

MOLECULAR ANALYSES OF MICROBIAL MATS

by

Kaliopi Bousse

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment
of the requirements for the Master of Science in Marine Studies

Summer 2018

© 2018 Kaliopi Bousse
All Rights Reserved

MOLECULAR ANALYSES OF MICROBIAL MATS

by

Kaliopi Bousses

Approved: _____
Jennifer F. Biddle, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____
Mark Moline, Ph.D.
Director of the School of Marine Science and Policy

Approved: _____
Estella E. Atekwana, Ph.D.
Dean of the College of Earth, Ocean and Environment

Approved: _____
Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

ACKNOWLEDGMENTS

I would like to thank Dr. Thomas Hanson and Dr. Adam Marsh (College of Earth Ocean and Environment) for accepting to be on my thesis committee, Sean McAllister and Andreas Teske for providing me with previously unpublished *Beggiatoa* genomes, Dr. Rosa Leon-Zayas (Willamette University) and Glenn Christman (College of Earth Ocean and Environment) for all the bioinformatics support, Dr. Jen Biddle for being a great advisor, and the Biddle lab as a whole for being such a fun and engaging group of scientists. I would also like to acknowledge the services from the University of Delaware biomix computer cluster system and Delaware Biotechnology Institute. This research was made possible due to Dr. Barbara MacGregor and the JGI Community Sequencing Proposal, for which she is the principle investigator (Proposal ID:503040).

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	vii
ABSTRACT	ix

Chapter

1	SUCCESSION IN MICROBIAL MAT ASSEMBLAGES	1
1.1	Introduction	1
1.2	Materials and methods.....	4
1.2.1	Sample collection.....	4
1.2.2	Genomic DNA extraction, 16S rRNA gene amplicon and metagenome sequencing	5
1.2.3	Sequence analysis	5
1.2.4	Metagenome assembly and annotation	6
1.3	Results	8
1.3.1	16S rRNA gene amplicon analysis	8
1.3.2	Metagenomic taxonomic affiliations and putative function	9
1.4	Discussion.....	23
2	UNCOVERING KEY DIFFERENCES IN PHYLOGENY AND FUNCTIONAL POTENTIAL OF BEGGIATOVA GENOMES.....	28
2.1	Introduction	28
2.2	Materials and methods.....	30
2.2.1	Genome collection	30
2.2.2	Metagenome-assembled genome analysis	31
2.3	Results	33
2.3.1	Phylogenetic affiliations of <i>Beggiatoa</i> genomes	33

2.3.2 Putative functional annotations and metabolic pathways detected in <i>Beggiatoa</i> genomes	34
2.4 Discussion.....	45
REFERENCES	57
Appendix	
PROKARYOTIC AND EUKARYOTIC COMMUNITY ASSEMBLAGES WITHIN A MICROBIALITE AND NODULES IN PAVILION LAKE, BRITISH COLUMBIA, CANADA	67

LIST OF TABLES

Table 1.1 OTUs present in microbial mats (FB and BB) and incoming water (IW) ...	15
Table 1.2 Metrics from metagenomes obtained from microbial mats grown on bone surfaces	16
Table 1.3 Taxonomic affiliations of MAGs obtained from FB and BB samples.....	17
Table 1.4 Average percent dissimilarity in functional potential between the groups of metagenomes	19
Table 1.5 Results of SIMPER test from epsilonproteobacterial and gammaproteobacterial MAGs	22
Table 2.1 Genomes and genomic bins related to the genus <i>Beggiatoa</i>	38
Table 2.2 Results of SIMPER test between deep and shallow <i>Beggiatoa</i> genomes	40
Table A.1 OTUs from "Other" Cyanobacteria	79
Table A.2 OTUs from "Other" Cyanobacteria in nodules.....	82

LIST OF FIGURES

Figure 1.1 Observed mat development.....	13
Figure 1.2 Percent relative abundance of Phylum-level OTUs in microbial mat samples.	13
Figure 1.3 Percent relative abundance of Genus-level OTUs within the gammaproteobacterial class from microbial mat samples.....	14
Figure 1.4 Percent relative abundance of Genus-level OTUs within the epsilonproteobacterial class from microbial mat samples	14
Figure 1.5 Cluster diagrams from the Bray-Curtis resemblance matrix of the OTU counts in the microbial mat samples	15
Figure 1.6 %GC vs Genome Size of retrieved MAGs	18
Figure 1.7 Heatmap of putative functions in microbial mat metagenomes.....	20
Figure 1.8 Heatmap of putative functions in microbial mat genomes retrieved from metagenomes	21
Figure 2.1 Phylogenetic affiliations of <i>Beggiatoa</i> genomes	39
Figure 2.2 KEGG annotated pathways that are present in the tested genomes	41
Figure 2.3 Heatmap of functional annotations present in the genomes after the Roary pangenome construction	42
Figure 2.4 Gene annotations as a function of number of genomes	43
Figure 2.5 Gene history of <i>Beggiatoa</i> genomes	44
Figure A.1 Shore based PAM and imaging of microbialite mound from 28 meters depth.....	75
Figure A.2 Shore based PAM and imaging of pink nodules from 21 meters depth.....	76
Figure A.3 Microelectrode profiles of modules and background biofilm.....	76

Figure A.4 Prokaryotic OTUs in microbialite mound from Three Poles	77
Figure A.5 Eukaryotic OTUs in microbialite mound from Three Poles	79
Figure A.6 Prokaryotic OTUs nodules and background biofilm form South Basin East.....	81
Figure A.7 Similarity percentage (SIMPER) of most contributing OTUS in the prokaryotic communities of green and pink nodules	82
Figure A.8 Eukaryotic OTUs nodules and background biofilm form South Basin East.....	84
Figure A.9 Similarity percentages (SIMPER) of most contributing OTUs in the eukaryotic communities of green and pink nodules.....	85
Figure A.9 Similarity percentages (SIMPER) of most contributing OTUs in the eukaryotic communities of green and pink nodules.....	85
Figure A.S1 Chao and Shannon indexes for the microbialite mound prokaryotic communities	94
Figure A.S2 Profiles of Alphaproteobacteria from microbialite mound 16S rRNA gene amplicon samples	94
Figure A.S3 Profiles of Cyanobacteria from microbialite mound 16S rRNA gene amplicon samples.....	95
Figure A.S4 Chao and Shannon indexes for the microbialite mound eukaryotic communities	95
Figure A.S5 Profiles of Alphaproteobacteria from nodule 16S rRNA gene amplicon samples.....	96
Figure A.S6 Profiles of Cyanobacteria from nodule 16S rRNA gene amplicon samples.....	96
Figure A.S7 Chao and Shannon index for the nodule and background biofilm prokaryotic communities.	97
Figure A.S8 Chao and Shannon index for the nodule and background biofilm eukaryotic communities	97

ABSTRACT

Whether found in shallow water sediments or at the bottom of the deep-sea, microbial mats are key participants in the recycling and utilization of important elements and compounds. Despite these mats being nearly ubiquitous, little research has focused on the early stages of their community assemblage. In chapter 1, I describe the development of microbial mats on two bones exposed to water incoming from a tidal river. Molecular analysis of these mats, including 16S rRNA gene amplicon and metagenomic sequencing, showed that gammaproteobacterial classes aided in the initial development of the mat, and were swiftly overtaken by epsilonproteobacterial classes. This observed successional change in mat members over time was not represented in the putative functions associated with the communities, as the potential for autotrophic and heterotrophic lifestyles overlapped throughout the duration of the mat development. Together these data suggest that functional redundancy may be a characteristic of these mat types, despite the shifts in mat community members. I wanted to further understand the phylogenetic and functional potential of one mat member in particular, the large filamentous sulfur-oxidizing *Beggiatoa*. Nine metagenomic bins were retrieved from the mat metagenomes and compared with marine and freshwater genomes. *Beggiatoa* genomes from shallow waters and their deep water counterparts appear to rely on different modes of growth and energy acquisition. Furthermore, pangenomic investigation suggests that shallow-dwelling *Beggiatoa* may be more suited to regulate

their responses to stimuli such as light changes, and that the gene history of these organisms may shift due to different habitats or geographic location.

Chapter 1

MICROBIAL SUCCESSION IN MICROBIAL MAT AGGREGATIONS

1.1 Introduction

Over the years, observations of successional changes and patterns in community development of plants and animals have been important topics of scientific research. However, efforts to understand how microbial communities form and change over time have only recently been made. This may be surprising, given the fact that the majority of diversity is microbial in nature (Pace, 1997). Therefore, it is to be expected that microbial communities form and develop over time in a fashion similar to those of higher-level organisms (Fierer, Nemergut, Knight, & Craine, 2010).

Biofilms and microbial mats are formations of microbial communities that develop on solid surfaces in which micro-scale physicochemical gradients are utilized for energy utilization and growth (Bolhuis, Cretoiu, & Stal, 2014). Microbial mats are capable of forming on organic and inorganic substrates and are classified based on the type of surface they grow on (Hladyz, Cook, Petrie, & Nielsen, 2010). For example, mats growing on wood surfaces are referred to as epixylic, while mats forming on rocks are referred to as epilithic (Hladyz et al., 2010; Lang et al., 2016). In different microbial mats pathways for energy acquisition and growth may vary, which can be reflected in what taxonomic affiliations are associated with mats on different substrates (Lang et al., 2016). Additionally, successional patterns have been observed in these types of microbial communities. For example, studies of simulated marine wood falls showed successional changes in the community associated with the

epixylon. Specifically, metagenomic analyses suggested that the initial colonizers might have been chemolithoheterotrophs, with later community members being chemoautotrophs (Kalenitchenko et al., 2016). Automated ribosomal intergenic spacer analysis (ARISA) and 16S rRNA gene pyrosequencing have revealed successional changes in the community composition of epilithic and epinecrotic (on the surface of deceased carcasses) microbial mats (Benbow, Pechal, Lang, Erb, & Wallace, 2015; Lang et al., 2016). On epinecrotic biofilms, inverse changes in the relative abundances of Proteobacteria and Firmicutes have occurred over the course of 21 or 42 days, depending on the season. Changes at the genus level have also been documented in these mat types (Benbow et al., 2015).

In addition to epilithic and epixylic mats, an additional phase of mat development is known from whale fall decomposition, and is characterized by colonization of chemosynthetic microorganisms in large microbial mats. This decomposition stage is also known as sulfophilic, as sulfides are in high concentrations near the carcass. This sulfide production is attributed to the sulfate-reducing bacteria that further aid in the decomposition of any remaining whale body tissue that macrofaunal lineages, such as polychaetes and crustaceans, have left behind (Smith, Glover, Treude, Higgs, & Amon, 2015). Once the bones are bare, the release of sulfide from the sulfate-reducing bacteria and dissolution of bone material is suspected to stimulate the formation and growth of dense sulfur-oxidizing bacterial mats that have been observed near whale falls (Smith et al., 2015).

Few studies have focused on the changes in taxonomic diversity and functional potential of microbial mats associated with whale falls. One study showed that whale falls seemed to be lacking in diversity and functional attributes when compared to

open ocean and soil communities (Tringe et al., 2005). The difficulty in obtaining samples from these ecosystems may be a reason why research is lacking in this field. What proves to be even more challenging is the study of successional patterns in taxonomy and putative metabolic function associated with mats from whale falls. Despite the lack of samples, whale falls are believed to be habitats of significant importance, as they may have served as hot spots for chemosynthetic vent and seep faunal lineages to disperse to (Smith et al., 2015). More studies on whale falls and the microbial mats associated with these environments would provide insight on how these ecosystems thrive.

In an attempt to provide more information regarding microbial mat development on bones in a more accessible manner than a whale fall, this chapter explores the formation and growth of a mat on pork bones in shallow waters. In this way, we can measure mat development on bones from whale falls and answer the following questions: On a temporal scale are there successional changes in the inhabitants of a bone-based microbial mat? And, if so, what putative metabolic functions are associated with such shifts in microbial communities? We employed 16S rRNA gene amplicon data, whole metagenomic investigations, and genomic binning techniques to answer the questions raised in this study. The results of this study showed that successional patterns in the community composition of microbial mats did occur. Over the course of mat development, initial colonizing groups within the class of Gammaproteobacteria were overtaken by genera from Epsilonproteobacteria. We were able to retrieve genomes with high completeness levels and determine putative functional pathways associated with these mat communities. The changes in

phylogeny did not alter the functional potential of the mats over time, suggesting that functional redundancy occurs in these types of microbial communities.

1.2 Material and Methods

1.2.1 Sample Collection

Two pork bones were cooked to remove flesh, boiled to remove further cartilage and connective tissues, then autoclaved and placed in a tank located at the Pollution Ecology Lab (PEL) of the University of Delaware (38.7894827, -75.1644337). One was labeled FB and the other BB (Figure 1.1). The tank at the PEL was exposed to a continuous flow of water from the mouth of the Broadkill River. Twelve days after the bones were placed into the tank, colonization became apparent once thin white mats appeared on the bone surfaces and then both mats were sampled twice a week from November 7, 2016 until December 1, 2016. Eight samples came from each bone mat (16 total). Samples from the first week were named BB1, BB2 for one bone and FB1 and FB2 for the second. Samples from the second week were BB3, BB4, FB3, and FB4. Samples for the third week were samples BB5, BB6, FB5, and FB6. Samples from the fourth and final week were BB7, BB8, FB7 and FB8. The surface of the mat was scraped with a sterilized spatula making sure not to sample from the previously sampled areas. Mat material was then transferred to a microcentrifuge tube and stored at -20 °C until DNA extractions were performed within two weeks from the last sampling date. Concurrently with sampling of microbial mats in this tank setup, samples were collected from the incoming water once a week for 12 weeks. Cells in the incoming water samples (labeled IW) were filtered onto a 0.2 mixed cellulose filter and stored at -20°C.

1.2.2 Genomic DNA extractions, 16S rRNA gene amplicon and metagenome sequencing

Genomic DNA was extracted from 0.2 g of mat and IW material using the MoBio PowerSoil kit (MoBio Valencia, CA) per the manufacturer's instructions. The extracted DNA was sent to the United States Department of Energy's Joint Genome Institute (JGI) for amplicon (mat and IW) and metagenomic (mat only) sequencing. For the 16SrRNA amplicon sequencing, the V4 variable region (515F/805R) was sequenced on a MiSeq Illumina sequencer (Illumina, San Diego, CA, USA), producing 300-bp paired-end reads (Caporaso et al., 2011). For the metagenomes DNA was sequenced on a HiSeq Illumina sequencer generating 150-bp paired-end reads (Illumina, San Diego, CA, USA), using JGI protocols. All sequences are publicly available through the Joint Genome Institute's Genome Portal interface (Proposal ID:503040).

1.2.3 Sequence Analysis

The 16S rRNA gene amplicon sequences from the two bones and IW were analyzed separately. The QIIME pipeline was used to demultiplex and quality filter the sequences to a minimum quality score of 29 (Caporaso et al., 2010). The SILVA database was used for taxonomic assignments of operational taxonomic units (OTU) at the 97% identity level (Quast et al., 2013). Unassigned and singleton OTUs were removed through QIIME, and rarefactions were performed prior to generating biom tables. The sequences for each community sample were rarefied to 221,647 (FB), 249,295 (BB), and 209,043(IW). Bray-Curtis dissimilarity matrices and cluster analyses were calculated using the statistical software package PRIMER-E (Clarke & Gorley, 2006).

1.2.4. Metagenome assembly and annotation

A total of 15 metagenomes were obtained from the bone mats. Sequencing of the metagenome for the last sampling day on the BB (BB8) bone failed. For each metagenome, forward and reverse reads were trimmed to a quality score of 20 using the wrapper script trim galore (Krueger, 2015). The remaining reads that passed quality control were assembled in SPADIS with the flag `--meta`, and k-mer sizes of 21, 33, 55 (Nurk et al., 2013).

Individual genomes were recovered from the assembled contigs with Maxbin (Wu, Tang, Tringe, Simmons, & Singer, 2014). The metagenome-assembled genomes (MAGs) were screened for contamination and completion using CheckM (Parks, Imelfort, Skennerton, Hugenholtz, & Tyson, 2015). MAGs with high contamination values were assessed through Vizbin and further manually refined until contamination was below 10% (Laczny et al., 2015). The MAGs with a minimum of 60% completeness and a maximum of 10% contamination were phylogenetically annotated with Phylosift (Darling et al., 2014).

Functional annotations were performed with Prokka on the assembled metagenomes, as well as all MAGs that were taxonomically identified as *Beggiatoa* (Seemann, 2014). The analysis for the *Beggiatoa* MAGs will be further discussed in chapter 2. In addition to the *Beggiatoa* MAGs, genomic annotations were performed on MAGs with the highest percent completeness values that were taxonomically identified as *Marinobacterium* (BB4_007), *Arcobacter* (FB1_004), and *Sulfurovum* (FB7_008). The functional predictions from Prokka were uploaded to GhostKoala in order to utilize the Kyoto Encyclopedia for Genes and Genomes (KEGG) Ontology (KO) system for annotations of the whole metagenomic assemblies and MAGs

(Kanehisa, Sato, & Morishima, 2016). More specifically, the script KEGG-decoder.py (<https://github.com/bjtully/BioData/tree/master/KEGGDecoder>) was used to determine if certain pathways and functions were complete according to the KO assignments (Tully, Wheat, Glazer, & Huber, 2018).

Using the same statistical software package PRIMER-E, SIMPER analyses were performed in order to determine which putative metabolic pathways might be more influential towards any observed differences in functional potential of the mats over time (Clarke & Gorley, 2006). The SIMPER analysis, which is short for “similarity percentage” calculates the average contribution of the annotated functions to the dissimilarity observed between pre-assigned groups. For the functional potential of whole metagenomic annotations, the groups assigned were based on the week of sampling, where week one included samples BB1, BB2, FB1 and FB2, week two included BB3, BB4, FB3 and FB4, week three included BB5, BB6, FB5 and FB6 and week four included BB7, FB7 and FB8. SIMPER analysis was also utilized to determine functional differences in four MAGs that were grouped based off of their taxonomic class affiliations. Specifically, the first group was labeled Epsilonproteobacteria, which had an *Arcobacter* (FB1_004) and a *Sulfurovum* (FB7_008) MAG. The second group was labeled Gammaproteobacteria and had a *Beggiatoa* MAG (BB1_001) and a *Marinobacterium* MAG (BB4_007). These genomes were selected as being the most complete MAGs that were representative of the majority population.

1.3 Results

1.3.1 16S rRNA gene amplicon analysis

At the phylum level, all 16S rRNA amplicon sequencing samples from both bones were most abundant in Proteobacteria with Bacteroidetes, Firmicutes and other organisms following in lower percent relative abundances (Figure 1.2). At lower phylogenetic levels, an opposing trend was observed between relative abundances of Gammaproteobacteria and Epsilonproteobacteria (Figure 1.3; Figure 1.4). More specifically, the maximum relative abundance of Gammaproteobacteria was observed at the start of sampling (40% in FB1 and 50% in BB1), while the minimum was observed in later sampling times (12% in FB8 and 14% in BB6; Figure 1.3). This decrease in Gammaproteobacteria was met by an increase in relative abundance of Epsilonproteobacteria found in samples from the last week; more than 60% of the total community in samples FB7, FB8, BB7, and BB8 was comprised of Epsilonproteobacteria (Figure 1.4).

Other notable trends included the initial high abundance of OTUs identified as *Leucothrix*, a genus from the Thiotrichacea family, in sample FB1 (10% of the total community), and its subsequent decrease over time (Figure 1.3). Similarly, in BB1 another Thiotrichaceon, known as *Candidatus Thiopilula* had a relative abundance of 13%, but showed a decreased presence in later samples. The Epsilonproteobacterial genus *Arcobacter* remained in relative abundances ranging from 55% (FB5) to 20% (FB8) and 12% (BB1) to 51% (BB5) (Figure 1.4). OTUs related to *Sulfurovum* showed an increase in relative abundance over time, from 0.1% (FB1) to 46% (FB8) and 0.7%(BB1) to 19% (BB7).

Community structures of the mats from both bones showed changes over time (Figure 1.5). In FB, FB1 branched separately from the rest of the samples, as did FB7 and FB8. Furthermore, FB2, FB3, and FB4 clustered separately from FB5 and FB6. While the community structure of BB was not identical to FB, the samples did group into similar clades, with BB1 and BB2 separating from the later samples and BB7 and BB8 separating from the earlier ones (Figure 1.5).

From the incoming water samples, OTUs related to *Arcobacter*, *Beggiatoa*, *Leucothrix*, and *Thiopilula* were identified and compared to mat amplicon data. There were observed differences in the relative abundances of these OTUs, based off of sampling source (microbial mat or incoming water; Table 1.1). OTUs related to the noted taxonomies were several percentages higher in relative abundance when they were identified in microbial mat samples than when they were found in incoming water samples. In samples from incoming water, the relative abundance of Gammaproteobacteria ranged from 12% to 30% (Table 1.1). The same class ranged from 12 to 49% in the microbial mats. The relative abundance of Epsilonproteobacteria ranged from 0.5% to 2% of the total community in the incoming water samples, but appeared in much higher abundances in the microbial mat samples.

1.3.2 Metagenomic taxonomic affiliations and putative function

Of the 16 samples from the microbial mats, 15 metagenomes were sequenced successfully, as a metagenome for sample BB8 was not produced, due to a bad sequence run. The remaining metagenomes that were constructed had a range of 20,563,976 (BB1) to 31,700,026 (FB6) paired-end reads after trimming and quality checking (Table 1.2). The longest contig was obtained from the FB1 assembled metagenome (544,363 bp) and the shortest one came from FB6 (161,251 bp). Sample

FB1 also had the highest maximum contig length to read ratio, while FB6 had the smallest, suggesting that the complexity of FB1 was the simplest and FB6 contained more complex biomass and/or degraded extracellular DNA.

A total of 160 MAGs were obtained from the metagenomes of the FB and BB microbial mats, ranging in completeness of 60.18% to 99.82% (Table 1.3; Figure 1.6). Most genomes were assigned to gammaproteobacterial (59) or epsilonproteobacterial (55) classes. MAGs related to Deltaproteobacteria and Bacteroidetes were also present. Interestingly, several of these taxa did not show any clear patterns of appearance over time. Many genomes appeared in almost all metagenomic samples, such as those assigned to *Marinobacterium* or *Arcobacter* genera. In fact, a total of 22 MAGs were related to *Marinobacterium* and 19 to *Arcobacter*. Other taxa appeared in earlier sampling times, disappeared and reappeared in metagenomes at later sampling times. From the BB metagenomes, *Beggiatoa* genomes appeared at least once during every sampling week. However, in the FB metagenomes, *Beggiatoa* MAGs did not appear in FB5, FB6, or FB8. A *Thiovulum* genome was retrieved during the second week of sampling in metagenome BB4, but none were found in metagenomes BB5 or BB6. A second genome appeared during the final week of sampling in the BB7 metagenome (Table 1.3).

Metagenomes from both bones had complete pathways for glycolysis and the citric acid (TCA) cycle, as well as carbon fixation pathways, such as the Calvin Benson Bassham (CBB) cycle and reverse TCA (rTCA) cycle (Figure 1.7). Pathways for nitrite oxidation, dissimilatory nitrate reduction to ammonium, and nitrogen fixation were complete as well. Sulfide oxidation and sulfur assimilation were also complete across all metagenomes. Pathways for methanogenesis were not complete,

with the exemption of methanogenesis via trimethylamine, which was present in all metagenomes. Most pathways for mixed acid fermentation were complete, except for the conversion of formate to carbon dioxide and hydrogen gas. Overall, the putative function of the metagenomic samples did not seem to show any change over time. This was confirmed through SIMPER analysis, in which the percent dissimilarity between the tested groups did not go above 6.44% (Table 1.4).

The MAGs related to Epsilonproteobacteria or Gammaproteobacteria were chosen as representatives of taxa that were most abundant in the 16SrRNA gene amplicon and metagenomic data and used to test the dissimilarity of the two classes. These four genomes were specifically selected since they were the most complete and/or had the lowest amount of contamination out of the remaining genomes related to the two classes. The gammaproteobacterial MAGs BB1_001 and BB4_007 were 97.28% and 99.82% complete respectively. The epsilonproteobacterial MAGs FB1_004 and FB7_008 were 99.29% and 99.59% complete, respectively. The putative functions of MAGs related to Gammaproteobacteria and Epsilonroteobacteria showed some key differences (Figure 1.8). The average dissimilarity between the MAGs of the two classes was 41.66 %. Certain pathways contributed more to the observed dissimilarity (Table 1.5). The epsilonproteobacterial MAG related to the *Sulfurovum* genus (FB7_008) was the only genome that contained a complete rTCA pathway (Figure 1.8). No other carbon-fixation pathways were identified in any of the four MAGs. While all four genomes showed the potential for dissimilatory nitrate reduction, the gammaproteobacterial *Marinobacterium* MAG (BB4_007) was the only one out of the four to have genes coding for the membrane-bound nitrate reductase, known as the Nar system. The other three coded genes for the periplasmic nitrate

reduction (Nap) system. BB4_007 was also the only MAG that contained the necessary genes to complete nitrite oxidation. Dissimilatory nitrate reduction to ammonium (DNRA) was complete only in the *Beggiatoa* MAG (BB1_001). Nitric oxide and nitrous oxide reduction were also complete in the same MAG. Nitrogen fixation, dissimilatory sulfate reduction and oxidation potential were present only in the two epsilonproteobacterial MAG. On the other hand, thiosulfate oxidation was complete in BB4_007, and only partially complete in BB1_001 and FB7_008. The potential for sulfide oxidation was complete in all four MAGs. Finally, the glyoxylate shunt was complete only in the two gammaproteobacterial bins (Figure 1.8). As representatives of the two classes that were most commonly associated with the mat communities of this study, the putative function of these MAGs suggest that chemoheterotrophic pathways may be utilized by the gammaproteobacterial members, while chemoautotrophic pathways may be more associated with certain epsilonproteobacterial members.

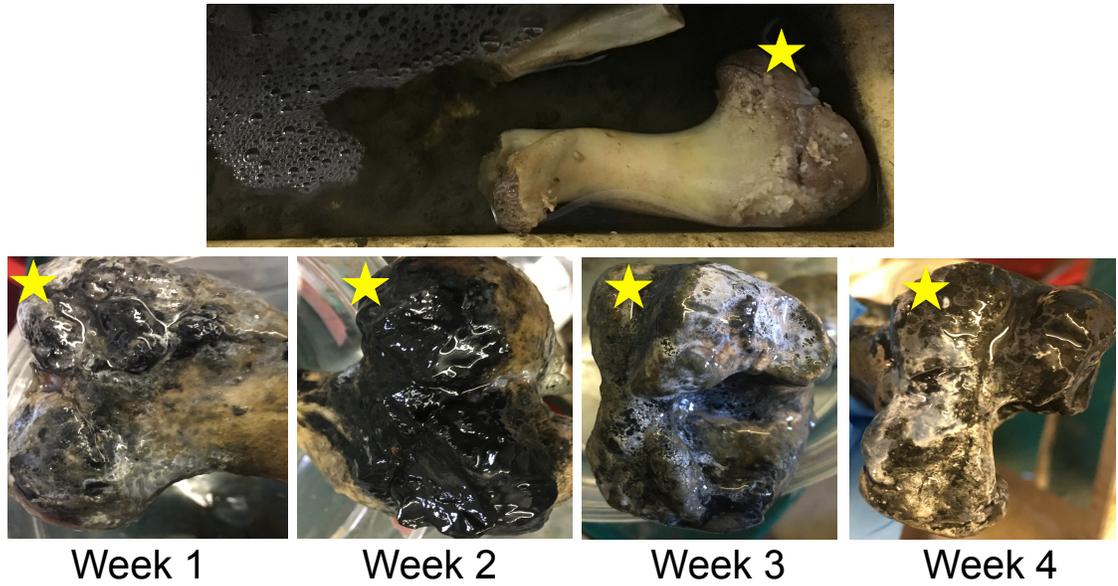


Figure 1.1 Observed mat development. The top panel shows the whole bone (BB) when it was first placed in the tank. The yellow star points to the surface of the bone that was photographed in the bottom panel. The images in the bottom panel show the presence of mat development during the four weeks of sampling.

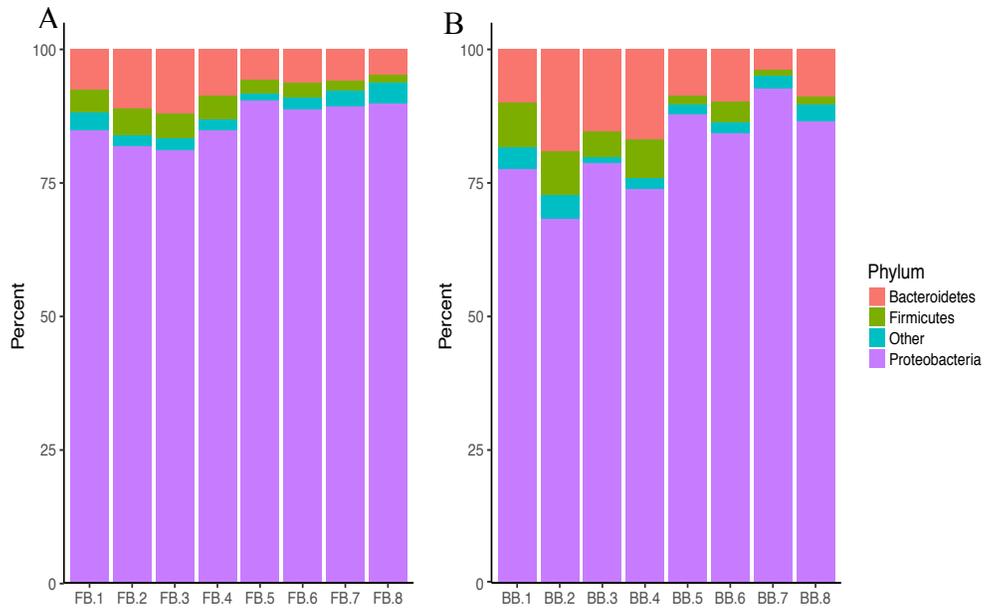


Figure 1.2 Percent relative abundance of Phylum-level OTUs in microbial mat samples. Panel A: FB samples (FB1 to FB8). Panel B: BB samples (BB1-BB8).

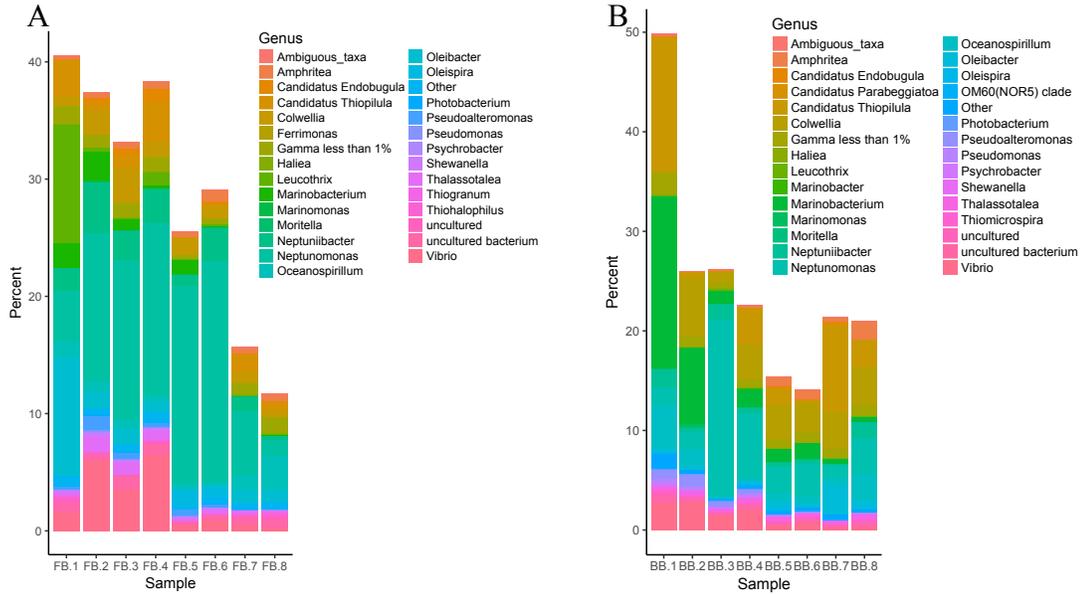


Figure 1.3 Percent relative abundance of genus-level OTUs within the gammaproteobacterial class from microbial mat samples. Panel A: FB samples (FB1 to FB8). Panel B: BB samples (BB1-BB8).

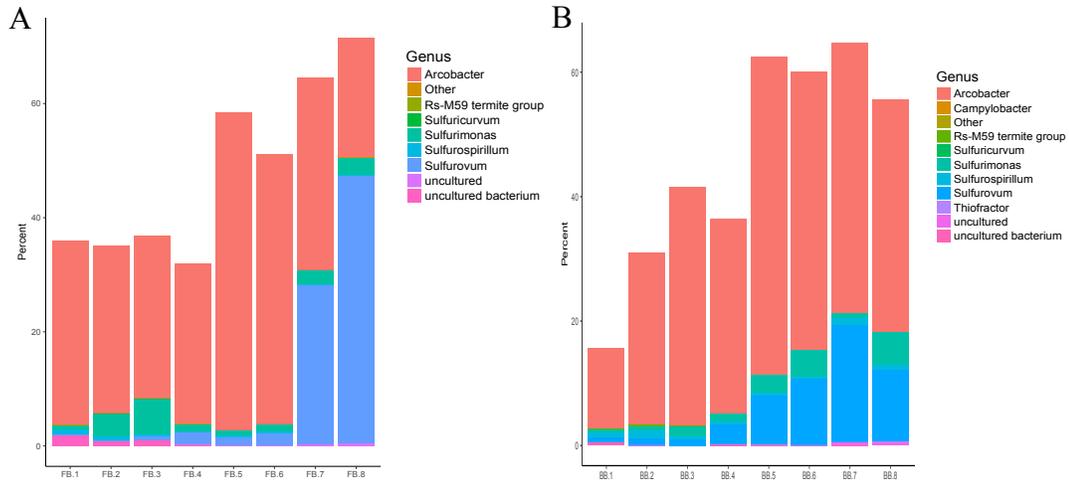


Figure 1.4 Percent relative abundance of genus-level OTUs within the epsilonproteobacterial class from microbial mat samples. Left panel: FB samples (FB1 to FB8). Right panel: BB samples (BB1-BB8).

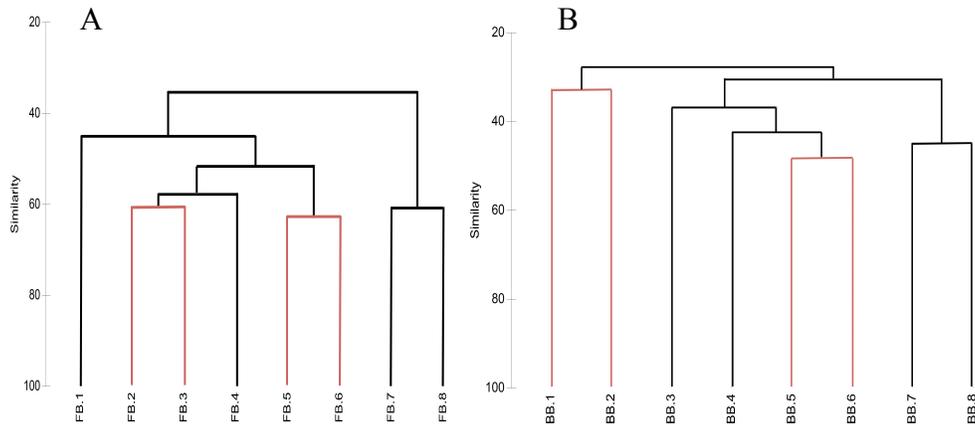


Figure 1.5 Cluster diagrams from the Bray-Curtis resemblance matrix of the OTU counts in the microbial mat samples. Panel A: FB samples. Panel B: BB samples. Count values were square-root transformed prior to calculating the Bray-Curtis resemblance matrix. Red-colored braches indicate that sample differences are not significant, while black-colored branches represent sample differences that are significant (Significance level: 5%).

Table 1.1 OTUs present in microbial mats (FB and BB) and incoming water samples (IW)

OTUs related to:	Max relative abundance in sample type(%)		
	FB	BB	IW
<i>Arcobacter</i>	55	51	0.008
<i>Beggiatoa</i>	0.04	0.24	0.002
<i>Leucothrix</i>	10	0.2	0.009
<i>Thiopilula</i>	1	13	0.001

Table 1.2 Metrics from metagenomes obtained from microbial mats grown on bones

Metagenome	Number of Unassembled Reads	Maximum Contig Length (bp)	Max Contig Length to Read Ratio (100X)
BB1	20,563,976	203,592	0.99
BB2	23,123,872	287,692	1.24
BB3	30,400,916	421,232	1.39
BB4	24,251,912	284,823	1.17
BB5	27,971,112	212,668	0.76
BB6	30,104,508	249,542	0.83
BB7	27,146,506	269,690	0.99
FB1	26,503,318	544,363	2.05
FB2	31,610,490	461,467	1.46
FB3	31,098,422	258,894	0.83
FB4	25,074,588	178,026	0.71
FB5	28,141,282	205,623	0.73
FB6	31,700,026	161,251	0.51
FB7	31,243,830	290,751	0.93
FB8	26,729,504	239,662	0.90

Table 1.3 Taxonomic affiliations of MAGs obtained from FB and BB samples

Phylogenetic Assignment	Number of MAGs	Presence in BB Samples	Presence in FB Samples
<i>Roseobacter</i>	1	-	FB3
Unassigned Bacteroidales	4	BB4, BB7	FB3
Unassigned Marinilabiliaceae	1	BB4	-
Unassigned Porphyromonadaceae	9	BB1, BB2, BB3, BB4, BB6	FB8
<i>Alkaliphilus</i>	1	-	FB5
<i>Sedimentibacter</i>	1	BB4	-
<i>Desulfobacter</i>	4	BB4	FB6, FB7, FB8
<i>Desulfococcus</i>	3	BB6	FB4, FB7
<i>Desulfuromonas</i>	7	BB2, BB5, BB7	FB2, FB5, FB7
Unassigned Desulfuromonadales	1	BB6	-
Unassigned Myxococcales	1	BB5	-
<i>Arcobacter</i>	19	BB2, BB3, BB4, BB5, BB6, BB7	FB1, FB2, FB3, FB5, FB6, FB7, FB8
<i>Sulfurimonas</i>	14	BB3, BB4, BB5, BB6,	FB2, FB3, FB4, FB7, FB8
<i>Sulfurospirillum</i>	3	-	FB2, FB3, BB7
<i>Sulfurovum</i>	15	BB3, BB4, BB5, BB6, BB7	FB4, FB5, FB6, FB7, FB8
<i>Thalassolituus</i>	11	BB1, BB5, BB6, BB7	FB1, FB2, FB3, FB4, FB5, FB6, FB7
<i>Thiovulum</i>	4	BB4, BB7	FB7, FB8
<i>Polaribacter</i>	3	-	FB1, FB2, FB3
Unclassified Flavobacteriales	7	BB3, BB3, BB4, BB5, BB6	-
<i>Beggiatoa</i>	10	BB1, BB2, BB4, BB5, BB7	FB1, FB2, FB4, FB7
<i>Colwellia</i>	5	BB2, BB4, BB5, BB6, BB7	-
Gamma IMCC1989	3	-	FB2, FB3, FB6
<i>Marinobacterium</i>	22	BB1, BB2, BB3, BB4, BB5, BB6,	FB1, FB2, FB3, FB4, FB5, FB6, FB7, FB8
<i>Marinomonas</i>	1	-	FB4
<i>Nitrosoccus</i>	1	-	FB3
<i>Thiomicrospira</i>	2	BB3, BB6	-
<i>Thiothrix</i>	1	-	FB1
Unclassified Alteromonadales	1	BB7	-
<i>Vibrio</i>	1	-	FB4
<i>Parcubacteria</i>	1	-	FB7
<i>Saprospira</i>	1	-	FB1
N/A	2	BB4	FB5



Figure 1.6 %GC vs Genome Size of retrieved MAGs. The plot shows which sample metagenome each genome came from, while colors represent the Phylsift taxonomic assignments. Font size represents percent completeness of each MAC according to CheckM, after refinement with Vizbin.

Table 1.4 Average percent dissimilarity in functional potential between the groups of metagenomes.

Groups¹	Average SIMPER Dissimilarity (%)
Week 1 and Week 2	4.76
Week 1 and Week 3	5.42
Week 2 and Week 3	4.49
Week 1 and Week 4	6.65
Week 2 and Week 4	6.44
Week 3 and Week 4	5.86

¹Group “Week 1” contained metagenomes FB1, FB2, BB1, and BB2. Group “Week 2” contained FB3, FB4, BB3, and BB4. “Week 3” contained FB5, FB6, and BB5, and BB6. Week 4 contained FB7, FB8, and BB7.

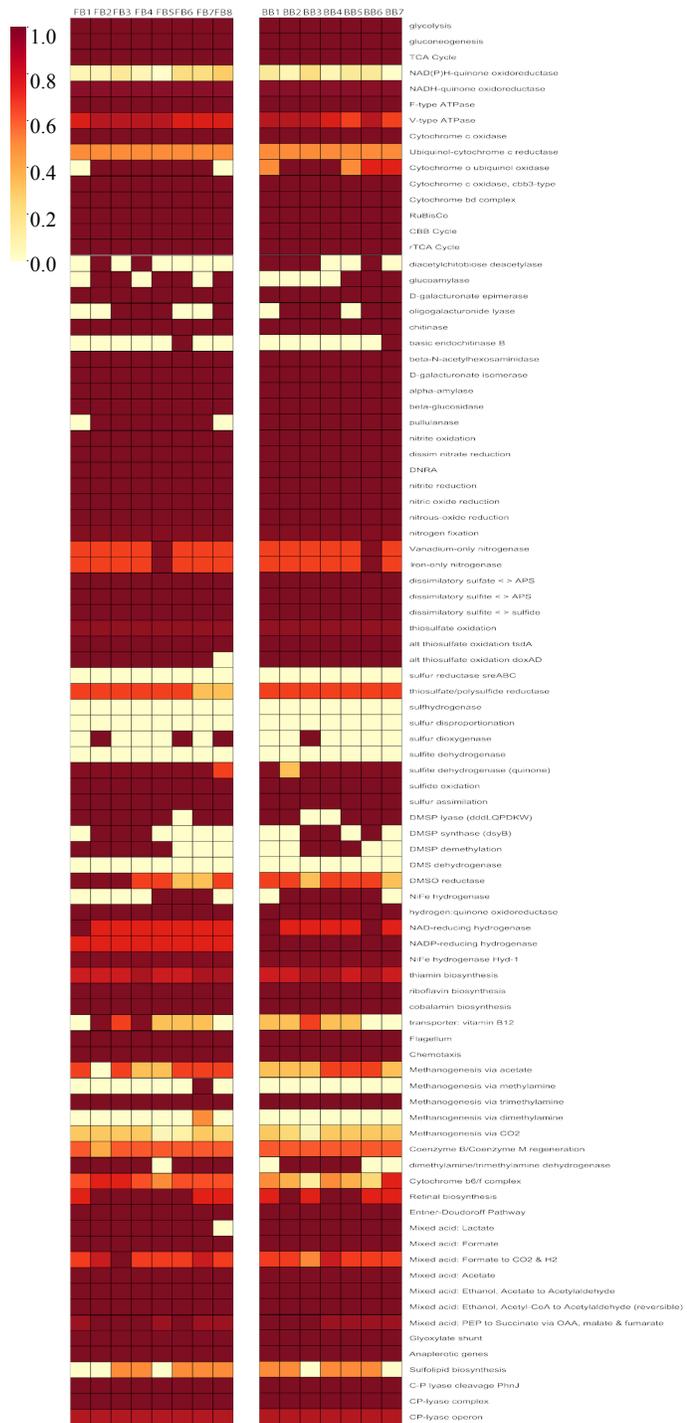


Figure 1.7 Heatmap of putative functions in microbial mat metagenomes. Left: FB metagenomes, FB1-FB8. Right: BB metagenomes, BB1-BB7.

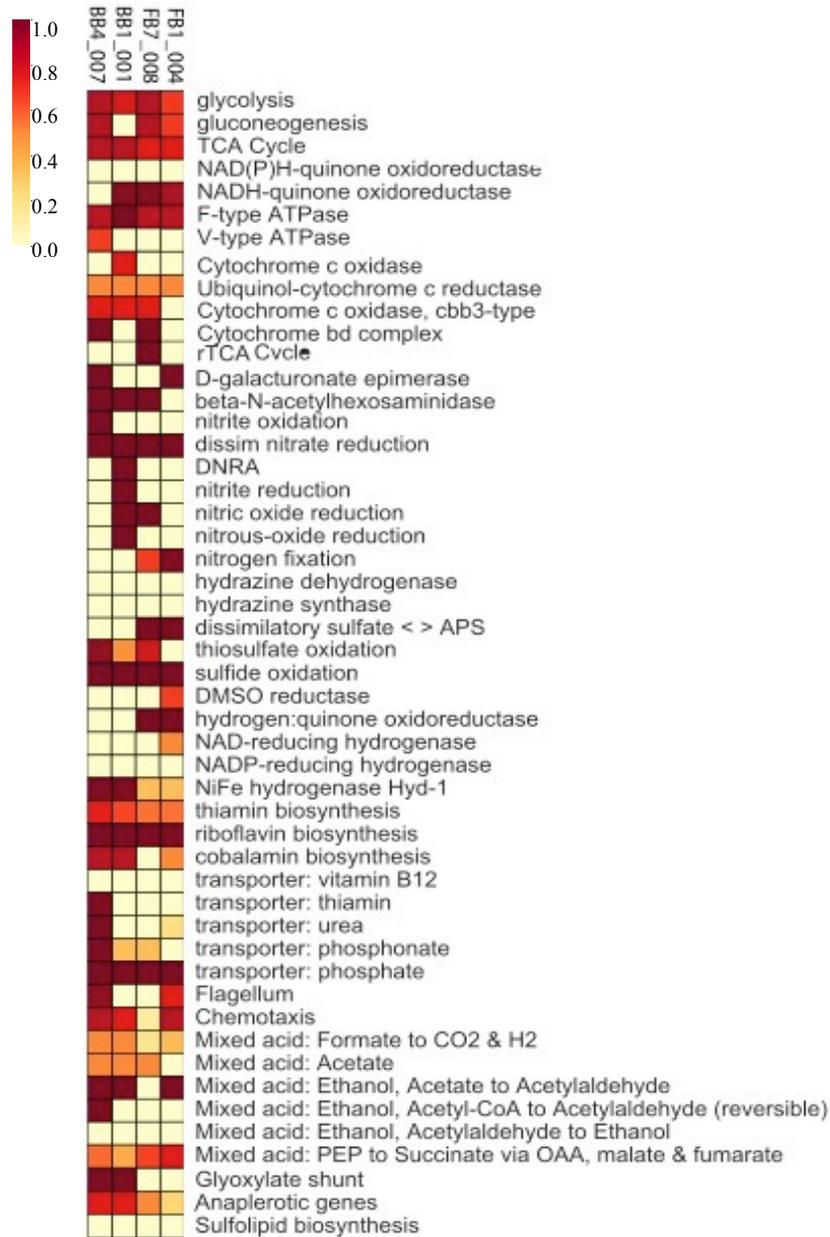


Figure 1.8 Heatmap of putative functions in microbial mat genomes retrieved from metagenomes. From left to right: BB4_007: *Marinobacterium* related genome. BB1_001: *Beggiatoa* related genome. FB7_008: *Sulfurovum* related genome, and FB1_004 is an *Arcobacter* related genome.

Table 1.5 Results of SIMPER test from epsilonproteobacterial and gammaproteobacterial MAGs

Within-Group Similarity (ϵ-proteobacteria)	Within-Group Similarity (γ-proteobacteria)	Between-Group Dissimilarity¹ (ϵ vs γ)
Dissim nitrate reduction (7.48) ¹	Beta-N-acetylhexosaminidase (5.71)	Dissimilatory sulfate <> APS(5.11)
Dissimilatory sulfate <> APS(7.48)	Dissim nitrate reduction (5.71)	Glyoxylate shunt (5.11)
Sulfide oxidation (7.48)	Sulfide oxidation (5.71)	Nitrogen fixation (4.23)
Hydrogen: quinone oxidoreductase (7.48)	Riboflavin biosynthesis (5.71)	NiFe hydrogenase Hyd-1 (3.37)
Riboflavin biosynthesis (7.48)	Transporter: phosphate	Cobalamin biosynthesis (3.2)
Transporter: phosphate(7.48)	Mixed acid: Ethanol, Acetate to Acetaldehyde (5.71)	DNRA (2.73)
NADH-quinone oxidoreductase (6.81)	Glyoxylate Shunt (5.71)	Nitrite reduction (2.73)
	NiFe hydrogenase Hyd-1 (5.65)	Nitrous-oxide reduction (2.73)
		Beta-N-acetylhexosaminidase (2.58)
		D-galacturonate (2.56)
		Nitric oxide reduction (2.56)
		Cytochrome bd complex (2.55)
		rTCA Cycle (2.53)
		Mixed acid: Ethanol, Acetate to Acetylaldehyde (2.53)
		Thiosulfate oxidation (2.44)
		Hydrogen:quinone oxidoreductase (5.11)

¹Numbers in parentheses represent percent contribution to within-group similarity or between-group dissimilarity

1.4 Discussion

The microbial community inhabiting the surface of bones changed during colonization and mat development. Specifically, the community profiles changed on a weekly basis. This was evident by the fact that samples taken within the same week branched together in the cluster diagrams (Figure 1.5). While Proteobacteria comprised the dominant phylum present on the surface of the bone (Figure 1.2), there were shifts in two classes throughout the course of this experiment. Gammaproteobacteria were the initial colonizers of the bone surface and remained high in relative abundances from week 1 (Samples FB1, 2 and BB1, 2) to approximately week 3 (FB5, FB6) (Figure 1.3). As time went by, this class decreased in relative abundance, allowing for Epsilonproteobacteria to dominate (Figure 1.4). A noteworthy difference between the findings of the present study and previous literature was the underrepresentation of Deltaproteobacteria. One stage described in whale fall studies is known as the sulfophilic stage (Smith et al., 2015). During this stage Deltaproteobacteria are present in high abundances and are thought to aid in the decomposition of additional whale tissue and bone lipids, as well as the increase in hydrogen sulfide via sulfate reduction (Deming et al., 1997; Goffredi & Orphan, 2010). Since the bones from the present study underwent thermal processing via autoclaving it is possible that the initial bone composition or lipid content may have been altered, preventing the anticipated bloom of sulfate reducing bacteria.

OTUs within the order of Thiotrichales were detected in the first day, followed by an immediate decrease in subsequent sampling. This boom-and-bust shift in the communities may have resulted from the increase in nutrient resource availability from the bone (Fowler & Winstanley, 2018). In the deep sea, a whale fall acts as a

localized source of nutrients that sparks the growth of microbial communities that are different from those that exist in sediments near the carcass (Goffredi & Orphan, 2010). In the presence of the pork bone, chemoheterotrophic Gammaproteobacteria may have been the first colonizers to use readily available carbon from the bones. OTUs related to *Leucothrix*, were abundant in FB1. *Leucothrix* is a chemoheterotrophic bacterium within the Thiotrichales order (Brock, 2006). On the other hand, OTUs related to the *Candidatus Thiopilula* spp. were most abundant in BB1. This candidate taxon has been found in mats associated with cold seeps. A metatranscriptomic analysis proposed that under anoxic conditions the cell type was able to oxidize sulfur through nitrate respiration, due to the expression of genes necessary for the oxidations of intracellular sulfur globules to sulfite and adenosine-5'-phosphosulfate, as well as several genes linked to nitrate reduction (Jones, Flood, & Bailey, 2015). In the present study, no MAGs associated with this organism was detected within the metagenomic datasets. Therefore, the potential preference of growth for this organism could not be identified from the available data.

In addition to the boom-and-bust community shifts that were observed in the two microbial mats studied, several microorganisms showed distinct shifts in abundance depending on the sample source. Aside from the OTUs related to *Leucothrix* and *Thiopilula* that were previously discussed, OTUs related to *Arcobacter* and *Beggiatoa* were in extremely low abundances (0.008% and 0.002%) in samples from incoming water (Table 1.1). The shift in abundance of *Arcobacter* OTUs increased up to 55% of the total community of the BB mat. This shift suggests that microorganisms may be present and capable of surviving in extremely low numbers in environments that are not suitable for their growth, as shown in experiments in which

serially diluted coastal seawater was used to inoculate wood samples. It was discovered that as little as one cell in 10L of seawater was able to function (D. Kalenitchenko, Le Bris, Peru, & Galand, 2018). Therefore, it is likely that the chemosynthetic *Arcobacter* maintained viability in incoming water and then increased in numbers once it came across the more favorable environment of the bone surface.

At the ecosystem level, metagenomics did not reveal any shifts in functional potential over time (Figure 1.7). In fact, the mat community in both bones contained complete pathways for heterotrophic and autotrophic growth, as well as aerobic and anaerobic respiration, which were observed across all sample times.

Epsilonproteobacteria may have been responsible for carbon fixation via the rTCA cycle, as shown in the genome FB7_008, which was related to *Sulfurovum*. Other Epsilonproteobacteria have been shown to fix carbon, such as *Sulfurimonas* and *Thiovulum* (Wirsen & Jannasch, 1978; Grote, Jost, Labrenz, Herndl, & Jürgens, 2008; Marshall, Blainey, Spormann, & Quake, 2012). Despite being in low relative abundances in the 16SrRNA amplicon data of this study, genomes related to these genera were also detected, so they may have been contributing to the organic carbon pool associated with the mat communities on the bone surfaces.

The two most abundant classes of the mat communities, Gammaproteobacteria and Epsilonproteobacteria, are known to play an active role in the oxidation of reduced sulfur compounds in a variety of environments, including hydrothermal vents, whale falls, wood falls and marine sediments (Deming, Reysenbach, Macko, & Smith, 1997; Yamamoto & Takai, 2011; Kalenitchenko et al., 2018; Wasmund, Mußmann, & Loy, 2017). Epsilonproteobacteria found in deep-sea environments have shown to contain genes necessary for the oxidation of sulfur using the Sox multienzyme

complex, a collection of enzymes that aid in the oxidation of inorganic sulfur compounds, as well as genes coding for polysulfide reductase (*psr*), which is necessary for polysulfide respiration (Yamamoto & Takai, 2011). The two epsilonproteobacterial MAGs FB1_004 and FB7_008 in the present study did not contain all the genes necessary for the Sox complex, or genes coding for Psr. However, genes for sulfide-quinone oxidoreductase (*sqr*) were found. Additionally, genes encoding for sulfate adenylyltransferase (*sat*) were present in the same MAGs, which is used for reverse sulfate reduction (Yamamoto & Takai, 2011). Genes for the complete Sox system were present in the *Marinobacterium* MAG, BB4_007. Furthermore, *sat* genes were not present in the two gammaproteobacterial genomes BB4_007 and BB1_001. However, *sqr* genes were detected in the gammaproteobacterial MAGs, suggesting that in these microbial mat communities, putative pathways for sulfur oxidation show some overlap between distinct classes.

Pathways for nitrogen metabolism showed some distinctions between the two classes. While all MAGs encoded genes for nitrate reduction the periplasmic nitrate reductase system (Nap) was found in the two epsilonproteobacterial MAGs, and the gammaproteobacterial MAG related to *Beggiatoa*. This pathway for respiratory nitrate reduction has been previously detected in deep-sea Epsilonproteobacteria and *Beggiatoa* filaments from marine sediments and studies have suggested that it is necessary for these organisms to respire under anaerobic conditions (Mußmann et al., 2007; Vetriani et al., 2014). Nitrogen fixation potential was complete in the MAG related to the *Arcobacter* genus. Other non-pathogenic members of this genus are known to have nitrogen-fixing potential (Pati et al., 2010; Roalkvam et al., 2015). The presence of the complete pathway in this MAG suggests that this genus may have

been playing a key role in supplying the microbial mats in this study with bioavailable nitrogen. In contrast, the MAG related to the genus *Beggiatoa* showed the potential for complete denitrification. Denitrification potential in *Beggiatoa* has been tested in hydrothermal systems, cold seeps, as well as freshwater sediments (Sweerts et al., 1990; Bowles & Joye, 2011; Bowles, Nigro, Teske, & Joye, 2012; Schutte et al., 2018). Therefore, it is possible that in the mats that grew on the bones of this study, *Beggiatoa*-related organisms further aided in the removal of nitrate from the mats.

Here we present findings on the successional patterns of microbial mats that grew on the surface of bones when exposed to water from the Broadkill River in Lewes, Delaware. Similar studies of microbial mat communities growing in whale fall habitats have focused on examining the taxonomic composition of these communities several months after bones were discovered or deposited at the sea floor (Goffredi & Orphan, 2010; Hilario et al., 2015). We were able to show that within the first few weeks of development, the microbial communities shifted from Gammaproteobacteria to Epsilonproteobacteria. Despite the changes in community dynamics of the mats, functional redundancy appears to be a characteristic of these ecosystems. Several of the colonizers were microorganisms that have been detected in chemosynthetic mats growing on whale falls, wood falls, cold seeps and hydrothermal vents. Therefore, the use of metagenomic techniques may be helpful in identifying key members of microbial mats from a variety of habitats in order to understand how they participate in the biogeochemical cycling of carbon, sulfur and nitrogen, thus furthering the knowledge of the elusive organisms associated with these types of ecosystems.

Chapter 2

UNCOVERING KEY DIFFERENCES IN PHYLOGENY AND FUNCTIONAL POTENTIAL OF BEGGIATOA GENOMES

2.1 Introduction

Of the many microorganisms inhabiting sulfur-oxidizing mats, *Beggiatoa* have is a genus with an extensive history, dating back to the first microbial ecologist Sergei Winogradsky, who keyed the term chemolithotrophy. By conducting a series of experiments on *Beggiatoa* filaments, he determined that this bacterium could utilize reduced inorganic compounds, such as hydrogen sulfide to gain energy (Dworkin, 2012). Because of this lifestyle, *Beggiatoa* filaments are found in environments that are rich in sulfur, such as salt marshes, marine and freshwater sediments, as well as hydrothermal vents and cold seeps (Strohl & Larkin, 1978; Jannasch, Nelson, & Wirsén, 1989; Mills et al., 2004; Mußmann et al., 2007).

Early attempts to characterize filaments of this genus described them based on morphological features, such as length and width size, as well as the presence of intracellular globules and their chemotactic response to opposing gradient levels of oxygen and hydrogen sulfide. Phylogenetic analyses of single cells revealed that members of this genus cluster into separate groups based on morphology, as well as environmental origin (Salman et al., 2011). For example the type species *Beggiatoa alba*, which was isolated from freshwater sediments, is less related to candidate genera of filaments that were retrieved from brackish or marine sediments (Salman, Bailey, & Teske, 2013). In addition to *B. alba*, single-cell sequencing analyses have revealed that the proposed family *Beggiatoaceae* is comprised of sub-clusters of large sulfur-oxidizing bacteria. Included in this family are filamentous bacteria that had been

identified as *Beggiatoa* in previous scientific literature, but based on phylogenetic data were separated into five separate branches (Salman et al., 2011). Two examples of these clusters are the candidate genera *Maribeggiatoa* that consists of filaments isolated from hydrothermal vents, and *Isobeggiatoa* that is comprised of filaments originating from brackish habitats.

The lack of pure cultures from marine *Beggiatoa* has made it difficult to advance these genera past the candidate level. However, whole genome amplification methods and metagenomic techniques have uncovered that these organisms contain the genetic tools necessary to cope with potential nutrient changes and shifts in availability of electron donors and acceptors within their environments. These include the potential to use different sulfur oxidation pathways, the capabilities of relying on either heterotrophic or autotrophic methods of growth, and finally, the possibility of nitrate respiration under oxygen-limited conditions (Mußmann et al., 2007; MacGregor, Biddle, Harbort, Matthysse, & Teske, 2013; Sharrar et al., 2017). It has also been suggested that the wide range of metabolic capabilities attributed to these organisms may have resulted from genetic exchange with other species that have been in close contact with *Beggiatoa* filaments (Mußmann et al., 2007; MacGregor, Biddle, & Teske, 2013), findings that may have not been possible without the use of genomic sequencing.

In this study, we sought to find more *Beggiatoa*-related genomes and determine what phylogenetic and putative metabolic functions were attributed to the genomes retrieved from metagenomes of microbial mats grown on bones that were deposited in high salinity river water. Putative functional pathways were compared to those of freshwater isolates and filaments from deep-sea mats, including two new

genomes from metagenomes and two new genomes from single filaments. A pangenome was constructed to discern what genes might be conserved or diminished across datasets derived from different habitats. Finally, taxonomic classifications of the annotated proteins from each genome were generated to understand patterns of gene history. Together, these findings show that location may influence the potential functional and regulatory capabilities that *Beggiatoa* genomes have, as well as how their genomes may be affected by other related organisms within their diverse habitats.

2.2 Materials and Methods

2.2.1 Genome collection

A total of 17 *Beggiatoa* genomes or metagenome-assembled genomes (MAGs) were used to answer the questions raised in this chapter (Table 2.1). Nine of these genomes were identified from the metagenomic samples described in Chapter 1. Only one genome (FB4.032) showed contamination levels higher than 10% and was refined through Vizbin (Laczny et al., 2015). The GOM8 (Gulf of Mexico) and GUB9 (Guaymas Basin) genomes came from freshly-collected filaments. Their reported widths were 35 μm for the orange filament GOM8 and 100-120 μm for the colorless filament GUB9 (Andreas Teske, personal communication). GOM8 and GUB9 were initially whole genome amplified (WGA) with the RepliG minikit, with modifications cited in (MacGregor, Biddle, & Teske, 2013), and then submitted for sequencing at the UT Austin Genomics Facility (Verena Salman, personal communication). The orange Guaymas Basin filament genome (boguay) was downloaded from JGI (GOLD project ID: Gp0006298). Two genomes came from known isolates *Beggiatoa leptomitiformis* (accession number: CP018889.1) and *Beggiatoa alba* (accession number: NZ_JH600070.1). The genome for a white filament in marine sediments from

Eckernforde Bay was downloaded from GenBank (ABBZ00000000 for pyrosequenced dataset). Two MAGs were provided from Sean McAllister. These genomes (SM.98 and SM.70) came from iron-oxidizing mats at the Mid-Atlantic Ridge Rainbow hydrothermal vent field, and at the Loihi Seamount hydrothermal vents respectively (Sean McAllister, personal communication).

2.2.2 Metagenome-assembled genome analysis

The nine MAGs from bones were analyzed as mentioned in Chapter 1. GOM8 and GUB9 underwent the same assembly (SPADES), genome assessment (CheckM) and refinement (Vizbin) as all other MAGs in Chapter 1. To obtain bins SM.98 and SM.70 Spades assemblies from each library were binned separately with Maxbin, Metabat, Concoct, and Binsanity and the best of the resulting MAGs were picked using the dereplication, aggregation, and scoring tool (Alneberg et al., 2014; Kang, Froula, Egan, & Wang, 2015; Wu et al., 2014; Graham, Heidelberg, & Tully, 2017). Completeness/contamination from the resulting bins were assessed using CheckM (Parks et al., 2015). Phylogenetic affiliations of the bone MAGs as well as, GOM8, GUB9, SM.98, and SM.70 were done with Phylosift (Darling et al., 2014).

Putative functional annotations were obtained through Prokka (Seemann, 2014). All 17 genome assemblies were uploaded to the Department of Energy Systems Biology Knowledgebase (KBase) platform (Arkin et al., 2016) to use the “Species Tree Construction” application in order to place the genomes into a tree based on 49 highly conserved Clusters of Orthologous Groups (COG) families (www.github.com/kbaseapps/SpeciesTreeBuilder/tree). The 16S rRNA gene could not be found in some of the genome assemblies retrieved from the metagenomes. Specifically, 16S rRNA genes were only found in BB1.001, BB4.002, FB7.010, and

FB4.032. Therefore, a search for other ribosomal protein annotations that were shared across all the genomes from the bones, as well as the additional genomes used in this study, was performed to infer further about their phylogenetic affiliations. Annotations for the 30S ribosomal protein S16 were used, as it was found in nearly all the genomes, excluding SM.70. The annotated protein sequences were aligned in MEGA7 using MUSCLE and a neighbor-joining tree was constructed (Kumar, Stecher, & Tamura, 2016).

As in Chapter 1, the annotated genomes were uploaded to GhostKOALA and underwent the same analysis with the KEGG decoder script (Tully et al., 2018) to determine pathways that were complete in each genome. Additionally, using the GFF3 output from Prokka for the 17 genomes listed, a pangenome was constructed using Roary (Page et al., 2015). A pangenome for the solely for bone MAGs (excluding FB4.032) was also constructed. The default settings were used. The presence-absence matrix of functional annotations in each genome were collected and used to generate a heatmap in order to identify if certain annotations showed a differential presence pattern in the *Beggiatoa* genomes.

The outputs from GhostKOALA and Roary were subjected to SIMPER tests to identify pathways or functional annotations that mostly contributed to differences between the *Beggiatoa* genomes (Clarke & Gorley, 2006). The groups in which the genomes were separated into were “Deep” and “Shallow”. The deep group included genomes from hydrothermal vents or seep systems, while the shallow group included the two isolates (b.alba, b.lep), the bins from the pork bone metagenomes, and the genome from Eckenforde Bay marine sediments (ABBZPS).

The predicted protein sequences for each genome were compared against the NCBI nr protein database through Diamond (Buchfink, Xie, & Huson, 2014). The output was used in DarkHorse to determine possible relatives of the *Beggiatoa* genomes that may have contributed to horizontal gene transfer (Podell & Gaasterland, 2007). The taxonomic affiliations of the protein predictions for the genomes were collected for Multidimensional Scaling using PRIMER (Clarke & Gorley, 2006).

2.3 Results

2.3.1 Phylogenetic affiliations of *Beggiatoa* genomes

Two phylogenetic trees were constructed to determine how the *Beggiatoa* genomes of this study are related (Figure 2.1). In the 30S ribosomal protein S16 tree, eight of the nine genomes from the bone microbial mat metagenomes formed a cluster that was separate from the species isolates. The ninth one, FB4.032, as well as SM.98 did not cluster with any known *Beggiatoa* genomes. The GOM8 and GUB9 genomes from hydrothermal vent filaments formed a separate cluster with the Boguay filament from Guaymas Basin. The ABBZPS genome from marine sediments in Eckenforde Bay was in the same clade. The sequence for the 30S ribosomal protein S16 subunit could not be found in SM.70. Therefore, this genome was excluded from the tree construction (Figure 2.1A).

Similar clusters were seen in the species tree resulting from concatenated COGs from the KBASE application “Insert Genomes into Species Tree”. All pork bone genomes formed a separate clade, except for FB4.032, which branched out on its own. GUB9 and Boguay, along with GOM8 from the Gulf of Mexico and ABBZPS

separated out from the *Beggiatoa alba* and *Beggiatoa leptomitiformis* branch. The SM.98 and SM.70 separated out from all other *Beggiatoa* genomes (Figure 2.1B).

2.3.2 Putative functional annotations and metabolic pathways detected in *Beggiatoa* genomes

The *Beggiatoa*-related genomes showed variations in pathways that were complete. The average dissimilarity between shallow and deep genomes was 37.5%. According to the SIMPER test, the pathways and annotations contributing most to the dissimilarity between the two groups of genomes were related to anabolic pathways, carbon fixation, as well as sulfur and nitrogen metabolism (Table 2.2).

The complete set of genes necessary for glycolysis was not present in any of the genomes (Figure 2.2). This observation was made in all bins, regardless of percent completeness of the genomes (Table 2.1). The only carbon fixation pathway found in these genomes was the Calvin Benson Bassham (CBB) cycle, which was complete in all the deep genomes except for SM.98 and SM.70. The CBB cycle was also complete in the b.lep and b.alba isolate genomes, the ABBZPS and the FB4.032 genomes (Figure 2.2). No other genomes from the bone microbial mat metagenomes had genes for carbon fixation pathways. Dissimilatory nitrate reduction to ammonium (DNRA) was complete in all the pork bone retrieved genomes, the *Beggiatoa* isolates, as well as GOM8, and SM.98. Nitrite reduction, nitric oxide reduction and nitrous oxide reduction were complete in the bone genomes, except for FB4.032. ABBZPS did not contain genes for nitrous oxide reduction, while the isolate genomes, as well as SM.70 and GUB9 did not contain any genes for denitrification. From the deep genomes, only GOM8 could go to complete denitrification. Nitrite and nitric oxide reduction were complete in boguay, but nitrous oxide reduction was not. On the other hand, nitrous

oxide reduction and nitrite reduction were complete in SM.98, and nitric oxide reduction was not. Nitrogen fixation pathways were complete for *b.lep* and *b.alba*. In FB4.032, only the *nifD* gene was present. The dissimilatory sulfate reduction and oxidation pathway was complete in ABBZPS, GOM8 and boguay. The reversible conversion of sulfate to adenosine phosphosulfate (APS) was present in GUB9 and SM.98 (Figure 2.2). The reversible conversion of sulfite to sulfide was also present in GUB9. Thiosulfate oxidation was not complete in any genome. However combinations of the SOX system were present in all the genomes except for SM.70. Sulfide oxidation was complete in all the genomes except GOM8. However, NiFe hydrogenase Hyd-1 was present in the pork bone genomes, except for FB4.032. Transporter systems were not complete, with exception of the phosphate transporter system. The glyoxylate shunt was complete in the isolate genomes, as well as the pork bone bins, except for FB4.032 (Figure 2.2).

From the pangenome constructed in Roary, noticeable differences in counts of several other functional annotations were found (Figure 2.3). With the exception of FB4.032, all the bone genomes had high counts of adenylate cyclase 1, the regulatory protein RpfC, the autoinducer 2 sensor kinas LuxQ, and the histidine-protein kinase BarA. Counts for these functional annotations were also high in ABBZPS, and GOM8. In these two genomes, high counts for functional annotations related to the endonuclease VapC and 3'3-cGAMP-specific phosphodiesterase 2 were also observed. *B.lep* and *B.alba* had lower counts of annotations for adenylate cyclase and the regulatory protein RpfC, but higher counts for the aerobic respiration control sensor protein ArcB, and the phosphoserine phosphatase RsbU. The colicin I receptor had more annotations in several of the bone MAGs, but did not have a high presence

in other genomes. The genomes FB4.032, SM.70, and SM98 showed little to no presence of the functional annotations listed in the heatmap (Figure 2.3). When all 17 genomes were used, the total number of genes, and the number of unique genes increased as more genomes were added (Figure 2.4A;Figure 2.4B). When generating a pangenome for eight of the nine genomes (excluding FB4.032) from the bone microbial mats, the amount of conserved genes, as well as unique genes did not increase (Figure 2.4C;Figure 2.4D). When FB4.032 was added in the pork bone microbial mat pangenome, results were similar to those generated by the 17-genome dataset (data not shown).

Differences were also observed in the gene history of the putative functions attributed to the genomes of this study. Multidimensional scaling of the related taxonomies as shown through the DarkHorse analysis indicate that most of the bone genomes had similar taxonomic classifications of annotated proteins, except for FB4.032 (Figure 2.5A). The classifications for SM.70, SM.98, GOM8, GUB9, and boguay, were less than 75% similar to the pork bone bins, and the isolate and ABBZPS classifications were completely different from the other genomes. For the isolate genomes and ABBZPS, a minimum of 98% of the taxonomic classifications was attributed to the order of Thiotrichales. Thiotrichales was also responsible for 69% to 73% of the annotations in GUB9, GOM8 and boguay, 40% to 49% of the classifications in the bone MAGs, but only 4% in SM.70 and SM.98. In these two genomes, Chromatiales comprised 20% and 25% of the classifications. In the remaining genomes, Chromatiales only contributed approximately 4% of the taxonomic classifications. Unclassified Gammaproteobacteris and Methylococcales were in higher relative abundances in SM.70 and SM.98 than in other genomes as

well. The highest relative abundances of annotations attributed to Cyanobacteria were found in GOM8, GUB9 and boguay (Figure 2.5B).

Table 2.1 Genomes and genomic bins related to the genus *Beggiatoa*.

Source of Bin/Genome	Bin/Genome name	Completeness	Contamination	Habitat
FB1	FB1.005	97.1	4.81	Mat on pork bone
FB2	FB2.015	93.53	2.23	Mat on pork bone
FB4	FB4.032	61.65	0	Mat on pork bone
FB4	FB4.003	98.28	5.02	Mat on pork bone
FB7	FB7.010	95.69	2.32	Mat on pork bone
BB1	BB1.001	97.28	1.7	Mat on pork bone
BB4	BB4.002	97.1	1.73	Mat on pork bone
BB5	BB5.009	94.3	2.93	Mat on pork bone
BB7	BB7.004	95.47	1.35	Mat on pork bone
Sean McAllister	SM.98	98	0.8	Rainbow Hydrothermal field (Mid-Atlantic Ridge)
Sean McAllister	SM.70	70	8.2	Loihi Seamount vent
Orange filament	GOM8	73.878	5.174	Cold seep system in Gulf of Mexico
White filament	GUB9	68.669	2.603	Hydrothermal system in Guaymas Basin
Orange filament	boguay	90.96	2.62	Hydrothermal system in Guaymas Basin (PRJNA19285)
White filament	ABBZPS	89.76	8.66	Marine sediment in Eckernforde Bay (ABBZ00000000)
<i>Beggiatoa leptomitiformis</i>	b.lep	98.45	1.45	Isolate (PRJNA358423)
<i>Beggiatoa alba</i>	b.alba	98.81	1.82	Isolate (PRJNA62137)

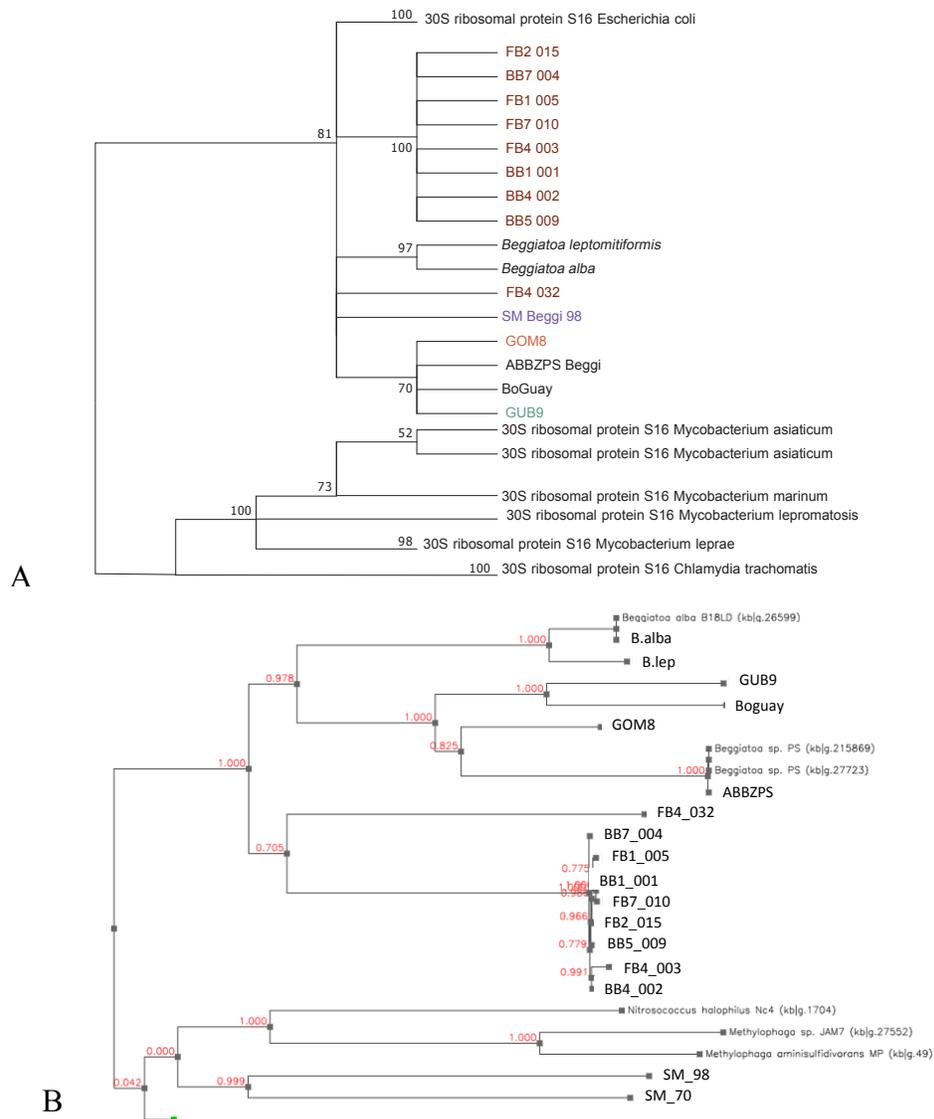


Figure 2.1 Phylogenetic affiliations of *Beggiatoa* genomes. Panel A: Neighbor-joining phylogenetic tree of the S16 subunit of the 30S ribosomal protein. The translated amino acid sequences were analyzed in MEGA7 and aligned with MUSCLE, with 200 bootstraps. Panel B: Species Tree from KBASE *Insert Genomes into Species Tree* app. This app selects an appropriate set of reference alignments based on 49 highly conserved clusters of orthologous groups (COG) families. Those sequences are extracted from the query genomes and concatenated before using FastTree2 to construct a tree.

Table 2.2 Results of SIMPER test between deep and shallow *Beggiatoa* genomes

Within-Group Similarity (Shallow)	Within-Group Similarity (Deep)	Between-Groups Dissimilarity (Shallow vs Deep)
Sulfide oxidation (5.42)	Beta-N-acetylhexosaminidase (7.31)	Glyoxylate shunt (4.84)
Transporter: phosphate (5.16)	NADH-quinone oxidoreductase (6.39)	Dissimilatory sulfate < > APS (4.17)
F-type ATPase (4.77)	TCA Cycle (5.12)	NiFe hydrogenase Hyd-1 (4.02)
NADH-quinone oxidoreductase (4.67)	Sulfide oxidation (4.6)	DNRA (3.57)
DNRA(4.61)	Riboflavin biosynthesis (4.51)	RuBisCo (3.55)
Riboflavin biosynthesis (4.51)	Glycolysis (4.49)	CBB Cycle (3.38)
Beta-N-acetylhexosaminidase (4.43)	RuBisCo (4.46)	Nitric oxide reduction (3.36)
NiFe hydrogenase Hyd-1 (4.24)	Transporter: phosphate (4.41)	Cytochrome bd complex (3.27)
TCA Cycle (4.02)	Chemotaxis (3.94)	Dissimilatory sulfite < > sulfide (3.21)
Glycolysis (3.7)	Dissim nitrate reduction (3.88)	Nitrous-oxide reduction (3.19)
Glyoxylate shunt (3.65)	Dissimilatory sulfate < > APS (3.88)	Sulfite dehydrogenase (quinone) (3.13)
Chemotaxis (3.57)		Nitrite reduction (2.8)
		Mixed acid: ethanol, acetate to acetylaldehyde (2.7)
		D-galacturonate epimerase (2.64)
		Cytochrome c oxidase (2.62)

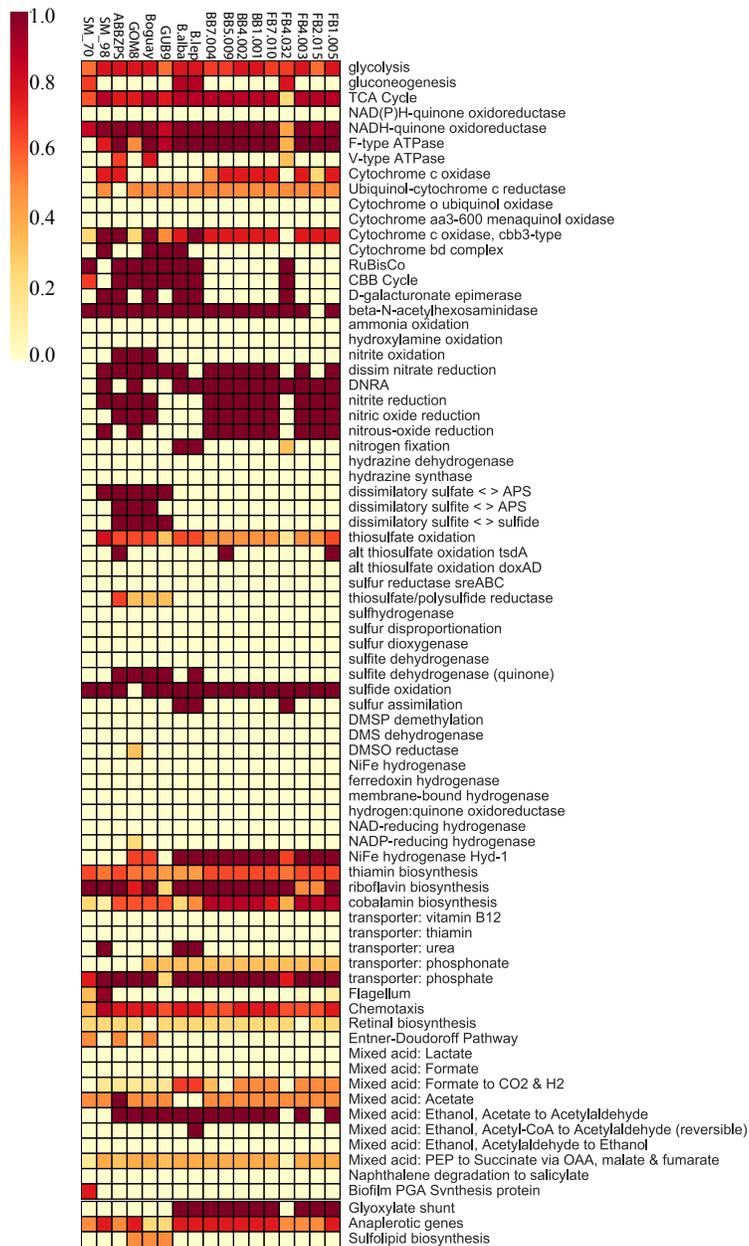


Figure 2.2 KEGG annotated pathways that are present in the tested genomes. Dark red boxes indicated that the pathway is complete. Lighter red are less complete pathways, and yellow are pathways that are not present within the genomes tested.

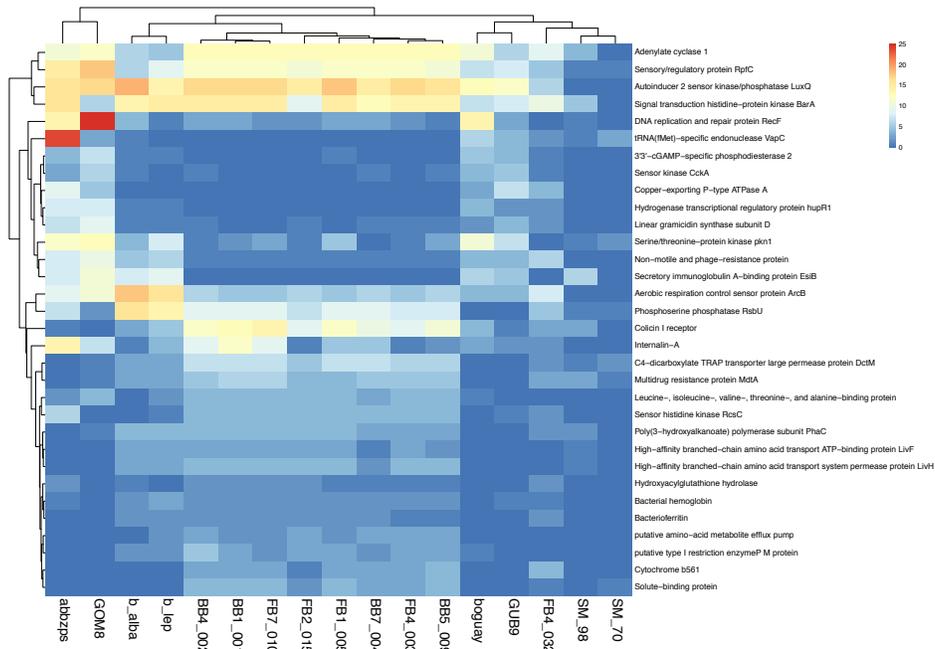


Figure 2.3 Heatmap of functional annotations present in the genomes, after the Roary pangenome construction. Warmer colors represent functional annotations with higher presence values. Cooler colors represent annotations with lower presence values. The annotations that were selected for plotting were determined after a SIMPER test based on “Deep” and “Shallow” *Beggiatoa* genomes showed which annotations contributed to 50% of the between-group dissimilarity.

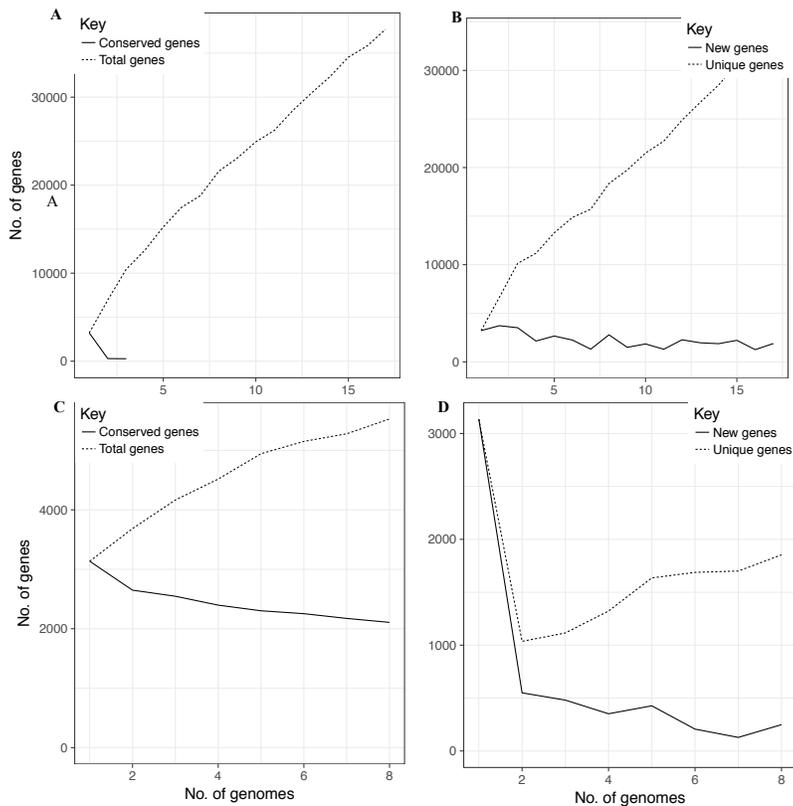


Figure 2.4 Gene annotations as a function of number of genomes. Panel A and B: Pangenome for all *Beggiatoa* genomes. Numbers of conserved and total genes versus number of genomes (17 total) are plotted in A and in Panel B new and unique genes are plotted versus number of genomes. Panel C and D: Pangenome for the *Beggiatoa* genomes from the pork bone microbial mats, excluding FB4.32. Numbers of conserved and total genes versus number of genomes (8 total).

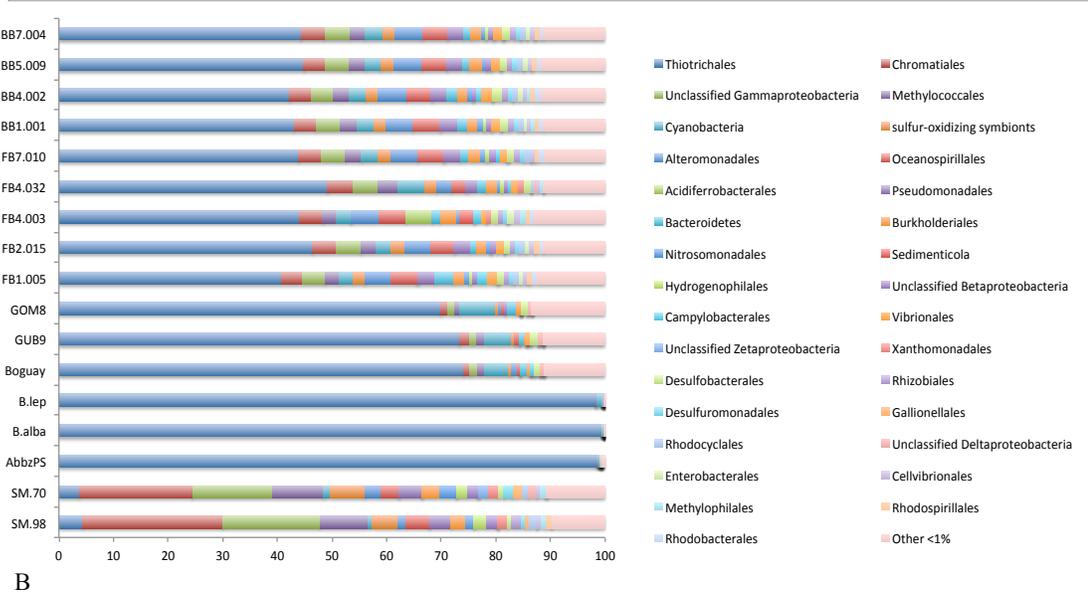


Figure 2.5 Gene history of *Beggiatoa* genomes. Panel A: MDS plot of taxonomic classifications of the annotated proteins in the *Beggiatoa* genomes of this study, as determined by DarkHorse analysis. The abundance values were square root transformed and a Bray-Curtis dissimilarity matrix was constructed. The green circle represents 60% similarity, and the blue one represents 75% similarity. Panel B: Percent relative abundance of the taxonomic classifications.

2.4 Discussion

The filamentous sulfide-oxidizing bacterium *Beggiatoa* has been found in microbial mats across a variety of habitats worldwide. Currently, only freshwater filaments are maintained in pure cultures (Mezzino, Strohl, & Larkin, 1984; Grabovich et al., 1998). While lacking in cultures, genomes from filaments originating from marine and brackish habitats have uncovered the wide physiological potential that can be attributed to these organisms (Mußmann et al., 2007; MacGregor, Biddle, & Teske, 2013; Flood, Bailey, & Biddle, 2014; Sharrar et al., 2017; Schutte et al., 2018).

We were able to add 11 new genomes of *Beggiatoa* to seven existing genomes to allow a comparative study of functional potential and taxonomic divergence across these lineages. Since 16S rRNA genes could not be found in all the genomes, the translated protein sequences for the 30S ribosomal protein S16 subunit was used for the construction of one of the phylogenetic trees, as it was one of the most widespread ribosomal proteins in the genomes. Difficulties in obtaining 16S rRNA genes from large sulfur bacteria have been noted previously (Salman et al., 2013). As a second for phylogenetic inference, the “species tree” generated through the online tool KBase is also presented (Figure 2.1B). The reference set of COGs used by this application already includes the genomes for *Beggiatoa alba* and the pyrosequenced filament from Baltic sea sediments (noted as ABBZPS in this study, and as *Beggiatoa* sp. PS in the KBase reference). Therefore, since the b.alba and ABBZPS genomes were also included as part of the query genomes, they appeared as duplicates in the tree (Figure 2.1B). These duplicated genomes allow for some added confidence regarding the overall positioning of the other genomes in the species tree. On the phylogenetic level,

most of the pork-bone originating genomes lie intermediately between the pure freshwater isolates and the marine genomes (Figure 2.1A; Figure 2.1B). Interestingly, the genome FB4.032 appears to be less related to the other *Beggiatoa* MAGs that inhabited the bone microbial mat suggesting that, even in a single mat, differentially related *Beggiatoa* may co-exist. This has been observed where orange and white *Beggiatoa* filaments at hydrothermal systems in the Guaymas Basin can cohabit the same mat, in which they distribute themselves as a result of different thermal tolerances, and electron donor availabilities (McKay et al., 2012; Schutte et al., 2018). Interestingly, GUB9, a large (~100 μm) white filament from the Guaymas Basin, was placed close to the thin (35 μm) orange filament Boguay from the same location (width of boguay reported in MacGregor et al., 2013). Orange filaments from this vent system are known as the Candidate genus *Maribeggiatoa*, but some studies show that the large white filaments are more distantly related (Salman et al., 2013). GOM8, a thin (~35 μm) orange filament from a seep system located in the Gulf of Mexico, was not placed in the same clade as boguay in the species tree, even though they were the same size (Figure 2.1B). This could suggest that orange and colorless filaments from the same geographic locations or hydrothermal systems are more closely related to each other than to pigmented and non-pigment filaments from different areas. However, more filament samples need to be sequenced to confirm this suggestion. Aside from geography and seep system, mat type may lead to further phylogenetic differentiation; SM.98 and SM.70 were two genomes retrieved from iron-oxidizing mats that were even more distant from the other *Beggiatoa* genomes used in this study (Figure 2.1B). Iron-oxidizing and other microbial mats may hold novel *Beggiatoa*-

related bacteria that have yet to be discovered and may be good targets for future exploration.

While new taxonomic lineages may still exist, it is unclear how much genetic novelty may be found within known *Beggiatoa* lineages. To determine this, we applied a pangenome investigation. Traditionally, the concept of a pangenome has been applied to pathogens in order to identify and understand genomic diversity between virulent and non-virulent strains of bacteria (Tettelin et al., 2005; Donati et al., 2010; Kaas, Friis, Ussery, & Aarestrup, 2012). More recent attempts have focused on identifying core and accessory genes of archaea and bacteria that are environmentally relevant (Beck, Knoop, & Steuer, 2018; Delmont & Eren, 2018; Moldovan & Gelfand, 2018; Reno, Held, Fields, Burke, & Whitaker, 2009). Given the large attempts to obtain genomic sequences of *Beggiatoa* filaments from diverse habitats, constructing a pangenome may provide further insight on how these organisms have evolved over time.

Pangenomes can be broadly separated into two groups: closed, and open (Vernikos, Medini, Riley, & Tettelin, 2015). A closed pangenome, such as that of *Pseudomonas aeruginosa*, is one in which the addition of new sequenced genomes does not add new genes to the genetic repertoire of the organism that is being studied. (Mosquera-Rendón et al., 2016). An open pangenome, such as that of *Streptococcus agalactiae*, is characterized by an expansion of the number of newly found genes with the increase in number of sequenced organisms (Tettelin et al., 2005).

The *Beggiatoa* pangenome presented in this study resembles what would be considered an open pangenome, when including *Beggiatoa* from different environments (Figure 2.4A). However, the pangenome generated from the eight

genomes of the bone microbial mats resembles that of a closed pangenome (Figure 2.4C). It appears that pangenomes from different sub-types of *Beggiatoa* are distinct from each other. In the archaeon *Sulfolobus islandicus*, Reno et al (2009) proposed that the pangenome of this organism is organized based on biogeography, and that genetic flow between separated populations was limited. In the class of Epsilonproteobacteria, Zhang and Sievert (2014) suggested that environmental conditions may have influenced the structure of the class's pangenome and may have led to speciation events in different habitats, such as vent and non-vent systems. Within *Beggiatoa*, the presence:absence matrix of functional annotations seems to separate based on both geography and habitat (Figure 2.3). The genomes from the pork bone microbial mats, excluding FB4.032, show very similar patterns in functional gene presence, as do both of the isolates. The genomes from different vent systems are more distant. The filaments from the Guaymas Basin vent system (GUB9 and boguay) are more distant from the filament isolated at the vent system in the Gulf of Mexico (GOM8), as are SM.70 and SM.98 which originate from the metagenomes of iron-oxidizing mats of Loihi Seamount ridge and Rainbow hydrothermal field, respectively. In order to further explore the concept of a closed *Beggiatoa* pangenome, we would need to sequence more environments and habitats in which they are found.

Counts of gene annotations necessary for regulatory processes, such as those encoding for adenylate cyclase were high in the genomes from the pork bone microbial mats. This family of adenylate cyclases (AC) is known to convert ATP into cyclic AMP (cAMP), a secondary messenger used for signal transduction in eukaryotic and prokaryotic organisms (Cooper, 2003). Some AC's are stimulated by blue light, and increase the conversion of ATP to cAMP (Iseki et al., 2002). Recently,

a DNA sequence encoding for one such photoactivated AC was identified in the pyrosequenced genome of the same *Beggiatoa* filament used in this study (labeled ABBZPS). In *E.coli* cells the expression of this protein increased cAMP production by two fold when exposed to blue-light irradiance (Ryu, Moskvina, Siltberg-Liberles, & Gomelsky, 2010). This specific AC is being used in higher-level organisms, due to its small size and better performance in cAMP production than the equivalent eukaryotic AC (Stierl et al., 2011). In fact, it has been proposed as a future tool in the field of optogenetic therapy (Kim, Folcher, Baba, & Fussenegger, 2015). Given the proposed method of action, this AC may function as a physiological response to the exposure of *Beggiatoa* to light. In early experimental observations, natural mats and cultivated filaments that were exposed to blue light showed a negative phototactic response, moving away from the light source (Schwedt, 2011). Mat disruption was also observed when cAMP was added to the medium used in cultivation experiments with the marine strain *Beggiatoa* sp. 35Flor (Nelson & Castenholz, 1982; Schwedt, 2011). In nature, the shallow-dwelling *Beggiatoa* filaments are exposed to more light than deep-dwelling filaments, which could explain why more genes encoding for adenylate cyclase are present in the pork bone microbial mat genomes. One characteristic of the spontaneous oxidation of sulfide is the production of sulfur- and oxygen- free radicals, resulting in chemiluminescence (Tapley, Buettner, & Shick, 1999). Schwedt (2011) proposed that this chemiluminescence might influence mat development and filament migration. Chemiluminescence during sulfide oxidation may be the reason light emission is sometimes observed at hydrothermal vents (Tapley et al., 1999). Therefore, some genomes, such as GOM8 and boguay that came from vent systems, may be able to regulate their movement patterns via adenylate cyclase in a similar

fashion to their shallow relatives. This could suggest that gene regulation, and not overall gene function, may contribute to differences in *Beggiatoa* genomes.

The potential for quorum sensing in *Beggiatoa* may also be present, as suggested by the high number of genes annotated for the sensory/regulatory protein RpfC and the kinase/phosphatase LuxQ (Figure 2.3). In *Xanthomonas campestris*, RpfC was proposed to directly interact with the diffusible signaling factor (DSF) altering the levels of cyclic diguanylate (c-di-GMP) available in the cell. In conditions of high cell densities and increased DSF, levels of c-di-GMP became reduced, allowing for Clp proteins to function and aid in biofilm dispersal (Srivastava & Waters, 2012). In *Vibrio fischeri* it was proposed that LuxQ positively affects biofilm formation through its kinase activity. Specifically, mutant strains of *V.fischeri* that did not express LuxQ, showed a delayed biofilm production (Ray & Visick, 2013). These two types of quorum sensing systems and how they relate to dispersal and mat formation in *Beggiatoa* have yet to be explored. The bone genomes, excluding FB4.032 had high counts of annotations for both types of quorum sensing pathways, as did ABBZPS and GOM8, while b.lep, b.alba, boguay and GUB9 had high counts for LuxQ, which suggests that these two methods of gene regulation may differ in *Beggiatoa* genomes.

In addition to the potential of quorum sensing pathways potentially being utilized for mat formation, certain *Beggiatoa* could possibly rely on other methods to regulate biofilm production. In the two isolates b.lep and b.alba, gene copies for phosphoserine phosphatase RsbU were elevated. The function of this phosphatase has not been studied in *Beggiatoa*. However, in *Bacillus subtilis* the expression of the σ stress factor is positively regulated by RsbU, which was shown to be important for

biofilm production (Delumeau et al., 2004; Nagórska, Hinc, Strauch, & Obuchowski, 2008). Understanding how RsbU and the stress factor σ^b regulate *Beggiatoa* cultures necessitates further research.

The gene annotations from the pangenome suggest that certain *Beggiatoa* genomes may also switch between different metabolic pathways. Genes encoding for kinase barA were also in high abundances in a majority of the genomes (Figure 2.3). This protein is part of a two-component system with UvrY, which in *E.coli* is necessary to switch between gluconeogenesis and glycolysis (Pernestig et al., 2003). Within the *Beggiatoa* genomes, genes encoding for UvrY were found in the two isolates b.alba and b.lep, SM.98, SM.70 and ABBZPS. The function of this system still remains unknown within *Beggiatoa*, but could be studied via competition experiments using freshwater cultures.

The functional potential of these genomes differed as well. Carbon fixation via RuBisCo and the CBB cycle were complete in the isolates, the deep genomes, and only FB4.032 from the bone mats (Figure 2.2). While RuBisCo was present in SM.70, not all genes necessary for a complete CBB cycle were. However, it is possible that genes may be lacking due to a smaller completeness level. In previous literature, autotrophic growth has been shown for *B. leptomitiformis* as well as the marine strain *Beggiatoa* MS-81-6 (Douglas C. Nelson & Jannasch, 1983; Patrinskaya, Grabovich, Muntyan, & Dubinina, 2001). For *Beggiatoa alba*, it has been reported that its growth occurs mixotrophically (Gude, Strohl, & Larkin, 1981). Uptake of CO₂ and carbon fixation was also reported to be present in *Beggiatoa* spp. taken from microbial mats of the Guaymas Basin vent system (Nelson, Wirsen, & Jannasch, 1989). The genomes of boguay and ABBZPS have been extensively analyzed confirming the potential for

autotrophic growth for these filaments as well (MacGregor, Biddle, & Teske, 2013; Mußmann et al., 2007). The lack of carbon fixation genes in the majority of the bone MAGs may be due to the fact that they came from metagenomic datasets, and the fact that their genomes are not complete. However, FB4.032, a bone mat genome showed the potential for carbon fixation, despite being only 60% complete (Table 2.1; Figure 2.2). With the remaining bone mat genomes ranging from 93% to 98% complete, it could be more likely that heterotrophic growth is indeed a characteristic of the remaining *Beggiatoa* genomes found in the bone mat metagenomes.

With the exception of GOM8, all other genomes contained *sqr* and *fcc* genes which allow for the oxidation of sulfide to elemental sulfur. In GOM8, it is likely that these genes may be present in parts of the genome that were not sequenced. The potential for using the reverse dissimilatory sulfate reduction pathway (rDsr) was complete in ABBZPS, boguay and GOM8. In SM.98 only a *sat* gene could be found, while GUB9 was missing *AprAB*. Mußmann et al (2007) suggested that *Beggiatoa* utilizes the rDsr pathway in times of non-overlapping gradients of sulfide and oxygen in their environments. However it may be likely that some *Beggiatoa* lack the genetic machinery necessary to use the rDsr pathway and are limited to surviving in habitats where sulfide and oxygen do overlap. This may be a characteristic of deep-sea dwelling *Beggiatoa* that some, but not all, shallow *Beggiatoa* filaments can perform. For example even though ABBZPS was considered to be a shallow filament in this study, b.lep, b.alba and the bone mat genomes did not show the potential for using two-step method of sulfur oxidation.

Dissimilatory nitrate reduction was complete in seven of the bone mat genomes, the isolates, GUB9, GOM8, ABBZPS, and SM.98. However, the deep

genomes encoded for the membrane-bound nitrate reductase (NarGH) and the shallow ones contained genes only for the periplasmic (NapAB). ABBZPS contained genes for both, as was also previously determined (Mußmann et al., 2007). Therefore, while nitrate respiration is a common characteristic of *Beggiatoa*, the method of using nitrate as an electron donor (via Nar or Nap, or both) may depend on the environment in which filaments are present. The remaining denitrification pathways were only fully complete in pork-bone genomes evident by the presence of *nirS*, *norBC* and *nosZ*. Denitrification potential was also shown in *Beggiatoa* mats from the same marine sediment that ABBZPS was taken from. Additionally, white filaments from the Guaymas Basin were able to use both denitrifying and ammonium producing pathways to reduce nitrate, while their orange neighbors only used the pathway for DNRA (Mußmann et al., 2007; Schutte et al., 2018). It was further suggested that the nitrate reduction pathway utilized by white filament mats might have been dependent on the availability of hydrogen sulfide from hydrothermal flow. Under conditions of low hydrogen sulfide, these mats performed denitrification and under high sulfide levels they performed DNRA. Since denitrification and DNRA were complete in the genomes of the bone *Beggiatoa*, future explorations of nitrate reductive pathways in incubation experiments may be of benefit in order to understand the fate of nitrogen in their habitat and to determine if the regulation of nitrate reduction pathways is an ability they share with their deep-dwelling relatives.

Aside from using sulfide or nitrate as electron donors, it is possible that the genomes from the pork-bone mats are capable of using hydrogen gas as an energy source. Genes encoding for the subunits of Ni-Fe hydrogenase Hyd-1 were present in pork-bone genomes and the two isolates suggesting the ability to oxidize molecular

hydrogen (Figure 2.2). Genes encoding for Ni-Fe hydrogenases were also reported in a *Beggiatoa* genome from an isolated sinkhole in Lake Huron, and the *Beggiatoa* strain 35Flor has been shown to oxidize H₂ under aerobic and anaerobic conditions (Kreutzmann & N. Schulz-Vogt, 2016; Sharrar et al., 2017).

Finally, the potential for using the glyoxylate shunt was present in the two isolates and bone genomes (except for FB4.032). The glyoxylate shunt allows for microorganisms to grow on acetate (Berg, Tymoczko, & Stryer, 2007). It has been shown that under lithotrophic conditions, in *Beggiatoa leptomitiformis*, the glyoxylate shunt is stimulated by a tetrameric malate dehydrogenase. Under organotrophic growth using lactate, the use of the TCA cycle is favored (Eprintsev et al., 2004). Since the deep dwelling genomes showed the potential for carbon fixation, the glyoxylate shunt may not be necessary in order to acquire carbon compounds from their environments. On the other hand, as representatives of shallow-dwelling *Beggiatoa*, the bone genomes may prefer to grow heterotrophically and could utilize the glyoxylate shunt to metabolize acetate.

Overall, putative functions of the genomes studied, show that there may be some distinct lifestyles for *Beggiatoa* depending on their location. Deep-dwelling *Beggiatoa* genomes may use carbon fixation pathways to generate biomass, but most shallow ones may have to rely solely on external carbon sources such as acetate. Other filaments, such as ABBZPS, may be capable of relying on both lifestyles. Likewise, sulfur oxidation and even nitrate respiration have been noted as characteristics of *Beggiatoa* filaments. However, the expected pathways for both of these metabolisms may be different in deep and shallow *Beggiatoa*. Further sequencing attempts of single

filaments, as well as metagenomes of mats from vent systems and shallow habitats may aid in resolving questions regarding how to classify and characterize this group of sulfur oxidizing bacteria.

Within microbial mats, *Beggiatoa* filaments are not isolated organisms. Instead they spend extended periods of time interacting with many other microbial inhabitants. It may be possible that this coexistence of bacteria, has allowed for genetic flow between mat community members. The possibility for genetic exchange in the boguay filament has already been described in detail by (MacGregor et al., 2013) . Sensory and signal transduction genes of this filament were related to Cyanobacteria, suggesting that genetic exchange between different filamentous bacteria may have occurred over evolutionary timescales. Cyanobacteria-related genes were identified in ABBZPS as well (Mußmann et al., 2007). In our study, it appeared that taxonomic imprints on the *Beggiatoa* genome depended on mat type or habitat. Gene annotations from the two filaments from seep systems, GOM8 and GUB9 showed similar taxonomic affiliations with boguay, while the pork bone mat genomes presented a different pattern of taxonomic origins. Even more different were the two genomes SM.98 and SM.70, which came from iron oxidizing mats. These different gene origins may result from differences in the community composition of different mats. Further investigations of *Beggiatoa* filaments may have to take into deeper consideration the community structure of their originating habitat, as it may have significant implications on the evolutionary history of these organisms.

The diversity of the large sulfur oxidizing bacterium *Beggiatoa* has been a topic of study for several years. Despite the difficulties in obtaining isolated cultures,

the advancement in sequencing technologies and bioinformatic algorithms has provided insight into how these organisms may function. Here we present near complete *Beggiatoa* genomes from metagenomic data of mats grown on bones in a natural system. We show that within this specific system we sufficiently represent the genome content of *Beggiatoa*, but when we compare this to other environments, we see that there is much more to be discovered about *Beggiatoa* in general. Function varies widely within the *Beggiatoa* and may be related to environmental pressures. Future searches for these organisms should include a wider range of habitats and environments than have been previously studied (i.e, iron-oxidizing mats or shallow water environments) in order to fully sample the diversity of this genus. Such efforts may present an even more phylogenetically and functionally diverse group of filamentous *Beggiatoa* than previously suggested, and may answer questions regarding the evolution and dispersal of this genus.

REFERENCES

- Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., ... Quince, C. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods*, *11*, 1144. Retrieved from <http://dx.doi.org/10.1038/nmeth.3103>
- Arkin, A. P., Stevens, R. L., Cottingham, R. W., Maslov, S., Henry, C. S., Dehal, P., ... Yoo, S. (2016). The DOE Systems Biology Knowledgebase (KBase). *BioRxiv*, 096354. <https://doi.org/10.1101/096354>
- Beck, C., Knoop, H., & Steuer, R. (2018). Modules of co-occurrence in the cyanobacterial pan-genome reveal functional associations between groups of ortholog genes. *PLoS Genetics*, *14*(3), 1–20. <https://doi.org/10.1371/journal.pgen.1007239>
- Benbow, M. E., Pechal, J. L., Lang, J. M., Erb, R., & Wallace, J. R. (2015). The Potential of High-throughput Metagenomic Sequencing of Aquatic Bacterial Communities to Estimate the Postmortem Submersion Interval. *Journal of Forensic Sciences*, *60*(6), 1500–1510. <https://doi.org/10.1111/1556-4029.12859>
- Bennett, A., & Bogorad, L. (1973). Complementary chromatic adaptation in a filamentous blue-green alga. *The Journal of Cell Biology*, *58*, 419–435. <https://doi.org/10.1083/jcb.58.2.419>
- Berg, J., Tymoczko, J., & Stryer, L. (2007). *Biochemistry*. *Biochemistry*. <https://doi.org/10.1007/978-3-8274-2989-6>
- Bolhuis, H., Cretoiu, M. S., & Stal, L. J. (2014). Molecular ecology of microbial mats. *FEMS Microbiology Ecology*, *90*(2), 335–350. <https://doi.org/10.1111/1574-6941.12408>
- Bowles, M., & Joye, S. (2011). High rates of denitrification and nitrate removal in cold seep sediments. *ISME Journal*, *5*(3), 565–567. <https://doi.org/10.1038/ismej.2010.134>
- Bowles, M. W., Nigro, L. M., Teske, A. P., & Joye, S. B. (2012). Denitrification and environmental factors influencing nitrate removal in Guaymas Basin hydrothermally altered sediments. *Frontiers in Microbiology*, *3*(OCT), 1–11. <https://doi.org/10.3389/fmicb.2012.00377>
- Brady, A. L., Slater, G. F., Omelon, C. R., Southam, G., Druschel, G., Andersen, D. T., ... Lim, D. S. S. (2010). Photosynthetic isotope biosignatures in laminated micro-stromatolitic and non-laminated nodules associated with modern, freshwater microbialites in Pavilion Lake, B.C. *Chemical Geology*, *274*(1–2), 56–67. <https://doi.org/10.1016/j.chemgeo.2010.03.016>
- Buchfink, B., Xie, C., & Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, *12*, 59. Retrieved from

- <http://dx.doi.org/10.1038/nmeth.3176>
- Burne, R. V., & Moore, L. S. (1987). Linked references are available on JSTOR for this article : Microbialites : Organosedimentary Deposits of Benthic Microbial Communities. *PALAIOS*, 2(3), 241–254.
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., ... Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7, 335. Retrieved from <http://dx.doi.org/10.1038/nmeth.f.303>
- Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., ... Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, 108(Supplement_1), 4516–4522. <https://doi.org/10.1073/pnas.1000080107>
- Chan, O. W., Bugler-Lacap, D. C., Biddle, J. F., Lim, D. S., McKay, C. P., & Pointing, S. B. (2014). Phylogenetic diversity of a microbialite reef in a cold alkaline freshwater lake. *Canadian Journal of Microbiology*, 60(6), 391–398. <https://doi.org/10.1139/cjm-2014-0024>
- Clarke, K. R., & Gorley, R. N. (2006). PRIMER v6: User Manual/Tutorial. *PRIMER-E, Plymouth UK*, 192 p. <https://doi.org/10.1111/j.1442-9993.1993.tb00438.x>
- Cooper, D. M. (2003). Regulation and organization of adenylyl cyclases and {cAMP}. *Biochem. J.*, 375(Pt 3), 517–529. <https://doi.org/10.1042/BJ20031061>
- Couradeau, E., Benzerara, K., Moreira, D., Gérard, E., Kaźmierczak, J., Tavera, R., & López-García, P. (2011). Prokaryotic and eukaryotic community structure in field and cultured microbialites from the alkaline Lake Alchichica (Mexico). *PLoS ONE*, 6(12). <https://doi.org/10.1371/journal.pone.0028767>
- Darling, A. E., Jospin, G., Lowe, E., Matsen, F. A., Bik, H. M., & Eisen, J. A. (2014). PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ*, 2, e243. <https://doi.org/10.7717/peerj.243>
- Delmont, T. O., & Eren, A. M. (2018). Linking pangenomes and metagenomes: the *Prochlorococcus* metapangenome. *PeerJ*, 6, e4320. <https://doi.org/10.7717/peerj.4320>
- Deming, J. W., Reysenbach, A. L., Macko, S. A., & Smith, C. R. (1997). Evidence for the microbial basis of a chemoautotrophic invertebrate community at a whale fall on the deep seafloor: Bone-colonizing bacteria and invertebrate endosymbionts. *Microscopy Research and Technique*, 37(2), 162–170. [https://doi.org/10.1002/\(SICI\)1097-0029\(19970415\)37:2<162::AID-JEMT4>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1097-0029(19970415)37:2<162::AID-JEMT4>3.0.CO;2-Q)
- Donati, C., Hiller, N. L., Tettelin, H., Muzzi, A., Croucher, N. J., Angiuoli, S. V., ... Masignani, V. (2010). Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. *Genome Biology*, 11(10), R107. <https://doi.org/10.1186/gb-2010-11-10-r107>
- Dupraz, C., Reid, R. P., Braissant, O., Decho, A. W., Norman, R. S., & Visscher, P. T. (2009). Processes of carbonate precipitation in modern microbial mats. *Earth-*

- Science Reviews*, 96(3), 141–162. <https://doi.org/10.1016/j.earscrev.2008.10.005>
- Dworkin, M. (2012). Sergei Winogradsky: A founder of modern microbiology and the first microbial ecologist. *FEMS Microbiology Reviews*, 36(2), 364–379. <https://doi.org/10.1111/j.1574-6976.2011.00299.x>
- Eprintsev, A. T., Falaleeva, M. I., Grabovich, M. Y., Parfenova, N. V., Kashirskaya, N. N., & Dubinina, G. A. (2004). The role of malate dehydrogenase isoforms in the regulation of anabolic and catabolic processes in the colorless sulfur bacterium *Beggiatoa leptomitiformis* D-402. *Microbiology*, 73(4), 367–371. <https://doi.org/10.1023/B:MICI.0000036977.79822.7a>
- Fierer, N., Nemergut, D., Knight, R., & Craine, J. M. (2010). Changes through time: Integrating microorganisms into the study of succession. *Research in Microbiology*, 161(8), 635–642. <https://doi.org/10.1016/j.resmic.2010.06.002>
- Flood, B. E., Bailey, J. V., & Biddle, J. F. (2014). Horizontal gene transfer and the rock record: Comparative genomics of phylogenetically distant bacteria that induce wrinkle structure formation in modern sediments. *Geobiology*, 12(2), 119–132. <https://doi.org/10.1111/gbi.12072>
- Fowler, A. C., & Winstanley, H. F. (2018). Microbial dormancy and boom-and-bust population dynamics under starvation stress. *Theoretical Population Biology*, 120, 114–120. <https://doi.org/10.1016/j.tpb.2018.02.001>
- Goffredi, S. K., & Orphan, V. J. (2010). Bacterial community shifts in taxa and diversity in response to localized organic loading in the deep sea. *Environmental Microbiology*, 12(2), 344–363. <https://doi.org/10.1111/j.1462-2920.2009.02072.x>
- Grabovich, M. Y., Dubinina, G. A., Lebedeva, V. Y., & Churikova, V. V. (1998). Mixotrophic and lithoheterotrophic growth of the freshwater filamentous sulfur bacterium *Beggiatoa leptomitiformis* D-402. *Mikrobiologiya*, 67, 383–388.
- Graham, E. D., Heidelberg, J. F., & Tully, B. J. (2017). BinSanity: unsupervised clustering of environmental microbial assemblies using coverage and affinity propagation. *PeerJ*, 5, e3035. <https://doi.org/10.7717/peerj.3035>
- Grote, J., Jost, G., Labrenz, M., Herndl, G. J., & Jürgens, K. (2008). Epsilonproteobacteria represent the major portion of chemoautotrophic bacteria in sulfidic waters of pelagic redoxclines of the Baltic and Black Seas. *Applied and Environmental Microbiology*, 74(24), 7546–7551. <https://doi.org/10.1128/AEM.01186-08>
- Gude, H., Strohl, W. R., & Larkin, J. M. (1981). Mixotrophic and heterotrophic growth of *Beggiatoa alba* in continuous culture. *Archives of Microbiology*, 129, 357–360.
- Hilario, A., Cunha, M. R., Génio, L., Marçal, A. R., Ravara, A., Rodrigues, C. F., & Wiklund, H. (2015). First clues on the ecology of whale falls in the deep Atlantic Ocean: Results from an experiment using cow carcasses. *Marine Ecology*, 36(S1), 82–90. <https://doi.org/10.1111/maec.12246>
- Hladysz, S., Cook, R. A., Petrie, R., & Nielsen, D. L. (2010). Influence of substratum on the variability of benthic biofilm stable isotope signatures: Implications for energy flow to a primary consumer. *Hydrobiologia*, 664(1), 135–146.

- <https://doi.org/10.1007/s10750-010-0593-0>
- Iseki, M., Matsunaga, S., Murakami, A., Ohno, K., Shiga, K., Yoshida, K., ... Watanabe, M. (2002). A blue-light-activated adenylyl cyclase mediates photoavoidance in *Euglena gracilis*. *Nature*, *415*(6875), 1047–1051. <https://doi.org/10.1038/4151047a>
- Jannasch, H. W., Nelson, C. D., & Wirsén, C. O. (1989). Massive Natural Occurrence of Unusually Large Bacteria (*Beggiatoa* sp.) at a Hydrothermal Deep-Sea Vent Site. *Nature*, *340*(6231), 301–303. <https://doi.org/10.1038/340301a0>
- Jones, D. S., Flood, B. E., & Bailey, J. V. (2015). Metatranscriptomic analysis of diminutive Thiomargarita-like bacteria (“*Candidatus Thiopilula*” spp.) from abyssal cold seeps of the Barbados accretionary prism. *Applied and Environmental Microbiology*, *81*(9), 3142–3156. <https://doi.org/10.1128/AEM.00039-15>
- Kaas, R. S., Friis, C., Ussery, D. W., & Aarestrup, F. M. (2012). Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse *Escherichia coli* genomes. *BMC Genomics*, *13*(1), 1. <https://doi.org/10.1186/1471-2164-13-577>
- Kalenitchenko, D., Le Bris, N., Peru, E., & Galand, P. E. (2018). Ultrarare marine microbes contribute to key sulphur-related ecosystem functions. *Molecular Ecology*. <https://doi.org/10.1111/mec.14513>
- Kalenitchenko, D., Dupraz, M., Le Bris, N., Petetin, C., Rose, C., West, N. J., & Galand, P. E. (2016). Ecological succession leads to chemosynthesis in mats colonizing wood in sea water. *ISME Journal*, *10*(9), 2246–2258. <https://doi.org/10.1038/ismej.2016.12>
- 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. <https://doi.org/10.1016/j.jmb.2015.11.006>
- Kang, D. D., Froula, J., Egan, R., & Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ*, *3*, e1165. <https://doi.org/10.7717/peerj.1165>
- Kim, T., Folcher, M., Baba, M. D. El, & Fussenegger, M. (2015). A synthetic erectile optogenetic stimulator enabling blue-light-inducible penile erection. *Angewandte Chemie - International Edition*, *54*(20), 5933–5938. <https://doi.org/10.1002/anie.201412204>
- Konishi, Y., Prince, J., & Knott, B. (2001). The fauna of thrombolitic microbialites, Lake Clifton, Western Australia. *Hydrobiologia*, *457*(1985), 39–47. <https://doi.org/10.1023/A:1012229412462>
- Kreutzmann, A.-C., & N. Schulz-Vogt, H. N. (2016). Oxidation of Molecular Hydrogen by a Chemolithoautotrophic ABSTRACT. *Applied and Environmental Microbiology*, *82*(8), 2527–2536. <https://doi.org/10.1128/AEM.03818-15>
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and*

- Evolution*, 33(7), 1870–1874. <https://doi.org/10.1093/molbev/msw054>
- Laczny, C. C., Sternal, T., Plugaru, V., Gawron, P., Atashpendar, A., Margossian, H. H., ... Wilmes, P. (2015). VizBin - an application for reference-independent visualization and human-augmented binning of metagenomic data. *Microbiome*. London. <https://doi.org/10.1186/s40168-014-0066-1>
- Lang, J., Erb, R., Pechal, J., Wallace, J., McEwan, R., & Benbow, M. (2016). Microbial Biofilm Community Variation in Flowing Habitats: Potential Utility as Bioindicators of Postmortem Submersion Intervals. *Microorganisms*, 4(1), 1. <https://doi.org/10.3390/microorganisms4010001>
- MacGregor, B. J., Biddle, J. F., Harbort, C., Matthyse, A. G., & Teske, A. (2013). Sulfide oxidation, nitrate respiration, carbon acquisition, and electron transport pathways suggested by the draft genome of a single orange Guaymas Basin Beggiatoa (Cand. Maribeggiatoa) sp. filament. *Marine Genomics*, 11, 53–65. <https://doi.org/10.1016/j.margen.2013.08.001>
- MacGregor, B. J., Biddle, J. F., & Teske, A. (2013). Mobile elements in a single-filament orange guaymas basin beggiatoa (“Candidatus maribeggiatoa”) sp. draft genome: Evidence for genetic exchange with cyanobacteria. *Applied and Environmental Microbiology*, 79(13), 3974–3985. <https://doi.org/10.1128/AEM.03821-12>
- Marshall, I. P. G., Blainey, P. C., Spormann, A. M., & Quake, S. R. (2012). A single-cell genome for Thiovulum sp. *Applied and Environmental Microbiology*, 78(24), 8555–8563. <https://doi.org/10.1128/AEM.02314-12>
- McKay, L. J., MacGregor, B. J., Biddle, J. F., Albert, D. B., Mendlovitz, H. P., Hoer, D. R., ... Teske, A. P. (2012). Spatial heterogeneity and underlying geochemistry of phylogenetically diverse orange and white Beggiatoa mats in Guaymas Basin hydrothermal sediments. *Deep-Sea Research Part I: Oceanographic Research Papers*, 67, 21–31. <https://doi.org/10.1016/j.dsr.2012.04.011>
- McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, 8(4). <https://doi.org/10.1371/journal.pone.0061217>
- Merz, M. U. E. (1992). The biology of carbonate precipitation by cyanobacteria. *Facies*, 26(1), 81–101. <https://doi.org/10.1007/BF02539795>
- Mezzino, M. J., Strohl, W. R., & Larkin, J. M. (1984). Characterization of Beggiatoa alba. *Archives of Microbiology*, 137(2), 139–144. <https://doi.org/10.1007/BF00414455>
- Mills, H. J., Mills, H. J., Martinez, R. J., Martinez, R. J., Story, S., Story, S., ... Sobecky, P. a. (2004). Identification of Members of the Metabolically Active Microbial Populations Associated with. *Society*, 70(9), 5447–5458. <https://doi.org/10.1128/AEM.70.9.5447>
- Moldovan, M. A., & Gelfand, M. S. (2018). Pangenomic definition of prokaryotic species and the phylogenetic structure of Prochlorococcus spp. *Frontiers in Microbiology*, 9(MAR), 1–11. <https://doi.org/10.3389/fmicb.2018.00428>
- Mosquera-Rendón, J., Rada-Bravo, A. M., Cárdenas-Brito, S., Corredor, M.,

- Restrepo-Pineda, E., & Benítez-Páez, A. (2016). Pangenome-wide and molecular evolution analyses of the *Pseudomonas aeruginosa* species. *BMC Genomics*, *17*(1), 1–14. <https://doi.org/10.1186/s12864-016-2364-4>
- Mußmann, M., Hu, F. Z., Richter, M., De Beer, D., Preisler, A., Jørgensen, B. B., ... Ehrlich, G. D. (2007). Insights into the genome of large sulfur bacteria revealed by analysis of single filaments. *PLoS Biology*, *5*(9), 1923–1937. <https://doi.org/10.1371/journal.pbio.0050230>
- Nagórska, K., Hinc, K., Strauch, M. A., & Obuchowski, M. (2008). Influence of the sigB stress factor and yxaB, the gene for a putative exopolysaccharide synthase under sigB control, on biofilm formation. *Journal of Bacteriology*, *190*(10), 3546–3556. <https://doi.org/10.1128/JB.01665-07>
- Nelson, D. C., & Castenholz, R. W. (1982). Light responses of *Beggiatoa*. *Archives of Microbiology*, *131*(2), 146–155. <https://doi.org/10.1007/BF01053997>
- Nelson, D. C., & Jannasch, H. W. (1983). Chemoautotrophic growth of a marine *Beggiatoa* in sulfide-gradient cultures. *Archives of Microbiology*, *136*(4), 262–269. <https://doi.org/10.1007/BF00425214>
- Nelson, D. C., Wirsén, C. O., & Jannasch, H. W. (1989). Characterization of Large, Autotrophic *Beggiatoa* spp. Abundant at Hydrothermal Vents of the Guaymas Basin. *Applied and Environmental Microbiology*, *55*(11), 2909–2917.
- Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., ... Pevzner, P. A. (2013). Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. In M. Deng, R. Jiang, F. Sun, & X. Zhang (Eds.), *Research in Computational Molecular Biology* (pp. 158–170). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Omelon, C. R., Brady, A. L., Slater, G. F., Laval, B., Lim, D. S. S., & Southam, G. (2013). Microstructure variability in freshwater microbialites, Pavilion Lake, Canada. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *392*, 62–70. <https://doi.org/10.1016/j.palaeo.2013.08.017>
- Pace, N. R. (1997). A molecular view of microbial diversity and the biosphere. [Review] [52 refs]. *Science*, *276*(5313), 734–740.
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., ... Parkhill, J. (2015). Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, *31*(22), 3691–3693. <https://doi.org/10.1093/bioinformatics/btv421>
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from. *Cold Spring Harbor Laboratory Press Method*, *1*, 1–31. <https://doi.org/10.1101/gr.186072.114>
- Pati, A., Gronow, S., Lapidus, A., Copeland, A., del Rio, T. G., Nolan, M., ... Kyrpides, N. C. (2010). Complete genome sequence of *Arcobacter nitrofigilis* type strain (CI T). *Standards in Genomic Sciences*, *2*(3), 300–308. <https://doi.org/10.4056/sigs.912121>
- Patrinskaya, V. Y., Grabovich, M. Y., Muntyan, M. S., & Dubinina, G. A. (2001).

- Lithoautotrophic growth of the freshwater colorless sulfur bacterium *Beggiatoa* “leptomitiformis” D-402. *Microbiology*, 70(2), 145–150.
<https://doi.org/10.1023/A:1010469211447>
- Pernestig, A. K., Georgellis, D., Romeo, T., Suzuki, K., Tomenius, H., Normark, S., & Melefors, Ö. (2003). The *Escherichia coli* BarA-UvrY two-component system is needed for efficient switching between glycolytic and gluconeogenic carbon sources. *Journal of Bacteriology*, 185(3), 843–853.
<https://doi.org/10.1128/JB.185.3.843-853.2003>
- Podell, S., & Gaasterland, T. (2007). DarkHorse: A method for genome-wide prediction of horizontal gene transfer. *Genome Biology*, 8(2).
<https://doi.org/10.1186/gb-2007-8-2-r16>
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1), 590–596.
<https://doi.org/10.1093/nar/gks1219>
- Ray, V. A., & Visick, K. L. (2013). LuxU Connects Quorum Sensing to Biofilm Formation in *Vibrio fischeri*, 86(4), 954–970.
<https://doi.org/10.1111/mmi.12035>
- Reno, M. L., Held, N. L., Fields, C. J., Burke, P. V., & Whitaker, R. J. (2009). Biogeography of the *Sulfolobus islandicus* pan-genome. *Proceedings of the National Academy of Sciences*, 106(44), 18873–18873.
<https://doi.org/10.1073/pnas.0911400106>
- Roalkvam, I., Drønen, K., Stokke, R., Daae, F. L., Dahle, H., & Steen, I. H. (2015). Physiological and genomic characterization of *arcobacter anaerophilus* IR-1 reveals new metabolic features in epsilonproteobacteria. *Frontiers in Microbiology*, 6(SEP), 1–12. <https://doi.org/10.3389/fmicb.2015.00987>
- Russell, J. A., Brady, A. L., Cardman, Z., Slater, G. F., Lim, D. S. S., & Biddle, J. F. (2014). Prokaryote populations of extant microbialites along a depth gradient in Pavilion Lake, British Columbia, Canada. *Geobiology*, 12(3), 250–264.
<https://doi.org/10.1111/gbi.12082>
- Ryu, M. H., Moskvina, O. V., Siltberg-Liberles, J., & Gomelsky, M. (2010). Natural and engineered photoactivated nucleotidyl cyclases for optogenetic applications. *Journal of Biological Chemistry*, 285(53), 41501–41508.
<https://doi.org/10.1074/jbc.M110.177600>
- Salman, V., Amann, R., Girnath, A. C., Polerecky, L., Bailey, J. V., Høgslund, S., ... Schulz-Vogt, H. N. (2011). A single-cell sequencing approach to the classification of large, vacuolated sulfur bacteria. *Systematic and Applied Microbiology*, 34(4), 243–259. <https://doi.org/10.1016/j.syapm.2011.02.001>
- Salman, V., Bailey, J. V., & Teske, A. (2013). Phylogenetic and morphologic complexity of giant sulphur bacteria. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 104(2), 169–186.
<https://doi.org/10.1007/s10482-013-9952-y>
- Schutte, C. A., Teske, A., MacGregor, B. J., Salman Carvalho, V., Lavik, Gaute, ...

- De Beer, D. (2018). Filamentous giant Beggiatoaceae from Guaymas Basin are capable of both denitrification and dissimilatory nitrate reduction to ammonium (DNRA). *Applied and Environmental Microbiology*.
<https://doi.org/10.1128/AEM.03397-16>
- Schwedt, A. (2011). Physiology of a marine Beggiatoa strain and the accompanying organism *Pseudovibrio* sp. – a facultatively oligotrophic bacterium. *PhD Thesis*, (September).
- Seckbach, J., & Aharon, O. (2010). *Microbial Mats Modern and Ancient Microorganisms in Stratified Systems*.
- Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, 30(14), 2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>
- Sharrar, A. M., Flood, B. E., Bailey, J. V., Jones, D. S., Biddanda, B. A., Ruberg, S. A., ... Dick, G. J. (2017). Novel large sulfur bacteria in the metagenomes of groundwater-fed chemosynthetic microbial mats in the Lake Huron basin. *Frontiers in Microbiology*, 8(MAY), 1–15.
<https://doi.org/10.3389/fmicb.2017.00791>
- Smith, C. R., Glover, A. G., Treude, T., Higgs, N. D., & Amon, D. J. (2015). Whale-Fall Ecosystems: Recent Insights into Ecology, Paleoecology, and Evolution. *Annual Review of Marine Science*, 7(1), 571–596.
<https://doi.org/10.1146/annurev-marine-010213-135144>
- Srivastava, D., & Waters, C. M. (2012). A tangled web: Regulatory connections between quorum sensing and cyclic Di-GMP. *Journal of Bacteriology*, 194(17), 4485–4493. <https://doi.org/10.1128/JB.00379-12>
- Stierl, M., Stumpf, P., Udvari, D., Gueta, R., Hagedorn, R., Losi, A., ... Hegemann, P. (2011). Light modulation of cellular cAMP by a small bacterial photoactivated adenylyl cyclase, bPAC, of the soil bacterium *Beggiatoa*. *Journal of Biological Chemistry*, 286(2), 1181–1188. <https://doi.org/10.1074/jbc.M110.185496>
- Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D. M., Breiner, H. W., & Richards, T. A. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular Ecology*, 19(SUPPL. 1), 21–31. <https://doi.org/10.1111/j.1365-294X.2009.04480.x>
- Strohl, W. R., & Larkin, J. M. (1978). Enumeration, isolation, and characterization of *Beggiatoa* from fresh water sediments. *Appl. Environ. Microbiol.*, 36(5), 755–770.
- Sweerts, J. P. R. A., De Beer, D., Nielsen, L. P., Verdouw, H., Van Den Heuvel, J. C., Cohen, Y., & Cappenberg, T. E. (1990). Denitrification by sulphur oxidizing *Beggiatoa* spp. mats on freshwater sediments. *Nature*.
<https://doi.org/10.1038/344762a0>
- Tapley, D. W., Buettner, G. R., & Shick, J. M. (1999). Free radicals and chemiluminescence as products of the spontaneous oxidation of sulfide in seawater, and their biological implications. *Biological Bulletin*, 196(1), 52–56.
<https://doi.org/10.2307/1543166>

- Tarhan, L. G., Planavsky, N. J., Laumer, C. E., Stolz, J. F., & Reid, R. P. (2013). Microbial mat controls on infaunal abundance and diversity in modern marine microbialites. *Geobiology*, *11*(5), 485–497. <https://doi.org/10.1111/gbi.12049>
- Tettelin, H., Masignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., ... Fraser, C. M. (2005). Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial “pan-genome.” *Proceedings of the National Academy of Sciences*, *102*(39), 13950–13955. <https://doi.org/10.1073/pnas.0506758102>
- Tringe, S. G., von Mering, C., Kobayashi, A. ;, Salamov, A. A., Chen, K. ;, Chang, H. W., ... Rubin, E. M. (2005). “Comparative metagenomics of microbial communities.” *Science* *308*.5721 (2005): 554-557. *Science*, *308*(5721), 4. <https://doi.org/10.1126/science.1107851>
- Tully, B. J., Wheat, C. G., Glazer, B. T., & Huber, J. A. (2018). A dynamic microbial community with high functional redundancy inhabits the cold, oxic seafloor aquifer. *ISME Journal*, *12*(1), 1–16. <https://doi.org/10.1038/ismej.2017.187>
- Vernikos, G., Medini, D., Riley, D. R., & Tettelin, H. (2015). Ten years of pan-genome analyses. *Current Opinion in Microbiology*, *23*, 148–154. <https://doi.org/10.1016/j.mib.2014.11.016>
- Vetriani, C., Voordeckers, J. W., Crespo-Medina, M., O’Brien, C. E., Giovannelli, D., & Lutz, R. A. (2014). Deep-sea hydrothermal vent Epsilonproteobacteria encode a conserved and widespread nitrate reduction pathway (Nap). *ISME Journal*, *8*(7), 1510–1521. <https://doi.org/10.1038/ismej.2013.246>
- Warden, J. G., Casaburi, G., Omelon, C. R., Bennett, P. C., Breecker, D. O., & Foster, J. S. (2016). Characterization of microbial mat microbiomes in the modern thrombolite ecosystem of lake Clifton, western Australia using shotgun metagenomics. *Frontiers in Microbiology*, *7*(JUL), 1–14. <https://doi.org/10.3389/fmicb.2016.01064>
- Wasmund, K., Mußmann, M., & Loy, A. (2017). The life sulfuric: microbial ecology of sulfur cycling in marine sediments. *Environmental Microbiology Reports*, *9*(4), 323–344. <https://doi.org/10.1111/1758-2229.12538>
- Weisburg, W. G., Barns, S. M., Pelletier, D. A., & Lane, D. J. (1991). 16S Ribosomal DNA Amplification for Phylogenetic Study. *Journal of Bacteriology*, *173*(2), 697–703. <https://doi.org/n.a>
- White, R. A., Power, I. M., Dipple, G. M., Southam, G., & Suttle, C. A. (2015). Metagenomic analysis reveals that modern microbialites and polar microbial mats have similar taxonomic and functional potential. *Frontiers in Microbiology*, *6*(SEP). <https://doi.org/10.3389/fmicb.2015.00966>
- Wirsen, C. O., & Jannasch, H. W. (1978). Physiological and morphological observations on *Thiovulum* sp. *Journal of Bacteriology*, *136*(2), 765–774. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=218603&tool=pmcentrez&rendertype=abstract>
- Wu, Y.-W., Tang, Y.-H., Tringe, S. G., Simmons, B. A., & Singer, S. W. (2014).

- MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome*, 2(1), 26. <https://doi.org/10.1186/2049-2618-2-26>
- Yamamoto, M., & Takai, K. (2011). Sulfur metabolisms in epsilon- and gamma-proteobacteria in deep-sea hydrothermal fields. *Frontiers in Microbiology*, 2(SEP), 1–8. <https://doi.org/10.3389/fmicb.2011.00192>
- Zhang, Y., & Sievert, S. M. (2014). Pan-genome analyses identify lineage- and niche-specific markers of evolution and adaptation in Epsilonproteobacteria. *Frontiers in Microbiology*, 5(MAR), 1–13. <https://doi.org/10.3389/fmicb.2014.00110>

Appendix

PROKARYOTIC AND EUKARYOTIC COMMUNITY ASSEMBLAGES WITHIN A MICROBIALITE AND NODULES IN PAVILION LAKE, BRITISH COLUMBIA, CANADA

Abstract

An oligotrophic and alkaline lake, Pavilion Lake in British Columbia hosts a wide range of morphologically distinct microbialites that often have small nodular mesostructures. It remains unknown how light levels may affect the growth and distribution patterns of morphologically distinct microbialites within the lake basin. As such, we investigated the active growth surfaces and microbial inhabitants of the microbialites. Pulse amplitude modulation (PAM) fluorometry was used to estimate photosynthetic activity associated with different regions of the surface of the mounds and nodules. Chlorophyll fluorescence was present along whole microbialite mounds, and microelectrode profiling showed that the nodular types and the background they were found on produced elevated levels of oxygen and pH. These findings suggest that growth is distributed across both the nodules and the underlying biofilm. 16S rRNA and 18S rRNA gene amplicon sequencing from direct samples of nodules and microbialite surfaces provided a phylogenetic fingerprint of the prokaryotic and eukaryotic microbial inhabitants. Variations in the prokaryotic communities of the different morphologies appear at the phylum level, with the dominating groups being Cyanobacteria and Alphaproteobacteria. *Leptolyngbya* had a higher relative abundance in green nodules than in pink ones. Eukaryotic communities also showed

variations with morphology. Relative abundance of green algae dropped in the communities found on nodules, while families of nematodes and dinoflagellates increased. In summary, active growth on microbialite mounds in Pavilion Lake is not limited to nodules, as the background surfaces are also associated with growth. Furthermore, the impact of eukaryotic organisms on microbialite structures and their association with these alkaline surfaces is yet to be determined.

Introduction

The trapping and precipitation of sedimentary grains by prokaryotic organisms can lead to the formation of large structures known as microbialites, a term that engulfs both structured (stromatolite) and disordered (thrombolite) mats (Burne & Moore, 1987). Carbonate precipitation within these mats is achieved by balancing metabolic activities of photosynthetic and heterotrophic microorganisms (Dupraz et al., 2009). These biological formations can show a wide range of different morphological types and features, yet little is known about their mode of overall growth, and how growth can be tied to morphological features. The growth mechanisms of these structures involve an initial accumulation of polysaccharides known as extracellular polymeric substances (EPS) that cyanobacteria use to adhere to surfaces, trap trace metals and for general protection. Once enough EPS is introduced into the system to create a sheath over the mat, the cyanobacteria actively start pumping bicarbonate into their cells. Bicarbonate ions are then converted into carbon dioxide and hydroxyl ions. The produced carbon dioxide goes towards photosynthesis, and the hydroxyl ions diffuse back to the sheath where they either become neutralized by H^+ ions, or react with other bicarbonate ions to form water and carbonate. The

newly formed carbonate ions attract calcium ions, which precipitate out in the form of calcium carbonate in the sheath (Merz, 1992).

Microbialite mats are distributed globally in a wide range of aquatic environments (Seckbach & Aharon, 2010). One such environment includes Pavilion Lake. An oligotrophic and alkaline lake, Pavilion Lake in British Columbia (location coordinates: 50°51'52.77"N, 121°44'15.23"W) hosts a wide range of morphologically distinct microbialite mounds (for map of Pavilion Lake, see (Russell et al., 2014). While initial information on the microbiology of Pavilion Lake microbialites is available, it is still unknown how light levels in these environments may affect the growth and distribution patterns of these photosynthetically driven mats within lake basins, and what communities make up these structures in areas other than those previously sampled. Distinct features of microbialite beds in Pavilion Lake are pink or green nodules that develop directly on the surface of microbialite mounds. Research has shown that the photosynthetic activity and color differentiations in these nodules may be due to varying abundances of cyanobacterial types found in these two forms of mesostructures (Brady et al., 2010; Russell et al., 2014). The microbialites found in Pavilion Lake are very diverse in morphology and distribution. This study area is unique, as it shows that microbialites can develop in non- extreme areas that support fish, plants and other living organisms.

Several studies have been conducted to understand the microbial diversity in microbialites and how microbe-mineral interactions in these systems may be driving the observed morphology. However, these dynamics still remain undiscovered. Furthermore, even less is known about the eukaryotic microorganisms that are found in microbialite systems and how their presence may have an influence on microbialite

growth. Molecular tools, along with measurements of fluorometry can shed light on the associated microbial communities and aid in gaining a better understanding of the driving forces behind the distribution and diversity of freshwater microbialites. Here, we used pulse amplitude modulation (PAM) fluorometry to estimate photosynthetic activity associated with different regions of the surface of microbialite mounds and nodules found in Pavilion Lake. The use of 16S rRNA and 18S rRNA gene amplicon sequencing from direct samples of nodules and microbialite surfaces was applied on natural samples to obtain a phylogenetic fingerprint of the microbial inhabitants. By applying this approach, our goals were to determine i) if the presence of active microorganisms was limited to specific areas of microbialites and ii) the organisms associated with growing regions of microbialites were representative of prokaryotic groups, eukaryotic groups, or both.

Methods

PAM measurements and imaging

A 20 cm (8") long microbialite mound at 28 meters depth from Three Poles(50°51'59.23"N, 121°44'8.33"W) was assessed for photosynthetic activity using shore-based pulse amplitude modulation (PAM) fluorometry. In order to determine the source of the photosynthetic signal on the surface of the mound, images under red illumination were obtained, while near infrared images were used to better depict the structure of the mound. Maximum quantum yields of photosystem II (PSII) was measured to find any localized areas of photosynthetic activity. Similar imaging techniques were used to assess photosynthetic activity in nodular structures and the underlying biofilm from microbialites at 18 meters depth, South Basin.

Cautions with PAM measurements

PAM measurements were obtained by determining two variables; the fluorescence yield of the target ambient irradiance (F) and maximum fluorescence yield under ambient irradiance, measured at the end of a 0.6 s saturating light pulse (F_m'). The yield of PSII under ambient conditions, the effective yield of PSII (Y_{II}) was estimated as $(F_m' - F):F_m'$. The relative photosynthetic electron transport rate (rETR) through photosystem (PSII) was estimated as $rETR = Y_{II} \cdot E$, where E is the available irradiance (Hofstraat et al. 1994). The estimation for the maximum yield of PSII (ϕ) is obtained by measurements that are made under fully dark-adapted conditions. When cells are exposed to light, the maximum yield begins to decline, or be “quenched”, due to increasing saturation of the photochemical mechanisms by light, and by non-photochemical protective mechanisms that the cells activate to protect them from over-oxidation of the intracellular medium. The overall efficiency of light utilization (Y_{II}) falls; this is normal and results in the characteristic saturation, or even inhibition, of photosynthetic rate at high irradiance. Y_{II} often remains constant at increasing but still low irradiance, and this indicates that the cells are able to utilize these low irradiances at optimal efficiency and rETR is primarily driven by irradiance.

Pulse amplitude modulation measures rETR, which is indicative of photosynthetic activity and not photosynthetic rate, as the latter is represented by accumulation of carbon. Electrons that are transported through PSII leads to intermediate reducing agents that may or may not be used for carbon fixation in Cyanobacteria.

The devices used for PAM measurements are optical devices, which work best for surface samples, rather than detecting activity found deeper in three-dimensional

structures. Therefore the measurements obtained from this method may be biased towards the biomass distribution found only on the surface of the microbialites and nodules. Furthermore, PAM imaging was performed with blue measuring lights, which is absorbed best by chlorophyll and carotenoids that are found in light-harvesting complexes. Light capturing pigments, such as phycobillins that are used by cyanobacteria do not absorb blue light well.

Microelectrode profiling

Nodules and background biofilm from 26 meters depth at Three Poles were sampled for oxygen and pH microelectrode profiling as additional assessments for photosynthetic activity. These profiles were used to understand the activity levels associated with the internal layers of the nodules and surface biofilm.

Measurements were done at shore. The nodular and background samples were equilibrated for 30 minutes in lake water at 15 °C and at an irradiance that was representative of the depth they were collected from ($50 \mu\text{mol m}^{-2} \text{s}^{-1}$). A micromanipulator and sensors with diameter tips from 25 to 50 μm diameter were used for profiling from outside of the boundary layer. Measurements were taken at 100 μm intervals until either the hard microbialite surface was reached for the background biofilm, or to a depth of 2-4 mm into the nodules, depending on their estimated size. The exact position of the sample surface was difficult to observe. Therefore, it was deduced based on the profiles that were obtained. Furthermore, as the electrode was lowered, it may have been compressed by the biofilm, before penetration of the tip occurred. Finally, due to the small scale of these measurements (holes with 0.05 mm diameter), there may be some constraints in describing whole nodule samples, which have diameters of roughly 10mm.

DNA extraction of nodular types from South Basin East (SBE) and the microbialite mound from South Basin West (SBW)

Nodules were removed from microbialites in South Basin East (50°50'48.43"N, 121°42'39.65"W) from a depth of 18 meters (58ft) depth. Duplicate samples were taken for each nodule type, as well as from the underlying portion of the microbialite on which the nodules were found. Microbialite samples were taken from a whole structure located in South Basin West (50°50'52.73"N, 121°42'41.38"W) at a depth of 26 meters (88ft). Specifically the structure was sampled in three separate zones, the top, middle and bottom. Extraction of genetic material from the SBE nodular and background biofilm samples were performed as written in (Russell et al., 2014). For each height of the microbialite, genetic material was isolated from an average of 0.17g of sample (ranging from 0.069 g to 0.358 g). Extraction of DNA was performed using the MoBio PowerSoil DNA Isolation kit (MoBio Labs, Carlsbad, CA, USA) and following the manufacturer's protocol. The lysis step was carried out using a homogenizer and subjecting each sample to two intervals of homogenization for 20 seconds with a one-minute rest period between intervals. A control was used during the extraction process. Average DNA yield was 21.4 ng/uL (ranging from 16.6 ng/uL to 26.6 ng/uL). For PCR amplification, library prep and sequencing on the Illumina MiSeq platform, purified DNA was sent to Mr. DNA(www.mrdnalab.com, Shallowater, TX, USA). The 27F bacterial primer sequence was used for the amplification of the 16S rRNA gene of the prokaryotic communities (Weisburg, Barns, Pelletier, & Lane, 1991), while the V4 region of the 18S rRNA gene was amplified for the examination of the eukaryotic communities (Stoeck et al., 2010).

The raw sequences were analyzed through the QIIME pipeline and OTU tables were generated at 97% identity and SILVA database was used for taxonomic assignments (Caporaso et al., 2010; Quast et al., 2013). After demultiplexing and quality filtering the reads, samples were rarefied to avoid sampling bias. Singleton, chloroplast and unassigned sequences were removed from the 16S amplicon data. The samples were rarefied to 29037 sequences and 33481 sequences, which was the smallest sequence count per sample from the prokaryotic microbialite mound and nodule samples respectively. Singletons, unassigned and sequences assigned to taxa without standing nomenclature were removed from the 18S rRNA gene amplicon data and then rarefied to 46802 and 30249 sequences for the eukaryotic microbialite mound and nodule samples respectively.

Statistical Analyses

Chao and Shannon richness indices were calculated in the QIIME pipeline (Caporaso et al., 2010). The 'R' package 'phyloseq' was used to plot the richness indices, and calculate the multidimensional scaling (MDS) plots (McMurdie & Holmes, 2013). Relative percent abundances were plotted in R at different taxonomic levels. PRIMER-E 6th Edition was used to calculate Bray-Curtis similarity matrices and cluster analyses (Clarke & Gorley, 2006). Finally, in order to identify OTUS that may be driving differences observed in pink and green nodules the SIMPER (similarity percentage) program in PRIMER was used to determine major OTUs contributing to the dissimilarity between the two nodule types (Clarke & Gorley, 2006).

Results

Active growth measurements on microbialites

Images of the structure and nodules were taken under near-infrared light (Figures 1A, 2A). Shore based imaging from the deep mound showed that the microbialite at Three Poles was covered in chlorophyll. The only exception was the portion of the mound that was chipped off on the top left part during sampling (Figure 1B). The PAM fluorometer measurements, suggested that the chlorophyll found on the mound is active across the whole microbialite (Figure 1C).

In a similar fashion, chlorophyll fluorescence was apparent on the pink nodules, as well as the background biofilm samples that were taken from South Basin (Figure 2B). Moreover, comparative PAM measurements suggest the background biofilm-associated chlorophyll may be more active than the nodule-associated chlorophyll (Figure 2C). Microelectrode profiling of the nodules and background biofilm show the production of oxygen (Figure 3A, 3B) and elevated pH levels in both structural types (Figure 3C, 3D).

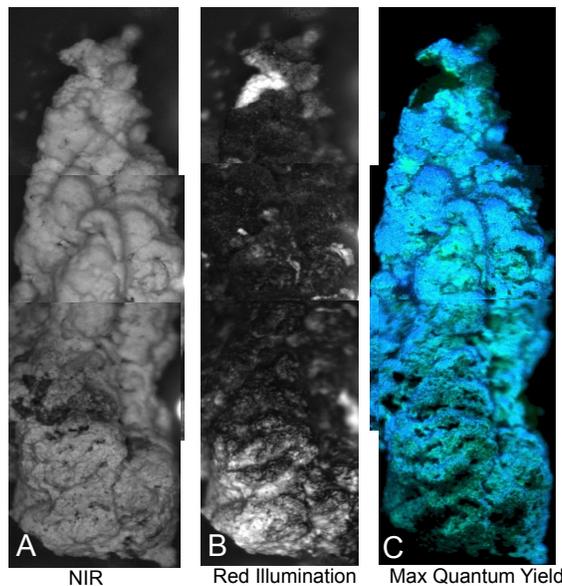


Figure A.1. Shore based PAM and imaging of microbialite mound from 28 meters depth, Three Poles. Panel A: Near Infrared. Panel B: Red illumination. Panel C: Maximum Quantum Yield.

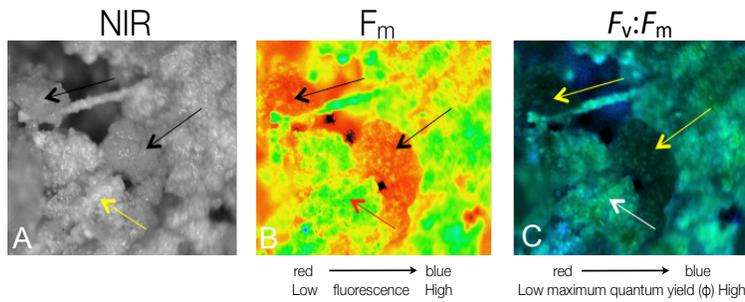


Figure A.2. Shore based PAM and imaging of pink nodules from 21 meters depth, South Basin Panel A: Near Infrared. Panel B: Fluorescence. Panel C: Maximum quantum yield.

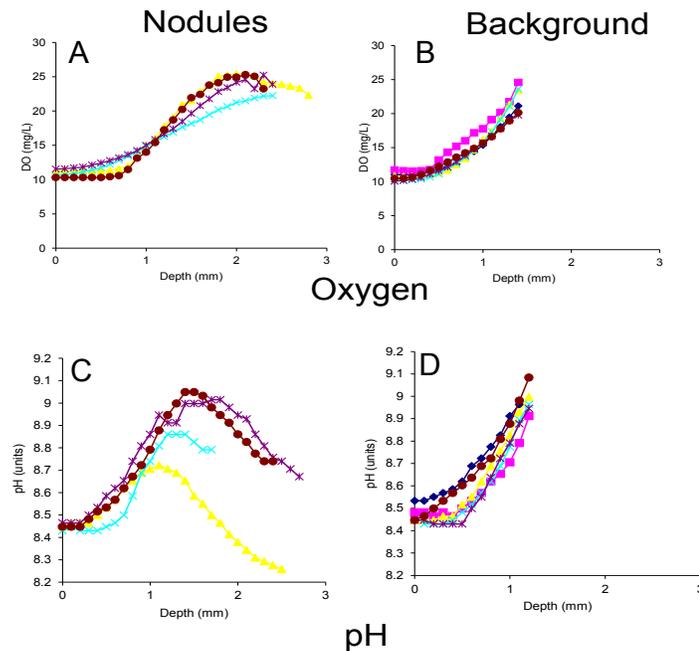


Figure A.3. Microelectrode profiles of nodules and background biofilm. Panel A: Dissolved oxygen measured in nodules. Panel B: Dissolved oxygen measured in background biofilm. Panel C: pH measured in nodules. Panel D: pH measured in background biofilm

Prokaryotic community structure of a microbialite

Overall, Proteobacteria, Cyanobacteria, Bacteroidetes, Chloroflexi, Nitrospira and Acidobacteria characterized the communities from the deep microbialite from Three Poles were (Figure 4A). Together, these phyla represented 97% of the relative abundance of the top region, 95% of the middle region and 94% of the bottom.

Proteobacterial abundances increased going down the length of the microbialite from 25%, to 39%, and 43% respectively. On the other hand, Cyanobacteria showed an opposite trend, with a higher abundance at the top (31%) and a lower abundance at the bottom (25%). Planctomycetes, Bacteroidetes and Chloroflexi also dropped in relative abundances at the bottom of the microbialites.

A cluster-based analysis indicated that the bottom community was different than the top and bottom of the microbialite (Figure 4B). Chao indices of the three communities suggest the bottom community may have higher richness, while the middle community lagged in both richness and diversity. (Figure S1).

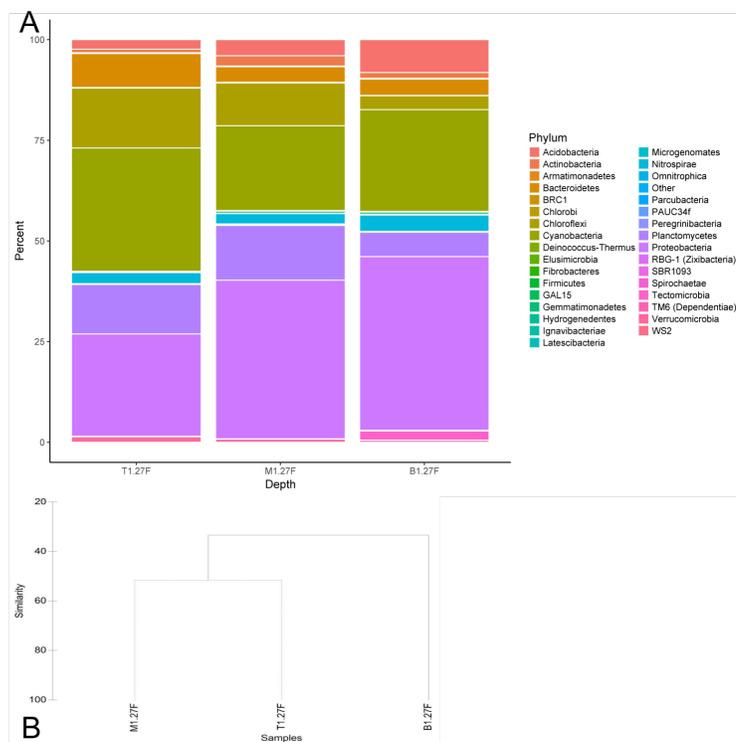


Figure A.4. Prokaryotic OTUs in microbialite mound from Three Poles. Panel A: Percent relative abundance of taxonomic profiles from 16S rRNA gene of mound. T1.27F: Top; M1.27F: Middle; B1.27F: Bottom. Panel B: Cluster diagram from Bray-Curtis resemblance matrix of 16S rRNA genes after square root-transformation. Red branches suggest no significant difference between samples.

Alphaproteobacteria were the most abundant class of the Proteobacteria, comprising 21% of the top community, 31% and 30% of the middle and bottom communities respectively. More specifically, this class was characterized by the pink non-sulfur and phototrophic bacterial orders Rhizobiales, Rhodospiralles and Rhodobacterales (Figure S2). In the Cyanobacterial phylum, Leptolyngbya were most abundant in the top and middle communities (Figure S3). The SILVA database did not assign higher cyanobacterial taxonomies to 282 OTUs, and were assigned to the group “Other” (Figure S3). These OTUs were characterized using BLAST. Table 1 depicts the blastn results from the most abundant of these OTUs. This group was mainly characterized by the Nostocales order.

Community structure of eukaryotic microorganisms of a microbialite

Chlorophytes (green algae) were the most representative eukaryotic group in all communities of the deep microbialite mound (Figure 5A). The relative abundance of this group dropped from 81% at top and 78% in middle, to 42% at the bottom of the mound. Florideophycidae (red algae) increased in abundance at the bottom. More specifically, this group comprised 1.2% and 2.4% at the top and middle of the mound, and jumped to 38% at the bottom. Ochrophyta also increased in the bottom sample and were mostly comprised of Dictyotales (brown algae).

Cluster analyses suggested a significant difference between the bottom community of eukaryotic microbes in comparison to the top and middle communities (Figure 5B). In a similar manner to the prokaryotic communities of the microbialite, richness was highest in the bottom eukaryotic community and dropped in the middle and top communities (Figure S4).

#OTU ID	T1.27F	M1.27F	B1.27F	SILVA Taxonomy	Blastn
denovo45119	1372	915	116	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo5977	1584	790	33	Bacteria; Cyanobacteria; Cyanobacteria	<i>Cylindrospermum stagnale</i>
denovo41842	346	356	21	Bacteria; Cyanobacteria; Cyanobacteria	<i>Anabaena cylindrica</i>
denovo15836	137	126	8	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo46267	150	72	13	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo14427	51	58	3	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo42871	25	18	1	Bacteria; Cyanobacteria; Cyanobacteria	<i>Gloeobacter kilauensis</i>
denovo35054	31	17	1	Bacteria; Cyanobacteria; Cyanobacteria	<i>Anabaena cylindrica</i>
denovo38981	2	15	47	Bacteria; Cyanobacteria; Cyanobacteria	<i>Anabaena cylindrica</i>
denovo25127	16	15	1	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo318	31	10	1	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo11222	2	4	0	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo19966	25	3	1	Bacteria; Cyanobacteria; Cyanobacteria	<i>Cylindrospermum stagnale</i>
denovo17377	39	3	1	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo15714	0	2	7	Bacteria; Cyanobacteria; Cyanobacteria	<i>Anabaena cylindrica</i>
denovo18059	1	1	38	Bacteria; Cyanobacteria; Cyanobacteria	<i>Anabaena cylindrica</i>
denovo10779	2	1	0	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo10276	0	1	0	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo1103	5	0	0	Bacteria; Cyanobacteria; Cyanobacteria	<i>Anabaena cylindrica</i>
denovo10963	2	0	0	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo40443	1	0	22	Bacteria; Cyanobacteria; Cyanobacteria	<i>Stanieria cyanosphaera</i>

Table A.1. OTUs from “Other” Cyanobacteria. OTUs with high abundances were characterized using BLAST and the 16S ribosomal RNA database.

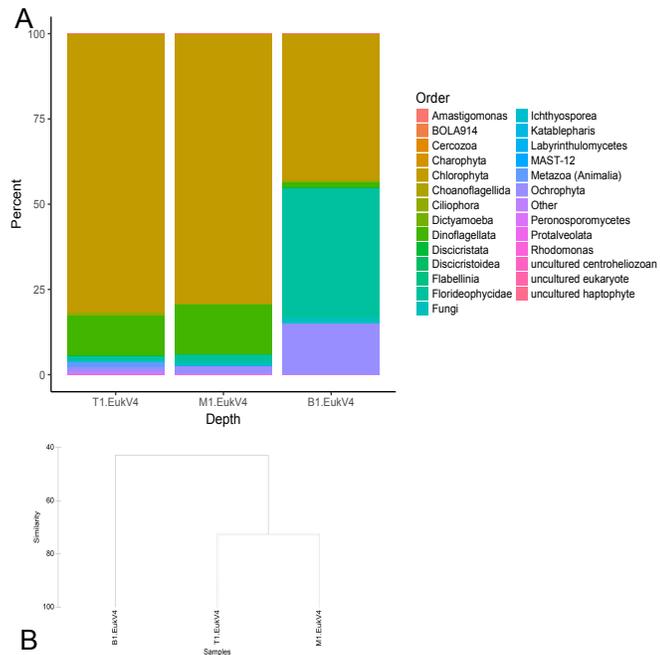


Figure A.5. Eukaryotic OTUs in microbialite mound from Three Poles. Panel A: Percent relative abundance of taxonomic profiles from 18S rRNA genes of mound. T1.EukV4: Top; M1.EukV4: Middle; B1.EukV4: Bottom. Panel B: Cluster diagram from Bray-Curtis resemblance matrix of 18S rRNA gene amplicons after square root-transformation. Red branches suggest no significant difference between samples.

Prokaryotic communities associated with nodules and underlying biofilm

Chloroflexi, Bacteroidetes, Planctomycetes, Cyanobacteria and Proteobacteria comprised the majority of the communities in all three sample types. These phyla represented about 86% of the communities from the background biofilm replicates and 96% of the communities in the green and pink nodules. The relative abundances of Proteobacteria and Cyanobacteria were similar in the background biofilm and green nodules. In the pink nodules, Proteobacteria were slightly more abundant than the Cyanobacteria (Figure 6A).

Much like the microbialite mound from Three Poles, the proteobacterial class that dominated the nodular and biofilm communities were Alphaproteobacteria. The relative abundance of Rhizobiales was highest in the background (10% in both replicates) and green nodular communities (12% in both replicates), but slightly dropped in the pink nodules (7% and 6% in each replicate). Sphingomonadales had higher relative abundances in the nodules than in the background (13% in green nodules, 6% and 11% in pink nodules, 0.8% in background). Rhodobacterales, Parvularculares and Caulobacterales showed highest relative abundances in the pink nodules than in the green nodules and background biofilm (Figure S5).

The relative abundance of Leptolyngbya was highest in the green nodule communities and lowest in the pink nodules (14% and 29% in the green nodules, and 7% in each of the pink nodules). Of all the Cyanobacterial OTUs, 548 were assigned to “Other”. BLAST against the 16S rRNA database showed that the most abundant of these OTUs fell within the orders of Nostocales and Gleobacterales (Table 2). Cluster analyses from the Bray-Curtis resemblance matrix indicated that the prokaryotic communities from each sample type represented distinct groups (Figure 6B). Chao and

Shannon indices of the communities suggest that the background biofilm communities may be richer and more diverse than the green and pink nodules (Figure S7).

The average dissimilarity between pink and green prokaryotic nodule communities was calculated to be 66.6% in PRIMER. Thirty six OTUs were highly contributing to the dissimilarity between pink and green nodules, according to the SIMPER results (Figure 7). The taxonomy of these OTUs was inferred either by the annotation based on the QIIME pipeline and the SILVA database, or by BLAST results for those OTUs that were not assigned higher taxonomy through QIIME. Nostocales represented seven of the cyanobacterial OTUs, while the remaining were Leptolyngbya and Synechococcus. The 13 alphaproteobacterial OTUs, were Rhizobiales, Rhodospirillales, Sphingomonadales, Caulobacterales, Parvularculales, and Rhodobacterales. The other two Proteobacteria were a Deltaproteobacterium and a Gammaproteobacterium. The remaining OTUs were Gemmatimonadetes, Bacteroidetes, Chloroflexi, Planctomycetes and Verrucomicrobia.

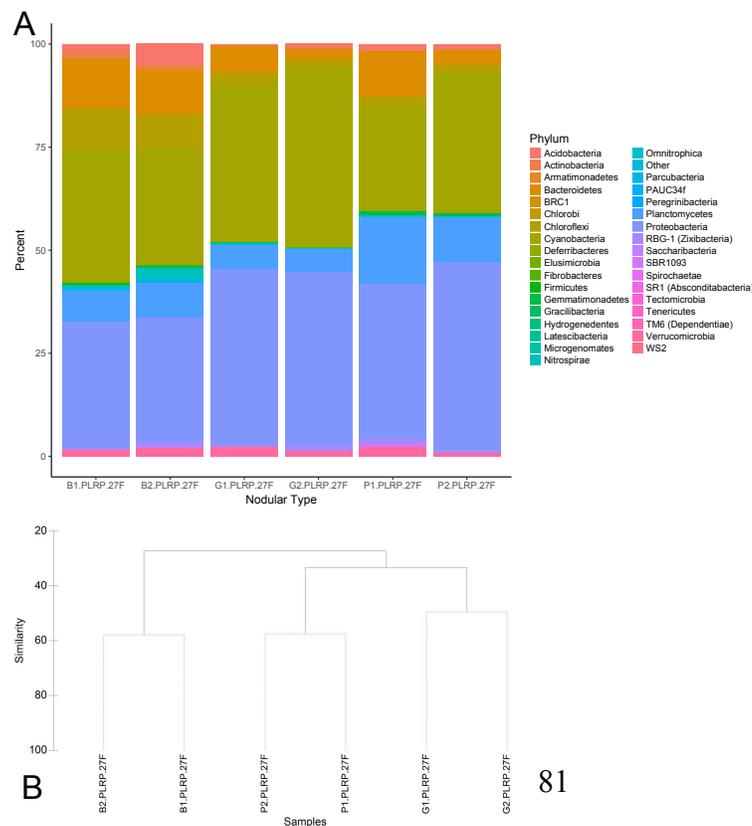


Figure A.6. Prokaryotic OTUs nodules and background biofilm form South Basin East. Panel A: Percent relative abundance of taxonomic profiles from 16S rRNA genes of nodular samples. B1.PLRP.27F, B2.PLRP.27F: Background biofilm; G1.PLRP.27F, G2.PLRP.27F: green nodules; P1.PLRP.27F, P2.PLRP.27F: pink nodules. Panel B: Cluster diagram from Bray-Curtis resemblance matrix of 16S rRNA gene after square root-transformation. Red branches suggest no significant difference between samples.

#OTU ID	B1.PLRP. 27F	B2.PLRP. 27F	P1.PLRP. 27F	P2.PLRP. 27F	G1.PLRP. 27F	G2.PLRP. 27F	SILVA Taxonomy	Blastn
denovo60758	1219	1296	250	80	117	61	Bacteria; Cyanobacteria; Cyanobacteria	<i>Cylindrospermum stagnale</i>
denovo39202	242	356	1912	5204	1101	654	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo27128	31	306	15	13	1750	926	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo55112	222	124	41	20	107	33	Bacteria; Cyanobacteria; Cyanobacteria	<i>Anabaena cylindrica</i>
denovo61375	117	60	19	11	2	1	Bacteria; Cyanobacteria; Cyanobacteria	<i>Gleobacter kilauensis</i>
denovo39272	2	47	1	5	216	110	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo53454	32	44	5	0	1	0	Bacteria; Cyanobacteria; Cyanobacteria	<i>Cylindrospermum stagnale</i>
denovo22600	5	28	13	7	88	30	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo63836	3	26	16	18	55	73	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo40675	1	17	1	0	21	4	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo71883	39	16	0	0	0	1	Bacteria; Cyanobacteria; Cyanobacteria	<i>Cylindrospermum stagnale</i>
denovo15607	1	16	0	0	71	40	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo10375	20	15	0	0	0	0	Bacteria; Cyanobacteria; Cyanobacteria	<i>Chamaesiphon minutus</i>
denovo26621	12	12	0	0	0	2	Bacteria; Cyanobacteria; Cyanobacteria	<i>Anabaena cylindrica</i>
denovo54153	9	12	5	0	1	0	Bacteria; Cyanobacteria; Cyanobacteria	<i>Cylindrospermum stagnale</i>
denovo18296	7	12	0	1	0	0	Bacteria; Cyanobacteria; Cyanobacteria	<i>Cylindrospermum stagnale</i>
denovo34095	19	10	2	2	0	2	Bacteria; Cyanobacteria; Cyanobacteria	<i>Cylindrospermum stagnale</i>
denovo49209	0	10	0	0	9	0	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo72166	1	9	0	2	0	1	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo57560	1	8	1	1	22	16	Bacteria; Cyanobacteria; Cyanobacteria	<i>Nodularia spumigena</i>
denovo72059	7	7	0	1	0	0	Bacteria; Cyanobacteria; Cyanobacteria	<i>Cylindrospermum stagnale</i>

Table A.2. OTUs from the “Other” Cyanobacteria in nodules. The most abundant OTUs were characterized using BLAST and the 16S ribosomal RNA database.

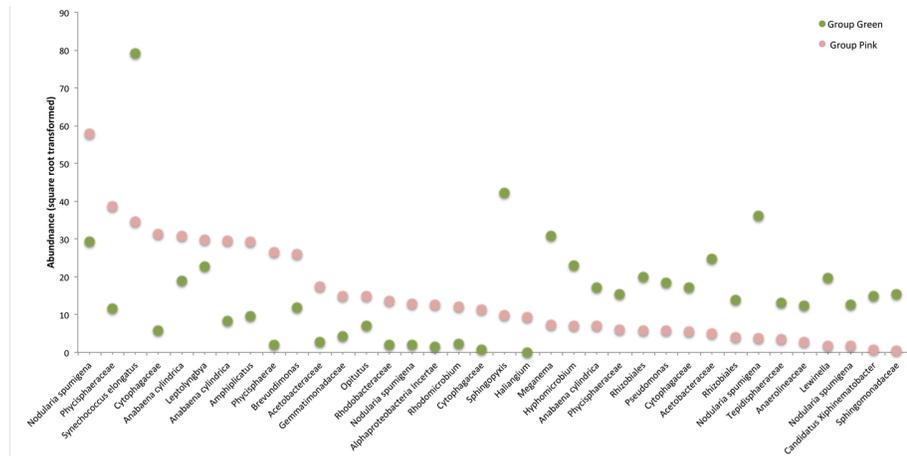


Figure A.7. Similarity percentages (SIMPER) of most contributing OTUs in the prokaryotic communities of green and pink nodules based on Bray-Curtis dissimilarity matrices. The y-axis represents the square root transformation of the abundance values of each OTU in the two groups (pink and green nodules)

Eukaryotic community structure from nodules and background biofilm from SBE

The eukaryotic microorganisms that were associated with nodules were Dinoflagellates (34% and 30% in the background biofilm, 51% and 56% in green nodules, 44% and 61% in pink nodules) and Metazoa (32% and 11% in background biofilm, 42% and 24% in green nodules, 46% and 28% in pink nodules). Of the dinoflagellates, Thoracosphaeraceae comprised 31% and 27% of the background communities, 49% and 50% of the green nodules, and in the pink nodules 39% and 53%. The nematode classes, Enoplia and Chromadorea were the dominant taxa of the Metazoa. In the background communities, Enoplia comprised 28% and 8% of the total communities. In the nodules, OTUs from these nematodes comprised 33% and 12% (green), and 40% and 20% (pink). OTUs related to Chromadorea comprised 3% and 2% of the background communities, 8% and 11% of the green nodular communities, and 5% and 7% of the pink nodular communities. OTUs related to green algae (Chlorophyta) represented 25% and 50% of the background biofilm communities, and only up to 2% of the nodular communities (Figure 8A). Eukaryotic communities associated with nodular samples also separated into distinct groupings, much like the prokaryotic communities (Figure 8B). The calculated Chao and Shannon indexes were lower for the green nodular communities. The number of observed species was highest in the pink nodules (Figure S8).

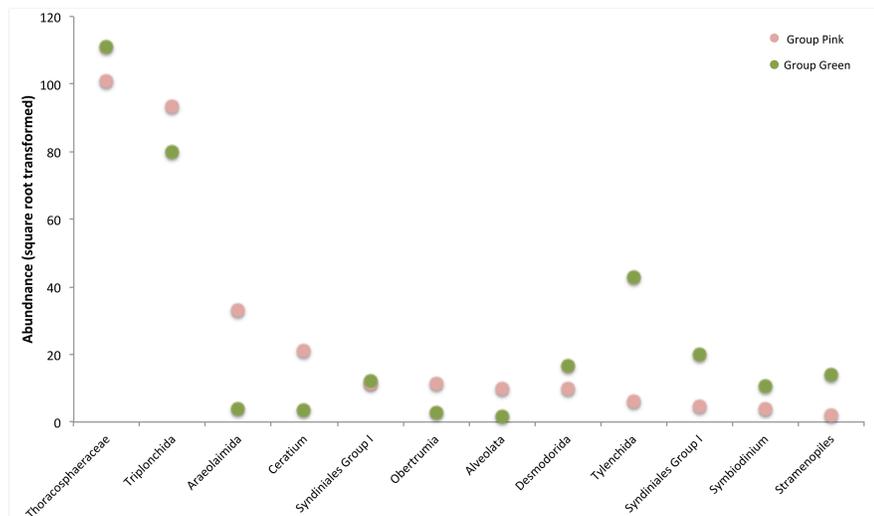


Figure A.9. Similarity percentages (SIMPER) of most contributing OTUs in the eukaryotic communities of green and pink nodules. The y-axis represents the square root transformation of the abundance values of each OTU in the two groups (pink and green nodules)

The average dissimilarity between pink and green prokaryotic nodule communities was calculated to be 41.97% in PRIMER. The abundances of 12 OTUs that contributed the highest percentage dissimilarity and totaled up to 10% of the cumulative dissimilarity were plotted (Figure 9). Seven of the 31 OTUs were Nostocales (Cyanobacteria), five were Rhizobiales (Alphaproteobacteria), and four were Sphingobacteriales (Bacteroidetes). Five of the OTUs were dinoflagellates, four were nematodes, and one was a ciliate. The remaining two were an unclassified Alveolate and an unclassified Stramenopile.

Discussion

Photosynthetic activity and growth are observed across whole mounds, as well as nodular microstructures

The aim of this study was to assess photosynthetic activity and community structures of prokaryotic and eukaryotic microbes associated with various morphological types of microbialites found in Pavilion Lake, BC Canada. Fluorometry and microelectrode profiles were obtained to understand which regions of the microbialite were actively photosynthesizing, while sequencing was performed to gain insight on community composition. The deep (28m) microbialite mound from Three Poles showed a strong photosynthetic signature (Figure 1). Red illumination indicated that the mound was covered in chlorophyll and the maximum quantum yield suggested the whole surface of the structure might be associated with eukaryotic-type photosynthesis. Such activity is consistent with previous studies in which scanning electron images showed the presence of intertwined algae and cyanobacteria on the surface of a PLRP microbialite from 25m depth (Omelon et al., 2013).

Fluorescence and maximum quantum yields from pink nodules and background biofilm at 21m depth showed that chlorophyll-associated fluorescence was lower on the surface of the nodules in comparison to the underlying area (Figure 2). Furthermore, maximum quantum yield was also slightly lower on the nodules. It should be noted that the fluorescence yields may vary according to the pigments and their concentrations associated with the surface layers. The I-PAM instrument emits blue light, which is well absorbed by chlorophyll-a in comparison to other pigments such as phycoerythrin and phycocyanin. Therefore, this method of assessing photosynthetic activity may be biased towards organisms that contain and utilize

chlorophyll-a as their primary light-harvesting pigments (such as diatoms and green algae). The concentration and distribution of chlorophyll-a may also influence the observed differences. Because light is used to assess activity, not much can be inferred about photosynthesis associated with deeper layers of the nodules, since light from the instrument cannot penetrate into the deeper photosynthetic layers of these microstructures. Therefore, if chlorophyll-a is mostly aggregated in a thin layer closer to the surface, the fluorescence yield will be higher in comparison to the same amount of chlorophyll-a that may be more deeply distributed in the thick layers of the nodules.

The microelectrode profiles from nodules and underlying biofilm indicate that oxygen production and elevated pH levels occur within the nodules and in the underlying biofilm. In previous studies, evidence for the biological role in carbonate precipitation had been made via $\delta^{13}\text{C}$ carbonate values measurements from green and pink nodules (Brady et al., 2010). Here we show that the pH levels measured in the underlying biofilm are high enough to promote the precipitation of carbonate. Therefore, microbialite growth may not be limited to the nodular structures since the background biofilm has the ability to also contribute to the carbonate precipitation through biotic activity.

Prokaryotic and eukaryotic communities shift along a single microbialite

Overall, the prokaryotic communities of the microbialite from SBW are comparable to profiles from microbialites obtained from a similar depth (28m), but different location within the lake, that were analyzed in previous studies (Russell et al., 2014). Proteobacteria and Cyanobacteria are the dominating phyla at all regions of the microbialite. The trends of the relative abundances are opposing, with Cyanobacteria decreasing in abundance going down the mound, and Proteobacteria

increasing. Rhizobiales, and the purple phototrophic Rhodobacter and Rhodospirillales are the major groups of the Alphaproteobacteria. Major representatives of Cyanobacteria include the genus *Leptolyngbya* (Subsection III) in the top and middle communities, while *Phormidium* increases in the bottom community. High abundances classify as “Other” Cyanobacteria, similarly to cyanobacterial profiles from microbialites sampled at Three Poles (Russell et al., 2014). Based on BLAST results, the most abundant of these OTUs are classified in the order Nostocales, confirming the association of this group with microbialite mounds found in alkaline lakes (Couradeau et al., 2011). Statistically, the communities differ at the three regions of the mound (Figure 4.B). The presence of phototrophic groups in all three regions may indicate that at the top of the mound Cyanobacteria are the primary drivers of prokaryotic photosynthesis, but are replaced by Alphaproteobacteria towards the bottom of the microbialite. This regional succession in phylogeny may be how the prokaryotic communities of microbialites are able to continue growing under conditions that do not favor cyanobacterial photosynthesis. While the fluorometric data presented in this paper show that a eukaryotic-type photosynthesis is widespread (Figure 1), the lack of data accounting for photosynthetic activity attributed to prokaryotes prevents any further suggestions regarding the contribution of non-eukaryotic photosynthesis (such as alphaproteobacterial and cyanobacterial photosynthesis) to microbialite growth.

The eukaryotic communities from the SBW microbialite show successional trends as well, which should be expected, as they respond to the incoming light that penetrates the water column. The top and middle of the microbialite have higher abundances of dinoflagellates, which have been found in association with

microbialites from other locations within the lake (White, Power, Dipple, Southam, & Suttle, 2015). Since dinoflagellates are mixotrophic, the role they play in microbialite development is hard to discern. While green algae (Chlorophyta) are the dominant order across all regions, the relative abundance of this groups decreases at the bottom of the microbialite, where red (Florideophycidae) and brown algae (Ochrophyta) become more apparent (Figure 5A). These algal types contain the accessory pigments fucoxanthin and phycoerythrin, which absorb blue-green light. If light levels are lower at the bottom of the microbialite, the heightened presence of red and brown algae may be allowed due to their higher content in pigments that absorb deeper penetrating light.

Overall, the community profile of the SBW microbialite is more diverse in prokaryotic composition than in eukaryotic microbe composition (Figure S1; Figure S4). Previous studies that relied on T-RFLP analyses suggested that Pavilion Lake microbialites from Three Poles and Willow Point presented little to no change in bacterial and eukaryotic communities across the mounds that were sampled (Chan et al., 2014). The findings of the present study may provide a better resolution to assess spatial changes in microbial community assemblages across a microbialite mound. Photosynthetic microorganisms found in both the eukaryotic and prokaryotic domains are the dominant groups across the whole microbialite (Figure 4; Figure 5). Likely the ecophysiological habitat of photosynthesizers shifts based on changing environmental conditions across the microbialite. For example, lower light levels or shading may allow for purple phototrophic bacteria or red algae to outcompete cyanobacteria and green algae by utilizing different wavelengths of light towards the base of the microbialite. This suggests that all areas of the microbialite may photosynthesize, and microbialite growth is not restricted to vertical orientations.

Distinct prokaryotes and eukaryotic microbes comprise different nodule types and underlying biofilm

While photosynthetic (e.g. Cyanobacteria) and photoheterotrophic Alphaproteobacteria are present across nodules and the underlying biofilm, heterotrophic bacteria (Sphingomonadales) are associated with nodular types (Figure S3). Heterotrophic respiration of organic matter leads to CO₂ production, which under the high pH levels that are sustained by photosynthetic activity (Figure 3) allows for CaCO₃ precipitation to occur (Omelson et al., 2013). Therefore, different organisms may share the role of carbonate precipitation, through different metabolic capabilities, with nodules supporting heterotrophic calcification (due to respiration from - Sphingomonadales) and background biofilm supporting phototrophic calcification (due to photosynthetic activity from Cyanobacteria and algal groups).

The eukaryotic microbial communities are dominated by few groups that change across sample types. More specifically, the oversaturation of green algae is not apparent in nodules, as this group has a higher relative abundance in the background biofilm (Figure 8A). Dinoflagellates and nematodes (the primary representative of metazoa) are more abundant in green and pink nodules. While the OTU data presented in this study suggest that there may be an association of metazoa with the nodules, these organisms have not yet been microscopically visualized in these sample types. Therefore, at this point, the underlying reason why OTUs from these organisms are more closely associated with the nodule types remains unknown.

Differences in nodular mesostructures may be driven by specific prokaryotic and eukaryotic OTUs

Pink and green nodular communities form distinct groupings for bacteria (Figure 6B) and eukaryotes (Figure 8B). This suggests that the observed differences in these structures may be due to different inhabitants whose presence may impact the physical appearance and growth of the nodules. OTUs assigned to the Nostocales order, such as *Nodularia spumigena*, and *Anabeana cylindrica* were major contributing species that had higher abundance values in the pink nodules (Figure 7). Previous studies have shown that *Tolypothrix* OTUs were observed only in samples of pink nodules and it was suggested that this species contributed to the color difference between nodules, via its ability to change pigment usage based on available light levels (Russell et al., 2014). No such OTUs were found in the nodules sampled for the present study. However, pigment usage adjustment, also known as complementary chromatic adaptation, is an ability that has been attributed to other Cyanobacteria in the Nostocales order and not solely to the genus *Tolypothrix* (Bennett & Bogorad, 1973). Therefore, the resulting pink color in some nodules may be due to an overabundance of Nostocales in general. Other OTUs that contributed to the dissimilarity included several Rhizobiales, such as *Sphingopyxis*, *Meganema*, and *Hyphomicrobium* (Figure 7). The abundances of these OTUs were higher in green nodules than they were in the pink. These organisms are known members of microbialites and other lithifying mats (Warden et al., 2016). (Russell et al., 2014) showed that Rhizobiales were present in green nodules. It may be possible that the nodules appear different, because Cyanobacteria that are capable of

complementary chromatic adaptation are outcompeting Rhizobiales in the pink nodules.

The major contributing eukaryotic OTUs to nodular communities included Nematodes, Dinoflagellates and other Alveolates (Figure 9). Such organisms are known fauna in many microbialite environments (Konishi, Prince, & Knott, 2001). In the context of eukaryotes and their presence on different nodules, there is no clear preference or absolute association between eukaryote group and type of nodule, as representatives of these three phyla are found in both pink and green nodules. However, a study of faunal abundances associated with different microbialite types found in the Exuma Cays, Bahamas, indicated faunal densities were highest in mat types rich in the Nostocale *Dichothrix* (Tarhan, Planavsky, Laumer, Stolz, & Reid, 2013). This microbialite mat type appeared tuft-like, similar to the pink and green nodules of Pavilion Lake. It is possible that both nodule types of accumulated cyanobacteria and other prokaryotic organisms are popular grazing sites for higher-level organisms than the encrusting background that the tufts are found on.

Conclusions

The findings of this study suggest that the photosynthetic activity of Pavilion Lake microbialites is not limited to a single type of microbialite surface. In fact, photosynthesis appears to be occurring along the whole length of a mound, on nodules, as well as the background biofilm the nodular tufts are attached to. This may be possible due to different groups of photosynthetic organisms appearing in succession along the length of a microbialite or within different mesostructures. The organisms found at different regions of the mound may be a determining factor of the overall morphology of the microbialite, as they grow and precipitate calcium

carbonate. While the exact mechanism for this variation still remains unknown, it is possible that oxygen-consuming microorganisms may also play a role in microbialite growth. Finally, eukaryotic organisms seem to be more closely associated with the pink and green nodules. Studies that incorporate activity measurements and infaunal community investigations of nodules will provide further insight regarding the impact of eukaryotes on the development of the nodular mesostructures that are associated with the microbialites of Pavilion Lake.

Supplementary Figures for Appendix Chapter

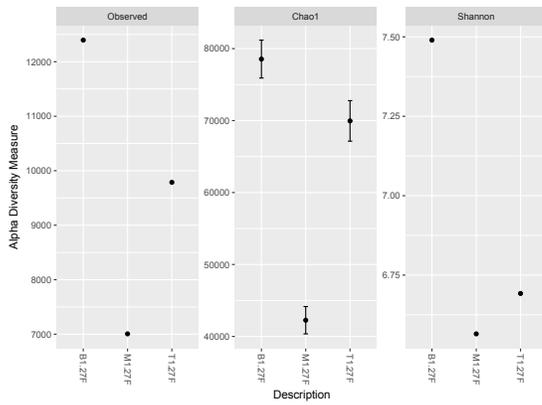


Figure A.S1. Chao and Shannon indexes for the microbialite mound prokaryotic communities.

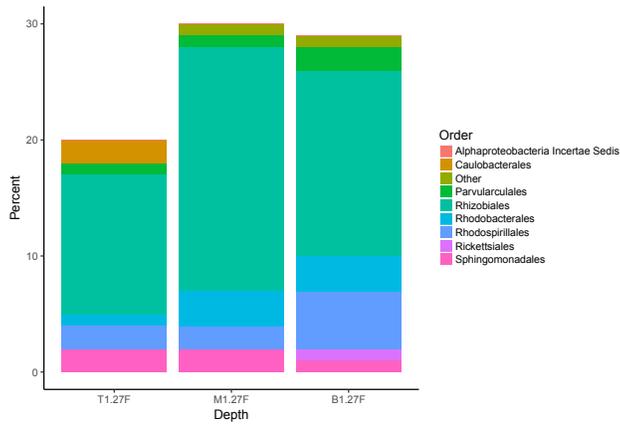


Figure A.S2. Profiles of Alphaproteobacteria from microbialite mound 16S rRNA gene amplicon samples. The percent abundance on the y-axis is the relative abundance from the total community.

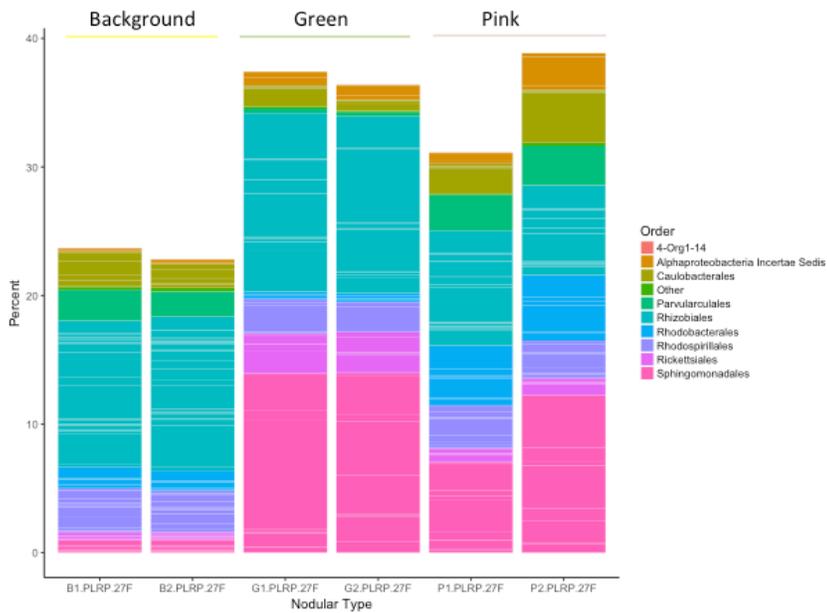


Figure A.S5. Profiles of Alphaproteobacteria from nodule 16S rRNA gene amplicon samples. The percent abundance on the y-axis is the relative abundance from the total community.

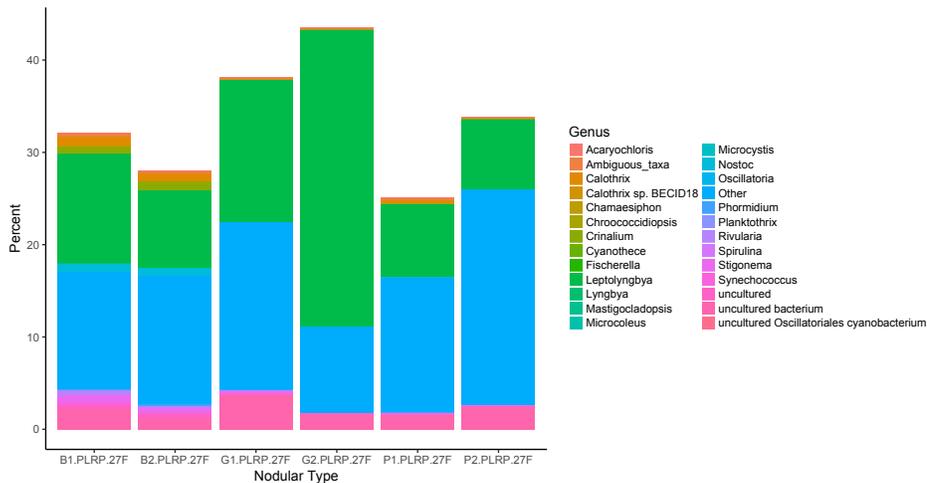


Figure A.S6. Profiles of Cyanobacteria from nodule 16S rRNA gene amplicon samples. The percent abundance on the y-axis is the relative abundance from the total community.

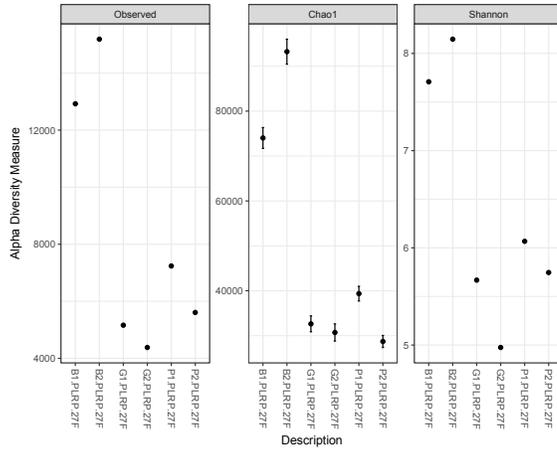


Figure A.S7. Chao and Shannon index for the nodule and background biofilm prokaryotic communities.

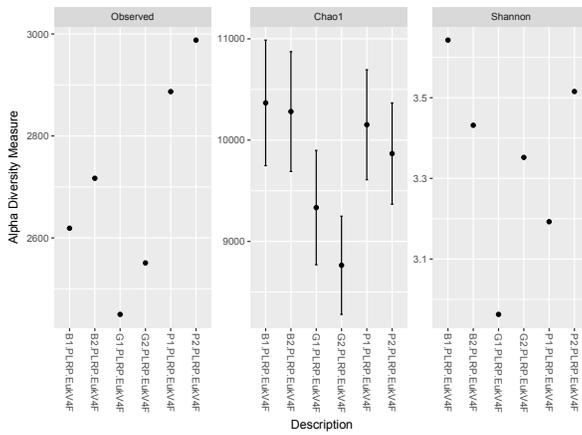


Figure A.S8. Chao and Shannon index for the nodule and background biofilm eukaryotic communities.

References

- Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., ... Quince, C. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods*, *11*, 1144. Retrieved from <http://dx.doi.org/10.1038/nmeth.3103>
- Arkin, A. P., Stevens, R. L., Cottingham, R. W., Maslov, S., Henry, C. S., Dehal, P., ... Yoo, S. (2016). The DOE Systems Biology Knowledgebase (KBBase). *BioRxiv*, 096354. <https://doi.org/10.1101/096354>
- Beck, C., Knoop, H., & Steuer, R. (2018). Modules of co-occurrence in the cyanobacterial pan-genome reveal functional associations between groups of ortholog genes. *PLoS Genetics*, *14*(3), 1–20. <https://doi.org/10.1371/journal.pgen.1007239>
- Benbow, M. E., Pechal, J. L., Lang, J. M., Erb, R., & Wallace, J. R. (2015). The Potential of High-throughput Metagenomic Sequencing of Aquatic Bacterial Communities to Estimate the Postmortem Submersion Interval. *Journal of Forensic Sciences*, *60*(6), 1500–1510. <https://doi.org/10.1111/1556-4029.12859>
- Bennett, A., & Bogorad, L. (1973). Complementary chromatic adaptation in a filamentous blue-green alga. *The Journal of Cell Biology*, *58*, 419–435. <https://doi.org/10.1083/jcb.58.2.419>
- Berg, J., Tymoczko, J., & Stryer, L. (2007). *Biochemistry*. *Biochemistry*. <https://doi.org/10.1007/978-3-8274-2989-6>
- Bolhuis, H., Cretoiu, M. S., & Stal, L. J. (2014). Molecular ecology of microbial mats. *FEMS Microbiology Ecology*, *90*(2), 335–350. <https://doi.org/10.1111/1574-6941.12408>
- Bowles, M., & Joye, S. (2011). High rates of denitrification and nitrate removal in cold seep sediments. *ISME Journal*, *5*(3), 565–567. <https://doi.org/10.1038/ismej.2010.134>
- Bowles, M. W., Nigro, L. M., Teske, A. P., & Joye, S. B. (2012). Denitrification and environmental factors influencing nitrate removal in Guaymas Basin hydrothermally altered sediments. *Frontiers in Microbiology*, *3*(OCT), 1–11. <https://doi.org/10.3389/fmicb.2012.00377>
- Brady, A. L., Slater, G. F., Omelon, C. R., Southam, G., Druschel, G., Andersen, D. T., ... Lim, D. S. S. (2010). Photosynthetic isotope biosignatures in laminated micro-stromatolitic and non-laminated nodules associated with modern, freshwater microbialites in Pavilion Lake, B.C. *Chemical Geology*, *274*(1–2), 56–67. <https://doi.org/10.1016/j.chemgeo.2010.03.016>
- Buchfink, B., Xie, C., & Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, *12*, 59. Retrieved from <http://dx.doi.org/10.1038/nmeth.3176>
- Burne, R. V., & Moore, L. S. (1987). Linked references are available on JSTOR for this article : Microbialites : Organosedimentary Deposits of Benthic Microbial Communities. *PALAIOS*, *2*(3), 241–254.
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello,

- E. K., ... Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7, 335. Retrieved from <http://dx.doi.org/10.1038/nmeth.f.303>
- Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., ... Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, 108(Supplement_1), 4516–4522. <https://doi.org/10.1073/pnas.1000080107>
- Chan, O. W., Bugler-Lacap, D. C., Biddle, J. F., Lim, D. S., McKay, C. P., & Pointing, S. B. (2014). Phylogenetic diversity of a microbialite reef in a cold alkaline freshwater lake. *Canadian Journal of Microbiology*, 60(6), 391–398. <https://doi.org/10.1139/cjm-2014-0024>
- Clarke, K. R., & Gorley, R. N. (2006). PRIMER v6: User Manual/Tutorial. *PRIMER-E, Plymouth UK*, 192 p. <https://doi.org/10.1111/j.1442-9993.1993.tb00438.x>
- Cooper, D. M. (2003). Regulation and organization of adenylyl cyclases and {cAMP}. *Biochem. J.*, 375(Pt 3), 517–529. <https://doi.org/10.1042/BJ20031061>
- Couradeau, E., Benzerara, K., Moreira, D., Gérard, E., Kaźmierczak, J., Tavera, R., & López-García, P. (2011). Prokaryotic and eukaryotic community structure in field and cultured microbialites from the alkaline Lake Alchichica (Mexico). *PLoS ONE*, 6(12). <https://doi.org/10.1371/journal.pone.0028767>
- Darling, A. E., Jospin, G., Lowe, E., Matsen, F. A., Bik, H. M., & Eisen, J. A. (2014). PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ*, 2, e243. <https://doi.org/10.7717/peerj.243>
- Delmont, T. O., & Eren, A. M. (2018). Linking pangenomes and metagenomes: the *Prochlorococcus* metapangenome. *PeerJ*, 6, e4320. <https://doi.org/10.7717/peerj.4320>
- Deming, J. W., Reysenbach, A. L., Macko, S. A., & Smith, C. R. (1997). Evidence for the microbial basis of a chemoautotrophic invertebrate community at a whale fall on the deep seafloor: Bone-colonizing bacteria and invertebrate endosymbionts. *Microscopy Research and Technique*, 37(2), 162–170. [https://doi.org/10.1002/\(SICI\)1097-0029\(19970415\)37:2<162::AID-JEMT4>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1097-0029(19970415)37:2<162::AID-JEMT4>3.0.CO;2-Q)
- Donati, C., Hiller, N. L., Tettelin, H., Muzzi, A., Croucher, N. J., Angiuoli, S. V., ... Massignani, V. (2010). Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. *Genome Biology*, 11(10), R107. <https://doi.org/10.1186/gb-2010-11-10-r107>
- Dupraz, C., Reid, R. P., Braissant, O., Decho, A. W., Norman, R. S., & Visscher, P. T. (2009). Processes of carbonate precipitation in modern microbial mats. *Earth-Science Reviews*, 96(3), 141–162. <https://doi.org/10.1016/j.earscirev.2008.10.005>
- Dworkin, M. (2012). Sergei Winogradsky: A founder of modern microbiology and the first microbial ecologist. *FEMS Microbiology Reviews*, 36(2), 364–379. <https://doi.org/10.1111/j.1574-6976.2011.00299.x>
- Eprintsev, A. T., Falaleeva, M. I., Grabovich, M. Y., Parfenova, N. V., Kashirskaya,

- N. N., & Dubinina, G. A. (2004). The role of malate dehydrogenase isoforms in the regulation of anabolic and catabolic processes in the colorless sulfur bacterium *Beggiatoa leptomitiformis* D-402. *Microbiology*, *73*(4), 367–371. <https://doi.org/10.1023/B:MICI.0000036977.79822.7a>
- Fierer, N., Nemergut, D., Knight, R., & Craine, J. M. (2010). Changes through time: Integrating microorganisms into the study of succession. *Research in Microbiology*, *161*(8), 635–642. <https://doi.org/10.1016/j.resmic.2010.06.002>
- Flood, B. E., Bailey, J. V., & Biddle, J. F. (2014). Horizontal gene transfer and the rock record: Comparative genomics of phylogenetically distant bacteria that induce wrinkle structure formation in modern sediments. *Geobiology*, *12*(2), 119–132. <https://doi.org/10.1111/gbi.12072>
- Fowler, A. C., & Winstanley, H. F. (2018). Microbial dormancy and boom-and-bust population dynamics under starvation stress. *Theoretical Population Biology*, *120*, 114–120. <https://doi.org/10.1016/j.tpb.2018.02.001>
- Goffredi, S. K., & Orphan, V. J. (2010). Bacterial community shifts in taxa and diversity in response to localized organic loading in the deep sea. *Environmental Microbiology*, *12*(2), 344–363. <https://doi.org/10.1111/j.1462-2920.2009.02072.x>
- Grabovich, M. Y., Dubinina, G. A., Lebedeva, V. Y., & Churikova, V. V. (1998). Mixotrophic and lithoheterotrophic growth of the freshwater filamentous sulfur bacterium *Beggiatoa leptomitiformis* D-402. *Mikrobiologiya*, *67*, 383–388.
- Graham, E. D., Heidelberg, J. F., & Tully, B. J. (2017). BinSanity: unsupervised clustering of environmental microbial assemblies using coverage and affinity propagation. *PeerJ*, *5*, e3035. <https://doi.org/10.7717/peerj.3035>
- Grote, J., Jost, G., Labrenz, M., Herndl, G. J., & Jürgens, K. (2008). Epsilonproteobacteria represent the major portion of chemoautotrophic bacteria in sulfidic waters of pelagic redoxclines of the Baltic and Black Seas. *Applied and Environmental Microbiology*, *74*(24), 7546–7551. <https://doi.org/10.1128/AEM.01186-08>
- Gude, H., Strohl, W. R., & Larkin, J. M. (1981). Mixotrophic and heterotrophic growth of *Beggiatoa alba* in continuous culture. *Archives of Microbiology*, *129*, 357–360.
- Hilario, A., Cunha, M. R., Génio, L., Marçal, A. R., Ravara, A., Rodrigues, C. F., & Wiklund, H. (2015). First clues on the ecology of whale falls in the deep Atlantic Ocean: Results from an experiment using cow carcasses. *Marine Ecology*, *36*(S1), 82–90. <https://doi.org/10.1111/maec.12246>
- Hladyz, S., Cook, R. A., Petrie, R., & Nielsen, D. L. (2010). Influence of substratum on the variability of benthic biofilm stable isotope signatures: Implications for energy flow to a primary consumer. *Hydrobiologia*, *664*(1), 135–146. <https://doi.org/10.1007/s10750-010-0593-0>
- Iseki, M., Matsunaga, S., Murakami, A., Ohno, K., Shiga, K., Yoshida, K., ... Watanabe, M. (2002). A blue-light-activated adenylyl cyclase mediates photoavoidance in *Euglena gracilis*. *Nature*, *415*(6875), 1047–1051. <https://doi.org/10.1038/4151047a>

- Jannasch, H. W., Nelson, C. D., & Wirsén, C. O. (1989). Massive Natural Occurrence of Unusually Large Bacteria (*Beggiatoa* sp.) at a Hydrothermal Deep-Sea Vent Site. *Nature*, *340*(6231), 301–303. <https://doi.org/10.1038/340301a0>
- Jones, D. S., Flood, B. E., & Bailey, J. V. (2015). Metatranscriptomic analysis of diminutive Thiomargarita-like bacteria (“*Candidatus Thiopilula*” spp.) from abyssal cold seeps of the Barbados accretionary prism. *Applied and Environmental Microbiology*, *81*(9), 3142–3156. <https://doi.org/10.1128/AEM.00039-15>
- Kaas, R. S., Friis, C., Ussery, D. W., & Aarestrup, F. M. (2012). Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse *Escherichia coli* genomes. *BMC Genomics*, *13*(1), 1. <https://doi.org/10.1186/1471-2164-13-577>
- Kalenitchenko, D., Le Bris, N., Peru, E., & Galand, P. E. (2018). Ultrarare marine microbes contribute to key sulphur-related ecosystem functions. *Molecular Ecology*. <https://doi.org/10.1111/mec.14513>
- Kalenitchenko, D., Dupraz, M., Le Bris, N., Petetin, C., Rose, C., West, N. J., & Galand, P. E. (2016). Ecological succession leads to chemosynthesis in mats colonizing wood in sea water. *ISME Journal*, *10*(9), 2246–2258. <https://doi.org/10.1038/ismej.2016.12>
- 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. <https://doi.org/10.1016/j.jmb.2015.11.006>
- Kang, D. D., Froula, J., Egan, R., & Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ*, *3*, e1165. <https://doi.org/10.7717/peerj.1165>
- Kim, T., Folcher, M., Baba, M. D. El, & Fussenegger, M. (2015). A synthetic erectile optogenetic stimulator enabling blue-light-inducible penile erection. *Angewandte Chemie - International Edition*, *54*(20), 5933–5938. <https://doi.org/10.1002/anie.201412204>
- Konishi, Y., Prince, J., & Knott, B. (2001). The fauna of thrombolitic microbialites, Lake Clifton, Western Australia. *Hydrobiologia*, *457*(1985), 39–47. <https://doi.org/10.1023/A:1012229412462>
- Kreutzmann, A.-C., & N. Schulz-Vogt, H. N. (2016). Oxidation of Molecular Hydrogen by a Chemolithoautotrophic ABSTRACT. *Applied and Environmental Microbiology*, *82*(8), 2527–2536. <https://doi.org/10.1128/AEM.03818-15>. Editor
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. <https://doi.org/10.1093/molbev/msw054>
- Laczny, C. C., Sternal, T., Plugaru, V., Gawron, P., Atashpendar, A., Margossian, H. H., ... Wilmes, P. (2015). VizBin - an application for reference-independent visualization and human-augmented binning of metagenomic data. *Microbiome*. London. <https://doi.org/10.1186/s40168-014-0066-1>

- Lang, J., Erb, R., Pechal, J., Wallace, J., McEwan, R., & Benbow, M. (2016). Microbial Biofilm Community Variation in Flowing Habitats: Potential Utility as Bioindicators of Postmortem Submersion Intervals. *Microorganisms*, 4(1), 1. <https://doi.org/10.3390/microorganisms4010001>
- MacGregor, B. J., Biddle, J. F., Harbort, C., Matthyse, A. G., & Teske, A. (2013). Sulfide oxidation, nitrate respiration, carbon acquisition, and electron transport pathways suggested by the draft genome of a single orange Guaymas Basin *Beggiatoa* (Cand. *Maribeggiatoa*) sp. filament. *Marine Genomics*, 11, 53–65. <https://doi.org/10.1016/j.margen.2013.08.001>
- MacGregor, B. J., Biddle, J. F., & Teske, A. (2013). Mobile elements in a single-filament orange guaymas basin *beggiatoa* (“*Candidatus maribeggiatoa*”) sp. draft genome: Evidence for genetic exchange with cyanobacteria. *Applied and Environmental Microbiology*, 79(13), 3974–3985. <https://doi.org/10.1128/AEM.03821-12>
- Marshall, I. P. G., Blainey, P. C., Spormann, A. M., & Quake, S. R. (2012). A single-cell genome for *Thiovulum* sp. *Applied and Environmental Microbiology*, 78(24), 8555–8563. <https://doi.org/10.1128/AEM.02314-12>
- McKay, L. J., MacGregor, B. J., Biddle, J. F., Albert, D. B., Mendlovitz, H. P., Hoer, D. R., ... Teske, A. P. (2012). Spatial heterogeneity and underlying geochemistry of phylogenetically diverse orange and white *Beggiatoa* mats in Guaymas Basin hydrothermal sediments. *Deep-Sea Research Part I: Oceanographic Research Papers*, 67, 21–31. <https://doi.org/10.1016/j.dsr.2012.04.011>
- McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, 8(4). <https://doi.org/10.1371/journal.pone.0061217>
- Merz, M. U. E. (1992). The biology of carbonate precipitation by cyanobacteria. *Facies*, 26(1), 81–101. <https://doi.org/10.1007/BF02539795>
- Mezzino, M. J., Strohl, W. R., & Larkin, J. M. (1984). Characterization of *Beggiatoa alba*. *Archives of Microbiology*, 137(2), 139–144. <https://doi.org/10.1007/BF00414455>
- Mills, H. J., Mills, H. J., Martinez, R. J., Martinez, R. J., Story, S., Story, S., ... Sobecky, P. a. (2004). Identification of Members of the Metabolically Active Microbial Populations Associated with. *Society*, 70(9), 5447–5458. <https://doi.org/10.1128/AEM.70.9.5447>
- Moldovan, M. A., & Gelfand, M. S. (2018). Pangenomic definition of prokaryotic species and the phylogenetic structure of *Prochlorococcus* spp. *Frontiers in Microbiology*, 9(MAR), 1–11. <https://doi.org/10.3389/fmicb.2018.00428>
- Mosquera-Rendón, J., Rada-Bravo, A. M., Cárdenas-Brito, S., Corredor, M., Restrepo-Pineda, E., & Benítez-Páez, A. (2016). Pangenome-wide and molecular evolution analyses of the *Pseudomonas aeruginosa* species. *BMC Genomics*, 17(1), 1–14. <https://doi.org/10.1186/s12864-016-2364-4>
- Mußmann, M., Hu, F. Z., Richter, M., De Beer, D., Preisler, A., Jørgensen, B. B., ... Ehrlich, G. D. (2007). Insights into the genome of large sulfur bacteria revealed

- by analysis of single filaments. *PLoS Biology*, 5(9), 1923–1937.
<https://doi.org/10.1371/journal.pbio.0050230>
- Nagórska, K., Hinc, K., Strauch, M. A., & Obuchowski, M. (2008). Influence of the sigB stress factor and yxaB, the gene for a putative exopolysaccharide synthase under sigB control, on biofilm formation. *Journal of Bacteriology*, 190(10), 3546–3556. <https://doi.org/10.1128/JB.01665-07>
- Nelson, D. C., & Castenholz, R. W. (1982). Light responses of Beggiatoa. *Archives of Microbiology*, 131(2), 146–155. <https://doi.org/10.1007/BF01053997>
- Nelson, D. C., & Jannasch, H. W. (1983). Chemoautotrophic growth of a marine Beggiatoa in sulfide-gradient cultures. *Archives of Microbiology*, 136(4), 262–269. <https://doi.org/10.1007/BF00425214>
- Nelson, D. C., Wirsén, C. O., & Jannasch, H. W. (1989). Characterization of Large, Autotrophic Beggiatoa spp. Abundant at Hydrothermal Vents of the Guaymas Basin. *Applied and Environmental Microbiology*, 55(11), 2909–2917.
- Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., ... Pevzner, P. A. (2013). Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. In M. Deng, R. Jiang, F. Sun, & X. Zhang (Eds.), *Research in Computational Molecular Biology* (pp. 158–170). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Omelon, C. R., Brady, A. L., Slater, G. F., Laval, B., Lim, D. S. S., & Southam, G. (2013). Microstructure variability in freshwater microbialites, Pavilion Lake, Canada. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 392, 62–70. <https://doi.org/10.1016/j.palaeo.2013.08.017>
- Pace, N. R. (1997). A molecular view of microbial diversity and the biosphere. [Review] [52 refs]. *Science*, 276(5313), 734–740.
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., ... Parkhill, J. (2015). Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, 31(22), 3691–3693. <https://doi.org/10.1093/bioinformatics/btv421>
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from. *Cold Spring Harbor Laboratory Press Method*, 1, 1–31. <https://doi.org/10.1101/gr.186072.114>
- Pati, A., Gronow, S., Lapidus, A., Copeland, A., del Rio, T. G., Nolan, M., ... Kyrpides, N. C. (2010). Complete genome sequence of Arcobacter nitrofigilis type strain (CI T). *Standards in Genomic Sciences*, 2(3), 300–308. <https://doi.org/10.4056/sigs.912121>
- Patrinskaya, V. Y., Grabovich, M. Y., Muntyan, M. S., & Dubinina, G. A. (2001). Lithoautotrophic growth of the freshwater colorless sulfur bacterium Beggiatoa “leptomitiformis” D-402. *Microbiology*, 70(2), 145–150. <https://doi.org/10.1023/A:1010469211447>
- Pernestig, A. K., Georgellis, D., Romeo, T., Suzuki, K., Tomenius, H., Normark, S., & Melefors, Ö. (2003). The Escherichia coli BarA-UvrY two-component system is

- needed for efficient switching between glycolytic and gluconeogenic carbon sources. *Journal of Bacteriology*, 185(3), 843–853.
<https://doi.org/10.1128/JB.185.3.843-853.2003>
- Podell, S., & Gaasterland, T. (2007). DarkHorse: A method for genome-wide prediction of horizontal gene transfer. *Genome Biology*, 8(2).
<https://doi.org/10.1186/gb-2007-8-2-r16>
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1), 590–596.
<https://doi.org/10.1093/nar/gks1219>
- Ray, V. A., & Visick, K. L. (2013). LuxU Connects Quorum Sensing to Biofilm Formation in *Vibrio fischeri*, 86(4), 954–970.
<https://doi.org/10.1111/mmi.12035.LuxU>
- Reno, M. L., Held, N. L., Fields, C. J., Burke, P. V., & Whitaker, R. J. (2009). Biogeography of the *Sulfolobus islandicus* pan-genome. *Proceedings of the National Academy of Sciences*, 106(44), 18873–18873.
<https://doi.org/10.1073/pnas.0911400106>
- Roalkvam, I., Drønen, K., Stokke, R., Daae, F. L., Dahle, H., & Steen, I. H. (2015). Physiological and genomic characterization of *arcobacter anaerophilus* IR-1 reveals new metabolic features in epsilonproteobacteria. *Frontiers in Microbiology*, 6(SEP), 1–12. <https://doi.org/10.3389/fmicb.2015.00987>
- Russell, J. A., Brady, A. L., Cardman, Z., Slater, G. F., Lim, D. S. S., & Biddle, J. F. (2014). Prokaryote populations of extant microbialites along a depth gradient in Pavilion Lake, British Columbia, Canada. *Geobiology*, 12(3), 250–264.
<https://doi.org/10.1111/gbi.12082>
- Ryu, M. H., Moskvina, O. V., Siltberg-Liberles, J., & Gomelsky, M. (2010). Natural and engineered photoactivated nucleotidyl cyclases for optogenetic applications. *Journal of Biological Chemistry*, 285(53), 41501–41508.
<https://doi.org/10.1074/jbc.M110.177600>
- Salman, V., Amann, R., Girnath, A. C., Polerecky, L., Bailey, J. V., Høglund, S., ... Schulz-Vogt, H. N. (2011). A single-cell sequencing approach to the classification of large, vacuolated sulfur bacteria. *Systematic and Applied Microbiology*, 34(4), 243–259. <https://doi.org/10.1016/j.syapm.2011.02.001>
- Salman, V., Bailey, J. V., & Teske, A. (2013). Phylogenetic and morphologic complexity of giant sulphur bacteria. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 104(2), 169–186.
<https://doi.org/10.1007/s10482-013-9952-y>
- Schutte, C. A., Teske, A., MacGregor, B. J., Salman Carvalho, V., Lavik, Gaute, ... De Beer, D. (2018). Filamentous giant Beggiatoaceae from Guaymas Basin are capable of both denitrification and dissimilatory nitrate reduction to ammonium (DNRA). *Applied and Environmental Microbiology*.
<https://doi.org/10.1128/AEM.03397-16>
- Schwedt, A. (2011). Physiology of a marine Beggiatoa strain and the accompanying

- organism *Pseudovibrio* sp . – a facultatively oligotrophic bacterium. *PhD Thesis*, (September).
- Seckbach, J., & Aharon, O. (2010). *Microbial Mats Modern and Ancient Microorganisms in Stratified Systems*.
- Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>
- Sharrar, A. M., Flood, B. E., Bailey, J. V., Jones, D. S., Biddanda, B. A., Ruberg, S. A., ... Dick, G. J. (2017). Novel large sulfur bacteria in the metagenomes of groundwater-fed chemosynthetic microbial mats in the Lake Huron basin. *Frontiers in Microbiology*, *8*(MAY), 1–15. <https://doi.org/10.3389/fmicb.2017.00791>
- Smith, C. R., Glover, A. G., Treude, T., Higgs, N. D., & Amon, D. J. (2015). Whale-Fall Ecosystems: Recent Insights into Ecology, Paleoecology, and Evolution. *Annual Review of Marine Science*, *7*(1), 571–596. <https://doi.org/10.1146/annurev-marine-010213-135144>
- Srivastava, D., & Waters, C. M. (2012). A tangled web: Regulatory connections between quorum sensing and cyclic Di-GMP. *Journal of Bacteriology*, *194*(17), 4485–4493. <https://doi.org/10.1128/JB.00379-12>
- Stierl, M., Stumpf, P., Udvari, D., Gueta, R., Hagedorn, R., Losi, A., ... Hegemann, P. (2011). Light modulation of cellular cAMP by a small bacterial photoactivated adenyl cyclase, bPAC, of the soil bacterium *Beggiatoa*. *Journal of Biological Chemistry*, *286*(2), 1181–1188. <https://doi.org/10.1074/jbc.M110.185496>
- Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D. M., Breiner, H. W., & Richards, T. A. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular Ecology*, *19*(SUPPL. 1), 21–31. <https://doi.org/10.1111/j.1365-294X.2009.04480.x>
- Strohl, W. R., & Larkin, J. M. (1978). Enumeration, isolation, and characterization of *Beggiatoa* from fresh water sediments. *Appl. Environ. Microbiol.*, *36*(5), 755–770.
- Sweerts, J. P. R. A., De Beer, D., Nielsen, L. P., Verdouw, H., Van Den Heuvel, J. C., Cohen, Y., & Cappenberg, T. E. (1990). Denitrification by sulphur oxidizing *Beggiatoa* spp. mats on freshwater sediments. *Nature*. <https://doi.org/10.1038/344762a0>
- Tapley, D. W., Buettner, G. R., & Shick, J. M. (1999). Free radicals and chemiluminescence as products of the spontaneous oxidation of sulfide in seawater, and their biological implications. *Biological Bulletin*, *196*(1), 52–56. <https://doi.org/10.2307/1543166>
- Tarhan, L. G., Planavsky, N. J., Laumer, C. E., Stolz, J. F., & Reid, R. P. (2013). Microbial mat controls on infaunal abundance and diversity in modern marine microbialites. *Geobiology*, *11*(5), 485–497. <https://doi.org/10.1111/gbi.12049>
- Tettelin, H., Masignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., ... Fraser, C. M. (2005). Genome analysis of multiple pathogenic isolates of

- Streptococcus agalactiae: Implications for the microbial “pan-genome.” *Proceedings of the National Academy of Sciences*, 102(39), 13950–13955. <https://doi.org/10.1073/pnas.0506758102>
- Tringe, S. G., von Mering, C., Kobayashi, A. ;, Salamov, A. A., Chen, K. ;, Chang, H. W., ... Rubin, E. M. (2005). “Comparative metagenomics of microbial communities.” *Science* 308.5721 (2005): 554-557. *Science*, 308(5721), 4. <https://doi.org/10.1126/science.1107851>
- Tully, B. J., Wheat, C. G., Glazer, B. T., & Huber, J. A. (2018). A dynamic microbial community with high functional redundancy inhabits the cold, oxic seafloor aquifer. *ISME Journal*, 12(1), 1–16. <https://doi.org/10.1038/ismej.2017.187>
- Vernikos, G., Medini, D., Riley, D. R., & Tettelin, H. (2015). Ten years of pan-genome analyses. *Current Opinion in Microbiology*, 23, 148–154. <https://doi.org/10.1016/j.mib.2014.11.016>
- Vetriani, C., Voordeckers, J. W., Crespo-Medina, M., O’Brien, C. E., Giovannelli, D., & Lutz, R. A. (2014). Deep-sea hydrothermal vent Epsilonproteobacteria encode a conserved and widespread nitrate reduction pathway (Nap). *ISME Journal*, 8(7), 1510–1521. <https://doi.org/10.1038/ismej.2013.246>
- Warden, J. G., Casaburi, G., Omelon, C. R., Bennett, P. C., Breecker, D. O., & Foster, J. S. (2016). Characterization of microbial mat microbiomes in the modern thrombolite ecosystem of lake Clifton, western Australia using shotgun metagenomics. *Frontiers in Microbiology*, 7(JUL), 1–14. <https://doi.org/10.3389/fmicb.2016.01064>
- Wasmund, K., Mußmann, M., & Loy, A. (2017). The life sulfuric: microbial ecology of sulfur cycling in marine sediments. *Environmental Microbiology Reports*, 9(4), 323–344. <https://doi.org/10.1111/1758-2229.12538>
- Weisburg, W. G., Barns, S. M., Pelletier, D. A., & Lane, D. J. (1991). 16S Ribosomal DNA Amplification for Phylogenetic Study. *Journal of Bacteriology*, 173(2), 697–703. <https://doi.org/n.a>
- White, R. A., Power, I. M., Dipple, G. M., Southam, G., & Suttle, C. A. (2015). Metagenomic analysis reveals that modern microbialites and polar microbial mats have similar taxonomic and functional potential. *Frontiers in Microbiology*, 6(SEP). <https://doi.org/10.3389/fmicb.2015.00966>
- Wirsen, C. O., & Jannasch, H. W. (1978). Physiological and morphological observations on *Thiovulum* sp. *Journal of Bacteriology*, 136(2), 765–774. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=218603&tool=pmcentrez&rendertype=abstract>
- Wu, Y.-W., Tang, Y.-H., Tringe, S. G., Simmons, B. A., & Singer, S. W. (2014). MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome*, 2(1), 26. <https://doi.org/10.1186/2049-2618-2-26>
- Yamamoto, M., & Takai, K. (2011). Sulfur metabolisms in epsilon- and gamma-proteobacteria in deep-sea hydrothermal fields. *Frontiers in Microbiology*,

2(SEP), 1–8. <https://doi.org/10.3389/fmicb.2011.00192>

Zhang, Y., & Sievert, S. M. (2014). Pan-genome analyses identify lineage- and niche-specific markers of evolution and adaptation in Epsilonproteobacteria. *Frontiers in Microbiology*, 5(MAR), 1–13. <https://doi.org/10.3389/fmicb.2014.00110>