Metagenomic Profiling of Nitrogen Cycling Potential in Caribbean Sponges

By

Luis Zuniga Acuna

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Biology

Middle Tennessee State University

April, 2024

Thesis Committee:

Cole Easson, Chair

Jessica Arbour

Cara Fiore

ABSTRACT

The high species diversity of coral reefs is in part possible because of the efficient recycling of essential nutrients facilitated by resident organisms and their microbiomes. Sponges and their symbionts play a role in the recycling of nutrients and help ensure productive energy flow, allowing coral reefs to thrive. Here we present a metagenomic analysis of common and abundant sponges from the Caribbean, where microbial metagenome assembled genomes (MAGs) were assembled from shotgun sequencing data and were annotated for metabolic function. We focused on the completeness of nitrogen cycling pathways and the presence of genes from these pathways because bioavailability of nitrogen compounds can be essential in maintaining high biodiversity on coral reefs. We recovered high-quality genomes that spanned twenty-five unique taxa, with eleven of them contributing to the nitrogen cycling potential of the sponge holobiont. Insights on how common reef animals such as sponges interact with biologically important compounds may be a key component in understanding how coral reefs will continue to change in the coming decades.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICIES

ABBREVIATIONS

ANAMMOX – anerobic ammonia oxidation

ANR – assimilatory nitrate reduction

DNR – dissimilatory nitrate reduction

MAG – metagenome assembled genome

NGS – next-generation sequencing

nMDS – non-metric multidimensional scaling

KO – Kegg ortholog

mRNA – messenger RNA

Background & Introduction

Coral Reefs as a System

Coral reefs are regarded as some of the Earth's most biologically diverse and endangered ecosystems (Knowlton et al. 2010, Hughes 1994). Despite being productive geological and biological systems, they are sometimes located in low-nutrient environments. This was pointed out by Charles Darwin (Darwin 1842), when he noted that the conditions necessary for a coral reef system to thrive are quite particular. For example, many reef building corals inhabit a relatively narrow temperature range of 23ºC-29ºC, although some can withstand slightly lower or higher temperatures (Guan et al 2015). These observations have led scientists to speculate on how these biodiverse systems formed, and the underlying mechanisms that maintain them. Given the complexity of coral reef systems, one may conclude that they are held at a stable equilibrium mediated by competition between co-existing species and tight niche differentiation (Knowlton and Jackson 1994). However, Connell (1978) posits the "intermediate disturbance hypothesis", which states that high diversity in tropical landscapes emerges from disturbances of varying degrees (windstorms, insect plagues, wave storms, freshwater floods). These intermediate disturbances allow for the periodic rearrangement of species structure because they often free up important resources such as space and cover. Differential recruitment based on efficient exploitation of resources can then allow for increased diversity over time. Of course, this does not mean that "equilibrium" and "non-equilibrium" mechanisms are mutually exclusive. Reef diversity can be maintained by niche differentiation and/or intermediate disturbance, but Connell (1978) suggests intermediate disturbance may be more likely, because there may be too many species with similar strategies to feasibly partition resources at the fine scales that would be

required. Given the increased threats coral reefs have been faced with in the last several decades, it is important to consider the general mechanisms that allow for the upkeep of biodiversity.

Coral reefs are formed by the polyps of reef-building stony corals (Sheppard 2018). As the CO_2 exhaled by a coral polyp is dissolved into the water, it can take form of carbonate ions (CO_3^{2-}). CO_3^{2-} can then react with nearby Ca^{2+} ions to form solid calcium carbonate ($CaCO_3$). Over geological time, and with many individual participating, the $CaCO_3$ excreted by those corals has created the coral reef ecosystems of today (Helman et al. 2008). While they are the namesake of these revered environments, other organisms also play role in the physical building of coral reefs. Marine sponges are known to help act as a "glue" that can assist in adhering corals to their substratum (Wulff 2013). Other marine calcifiers such as mollusks can also contribute the net calcification of reef systems. Indeed, the combined effort of all these organisms over millions of years should be cited for how we observe coral reefs today (Pandolfi 2011).

Coral reefs not only serve as habitat space for a diverse community of flora and fauna but also provide protection for shores from wave action (Harris et al. 2018). In some cases, coastal communities face a considerable increased risk of wave damage as reefs become less prominent (van Zaten et al 2014). Coral reefs also provide other goods and services to humans. Some of these goods and services are food for coastal populations, environmental services such as waste assimilation, or biogeochemical services such as nitrogen fixation (Cesar 2002). Coral reefs also serve economic purposes for those areas where they are located, bringing in revenue from tourism and fisheries (Sheppard 2018). For example, coral reefs bring in millions of dollars for over 100 jurisdictions in the world for their recreational and aesthetic value (Spalding et al. 2017). While fisheries and an excess of tourism can negatively impact reefs, these factors may

also be incentives to support sustainable management and conservation, given their economic value to these coastal areas.

Despite their environmental and economic importance, coral reefs have steadily been degrading over the past several decades (Pandolfi et al 2003). This is a complex consequence of anthropogenic impacts from pollution and over-fishing, and of global influences such as the increase of the average temperature of the Earth due to climate change (Hughes 1994, Hoegh-Guldernberg 1999). The CO_2 generated by human activity has not only made the planet's oceans warmer but has also made them more acidic (Hönisch et al 2012). Given the narrow pH range that allows for life, a decrease of the ocean's pH can be detrimental to much of marine life. These and many other factors combined have led to several large-scale bleaching events of many of Earth's reefs (Brown 1997, Douglas 2003), killing off and endangering many species of corals and affecting reef dynamics on a larger scale. Coral reefs aren't the only marine systems that are changing, but the Earth's Ocean in general face several different challenges. These changes include decreased ocean productivity, altered food-web dynamics, reduced abundance of habitat forming species, a shift in species distributions, and a greater incidence of disease (Hoegh-Guldberg & Bruno 2010). Given the magnitude and complexity of these issues, there is still uncertainty about the spatial and temporal scales of these changes. However, there are widespread efforts to factor the changing conditions of Earth into models of the ocean (Frazão Santos et al. 2020, Winton et al. 2013).

Coral bleaching, or the die-offs of corals due to a loss of their photosynthetic symbionts is often the mark of an "unhealthy" reef. This is a phenomenon where external stressors (temperature, pH) make a coral host unsuitable for its symbionts, causing a mass expulsion of the photosynthetic algae from their hosts, in turn leading to the degradation of corals since many of

them are dependent on the symbiotic relationship (Lesser 1997, Brown & Ogden 1993). Once large amounts of corals begin to die, biodiversity suffers, as corals are an immense part of coral reef architecture that many other animals use for shelter. Overall reef productivity also decreases since coral-algal symbioses contribute to much of a reef's primary production (Hatcher 1990). While the dynamics of coral reefs are altered because of anthropogenic impacts, climate change, and mass community die-offs, the exact fate of reef systems is still not known. It is important to remember that these marine systems change over geological time, and the studies of the last ~100 years may not necessarily be enough to completely understand the bigger picture of coral reefs (Hughes & Connell 1994). Still, efforts are being made to predict the future of coral reef systems.

In some cases, "phase-shifts" from coral-dominated communities to algal-dominated communities have been observed and are of concern (McManus & Polsenberg 2004). A phase-shift is the idea that biotic and abiotic changes in a system can cause a shift in the abundances and composition of species (such as a coral reef being overtaken by algae), but with the caveat that if conditions return to "normal", species composition and abundances would also return to "normal" (Dudgeon et al. 2010). Along with phase-shifts comes the concept of an alternative stable state, which is when a single set of environmental conditions supports more than one community state, where there is presumably some extrinsic factor that causes the shift from one state to another. Dudgeon et al. (2010) suggests that many documented algae-dominated stable-states are phase-shifts where coral-dominated communities can recover given enough time or alleviation of some stressor. For example, Discovery Bay, Jamaica experienced several coral bleaching events in conjunction with a loss of the herbivorous sea urchin *Diadema antillarum* which led to macroalgae becoming the dominant group in the 1980's. *D. antillarum* was critical

in keeping algal cover to low because of its grazing of algae. Eventually, in the 1990's, *D. antillarum* populations recovered, and benthic communities returned to a coral-dominated state by the mid 2000's. While some concluded that the algae-dominated community was a new stable-state, it showed to be a "phase-shift" once conditions returned to "normal". Of course, other vital reef community members have suffered die-offs in the last few decades, such as the poriferans. Two mass sponge die-off events were recorded in 2008 and 2009 in the Mediterranean, where sponges who hosted photosynthesizing symbionts seemed to be more drastically affected in contrast with species who hosted few or no photosynthesizing microbes (Cebrian et al 2011). This disproportionate die-off of sponges suggested rising sea temperatures were a culprit, much like the case of algae expulsion by corals. Another case of sponge-die offs related to recurring picoplankton blooms was recorded between 1991 and 2006 in the Florida Keys (Stevely et al. 2011). This 15-year long monitoring concluded that sponge communities can eventually recover from these events, albeit very slowly. Some species recovered at a consistent rate, and others had continually changed in abundance spatially and temporally throughout the study period. Although sponge communities can suffer similarly to coral communities, they may be more resistant to ocean acidification and rising sea temperatures, with some studies hypothesizing sponge-dominated communities emerging in the future (Bell et al. 2013).

This and other studies highlight the complexity of the issues coral reefs face, and the necessity for long-term studies to perhaps predict the trajectories coral reefs may follow. While the outlook of many of Earth's coral reefs may look grim, it is increasingly important to further our understanding of these dynamic ecosystems as they continue to change in the coming decades.

Nutrient Cycling & Energy Movement

A reef's exceptional diversity and productivity are partly maintained through efficient recycling of nutrients, where the existing cycles effectively keep nutrients within the reef system (Atkinson & Falter 2003). These biogeochemical fluxes are often facilitated by different organisms, with microorganisms mediating a large part of critical chemical reactions (Falkowski et al. 2008, Madsen 2011). For example, some of these nutrient transformations involve the reduction, oxidation, and remineralization of nitrogen compounds, which are processes that are part of the greater global nitrogen cycle (Gruber 2008, Shlesinger & Bernhardt 2013). Nitrogen fixation and denitrification are of particular interest given the biological importance of nitrogen-containing biomolecules and limited bioavailability of nitrogen on reefs (Wiebe et al. 1975, Larkum et al. 1988, Fiore et al. 2010). Indeed, a similar story can be told for the other major nutrients in marine systems: carbon, phosphorus, and sulfur compounds (Shlesinger and Bernhardt 2013).

In addition to driving biogeochemical cycles, microorganisms are also able to transfer dissolved organic matter (DOM) to higher trophic levels through the hypothesized "microbial loop" (Azam et al. 1983, Fenchel 2008). While microorganisms can underpin these networks of energy movement, larger reef residents can also play a role. Corals are hypothesized to contribute to the cycling of organic matter in reefs through their release of coral mucus (Wild et al. 2004). As coral mucus is released, it can trap smaller particles and create a denser aggerate, which could then be consumed by filter feeders (Bythell & Wild 2011). Similarly, cryptic reef sponges are posited to return particulate and dissolved organic matter (POM and DOM) in the form of detritus to other reef dwellers through the "sponge loop" (de Goeij et al. 2013). The sponge loop proposes that given the large volumes of water sponges filter every day, from which

they are presumably filtering particles for biomass, many sponges might be much larger. Stable isotope experiments and microscopic analysis revealed that sponges are constantly sloughing off cells as detritus which can be eaten by detritovores (Alexander et al 2014, McMurray et al. 2018). Bacterioplankton also uptake dissolved organic matter (DOM), but their uptake rates did not fully explain DOM removal on Caribbean and Indo-Pacific reefs (de Goeij & van Duyl 2007). This begins to reveal a complex web of interactions between the movement of energy and nutrients, partly mediated by the microbes and animals on coral reefs.

The Nitrogen Cycle

The nitrogen cycle is one of many biogeochemical processes such as the carbon, sulfur, or water cycle. More specifically, the nitrogen cycle is a type of nutrient cycle that outlines how nitrogen becomes bioavailable, is processed by living organisms, and eventually returned to an abiotic compartment, such as the atmosphere. Six major nitrogen cycle pathways are described as contributing to the cycling of nitrogen: nitrification, dentification, nitrogen fixation, assimilatory nitrate reduction, dissimilatory nitrate reduction, and anerobic ammonia-oxidation (ANAMMOX). Each of these pathways can be thought of a set of steps where one chemical species is transformed into another by enzymes that are encoded for by various genes or gene clusters. Nitrogen in its inert gaseous form first becomes bioavailable to living organisms through nitrogen fixation, which reduces diatomic nitrogen into “fixed” ammonia (Figure 1A). Nitrogen fixation can only be carried out by select prokaryotes called diazotrophs. These organisms can be free-living or in symbiosis with eukaryotes. From here, ammonia could be converted into nitrate through nitrification (Figure 1B), or it could react with nitrite in the ANAMMOX pathway to be converted back into inert nitrogen (Figure 1C). Nitrate can be “denitrified” through the denitrification pathway (Figure 1D), another pathway that returns

bioavailable nitrogen to an abiotic compartment. Nitrate can also be reduced back into ammonia through either assimilatory nitrate reduction (ANR) (Figure 1E) or dissimilatory nitrate reduction (DNR) (Figure 1F). In these two reduction processes, the ammonia generated by ANR will be used to assimilation into biomass, and DNR generates ammonia as a cellular waste product. Since nitrogen containing compounds such as nucleic acids and amino acids are essential to life, these processes have been well studied (Fowler et al. 2012). For example, the discovery of ammonia-oxidizing bacteria has led to applications in wastewater treatment, where the bacteria are used to remove ammonia from wastewater (Kartal et al. 2010). From an ecological perspective, these six pathways facilitated by microorganisms who can be in symbiosis with higher organisms provide insights into sources and sinks of nitrogen in an ecosystem.

Figure 1: Schematic of the six nitrogen cycle pathways.

Symbiosis

While it is true that microorganisms are mediating biochemical reactions through their metabolism wherever they are in the ambient water, these reactions are not exclusive to the open ocean. In many cases, a community of microbes can be in symbiotic relationships with a variety of larger organisms (Cavanaugh 1994, Apprill 2017). The idea of symbiosis has been a topic of biological study for well over 100 years but became more relevant in the 1960's when DNA could be more rigorously studied. Symbiotic associations as we know them today can range from the macroscopic to the microscopic. A well-known example of a symbiotic association in the ocean is that of clown fish and sea anemones, where clown fish can take refuge within an anemone's tentacle since they are adapted to be immune to the cnidarian's toxin. In "return" the fish can benefit the anemone by dropping food or feces. This is known as a mutualistic relationship because both organisms presumably benefit from the association. While symbiosis is often thought of as beneficial, it exists on a gradient spanning from mutualistic, commensal, and parasitic. An example of a parasitic association is between cymothoid isopods and fish. These isopods will attach to the skin of fish and feed on its blood, and in some cases the isopod will make its home inside the mouth of a fish, replacing its tongue. Both biological associations are considered symbioses since they are the result of long-term evolutionary processes. The organism pairs need not know they are "helping" or "hurting" one another, only that over time these interactions emerged between species because they would presumably increase the fitness of one or both parties.

On coral reefs, symbiosis has been a relevant topic of study, especially those symbioses between invertebrates and their microbiomes, namely corals and sponges. Given their predominance on reefs, understanding these invertebrates' symbiotic interactions could pave the

way for understanding other animal-microbiome associations. A classic example of symbiosis on coral reefs is the relationship between corals and photosynthetic dinoflagellates (Venn et al. 2008). These symbiotic organisms, or symbionts, are of the family Symbiodinaceae (LaJeunesse et al. 2018) and provide their host with resources such as sugars while the host may provide CO_2 and nitrogen compounds to the algae (Lema et al. 2012, Muller-Parker et al. 2015). Marine sponges are also known to harbor symbiotic microorganisms (Lee et al. 2001, Webster & Taylor 2012), such as cyanobacteria capable of converting elemental nitrogen into biochemically available ammonium (Freeman and Thacker 2011, Freeman et al. 2013, Wilkinson & Fay 1979). These types of symbioses exist widely in marine systems (Apprill 2017), and the presence of a microbial symbiont can sometimes be vital to the success of the host. In many cases, the "combined effort" of a microbial symbiont and its host allows both organisms to exploit their given environmental conditions more efficiently (Fiore et al. 2010, Glasl et al. 2016). Similarly to the macroscopic symbioses' examples, these microbial symbioses have emerged from long-term associations, and their exact origins perhaps cannot be pinpointed. One recurring pattern in microbial symbioses is that there seems to be a "species-specificity" in the microbiomes profiled in individuals of a species, ranging from phytoplankton to non-human primates (Yildirim et al. 2010, Vecchi et al. 2018, Jackrel et al. 2021). This brings up the question of how exactly a specific microbial community is maintained throughout successive generations, which is an ongoing topic of study in microbiome research. Given the recognition of the microbiome in animal and plant health, there has been a recent advocacy for considering the role of the microbiome in speciation through the concept of the holobiont (Bordenstein and Theis 2015). The holobiont concept posits that a host and its microbiome (and all their genes) should be studied together when considering ecological and evolutionary questions. While this makes sense

on the surface, it comes with its own set of challenges, and the holobiont concept is perhaps suited for some questions but not all. For example, one assumption that the holobiont concept seems to take is that hosts have evolved to select for favorable microbiome members. However, as Moran & Solan (2015) point out in response to Bordenstein & Theis (2015), hosts may adapt to reliable symbionts in the same way they would adapt to any other abiotic variable, and the host is not necessarily selecting for specific symbionts over evolutionary time. Given the important role some symbionts play in host health, Bordenstein & Theis (2015) suggest using the holobiont as the primary unit of selection. Moran & Solan (2015) challenge this by reminding the reader that selection can vary greatly at different levels of biological organization, and while in some cases host-symbiont fitness interests may align, it is likely an oversimplification to call the holobiont the primary unit of selection for most systems.

This study will focus on microbial symbioses in common Caribbean sponge species, in the context of how the host-symbiont metabolism could play an ecological role in the cycling of often limiting nitrogen compounds on coral reefs.

Sponges and Nutrient Cycling

Marine sponges are known to host diverse assemblages of microbes (Thacker and Freeman 2012, Taylor and Webster 2012, Easson and Thacker 2014), and are common members of coral reef ecosystems. The poriferans are found all over the world (Van Soest et al. 2012) in freshwater and the ocean, and their ecological success in different environments is a topic of interest in the field of sponge ecology. This has led some to wonder, how is it that sponges have been able to thrive in different aquatic environments all over the Earth? One-way sponges may survive in different geographies could be by access to an environment's resources mediated by members of their microbiomes (Freeman et. al 2021). This expansion of available sources of

energy for the sponge holobiont (the combination of the host and the many microbial species living inside it) may allow it to access different nutritional niches that would not be available otherwise (Zhang et al 2019). On coral reefs, this idea of accessing different pools of nutrients could be especially important since some essential nutrients might be limited in a reef system (Bristow et al. 2017). The interactions that arise between a sponge host's metabolism, and their microbiome's metabolism can be difficult to resolve given the complexity that arises when studying host-symbiont associations (Thomas et al. 2016). Many genomic approaches have been taken in recent years with questions regarding connections between the two holobiont compartments, and their connections to biogeochemical cycles (Li et al. 2016, Engelberts et al. 2020, Karimi et al. 2018). For instance, Maldonado et al. (2012) found that deep-sea glass sponge holobionts in Nova Scotia may significantly affect ammonium and nitrate concentrations through various nitrogen cycle pathways detected with metagenome-assembled genomes (MAGs). A study on Amazon River plume nutrients suggests that sponges with high-microbial diversity appeared to rely on their symbionts for nutrients derived from anaerobic processes and nitrogen cycling metabolites (de Menezes et al. 2022). These studies suggest that the sponge holobiont could perhaps play a role in the nutrient cycling that happens on coral reefs (Zhang et al. 2019). Of the existing biogeochemical cycles where the ocean plays a role, the nitrogen cycle is one of the most widely studied (Zehr & Kudela 2011, Fowler et al 2013). A review of nitrogen transformations in marine symbioses by Fiore et al (2010) summarizes the potential of nitrogen fixation and other pathways present in different host-symbiont systems (corals, sponges, diatoms). In sponges, both molecular techniques and stable-isotope markers (Southwell et al 2008) have been employed to detect genes and molecules that indicate nitrification and anaerobic ammonia-oxidation (ANAMMOX). Freeman et al (2020) used stable-isotope measurements to

suggest that sponge symbionts also can facilitate assimilatory nitrate reduction, which would provide the sponge host with bioavailable ammonia. One way a sponge host could use ammonia is as a precursor in the biosynthesis of the amino acid arginine. This connection between the sponge holobiont and nutrient cycling on coral reefs is one of the motivations of this study.

Study Aims

The objective of this study is to determine whether there is convergence or divergence in nitrogen metabolisms mediated by the microbial communities found within co-occurring sponge species. This was done by assessing the functional potential of our sampled communities through metagenomic sequencing where draft metagenome-assembled genomes (MAGs) and assembled contigs were annotated for genes that contribute to the six nitrogen cycle processes. MAGs can home in on specific taxa who may be driving nitrogen cycle processes, whereas annotated assembled contigs are a more coarse-grained view of functional potential of an entire community rather than a handful of taxa. Since sponge microbiomes are known to be species-specific in their community structure (Easson and Thacker 2014, Reveillaud et al. 2014, Thomas et al. 2016), we expect sponges with more similar microbiomes to host similar metabolic activities. Conversely, we expect to see a divergence in community structure to be associated with differences in metabolic potential in different species. The measures used to determine differences in functional potential between sponge species were abundances of nitrogen cycle genes found in both MAGs and contigs.

Materials & Methods

Metagenomics as a Tool

In this study, metagenomics is synonymous with environmental genomics. This is because metagenomic analyses start from a mixed environmental sample (Wooley et al. 2010), and the sequencing preparation does not target any one marker gene such as 16S and 18S rRNA gene surveys do. Instead, an entire mixed microbial community is sampled and sequenced after a filtering step to select for the community one is interested in. If one is studying the microbiome from an animal, it is likely that separating host from microbe would beneficial before sequencing. With this data, one can carry out different types of analyses, and in this case the goal is to recover metagenome assembled genomes (MAG's) and assign functional potential to the captured microbial communities.

One of the main advantages of using metagenomics is that one can obtain information from microorganisms that cannot be cultured in a laboratory. This can be more relevant when studying environments such as the ocean, freshwater, and soil systems (Biers et al. 2009, Edge et al. 2020). In contrast to analyses of well-established marker genes, metagenomics attempts to look at entire genomes rather than specific gene loci. This does not come without some drawbacks, as metagenomic sequencing and analyses often require more computational resources to carry out. In addition, the field of metagenomics is rapidly evolving, and there is not s singular well-established method in carrying out these studies. Often, one chooses between many different tools that accomplish similar tasks. Metagenomics was chosen for this study because it can examine a larger portion of microbial genomes, and because our question necessitates information that is not present in traditional marker gene analyses.

Reconciliation of examining DNA and not RNA

One key factor to keep in mind when drawing conclusions from metagenomic studies is that since we will be working with gene profiles and not mRNA, we cannot directly infer metabolic function from our gene analyses. To be more confident in those potential metabolic functions, this study would need to be paired with a quantification of mRNA transcripts, which could provide a relationship between protein encoding genes and metabolic processes. This is a limiting factor of only using metagenomics to assess metabolic potential reviewed by Rocca et al. (2014). Rocca et al (2014) noted a significant positive correlation between gene abundance and process in studies that included both abundance and functional analyses. This suggests that perhaps gene abundances in combination with functional data (obtained from a database such as the Kyoto Encyclopedia of Genes and Genomes) can lend more credence to the conclusions drawn. This can be implemented into the study by examining the strength of correlation between gene abundances and functional data. While this approach could strengthen our conclusions, it is still important to remember its limitations.

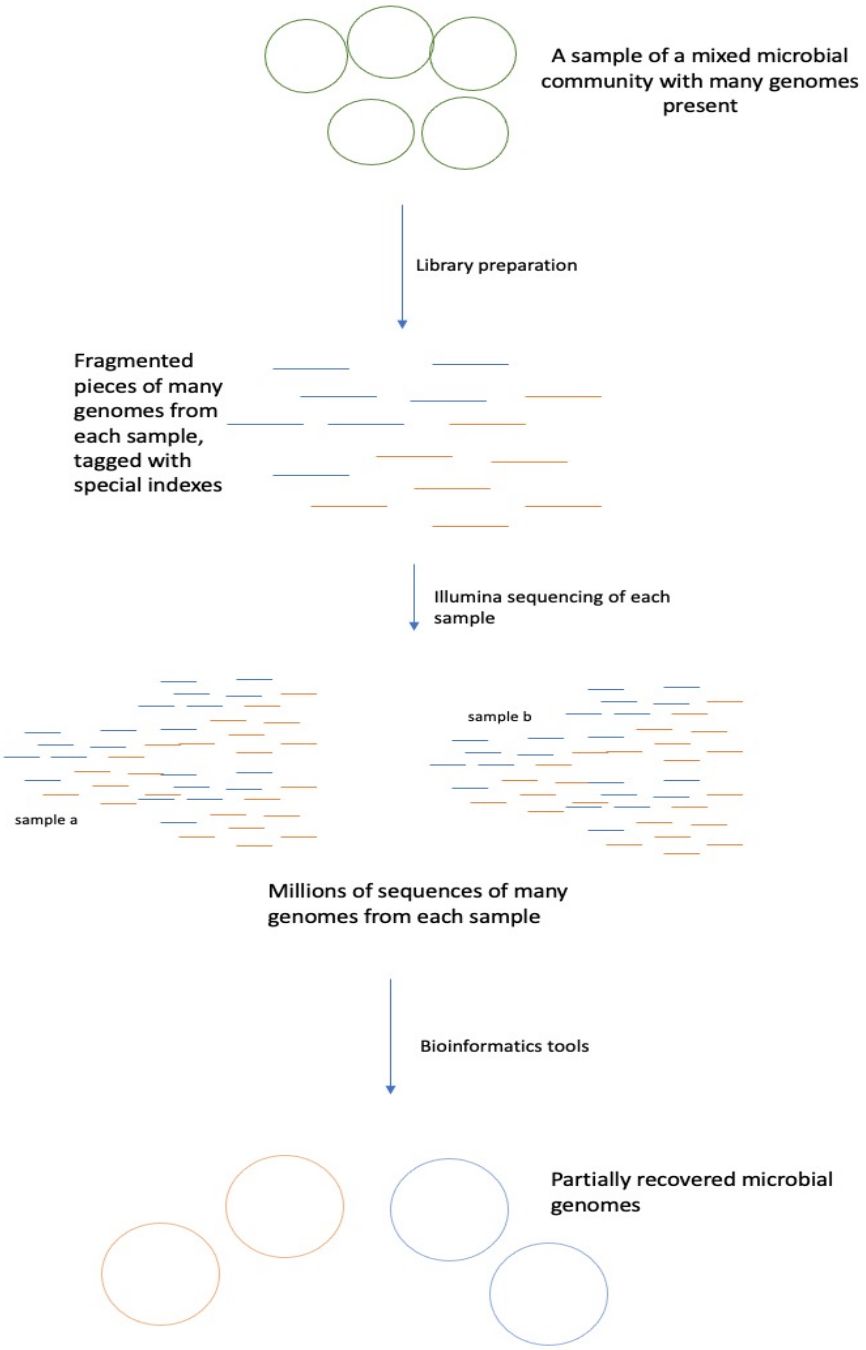


Figure 2. A general overview of the metagenomic method.

General Overview of Methods

Samples have been collected from a designated coral reef site in Summerland Key, Florida, USA. Given the results of a preliminary study (see next section), DNA was extracted using the QIAmp kit, which will deplete host DNA and recover microbiome DNA. The Illumina HiSeq was be used to sequence the recovered DNA.

Upon retrieving raw sequence data, sequence reads were processed through a 'bioinformatics pipeline'. In this study, a pipeline refers to a set of software programs that will be applied in a specific order to our data, where the output from one software is the input of the next software.

Preliminary Study to Determine Optimal DNA Recovery Method

The DNA extraction method chosen for this study can have an impact on the sequence reads obtained downstream. High-levels of host DNA present in the final sample that is sequenced can make it difficult to achieve the sequencing depth and coverage necessary to practically assemble metagenomes (Pereira-Marques et al. 2019).

For this reason, a preliminary sequencing run was conducted June 2022 testing different DNA extraction methods with varying levels of selectivity for host and microbiome DNA. This preliminary run tested three different DNA extraction protocols to determine which method would deplete the most host DNA. Statistical analyses could not be carried out on the results of this preliminary run due to a lack of replicates. The three DNA extraction methods were each tested once on a subsample of the three individuals, totaling 9 samples that were sequenced. The three DNA extraction methods tested were as follows.

DNA Extraction Method 1 – QIAmp kit + cell separation

A physical cell separation following methods from Freeman and Thacker 2011, followed by a DNA extraction using the QIAmp DNA Microbiome kit (QIAmp kit). This QIAmp kit separates and then depletes host DNA (Wen et al. 2016, Heravi et al. 2020). This separation will concentrate microbial DNA into a pellet. A chemical separation of host DNA is then preformed on the pellet using the QIAmp kit.

DNA Extraction Method 2 – DNEasy kit + cell separation

A DNA extraction using the DNeasy PowerLyzer PowerSoil (DNeasy) kit on a pellet obtained from a physical cell separation. The DNeasy kit extracts all nucleic acids from a sample, with no selectivity for host or microbiome DNA.

DNA Extraction Method 3: QIAmp kit only.

A DNA extraction on a sample of sponge tissue using the QIAmp kit (without physical separation of host DNA).

Table 1. Table of observed proportions of prokaryotic and eukaryotic reads in 9 samples using Kaiju (3 individuals per method)

Method	Average proportion of all reads with classified taxonomy	Average proportion of prokaryotic reads (from total classified reads)	Average proportion of eukaryotic reads (from total classified reads)
QIAmp kit + cell separation	~0.79	~0.56	~0.056
DNEasy kit + cell separation	~0.39	~0.63	~0.22
QIAmp kit	~0.90	~0.93	~0.016

Results of Preliminary Study for Optimal DNA Recovery

To determine which DNA extraction method resulted in the optimal bacterial DNA recovery, it was necessary to obtain a taxonomic profile of the communities present in our samples. The goal was to obtain a ratio of eukaryotes to prokaryotes found in our samples, and then to use this ratio as a rough proxy for host DNA to microbiome DNA. While not all eukaryotic DNA sequenced necessarily belonged to the host, we were working under the assumption that this proxy ratio would give us a reasonable estimate of the amount of host DNA depleted through the extractions.

Taxonomic classification was performed by Kaiju (Menzel et al. 2016) via their free web server (https://kaiju.binf.ku.dk/server). The results from Kaiju listed in Table 1 suggest that the QIAmp kit on its own returned the most prokaryotic DNA, in addition to also classifying the most reads. While the combination of cell separation and the QIAmp kit also returned decent

overall classification, the prokaryote/eukaryote ratio (P/E) was smaller than observed in the former method. The DNeasy kit with a cell separation, which had no selectivity for microbiome/host DNA, returned the lowest overall classification and P/E. Based off these results, the QIAmp kit on its own was used to extract DNA from samples.

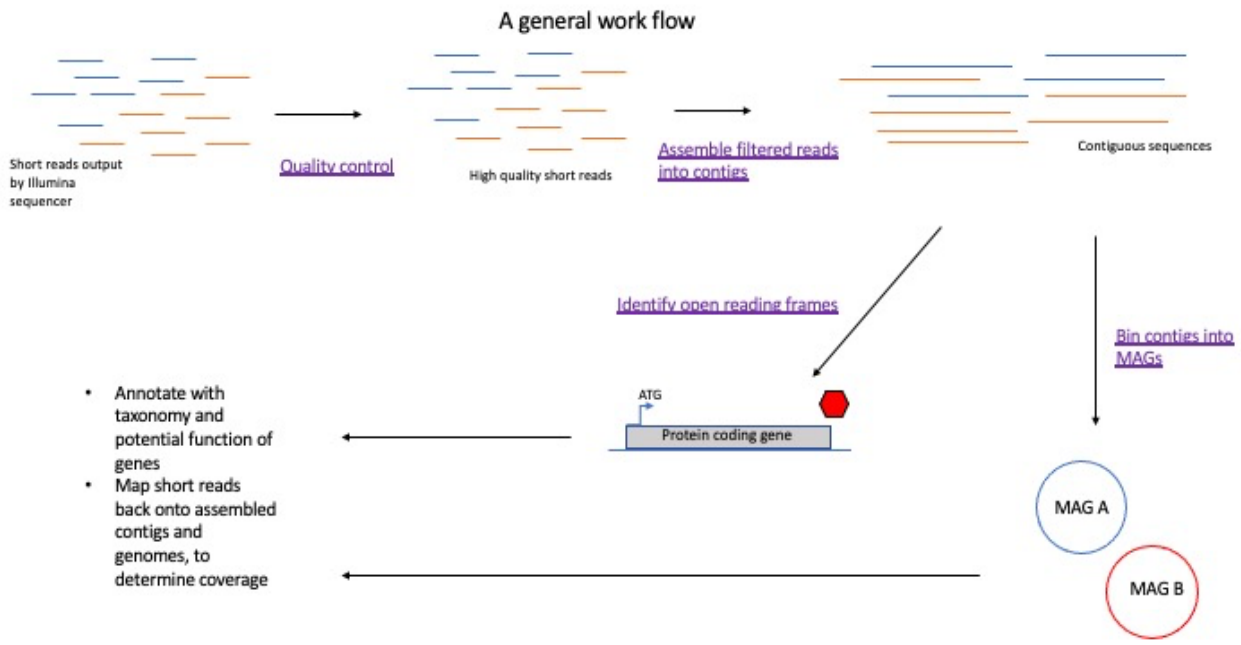


Figure 3. A general workflow for working with our metagenomic reads to assemble metagenomes.

General Steps of the Workflow

1. Quality control

Low quality reads were filtered out and adapter sequences were trimmed from the raw demultiplexed data. This was done using fastp version 0.23.2 (Chen et al 2018, Supplementary File 1). Each of these steps of the workflow was carried out on the Voltron Cluster, a Linux based computing cluster hosted by the MTSU chemistry department. This cluster required the user to write a job script to reserve shared resources and computing time. All job scripts which document terminal commands used are included as supplementary files.

2. Contig assembly

Filtered short reads were assembled into longer nucleotide sequences called contigs. For this study, SPAdes version 3.15.0 (Bankevich et al 2012) was used to assemble contigs from short reads output by fastp (Supplementary File 2).

3. Gene prediction

Gene prediction can be carried out by identifying open reading frames (ORFs), which allows us to identify likely protein coding sequences (Sieber et al 2018). Once we identified genes, they were annotated. Annotation identifies the protein encoded by each gene by searching our ORF's against a chosen database. Prodigal version 2.6.3 (Hyatt et al 2010) to identify ORF's (Supplementary File 3).

4. Functional annotation with KEGG and extracting nitrogen cycle gene annotations

With predicted ORF's, some functional annotations were run. Here, we used KEGG's free "GhostKoala" server (Kanehisa et al 2016). Protein sequences were uploaded, and an email was sent when the job was ready to start and when it was complete. Our jobs were run with the "genus_prokaryotes + family_eukaryotes" database. Due to KEGG's upload limit, outputs from prodigal were broken down into different files, and thus the annotation files were broken down into several files. KOALA-formatter (https://github.com/Arkadiy-Garber/KOALA-formatter) was used to consolidate the individual files into one summary file. With these annotation summaries, it was necessary to extract only annotations involved with nitrogen cycle processes. To do this, the in-house script "koala_sort.py" (Supplementary File 4) was used to filter out the genes outlined in Figure 1 from the summary files. The genes were then counted by which process they came from, and heatmaps were generated to visualize the abundances of the genes in each sample (Supplementary File 5). Recalling that these heatmaps represent abundances from contigs and not MAGs, this helps us visualize overall functional potential from the sponge holobiont.

5. Sequence alignment

For our sequence alignments, short paired-end reads were mapped onto contigs built with SPAdes. This outputs a .sam file (sequence alignment map), which contains information such as depth of coverage obtained, and is necessary for binning MAG's downstream. Sequence alignment was done using minimap2 version 2.25 (Li 2018, Supplementary file 6). In their SAM format, alignment files are large and slow to work with, and so they are often converted into BAM (binary alignment map), or a binary representation of the counterpart SAM file. After

generating the BAM file, the file is indexed, which allows for fast access to sequences in the file. This was done using samtools version 1.16.1 (Danecek et al 2021, Supplementary File 7).

7. Genome binning, Quality Checks, and Assigning Taxonomy, and Annotating MAGs

In the binning step, contigs were grouped into draft genomes. For this study, metabinner version 1.4.4 was used (Wang et al 2023). Once metagenomes were obtained, we needed to ensure that they were of high quality. The two key metrics used to assess the reliability of our genomes are completeness and contamination/redundancy (Parks et al 2015). Completeness refers to an estimated proportion of the genome that we've recovered, and contamination is a proportion of how much of the genome is likely contaminated with DNA from other microorganisms. These two metrics are determined by the presence of single-copy genes. In this study, we considered a genome high-quality if it was greater than 80% complete and less than 5% contaminated. Binning and quality assessment were carried out by metabinner, which runs CheckM version 1.2.2 (Parks et al 2015) after it generates bins (Supplementary File 8). To assign taxonomy to the bins, GTDB-Tk version 2.1.1 (Chaumeil et al 2020) was used (Supplementary File 9). The script "count_bins.py" (Supplementary File 10) was written to count how many times each taxon appeared in each sample, and the relative abundance of taxa in the MAGs was visualized with R from this count table (Supplementary File 11). Annotating MAGs was similar to the contigs, with protein files being uploaded to GhostKoala and then being sorted with sort_koala.py. With the MAG's, we further extracted only nitrogen cycle gene sequences with "grep_seq.pbs" (Supplementary File 12). The script "change_headers.py" was written to change the header of each sequence such that they were simplified but remained unique (Supplementary File 13). An additional python script called "make_table.py" was written to generate

presence/absence tables of each nitrogen cycle gene present in each sample (Supplementary File 14).

8. Multiple Sequence Alignments and Generating Phylogenetic Tree Files

To observe relationships between the nitrogen cycle genes extracted from the python scripts a multiple sequence alignment was conducted using MAFFT version 7.505 (Katoh & Standley 2013) hosted on the CIPRES Science Gateway server, a free bioinformatics tool website hosted by UC San Diego (Miller et al 2010). Fasta files containing sequences of a gene generated by "change_headers.py" for all samples were uploaded to CIPRES, and default parameters were used with the data type being "nucleotide". Phylogenetic trees were generated from the MAFFT files using RAxML-NG version 1.2.0 (Kozlov et al 2019, Supplementary File 12), using the "GTR+G" model option, and tree files were output as newick tree files.

9. Faiths Phylogenetic Diversity Metric and Multivariate Analyses with R

The purpose of the phylogenetic analyses was to determine which nitrogen cycle genes drove dissimilarities across our species. This was done by conducting sequence alignments followed by the creation of maximum likelihood phylogenetic trees from nitrogen cycle genes and calculating faiths phylogenetic diversity (FPD), which considers a phylogeny and a presence absence table to tell us about the diversity of nitrogen cycle genes in our species. These tables of diversity metrics were then used to generate nMDS plots and dendrograms to visualize dissimilarities. For the purposes of this study, we will define a group as samples with a dissimilarity of less than 20%. R packages picante, vegan, ggplot2, and tidyverse were used to generate faith's phylogenetic diversity metric (PD), conduct dimensional reductions based off the metric, and visualize dissimilarity between samples with NMDS plots (Supplementary File 13).

Faiths PD considers the presence/absence of nitrogen cycle genes present in our samples and their relatedness from the RAxML-NG phylogenies to assess the diversity of genes in sponge and water samples. The envfit function from the vegan package was also used to visualize vectors associated with different genes, and the Bray-Curtis distance metric was used for visualizing dissimilarities.

Results

Results of metagenome assemblies

Metagenome assembled genomes (MAGs) were recovered from five of the seven sponge species and from filtered seawater samples (Figure 4). The number of high-quality MAGs, percent completeness, and percent contamination varied across samples (Table 2). Two sponge species, *I. birotulata*, and *V. rigida* samples, only produced four total MAGs, all of which had low completeness (max = 17.59%) thus they were not included in MAG analyses. For this study, we've defined high quality as being greater than or equal 80% complete and less than or equal to 5% contaminated. Generally, 90% completeness is the standard for high quality MAGs (Bowers et al. 2017). However, the completeness cutoff was dropped to 80% so we could include several MAGs that were between 88-89% complete and less than 5% contaminated. Before distilling MAGs down to a list that met the high-quality standard, there were a total of 101 MAGs. From these 101 MAGs, 46 were high-quality and 55 were medium-quality (Supplementary Table 1). A MAG was considerd medium quality if it was not high-quality and was greater than or equal to 50% complete and less than or equal to 10% contaminated. Although the medium-quality MAGs were not used in the analyses their quality and assigned taxonomy are included in Supplementary Table 1.

Table 2: Number of high-quality MAGs from each sample, with average completeness and contamination values.

Sample	Number of high-quality MAG's	Average Completeness (%)	Average Contamination (%)
Acau1	8	91.5	2.2
Acau2	8	96.1	1.4
Acon1	5	92	1.6
Acon2	8	93.6	2.3
Acon3	1	93.8	2.5
Acra1	1	89.3	2
Aful2	3	91.4	1.6
Xmut1	3	97.2	2.4
Xmut2	3	94	1.5
Wfilter1	3	88.5	2
Wfilter2	1	96.6	3.4
Wfilter3	2	94.4	3.6

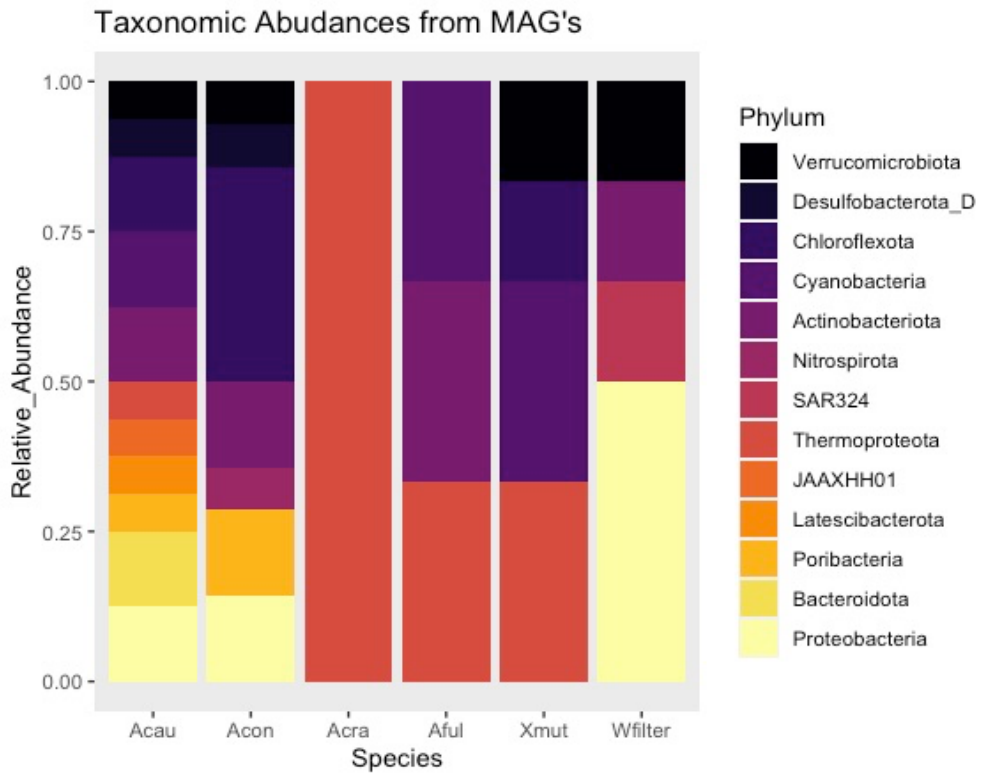


Figure 4: Taxonomic profiles of 46 high-quality MAG's across five sponge species and seawater samples. MAGs are classified to phylum.

The high-quality MAG's spanned twelve bacterial phyla and one archaeal phylum. *A. cauliformis* and *A. conifera* showed the highest MAG richness (11 and 7 different phyla for *A. cauliformis* and *A. conifera* respectively) in assigned taxonomy and returned the most MAGs, comprising more than 50% of all high-quality MAG's. Proteobacteria was the most common phylum across all samples, appearing in seven of the twelve samples. Both Alphaproteobacteria and Gammaproteobacteria classes were represented in these seven samples. Four Cyanobacteria genomes from the genus *Synechococcus* were recovered from *X. muta*, *A. fulva*, and *A. cauliformis*, all of which have previously been recorded to host *S. spongiarum* as an abundant photosymbiont (Erwin & Thacker 2008). Six representatives of the archaeal phylum Thermoproteota were found in *A. cauliformis*, *X. muta*, *A. fulva*, and *A. crassa* all were assigned to the same strain from the class Thaumarchaeota who are free-living ammonia-oxidizers (Prosser 1990, Pester et al. 2011). Since several of the same MAGs appeared in multiple samples, we have standardized the nomenclature and condensed them into 26 unique bins for the 46 high quality MAGs that were recovered (Supplementary Table 2). MAGs that were classified down to the same genus were considered the same MAG in the condensed naming system.

A. cauliformis (n=2) replicates, yielded 16 high-quality MAG's (Supplementary Table 1) that represented eleven bacterial phyla and one archaeal phylum. (Supplementary Table 2). Genes for the processes of nitrification, denitrification, ANR, and DNR were detected in six separate MAGs recovered from this species. A near complete set of genes for the denitrification pathway were found in an Alphaproteobacterial symbiont (Bin 19). These are the genes *NarGHI*, *NirK*, and *NosZ* which facilitate the first, second, and fourth steps of denitrification, respectively. For nitrification, all genes necessary were observed in three different MAGs. The Thaumarchaeota MAG (Bin 1) contained AmoCAB, which catalyzes the first transformation

from ammonia to hydroxylamine. Two Actinobacterial MAGs (Bin2 and Bin3, possessed Hao, which codes for hydroxylamine dehydrogenase and converts hydroxylamine into nitrite. The Alphaproteobacteria MAG (Bin19), also had NxrAB, which catalyzes the final step of nitrification of hydrodoxylamine into nitrite. For ANR, the genes NarB and NirA which catalyze the first and second steps of ANR, respectively, were detected in a Cyanobacteria MAG (Bin12), two Poribacteria MAGs (Bin17 and Bin18), and two Verucomicrobiota MAGs (Bin 24 and Bin25). These ANR genes represent all the genes necessary to carry out the complete ANR pathway. Finally, for DNR, NarGHI, which facilitates the first step of DNR of converting nitrate to nitrite was detected in the Alphaproteobacteria MAG (Bin 19).

Agelas conifera

Our three *A. conifera* samples returned 14 high-quality MAGs (Supplementary Table 1) that represented seven bacterial phyla (Supplementary Table 2). MAGs from replicates of *A. conifera* contained genes for denitrification, nitrification, ANAMMOX, ANR, and DNR. Nearly all these genes were found in a single Nitrospirota MAG (Bin 16). This included *NarGHI*, *NirK*, and *NxrAB* which are involved in denitrification, DNR, and nitrification respectively. Indeed, the members of the phylum Nitrospirota found in marine and other aquatic environments are heavily involved in the nitrogen cycle (Daims et al. 2000, Daims & Wagner 2018). One MAG that was recovered from two A. conifera replicates, Poribacteria (Bin17), was originally discovered in sponges and has been identified in sponge microbiome studies (Fieseler et al. 2004, Podell et al. 2019). Our annotations for this Poribacteria MAG revealed the NarB gene, which initiates the first conversion of nitrate to nitrite in ANR. Other high-quality MAGs from the three A. conifera replicates included the phyla Chloroflexota, Desulfobacterota, but these phyla also did not reveal nitrogen cycle metabolism in our annotations.

Aiolochroia crassa

Our *A. crassa* sample returned one high-quality MAG, which was identified to the class Thaumarchaeota (Bin 1), for which our annotations revealed the potential to catalyze the first steps of nitrification with the *amoCAB* gene.

Aplysina fulva

A. fulva returned three high-quality bins: Bin1, Bin2, and Bin12. These bins belonged to the phyla Thermoproteota, Actinobacteriota, and Cyanobacteria respectively. The Thaumarchaeota MAG (Bin 1) did reveal potential for nitrogen cycle metabolism with *amoCAB* in denitrification and *nirK* in denitrification, much like our other samples who hosted this taxon. An Actinobacterial MAG (Bin 2) revealed potential for nitrification through the possession of the *Hao* gene, facilitating the second step of nitrification. A Cyanobacteria MAG (Bin12) possessed the *NarB* and *NirA* genes, which carry out the first and second steps of ANR, respectively.

Xestospongia muta

Two X. *muta* samples produced three unique MAGs. These were Bin1, Bin12, and Bin24. They were identified as coming from the phyla Thaumracheota, Cyanobacteria, and Verrucomicrobia respectively. The Thaumarchaeota MAG (Bin 1) possessed the amoCAB gene that participates in nitrification, and the *nirK* gene that participates in denitrification. The Cyanobacteria MAG (Bin 10) had the *NarB* and *NirA* genes from ANR. The Verrucomicrobia (Bin 24) possessed *NxrAB* from nitrification, *NarGHI*, *NirK*, and *NosZ* from denitrification, and *NarGHI* from DNR.

Water filters

The filtered water samples returned three unique high-quality MAGs: Bin2, Bin22, and Bin25. These were classified as Actinobacteriota, SAR324, and Verrucomicrobia respectively. Our annotations for these seawater samples revealed nitrogen cycle potential in the Actinobacteria MAG (Bin 2) which possessed the *Hao* gene necessary for an intermediate step in nitrification. SAR324 (Bin 22) is widely distributed group appearing in many different environments (Sheik et al. 2014) whose metabolic potential has been studied in terms of nitrogen and sulfur (Malfertheiner et al. 2022). In our annotations, the SAR324 MAG also possessed the *Hao* gene from nitrification. Lastly, the Verrucomicrobiota MAG (Bin 25) possessed *NapAB* from denitrification, *NirA* from ANR, and *NapAB* from DNR.

Results of Completeness of N cycle pathways in each species using assembled contigs

In this part of the results, we will look at the nitrogen cycling potential for the entire microbiome of an individual sponge rather than focusing on individual taxa. We did this by annotating contigs assembled from short reads. Specifically, we will look at whether nitrification, denitrification, ANR, and DNR are complete pathways in our species. Complete is defined as possessing all the genes necessary to carry out one of the four nitrogen cycle processes. (Figure 1).

Nitrification

Nitrification is a biological process that transforms ammonia into nitrate, with hydroxylamine and nitrite as intermediates between ammonia and nitrite. Our annotations revealed that *A. cauliformis*, *A. conifera*, *A. crassa*, *A. fulva*, *V. rigida*, and *X. muta* possessed all the genes necessary for complete nitrification, with only *I. birotulata* missing the *amoCAB* gene. Although most species showed a complete pathway, *A. cauliformis*, *A. conifera*, and *A. crassa*

had the highest abundance of nitrification genes, which suggests these three species may have the highest nitrification potential from our species list.

Denitrification

Denitrification involves the complete reduction of nitrate into elemental nitrogen, which is carried out by four different enzymes. All samples including filtered seawater showed complete sets of genes for the denitrification pathway with *A. cauliformis*, *A. conifera*, *A. crassa*, and *X. mut* showing a high abundance of all genes except *nosZ*, which catalyzes the final step of denitrification of turning nitrous oxide into nitrogen (Figure 6). Denitrification gene counts were lower in I. birotulata and V. rigida than in filtered seawater, suggesting a low potential for denitrification in these two species.

Dissimilatory Nitrate Reduction

Dissimilatory nitrate reduction (DNR) is the process of reducing nitrate to ammonia, with nitrite as an intermediate. In contrast to denitrification which removes bioavailable nitrogen from a system by producing nitrogen gas, DNR keeps bioavailable nitrogen in a system by producing ammonia could be used by other microorganisms. Like denitrification, all samples showed complete sets of genes for DNR except for I. birotulata who was missing the nirD gene. In general, nirD was also the least abundant gene in all samples (Figure 7). Interestingly, the filtered seawater had the highest counts for all DNR genes, suggesting an importance to preserve some bioavailable nitrogen in the water column.

Assimilatory Nitrate Reduction

Assimilatory nitrate reduction (ANR) is a very similar process to DNR, with the only differences being that different enzymes are used in the second to convert nitrite to ammonia, and presumably the ammonia products in this process are assimilated into biomass rather than

expelled as waste. For ANR, all samples showed presence of *narB* and *nirA*, the two genes necessary for the first and second steps of ANR respectively. Like DNR, *I. birotulata* and *V. rigida* had the lowest abundances in *narB* and *nirA*, suggesting a lower potential than the other species. *A. cauliformis* and *X. muta* had the highest abundances of the ANR genes, with *A. cauliformis* being like filtered seawater in terms of gene counts (Figure 8).

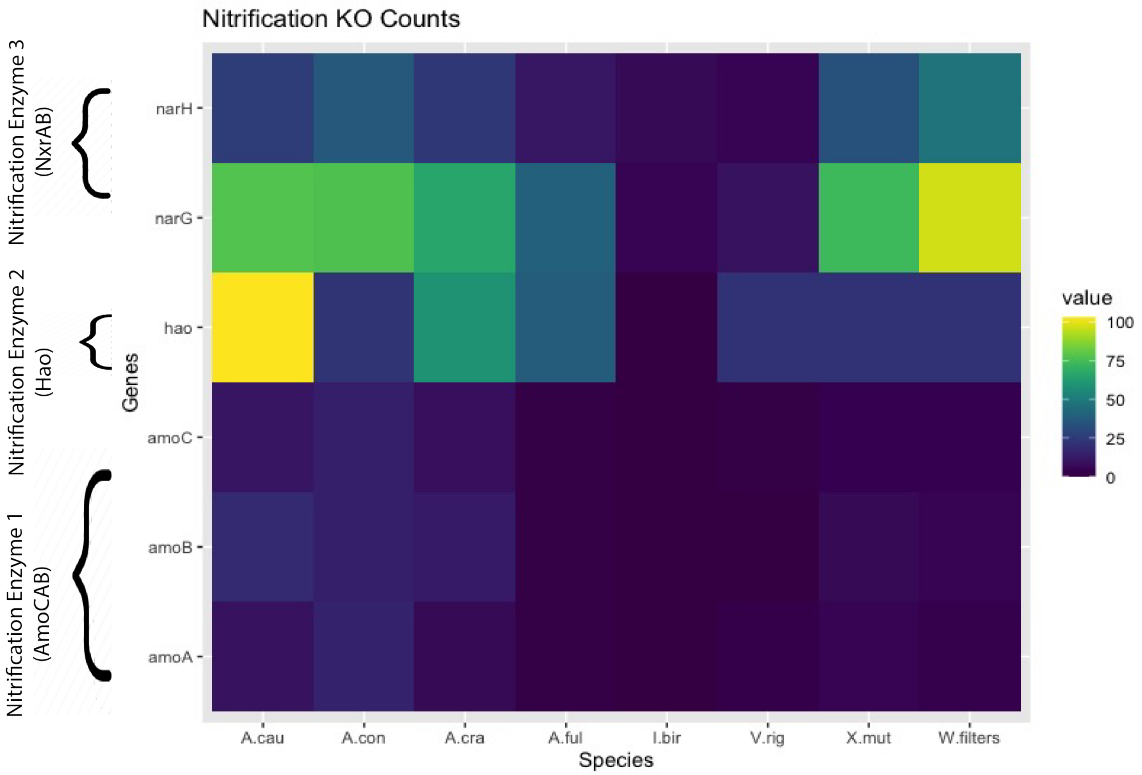
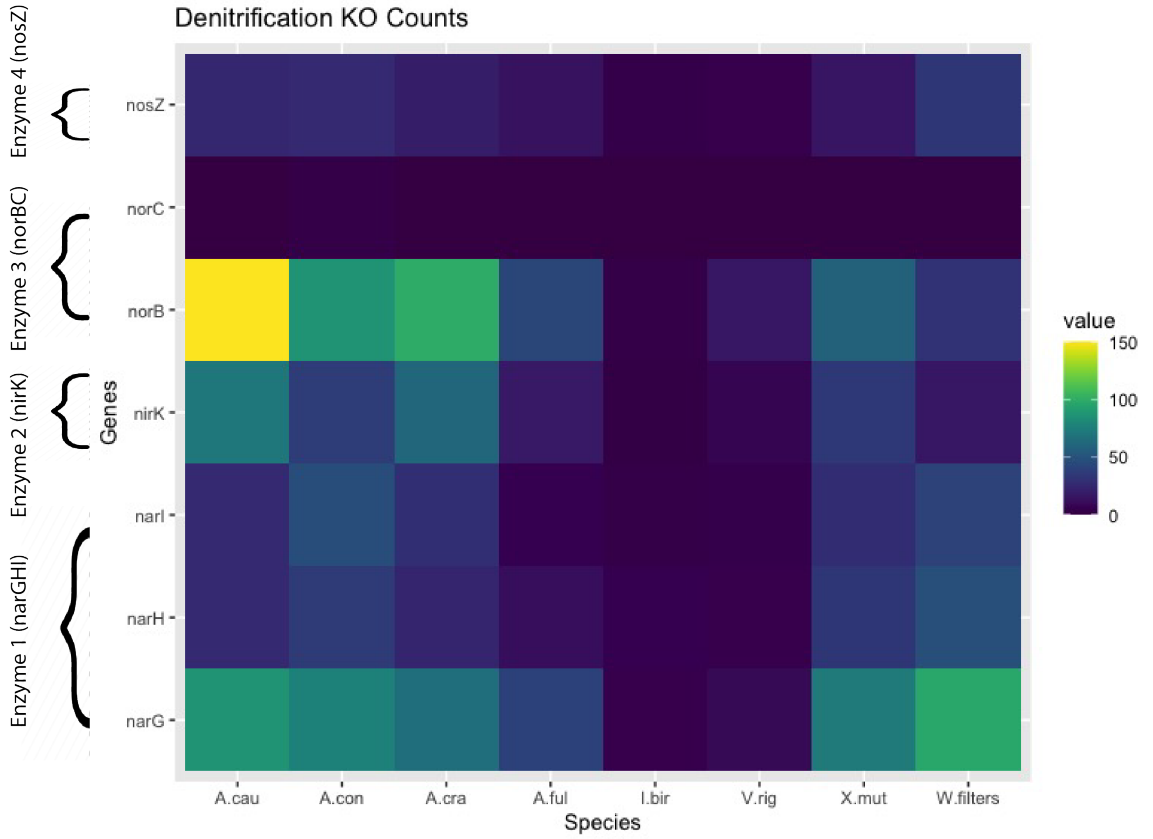


Figure 5: A) Heatmap of nitrification gene abundances in each species. B) Linear schematic of nitrification pathway.



Denitrification

Figure 6: A) Heatmap of denitrification gene abundances in each species. B) Linear schematic of denitrification pathway.

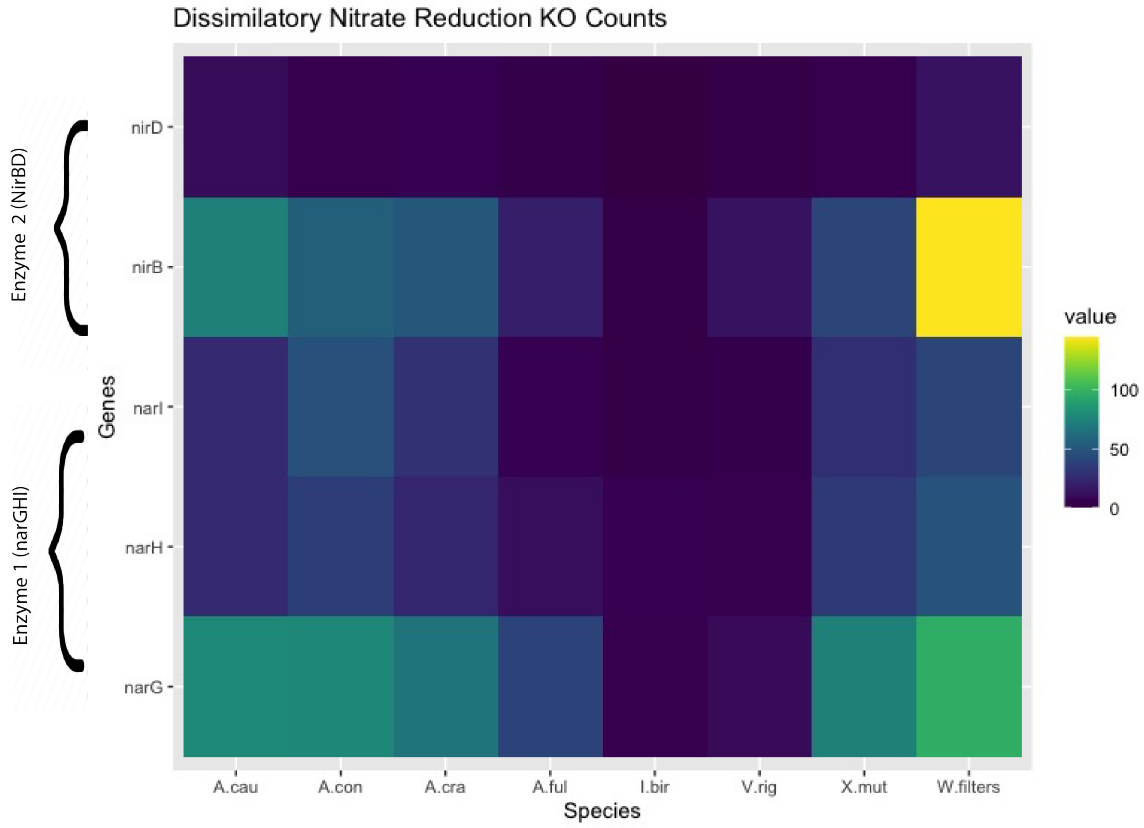
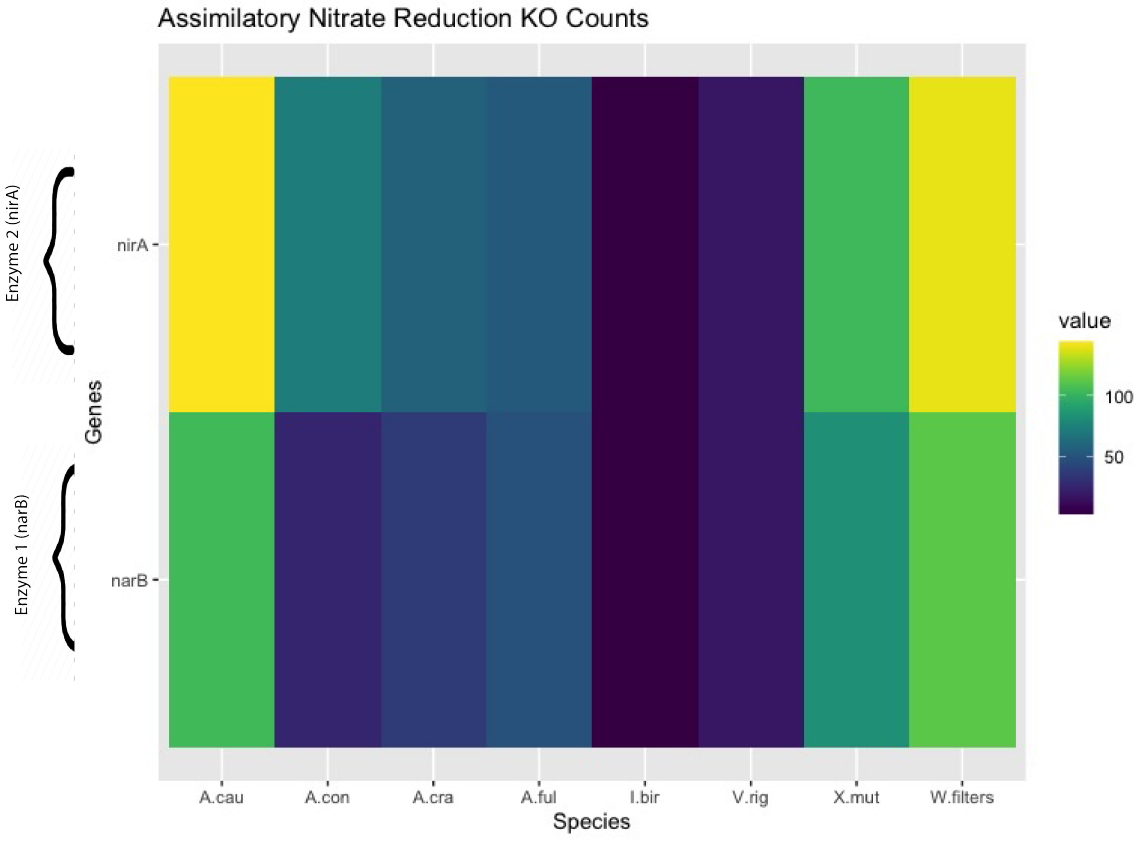


Figure 7: A) Heatmap of dissimilatory nitrate reduction gene abundances in each species. B) Linear schematic of dissimilatory nitrate reduction pathway.



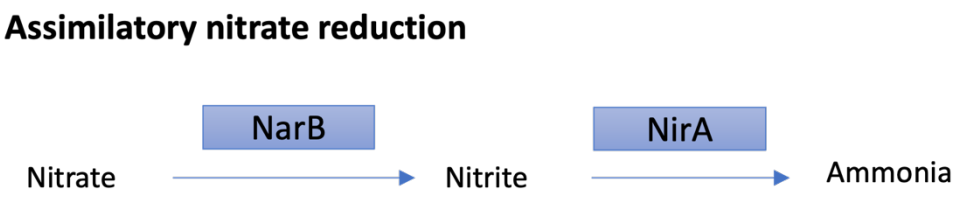


Figure 8: A) Heatmap of assimilatory nitrate reduction gene abundances in each species. B) Linear schematic of assimilatory nitrate reduction pathway.

Results of Phylogenetic Analyses

Nitrification

A dendrogram visualizing Bray-Curtis dissimilarity of nitrification genes showed a split into three main groups. *V. rigida* grouped by itself, *A. conifera* plus one *A. cauliformis* replicate, and the rest of the samples. At the 20% threshold, nearly all samples grouped by species except for *A. cauliformis*. Fitting environmental vectors onto an nMDS revealed that *amoA* (r^2=0.66), *amoB* (r^2=0.41), and *hao* (r^2=0.44) were correlated with the dissimilarities observed. *AmoA/amoB* and *hao* catalyze the first two critical steps in nitrification, respectively. This result agrees with the pathway completeness analysis of nitrification where *A. cauliformis*, *A. conifera*, *A. crassa* had the highest functional potential for nitrification and a high abundance of *AmoA/AmoB* and *hao* genes. Water samples also grouped together but lacked nitrification genes in our annotations (Figure 9).

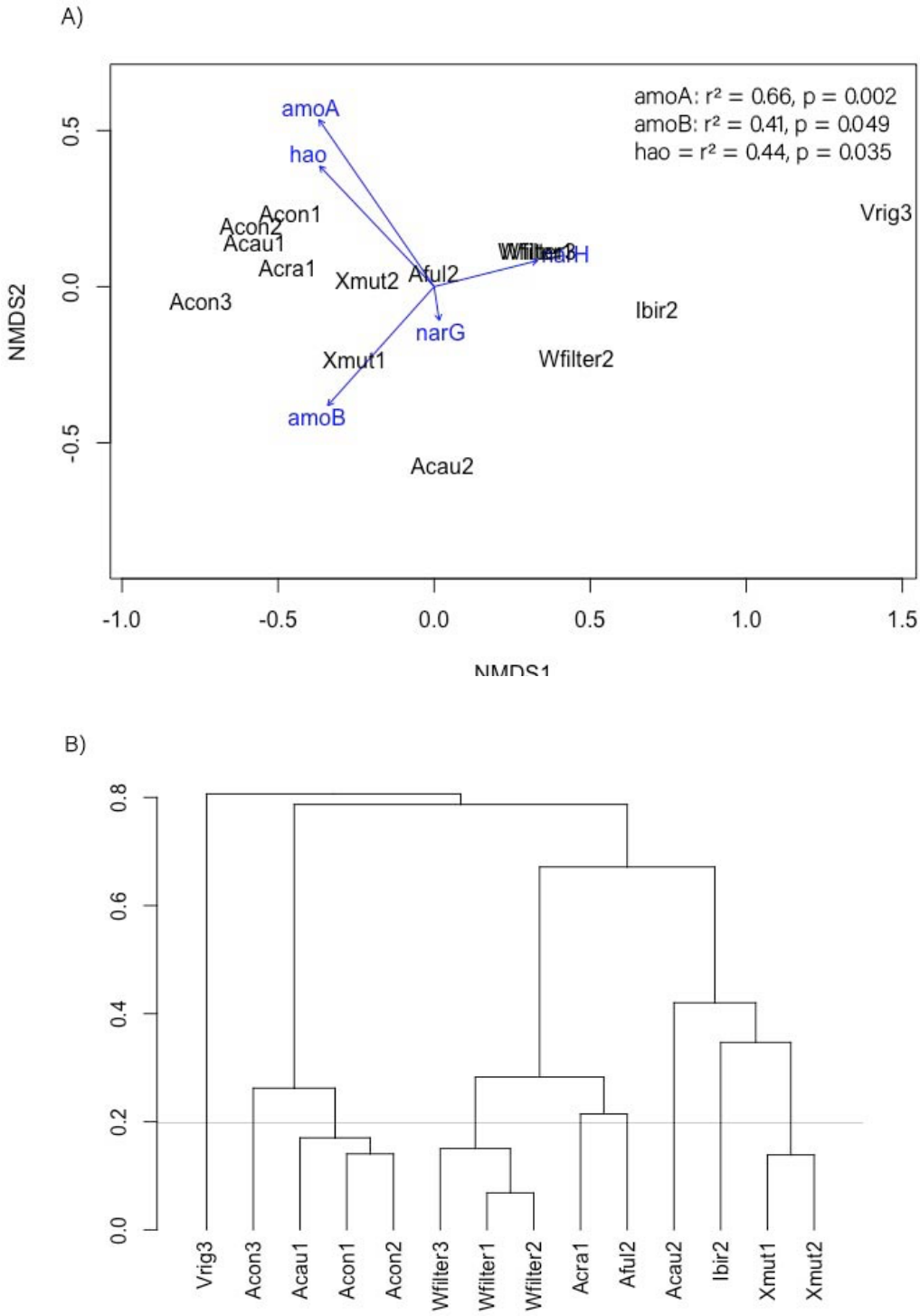


Figure 9: A) nMDS plot of nitrification genes B) A dendrogram based on Bray-Curtis dissimilarity of nitrification genes.

Denitrfication

Clustering of dentification genes showed most samples grouped by species, except for one *A. conifera* replicate that grouped with the *I. birotulata* sample (Figure 10B). Our *A. fulva* sample also grouped with the sweater samples. A dimensional reduction of species dissimilarities showed vectors for *narG* (r^2=0.64), *narH* (r^2=0.82, p = 0.001), and *nosZ* (r^2=0.58) were correlated with the diversity of denitrification genes in our samples (Figure 10A). *NarG* and *narH* help form the *narGHI* enzyme which facilitates the first step of denitrification. *NosZ* is solely responsible for the final step of denitrification where nitrous oxide is transformed into nitrogen gas. Considering the completeness analysis of denitrification where there were similar abundances of denitrification genes across all samples, this may suggest that the first and last steps of denitrification are of particular importance and may be limiting steps of the pathway.

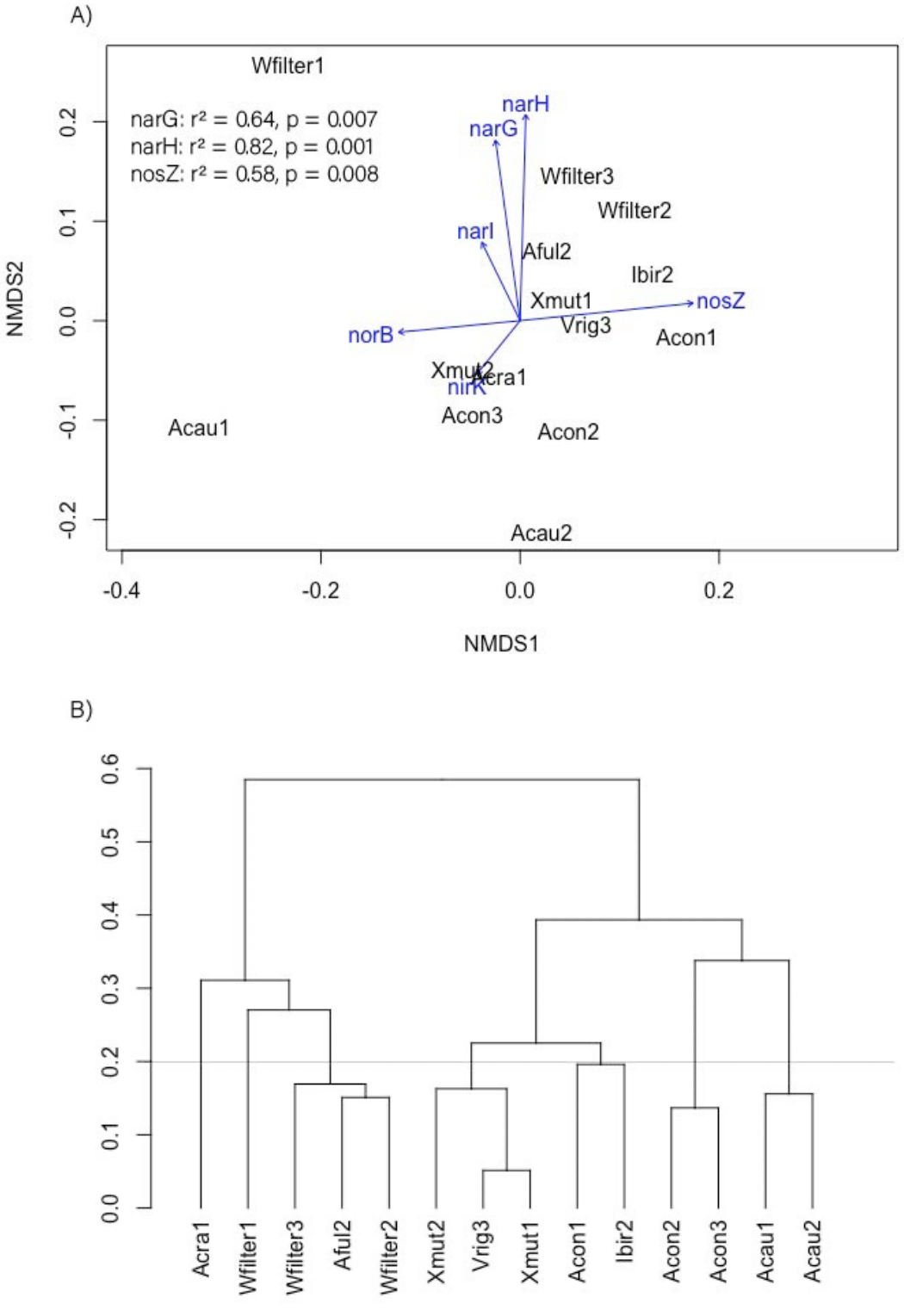


Figure 10: A) nMDS plot of denitrification genes B) A dendrogram based on Bray-Curtis dissimilarity of denitrification genes.

Dissimilatory Nitrate Reduction

A dendrogram of DNR gene dissimilarities revealed some grouping by species, although *A. conifera*, *A. cauliformis*, and *X. muta* replicates did not all cluster together, and our *A. crassa* sample clustered with the seawater samples (11B). The genes that seemed to drive the most dissimilarity in the ordination were *nirD* (r^2=0.64), and *narH* (r^2=0.64) (Figure 11A). *NarH* is part of the *narGHI* enzyme that carries out the first step of DNR. *NirD* is part of the nirBD complex that carries out the second step of DNR. This suggest that the presence of *narGHI* in sponge symbionts may be quite important, since our results show that *narGHI* it contributed to dissimilarities in both DNR, and denitrification.

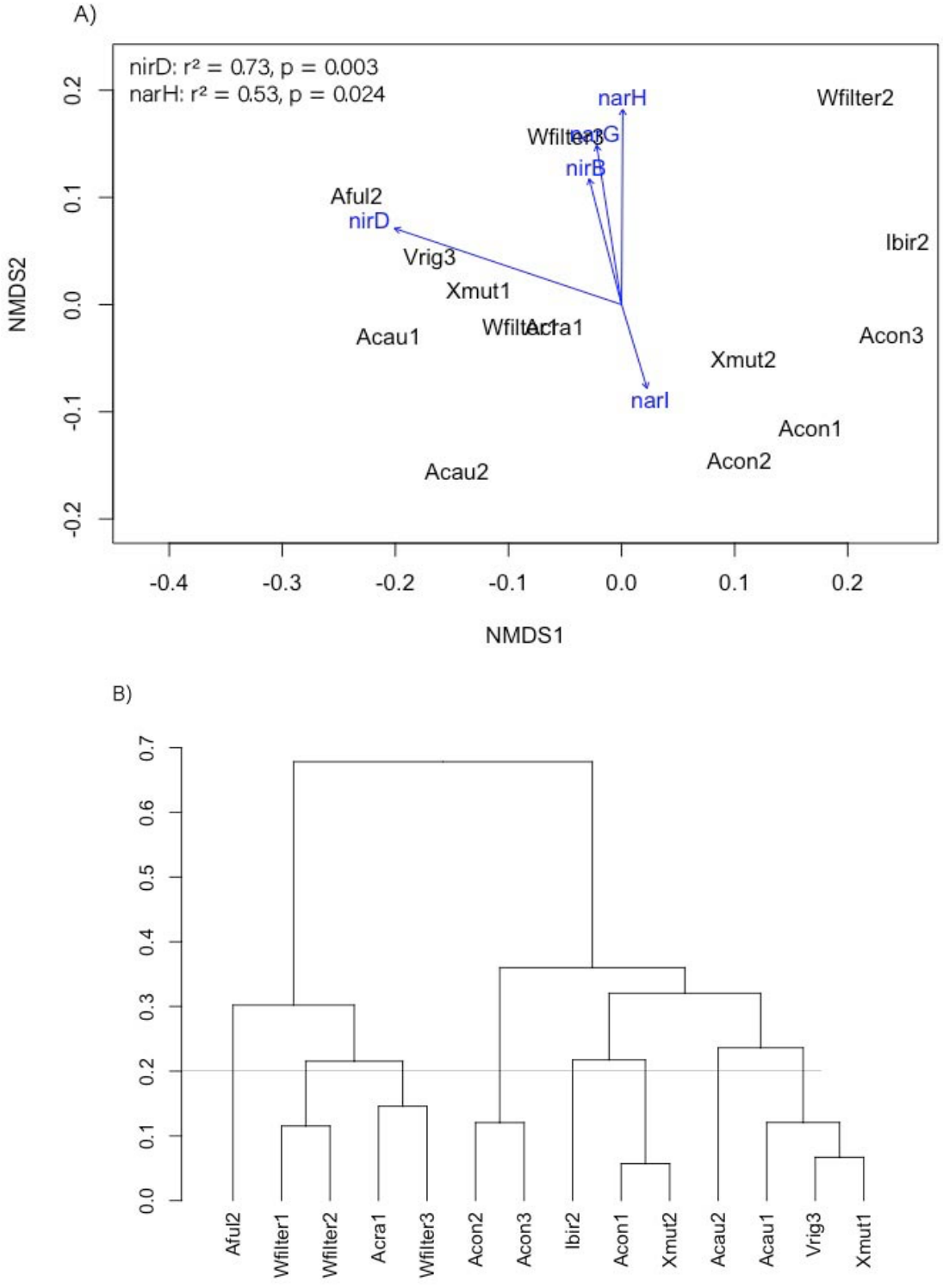


Figure 11: A) nMDS plot of dissimilatory nitrate reduction genes B) A dendrogram based on Bray-Curtis dissimilarity of dissimilatory nitrate reduction genes.

Assimilatory Nitrate Reduction

Visualizing clustering of Bray-Curtis Dissimilarity in ANR genes showed all sponge samples clustered by species and one seawater sample clustered separate from the other two (12 B). Fitting environmental vectors onto the nMDS ordination for ANR genes did not reveal significant correlation in diversity from any of the genes, as the magnitude for the vectors was similar and their directions were all different (12A). The nMDS visualization for ANR genes showed a small clustering of *A. cauliformis* and *X. muta* replicates (12A). In this case, there are only three genes that participate in ANR. These were *narB* and *nasAB* both of which have the potential for catalyze the first step of ANR, and nirA that contributes to the second step of ANR. Ammonia produced from ANR is usually assimilated into the cell, and so it may be important to have more than one enzyme that can catalyze the first step ANR according to our annotations.

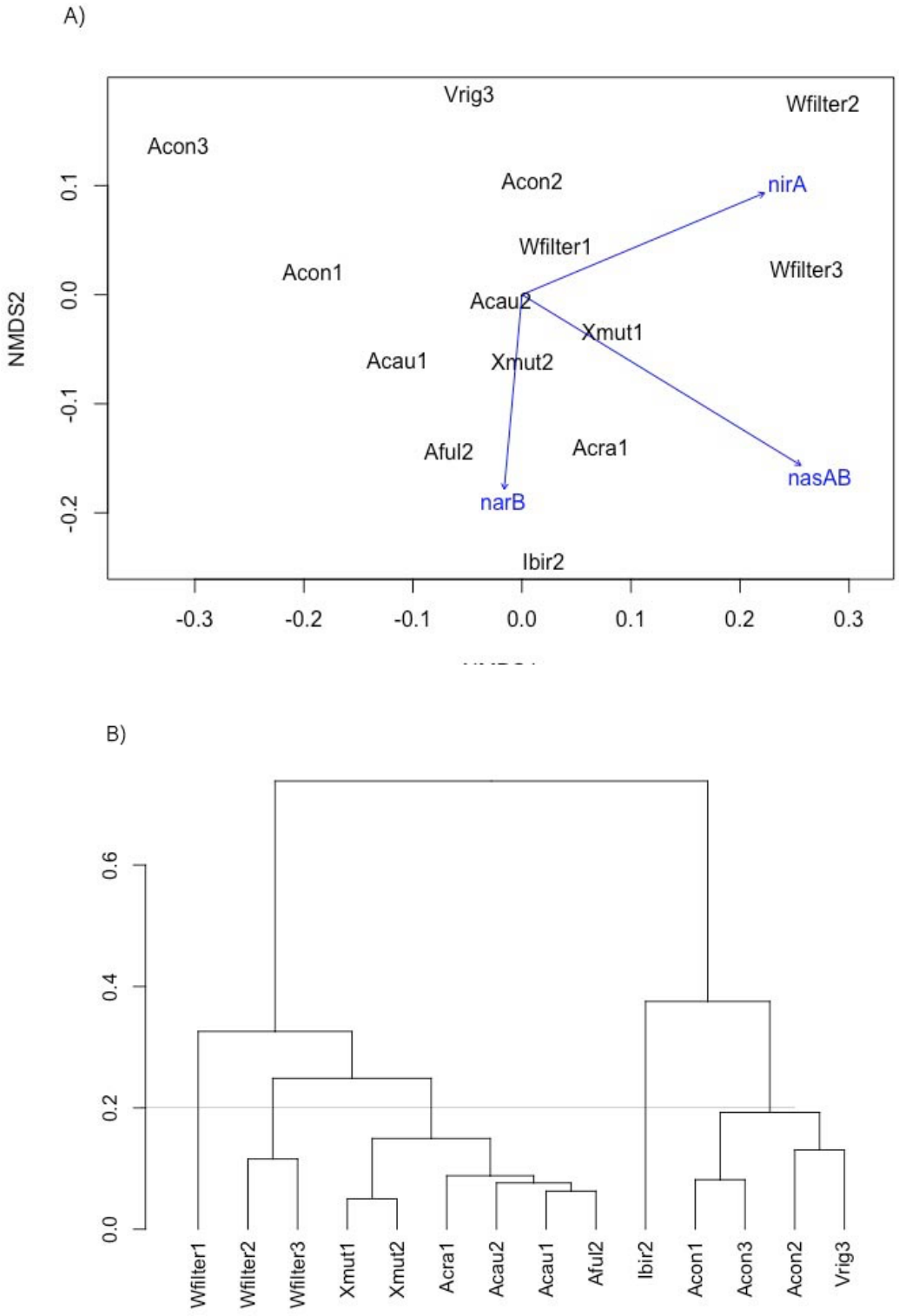


Figure 12: A) nMDS plot of assimilatory nitrate reduction genes B) A dendrogram based on Bray-Curtis dissimilarity of assimilatory nitrate reduction genes.

Discussion

Sponges are common community members for different types of marine ecosystems, and their connection to different nutrient cycles has been well-studied (Maldonado et al. 2012, Zhang et al. 2019). Common approaches to these studies include *in-situ* labeled isotope tracers to track different forms carbon and nitrogen displaced by sponge pumping action (e.g. Weisz et al. 2007, Freeman et al. 2020) along with gene-level studies that look at the metabolic functions of microorganisms who make up sponge microbiomes (reviewed in Fiore et al. 2010). The current study took a metagenomic approach to detecting nitrogen cycling potential in the microbiomes of seven common Caribbean sponge species and ambient seawater to investigate their potential as nitrogen cyclers and determine whether there was convergence or divergence in nitrogen cycling potential among these common, co-occurring species.

Using metagenome assembled genomes (MAGs) to detect nitrogen cycling potential

MAGs were assembled for this study because it gives one the opportunity to examine near-complete microbial genomes rather than relying on amplification of specific marker-genes (Pérez-Cobas et al. 2020). Considering the high-sequencing depth and coverage needed to recover whole genomes, MAGs that are high-quality represent microorganisms present in high abundances who may be important symbionts. The taxa in our samples that contributed to nitrogen cycling potential included representatives of the phyla Thaumarchaeota, Actinobacteria, Cyanobacteria, Nitrospirota, Proteobacteria, Poribacteria, SAR324, and Verrucomicrobiota, who have been recorded as community members in previous sponge microbiome studies (Weigel & Erwin 2017, Haber et al. 2021). These taxa have been credited as contributing to different nitrogen cycle processes (Zhang et al. 2013, Haber et al. 2021, Glasl et al. 2023), and our annotations for these MAGs supported this previous research and revealed potential for

nitrification, denitrification, ANAMMOX, assimilatory nitrate reduction (ANR) and dissimilatory nitrate reduction (DNR).

Haber et al. (2021) recently compared 11 high-quality archaeal genomes of Thaumarchaeota symbionts in sponges and found that nearly all of them possessed the *amoABC* gene cluster, which our 5 Thaumarchaeota MAGs found in *A. cauliformis*, *A. fulva*, *A. crassa*, and in both *X. muta* samples also possessed. Thaumarchaeota was a phylum that was proposed by Brochier-Armanet et al. (2008), and there are currently four described species within this phylum (Brochier-Armanet et al. 2011). Our Thaumarchaeota MAGs were classified down to the family Nitrosopumilaceae, who are often referred to as ammonia-oxidizing archaea (AOA) (Pester et al. 2011). AmoABC is part of the nitrification pathway, and Engelberts et al. (2020) found that this first step of nitrification facilitated by *AmoABC* is almost exclusive to Thaumarchaeota in sponges. Our data suggests that Thaumarchaeota is likely an important sponge-symbiont because it was present in four of seven species, and their recently recorded widespread distribution may imply a significant importance throughout the ocean (Zhong et al. 2020, Aylward & Santoro 2020). In addition to being recorded in sponges, Thaumarchaeota have also been found in corals (Tandon et al. 2023), and free living in the deep ocean through metagenomics studies (Zhong et al. 2020). In the open ocean, AOA carrying out ANAMMOX are also abundant in oxygen-minimum zones (Pitcher et al. 2011). *X. muta* (order Haplosclerida), and *A. cauliformis* and *A. fulva* (order Verongiida). are quite distantly related, and so recording the same archaeal symbiont in these samples could imply a deep evolutionary history between the Thaumarchaeota and certain sponge lineages.

Nitrospirota was recovered from both *A. conifera* samples, with one MAG being classified as high-quality and one as medium quality (Supplementary Table 1). This taxon

contained gene pathways for nitrification, and denitrification, showing high functional potential for nitrogen cycling. Nitrospirota has previously been detected in other sponge species used in this study such as *X. muta* (Engelberts et al. 2020) even though we did not sequence them as MAGs from our *X. muta* samples. It is possible that higher sequencing depth could have recover additional Nitrospirota MAGs in our samples, which is not out of the ordinary for metagenome studies (Pereira-Marques et al. 2019). Nitrospirota are found in many different environments such as wastewater treatment plants and terrestrial systems (Mehrani et al. 2020, Mosley et al. 2022). Their wide-spread distribution and our results suggests that Nitrospirota likely contributes to the nitrogen cycle on a global scale.

A different MAG with a similar functional potential to Nitrospirota was Bin19, which was only identified down to the class Alphaproteobacteria (Bin 19). This MAG was recovered from a single *A. cauliformis* sample. Different Alphaproteobacteria are known to contribute to nitrogen and carbon cycling (Hunter et al. 2006), and Karimi et al. (2018) recently looked at Alphaproteobacterial symbionts in sponges and found functional potentials for carbon, nitrogen, sulfur, and phosphorus metabolism. Karimi et al. (2018) detected the *NarGHI* gene in Alphaproteobacteria from two *Aplysina* species from the Mediterranean using a MAG-based approach. Our high-quality alphaproteobacterial MAG also possessed *NarGHI*, as well as *NirK* and *NosZ*. These three enzymes nearly complete the denitrification pathway, only missing the *NorBC* gene. The absence of *NorBC* means that the denitrification process stops at nitric oxide (Figure 3). There are two potential explanations for this absence: 1) *NorBC* plays a role in regulating metamorphosis in certain invertebrates, and 2) our KEGG annotations did not detect *NorBC* even though it may have been present. In support of the first explanation Ueda et al. (2016) showed that nitric oxide plays a role in inducing or repressing metamorphosis in the

marine sponge *Amphimidon queenslandica*. In addition to regulating metamorphosis, Nemoy et al (2021) suggested it could also regulate sponge regeneration. Perhaps the absence of NorBC in our MAGs is because of nitric oxide's physiological role. Indeed, nitric oxide is a well-known signaling molecule in many other invertebrates and even mammals (Jacklet 1997, Dai et al. 2013), but at present, whether it has a broader role in sponges is not well understood.

In two of our water samples and *A. fulva* the taxon classified in order SAR324 (Bin 20) possessed potential for nitrification with the *Hao* gene (Supplementary Table 4). SAR324 is a clade classified within the class Deltaproteobacteria, and studies have shown its ubiquity as a free-living organism in the ocean (Boeuf et al. 2021, Malfertheiner et al. 2022). Although not observed in our SAR324 MAG (Bin 20), Boeuf et al. (2021) detected evidence for denitrification through the presence of *NirK* (which converts nitrite to nitric oxide) and *NosZ* (nitrous oxide to nitrogen) in their metapangenomics study. SAR324's relationship with sponges, remains understudied at present, however, Cleary et al. (2023) recently noted that SAR324 seemed to be more abundant in calcareous sponges versus non-calcareous sponges using an amplicon-level approach. Given this finding, it is interesting that *A. fulva*, a Demosponge, hosted SAR324 in a large enough abundance for us to recover it as a high-quality MAG. While the connection between SAR324, sponges, and nitrogen cycling is largely unclear, our annotations suggest SAR324 may contribute to the nitrification potential of the sponge holobiont.

Two Actinobacteria classified down to the class *Acidimicrobiia* were found in *A. fulva* and seawater (Bin 2) and in *A. cauliformis* (Bin 3). Actinobacteria are known nitrogen fixers who can live in symbiosis with actinorhizorial plants in soil systems (Gtari et al. 2012), although our annotations did not reveal any nitrogen fixation pathways for these MAGs. Instead, our annotations showed the presence of the *Hao* gene in both Acitnobacteria MAGs, suggesting a

potential for nitrification within the sponge holobiont. In marine systems much of the focus on Actinobacteria has been in their production of novel natural products (Sheikh et al. 2018), which may be connected to chemical defenses that some sponges have. Indeed, both *A. cauliformis* and *A. fulva* have been recorded as possessing chemical defense systems through secondary metabolites (Thoms et al. 2006, Freeman & Gleason 2010), but it is currently unclear whether Actinobacteria symbionts are responsible for these defensive chemicals.

Cyanobacteria MAGs (Bin 12) classified down to the genus *Synechococcus* were found in *A. cauliformis*, *A. fulva*, and *X. muta*. Our cyanobacterial MAGs possessed genes for ANR, where nitrate is reduced in two steps to ammonia, by *NarB* and *NirA* respectively. There is experimental evidence that cyanobacteria in sponges can fix nitrogen by way of the acetylene reduction assay, which detects nitrogenase activity, the sole enzyme needed to convert atmospheric nitrogen to bioavailable ammonia (Wilkinson & Fay 1979). Our annotations did not reveal the presence of the nitrogenase gene cluster. The significance of ANR lies in that the ammonia resulting from the two-step reduction of nitrate can be used for amino acid or nucleotide biosynthesis by the bacteria. In bacteria, this ammonia can also be used for arginine biosynthesis. Many invertebrates cannot produce their own arginine (Song et al. 2021), and so it is possible the ANR process in sponge symbionts could benefit the host by providing a source of arginine. Interestingly, arginine is also a precursor in the synthesis of nitric oxide (Bredt 1999), which plays a role in larval settlement and metamorphosis of invertebrates as previously noted (Song et al. 2021).

Two Poribacteria MAGs were detected in our species. Bin17 found in *A. conifera*, which possessed the NarB gene from the ANR pathway. Bin18, present in *A. cauliformis*, had the Hao gene from nitrification and the NirA gene from ANR. Poribacteria are a candidate phylum

originally discovered in sponges (Fieseler et al. 2004) but have yet to be isolated in culture. Siegl et al. (2011) used pyrosequencing of single-amplified genomes (SAGs) to characterize Poribacteria from the Mediterranean sponge *Aplysina aerophoba* and detected genes for denitrification and ANR. Our annotations did not reveal potential for denitrification but did for ANR in both Poribacteria MAGs through the NarB and NirA genes. Analyses from Siegl et al. (2010) also revealed sponge-specific polyketide synthases and adhesion proteins, implying that the Poribacteria and their sponge hosts have developed symbiotic mechanisms over evolutionary time.

Verrucomicrobia MAGs were recovered from *A. cauliformis*, *X. muta*, and seawater. In *A. cauliformis* and *X. muta* the Verrucomicrobia MAG (Bin24) contained genes from several processes: *NxrAB* from the final step of denitrification, *NarGHI* from the first step of denitrification, and *NarB* along with *NarGHI* from ANR and DNR respectively. The Verrucomicrobia MAG recovered from the seawater (Bin 25) contained the NapAB gene that participates in both denitrification and DNR, and the NirA gene from ANR. Compared to the other MAGs in this study with nitrogen cycling potential, Verrucomicrobia are not as well studied. Verrucomicrobia is quite common in soil and marine environments (Bergmann et al. 2011), and Khadem et al. (2010) used the acetylene reduction assay and soil growth experiments to show that their Verrucomicrobial strain was capable of fixing nitrogen. Sizikov et al. (2020) recently undertook a characterization of Verrucomicrobia found in the Mediterranean sponge *Petrosia ficiformis* and seawater from the East Mediterranean Sea and found an enrichment of toxin-antitoxin systems in the symbiotic strain of Verrucomicrobia, suggesting that the sponge host may use secondary metabolites from the symbiont for chemical defense. Indeed, a handful of other studies have focused on the geographical distribution and potential for secondary

metabolites related to Verrucomicrobia in sponges (Kasai et al. 2007, Ruocco et al. 2021), but their connection to nitrogen fluxes in context of the sponge holobiont remains unclear. It seems this symbiont has not been successfully isolated in culture as of the present study, and perhaps similar nitrogen fixation techniques such as those used in soil systems may be necessary to further elucidate the nitrogen cycling potential of this seemingly ubiquitous prokaryote.

While we chose not to infer metabolic potential from medium quality MAGs due to them not meeting the high-quality threshold, although the assigned taxonomy for these MAGs can still be informative. For many of the MAGs with potential for nitrogen cycling, such as the taxa SAR324, Thaumarchaeota, and Nitrospirota, we recovered medium quality MAGs in addition to the ones discussed (Supplementary Table 1). Annotation profiles for these medium quality examples were nearly identical to their high-quality counterparts, thus bolstering the nitrogen cycling potential of the host-symbiont complex (the holobiont), assuming the taxonomic assignments are accurate.

Using MAGs to study the metabolic potential of highly abundant symbionts can be quite informative given adequate quality of the MAGs. It is possible to infer entire pathways such as denitrification in our high-quality alphaproteobacterial MAG. In our high-quality MAGs the nitrification and denitrification pathways were the most well represented, and the taxa facilitating these processes were spread between different species. In some cases, such as with our Nitrospirota MAG, it was detected more than once although not always meeting our quality criteria to include in the results. Our data suggests some taxa such as Thaumarchaeota, Nitrospirota, and Alphaproteobacteria do play a key role in the cycling of nitrogen compounds in different sponge species. When looking at the species these notable taxa were present in, *A. cauliformis*, *A. fulva*, *A. crassa*, and *X. muta* all hosted potential for nitrification through the

archaea Thaumarchaeota, suggesting functional convergence towards nitrification through a symbiosis with this archaeon. However, categorizing divergence and convergence is more nuanced than this single example, given that the potential for ANAMMOX was not isolated to one prokaryote, but seen in Thaumarchoeta, Nitrosiporota, and Alphaproteobacteria and found in multiple sponge species (Supplementary Table 4). This observation shows some divergence in symbiont capability across sponge species.

Phylogenetic diversity of nitrogen cycle genes

For the processes analyzed here: nitrification, denitrification, and DNR, samples from the same species grouped more closely together than samples of different species, but these groupings varied by process. For example, in our visualizations of nitrification genes (Figure 9) *A. cauliformis* samples grouped separately in contrast to denitrification, where they grouped very closely (Figure 10). Another example of this shifting of groups can be seen with *X. muta,* where both samples grouped separately in DNR visualizations (Figure 11) but together in ANR (Figure 12). This may suggest that while the diversity of nitrogen cycle genes is quite similar in samples from the same species, which is in line with the idea of species-specific microbiomes, there can also be variations within species, possibly due to genetic diversity of the host (et al. 2022). Slight variations within species could be due to things such as environmental gradients of nutrients where samples are collected (Zhang et al. 2019). It appears that our species have varying functional potentials for nitrogen cycling, with slight variations within species driven by specific genes.

Nitrification is one of the most well-studied of the nitrogen cycle processes and is relevant in not just marine environments, but soil and freshwater systems as well (Schmidt 1982, Bowden 1986). Nitrification involves the oxidation of ammonia to nitrate, where ammonia is

used as an energy source for ammonia-oxidizing bacteria (AOB) and ammonia-oxidizing archaea (AOA). Both AOB and AOA have been detected in sponges from deep-sea, tropical, and arctic environments (Ribes et al. 2012, Cardoso et al. 2013, Tian et al. 2016), and so it is likely that the sponge microbiome contributes to local nitrogen fluxes in these various environments. Nitrification has also been observed in in-situ and incubation experiments that used stable isotopes of nitrogen to track ammonia-oxidation occurring in sponges, which included three of our species: *A cauliformis*, *V. rigida*, and *X. muta* (Southwell et al. 2008). While both *A. cauliformis* and *X. muta* hosted nitrification genes, they did not group as closely as a pair such as *A. cauliformis* and *A. conifera*, and perhaps they may use slightly different versions of genes. In our data, the most widely seen nitrifier was Thaumarchaeota, which is in concordance with other studies on nitrogen fluxes in sponges who have measured a high abundance of AOA (Cardoso et al. 2013), and specifically Engelberts et al. (2020) who noted the link between the first step of nitrification facilitated by Thaumarchaeota.

Denitrification can be thought of as completing the cycle of taking dissolved organic nitrogen (DIN) and returning it to the atmosphere as nitrogen gas starting with nitrate, the product of nitrification. In our data, genes from denitrification drove the most dissimilarities in *A. cauliformis*, *A. conifera*, *A. crasssa*, *A. fulva*, and *X. muta*. The presence of denitrifying bacteria has been previously detected in sponges in stable-isotope incubation experiments (Shläppy et al. 2010) and gene-amplification studies (Rooks et al. 2020). Zhang et al. (2013) looked at phylogenetic diversity of key genes in six sponges and found that these genes likely came from an Alphaproteobacteria. In our own data, we found those genes in Alphaproteobacteria and Nitrospirota. Shläppy et al. (2010) noted in their detection of nitrogen gas from sponge incubation experiments that nitrogen gas produced came almost exclusively

from denitrification and not from ANAMMOX, due to how they labeled nitrogen compounds in their experimental design. This and our data suggest that the connection between sponges' production of nitrogen gas is likely facilitated by denitrifying bacteria, and that they are likely part of the Proteobacteria classes. Fiore et al. (2013) used stable isotopes to track dissolved inorganic nitrogen in X. muta and found evidence for denitrification, which is in line with our genes extracted from MAGs and contigs. It should be noted that at the end of the denitrification pathway, nitrogen is no longer bioavailable, and unless nitrogen fixation is also detected it is assumed the nitrogen has left the system until it is fixed by nitrogen fixing bacteria. Detecting both nitrification and nitrogen fixation within a sponge would imply that the cycle is "closed", and nitrogen could be fixed and converted back into its bioavailable states.

The final process that drove dissimilarity among species was dissimilatory nitrate reduction (DNR), and these genes occurred in *A. cauliformis*, *A. conifera*, *A. crasssa*, *A. fulva*, and *X. muta*. O'brien et al. (2023) recently used metagenomics and metatranscriptomics to investigate symbiont pathways in common coral reef sponges and detected high expression levels for nitrate reduction through transcripts for the NirD gene which suggested that their samples were reducing nitrates at the time of collection. Although our study was not paired with analyses of mRNA transcripts, it seems that the presence of NirD suggests the metabolic potential of DNR in our samples. This could be important for the sponge host as it produces bioavailable ammonia that could be involved in production of amino acids (Mohiuddin & Khattar 2019). Detection of DNR in these different species from the Caribbean suggests that these common sponge holobionts could act as sources of bioavailable nitrogen by generating ammonia.

Our phylogenetic analyses revealed that nitrification, denitrification, and DNR genes drove the most dissimilarities in the species all sponge species except *I. birotualta* and *V. rigida* with nitrification being the most prevalent pathway. A high functional potential for nitrification implies symbionts could be removing ammonia, possibly excreted by the host. Our data shows that Thaumarchaeota are key players in sponge-mediated nitrification, concordant with previous research (Brochier-Armanet et al. 2012, Haber et al. 2021). Denitrification also drove dissimilarities in our data, supporting previous research that detected nitrogen gas "exhaled" by sponges (Shläppy et al. 2010, Rooks et al. 2020). Nitrogen fixation, which could fix the nitrogen released by denitrifying bacteria, was not detected in our samples but has been detected in cyanobacteria hosted by sponges using the acetylene reduction assay (Wilkinson & Fay 1979). Our data suggests that while many of these species overlap in some of the taxa they host, there was convergence across disparate sponge species in terms of nitrification and denitrification, since five of our seven species hosted genes for these important processes. A higher replication of species and higher sequencing depth may help in better identifying important sponge-symbionts that contribute to these significant processes, further establishing some species as sources or sinks of dissolved inorganic nitrogen and dissolved organic nitrogen. Future studies coupling metatranscriptomics and stable-isotope components will be important in narrowing down the functional potential and solidifying the connection between sponges, their microbiomes, and nitrogen cycling.

Contigs and MAGs to detect functional potential: is one better than the other?

Annotating MAGs gives one a "whole genome" view of functional potential given sufficient MAG quality, which was important in this study because we were trying to detect various nitrogen cycle genes. Keeping in mind the high sequencing depth and coverage needed

to recover MAGs, analyses of recovered genomes can only give information over the highly abundant microbes. This is not to say that contigs should be disregarded in metagenomic analyses in favor of MAGs, but that they are a more "zoomed out" in contrast to MAGs that home in on specific taxa and their full metabolic potentials. For example, both contigs and MAGs showed that nitrification and denitrification were the most prevalent pathways in our species. One difference between the contigs and MAGs was in a process such as ANAMMOX, where there were a few annotations for ANAMMOX genes in the contigs (Supplementary Table 4), and even less or none in the MAGs. Perhaps the discrepancies in the annotations between both data types suggests that, at the contig level, some pathways may be facilitated by many symbionts, and thus the annotations were not recoverable due to the noise of all the shorter reads. In contrast, the processes of nitrification and denitrification were present in five of twenty-five unique MAGs (Supplementary Table 4), suggesting that these processes occur at higher rates in relatively fewer taxa, allowing our sequencing depth to recover several high-quality MAGs from these prominent community members once some contigs are filtered out. Metagenomics is a field that is still emerging in terms of methodologies, challenges, and implications (Setubal 2021), but has been useful in detecting and investigating previously uncultured microbes that play key roles in global nutrient cycles in marine sponges. Annotating high-quality contigs can still be useful to quickly profile a "big picture" of nitrogen cycling potential, but taking the time to assemble MAGs allows one to home in on specific taxa and genes if interest. Future studies could focus on further identification and profiling of nitrification and denitrification mediated by sponge microbiomes and their contributions to local nitrogen fluxes on coral reefs. As the idea of what a healthy reef is continues to change in the coming decades, investigating the ecological role of

common reef animals such as sponges will help paint a better picture of the future of coral reefs and other marine ecosystems.

REFERENCES

Alexander, B. E., Liebrand, K., Osinga, R., van der Geest, H. G., Admiraal, W., Cleutjens, J. P., ... & de Goeij, J. M. (2014). Cell turnover and detritus production in marine sponges from tropical and temperate benthic ecosystems. *PLoS One*, *9*(10), e109486.

Apprill, A. (2017). Marine animal microbiomes: toward understanding host–microbiome interactions in a changing ocean. *Frontiers in Marine Science*, *4*, 222. 10.3389/fmars.2017.00222

Atkinson M.K., Falter J. L. (2003). *Coral Reefs* in *Biogeochemistry of Marine Systems*, K. D. Black, G. B. Shimmield, Eds. Blackwell Publishing. 40-64.

Aylward, F. O., & Santoro, A. E. (2020). Heterotrophic Thaumarchaea with small genomes are widespread in the dark ocean. *MSystems*, *5*(3), 10-1128.

Azam, F., Fenchel, T., Field, J. G., Gray, J. S., Meyer-Reil, L. A., & Thingstad, F. (1983). The ecological role of water-column microbes in the sea. *Marine ecology progress series*, *10,* 257-263. 10.3354/meps010257

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., ... & Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology*, *19*(5), 455-477.

Bell, J. J., Davy, S. K., Jones, T., Taylor, M. W., & Webster, N. S. (2013). Could some coral reefs become sponge reefs as our climate changes?. *Global change biology*, *19*(9), 2613-2624.

Bergmann, G. T., Bates, S. T., Eilers, K. G., Lauber, C. L., Caporaso, J. G., Walters, W. A., ... & Fierer, N. (2011). The under-recognized dominance of Verrucomicrobia in soil bacterial communities. *Soil Biology and Biochemistry*, *43*(7), 1450-1455.

Biers, E. J., Sun, S., & Howard, E. C. (2009). Prokaryotic genomes and diversity in surface ocean waters: interrogating the global ocean sampling metagenome. Applied and environmental microbiology, 75(7), 2221-2229.

Boeuf, D., Eppley, J. M., Mende, D. R., Malmstrom, R. R., Woyke, T., & DeLong, E. F. (2021). Metapangenomics reveals depth-dependent shifts in metabolic potential for the ubiquitous marine bacterial SAR324 lineage. *Microbiome*, *9*, 1-18.

Bordenstein, S. R., & Theis, K. R. (2015). Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS biology*, *13*(8), e1002226.

Bowden, W. B. (1986). Nitrification, nitrate reduction, and nitrogen immobilization in a tidal freshwater marsh sediment. *Ecology*, *67*(1), 88-99.

Bowers, R. M., Kyrpides, N. C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B. K., ... & Woyke, T. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature biotechnology*, *35*(8), 725-731.

Bristow, L. A., Mohr, W., Ahmerkamp, S., & Kuypers, M. M. (2017). Nutrients that limit growth in the ocean. *Current Biology*, *27*(11), R474-R478. 10.1016/j.cub.2017.03.030

Brochier-Armanet, C., Boussau, B., Gribaldo, S., & Forterre, P. (2008). Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nature Reviews Microbiology*, *6*(3), 245-252.

Brochier-Armanet, C., Gribaldo, S., & Forterre, P. (2012). Spotlight on the Thaumarchaeota. *The ISME journal*, *6*(2), 227-230.

Brown, B. E. (1997). Coral bleaching: causes and consequences. *Coral reefs*, *16*(1), S129-S138.

Brown, B. E., & Ogden, J. C. (1993). Coral bleaching. *Scientific American*, *268*(1), 64-70.

Bythell, J. C., & Wild, C. (2011). Biology and ecology of coral mucus release. *Journal of Experimental Marine Biology and Ecology*, *408*(1-2), 88-93.

Cardoso, J. F., van Bleijswijk, J. D., Witte, H., & van Duyl, F. C. (2013). Diversity and abundance of ammonia-oxidizing Archaea and Bacteria in tropical and cold-water coral reef sponges. *Aquatic Microbial Ecology*, *68*(3), 215-230.

Cavanaugh, C. M. (1994). Microbial symbiosis: patterns of diversity in the marine environment. *American Zoologist*, *34*(1), 79-89. 10.1093/icb/34.1.79

Cebrian, E., Uriz, M. J., Garrabou, J., & Ballesteros, E. (2011). Sponge mass mortalities in a warming Mediterranean Sea: are cyanobacteria-harboring species worse off?. *PLoS One*, *6*(6), e20211.

Cesar, H. S. (2002). Coral reefs: their functions, threats and economic value.

Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., & Parks, D. H. (2020). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database.

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884-i890.

Cleary, D. F. R., de Voogd, N. J., Stuij, T. M., Swierts, T., Oliveira, V., Polónia, A. R. M., ... & Coelho, F. J. R. C. (2023). A study of sponge symbionts from different light habitats. *Microbial Ecology*, 1-19.

Compeau, P. E., Pevzner, P. A., & Tesler, G. (2011). How to apply de Bruijn graphs to genome assembly. *Nature biotechnology*, *29*(11), 987-991. 10.1038/nbt.2023

Dai, Z., Wu, Z., Yang, Y., Wang, J., Satterfield, M. C., Meininger, C. J.

Daims, H., & Wagner, M. (2018). Nitrospira. *Trends in Microbiology*, *26*(5), 462-463.

Daims, H., Nielsen, P. H., Nielsen, J. L., Juretschko, S., & Wagner, M. (2000). Novel Nitrospira-like bacteria as dominant nitrite-oxidizers in biofilms from wastewater treatment plants: diversity and in situ physiology. *Water science and technology*, *41*(4-5), 85-90.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., ... & Li, H. (2021). Twelve years of SAMtools and BCFtools. *Gigascience*, *10*(2), giab008.

Darwin, C. (2020). The structure and distribution of coral reefs. In *The Structure and Distribution of Coral Reefs*. University of California Press.

De Goeij, J. M., Van Oevelen, D., Vermeij, M. J., Osinga, R., Middelburg, J. J., de Goeij, A. F., & Admiraal, W. (2013). Surviving in a marine desert: the sponge loop retains resources within coral reefs. *Science*, *342*(6154), 108-110. 10.1126/science.1241981

De Menezes, T. A., De Freitas, M. A., Lima, M. S., Soares, A. C., Leal, C., Busch, M. D. S., ... & Thompson, F. L. (2022). Fluxes of the Amazon River plume nutrients and microbes into marine sponges. *Science of The Total Environment*, *847*, 157474.

Diaz, M. C., & Ward, B. B. (1997). Sponge-mediated nitrification in tropical benthic communities. *Marine Ecology Progress Series*, *156*, 97-107.

Douglas, A. E. (2003). Coral bleaching—how and why?. *Marine pollution bulletin*, *46*(4), 385-392.

Dudgeon, S. R., Aronson, R. B., Bruno, J. F., & Precht, W. F. (2010). Phase shifts and stable states on coral reefs. *Marine Ecology Progress Series*, *413*, 201-216.

Easson, C. G., & Thacker, R. W. (2014). Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges. *Frontiers in microbiology*, *5*, 532. 10.3389/fmicb.2014.00532

Edge, T. A., Baird, D. J., Bilodeau, G., Gagné, N., Greer, C., Konkin, D., ... & Macklin, J. (2020). The Ecobiomics project: Advancing metagenomics assessment of soil health and freshwater quality in Canada. *Science of The Total Environment*, *710*, 135906.

Engelberts, J. P., Robbins, S. J., de Goeij, J. M., Aranda, M., Bell, S. C., & Webster, N. S. (2020). Characterization of a sponge microbiome using an integrative genome-centric approach. *The ISME journal*, *14*(5), 1100-1110. 10.1038/s41396-020-0591-9

Engelberts, J. P., Robbins, S. J., de Goeij, J. M., Aranda, M., Bell, S. C., & Webster, N. S. (2020). Characterization of a sponge microbiome using an integrative genome-centric approach. *The ISME Journal*, *14*(5), 1100-1110.

Erwin, P. M., & Thacker, R. W. (2008). Cryptic diversity of the symbiotic cyanobacterium Synechococcus spongiarum among sponge hosts. *Molecular Ecology*, *17*(12), 2937-2947.

Falkowski, P. G., Fenchel, T., & Delong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *science*, *320(5879)*, 1034-1039. 10.1126/science.1153213

Fenchel, T. (2008). The microbial loop–25 years later. *Journal of Experimental Marine Biology and Ecology*, *366*(1-2), 99-103. 10.1016/j.jembe.2008.07.013

Fenchel, Tom & Field, John & Gray, J.S. & Meyer-Reil, L.A. & Thingstad, T.Frede. (1983). The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.*. 10. 257-263.

Fieseler, L., Horn, M., Wagner, M., & Hentschel, U. (2004). Discovery of the novel candidate phylum "Poribacteria" in marine sponges. *Applied and environmental microbiology*, *70*(6), 3724-3732.

Fiore, C. L., Baker, D. M., & Lesser, M. P. (2013). Nitrogen biogeochemistry in the Caribbean sponge, Xestospongia muta: a source or sink of dissolved inorganic nitrogen?. *PLoS One*, *8*(8), e72961.

Fiore, C. L., Jarett, J. K., Olson, N. D., & Lesser, M. P. (2010). Nitrogen fixation and nitrogen transformations in marine symbioses. *Trends in microbiology*, *18*(10), 455-463

Fowler, D., Coyle, M., Skiba, U., Sutton, M. A., Cape, J. N., Reis, S., ... & Voss, M. (2013). The global nitrogen cycle in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1621), 20130164.

Frazão Santos, C., Agardy, T., Andrade, F., Calado, H., Crowder, L. B., Ehler, C. N., ... & Rosa, R. (2020). Integrating climate change in ocean planning. *Nature Sustainability*, *3*(7), 505-516.

Freeman, C. J., & Gleason, D. F. (2010). Chemical defenses, nutritional quality, and structural components in three sponge species: *Ircinia felix*, *I. campana*, and *Aplysina fulva*. *Marine Biology*, *157*, 1083-1093.

Freeman, C. J., & Thacker, R. W. (2011). Complex interactions between marine sponges and their symbiotic microbial communities. *Limnology and Oceanography*, *56*(5), 1577-1586. 10.4319/lo.2011.56.5.1577

Freeman, C. J., Easson, C. G., Matterson, K. O., Thacker, R. W., Baker, D. M., & Paul, V. J. (2020). Microbial symbionts and ecological divergence of Caribbean sponges: a new perspective on an ancient association. *The ISME Journal*, *14*(6), 1571-1583.

Freeman, C. J., Thacker, R. W., Baker, D. M., & Fogel, M. L. (2013). Quality or quantity: is nutrient transfer driven more by symbiont identity and productivity than by symbiont abundance?. *The ISME journal*, *7*(6), 1116-1125.

Freeman, C. J., Easson, C. G., Fiore, C. L., & Thacker, R. W. (2021). Sponge–microbe interactions on coral reefs: multiple evolutionary solutions to a complex environment. *Frontiers in marine science*, *8*, 705053.

Guan, Y., Hohn, S., & Merico, A. (2015). Suitable environmental ranges for potential coral reef habitats in the tropical ocean. *PloS one*, *10*(6), e0128831.

Glasl, B., Herndl, G. J., & Frade, P. R. (2016). The microbiome of coral surface mucus has a key role in mediating holobiont health and survival upon disturbance. *The ISME journal*, *10*(9), 2280-2292.

Glasl, B., Luter, H. M., Damjanovic, K., Kitzinger, K., Mueller, A. J., Mahler, L., ... & Wagner, M. (2023). Novel nitrifying symbiont lineages are vertically inherited and widespread in marine sponges. *bioRxiv*, 2023-12.

Gruber, N. (2008). *The marine nitrogen cycle: overview and challenges*. in *Nitrogen in the marine environment*, *2*, 1-50.

Gtari, M., Ghodhbane-Gtari, F., Nouioui, I., Beauchemin, N., & Tisa, L. S. (2012). Phylogenetic perspectives of nitrogen-fixing actinobacteria. *Archives of microbiology*, *194*, 3-11.

Haber, M., Burgsdorf, I., Handley, K. M., Rubin-Blum, M., & Steindler, L. (2021). Genomic insights into the lifestyles of Thaumarchaeota inside sponges. *Frontiers in Microbiology*, *11*, 622824.

Harris, D. L., Rovere, A., Casella, E., Power, H., Canavesio, R., Collin, A., ... & Parravicini, V. (2018). Coral reef structural complexity provides important coastal protection from waves under rising sea levels. *Science advances*, *4*(2), eaao4350.

Hatcher, B. G. (1990). Coral reef primary productivity. A hierarchy of pattern and process. *Trends in Ecology & Evolution*, *5*(5), 149-155.

Helman, Y., Natale, F., Sherrell, R. M., LaVigne, M., Starovoytov, V., Gorbunov, M. Y., & Falkowski, P. G. (2008). Extracellular matrix production and calcium carbonate precipitation by coral cells in vitro. *Proceedings of the National Academy of Sciences*, *105*(1), 54-58.

Heravi, F. S., Zakrzewski, M., Vickery, K., & Hu, H. (2020). Host DNA depletion efficiency of microbiome DNA enrichment methods in infected tissue samples. *Journal of microbiological methods*, *170*, 105856.

Hoegh-Guldberg Ove (1999) Climate change, coral bleaching and the future of the world's coral reefs. *Marine and Freshwater Research,* 50(8), 839-866. 10.1071/MF99078

Hoegh-Guldberg, O., & Bruno, J. F. (2010). The impact of climate change on the world's marine ecosystems. *Science*, *328*(5985), 1523-1528.

Hönisch, B., Ridgwell, A., Schmidt, D. N., Thomas, E., Gibbs, S. J., Sluijs, A., ... & Williams, B. (2012). The geological record of ocean acidification. *science*, *335(6072)*, 1058-1063. https://doi.org/10.1126/science.1208277

Hughes, T.P. (1994). Catastrophes, phase shifts, and large-scale degradation of a Caribbean coral reef. *Science*, *265(5178)*, 1547-1551. 10.1126/science.265.5178.1547

Hughes, T. P., & Connell, J. H. (1999). Multiple stressors on coral reefs: A long-term perspective. *Limnology and oceanography*, *44*(3part2), 932-940.

Hunter, E. M., Mills, H. J., & Kostka, J. E. (2006). Microbial community diversity associated with carbon and nitrogen cycling in permeable shelf sediments. *Applied and environmental microbiology*, *72*(9), 5689-5701.

Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics*, *11*, 1-11.

Jacklet, J. W. (1997). Nitric oxide signaling in invertebrates. *Invertebrate Neuroscience*, *3*, 1-14.

Jackrel, S. L., Yang, J. W., Schmidt, K. C., & Denef, V. J. (2021). Host specificity of microbiome assembly and its fitness effects in phytoplankton. *The ISME journal*, *15*(3), 774-788.

Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of molecular biology*, *428*(4), 726-731.

Khadem, A. F., Pol, A., Jetten, M. S., & Op den Camp, H. J. (2010). Nitrogen fixation by the verrucomicrobial methanotroph 'Methylacidiphilum fumariolicum'SolV. *Microbiology*, *156*(4), 1052-1059.

Karimi, E., Slaby, B. M., Soares, A. R., Blom, J., Hentschel, U., & Costa, R. (2018). Metagenomic binning reveals versatile nutrient cycling and distinct adaptive features in alphaproteobacterial symbionts of marine sponges. *FEMS Microbiology Ecology*, *94*(6), fiy074.

Kartal, B., Kuenen, J. V., & Van Loosdrecht, M. C. M. (2010). Sewage treatment with anammox. *Science*, *328*(5979), 702-703.

Kasai, H., Katsuta, A., Sekiguchi, H., Matsuda, S., Adachi, K., Shindo, K., ... & Shizuri, Y. (2007). Rubritalea squalenifaciens sp. nov., a squalene-producing marine bacterium belonging to subdivision 1 of the phylum 'Verrucomicrobia'. *International journal of systematic and evolutionary microbiology*, *57*(7), 1630-1634.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, *30*(4), 772-780.

Kelly, J. B., Carlson, D. E., Low, J. S., & Thacker, R. W. (2022). Novel trends of genome evolution in highly complex tropical sponge microbiomes. *Microbiome*, *10*(1), 164.

Knowlton, N., & Jackson, J. B. (1994). New taxonomy and niche partitioning on coral reefs: jack of all trades or master of some?. *Trends in Ecology & Evolution*, *9*(1), 7-9.

Knowlton, N., Brainard, R.E., Fisher, R., Moews, M., Plaisance, L. and Caley, M.J. (2010). *Coral Reef Biodiversity* in *Life in the World's Oceans: Diversity, Distribution, and Abundance*. A.D. McIntyre (Ed.) 10.1002/9781444325508.

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, *35*(21), 4453-4455.

LaJeunesse, T. C., Parkinson, J. E., Gabrielson, P. W., Jeong, H. J., Reimer, J. D., Voolstra, C. R., & Santos, S. R. (2018). Systematic revision of Symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Current Biology*, *28*(16), 2570-2580.

Larkum, A. W. D., Kennedy, I. R., & Muller, W. J. (1988). Nitrogen fixation on a coral reef. *Marine Biology*, *98*, 143-155.

Lee, Y. K., Lee, J. H., & Lee, H. K. (2001). Microbial symbiosis in marine sponges. *Journal of Microbiology*, *39*(4), 254-264.

Lesser, M. P. (1997). Oxidative stress causes coral bleaching during exposure to elevated temperatures. *Coral reefs*, *16*, 187-192.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094-3100.

Li, Z., Wang, Y., Li, J., Liu, F., He, L., He, Y., & Wang, S. (2016). Metagenomic analysis of genes encoding nutrient cycling pathways in the microbiota of deep-sea and shallow-water sponges. *Marine Biotechnology*, *18*(6), 659-671. 10.1007/s10126-016-9725-5

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic acids research*, *44*(D1), D457-D462.

Madsen E. L. (2011). Microorganisms and their roles in fundamental biogeochemical cycles. *Current opinion in biotechnology*, *22*(3), 456–464.

Maldonado, M., Ribes, M., & van Duyl, F. C. (2012). Nutrient fluxes through sponges: biology, budgets, and ecological implications. *Advances in marine biology*, *62*, 113-182.

Malfertheiner, L., Martínez-Pérez, C., Zhao, Z., Herndl, G. J., & Baltar, F. (2022). Phylogeny and metabolic potential of the candidate phylum SAR324. *Biology*, *11*(4), 599.

McManus, J. W., & Polsenberg, J. F. (2004). Coral–algal phase shifts on coral reefs: ecological and environmental aspects. *Progress in Oceanography*, *60*(2-4), 263-279. 10.1016/j.pocean.2004.02.014

McMurray, S. E., Stubler, A. D., Erwin, P. M., Finelli, C. M., & Pawlik, J. R. (2018). A test of the sponge-loop hypothesis for emergent Caribbean reef sponges. *Marine Ecology Progress Series*, *588*, 1-14. 10.3354/meps12466

Mehrani, M. J., Sobotka, D., Kowal, P., Ciesielski, S., & Makinia, J. (2020). The occurrence and role of Nitrospira in nitrogen removal systems. *Bioresource technology*, *303*, 122936.
Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature communications*, *7*(1), 1-9.

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010, November). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *2010 gateway computing environments workshop (GCE)* (pp. 1-8). Ieee.

Mohiuddin, S. S., & Khattar, D. (2019). Biochemistry, ammonia.

Moran, N. A., & Sloan, D. B. (2015). The hologenome concept: helpful or hollow?. *PLoS biology*, *13*(12), e1002311.

Mosley, O. E., Gios, E., Close, M., Weaver, L., Daughney, C., & Handley, K. M. (2022). Nitrogen cycling and microbial cooperation in the terrestrial subsurface. *The ISME Journal*, *16*(11), 2561-2573.

Nemoy, P., Spanier, E., & Angel, D. L. (2021). Nitrification activity of the sponge Chondrosia reniformis under Elevated concentrations of ammonium. *Frontiers in Marine Science*, *7*, 607979.

O'Brien, P. A., Tan, S., Frade, P. R., Robbins, S. J., Engelberts, J. P., Bell, S. C., ... & Bourne, D. G. (2023). Validation of key sponge symbiont pathways using genome-centric metatranscriptomics. *Environmental Microbiology*.

Pandolfi, J. M., Bradbury, R. H., Sala, E., Hughes, T. P., Bjorndal, K. A., Cooke, R. G., ... & Jackson, J. B. (2003). Global trajectories of the long-term decline of coral reef ecosystems. *Science*, *301*(5635), 955-958.

Pandolfi, J. M. (2011). The paleoecology of coral reefs. *Coral reefs: an ecosystem in transition*, 13-24.

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research*, *25*(7), 1043-1055. 10.1101/gr.186072.114

Pereira-Marques, J., Hout, A., Ferreira, R. M., Weber, M., Pinto-Ribeiro, I., Van Doorn, L. J., ... & Figueiredo, C. (2019). Impact of host DNA and sequencing depth on the taxonomic resolution of whole metagenome sequencing for microbiome analysis. *Frontiers in microbiology*, *10*, 1277.

Pérez-Cobas, A. E., Gomez-Valero, L., & Buchrieser, C. (2020). Metagenomic approaches in microbial ecology: an update on whole-genome and marker gene sequencing analyses. *Microbial genomics*, *6*(8).

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

Pitcher, A., Villanueva, L., Hopmans, E. C., Schouten, S., Reichart, G. J., & Sinninghe Damsté, J. S. (2011). Niche segregation of ammonia-oxidizing archaea and anammox bacteria in the Arabian Sea oxygen minimum zone. *The ISME journal*, *5*(12), 1896-1904.

Podell, S., Blanton, J. M., Neu, A., Agarwal, V., Biggs, J. S., Moore, B. S., & Allen, E. E. (2019). Pangenomic comparison of globally distributed Poribacteria associated with sponge hosts and marine particles. *The ISME journal*, *13*(2), 468-481.

Prosser, J. (1990). Autotrophic nitrification in bacteria. *Advances in microbial physiology*, *30*, 125-181.

Reveillaud, J., Maignien, L., Murat Eren, A., Huber, J. A., Apprill, A., Sogin, M. L., & Vanreusel, A. (2014). Host-specificity among abundant and rare taxa in the sponge microbiome. *The ISME journal*, *8*(6), 1198–1209. 10.1038/ismej.2013.227

Ribes, M., Jiménez, E., Yahel, G., López-Sendino, P., Diez, B., Massana, R., ... & Coma, R. (2012). Functional convergence of microbes associated with temperate marine sponges. *Environmental microbiology*, *14*(5), 1224-1239.

Rocca, J. D., Hall, E. K., Lennon, J. T., Evans, S. E., Waldrop, M. P., Cotner, J. B., ... & Wallenstein, M. D. (2015). Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed. *The ISME journal*, *9*(8), 1693-1699.

Rooks, C., Fang, J. K. H., Mørkved, P. T., Zhao, R., Rapp, H. T., Xavier, J. R., & Hoffmann, F. (2020). Deep-sea sponge grounds as nutrient sinks: denitrification is common in boreo-Arctic sponges. *Biogeosciences*, *17*(5), 1231-1245.

Ruocco, N., Esposito, R., Bertolino, M., Zazo, G., Sonnessa, M., Andreani, F., ... & Costantini, M. (2021). A metataxonomic approach reveals diversified bacterial communities in antarctic sponges. *Marine drugs*, *19*(3), 173.

Schläppy, M. L., Schöttner, S. I., Lavik, G., Kuypers, M. M., de Beer, D., & Hoffmann, F. (2010). Evidence of nitrification and denitrification in high and low microbial abundance sponges. *Marine biology*, *157*, 593-602.

Schlesinger, W. H., & Bernhardt, E. S. (2013). *Biogeochemistry: An Analysis of Global Change.* (3rd ed). Elsevier.

Schmidt, E. L. (1982). Nitrification in soil. *Nitrogen in agricultural soils*, *22*, 253-288.

Setubal, J. C. (2021). Metagenome-assembled genomes: concepts, analogies, and challenges. *Biophysical Reviews*, *13*(6), 905-909.

Sheik, C. S., Jain, S., & Dick, G. J. (2014). Metabolic flexibility of enigmatic SAR 324 revealed through metagenomics and metatranscriptomics. *Environmental microbiology*, *16*(1), 304-317.

Sheikh, M. A., Rathore, D. S., Gohel, S. D., & Singh, S. P. (2018). Marine Actinobacteria Associated With The Invertebrate Hosts: A Rich Source Of Bioactive Compounds: A Review. *Journal of Cell & Tissue Research*, *18*(1).

Sheppard, C.R.C., Davy S.K., Pilling, G.M., Graham, N.A.J. (2018). *The Biology of Coral Reefs.* Oxford University Press. (ed. 2).

Sieber, P., Platzer, M., Schuster S. (2018). The Definition of Open Reading Frame Revisited. *Trends in Genetics* 34, 167-170. 10.1016/j.tig.2017.12.009

Siegl, A., Kamke, J., Hochmuth, T., Piel, J., Richter, M., Liang, C., ... & Hentschel, U. (2011). Single-cell genomics reveals the lifestyle of Poribacteria, a candidate phylum symbiotically associated with marine sponges. *The ISME journal*, *5*(1), 61-70.

Sizikov, S., Burgsdorf, I., Handley, K. M., Lahyani, M., Haber, M., & Steindler, L. (2020). Characterization of sponge-associated Verrucomicrobia: microcompartment-based sugar utilization and enhanced toxin–antitoxin modules as features of host-associated Opitutales. *Environmental Microbiology*, *22*(11), 4669-4688.

Song, H., Hewitt, O. H., & Degnan, S. M. (2021). Arginine biosynthesis by a bacterial symbiont enables nitric oxide production and facilitates larval settlement in the marine-sponge host. *Current Biology*, *31*(2), 433-437.

Southwell, M. W., Popp, B. N., & Martens, C. S. (2008). Nitrification controls on fluxes and isotopic composition of nitrate from Florida Keys sponges. *Marine Chemistry*, *108*(1-2), 96-108.

Spalding, M., Burke, L., Wood, S. A., Ashpole, J., Hutchison, J., & Zu Ermgassen, P. (2017). Mapping the global value and distribution of coral reef tourism. *Marine Policy*, *82*, 104-113.

Stevely, J. M., Sweat, D. E., Bert, T. M., Sim-Smith, C., & Kelly, M. (2011). Sponge mortality at Marathon and Long Key, Florida: patterns of species response and population recovery.

Tandon, K., Ricci, F., Costa, J., Medina, M., Kühl, M., Blackall, L. L., & Verbruggen, H. (2023). Genomic view of the diversity and functional role of archaea and bacteria in the skeleton of the reef-building corals Porites lutea and Isopora palifera. *GigaScience*, *12*, giac127.

Taylor, M. W., Radax, R., Steger, D., & Wagner, M. (2007). Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. *Microbiology and molecular biology reviews*, *71*(2), 295-347.

Thacker, R. W., & Freeman, C. J. (2012). Sponge-microbe symbioses: recent advances and new directions. *Advances in marine biology*, *62*, 57–111. 10.1016/B978-0-12-394283-8.00002-3

Thomas, T., Moitinho-Silva, L., Lurgi, M., Björk, J. R., Easson, C., Astudillo-García, C., ... & Webster, N. S. (2016). Diversity, structure and convergent evolution of the global sponge microbiome. *Nature communications*, *7*(1), 1-12. 10.1038/ncomms11870

Thoms, C., Ebel, R., & Proksch, P. (2006). Activated chemical defense in Aplysina sponges revisited. *Journal of chemical ecology*, *32*, 97-123.

Tian, R. M., Sun, J., Cai, L., Zhang, W. P., Zhou, G. W., Qiu, J. W., & Qian, P. Y. (2016). The deep-sea glass sponge L ophophysema eversa harbours potential symbionts responsible for the nutrient conversions of carbon, nitrogen and sulfur. *Environmental Microbiology*, *18*(8), 2481-2494.

Ueda, N., Richards, G. S., Degnan, B. M., Kranz, A., Adamska, M., Croll, R. P., & Degnan, S. M. (2016). An ancient role for nitric oxide in regulating the animal pelagobenthic life cycle: evidence from a marine sponge. *Scientific reports*, *6*(1), 37546.

Van Soest, R. W., Boury-Esnault, N., Vacelet, J., Dohrmann, M., Erpenbeck, D., De Voogd, N. J., ... & Hooper, J. N. (2012). Global diversity of sponges (Porifera). *PLoS one*, *7*(4).

van Zanten, B. T., van Beukering, P. J., & Wagtendonk, A. J. (2014). Coastal protection by coral reefs: A framework for spatial assessment and economic valuation. *Ocean & coastal management*, *96*, 94-103.

Vecchi, M., Newton, I. L., Cesari, M., Rebecchi, L., & Guidetti, R. (2018). The microbial community of tardigrades: environmental influence and species specificity of microbiome structure and composition. *Microbial Ecology*, *76*, 467-481.

Venn, A. A., Loram, J. E., & Douglas, A. E. (2008). Photosynthetic symbioses in animals. *Journal of experimental botany*, *59*(5), 1069-1080.

Wang, Z., Huang, P., You, R., Sun, F., & Zhu, S. (2023). MetaBinner: a high-performance and stand-alone ensemble binning method to recover individual genomes from complex microbial communities. *Genome Biology*, *24*(1), 1.

Webster, N. S., & Taylor, M. W. (2012). Marine sponges and their microbial symbionts: love and other relationships. *Environmental microbiology*, *14*(2), 335–346. 10.1111/j.1462-2920.2011.02460.x

Weigel, B. L., & Erwin, P. M. (2017). Effects of reciprocal transplantation on the microbiome and putative nitrogen cycling functions of the intertidal sponge, Hymeniacidon heliophila. *Scientific reports*, *7*(1), 43247.

Weisz, J. B., Hentschel, U., Lindquist, N., & Martens, C. S. (2007). Linking abundance and diversity of sponge-associated microbial communities to metabolic differences in host sponges. *Marine Biology*, *152*, 475-483.

Wen, Y., Xiao, F., Wang, C., & Wang, Z. (2016). The impact of different methods of DNA extraction on microbial community measures of BALF samples based on metagenomic data. *American Journal of Translational Research*, *8*(3), 1412.

Wiebe, W. J., Johannes, R. E., & Webb, K. L. (1975). Nitrogen fixation in a coral reef community. *Science*, *188*(4185), 257-259.

Wild, C., Huettel, M., Klueter, A., Kremb, S. G., Rasheed, M. Y., & Jørgensen, B. B. (2004). Coral mucus functions as an energy carrier and particle trap in the reef ecosystem. *Nature*, *428*(6978), 66-70. 10.1038/nature02344

Wilkinson, C. R., & Fay, P. (1979). Nitrogen fixation in coral reef sponges with symbiotic cyanobacteria. *Nature*, *279*(5713), 527-529.

Winton, M., Griffies, S. M., Samuels, B. L., Sarmiento, J. L., & Frölicher, T. L. (2013). Connecting changing ocean circulation with changing climate. *Journal of climate*, *26*(7), 2268-2278.

Wooley, J. C., Godzik, A., & Friedberg, I. (2010). A primer on metagenomics. *PLoS computational biology*, *6*(2), e1000667.

Wulff, J. (2013). Recovery of sponges after extreme mortality events: morphological and taxonomic patterns in regeneration versus recruitment. *Integrative and comparative biology*, *53*(3), 512-523.

Yildirim, S., Yeoman, C. J., Sipos, M., Torralba, M., Wilson, B. A., Goldberg, T. L., ... & Nelson, K. E. (2010). Characterization of the fecal microbiome from non-human wild primates reveals species specific microbial communities. *PloS one*, *5*(11), e13963.

Zehr, J. P., & Kudela, R. M. (2011). Nitrogen cycle of the open ocean: from genes to ecosystems. *Annual review of marine science*, *3*, 197-225.

Zhang, F., Jonas, L., Lin, H., & Hill, R. T. (2019). Microbially mediated nutrient cycles in marine sponges. *FEMS microbiology ecology*, *95*(11). 10.1093/femsec/fiz155

Zhang, X., He, L., Zhang, F., Sun, W., & Li, Z. (2013). The different potential of sponge bacterial symbionts in N2 release indicated by the phylogenetic diversity and abundance analyses of denitrification genes, nirK and nosZ. *PLoS One*, *8*(6), e65142.

Zhong, H., Lehtovirta-Morley, L., Liu, J., Zheng, Y., Lin, H., Song, D., ... & Zhang, X. H. (2020). Novel insights into the Thaumarchaeota in the deepest oceans: their metabolism and potential adaptation mechanisms. *Microbiome*, *8*, 1-16.

Appendix 1 – Supplementary File 1

https://github.com/la-zuniga/LuisZ_MS_supplemental/tree/main/fastp

Appendix 2 – Supplementary File 2

https://github.com/la-zuniga/LuisZ_MS_supplemental/tree/main/spades

Appendix 3 – Supplementary File 3

https://github.com/la-zuniga/LuisZ_MS_supplemental/tree/main/prodigal

Appendix 4 – Supplementary File 4

https://github.com/la-zuniga/LuisZ_MS_supplemental/blob/main/koala_sort.py

Appendix 5 – Supplementary File 5

https://github.com/la-zuniga/LuisZ_MS_supplemental/blob/main/heatmaps.R

Appendix 6 – Supplementary File 6

https://github.com/la-zuniga/LuisZ_MS_supplemental/tree/main/minimap2

Appendix 7 – Supplementary File 7

https://github.com/la-zuniga/LuisZ_MS_supplemental/tree/main/samtools

Appendix 8 – Supplementary File 8

https://github.com/la-zuniga/LuisZ_MS_supplemental/tree/main/metabinner

Appendix 9 – Supplementary File 9

https://github.com/la-zuniga/LuisZ_MS_supplemental/tree/main/gtdbtk

Appendix 10 – Supplementary File 10

https://github.com/la-zuniga/LuisZ_MS_supplemental/blob/main/count_bins.py

Appendix 11 – Supplementary File 11

https://github.com/la-zuniga/LuisZ_MS_supplemental/blob/main/bin_tax_viz.R

Appendix 12 – Supplementary File 12

https://github.com/la-zuniga/LuisZ_MS_supplemental/tree/main/grep

Appendix 13 - Supplementary File 13

https://github.com/la-zuniga/LuisZ_MS_supplemental/blob/main/change_headers.py

Appendix 14 - Supplementary File 14

https://github.com/la-zuniga/LuisZ_MS_supplemental/blob/main/make_table.py

Appendix 15 - Supplementary File 15

https://github.com/la-zuniga/LuisZ_MS_supplemental/tree/main/raxml

Appendix 16 - Supplementary File 16

https://github.com/la-zuniga/LuisZ_MS_supplemental/blob/main/FaithsPD.R

Appendix 17 - Supplementary Table 1

https://github.com/la-zuniga/LuisZ_MS_supplemental/blob/main/supplementary_table_1.csv

Appendix 18 - Supplementary Table 2

https://github.com/la-zuniga/LuisZ_MS_supplemental/blob/main/supplementary_table_2.xlsx

Appendix 19 - Supplementary Table 3

https://github.com/la-zuniga/LuisZ_MS_supplemental/blob/main/supplementary_table_3.xlsx

Appendix 20 - Supplementary Table 4

https://github.com/la-zuniga/LuisZ_MS_supplemental/blob/main/supplementary_table_4.xlsx