are well suited as scavengers in low-NO_2 and O_2 environments [155, 265]. Additionally, it makes sense that enhanced DNA repair using “uvrD-related helicases” and “two cell division clusters relating to chromosome partitioning” were also top differentiating features in the mussel habitat. DNA repair and recombination biomarkers suggest that *N. moscoviensis* may have dominated the mussel habitat sediment by way of genetic

diversity. To this point, flexible metabolism and a mixotrophic lifestyle has enabled *N. moscoviensis* to be ecologically successful in many niches [33].

We demonstrated that C metabolic processes of *N. moscoviensis* were impacted in the mussel habitat, as evidenced by Calvin cycle and glycogen metabolism biomarkers while the no-mussel treatments contained C stress protein biomarkers. C starvation protein A (cstA) is a membrane protein predicted to be involved in peptide uptake when C availability is low [266]. Experimental evidence has also shown the *cstA* gene is upregulated during C starvation [267], and allows for increased cellular growth by importing peptides as sources of C and N [268]. It is possible that NOB-like *Nitrospira* were C-stressed without mussels and is further supported by the glutathione-regulated potassium efflux gene (*kefC*) and methylglyoxal metabolism biomarkers. Evidence suggests that glutathione activates KefC potassium channels to protect against the toxic effects of methylglyoxal, and often occurs because of limited phosphate or excessive C concentrations [269, 270]. Previous studies have shown that mussel biodeposits contain fairly consistent C:N ratios near 11 [271], so further research should address if the mussel habitat contained a more consistent C:N ratio and led to reduced C-stress biomarkers in *N. moscoviensis*.

Furthermore, it is possible that NOB *Nitrospira* had an enhanced niche through diverse C sources in the mussel habitat. Formate hydrogenase enzymes were a highly abundant N-cycling SEED category, while the gene families nirK, ure and urt, represented the lowest N-cycling genetic potential. These results suggest that the NOB *Nitrospira* had greater genetic potential to degrade to formate than by reciprocal feeding of mussel-derived urea. Taken together, these results signify *N. moscoviensis* in the

mussel habitat had greatest potential for carbon degradation, NO_2^- oxidation, and likely thrived from fermentation products common in an oxic-anoxic niche.

4.5.2 N impacted the genomic potential of comammox *Nitrospira*

Our results from the mussel habitat showed that *Ca*. Nitrospira inopinata had the largest genomic potentials for NH_3 oxidation, urea decomposition, and NO_2^- redox reactions. In comparison to *N. moscoviensis*, *Ca*. Nitrospira inopinata was genetically equipped to obtain electrons from multiple N compounds. Furthermore, since the comammox genome contained large abundances of N-cycling genes compared to *N. moscoviensis*, this indicates that these two organisms were successful in the mussel habitat for different metabolic functions. In the mussel habitat, *Ca*. Nitrospira inopinata had an order of magnitude greater capacity for urea decomposition than *Nitrospira moscoviensis*, with the most evident differences shown for urease proteins and urea transporters. These results suggest that commamox were likely degrading urea and oxidizing NH_3, and supports previous findings that freshwater mussels indirectly increase concentrations of NO_2^- and NO_3^- when housed in high nutrient waters [13], similar to conditions found in the UMR.

The comammox genome was marked with multidrug resistance efflux pumps (cmeB/acrB), and metal resistance (Zn and Cu) in the mussel habitat and suggests there may have been a selective pressure to use defense mechanisms. These multidrug resistance genes are often a response to one or more substrates and enables resistance towards numerous antimicrobial substrates including antibiotics and heavy metals [272, 273]. Furthermore, our detected biomarkers for increased Phd-Doc toxin-antitoxin genes, are attributed with biochemical processes including antibiotic resistance [274]. One

explanation for these findings may be related to mussels hosting antibiotic resistant symbionts [275, 276], multidrug-resistant pathogens, [277], and enhanced horizontal gene transfer within their gut microbial community [278]. Previous research has also shown that mussels bioaccumulate various metals [279, 280], including elements found in this biomarker study, copper and zinc [281]. It is possible that horizontal gene transfer of multidrug resistance was facilitated by symbionts in mussel biodeposits, or from the decay of mussel tissue after death. However, this study was not designed to pinpoint antimicrobial stressors in the mussel habitat, so we cannot definitively say that commamox *Nitrospira* obtained these features as a direct result of their niche.

The biomarker analysis also revealed that phosphorus metabolism was a distinctive feature of comammox *Nitrospira* in the mussel habitat. This may be influenced by mussel excretion products regenerating phosphate in ecosystems, but ultimately is dependent on the suspended food available to mussels [271]. Regeneration of phosphate in mussel habitat sediments would give comammox *Nitrospira* a selective advantage over AOA and AOB which do not encode an alkaline phosphatase [282]. Another biomarker potentially explaining the success of comammox *Nitrospira* was the NiFe hydrogenase maturation protein (hypF) found in the mussel habitat. Recent studies have reported that comammox have the potential to oxidize H_2 and use sulfur as an electron acceptor in anaerobic conditions [239], and this further emphasizes the importance of metabolic versatility in the UMR mussel habitat.

4.5.3 NOB and comammox *Nitrospira* coexist in a mussel habitat niche

Several genomic properties enable *Ca*. Nitrospira inopinata to outcompete other microorganisms, including high affinities for NH_3 transport and oxidation, urea transport,

and an alkaline phosphatase enabling comammox *Nitrospira* to compete in low P environments [32, 282, 283]. The importance of scavenging urea and NH_3 is emphasized by the fact that *Ca*. Nitrospira inopinata cannot survive on NO_2^- alone, since the organism lacks the ability to utilize NO_2^- as a source of N [199]. Although comammox *Nitrospira* would have the advantage over NOB *Nitrospira* by scavenging NH_3 and urea, both organisms could occupy the same niche due to their metabolic flexibility. NOB *Nitrospira* thrive in oxic-anoxic niches where formate is produced by fermentation, and would not be outcompeted by comammox *Nitrospira* clade A, which lack a formate dehydrogenase enzyme [283]. Furthermore, *N. moscovinesis* can simultaneously produce and consume NO_2^-, by coupling formate oxidation to NO_3^- reduction while aerobically oxidizing NO_2^- [33]. This may explain our findings that NOB *Nitrospira* had high genomic potentials for NO_2^-oxidation (nxr) NO_3^- reduction (napG), and formate hydrogenation (hyf, fds).

Surprisingly, we detected an order of magnitude greater *nirK* compared to *nxr* in *Ca*. Nitrospira inopinata and signifies that NO would be a major product of nitrification. Although previous studies have not documented gaseous NO_x production from *Nitrospira* nirK, comammox organisms do have the genetic capability for NO_3^- reduction to NO_2^- and NO [239]. Furthermore, *N. moscoviensis* showed genomic potential to respond to and degrade NO and N_2O, potentially removing the major gaseous products of nitrification. These results suggest that NOB- and comammox-*Nitrospira* were able to thrive in the mussel habitat because comammox were utilizing urea, oxidizing ammonia, and performing NO_2^- redox reactions, while NOB scavenged carbon sources and used N compounds as a source of electrons.

4.6 Conclusion

This research used genomic biomarkers to show that *Nitrospira moscoviensis* and *Candidatus Nitrospira* inopinata predominated an N-cycling oxic-anoxic niche in sediment of a mussel habitat. Our research showed that formate oxidation coupled to NO_3^- reduction by NOB *Nitrospira* may have enabled co-existence with comammox *Nitrospira* in this sediment niche. The mussel habitat harbored comammox *Nitrospira* with enhanced RND efflux transporters and metal resistance, phosphorus metabolism, and showed evidence of hydrogen oxidation, while decreasing the genomic potential for potassium homeostasis and oxidative stress. For *N. moscoviensis*, the mussel habitat contained greater abundances of translational control genes and heme utilization while the no-mussel treatment showed genomic evidence of carbon stress. Both NOB and comammox *Nitrospira* were marked by diverse metabolism in the mussel habitat and may have contributed towards the increased abundance of both organisms. More research is needed to determine the biogeochemical signatures of the mussel habitat that may be responsible for these various biomarkers. Ultimately, this study provided metagenomic evidence showing that niche partitioning and mixotrophic metabolism allowed NOB and commamox *Nitrospira* to coexist in mussel habitat sediment.

CHAPTER 5: CONCLUSIONS

5.1 Engineering applications

5.1.1 Modeling mussel influences on environmental microbiology

Modeling the fate of aquatic N in an agroecosystem with mussels and the subsequent transformations by N-cycling bacteria is needed to determine the large-scale impact of restoring mussel habitats on N sequestration. This computational model reinforced lab mesocosm experiments showing that mussels increase porewater NH_4^+ concentrations [13, 56]. These experiments also showed an increase in NO_2^- and NO_3^- in water overlying mussels [13, 56], suggesting that microorganisms were oxidizing the NH_4^+ added by mussels. To truly represent microbial activity, the updated model included microbial N transformations and heterotrophic consumption of oxygen.

Metagenomic sequencing of UMR mussel bed sediment revealed the presence of abundant N-transforming microorganisms [203]. Numerous nitrifiers (NH_4^+ and NO_2^- oxidizers) were detected, including those of the genera *Nitrospira*, *Nitrosocaldus*, *Nitrososphaera*, *Nitrococcus*, *Nitrobacter*, *Nitrosococcus*, and *Nitrosomonas*. Denitrifying organisms (NO_3^- and NO_2^- reducers) in the mussel sediment included species *denitrificans* and *nitritireducens* from numerous genera. Additionally, anaerobic ammonium oxidizing (anammox) microorganisms of the candidate genus *Brocadia* were present and suggested an oxic-anoxic interface niche was present in the sediment.

5.1.2 Methods

The existing model (STELLA version 8.0, ISEE Systems, Inc., Lebanon, New Hampshire) from Bril et al. (2016) simulated the fate of N from mussel excretion events in 30-minute increments for a total of 2160 hours [59]. Overall, the model simulated the transformation of N into phytoplankton, consumed as a food source for mussels, and was further excreted into sediment. The model also predicted inorganic N concentrations in porewater and overlying water from nitrification and denitrification activity. The performance of this updated model with sediment microbial communities tested the N concentrations in sediment compartments and overlying water in response to 1) N runoff and 2) mussel density.

It was assumed that oxygen transfer was not limiting in the oxic sediment zone, and that denitrification and anammox reactions did not occur in oxic sediment due to inhibition. Oxygen was not present in the anoxic zone, and nitrification was not included as a result. NH_4^+ was oxidized by nitrifiers and anammox in the oxic-anoxic transition zone and was only oxidized by anammox in anoxic sediment. Inorganic N diffusion rates within sediment were kept constant at 0.002. One limitation to this work was the assumption that diffusion dominated the transfer of N and ignored mixing from mussel bioturbation.

For simulations, the baseline inputs were 50 g mussel biomass, and river water concentrations of 0.1 mg-N/L NO_3^- and 1.14 mg-N/L NH_4^+, similar to pre-development UMR conditions [60]. Additional simulations were run to represent UMR conditions since the 1990's, including an increase in NO_3^- to 10 mg/L and reducing the mussel biomass by $1/10^{th}$ [60].

5.1.3 Results and discussion

Pre-industrial revolution conditions showed mussel excretion products (i.e. NH_4^+) being oxidized to NO_2^- by nitrifiers in oxic sediment and within in the oxic-anoxic transition zone. NO_2^- accumulated to almost 0.05 mg/L in anoxic sediment, suggesting that it was being consumed by anammox and denitrifiers at a slower rate than the diffusion rate from oxic sediment. Simulations showed a cyclical and inverse relationship between NH_4^+ and NO_2^- concentrations (not depicted) but both N species maintained concentrations below 0.05 mg/L-N. On the other hand, NO_3^- accumulated in the oxic-anoxic transition zone and suggests that NH_4^+ and NO_2^- oxidation outpaced diffusion of NO_3^- into anoxic sediment and/or denitrification was a limiting reaction. Interestingly, pre-development NO_3^- river water concentrations stabilized near 5 mg/L despite reaching lower equilibrium concentrations in each sediment component (**Figure 5.1**). Further investigation is needed to determine which factors limit the removal of aquatic N.

The next set of simulations aimed to replicate current UMR NO_3^- concentrations (near 10 mg/L). NO_3^- accumulated in oxic sediment and reached a lower concentration at equilibrium when compared to pre-development (**Figure 5.2**). Similar to pre-industrial conditions, there was a slight accumulation of NO_3^- in the oxic-anoxic transition zone but remained well below 1 mg/L in anoxic sediment (**Figure 5.2**). The most unexpected finding from this simulation was that NO_3^- decreased in abundance until reaching steady state near 2 mg/L (**Figure 5.2**) while pre-industrial NO_3^- accumulated to a concentration near 5 mg/L (**Figure 5.1**). Since mussel biomass was held constant in this portion of the simulation, it suggests that N removal was limited by primary productivity in pre-industrial conditions and not mussel filtration capacity.

Lastly, a 1/10th decrease in mussel biomass resulted in similar N concentrations in oxic porewater and anoxic porewater, but NO_3^- concentrations differed for overlying water and the oxic-anoxic transition sediment. Although NO_3^- concentrations in water decreased to about 2 mg/L, there were significant oscillations of ± 1 mg/L per 1 to 2-hours of simulation time (**Figure 5.3**). This may be due to a lag between phytoplankton production and mussel feeding events, but further investigation is needed to confirm this finding. NO_3^- concentrations also increased for the 2160-hour simulation in the oxic-anoxic transition sediment zone but remained below 1 mg/L in anoxic sediment (**Figure 5.3**). This suggests that the diffusion of nitrate between sediment compartments limited the transfer to deeper and anoxic zones, and likely limiting denitrification rates. However, it is important to note that this study was limited to diffusion and did not include mechanical sediment mixing from mussel bioturbation.

Future work should include adding a microbial biomass doubling time factor, since N-cycling microorganisms have a doubling time significantly less than the simulation time (90 days). Additionally, this model could more accurately represent microbial ecology by adding stocks and flows of organic carbon. This would enable a better representation of competition between anammox and denitrifying organisms, as a C/N ratio of 3 enables denitrifiers to be more competitive. Additionally, organic carbon has an oxygen demand associated with it, which would also change the dynamics of oxygen transfer between sediment components. For example, if carbonaceous oxygen demand outpaced that of nitrifiers, there would be an accumulation of NH_4^+ instead of being further transformed via denitrification and anammox. Lastly, once these

components are fine-tuned, the model should be scaled up to a larger UMR basin to predict results at an environmentally relevant scale.

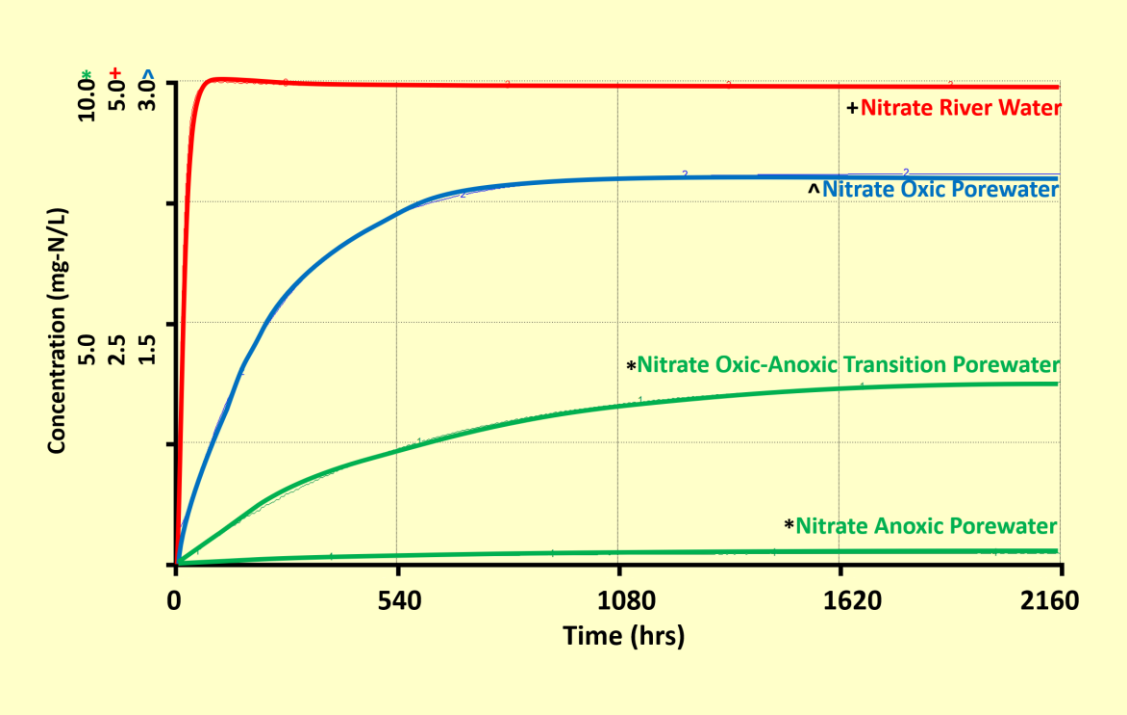


Figure 5.1 Simulation of pre-industrial revolution values for mussels and nitrate. Symbols next to line labels correspond to numerical values for the y-axis.

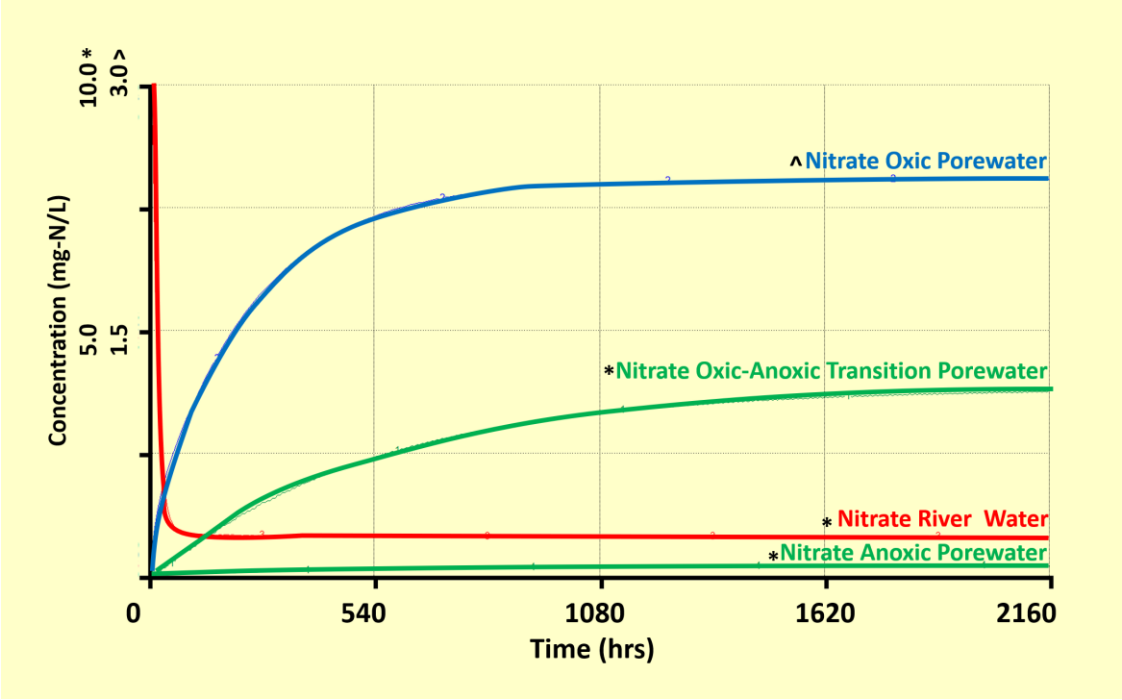


Figure 5.2 Simulation of initial nitrate concentrations of 10 mg/L and original mussel biomass. Symbols next to line labels correspond to numerical values for the y-axis.

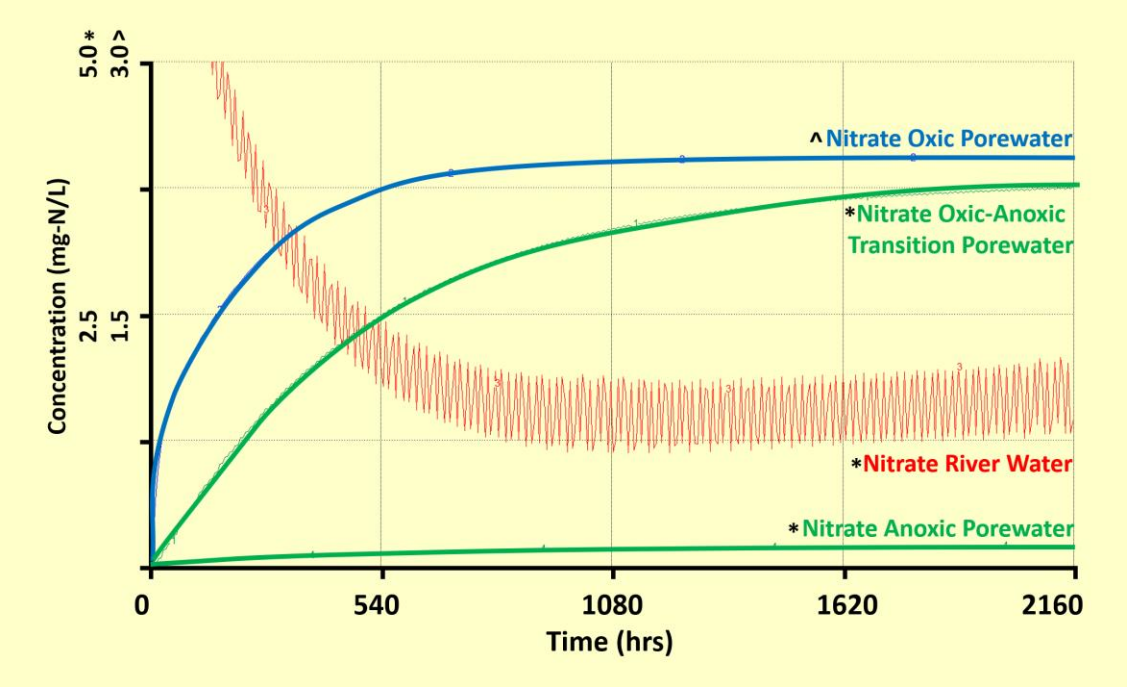


Figure 5.3 Simulation of post-industrial revolution values of $1/10^{th}$ reduction in mussel biomass and 10 mg/L nitrate. Symbols next to line labels correspond to numerical values for the y-axis.

5.1.4 Implications

This research can be used in science and engineering applications by informing models used to predict if aquatic N can be remediated by restoring native mussel habitats. Real world applications of this model include determining how mussels and sediment microbes respond to varying concentrations of N from agricultural runoff and how N concentrations vary with mussel density. This model reinforced lab mesocosm experiments showing that mussels increase porewater NH_4^+ concentrations [13, 56]. These experiments also showed an increase in NO_2^- and NO_3^- in water overlying mussels [13, 56], suggesting that microorganisms were oxidizing NH_4^+ added by mussels.

Overall, this investigation confirmed the intricate interaction of N transformations between mussels and sediment microbes and adding stocks and flows of microbiological activity specific to an oxygen gradient is necessary for an accurate N prediction tool. Model simulations showed river water reached an equilibrium nitrate concentration of 5 mg/L-N for pre-industrial conditions (**Figure 5.1**). Secondly, the pre-industrial mussel population adequately consumed phytoplankton and decreased river nitrate concentrations from 10 mg/L-N to about 1 mg/L-N (**Figure 5.2**). However, the post-industrial values of nitrate (10 mg/L-N) and a $1/10^{th}$ reduction in mussel biomass resulted in less stable, yet reduced nitrate concentrations (**Figure 5.3**).

Comparing pre- and post-industrial revolution simulations showed that increasing NO_3^- concentrations of river water to 10 mg/L led to lower a lower NO_3^- concentration at equilibrium. These results do not suggest that we continue with "business as usual" for N runoff but does demonstrate that mussels can sequester N into sediment by modeling biogeochemical processes. Future improvements should include rates of mechanical

mixing from mussel bioturbation, a microbial biomass doubling time, stocks and flows of organic carbon and carbonaceous oxygen demand. Lastly, this model should be scaled up to represent the UMR basin to simulate the fate and transport of N at an environmentally relevant scale. Ultimately, a computational model of N biogeochemistry would be useful as a policy making tool to incentivize mussel bed restoration and to determine the fate of N pollution in the UMR.

5.2 Future research proposal

Future work should determine *how* and *why* freshwater mussels impact microbial communities (see **Figure 5.4**). This research would elucidate the biogeochemical interactions between mussels and microorganisms via an "interactome" approach by combining metabolomics, metagenomics, and transcriptomics.

The goals are:

1) Characterize the metabolic products contained in mussel hemolymph, biodeposits, and tissues.
2) Characterize the activity of the sediment microbial community and the microbes living on mussels by sequencing the respective transcriptomes

These research goals are informed by results presented in Chapters 3 and 4, particularly about finding increased abundances of sulfur-cycling endosymbionts and large abundances of antibiotic resistance genes. These results would indicate if freshwater mussels host an environmental hot-spot of antibiotic resistant genes and would also provide evidence of microbial activity needed to advocate for habitat restoration.

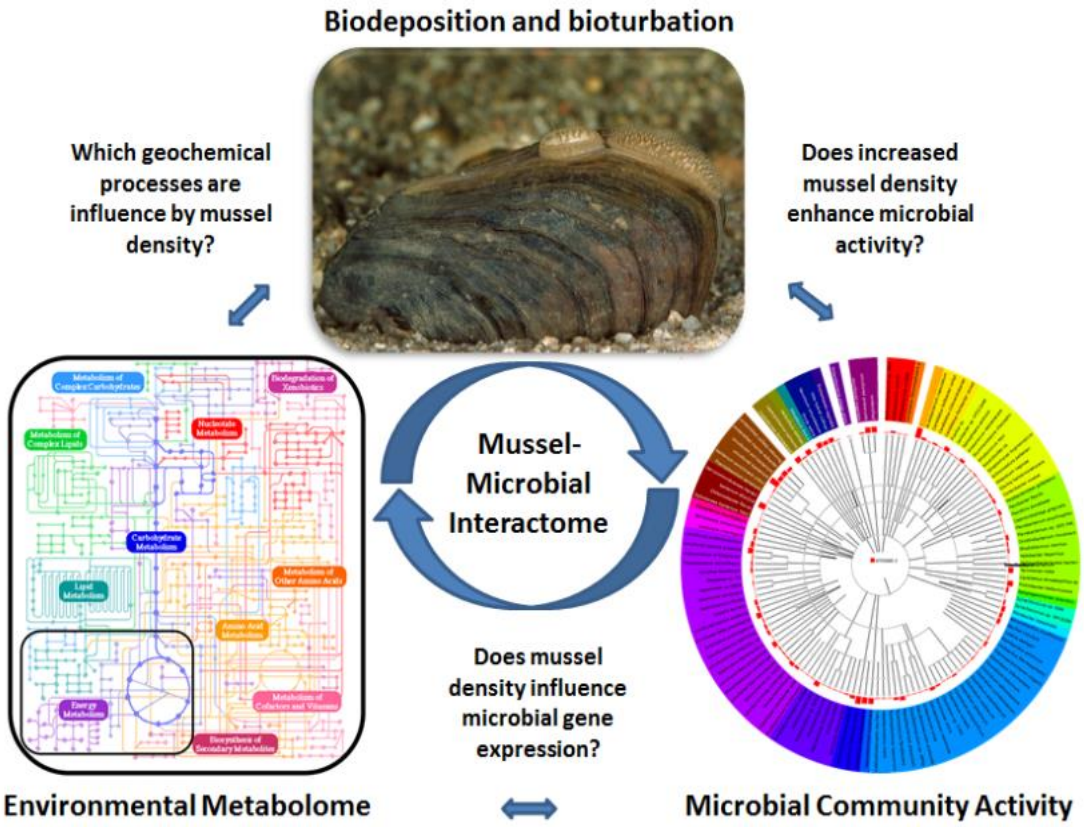


Figure 5.4 Proposed research to combine metabolomics, transcriptomics, and genomics of sediment microbial communities and native freshwater mussels.

5.3 Final conclusions

This research utilized high-throughput sequencing of sediment microbiomes to identify how native freshwater mussels impact microbial community structure and function. As far as we know, this is the first study to characterize microbial communities in UMR sediments with mussels. First, we performed qPCR of the anammox-specific 16S rRNA gene and revealed an increase in anammox bacteria abundance at sediment depths relevant to native mussel burrowing. 16S rRNA gene amplicon sequencing showed that sediment impacted by mussels contained a distinct microbial community and harbored a niche for microorganisms which thrive in oxic-anoxic interface niches. This sequencing approach also revealed that mussels increased the abundance of ammonia oxidizing bacteria classified as *Nitrosomonadaceae* and nitrite oxidizing bacteria (genus *Nitrospira*), while mussels decreased the abundance of ammonia oxidizing archaea classified in the genus *Candidatus* Nitrososphaera, and microorganisms which couple denitrification with methane oxidation.

Secondly, metagenomic shotgun sequencing revealed a large genomic potential for nitrate metabolism in UMR sediments and mussel presence resulted in significantly increased nitrification potential at the expense of genes responsible for nitrous oxide production. In support of these findings, *amoA* and *nxr* genes were the most predominant biomarkers of the mussel sediment and likely originated from comammox *Nitrospira* and NOB *Nitrospira*.

The third experiment used genomic biomarkers to show that *Nitrospira moscoviensis* and *Candidatus* Nitrospira inopinata were most responsible for differences in N-cycling microbial communities from mussels. This research showed that formate

oxidation coupled with NO_3^- reduction by NOB *Nitrospira* may have enabled co-existence with comammox *Nitrospira* in this sediment niche.

The experimental approaches used in chapters 3 and 4 differed in bioinformatics approach but obtained converging lines of evidence. Chapter 3 used a gene-centric approach to define the N-cycling metabolic potential of the entire microbial community while chapter 4 used a genome-centered approach to quantify species-specific metabolic pathways. Chapter 3 results indicated nitrification genes were most responsible for distinct microbial communities with mussels and the nitrification genes most increased also belonged to *Nitrospira* species. Similarly, chapter 4 resulted in the conclusion that genomes belonging to NOB and comammox *Nitrospira* were the most differentiating taxons and confirmed the species could coexist in the same niche by metabolic flexibility In combination with computational modeling, this research converges with previous evidence that native freshwater mussels enhance N-cycling of benthic organisms, and microbial activity is an explanation for the observed NO_3^- increases with mussels. Ultimately, this research is a first step in advocating for native mussel habitat restoration and adds "microbial N-cycling" to the list of ecosystem services performed by mussels.

APPENDICES

Supplementary material for Chapter 3

Table A3.1 Sequence accessions for each sample found at MG-RAST, ENA, and NCBI. MG-RAST projects detail sequence statistics from QA/QC and links this shotgun sequencing study with previous amplicon sequencing results.

Sample Name	MG-RAST ID	MG-RAST (16S rRNA amplicon)	European Nucleotide Archive	Sequence Read Archive (NCBI) BioSample Accession
	MG-RAST Project mgp21252	MG-RAST Project mgp18682	Study Accession PRJEB23134	BioProject ID PRJNA414922
No mussel (S1)	mgm4730047.3	mgm4705698.3	ERS1981943	SAMN06710719
No mussel (S2)	mgm4730043.3	mgm4705704.3	ERS1981944	SAMN06710724
No mussel (S3)	mgm4730044.3	mgm4705686.3	ERS1981945	SAMN06710731
No mussel (S4)	mgm4730042.3	mgm4705697.3	ERS1981946	SAMN06710729
Mussel (S5)	mgm4730045.3	mgm4705708.3	ERS1981947	SAMN06710710
Mussel (S6)	mgm4730048.3	mgm4705672.3	ERS1981948	SAMN06710716
Mussel (S7)	mgm4730041.3	mgm4705699.3	ERS1981949	SAMN06710705
Mussel (S8)	mgm4730046.3	mgm4705680.3	ERS1981950	SAMN06710725

Table A3.2 Numerous nitrification and urease protein clusters were differentially abundant with mussels and are listed in order of decreasing effect size. Clusters are labeled with UniRef90 ID and lowest common ancestor taxonomic classification.

N-cycling functional gene	Protein cluster (LCA)	LDA effect size	P-value	Treatment
nxrC	UniRef90_D8PI74 (*Nitrospira*)	3.32	0.018	Mussel
nxrB	UniRef90_D8PI40 (*Nitrospira*)	3.21	0.021	Mussel
nxrA	UniRef90_D8PI59 (*Nitrospira*)	3.16	0.021	Mussel
amoA	UniRef90_Q4FIQ2 (Bacteria)	3.07	0.043	Mussel
amoA	UniRef90_Q4ACX5 (Bacteria)	2.96	0.043	Mussel
nxrB	UniRef90_T2HGM4 (Bacteria)	2.96	0.043	Mussel
ureC	UniRef90_D2CR64 (unclassified)	2.96	0.021	Mussel
amoA	UniRef90_Q4ADF9 (Bacteria)	2.79	0.018	Mussel
nxrA	UniRef90_D8PI41 (*Nitrospira*)	2.75	0.021	Mussel
hao	UniRef90_B9V947 (*Nitrosomonas*)	2.46	0.043	Mussel
ureC	UniRef90_D2CR19 (unclassified)	2.43	0.043	Mussel
ureC	UniRef90_D2CQ18 (unclassified)	2.41	0.038	Mussel
amoA	UniRef90_Q04507 (Bacteria)	2.39	0.047	Mussel
amoA	UniRef90_O85112 (Bacteria)	2.37	0.047	Mussel
amoA	UniRef90_B5B293 (Bacteria)	2.33	0.047	Mussel
amoA	UniRef90_H7CHP0 (Archaea)	2.32	0.047	Mussel
amoA	UniRef90_E9JNT4 (Bacteria)	2.26	0.047	Mussel

Table A3.3 List of IMG bacterial and archaeal completed genomes and respective monooxygenase gene products used for multiple sequence alignments.

Genome Name	Gene Product	Locus Tag
Candidatus Nitrosocosmicus exaquare G61	amoA	Ga0175697_111208
Candidatus Nitrosomarinus catalina SPOT01	amoA	Ga0226285_11260
Candidatus Nitrosopelagicus brevis CN25	amoA	Ga0069311_11304
Candidatus Nitrosopumilus adriaticus NF5	amoA	Ga0077947_111922
Candidatus Nitrosopumilus koreensis AR1	amoA	NKOR_08170
Candidatus Nitrosopumilus piranensis D3C	amoA	Ga0077946_111587
Candidatus Nitrosopumilus piranensis D3C	amoA	Ga0077946_111897
Candidatus Nitrosopumilus sediminis AR2	amoA	NSED_08255
Candidatus Nitrososphaera evergladensis SR1	amoA	NTE_00961
Candidatus Nitrososphaera gargensis Ga9-2	amoA	Ngar_c25350
Candidatus Nitrosotalea devanaterra NDEV1	amoA	Ga0226572_11339
Cenarchaeum symbiosum A	amoA	CENSYa_0402
Nitrosopumilus maritimus SCM1	amoA	Nmar_1500
Nitrososphaera viennensis EN76	amoA	NVIE_027270
Thaumarchaeota archaeon MY3	amoA	Ga0114842_111096
Thaumarchaeota archaeon SAT1	amoA	Ga0077941_11426
Nitrosomonas communis Nm2	amoA	Ga0078022_11129
Nitrosomonas communis Nm2	amoA	Ga0078022_113140
Nitrosomonas europaea ATCC 19718	amoA	NE0944
Nitrosomonas europaea ATCC 19718	amoA	NE2063
Nitrosomonas eutropha C91	amoA	Neut_2077
Nitrosomonas eutropha C91	amoA	Neut_2318
Nitrosomonas sp. AL212	amoA	NAL212_0798
Nitrosomonas sp. AL212	amoA	NAL212_1387
Nitrosomonas sp. AL212	amoA	NAL212_2605
Nitrosomonas sp. IS79A3	amoA	Nit79A3_0472
Nitrosomonas sp. IS79A3	amoA	Nit79A3_1080
Nitrosomonas sp. IS79A3	amoA	Nit79A3_2885
Nitrosospira briensis C-128	amoA	F822DRAFT_0879
Nitrosospira briensis C-128	amoA	F822DRAFT_1681
Nitrosospira briensis C-128	amoA	F822DRAFT_2229
Nitrosospira multiformis ATCC 25196	amoA	Nmul_A0799
Nitrosospira multiformis ATCC 25196	amoA	Nmul_A2325
Nitrosospira multiformis ATCC 25196	amoA	Nmul_A2765
Candidatus Nitrospira inopinata ENR4	amoA	Ga0125266_112381
Nitrosococcus halophilus Nc4	amoA	Nhal_0676
Nitrosococcus oceani C-107	amoA	Noc_2502

Table A3.3 cont.

Genome Name	Gene Product	Locus Tag
Nitrosococcus watsoni C-113	amoA	Nwat_0632
Mycobacterium chubuense NBB4	pmoA	Mycch_5910
Mycobacterium rhodesiae NBB3	pmoA	MycrhN1_3041
Methylocystis sp. SC2	pmoA	BN69_0203
Methylocystis sp. SC2	pmoA	BN69_2827
Methylocystis sp. SC2	pmoA	BN69_3534
Methylococcus capsulatus Bath	pmoA	MCA1797
Methylococcus capsulatus Bath	pmoA	MCA2854
Methylomicrobium alcaliphilum 20Z	pmoA	MEALZ_0515
Methylomonas methanica MC09	pmoA	Metme_00037360
Methylovulum psychrotolerans HV10_M2	pmoA	Ga0226488_112611
Methylacidiphilum fumariolicum SolV	pmoA	Ga0069468_111471
Methylacidiphilum fumariolicum SolV	pmoA	Ga0069468_111639
Methylacidiphilum fumariolicum SolV	pmoA	Ga0069468_111642
Methylacidiphilum infernorum V4	pmoA	Minf_1507
Methylacidiphilum infernorum V4	pmoA	Minf_1510
Methylacidiphilum infernorum V4	pmoA	Minf_1590

Table A3.4 Biomarker gene families composing the dissimilatory nitrate reduction pathway. Specific protein clusters are listed in decreasing order of LDA effect size. The higher order pathway, denitrification, was not deemed differentially abundant.

N-cycling functional gene	Protein cluster (LCA)	LDA effect size	P-value	Treatment
narH	UniRef90_UPI00035DFE69 (*Methylosarcina*)	2.81	0.018	Mussel
narG	UniRef90_A9Y2C8 (Bacteria)	2.80	0.047	No mussel
narG	UniRef90_D5KJU2 (Bacteria)	2.78	0.047	No mussel
narG	UniRef90_B6E5A4 (Bacteria)	2.65	0.042	No mussel
narH	UniRef90_UPI0004057A01 (*Pleomorphomonas*)	2.57	0.018	No mussel
narG	UniRef90_B6E5P2 (Bacteria)	2.53	0.043	Mussel
narG	UniRef90_G3J0V4 (*Methylobacter*)	2.38	0.043	Mussel
narG	UniRef90_C0M0N4 (Bacteria)	2.21	0.021	No mussel
narG	UniRef90_H0TSP6 (*Bradyrhizobium*)	2.20	0.021	No mussel
narG	UniRef90_B1G0C2 (*Burkholderia*)	2.19	0.043	Mussel
narG	UniRef90_B1PUT8 (Bacteria)	2.19	0.043	Mussel

Table A3.5 Differentially abundant protein clusters within the parent category, DNRA.

N-cycling functional gene	Protein cluster (LCA)	LDA effect size	P-value	Treatment
nrfA	UniRef90_S4UCW4 (Bacteria)	3.22	0.014	No mussel
nrfA	UniRef90_S4UD05 (Bacteria)	2.79	0.047	No mussel
nirB	UniRef90_V5BZ33 (Methyloglobulus)	2.94	0.014	Mussel
nrfA	UniRef90_S4UG91 (Bacteria)	2.46	0.047	Mussel

Table A3.6 N-fixation functional genes were found to be differentially abundant within both treatments.

N-cycling functional gene	Protein cluster (LCA)	LDA effect size	P-value	Treatment
nifD	UniRef90_B4ULZ0 (*Anaeromyxobacter*)	2.65	0.043	No mussel
nifD	UniRef90_A7HET0 (*Anaeromyxobacter*)	2.64	0.043	No mussel
nifH	UniRef90_Q4PS30 (unclassified)	2.53	0.047	No mussel
nifH	UniRef90_E1QE30 (*Desulfarculus*)	2.43	0.020	No mussel
nifH	UniRef90_W5U1A9 (unclassified)	2.42	0.047	Mussel
nifH	UniRef90_G8B3A2 (Bacteria)	2.41	0.018	No mussel
nifH	UniRef90_B7ZGG9 (Alphaproteobacteria)	2.36	0.014	No mussel
nifH	UniRef90_Q8KKJ6 (Bacteria)	2.36	0.042	Mussel
nifH	UniRef90_B6DAC8 (Bacteria)	2.33	0.020	No mussel
nifD	UniRef90_D3RUE5 (Chromatiaceae)	2.30	0.047	Mussel
nifH	UniRef90_X2IXN1 (Bacteria)	2.30	0.047	Mussel
nifH	UniRef90_D3H5I8 (Bacteria)	2.28	0.047	No mussel
nifH	UniRef90_X2J0I1 (Bacteria)	2.28	0.047	No mussel
nifD	UniRef90_I3CFW7 (*Beggiatoa*)	2.24	0.047	Mussel

Table A3.7 Denitrification gene families found to be differentially abundant, despite the denitrification pathway not found to be statistically different.

N-cycling functional gene	Protein cluster (LCA)	LDA effect size	P-value	Treatment
Nitrous oxide reduction				
nosZ	UniRef90_K7XSE7 (Bacteria)	2.76	0.021	No mussel
nosZ	UniRef90_K7Y664 (Bacteria)	2.70	0.047	No mussel
nosZ	UniRef90_K7XTV0 (Bacteria)	2.68	0.047	No mussel
nosZ	UniRef90_K7XMD0 (Bacteria)	2.65	0.020	No mussel
nosZ	UniRef90_K7WLZ2 (unclassified)	2.65	0.047	No mussel
nosZ	UniRef90_B8R0E4 (Bacteria)	2.59	0.047	No mussel
nosZ	UniRef90_K7XU27 (Bacteria)	2.57	0.020	No mussel
nosZ	UniRef90_K7X6F4 (Bacteria)	2.57	0.047	No mussel
nosZ	UniRef90_B8R0A6 (Bacteria)	2.48	0.043	Mussel
nosZ	UniRef90_K7X032 (Bacteria)	2.45	0.038	No mussel
nosZ	UniRef90_A1KA74 (Azoarcus)	2.43	0.047	No mussel
nosZ	UniRef90_K7XCX6 (unclassified)	2.40	0.042	No mussel
nosZ	UniRef90_B8R0B8 (Bacteria)	2.37	0.043	Mussel
nosZ	UniRef90_S5UJ63 (Bacteria)	2.27	0.020	Mussel
nosZ	UniRef90_H9BVI3 (Bacteria)	2.24	0.047	No mussel
Nitric oxide reduction				
norB	UniRef90_F8S9B7 (Bacteria)	2.94	0.047	No mussel
norB	UniRef90_F8S960 (Bacteria)	2.87	0.038	No mussel
norB	UniRef90_F8S8X5 (Bacteria)	2.73	0.047	No mussel
norB	UniRef90_G2FJZ6 (Gammaproteobacteria)	2.49	0.047	Mussel
norB	UniRef90_I1ZIV3 (*Zoogloea*)	2.46	0.021	No mussel
norB	UniRef90_C5NT92 (*Ochrobactrum*)	2.39	0.047	No mussel
norB	UniRef90_V9TN78 (Proteobacteria)	2.36	0.047	Mussel
norB	UniRef90_L0PRW7 (unclassified)	2.35	0.047	Mussel
norB	UniRef90_Q84D93 (Bacteria)	2.21	0.042	Mussel
Nitrite reduction				
nirS	UniRef90_B2CG04 (Bacteria)	2.54	0.043	No mussel
nirS	UniRef90_A6YLH2 (Bacteria)	2.53	0.047	No mussel
nirS	UniRef90_Q59HK1 (Bacteria)	2.47	0.047	Mussel
nirS	UniRef90_G0Z4D9 (Bacteria)	2.44	0.021	No mussel
nirS	UniRef90_F6IBN8 (Bacteria)	2.41	0.047	Mussel
nirS	UniRef90_A0A024BRU3 (Bacteria)	2.38	0.047	Mussel

Table A3.7 cont.

N-cycling functional gene	Protein cluster (LCA)	LDA effect size	P-value	Treatment
nirS	UniRef90_Q2F353 (Bacteria)	2.37	0.047	Mussel
nirS	UniRef90_A0A024BR71 (Bacteria)	2.30	0.047	Mussel
nirS	UniRef90_Q6TBA1 (Bacteria)	2.22	0.020	No mussel
nirS	UniRef90_A0A024BS11 (Bacteria)	2.22	0.021	Mussel
nirS	UniRef90_G4XN08 (Bacteria)	2.19	0.042	Mussel

Supplementary material for Chapter 4

Table A4.1 Most abundant protein functions for *Nitrospira moscovinesis* metagenomic reads, listed as a percentage of total reads with a SEED classification.

Top 25% SEED subsystems			
No Mussels		**Mussel Habitat**	
Peptidoglycan Biosynthesis (mur/ddl genes)	3.00%	Peptidoglycan Biosynthesis (mur/ddl genes)	3.12%
Respiratory Complex I	2.65%	Respiratory Complex I	3.09%
Entner-Doudoroff Pathway	2.49%	Bacterial Cytoskeleton	2.22%
Bacterial Cytoskeleton	2.23%	Multidrug Resistance Efflux Pumps	2.18%
DNA-replication	2.17%	Entner-Doudoroff Pathway	1.91%
Multidrug Resistance Efflux Pumps	2.15%	DNA-replication	1.84%
DNA repair, bacterial	1.74%	DNA repair, bacterial	1.76%
5-FCL-like protein	1.60%	Folate Biosynthesis	1.72%
EC 6.1.1.- Ligases forming aminoacyl-tRNA and related compounds	1.60%	Cysteine Biosynthesis	1.70%
Biotin biosynthesis	1.55%	Proteasome bacterial	1.68%
Bacterial Chemotaxis	1.40%	EC 6.1.1.- Ligases forming aminoacyl-tRNA and related compounds	1.55%
Glutathione-regulated potassium-efflux system and associated functions	1.38%	YebC	1.39%
Bacterial motility:Gliding	1.33%	alpha carboxysome	1.37%
Top 25% SEED enzymes			
No Mussels		**Mussel Habitat**	
Acriflavin resistance protein	1.85%	Acriflavin resistance protein	1.92%
Gluconolactonase (EC 3.1.1.17)	1.18%	ATP-dependent protease (EC 3.4.21.53) Type I	1.67%
Carbon starvation protein A	1.14%	Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB	1.02%
Response regulator of zinc sigma-54-dependent two-component system	1.08%	Threonyl-tRNA synthetase (EC 6.1.1.3)	0.99%
Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB	1.00%	Carbon starvation protein A	0.92%
Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	0.98%	NADH-ubiquinone oxidoreductase chain G (EC 1.6.5.3)	0.92%
1-deoxy-D-xylulose 5-phosphate synthase (EC 2.2.1.7)	0.93%	Thymidylate synthase thyX (EC 2.1.1.-)	0.87%
Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)	0.91%	Gluconolactonase (EC 3.1.1.17)	0.87%
ATP-dependent protease La (EC 3.4.21.53) Type I	0.83%	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	0.85%
NADH-ubiquinone oxidoreductase chain G (EC 1.6.5.3)	0.80%	Response regulator of zinc sigma-54-dependent two-component system	0.81%
Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)	0.78%	NADH-ubiquinone oxidoreductase chain M (EC 1.6.5.3)	0.79%
Multimodular transpeptidase-transglycosylase (EC 2.4.1.129)	0.74%	Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)	0.79%

Table A4.1 cont.

Top 25% SEED enzymes cont.			
No Mussels		**Mussel Habitat**	
Adenylylsulfate kinase (EC 2.7.1.25)	0.70%	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)	0.77%
Signal transduction histidine kinase CheA (EC 2.7.3.-)	0.69%	ATP-dependent DNA helicase UvrD/PcrA, proteobacterial paralog	0.76%
Hypothetical transmembrane protein coupled to NADH-ubiquinone oxidoreductase homolog (chain L)	0.67%	Hypothetical transmembrane protein coupled to NADH-ubiquinone oxidoreductase homolog (chain L)	0.70%
diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)	0.66%	Adenylylsulfate kinase (EC 2.7.1.25)	0.70%
tRNA pseudouridine 13 synthase (EC 4.2.1.-)	0.66%	Single-stranded-DNA-specific exonuclease RecJ (EC 3.1.-.-)	0.69%
Soluble pyridine nucleotide transhydrogenase (EC 1.6.1.1)	0.65%	Deoxyhypusine synthase (EC 2.5.1.46)	0.67%
DNA mismatch repair protein MutS	0.64%	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8)	0.66%
Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4)	0.63%	Soluble pyridine nucleotide transhydrogenase (EC 1.6.1.1)	0.65%
Deoxyhypusine synthase (EC 2.5.1.46)	0.62%	Sulfate adenylyltransferase subunit 2 (EC 2.7.7.4)	0.65%
NADH-ubiquinone oxidoreductase chain M (EC 1.6.5.3)	0.61%	D-alanine--D-alanine ligase (EC 6.3.2.4)	0.64%
Acetylornithine aminotransferase (EC 2.6.1.11)	0.61%	Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)	0.63%
Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8)	0.59%	DNA mismatch repair protein MutS	0.61%
DNA recombination protein RmuC	0.59%	Signal transduction histidine kinase CheA (EC 2.7.3.-)	0.59%
Chaperone protein DnaJ	0.59%	TldD family protein, Actinobacterial subgroup	0.58%
Chromosomal replication initiator protein DnaA	0.58%	Thymidylate kinase (EC 2.7.4.9)	0.58%
Thymidylate synthase thyX (EC 2.1.1.-)	0.57%	diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)	0.57%
tRNA nucleotidyltransferase (EC 2.7.7.21) (EC 2.7.7.25)	0.55%	DNA recombination protein RmuC	0.55%
3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)	0.55%	Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)	0.55%
Superoxide dismutase [Mn] (EC 1.15.1.1)	0.54%	Acetylornithine aminotransferase (EC 2.6.1.11)	0.54%
TldD family protein, Actinobacterial subgroup	0.52%	Cell division protein FtsZ (EC 3.4.24.-)	0.54%
DNA polymerase III alpha subunit (EC 2.7.7.7)	0.52%		
Ribosomal protein S12p Asp88 (E. coli) methylthiotransferase	0.51%		

Table A4.2 Protein functions with statistically significant LDA effect sizes for *Nitrospira moscoviensis*.

Treatment	SEED classification	LDA effect size	P-value	*N. Moscoviensis* gene ID	N. moscoviensis gene name
Mussel Habitat	YebC-like protein (EC 6.1.1.3)	3.49	0.043	NITMOv2_0644; NITMOv2_0329	YchF and thrS
	hypothetical protein related to heme utilization (FIG039061)	3.33	0.021	-	-
	ATP dependent DNA helicase (UvrD/PcrA) proteobacterial paralog	3.28	0.043	NITMOv2_1186	PcrA
	Histidinol dehydrogenase (EC 1.1.1.23)	2.91	0.021	NITMOv2_0825	hisD
	Ribulose phosphate-3-epimerase (EC 5.1.3.1)	2.87	0.021	NITMOv2_0323	cbbE
	Ribonuclease H III (EC 3.1.26.4)	2.85	0.043	NITMOv2_0850	rnhC
	Type IV pilus biogenesis protein (PilQ)	2.84	0.043	NITMOv2_1246	pilQ
	Signal recognition particle subunit (Ffh) SRP54 (EC 3.A.5.1.1)	2.76	0.043	NITMOv2_3569	ffh
	4-alpha-glucanotransferase amylomaltase (EC 2.4.1.25)	2.74	0.043	NITMOv2_1162	malQ
	GMP synthase (glutamine-hydrolyzing) (EC 6.3.5.2)	2.68	0.021	NITMOv2_1344	guaA
	Alkaline phosphatase (EC 3.1.3.1)	2.67	0.021	NITMOv2_2620	PhoP
No Mussel	Glutathione regulated potassiumefflux system	3.31	0.021	NITMOv2_4520	CPA2 protein family (i.e. KefB)
	Phosphogluconolactonase	3.16	0.043	NITMOv2_0274	pgl
	Misc. protein function (COG2363)	3.10	0.021	-	integral membrane protein of unknown function DUF423 with Metl-like superfamily
	Superoxide dismutase Fe-Mn (EC 1.15.1.1)	3.10	0.043	NITMOv2_2805	sod
	Asparagine synthetase (glutamine-hydrolyzing) (EC.6.3.5.4)	3.06	0.021	NITMOv2_3400	asnB
	Glucose methanol choline (GMC) oxidoreductase NAD binding site	3.04	0.043	NITMOv2_4007	gmc
	Pyruvate decarboxylase (EC 4.1.1.1)	3.04	0.021	NITMOv2_3169	cfp
	Carbon starvation protein A	3.00	0.021	NITMOv2_0147	cstA
	Beta hexosaminidase (EC 3.2.1.52)	2.94	0.043	NITMOv2_0180	nagZ

Table A4.2 cont.

Treatment	SEED classification	LDA effect size	P-value	*N. Moscoviensis* gene ID	N. moscoviensis gene name
No Mussel cont.	Tryptophanase (EC 4.1.99.1)	2.83	0.021	NITMOv2_0067	tnaA
	Ribosomal large subunit pseudouridine synthase D (EC 4.2.1.70)	2.76	0.043	NITMOv2_0475	truD
	Endonuclease III (EC 4.2.99.18)	2.72	0.021	NITMOv2_4513	nth
	Xylulose-5-phosphate phosphoketolase (EC 4.1.2.9)	2.72	0.043	NITMOv2_2536	xfp
	Flagellar biosynthesis protein (FlhB)	2.72	0.043	NITMOv2_2196	flhB
	Acetoacetate metabolism regulatory protein (AtoC)	2.70	0.021	NITMOv2_1164; NITMOv2_1908; NITMOv2_2568; NITMOv2_3235; NITMOv2_3750; NITMOv2_4352	AtoC

Table A4.3 Most abundant protein functions for *Candidatus* Nitrospira inopinata metagenomic reads, listed as a percentage of total reads with a SEED classification.

Top 25% SEED subsystems			
No Mussels		**Mussel Habitat**	
Restriction-Modification System	5.56%	5-FCL-like protein	4.34%
Urea decomposition	4.00%	FOL Commensurate regulon activation	4.34%
5-FCL-like protein	3.92%	Biogenesis of c-type cytochromes	4.27%
Ammonia monooxygenase	3.65%	Ammonia monooxygenase	3.79%
Biogenesis of c-type cytochromes	3.48%	Urea decomposition	3.52%
FOL Commensurate regulon activation	3.26%	Restriction-Modification System	3.16%
Glycogen metabolism	2.56%	Glycogen metabolism	2.83%
Top 25% most abundant SEED enzymes			
No Mussels		**Mussel Habitat**	
RND efflux system, inner membrane transporter CmeB	3.26%	RND efflux system, inner membrane transporter CmeB	4.34%
Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)	3.10%	Ammonia monooxygenase C-subunit (EC 1.14.13.25)	2.29%
Ammonia monooxygenase C-subunit (EC 1.14.13.25)	2.51%	Alcohol dehydrogenase (EC 1.1.1.1)	1.83%
Osmosensitive K+ channel histidine kinase KdpD (EC 2.7.3.-)	1.80%	Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)	1.69%
Alcohol dehydrogenase (EC 1.1.1.1)	1.68%	1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)	1.51%
DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)	1.38%	Cytochrome c heme lyase subunit CcmF	1.45%
Cytochrome c heme lyase subunit CcmF	1.27%	Osmosensitive K+ channel histidine kinase KdpD (EC 2.7.3.-)	1.34%
1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)	1.24%	Cobalt-zinc-cadmium resistance protein CzcA	1.30%
Type III restriction-modification system methylation subunit (EC 2.1.1.72)	1.15%	Cytochrome c551 peroxidase (EC 1.11.1.5)	1.30%
Putative two-domain glycosyltransferase	1.13%	Glycogen phosphorylase (EC 2.4.1.1)	1.14%
Glycogen phosphorylase (EC 2.4.1.1)	1.10%	Ferredoxin	1.10%
Archaeal S-adenosylmethionine synthetase (EC 2.5.1.6)	1.08%	Putative two-domain glycosyltransferase	1.04%
Long-chain-fatty-acid Acetyl-CoA ligase (EC 6.2.1.3)	1.06%	Exopolyphosphatase (EC 3.6.1.11)	1.03%
Cobalt-zinc-cadmium resistance protein CzcA	1.05%	Archaeal S-adenosylmethionine synthetase (EC 2.5.1.6)	1.01%
DNA recombination protein RmuC	1.02%	[NiFe] hydrogenase metallocenter assembly protein HypF	0.98%
Cytochrome c551 peroxidase (EC 1.11.1.5)	0.95%	Urea ABC transporter, permease protein UrtB	0.93%
Urea ABC transporter, urea binding protein	0.95%	Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4)	0.92%

Table A4.4 Protein functions with statistically significant effect sizes for *Candidatus* Nitrospira inopinata.

Treatment	SEED classification	LDA effect size	P-value	*Ca*. N. inopinata ID	Gene name
Mussel Habitat	RND efflux systeminner membrane transporter (CmeB)	3.81	0.043	NITINOP_0165 NITINOP_0322	acrB
	NiFe hydrogenase metallocenter assembly protein (HypF)	3.21	0.043	NITINOP_0583	hypF
	Thiamine-monophosphate kinase	3.11	0.021	NITINOP_3242 NITINOP_3243	thiE; thiG
	Cytochrome c type biogenesis protein (CcmE) heme chaperone	3.00	0.043	NITINOP_0071	ccmE
	Lipopolysaccharide heptosyltransferase I (EC 2.4.1.-)	2.89	0.043	NITINOP_3119	putative ADP-heptose-LPS heptosyl-transferase
	Threonine synthase (EC 4.2.3.1)	2.45	0.014	NITINOP_0769	thrC
	Phosphoribosylformyl-glycinamidine synthase (EC 6.3.5.3)	2.37	0.014	NITINOP_1803	purL
	A/G specific adenine glycosylase (EC 3.2.2.-)	2.36	0.047	NITINOP_0052	mutS
	HflK membrane protease (EC 3.4.-.-)	2.35	0.047	NITINOP_0656 NITINOP_2485	ftsH
	Carbamoyl phosphate synthase large chain (EC 6.3.5.5)	2.35	0.047	NITINOP_2878	carB
	Sensor protein of zinc sigma-54-dependent two component system	2.33	0.047	NITINOP_0732	kdpD
	Periplasmic divalent cation tolerance protein (cutA)	2.23	0.047	NITINOP_1511	cutA
	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha (EC 2.5.1.54)	2.22	0.047	NITINOP_1287 NITINOP_2353	aroF
	Phosphoribosylformyl-glycinamidine synthase (EC 6.3.5.3)	2.22	0.047	NITINOP_1803	purL
	Cytidylate kinase (EC 2.7.4.25)	2.20	0.047	NITINOP_1290	cmk
	RND efflux systemmembrane fusion protein (CmeA)	2.19	0.047	NITNOP_0560 NITNOP_0488 NITNOP_2329	putative efflux transporter and membrane fusion protein
	Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19)	2.15	0.047	NITINOP_0113	hisI
	Translation initiation factor 2	2.15	0.047	NITINOP_0786	infB
	Integral inner membrane protein of type IV secretion complex	2.13	0.047	NITINOP_2980	virB

Table A4.4 cont.

Treatment	SEED classification	LDA effect size	P-value	*Ca*. N. inopinata ID	Gene name
No mussels	Lipopolysaccharide heptosyltransferase II	2.77	0.021	NITINOP_3124	gmhB
	Agmatinase (EC 3.5.3.11)	2.76	0.021	NITINOP_1173	speB
	NADH-quinone oxidoreductase chain G (EC 1.6.5.3)	2.66	0.021	NITINOP_1374	nuoG
	Quinolinate phosphoribosyltransferase; decarboxylating (EC 2.4.2.19)	2.63	0.021	NITINOP_2729	nadC
	Ribonucleotide reductase of class II coenzyme B12 dependent (EC 1.17.4.1)	2.57	0.021	NITINOP_2043	nrdJ
	Imidazole glycerol phosphate synthase amidotransferase subunit (EC 2.4.2.-)	2.51	0.021	NITINOP_0117	hisH
	Pseudouridine synthase family protein	2.44	0.021	NITNOP_1829	rluB
	Excinuclease ABC subunit C	2.37	0.021	NITINOP_1361	uvrC
	DNA recombination and repair protein (RecO)	2.37	0.042	NITINOP_2378	recO
	Heat shock protein	2.35	0.043	NITINOP_0609	Hsp20 family
	Potassium transporting ATPase (EC 3.6.3.12)	2.31	0.020	NITINOP_1503 NITINOP_1504	kdpA; kdpB
	DNA polymerase III (EC 2.7.7.7)	2.29	0.042	NITINOP_1079	dnaX

REFERENCES

1. Galloway, J.N., et al., *Nitrogen cycles: past, present, and future.* Biogeochemistry, 2004. **70**(2): p. 153-226.

2. USGS, *Nitrogen (fixed)- Ammonia.* Mineral Commodity Summaries, 2015: p. 112-113.

3. Frink, C.R., P.E. Waggoner, and J.H. Ausubel, *Nitrogen fertilizer: Retrospect and prospect*. Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(4): p. 1175-1180.

4. Diaz, R.J. and R. Rosenberg, *Spreading dead zones and consequences for marine ecosystems.* Science, 2008. **321**(5891): p. 926-929.

5. David, M.B., L.E. Drinkwater, and G.F. McLsaac, *Sources of Nitrate Yields in the Mississippi River Basin.* Journal of Environmental Quality, 2010. **39**(5): p. 1657-1667.

6. Rabalais, N.N.a.T., R. E, *2017 Shelfwide Cruise Press Release*, in *Press Release*. 2017, Louisana University Marine Consortium.

7. Newton, T.J., et al., *Population assessment and potential functional roles of native mussels in the Upper Mississippi River.* Aquatic Conservation-Marine and Freshwater Ecosystems, 2011. **21**(2): p. 122-131.

8. Hunter, P.L., Scott *The Top 10 Biggest Wastewater Treatment Plants*. 2012.

9. Strayer, D.L., *Understanding how nutrient cycles and freshwater mussels (Unionoida) affect one another.* Hydrobiologia, 2014. **735**(1): p. 277-292.

10. Newell, R.I.E., *Ecosystem influences of natural and cultivated populations of suspension-feeding bivalve molluscs: A review.* Journal of Shellfish Research, 2004. **23**(1): p. 51-61.

11. Vaughn, C.C. and C.C. Hakenkamp, *The functional role of burrowing bivalves in freshwater ecosystems.* Freshwater Biology, 2001. **46**(11): p. 1431-1446.

12. Matisoff, G., J.B. Fisher, and S. Matis, *Effects of benthic macroinvertebrates on the exchange of solutes between sediments and freshwater.* Hydrobiologia, 1985. **122**(1): p. 19-33.

13. Bril, J.S., et al., *Sensor data as a measure of native freshwater mussel impact on nitrate formation and food digestion in continuous-flow mesocosms.* Freshwater Science, 2014. **33**(2): p. 417-424.

14. Segata, N., et al., *Metagenomic microbial community profiling using unique clade-specific marker genes.* Nature Methods, 2012. **9**(8): p. 811-+.

15. Bruenderman, S., Sternburg, J., Barnhart, C., *Missouri's Freshwater Mussels.* Conservation Commission of the State of Missouri, 2002.

16. Sonthiphand, P. and J.D. Neufeld, *Evaluating Primers for Profiling Anaerobic Ammonia Oxidizing Bacteria within Freshwater Environments.* Plos One, 2013. **8**(3).

17. Stein, L.Y. and M.G. Klotz, *The nitrogen cycle.* Current Biology, 2016. **26**(3): p. R94-R98.

18. Crowe, S.A., et al., *Anammox, denitrification and fixed-nitrogen removal in sediments from the Lower St. Lawrence Estuary.* Biogeosciences, 2012. **9**(11): p. 4309-4321.

19. Kartal, B., et al., *Molecular mechanism of anaerobic ammonium oxidation.* Nature, 2011. **479**(7371): p. 127-U159.

20. Kuenen, J.G., *Anammox bacteria: from discovery to application.* Nature Reviews Microbiology, 2008. **6**(4): p. 320-326.

21. Kartal, B., et al., *Chapter 3 - Anammox—Growth Physiology, Cell Biology, and Metabolism*, in *Advances in Microbial Physiology*, R.K. Poole, Editor. 2012, Academic Press. p. 211-262.

22. Kartal, B., et al., *Anammox bacteria disguised as denitrifiers: nitrate reduction to dinitrogen gas via nitrite and ammonium.* Environmental Microbiology, 2007. **9**(3): p. 635-642.

23. Francis, C.A., J.M. Beman, and M.M.M. Kuypers, *New processes and players in the nitrogen cycle: the microbial ecology of anaerobic and archaeal ammonia oxidation.* Isme Journal, 2007. **1**(1): p. 19-27.

24. Penton, C.R., A.H. Devol, and J.M. Tiedje, *Molecular evidence for the broad distribution of anaerobic ammonium-oxidizing bacteria in freshwater and marine sediments.* Applied and Environmental Microbiology, 2006. **72**(10): p. 6829-6832.

25. Smith, R.L., et al., *Role of Anaerobic Ammonium Oxidation (Anammox) in Nitrogen Removal from a Freshwater Aquifer.* Environmental Science & Technology, 2015. **49**(20): p. 12169-12177.

26. Sonthiphand, P., M.W. Hall, and J.D. Neufeld, *Biogeography of anaerobic ammonia-oxidizing (anammox) bacteria.* Frontiers in Microbiology, 2014.

27. Sun, W., et al., *Molecular diversity and distribution of anammox community in sediments of the Dongjiang River, a drinking water source of Hong Kong.* Journal of Applied Microbiology, 2014. **116**(2): p. 464-476.

28. Wu, Y.C., et al., *Molecular Detection of Novel Anammox Bacterial Clusters in the Sediments of the Shallow Freshwater Lake Taihu.* Geomicrobiology Journal, 2012. **29**(9): p. 852-859.

29. Costa, E., J. Pérez, and J.-U. Kreft, *Why is metabolic labour divided in nitrification?* Trends in Microbiology, 2006. **14**(5): p. 213-219.

30. van Kessel, M., et al., *Complete nitrification by a single microorganism*. Nature, 2015. **528**(7583): p. 555-+.

31. Santoro, A.E., *The do-it-all nitrifier The discovery of bacteria that can oxidize both ammonia and nitrite upends a long-held dogma.* Science, 2016. **351**(6271): p. 342-343.

32. Kits, K.D., et al., *Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle.* Nature, 2017. **549**(7671): p. 269-+.

33. Koch, H., et al., *Expanded metabolic versatility of ubiquitous nitrite-oxidizing bacteria from the genus Nitrospira.* Proceedings of the National Academy of Sciences, 2015. **112**(36): p. 11371-11376.

34. Palatinszky, M., et al., *Cyanate as energy source for nitrifiers.* Nature, 2015. **524**(7563): p. 105-108.

35. Koch, H., et al., *Growth of nitrite-oxidizing bacteria by aerobic hydrogen oxidation.* Science, 2014. **345**(6200): p. 1052-1054.

36. Daims, H., S. Lücker, and M. Wagner, *A New Perspective on Microbes Formerly Known as Nitrite-Oxidizing Bacteria.* Trends in Microbiology, 2016. **24**(9): p. 699-712.

37. Dang, H.Y., et al., *Environmental Factors Shape Sediment Anammox Bacterial Communities in Hypernutrified Jiaozhou Bay, China*. Applied and Environmental Microbiology, 2010. **76**(21): p. 7036-7047.

38. Bi, Z., et al., *Effects of the C/N ratio and bacterial populations on nitrogen removal in the simultaneous anammox and heterotrophic denitrification process: Mathematic modeling and batch experiments.* Chemical Engineering Journal, 2015. **280**: p. 606-613.

39. Lu, S.M., et al., *Vertical Segregation and Phylogenetic Characterization of Ammonia-Oxidizing Bacteria and Archaea in the Sediment of a Freshwater Aquaculture Pond*. Frontiers in Microbiology, 2016. **6**.

40. Rasigraf, O., et al., *Metagenomic potential for and diversity of N-cycle driving microorganisms in the Bothnian Sea sediment*. Microbiologyopen, 2017. **6**(4): p. 13.

41. Ulloa, O., et al., *Microbial oceanography of anoxic oxygen minimum zones.* Proceedings of the National Academy of Sciences, 2012. **109**(40): p. 15996-16003.

42. Oulas, A., et al., *Metagenomics: Tools and Insights for Analyzing Next-Generation Sequencing Data Derived from Biodiversity Studies.* Bioinformatics and Biology Insights, 2015. **9**: p. 75-88.

43. Morgan, X.C. and C. Huttenhower, *Chapter 12: Human Microbiome Analysis.* PLOS Computational Biology, 2012. **8**(12): p. e1002808.

44. Morales, Y., et al., *Effects of substrate and hydrodynamic conditions on the formation of mussel beds in a large river*. Journal of the North American Benthological Society, 2006. **25**(3): p. 664-676.

45. Newton, T.J., et al., *Population assessment and potential functional roles of native mussels in the Upper Mississippi River.* Aquatic Conservation: Marine and Freshwater Ecosystems, 2011. **21**(2): p. 122-131.

46. Ikenberry, C.D., et al., *Nitrate-Nitrogen Export: Magnitude and Patterns from Drainage Districts to Downstream River Basins.* Journal of Environmental Quality, 2014. **43**(6): p. 2024-2033.

47. Schilling, K.E., C.F. Wolter, and E. McLellan, *Agro-hydrologic Landscapes in the Upper Mississippi and Ohio River Basins.* Environmental Management, 2015. **55**(3): p. 646-656.

48. Vaughn, C.C., S.J. Nichols, and D.E. Spooner, *Community and foodweb ecology of freshwater mussels.* Journal of the North American Benthological Society, 2008. **27**(2): p. 409-423.

49. Young, N.C., *Physical characterization of freshwater mussel habitats in upper mississippi river pool 16.* (Order No. 3229716). Available from ProQuest Dissertations & Theses Global. (305340701). Retrieved from http://proxy.lib.uiowa.edu/login?url=http://search.proquest.com/docview/305340701?accountid=14663, 2006.

50. Haag, W.R., *North American Freshwater Mussels: Natural History, Ecology, and Conservation*. 2012, New York, NY: Cambridge University Press.

51. Allen, W.R., *Studies of the Biology of Freshwater Mussels. II, The Nature and Degree of Response to Certain Physical and Chemical Stimuli.* Ohio Journal of Science, 1923. **23**(2).

52. Matteson, M.R., *Studies on the Natural History of the Unionidae.* The American Midland Naturalist, 1955. **53**(1): p. 126-145.

53. Allen, D.C. and C.C. Vaughn, *Burrowing behavior of freshwater mussels in experimentally manipulated communities.* Journal of the North American Benthological Society, 2009. **28**(1): p. 93-100.

54. Schwalb, A.N. and M.T. Pusch, *Horizontal and vertical movements of unionid mussels in a lowland river.* Journal of the North American Benthological Society, 2007. **26**(2): p. 261-272.

55. Newton, T.J., S.J. Zigler, and B.R. Gray, *Mortality, movement and behaviour of native mussels during a planned water-level drawdown in the Upper Mississippi River.* Freshwater Biology, 2015. **60**(1): p. 1-15.

56. Bril, J.S., et al., *Simulated mussel mortality thresholds as a function of mussel biomass and nutrient loading.* PeerJ, 2017.

57. Bayne, B.L., *Physiological Changes in Mytilus-Edulis-L Induced by Temperature and Nutritive Stress.* Journal of the Marine Biological Association of the United Kingdom, 1973. **53**(1): p. 39-58.

58. Atkinson, C.L., J.F. Kelly, and C.C. Vaughn, *Tracing Consumer-Derived Nitrogen in Riverine Food Webs.* Ecosystems, 2014. **17**(3): p. 485-496.

59. Bril, J.L., Kathryn; Just, Craig; Spak, Scott; Newton; Teresa, *Simulated Mussel Mortality Thresholds as a Function of Mussel Biomass and Nutrient Loading.* In Review, 2016.

60. Houser, J.N. and W.B. Richardson, *Nitrogen and phosphorus in the Upper Mississippi River: transport, processing, and effects on the river ecosystem.* Hydrobiologia, 2010. **640**(1): p. 71-88.

61. Hill, B.H., et al., *A Synoptic Survey of Nitrogen and Phosphorus in Tributary Streams and Great Rivers of the Upper Mississippi, Missouri, and Ohio River Basins.* Water Air and Soil Pollution, 2011. **216**(1-4): p. 605-619.

62. Hill, B.H., C. M. Elonen, T. M. Jicha, and D. W. Bolgrien. *Nutrient Chemistry and Microbial Activity in the Upper Mississippi River Basin: Stoichiometry and Downstream Patterns.* in *American Water Resources Association Annual Meeting*. 2008. New Orleans, LA: National Health and Environmental Effects Research Laboratory.

63. Mason, O.U., et al., *Mississippi River Plume Enriches Microbial Diversity in the Northern Gulf of Mexico.* Frontiers in Microbiology, 2016. **7**.

64. Millar, J.J., et al., *Particle-associated and cell-free extracellular enzyme activity in relation to nutrient status of large tributaries of the Lower Mississippi River.* Biogeochemistry, 2015. **124**(1-3): p. 255-271.

65. Hayatsu, M., K. Tago, and M. Saito, *Various players in the nitrogen cycle: Diversity and functions of the microorganisms involved in nitrification and denitrification.* Soil Science and Plant Nutrition, 2008. **54**(1): p. 33-45.

66. Burrell, P.C., C.M. Phalen, and T.A. Hovanec, *Identification of bacteria responsible for ammonia oxidation in freshwater aquaria.* Applied and Environmental Microbiology, 2001. **67**(12): p. 5791-5800.

67. Aakra, A., et al., *Detailed phylogeny of ammonia-oxidizing bacteria determined by rDNA sequences and DNA homology values.* International Journal of Systematic and Evolutionary Microbiology, 2001. **51**: p. 2021-2030.

68. Prosser, J.I., I.M. Head, and L.Y. Stein, *The Family Nitrosomonadaceae*, in *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*, E. Rosenberg, et al., Editors. 2014, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 901-918.

69. Leininger, S., et al., *Archaea predominate among ammonia-oxidizing prokaryotes in soils.* Nature, 2006. **442**(7104): p. 806-809.

70. Martens-Habbena, W., et al., *Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria.* Nature, 2009. **461**(7266): p. 976-U234.

71. Hatzenpichler, R., et al., *A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring.* Proceedings of the National Academy of Sciences, 2008. **105**(6): p. 2134-2139.

72. Koper, T.E., et al., *Urease-Encoding Genes in Ammonia-Oxidizing Bacteria.* Applied and Environmental Microbiology, 2004. **70**(4): p. 2342-2348.

73. Spang, A., et al., *The genome of the ammonia-oxidizing Candidatus Nitrososphaera gargensis: insights into metabolic versatility and environmental adaptations.* Environmental Microbiology, 2012. **14**(12): p. 3122-3145.

74. Prosser, J.I. and G.W. Nicol, *Relative contributions of archaea and bacteria to aerobic ammonia oxidation in the environment.* Environmental Microbiology, 2008. **10**(11): p. 2931-2941.

75. Bollmann, A., M.J. Bar-Gilissen, and H.J. Laanbroek, *Growth at low ammonium concentrations and starvation response as potential factors involved in niche differentiation among ammonia-oxidizing bacteria.* Applied and Environmental Microbiology, 2002. **68**(10): p. 4751-4757.

76. Daims, H., et al., *Complete nitrification by Nitrospira bacteria.* Nature, 2015. **528**(7583): p. 504-+.

77. Lucker, S., et al., *A Nitrospira metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(30): p. 13479-13484.

78. Thamdrup, B., *New Pathways and Processes in the Global Nitrogen Cycle*, in *Annual Review of Ecology, Evolution, and Systematics, Vol 43*, D.J. Futuyma, Editor. 2012. p. 407-428.

79. Trimmer, M. and P. Engstrom, *Distribution, Activity, and Ecology of Anammox Bacteria in Aquatic Environments.* Nitrification, 2011: p. 201-235.

80. Raghoebarsing, A.A., et al., *A microbial consortium couples anaerobic methane oxidation to denitrification.* Nature, 2006. **440**(7086): p. 918-921.

81. Welte, C.U., et al., *Nitrate- and nitrite-dependent anaerobic oxidation of methane.* Environmental Microbiology Reports, 2016. **8**(6): p. 941-955.

82. Ding, J., et al., *Experimental evaluation of the metabolic reversibility of ANME-2d between anaerobic methane oxidation and methanogenesis.* Applied Microbiology and Biotechnology, 2016. **100**(14): p. 6481-6490.

83. Haroon, M.F., et al., *Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage.* Nature, 2013. **500**(7464): p. 567-570.

84. Ettwig, K.F., et al., *Nitrite-driven anaerobic methane oxidation by oxygenic bacteria.* Nature, 2010. **464**(7288): p. 543-548.

85. Ettwig, K.F., et al., *Enrichment and Molecular Detection of Denitrifying Methanotrophic Bacteria of the NC10 Phylum.* Applied and Environmental Microbiology, 2009. **75**(11): p. 3656-3662.

86. Padilla, C.C., et al., *NC10 bacteria in marine oxygen minimum zones.* ISME J, 2016. **10**(8): p. 2067-2071.

87. Luesken, F.A., et al., *Simultaneous Nitrite-Dependent Anaerobic Methane and Ammonium Oxidation Processes.* Applied and Environmental Microbiology, 2011. **77**(19): p. 6802-6807.

88. Deutzmann, J.S. and B. Schink, *Anaerobic Oxidation of Methane in Sediments of Lake Constance, an Oligotrophic Freshwater Lake.* Applied and Environmental Microbiology, 2011. **77**(13): p. 4429-4436.

89. Hu, S.H., et al., *Enrichment of denitrifying anaerobic methane oxidizing microorganisms.* Environmental Microbiology Reports, 2009. **1**(5): p. 377-384.

90. Pfister, C.A., F. Meyer, and D.A. Antonopoulos, *Metagenomic Profiling of a Microbial Assemblage Associated with the California Mussel: A Node in Networks of Carbon and Nitrogen Cycling.* Plos One, 2010. **5**(5).

91. Kellogg, M.L., et al., *Denitrification and nutrient assimilation on a restored oyster reef.* Marine Ecology Progress Series, 2013. **480**: p. 1-19.

92. Pollet, T., et al., *Metabolic Activity and Functional Diversity Changes in Sediment Prokaryotic Communities Organically Enriched with Mussel Biodeposits.* Plos One, 2015. **10**(4).

93. Lohner, R.N., et al., *A comparison of the benthic bacterial communities within and surrounding Dreissena clusters in lakes.* Microbial Ecology, 2007. **54**(3): p. 469-477.

94. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Research, 1997. **25**(17): p. 3389-3402.

95. Wang, Q., et al., *Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy.* Applied and Environmental Microbiology, 2007. **73**(16): p. 5261-5267.

96. Wang, H.L., et al., *Assessing nitrogen transformation processes in a trickling filter under hydraulic loading rate constraints using nitrogen functional gene abundances.* Bioresource Technology, 2015. **177**: p. 217-223.

97. Caporaso, J.G., et al., *Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms.* Isme Journal, 2012. **6**(8): p. 1621-1624.

98. Navas-Molina, J.A., et al., *Advancing Our Understanding of the Human Microbiome Using QIIME*, in *Microbial Metagenomics, Metatranscriptomics, and Metaproteomics*, E.F. DeLong, Editor. 2013. p. 371-444.

99. Edgar, R.C., *Search and clustering orders of magnitude faster than BLAST.* Bioinformatics, 2010. **26**(19): p. 2460-2461.

100. Caporaso, J.G., et al., *PyNAST: a flexible tool for aligning sequences to a template alignment*. Bioinformatics, 2010. **26**(2): p. 266-267.

101. DeSantis, T.Z., et al., *Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB.* Applied and Environmental Microbiology, 2006. **72**(7): p. 5069-5072.

102. Price, M.N., P.S. Dehal, and A.P. Arkin, *FastTree 2-Approximately Maximum-Likelihood Trees for Large Alignments.* Plos One, 2010. **5**(3).

103. McMurdie, P.J. and S. Holmes, *phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data.* Plos One, 2013. **8**(4).

104. Chao, A. and C.-H. Chiu, *Nonparametric Estimation and Comparison of Species Richness*, in *eLS*. 2001, John Wiley & Sons, Ltd.

105. Li, M., et al., *Seasonal Dynamics of Anammox Bacteria in Estuarial Sediment of the Mai Po Nature Reserve Revealed by Analyzing the 16S rRNA and Hydrazine Oxidoreductase (hzo) Genes.* Microbes and Environments, 2011. **26**(1): p. 15-22.

106. Bray, J.R. and J.T. Curtis, *An Ordination of the Upland Forest Communities of Southern Wisconsin*. Ecological Monographs, 1957. **27**(4): p. 326-349.

107. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biology, 2014. **15**(12).

108. Bucci, J., et al., *Seasonal Changes in Microbial Community Structure in Freshwater Stream Sediment in a North Carolina River Basin.* Diversity, 2014. **6**(1): p. 18.

109. Zeng, J., et al., *Spatial distribution of bacterial communities in sediment of a eutrophic lake revealed by denaturing gradient gel electrophoresis and multivariate analysis.* Canadian Journal of Microbiology, 2008. **54**(12): p. 1053-1063.

110. Zhang, J.X., et al., *Distribution of sediment bacterial and archaeal communities in plateau freshwater lakes.* Applied Microbiology and Biotechnology, 2015. **99**(7): p. 3291-3302.

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

112. Dai, Y., et al., *Spatiotemporal variation of planktonic and sediment bacterial assemblages in two plateau freshwater lakes at different trophic status.* Applied Microbiology and Biotechnology, 2016. **100**(9): p. 4161-4175.

113. Fernandez, N.T., et al., *Changes in the composition and diversity of the bacterial microbiota associated with oysters (Crassostrea corteziensis, Crassostrea gigas and Crassostrea sikamea) during commercial production.* Fems Microbiology Ecology, 2014. **88**(1): p. 69-83.

114. Lee, P.O., et al., *Invasive dreissenid mussels and benthic algae in Lake Michigan: characterizing effects on sediment bacterial communities.* Fems Microbiology Ecology, 2015. **91**(1): p. 1-12.

115. Ngangbam, A.K., et al., *Characterization of Bacterial Communities Associated with the Tyrian Purple Producing Gland in a Marine Gastropod.* Plos One, 2015. **10**(10).

116. Frischer, M.E., et al., *Interactions between zebra mussels (Dreissena polymorpha) and microbial communities.* Canadian Journal of Fisheries and Aquatic Sciences, 2000. **57**(3): p. 591-599.

117. Trabal, N., et al., *Molecular Analysis of Bacterial Microbiota Associated with Oysters (Crassostrea gigas and Crassostrea corteziensis) in Different Growth Phases at Two Cultivation Sites.* Microbial Ecology, 2012. **64**(2): p. 555-569.

118. Neta, M.T.S., et al., *Microbiological quality and bacterial diversity of the tropical oyster Crassostrea rhizophorae in a monitored farming system and from natural stocks.* Genetics and Molecular Research, 2015. **14**(4): p. 15754-15768.

119. Lozupone, C.A. and R. Knight, *Global patterns in bacterial diversity.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(27): p. 11436-11440.

120. Fierer, N. and R.B. Jackson, *The diversity and biogeography of soil bacterial communities.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(3): p. 626-631.

121. Pfister, C.A., J.A. Gilbert, and S.M. Gibbons, *The role of macrobiota in structuring microbial communities along rocky shores.* Peerj, 2014. **2**.

122. Michaud, E., et al., *Spatial interactions in the Macoma balthica community control biogeochemical fluxes at the sediment-water interface and microbial abundances.* Journal of Marine Research, 2009. **67**(1): p. 43-70.

123. Richter, A., et al., *Association between the occurrence of the Thick-shelled River Mussel (Unio crassus) and macroinvertebrate, microbial, and diatom communities.* Freshwater Science, 2016. **35**(3): p. 922-933.

124. Grenz, C., et al., *Insitu Biochemical and Bacterial Variation of Sediments Enriched with Mussel Biodeposits.* Hydrobiologia, 1990. **207**: p. 153-160.

125. Hegaret, H., G.H. Wikfors, and S.E. Shumway, *Diverse feeding responses of five species of bivalve mollusc when exposed to three species of harmful algae.* Journal of Shellfish Research, 2007. **26**(2): p. 549-559.

126. Tenore, K.R. and W.M. Dunstan, *Comparison of feeding and biodeposition of three bivalves at different food levels.* Marine Biology, 1973. **21**(3): p. 190-195.

127. Cranford, P.J., et al., *Influence of mussel aquaculture on nitrogen dynamics in a nutrient enriched coastal embayment*. Marine Ecology Progress Series, 2007. **347**: p. 61-78.

128. Kristensen, E., et al., *What is bioturbation? The need for a precise definition for fauna in aquatic sciences.* Marine Ecology Progress Series, 2012. **446**: p. 285-302.

129. Laverock, B., et al., *Bioturbating shrimp alter the structure and diversity of bacterial communities in coastal marine sediments.* Isme Journal, 2010. **4**(12): p. 1531-1544.

130. Novais, A., et al., *Effects of the invasive clam Corbicula fluminea (Muller, 1774) on an estuarine microbial community.* Science of the Total Environment, 2016. **566**: p. 1168-1175.

131. McCall, P.L., M.J.S. Tevesz, and S.F. Schwelgien, *Sediment Mixing by Lampsilis Radiata Siliquoidea (Mollusca) from Western Lake Erie.* Journal of Great Lakes Research, 1979. **5**(2): p. 105-111.

132. Chen, J. and J.-D. Gu, *Faunal Burrows Alter the Diversity, Abundance, and Structure of AOA, AOB, Anammox and n-Damo Communities in Coastal Mangrove Sediments.* Microbial Ecology, 2017: p. 1-17.

133. McCall, P.L., G. Matisoff, and M.J.S. Tevesz, *The Effects of a Unionid Bivalve on the Physical, Chemical, and Microbial Properties of Cohesive Sediments from Lake Erie.* American Journal of Science, 1986. **286**(2): p. 127-159.

134. Shen, L.-d., et al., *Evidence for anaerobic ammonium oxidation process in freshwater sediments of aquaculture ponds.* Environmental Science and Pollution Research, 2016. **23**(2): p. 1344-1352.

135. Humbert, S., et al., *Molecular detection of anammox bacteria in terrestrial ecosystems: distribution and diversity.* ISME J, 2009. **4**(3): p. 450-454.

136. Kartal, B., et al., *Candidatus 'Brocadia fulgida': an autofluorescent anaerobic ammonium oxidizing bacterium*. FEMS Microbiology Ecology, 2008. **63**(1): p. 46-55.

137. Shen, L.D., et al., *Vertical distribution and activity of anaerobic ammonium-oxidising bacteria in a vegetable field.* Geoderma, 2017. **288**: p. 56-63.

138. Galán, A., et al., *Anammox bacteria and the anaerobic oxidation of ammonium in the oxygen minimum zone off northern Chile.* Deep Sea Research Part II: Topical Studies in Oceanography, 2009. **56**(16): p. 1021-1031.

139. Zhu, G., et al., *Anaerobic ammonia oxidation in a fertilized paddy soil.* The ISME Journal, 2011. **5**(12): p. 1905-1912.

140. Shen, L.D., et al., *Depth-specific distribution and importance of nitrite-dependent anaerobic ammonium and methane-oxidising bacteria in an urban wetland.* Soil Biology & Biochemistry, 2015. **83**: p. 43-51.

141. Oshiki, M., H. Satoh, and S. Okabe, *Ecology and physiology of anaerobic ammonium oxidizing bacteria.* Environmental Microbiology, 2016. **18**(9): p. 2784-2796.

142. Shen, J.P., Z.H. Xu, and J.Z. He, *Frontiers in the microbial processes of ammonia oxidation in soils and sediments.* Journal of Soils and Sediments, 2014. **14**(6): p. 1023-1029.

143. Zhu, X., et al., *Ammonia oxidation pathways and nitrifier denitrification are significant sources of N2O and NO under low oxygen availability.* Proceedings of the National Academy of Sciences, 2013. **110**(16): p. 6328-6333.

144. Zheng, Y.L., et al., *Community composition and activity of anaerobic ammonium oxidation bacteria in the rhizosphere of salt-marsh grass Spartina alterniflora.* Applied Microbiology and Biotechnology, 2016. **100**(18): p. 8203-8212.

145. Nie, S.A., et al., *Nitrogen loss by anaerobic oxidation of ammonium in rice rhizosphere.* Isme Journal, 2015. **9**(9): p. 2059-2067.

146. Green, D.S., B. Boots, and T.P. Crowe, *Effects of Non-Indigenous Oysters on Microbial Diversity and Ecosystem Functioning*. Plos One, 2012. **7**(10).

147. Shannon, J.M., et al., *Partial nitritation ANAMMOX in submerged attached growth bioreactors with smart aeration at 20 [degree]C.* Environmental Science: Processes & Impacts, 2015. **17**(1): p. 81-89.

148. Yang, J., et al., *Intermittent aeration in one-stage partial nitritation/anammox process.* Ecological Engineering, 2015. **75**: p. 413-420.

149. Hatzenpichler, R., *Diversity, Physiology, and Niche Differentiation of Ammonia-Oxidizing Archaea.* Applied and Environmental Microbiology, 2012. **78**(21): p. 7501-7510.

150. Nakagawa, T. and R. Takahashi, *Nitrosomonas stercoris sp. nov., a Chemoautotrophic Ammonia-Oxidizing Bacterium Tolerant of High Ammonium Isolated from Composted Cattle Manure.* Microbes and Environments, 2015. **30**(3): p. 221-227.

151. Pester, M., C. Schleper, and M. Wagner, *The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology.* Current Opinion in Microbiology, 2011. **14**(3): p. 300-306.

152. Jia, Z.J. and R. Conrad, *Bacteria rather than Archaea dominate microbial ammonia oxidation in an agricultural soil.* Environmental Microbiology, 2009. **11**(7): p. 1658-1671.

153. Di, H.J., et al., *Nitrification driven by bacteria and not archaea in nitrogen-rich grassland soils.* Nature Geosci, 2009. **2**(9): p. 621-624.

154. Zhalnina, K.V., et al., *Genome Sequence of Candidatus Nitrososphaera evergladensis from Group I.1b Enriched from Everglades Soil Reveals Novel Genomic Features of the Ammonia-Oxidizing Archaea.* PLOS ONE, 2014. **9**(7): p. e101648.

155. Nowka, B., H. Daims, and E. Spieck, *Comparison of Oxidation Kinetics of Nitrite-Oxidizing Bacteria: Nitrite Availability as a Key Factor in Niche Differentiation.* Applied and Environmental Microbiology, 2015. **81**(2): p. 745-753.

156. Maixner, F., et al., *Nitrite concentration influences the population structure of Nitrospira-like bacteria.* Environmental Microbiology, 2006. **8**(8): p. 1487-1495.

157. Schramm, A., et al., *Microenvironments and distribution of nitrifying bacteria in a membrane-bound biofilm.* Environmental Microbiology, 2000. **2**(6): p. 680-686.

158. Lam, P., et al., *Revising the nitrogen cycle in the Peruvian oxygen minimum zone.* Proceedings of the National Academy of Sciences, 2009. **106**(12): p. 4752-4757.

159. Pinto, A.J., et al., *Metagenomic Evidence for the Presence of Comammox Nitrospira-Like Bacteria in a Drinking Water System.* mSphere, 2016. **1**(1): p. e00054-15.

160. Chao, Y., et al., *Novel nitrifiers and comammox in a full-scale hybrid biofilm and activated sludge reactor revealed by metagenomic approach.* Applied Microbiology and Biotechnology, 2016. **100**(18): p. 8225-8237.

161. Zhang, R.J., B. Cui, and S.B. Huang, *Algae Consumption and Nitrate Removal in a Raw Water Transport System by Limnoperna fortunei and its Associated Microorganisms.* Water Environment Research, 2014. **86**(12): p. 2301-2308.

162. Hu, S., et al., *A laboratory investigation of interactions between denitrifying anaerobic methane oxidation (DAMO) and anammox processes in anoxic environments.* Scientific Reports, 2015. **5**: p. 8706.

163. Zhu, B., et al., *Anaerobic Oxidization of Methane in a Minerotrophic Peatland: Enrichment of Nitrite-Dependent Methane-Oxidizing Bacteria.* Applied and Environmental Microbiology, 2012. **78**(24): p. 8657-8665.

164. Winkler, M.K.H., et al., *Modelling simultaneous anaerobic methane and ammonium removal in a granular sludge reactor.* Water Research, 2015. **73**: p. 323-331.

165. van Bodegom, P., et al., *Methane oxidation and the competition for oxygen in the rice rhizosphere.* Applied and Environmental Microbiology, 2001. **67**(8): p. 3586-3597.

166. Ishii, S., et al., *Isolation of Oligotrophic Denitrifiers Carrying Previously Uncharacterized Functional Gene Sequences.* Applied and Environmental Microbiology, 2011. **77**(1): p. 338-342.

167. Sekiguchi, Y., et al., *Thermodesulfovibrio aggregans sp nov and Thermodesulfovibrio thiophilus sp nov., anaerobic, thermophilic, sulfate-reducing bacteria isolated from thermophilic methanogenic sludge, and emended description of the genus Thermodesulfovibrio.* International Journal of Systematic and Evolutionary Microbiology, 2008. **58**: p. 2541-2548.

168. Kirchman, D.L., *Processes in Microbial Ecology*. Processes in Microbial Ecology. 2012. 1-312.

169. He, S.M., et al., *Patterns in Wetland Microbial Community Composition and Functional Gene Repertoire Associated with Methane Emissions.* Mbio, 2015. **6**(3).

170. McKindsey, C.W., et al., *Influence of suspended and off-bottom mussel culture on the sea bottom and benthic habitats: a review*. Canadian Journal of Zoology-Revue Canadienne De Zoologie, 2011. **89**(7): p. 622-646.

171. Fdz-Polanco, F., et al., *New process for simultaneous removal of nitrogen and sulphur under anaerobic conditions.* Water Research, 2001. **35**(4): p. 1111-1114.

172. Strous, M., et al., *The anammox case - A new experimental manifesto for microbiological eco-physiology.* Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology, 2002. **81**(1-4): p. 693-702.

173. Schrum, H.N., et al., *Sulfate-reducing ammonium oxidation: A thermodynamically feasible metabolic pathway in subseafloor sediment.* Geology, 2009. **37**(10): p. 939-942.

174. Ali, M., et al., *The Increasing Interest of ANAMMOX Research in China: Bacteria, Process Development, and Application.* Biomed Research International, 2013.

175. Rios-Del Toro, E.E. and F.J. Cervantes, *Coupling between anammox and autotrophic denitrification for simultaneous removal of ammonium and sulfide by enriched marine sediments.* Biodegradation, 2016. **27**(2-3): p. 107-118.

176. Cai, J., J.X. Jiang, and P. Zheng, *Isolation and identification of bacteria responsible for simultaneous anaerobic ammonium and sulfate removal.* Science China-Chemistry, 2010. **53**(3): p. 645-650.

177. Rikmann, E., et al., *Sulfate-reducing anammox for sulfate and nitrogen containing wastewaters.* Desalination and Water Treatment, 2016. **57**(7): p. 3132-3141.

178. Mahmoudi, N., et al., *Microbial community composition and diversity in Caspian Sea sediments.* Fems Microbiology Ecology, 2015. **91**(1).

179. Arribas, L.P., et al., *Intertidal mussels as ecosystem engineers: their associated invertebrate biodiversity under contrasting wave exposures.* Marine Biodiversity, 2014. **44**(2): p. 203-211.

180. Borthagaray, A.I. and A. Carranza, *Mussels as ecosystem engineers: Their contribution to species richness in a rocky littoral community.* Acta Oecologica-International Journal of Ecology, 2007. **31**(3): p. 243-250.

181. Hartmann, J.T., et al., *Establishing mussel behavior as a biomarker in ecotoxicology.* Aquatic Toxicology, 2016. **170**: p. 279-288.

182. Lopes-Lima, M., et al., *Biology and conservation of freshwater bivalves: past, present and future perspectives.* Hydrobiologia, 2014. **735**(1): p. 1-13.

183. Chowdhury, G.W., A. Zieritz, and D.C. Aldridge, *Ecosystem engineering by mussels supports biodiversity and water clarity in a heavily polluted lake in Dhaka, Bangladesh.* Freshwater Science, 2016. **35**(1): p. 188-199.

184. Atkinson, C.L., et al., *Aggregated filter-feeding consumers alter nutrient limitation: consequences for ecosystem and community dynamics.* Ecology, 2013. **94**(6): p. 1359-1369.

185. Allen, D.C., et al., *Bottom-up biodiversity effects increase resource subsidy flux between ecosystems.* Ecology, 2012. **93**(10): p. 2165-2174.

186. Kroeze, C., A. Mosier, and L. Bouwman, *Closing the global N2O budget: A retrospective analysis 1500-1994.* Global Biogeochemical Cycles, 1999. **13**(1): p. 1-8.

187. Rockstrom, J., et al., *Planetary Boundaries: Exploring the Safe Operating Space for Humanity.* Ecology and Society, 2009. **14**(2): p. 33.

188. Ollivier, J., et al., *Nitrogen turnover in soil and global change.* FEMS Microbiology Ecology, 2011. **78**(1): p. 3-16.

189. Glibert, P.M., et al., *The Haber Bosch-harmful algal bloom (HB-HAB) link.* Environmental Research Letters, 2014. **9**(10): p. 13.

190. Sebilo, M., et al., *Long-term fate of nitrate fertilizer in agricultural soils.* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(45): p. 18185-18189.

191. NRC, *Annex 4, Nitrogen in agroecosystems*, in *A Framework for Assessing Effects of the Food System*, M.C.O. Nesheim, M.; Yih, P.T., Editor. 2015, National Academies Press (US): Washington DC, USA.

192. Burkart, M.R. and D.E. James, *Agricultural-nitrogen contributions to hypoxia in the Gulf of Mexico.* Journal of Environmental Quality, 1999. **28**(3): p. 850-859.

193. Lerch, R.N., et al., *Long-Term Agroecosystem Research in the Central Mississippi River Basin: Goodwater Creek Experimental Watershed and Regional Nutrient Water Quality Data.* Journal of Environmental Quality, 2015. **44**(1): p. 37-43.

194. Donner, S.D. and C.J. Kucharik, *Corn-based ethanol production compromises goal of reducing nitrogen export by the Mississippi River.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(11): p. 4513-4518.

195. Kreiling, R.M., et al., *Summer nitrate uptake and denitrification in an upper Mississippi River backwater lake: the role of rooted aquatic vegetation.* Biogeochemistry, 2011. **104**(1): p. 309-324.

196. Barrios, E., *Soil biota, ecosystem services and land productivity.* Ecological Economics, 2007. **64**(2): p. 269-285.

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

198. Boyd, E.S. and J.W. Peters, *New insights into the evolutionary history of biological nitrogen fixation.* Frontiers in Microbiology, 2013. **4**: p. 201.

199. Daims, H., et al., *Complete nitrification by Nitrospira bacteria.* Nature, 2015. **528**(7583): p. 504-509.

200. van Kessel, M.A.H.J., et al., *Complete nitrification by a single microorganism.* Nature, 2015. **528**: p. 555.

201. Hayatsu, M., K. Tago, and M. Saito, *Various players in the nitrogen cycle: Diversity and functions of the microorganisms involved in nitrification and denitrification.* Soil Science & Plant Nutrition, 2008. **54**(1): p. 33-45.

202. Zhang, W., et al., *Ecosystem services and dis-services to agriculture.* Ecological Economics, 2007. **64**(2): p. 253-260.

203. Black, E.M., M.S. Chimenti, and C.L. Just, *Effect of freshwater mussels on the vertical distribution of anaerobic ammonia oxidizers and other nitrogen-transforming microorganisms in upper Mississippi river sediment.* PeerJ, 2017. **5**: p. e3536.

204. Abubucker, S., et al., *Metabolic Reconstruction for Metagenomic Data and Its Application to the Human Microbiome.* PLOS Computational Biology, 2012. **8**(6): p. e1002358.

205. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Meth, 2012. **9**(4): p. 357-359.

206. Buchfink, B., C. Xie, and D.H. Huson, *Fast and sensitive protein alignment using DIAMOND.* Nat Meth, 2015. **12**(1): p. 59-60.

207. Li, W. and A. Godzik, *Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences.* Bioinformatics, 2006. **22**(13): p. 1658-1659.

208. Caspi, R., et al., *The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases.* Nucleic Acids Research, 2016. **44**(D1): p. D471-D480.

209. Segata, N., et al., *Metagenomic biomarker discovery and explanation.* Genome Biology, 2011. **12**(6): p. R60.

210. Kumar, S., G. Stecher, and K. Tamura, *MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.* Molecular Biology and Evolution, 2016. **33**(7): p. 1870-1874.

211. Edgar, R.C., *MUSCLE: multiple sequence alignment with high accuracy and high throughput*. Nucleic Acids Research, 2004. **32**(5): p. 1792-1797.

212. Waterhouse, A.M., et al., *Jalview Version 2—a multiple sequence alignment editor and analysis workbench.* Bioinformatics, 2009. **25**(9): p. 1189-1191.

213. Felsenstein, J., *Confidence-limits on Phylogenies- An Approach Using Bootstrap*. Evolution, 1985. **39**(4): p. 783-791.

214. Kreiling, R.M. and J.N. Houser, *Long-term decreases in phosphorus and suspended solids, but not nitrogen, in six upper Mississippi River tributaries, 1991-2014.* Environmental Monitoring and Assessment, 2016. **188**(8): p. 19.

215. David, M.B., et al., *Navigating the Socio-Bio-Geo-Chemistry and Engineering of Nitrogen Management in Two Illinois Tile-Drained Watersheds.* Journal of Environmental Quality, 2015. **44**(2): p. 368-381.

216. Garcia, A.M., et al., *Regional Effects of Agricultural Conservation Practices on Nutrient Transport in the Upper Mississippi River Basin.* Environmental Science & Technology, 2016. **50**(13): p. 6991-7000.

217. Hoellein, T.J., et al., *Contributions of freshwater mussels (Unionidae) to nutrient cycling in an urban river: filtration, recycling, storage, and removal.* Biogeochemistry, 2017. **135**(3): p. 307-324.

218. Zhang, X., et al., *Anthropogenic activities drive the microbial community and its function in urban river sediment*. Journal of Soils and Sediments, 2016. **16**(2): p. 716-725.

219. Wang, B.Z., et al., *Differential contributions of ammonia oxidizers and nitrite oxidizers to nitrification in four paddy soils.* ISME Journal, 2015. **9**(5): p. 1062-1075.

220. Lücker, S., et al., *A Nitrospira metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria*. Proceedings of the National Academy of Sciences, 2010. **107**(30): p. 13479-13484.

221. Altmann, D., et al., *In situ distribution and activity of nitrifying bacteria in freshwater sediment*. Environmental Microbiology, 2003. **5**(9): p. 798-803.

222. Zheng, X.F., et al., *Bacterial composition, abundance and diversity in fish polyculture and mussel-fish integrated cultured ponds in China*. Aquaculture Research, 2017. **48**(7): p. 3950-3961.

223. Shen, H., et al., *Response of the microbial community to bioturbation by benthic macrofauna on intertidal flats*. Journal of Experimental Marine Biology and Ecology, 2017. **488**: p. 44-51.

224. Altmann, D., et al., *Distribution and activity of nitrifying bacteria in natural stream sediment versus laboratory sediment microcosms*. Aquatic Microbial Ecology, 2004. **36**(1): p. 73-81.

225. Jäntti, H. and S. Hietanen, *The Effects of Hypoxia on Sediment Nitrogen Cycling in the Baltic Sea*. AMBIO, 2012. **41**(2): p. 161-169.

226. Lindemann, S., et al., *Effect of Eastern Oysters (Crassostrea virginica) and Seasonality on Nitrite Reductase Gene Abundance (nirS, nirK, nrfA) in an Urban Estuary*. Estuaries and Coasts, 2016. **39**(1): p. 218-232.

227. Dalsgaard, T., et al., *Oxygen at Nanomolar Levels Reversibly Suppresses Process Rates and Gene Expression in Anammox and Denitrification in the Oxygen Minimum Zone off Northern Chile*. mBio, 2014. **5**(6).

228. Zhang, X., et al., *Quantifying nitrous oxide fluxes on multiple spatial scales in the Upper Midwest, USA*. International Journal of Biometeorology, 2015. **59**(3): p. 299-310.

229. IPCC, *Climate Change 2007: the Physical Science Basis. Working Group I Report*, D.Q. S. Solomon, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, H.L. Miller (Eds.), Editor. 2007: New York, NY. p. 996.

230. Turner, P.A., et al., *Regional-scale controls on dissolved nitrous oxide in the Upper Mississippi River*. Geophysical Research Letters, 2016. **43**(9): p. 4400-4407.

231. Hallin, S., et al., *Relationship between N-cycling communities and ecosystem functioning in a 50-year-old fertilization experiment*. Isme Journal, 2009. **3**(5): p. 597-605.

232. Shange, R.S., et al., *Distinct Soil Bacterial Communities Revealed under a Diversely Managed Agroecosystem*. Plos One, 2012. **7**(7): p. 11.

233. Mulholland, P.J., et al., *Stream denitrification across biomes and its response to anthropogenic nitrate loading.* Nature, 2008. **452**: p. 202.

234. Gomez-Velez, J.D., et al., *Denitrification in the Mississippi River network controlled by flow through river bedforms.* Nature Geoscience, 2015. **8**: p. 941.

235. Strauss, E.A., et al., *Variability and regulation of denitrification in an Upper Mississippi River backwater.* Journal of the North American Benthological Society, 2006. **25**(3): p. 596-606.

236. Atkinson, C.L. and C.C. Vaughn, *Biogeochemical hotspots: temporal and spatial scaling of the impact of freshwater mussels on ecosystem function.* Freshwater Biology, 2015. **60**(3): p. 563-574.

237. Daebeler, A., et al., *Interactions between Thaumarchaea, Nitrospira and methanotrophs modulate autotrophic nitrification in volcanic grassland soil.* The Isme Journal, 2014. **8**: p. 2397.

238. Matantseva, O.V. and S.O. Skarlato, *Mixotrophy in microorganisms: Ecological and cytophysiological aspects.* Journal of Evolutionary Biochemistry and Physiology, 2013. **49**(4): p. 377-388.

239. Camejo, P.Y., et al., *Genome-Enabled Insights into the Ecophysiology of the Comammox Bacterium "Candidatus Nitrospira nitrosa".* mSystems, 2017. **2**(5): p. e00059-17.

240. Young, N.M., Andrew; Weber, Larry, *Hydrodynamic Investigation of Upper Mississippi River Freshwater Mussel Habitats*, in *Impacts of Global Climate Change*. 2005.

241. USACE, *Survey of freshwater mussels (Pelecypoda: Unionidae) at selected sites in Pools 11 through 24 of the Mississippi River*, in *EA Report 9031 Rock Island District, US Army Corps of Engineers, Rock Island, Illinois*. 1981, US Army Corps of Engineers.

242. USACE, *Resources inventory for the Upper Mississippi River Rock Island District, US Army Corps of Engineers, Rock Island, Illinois*. 1984, US Army Corps of Engineers.

243. Mirto, S., et al., *Microbial and Meiofaunal Response to Intensive Mussel-Farm Biodeposition in Coastal Sediments of the Western Mediterranean.* Marine Pollution Bulletin, 2000. **40**(3): p. 244-252.

244. Danovaro, R., et al., *Sustainable impact of mussel farming in the Adriatic Sea (Mediterranean Sea): evidence from biochemical, microbial and meiofaunal indicators.* Marine Pollution Bulletin, 2004. **49**(4): p. 325-333.

245. Andrews, S., *FastQC: a quality control tool for high throughput sequence data*. 2010: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

246. Huson, D.H., et al., *MEGAN analysis of metagenomic data.* Genome Research, 2007. **17**(3): p. 377-386.

247. Huson, D.H., et al., *MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data.* Plos Computational Biology, 2016. **12**(6).

248. O'Leary, N.A., et al., *Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation.* Nucleic Acids Research, 2016. **44**(D1): p. D733-D745.

249. Overbeek, R., et al., *The Subsystems Approach to Genome Annotation and its Use in the Project to Annotate 1000 Genomes.* Nucleic Acids Research, 2005. **33**(17): p. 5691-5702.

250. Okonechnikov, K., O. Golosova, and M. Fursov, *Unipro UGENE: a unified bioinformatics toolkit*. Bioinformatics, 2012. **28**(8): p. 1166-1167.

251. Carver, T., et al., *DNAPlotter: circular and linear interactive genome visualization.* Bioinformatics, 2009. **25**(1): p. 119-120.

252. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows–Wheeler transform*. Bioinformatics, 2009. **25**(14): p. 1754-1760.

253. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-2079.

254. Pfister, C.A., J.A. Gilbert, and S.M. Gibbons, *The role of macrobiota in structuring microbial communities along rocky shores.* PeerJ, 2014. **2**: p. e631.

255. McGill, B.J., et al., *Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework*. Ecology Letters, 2007. **10**(10): p. 995-1015.

256. Ferguson, G.P., D. McLaggan, and I.R. Booth, *Potassium channel activation by glutathione-S-conjugates in Escherichia coli: protection against methylglyoxal is mediated by cytoplasmic acidification.* Molecular Microbiology, 1995. **17**(6): p. 1025-1033.

257. Bott, C.B. and N.G. Love, *Implicating the Glutathione-Gated Potassium Efflux System as a Cause of Electrophile-Induced Activated Sludge Deflocculation.* Applied and Environmental Microbiology, 2004. **70**(9): p. 5569-5578.

258. Tötemeyer, S., et al., *From famine to feast: the role of methylglyoxal production in Escherichia coli.* Molecular Microbiology, 1998. **27**(3): p. 553-562.

259. Irzik, K., et al., *The KdpC subunit of the Escherichia coli K+-transporting KdpB P-type ATPase acts as a catalytic chaperone.* FEBS Journal, 2011. **278**(17): p. 3041-3053.

260. Cardenas, J.P., R. Quatrini, and D.S. Holmes, *Aerobic Lineage of the Oxidative Stress Response Protein Rubrerythrin Emerged in an Ancient Microaerobic, (Hyper)Thermophilic Environment*. Frontiers in Microbiology, 2016. **7**: p. 1822.

261. Zhang, Y., J. Lin, and Y. Gao, *In silico identification of a multi-functional regulatory protein involved in Holliday junction resolution in bacteria.* BMC Systems Biology, 2012. **6**(1): p. S20.

262. Teplyakov, A., et al., *Crystal Structure of the YchF Protein Reveals Binding Sites for GTP and Nucleic Acid.* Journal of Bacteriology, 2003. **185**(14): p. 4031-4037.

263. Zhou, X.L., et al., *Translational Quality Control by Bacterial Threonyl-tRNA Synthetases.* Journal of Biological Chemistry, 2016. **291**(40): p. 21208-21221.

264. Hershey, J.W.B., N. Sonenberg, and M.B. Mathews, *Principles of Translational Control: An Overview*. Cold Spring Harbor Perspectives in Biology, 2012. **4**(12).

265. Simonin, M., et al., *Coupling Between and Among Ammonia Oxidizers and Nitrite Oxidizers in Grassland Mesocosms Submitted to Elevated CO2 and Nitrogen Supply*. Microbial Ecology, 2015. **70**(3): p. 809-818.

266. Rasmussen, J.J., et al., *Campylobacter jejuni carbon starvation protein A (CstA) is involved in peptide utilization, motility and agglutination, and has a role in stimulation of dendritic cells.* Journal of Medical Microbiology, 2013. **62**: p. 1135-1143.

267. Dubey, A.K., et al., *CsrA regulates translation of the Escherichia coli carbon starvation gene, cstA, by blocking ribosome access to the cstA transcript.* Journal of Bacteriology, 2003. **185**(15): p. 4450-4460.

268. Vastermark, A., et al., *Expansion of the APC superfamily of secondary carriers.* Proteins-Structure Function and Bioinformatics, 2014. **82**(10): p. 2797-2811.

269. Ferguson, G.P., et al., *Methylglyoxal production in bacteria: suicide or survival?* Archives of Microbiology, 1998. **170**(4): p. 209-219.

270. Booth, I., *Glycerol and Methylglyoxal Metabolism*. EcoSal Plus, 2005.

271. Jansen, H.M., et al., *Seasonal variation in mineralization rates (C-N-P-Si) of mussel Mytilus edulis biodeposits.* Marine Biology, 2012. **159**(7): p. 1567-1580.

272. Anes, J., et al., *The ins and outs of RND efflux pumps in Escherichia coli.* Frontiers in Microbiology, 2015. **6**(587).

273. Lin, J., L. Overbye Michel, and Q. Zhang, *CmeABC Functions as a Multidrug Efflux System in Campylobacter jejuni.* Antimicrobial Agents and Chemotherapy, 2002. **46**(7): p. 2124-2131.

274. Liu, M., et al., *Bacterial addiction module toxin Doc inhibits translation elongation through its association with the 30S ribosomal subunit*. Proceedings of the National Academy of Sciences, 2008. **105**(15): p. 5885-5890.

275. Liu, J., J.H. Jung, and Y.H. Liu, *Antimicrobial Compounds from Marine Invertebrates-Derived Microorganisms.* Current Medicinal Chemistry, 2016. **23**(25): p. 2892-2905.

276. Cooke, M.D., *Antibiotic Resistance Among Coliform and Fecal Coliform Bacteria Isolated from the Freshwater Mussel Hydridella menziesii.* Antimicrobial Agents and Chemotherapy, 1976. **9**(6): p. 885-888.

277. da Silveira, C.S., O.V. de Sousa, and N.S. Evangelista-Barreto, *Propogation of Antimicrobial Resistant Salmonella spp. in Bivalve mollusks from estuary areas of Bahia, Brazil.* Revista Caatinga, 2016. **29**(2): p. 450-457.

278. Grevskott, D.H., et al., *Marine Bivalve Mollusks As Possible Indicators of Multidrug-Resistant Escherichia coli and Other Species of the Enterobacteriaceae Family*. Frontiers in Microbiology, 2017. **8**: p. 10.

279. Rzymski, P., et al., *Bioaccumulation of selected metals in bivalves (Unionidae) and Phragmites australis inhabiting a municipal water reservoir*. Environmental Monitoring and Assessment, 2014. **186**(5): p. 3199-3212.

280. Naimo, T.J., *A review of the effects of heavy metals on freshwater mussels.* Ecotoxicology, 1995. **4**(6): p. 341-362.

281. Jorge, M.B., et al., *Copper uptake, patterns of bioaccumulation and effects in glochida (larvae) of the freshwater mussel (Lampsilis cardium).* Environmental Toxicology and Chemistry: p. n/a-n/a.

282. Palomo, A., et al., *Comparative Genomics Sheds Light On Niche Differentiation And The Evolutionary History Of Comammox Nitrospira*. bioRxiv, 2017.

283. Hu, H.-W. and J.-Z. He, *Comammox—a newly discovered nitrification process in the terrestrial nitrogen cycle.* Journal of Soils and Sediments, 2017. **17**(12): p. 2709-2717.