COLONIZATION AND S(0) MINERALIZATION OF SULFUR OXIDIZING BIOFILMS IN THE FRASASSI CAVE SYSTEM

by

Michael J. Pavia

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Science in Biological Sciences

Summer 2018

COLONIZATION AND S(0) MINERALIZATION OF

SULFUR OXIDIZING BIOFILMS IN THE

FRASASSI CAVE SYSTEM

by

Michael J. Pavia

Approved: _____

Clara Chan, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____

E. Fidelma Boyd, Ph.D.
Chair of the Department of Biological Science

Approved: _____

George H. Watson, Ph.D.
Dean of the College of Arts and Sciences

Approved: _____

Douglas J. Doren, Ph.D.
Interim Vice Provost for Graduate and Professional Education

ACKNOWLEDGMENTS

First and foremost, I would like to thank all members of field team who traveled to the Frasassi Cave System and collected the samples I used for my thesis; These members include Dr. Clara Chan, Dr. Pauline Henri, and Dr. Jennifer Macalady. I would also like to thank the support this research received from the Open Geospatial Consortium.

I want to thank Clara, for her enthusiastic support of my scientific development, encouragement in trying new ideas, and always challenging me to consider different views of my data. As your student, I have grown immensely as a scientist. You have opened up my world to bioinformatics and an appreciation for scientific writing, which both will be indispensable as I continue my career. I would also like to thank members of my committee, Dr. E. Fidelma Boyd, for her diversity of expertise and encouragement, Dr. Shawn W. Polson, for his integral role in developing my bioinformatics skills, and to Dr. Erica M. Selva, who has played a larger role than just a committee member, and someone I feel I can now call a friend.

I must also thank the Chan Lab members, both present and past who have supported me throughout my time at the University of Delaware (UD). Special thanks to Pauline, who was less of a fellow member, and has been more a co-advisor. I cannot thank you enough for being with me every step of the way and challenging me more

than I thought possible. To Arkadiy Garber, Prasanna Joglekar, and Sean McAllister whose lunch time guidance and friendship have been major players in my development and enjoyment during my time at UD.

To Dr. Carlton Cooper, my friend and colleague, many thanks for believing and mentoring me in the various aspects of teaching at the college level. It is with his guidance I choose to peruse a career in higher education, and it would have never been possible if I had not met him. I am immensely grateful.

I would also like to thank the coaches and members of CrossFit Petram for providing me with a healthy avenue to funnel my stress and frustrations. You are a fabulous community, and I wish you all only the best.

To all of my friends at UD, Christine Chapman, Karen Christe, Nicholas Egan, Anders Kiledal, Colleen Pike, and Christopher Materna. I do not believe it would have been possible to make it through this arduous task without your constant support, and belief in me. The hours we have spent sitting at our kitchen table discussing science and life have truly allowed me to grow as a person, and the thanks can never be repaid.

Most importantly, my family. The ones who have and always will be my rock! I also want to dedicate this to my late grandfather. As, the only person in my family who pursued higher education it was amazing to have him in the beginning of my time here at UD, and I have used his memory to help get me through this past year.

In the words of Sasha Velour, “A thinking queen speaks to the heart. She’s stranger than fiction, better than art”. I may officially be done with course work required of a master’s student, but I’ll never stop learning and I’ll always be grateful

to the people who taught me how to learn. It is the breath of our knowledge and experiences that makes us different, and I am grateful to the UD community for shaping me into who I am today.

TABLE OF CONTENTS

LIST OF TABLES viii
LIST OF FIGURES x
ABSTRACT xii

Chapter

1 INTRODUCTION 1

1.1 Microbial Importance in the Sulfur Cycle 1
1.2 Taxonomically Diverse Sulfur Oxidizing Prokaryotes 2
1.3 Mechanism of Sulfide Oxidation 4
1.4 Frasassi Cave System and the Multi-omic Approach 7

2 MATERIALS & METHODS 10

2.1 Sample Acquisition and Preparation 10
2.2 Sequencing and Analysis of Bacteria 16S Ribosomal RNA Gene 10
2.3 Co-assembly of Three Metagenomes from the Pozzo dei Cristalli 11
2.4 Metagenome Sequencing and *De novo* Reconstruction of MAGs 12
2.5 Metatranscriptomic Library Preparation and Sequencing 14
2.6 Metatranscriptomic Analysis 15
2.7 Building the *Sulfurovum* Pangenomes and Defining Functional Clusters 15

3 RESULTS AND DISCUSSION 17

3.1 Geochemistry of Pozzo di Cristalli Stream and Colonization Observations 17
3.2 Metagenomic Sample Selection 20

3.2.1 Phylogeny of Community Composition and Metatranscriptome 25

3.3 Metabolic Functions of the Community 27
3.4 Pangenomic Analysis 33
3.5 Biogeochemical Niches of *Sulfurovum* and Other Dominant MAGs 38

3.5.1 Carbon 39

3.5.2 Oxygen 41
3.5.3 Nitrogen 43
3.5.4 Sulfur 45

3.6 Significance and Conclusion 47

REFERENCES 51

Appendix

A SUPPLEMENTARY MATERIAL 60

LIST OF TABLES

Table 1 *In situ* sulfide and oxygen concentrations during colonization. Measurements were made at stream source (Up) and between Net 2 and Net 3 (Down). Exact locations are indicated on Figure 1. At all sampling times pH and percent salinity was 7.1-7.2 and 1.3, respectively. Temperature remained stable throughout colonization at 13.7°C 19

Table 2 Microbial community morphologies of samples collected for molecular analysis 19

Table 3 Metagenome sequencing statistics. 22

Table 4 Statistics from the three assembly methods 23

Table 5 MAGs used for metatranscriptomic mapping. SL- single library SA – sub-assembly 24

Table 6 Statistics from metatranscriptome and mapping to MAGs. 25

Table 7 Percent of TPM coming from the dominant expressing organisms for sulfur oxidation genes 29

Table 8 Percent of TPM coming from the dominant expressing organisms for terminal oxidases 30

Table 9 Percent of TPM coming from the top three expressing MAGs for gene involved in carbon fixation. 32

Table 10 RAST verification of mis-annotated *sox* genes 36

Table 11 Percent of COG Category found within each pangenome MAG 38

Table 12 Presence and absence of key metabolic genes in dominant MAGs. Bolded letters indicate presence of only one subunit. 40

Table 13 Percent of expression by class expressing genes indicative of nitrogen cycling. 44

Table A.1 Protein sequences used for the concatenated single copy gene tree..... 60

Table A.2 The Rank abundance of genes in the cbb3 terminal oxidase. These ranks do not include hypothetical and structural ribosome genes..... 60

LIST OF FIGURES

Figure 1 The common Sox pathway (Modified from Grabarczyk & Berks, 2017) 6

Figure 2 (A) Installation of colonization nets within the Pozzo di Cristalli stream. Red X indicates the stream source, and the dashed line represents the flow of the stream. Yellow arrows indicate the locations of *in situ* geochemistry measurements. Green arrow indicates location of natural streamer collected for comparison to biofilms that colonized nets. (B) Macroscopic observation of biofilm colonization of Net 1 over time. All scales bars are 8cm. 18

Figure 3 Biofilm morphologies observed on Net 3 Day 17. 20

Figure 4 Microbial community structure based on the 16S sequences. (A) Tree based on the Yue & Clayton measure of dissimilatory between the structures of the biofilm communities. (B) Distribution of OTU and their taxonomic class. (*) indicates samples chosen for metagenomics. 21

Figure 5 Coverage and community expression of the MAGs within the streamer samples. (A) Metagenomic coverage (B) Total expression of MAGs 26

Figure 6 Normalized abundance and expression of pathways found within the streamer community for oxygen, nitrogen, phosphorous, carbon and sulfur metabolisms. Community is made up of the 30 MAGs. (A) Metagenomes were normalized by the single copy gene, DNA-directed RNA polymerase beta subunit gene (*rpoB*). Pathways were normalized to number of genes, and a minmax normalized abundance of this subset is plotted. (B) Metatranscriptome is plotted as expression (TPM), and pathways were normalized to number of genes. *ccoNOP* was divided by 10. 28

Figure 7 Maximum likelihood phylogenetic tree of 11 concatenated single-copy proteins showing the taxonomic placement of the *Campylobacterota* MAGs. Branch colors indicate environment form where genome or MAG was isolated. 34

Figure 8 Venn diagram of gene clusters and functional annotations shared by the seven Sulfurovum Pozzo dei Cristalli Metagenomes. (A) Unique gene clusters (9703) identified in the Anvi'o pangenomic pipeline. (B) Functional Pangenome based on 1906 unique COG annotations. 35

Figure 9 Log transformed TPM of terminal cytochrome c oxidases within the Halothiobacillaceae PC16-C1 and the two Sulfurovum (PC16-C8 and PC16-C9)..... 42

Figure 10 Log transformed TPM of sulfur oxidation genes with the *Sulfurovum* and *Halothiobacillaceae.* 46

Figure A.1 Community wide expression of the subunits involved in the pst system .. 61

Figure A.2 Pairwise comparison of Amino Acid Identity (AAI) within 281 MAGs generated from the three assembly methods. Each leaf on the dendrogram represents a different MAG comparison. Clusters of MAGs with >95% AAI identity are in yellow. 62

ABSTRACT

Elemental sulfur (S(0)) is an important geologic resource, yet we are still understanding the mechanisms by which sulfur oxidizing microbial communities facilitate its formation. The Pozzo dei Cristalli stream in the Frasassi Cave System is inhabited by *Sulfurovum* streamers rich in S(0), implying a microbial role in mineral formation. Using metagenomics and metatranscriptomics, we examined how the community functions to make S(0), and more specifically, to explore genetic diversity within the *Sulfurovum.* We sequenced metagenomes and 16S rRNA from colonization net samples, over seventeen days, that were placed in the air water interface of the Pozzo dei Cristalli steam. The community structure was simple with two taxa making up 75% to 95% of the community abundance: a Gammaproteobacteria *Halothiobacillaceae* (PC16-C1) and a Campylobacterota *Sulfurovum* (PC16-C2). After dereplication, we recovered thirty high-quality representative metagenome assembled genomes (MAGs), of which included seven *Sulfurovum.* From the metatranscriptomes, transcripts mapping to these MAGs show that the *Sulfurovum* conduct sulfur oxidation by multiple (differing by MAG), aerobic respiration, denitrification, and carbon fixation. *Sulfurovum* PC16-C2 was by far the most abundant and active in the streamer community, suggesting this population are the streamer architects and primary producers. This population, *Sulfurovum* PC16-C2,

highly expressed *sqr*, which would lead to S(0) formation, likely making PC16-C2 the primary contributor of S(0). To understand the role of the flanking *Sulfurovum* (seven in total), we performed a pangenomic to detect core (shared) and strain specific genes in the *Sulfurovum* MAGs. This revealed that 21% to 35% of genes found in the *Sulfurovum* MAGs are core, while 12% to 44% of genes were strain specific. All but one *Sulfurovum* MAG has *sqr*, but diversity was apparent in the gene content of the *sox* pathway. Remarkably, the seven Frasassi *Sulfurovum* MAGs span the phylogenetic diversity of *Sulfurovum*. Future work should focus on understanding if the Frasassi *Sulfurovum* are representative of their phylogenetic group; perhaps by performing a more comprehensive pangenomic analysis. In the case that the Frasassi *Sulfurovum* are characteristic, these MAGs could be used to understand what genes drive diversity and success of the cosmopolitan *Sulfurovum*, and fundamentally the geochemical niches that could support biogenic S(0) formation.

Chapter 1

INTRODUCTION

1.1 Microbial Importance in the Sulfur Cycle

Microbial communities play a significant role in the formation, consumption and immobilization of various biologically important elements. The concentration of sulfur in the Earth's crust is 260 ppm, and can be commonly found in -2, 0, or +6 oxidation states, represented by sulfide, elemental sulfur (S(0)), and sulfate (SO_4) (Taylor, 1964; Roy & Trudinger, 1970; Friedrich et al., 2005; Andreote et al. 2012). Environments with oxygen-sulfide chemoclines sulfur compounds with mixed oxidation states (eg. thiosulfate, -2 & +6) are found at lower concentrations (Roy & Trudinger, 1970). The variation of sulfur compounds gives rise to a complex geochemical cycle in which microbially-mediated biotransformations are poorly understood (Ivanov, 1968; Trudinger, 1970).

S(0) is a water insoluble intermediate redox species that can be found in environments like anoxic waters, marine sediments, terrestrial geothermal vents, and sulfidic caves (Toelsen & Jorgensen, 1982; Luther III et al., 1991; Nordstrom et al., 2005; Galdenzi et al., 2008). The formation of S(0) can occur through two reactions:

$$H_2S + \frac{1}{2}O_2 \rightarrow S(0) + H_2O \qquad \Delta G^{\circ\prime} = -420.7 \text{ kJ/rxn}$$

or:

$$S_2O_3 + \frac{1}{2}O_2 \rightarrow S(0) + SO_4 \qquad \Delta G^{\circ\prime} = -229.6 \text{ kJ/rxn}$$

The rate of this reaction by abiotic mechanisms is slow, occurring on the order of hours to days, which allows microbes to outcompete the geochemistry, and use reduced sulfur compounds like sulfide and thiosulfate in microaerophilic environments as sources of energy (Canfield, 2004; Dahl, 2017). Electrons captured from the oxidation of sulfide and thiosulfate are used to generate proton motive force in the respiratory chain where oxygen or nitrate can act as terminal electron acceptors (Hayes & Waldbauer, 2006). In many sulfur oxidizing bacteria, S(0) results as an intermediate product during the oxidation of sulfide or thiosulfate (Howarth & Jorgensen, 1984; Zhang & Millero, 1993). However, the full range of microbial taxa that drive this process and the role of ecological interactions in the production and consumption of S(0) is still poorly known.

1.2 Taxonomically Diverse Sulfur Oxidizing Prokaryotes

Lithotrophic sulfur oxidation is carried out by diverse photosynthetic and non-photosynthetic organisms. This microbial lifestyle is commonly found within the phylum Proteobacteria, namely in the classes of Alpha-, Beta-, and Gamma-subdivisions (Ghosh & Dam, 2009). Similarly, this mode of metabolism is also recognized in the Gram-positive genus *Sulfobacillus* (Fuchs et al.,1996), the archaea phylum *Crenarchaeota* (Johnson et al., 2008), and the recently reclassified phylum *Campylobacterota* (previously classified as *Epsilonproteobacteria*) (Waite et al.,

2017). Lithotrophic sulfur oxidation has also be suggested as one of the earliest metabolisms because of deeply branching sulfur oxidizers in both bacterial and archaeal lineages (Fuchs et al.,1996; Ghosh & Dam, 2009). Diversity in sulfur oxidizers has led to the evolution of many biochemical and molecular mechanisms that oxidize a wide range of reduced sulfur compounds.

Anoxygenic photolithotrophic sulfur oxidizers carry out light-dependent energy transfer with reduced sulfur compounds as electron donors. They are found in anoxic environments in the presence of light, such as thermally stratified meromictic lakes or saline lagoons (Frigaard & Dahl, 2008; Dahl, 2017). In most cases, S(0) is a common oxidation intermediate in these environments where sulfide concentrations are high. As a result, bacterial blooms of these anoxygenic photolithotrophs can produce on a regional scale deposition of S(0) that can be observed using satellites (Lavik et al. 2009).

Many non-photosynthetic sulfur oxidizers use molecular oxygen as a terminal electron acceptor (Ghosh & Dam, 2007). These oxygen dependent sulfur oxidizers can be found in wastewater, deep-sea hydrothermal vents, and sulfidic cave systems (Robertson & Keunen, 1983). However, some species, in particular, within the phylum *Campylobacterota* are facultatively anaerobic and also able to oxidize sulfur compounds using nitrate as their terminal electron acceptor (Campbell et al., 2006). In these ecosystems, devoid of light, chemolithotroph sulfur oxidizers have been suggested as the primary producers using the energy generated from sulfur oxidation to fix carbon (Campbell et al., 2003; Campbell et al., 2006).

Many photosynthetic and non-photosynthetic sulfur oxidizers form intra- or extracellular S(0) globules. These sulfur rich globules form from long chains of polymeric sulfur and can be comprised of compounds like cyclooctasulfur (S_8), polymeric sulfur (S_n), or polysulfides (S_n^{2-}). Intracellular S(0) globule formation is tightly regulated in *Allochromatium vinosum* (Stockdreher et al., 2012; Weissgerber et al., 2014). In this species, they act as an energy storage for further oxidation when electron donors are scarce. However, this may or may not be true for extracellular S(0) globule forming microbes, and it remains unclear whether these S(0) globules serve as energy reserves when electron donors are scarce, or if they are simply waste products from energy metabolism (Friedrich et al. 2005; Campbell et al., 2006; Macalady et al., 2006).

1.3 Mechanism of Sulfide Oxidation

Microbially mediated S(0) formation by sulfide oxidation can be carried out by two discrete periplasmic enzymes: flavocytochrome c (FccAB) and sulfide-quinone reductase (SQR) (Yamanaka et al., 1979; Shutz et al., 1999; Xia et al., 2017). Consisting of a larger flavoprotein (fccB), and a smaller hemoprotein (fccA) this soluble enzyme transfers electrons from sulfide to cytochromes (Kostanjececki et al., 2000). Localized in the inner membrane, *sqr* is a membrane bound single polypeptide with the catalytic site facing the periplasm and transferring electrons to the quinone pool via FAD (Marcia et al., 2009). Both of these enzymes are key players in the biotic formation of S(0) in the environment (Friedrich et al., 2005).

1.4 Mechanism of Thiosulfate Oxidation

A wide variety of sulfur oxidizing chemolithotrophs have the ability to oxidize thiosulfate by route of two known biochemical pathways, Tetrathionate intermediate pathway, and the thiosulfate oxidizing pathway (Sox) (Friedrich et al., 2005; Grimm & Dahl, 2008). Work in a number of organisms on Sox suggests this pathway can play a role in the production and break down of S(0) and other polysulfides (Rother et al., 2001; Eddie et al., 2013; Pjevac et al., 2014). To date, the reactions of the Sox enzymes has only been studied in *Paracoccus panotrophus*, so it is still unclear if these catalytic steps are shared throughout thiosulfate oxidizers. However, in the traditional sox pathway, thiosulfate is completely oxidized to sulfate in a four-step reaction (Sauve et al., 2007).

Located in the periplasmic space, the *sox* pathway is encoded by the conserved gene cluster *soxABCDXYZ*, each playing a distinct role in the oxidation of sulfur compounds (Kelly et al., 1997; Appia Amye et al., 2001; Bamford et al., 2002). Import of thiosulfate into the periplasm begins this multi-step enzymatic relay system by first covalently bonding of the sulfane atom, of thiosulfate, to SoxYZ catalyzed by SoxAX (Figure 1) (Quentmeier & Friedrich, 2001). Bound to SoxYZ the sulfone group is hydrolytically released as SO_4 by SoxB (Suave et al., 2007). SoxCD reduces the sulfane atom on SoxYZ, driven by a six-electron transfer through the dissociation of a water molecule (Suave et al., 2007). This newly formed sulfone group, bound to

SoxYZ, is ultimately cleaved by SoxB releasing another SO_4. Organisms lacking in genetic potential or expression of SoxCD do not reduce the sulfane atom bound to SoxYZ resulting in a buildup of polysulfides, which are ultimately deposited as S(0) (Quentmeier & Friedrich, 2001; Suave et al., 2007; Grimm & Dahl, 2008).

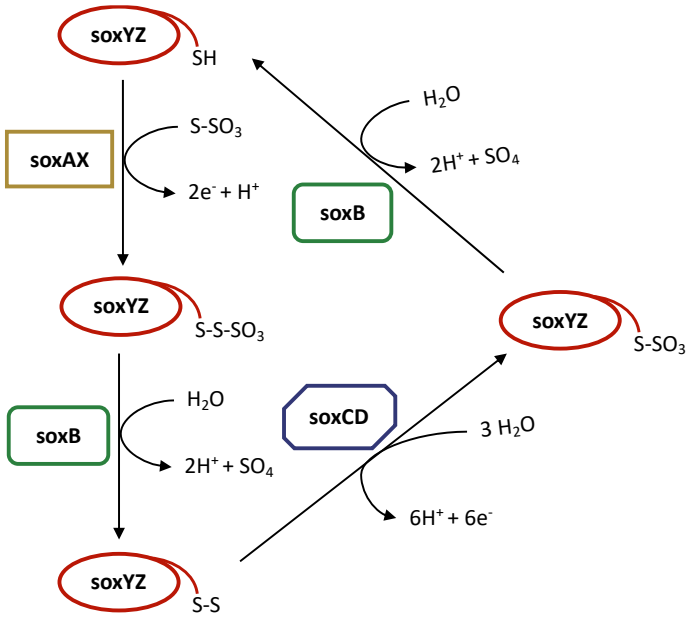

Figure 1 The common Sox pathway (Modified from Grabarczyk & Berks, 2017)

Further, *in situ* assays of the entire Sox pathway, with the exception of one enzymatic step have shown oxidation of sulfide, S(0), thiosulfate, and sulfite (Rother et al., 2001). Oxidation of thiosulfate and sulfite required both the presence of SoxYZ and SoxB, yet S(0) consumption was measured for all combinations of these Sox enzymes. The mechanism by which S(0) is produced and consumed by the Sox

pathway is still unknown, yet its presence within a genome may suggests variety in the source of reduced sulfur compounds used as electron donors.

1.4 Frasassi Cave System and the Multi-omic Approach

The subsurface sulfidic Frasassi cave ecosystem is exclusively sustained by lithotrophic sulfur oxidizing microbes, which makes it an ideal location to study how microbial communities produce and consume S(0) (Macalady et al., 2006; 2007; 2008; Jones et al., 2001). Floating in seasonally sulfidic ground water, these microbes form white streamers that are primarily sulfur by weight. It has been suggested that these streamers play a major role in the cave sulfur cycle, as well as may act as a store of S(0) within the cave (Macalady et al., 2006;2007;2008; Galdenzi et al., 2008; Hamilton et al., 2015).

A recent study by Hamilton et al. (2015) analyzed metagenomes from Pozzo di Cristalli stream, in the Frasassi cave system of Italy, and the neighboring locations Acquasanta Terme (cave) and Fissure Spring (aboveground). Metagenomics is a sequenced-based culture-independent analysis of the collective microbial genomes present in a given habitat (Riesenfeld et al., 2004; Simon and Daniel, 2009). With this method we can identify both the taxonomic abundance, as well as, the repertoire of genetic potential within a microbial community. Moreover, these genetic sequences can be "binned" based on sequence composition and read abundance to reconstruct individual microbial genomes from the community. Hamilton et al. were able to recover four high quality *Sulfurovum* metagenome assembled genomes (MAGs) that

contained homologs of *fccAB*, *sqr*, and incomplete *sox* pathways. The presence of these genes corresponded to the high-sulfide geochemistry of the water suggest that the *Sulfurovum* have potential for S(0) deposition within the streamer. The association of *Sulfurovum* with high sulfide environments and S(0) has also been observed in a wide range of other habitats, such as hydrothermal vents, glacial deposits, and tar seeps (Huber et al., 2010; Flores et al., 2011; Wright et al., 2013; Lormieres & Oger, 2017; Hamilton et al., 2015).

While the enzymatic mechanism by which the *Sulfurovum* deposit S(0) may be associated with *sqr*, *fccAB*, and/or the *sox* pathway, it has still not been shown if the Pozzo dei Cristalli *Sulfurovum* are coupling sulfur oxidation to oxygen or nitrate, both of which are limiting within the water (Hamilton et al., 2015). To understand the role of *Sulfurovum* in S(0) deposition we need to determine the enzymatic players, as well as understand the microbial consortium within the steamer. We took a metatranscriptomic approach to identify the expression profiles of the microbial community during streamer growth.

In this study, we identified potential key players of S(0) deposition, and potential players in the consumption of S(0) within streamer biofilms that colonized nets placed at the air-water interface within the Pozzo dei Cristalli stream. First, we sequenced six metagenomes from the colonization nets for *de novo* reconstruction of community member genomes. This was done to provide high quality MAGs for subsequent expression profiling, as well as to understand the genetic diversity within the *Sulfurovum* and other dominant organisms. For comparison of the genetic diversity

within the Frasassi Cave *Sulfurovum* we performed a pangenomic analysis to identify a core group of genes that allow this group of organisms to occupy and dominate the Pozzo dei Cristalli stream. The metagenomes and metatranscriptomes provide new insights into the cycling sulfur, carbon, oxygen and nitrogen, as well as, the overall diversity of *Sulfurovum* within the Frasassi Cave System.

Chapter 2

MATERIALS & METHODS

2.1 Sample Acquisition and Preparation

String nets were strung across the Pozzo di Cristalli stream in three locations, using Eppendorf tubes to keep the nets floating at the air-water interface. Over the course of seventeen days, streamers were sampled and placed in RNA Later and stored in a -20°C freezer. Samples were shipped on dry ice and stored at -80°C until nucleic acid extraction. Dissolved sulfide concentrations in the Pozzo dei Cristalli stream, were measured with a portable spectrophotometer using the methylene blue method (Hach Co., Loveland, CO, USA; Hach method 831, program 690). Within twelve hours of stream water collection, sulfate concentrations were measured at the Osservatorio Geologico di Coldigioco Geomicrobiology Lab using the manufacturer's instructions for the SulfaVer methods (Hach method 8051, program 680).

2.2 Sequencing and Analysis of Bacteria 16S Ribosomal RNA Gene

Genomic DNA was extracted from RNAlater (Invitrogen, United States) stored streamer samples using the FastDNA™ SPIN Kit for Soil (MP Biomedicals), following the manufacturer instructions. The V4-V5 regions of the 16S ribosomal

RNA gene (16S gene) were done using the MiSeq sequencing platform at the Centre for Comparative Genomics and Evolutionary Bioinformatics Integrated Microbiome Resource (CGEB IMR, Dalhousie University, Canada). The 16S gene data set was processed using the open-source pipeline in mothur (Schloss et al., 2009). Processed sequences were grouped into operational taxonomic units (OTU) with a 99% similarity threshold. We chose 99% in order to resolve closely related and recently diverged organisms of the *Sulfurovum* group (Fox et al., 1992; Nguyen et al., 2016). Representative sequences for each OTU were classified using the Bayesian classifier included in the mothur workflow (bootstrap cutoff 80%), using the SILVA taxonomy framework (Quast et al., 2012). A distance matrix was generated using a cutoff of 0.03 and used to calculate the Yue and Clayton theta, an index that measures the dissimilarity between sample communities based on the relative abundance of OTUs in samples (Yue & Clayton, 2005). DNA extraction and sequencing were conducted by Dr. Pauline Henri, while we both conducted data processing and analysis.

2.3 Co-assembly of Three Metagenomes from the Pozzo dei Cristalli

In order to identify expression profiles of specific taxa and the genetic diversity we needed to reconstruct genomes of the net colonizers. Our first approach was to co-assemble previously collected Pozzo dei Cristalli metagenomes from Hamilton et al., (2015). Raw sequencing libraries were downloaded from the NCBI Short Read Archive using the accession codes SRR1560850, SRR1560064, and SRR1559230 (Hamilton et al., 2015). Raw sequence reads were flashed and quality-

trimmed, using a quality control pipeline combining Trimmomatic and Flash (www.githib.com/mooreryan/qc). Quality controlled reads from the three metagenomes were merged and assembled using metaSPAdes v3.8.0 (Bankevich et al., 2015), with default settings. Contigs longer than 2000bp were binned using MetaBat v0.32.4 (Kang et al., 2015), GroopM v0.3.4 (Imelfort et al., 2014) and MaxBin v2.2.1 (Wu et al., 2014). CheckM v1.0.7 (Parks et al., 2015) was used to assess the completion and redundancy. Six MAGs were considered high quality (>90% complete and <10% redundant) and uploaded to Rapid Annotation using Subsystem Technology (RAST) (Aziz et al., 2008). All MAGs were phylogenetically distant from the streamer OTUs, identified by single copy gene sequence similarity. Because we were unable to reconstruct representative genomes from published metagenomes from the Pozzo dei Cristalli, we needed to build metagenomes from our net samples.

2.4 Metagenome Sequencing and *De novo* Reconstruction of MAGs

Genomic DNA was extracted as previously described in section 2.2. Samples were library prepped and sequenced using Illumina HiSeq 2 x 251 bp at University of Delaware Sequencing & Genotyping Center (Newark, DE). Metagenomic reads were quality controlled using the previously described pipeline (www.githib.com/mooreryan/qc). Two approaches were used to assemble the quality-controlled reads, both Single library and a co-assembly. In both approaches, metaSPAdes v3.8.8 (Bankevich et al., 2015) was used to assemble reads to contigs.

Quality controlled contigs, from both assembly approaches, over 2000bp were then binned by four different binning tools: MetaBAT (using the options very sensitive and super specific), MaxBin, CONCOCT and BinSanity (Alneberg et al., 2013; Graham et al., 2017; Kang et al., 2015; Wu et al., 2014). All MAGs produced were then put into DAS_Tool (Sieber et al., 2017) to dereplicate and extract the highest quality MAGs from each tool. The quality of MAGs obtained from DAS_Tool was assessed using the lineage workflow in CheckM v1.0.7 (Parks et al., 2015). MAGs identified as being over 90% complete were then annotated using RAST (Aziz et al., 2008), and their taxonomy was determined using Phylosift (Darling et al., 2014). Coverage of the MAGs, used as a proxy of abundance, was calculate by normalizing the coverage calculate by the jgi-pipeline by the length of the contigs.

Among all the samples, the dominant *Sulfurovum* MAG, representing 8 to 53% of the population, was consistently low quality across all samples and assembly approaches. Because this MAG is so dominant over all samples, it was important to retrieve an almost complete MAG. To improve it, we subsampled 1% of the single library quality-controlled reads three times and re-performed the steps from assembly to binning on the three subsampled datasets for each sample. The *Sulfurovum* MAGs obtained from the 1% subsampling had an increased completeness in 4 samples over the 6 total (data not shown).

We reconstructed 281 MAGs from the metagenomes and then clustered the MAGs based on their amino acid similarity using AAI (https://github.com/mooreryan/aai) (Figure A.2). Open reading frames (ORFs) in all

the MAGs were called using Prodigal (Hyatt et al., 2010). Although very stringent, this method is sensitive to the MAG completeness and we used Average Nucleotide Identity (ANI) to confirm the clusters using OrthoANI (Ouk Kim et al., 2016). AAI clustering produced 73 clusters of MAGs where the similarity was over 95% and considered to be the same species. Within each of these ANI clusters, a representative MAG was chosen based on a combination of high completeness and low contamination. The MAGs were then given the designation PC16, for **P**ozzo dei **C**ristalli stream, **16** for the year the samples were collected, and **C#**, which indicates what ANI cluster they represent. These representative MAGs were then annotated using RAST and KEGG Automatic Annotation Server (KAAS) (Moriya et al., 2007; Aziz et al., 2008).

2.5 Metatranscriptomic Library Preparation and Sequencing

RNA was extracted using the NucleoSpin RNA Isolation Kit (Machery-Nagel), with the modification of cell lysis via FastPrep Lysing Matrix E tubes (MP Bio), and the addition of SUPERase In RNase Inhibitor (Thermofisher). RNA was further treated with the TURBO DNA-Free Kit (Life Technologies) to more effectively remove DNA contamination. Ribosomal RNA was then removed using the MICROBExpress Bacterial mRNA Enrichment Kit (Thermo Fisher), and then concentrated using RNA Clean & Concentrator (Zymo Research). Each sample was run through a Fragment Analyzer (Advanced Analytical Technologies, Inc.) to provide the best estimate of RNA concentration and quality. Samples were sent for library

preparation and sequencing by Illumina HiSeq 1x151 at University of Delaware Sequencing & Genotyping Center (Newark, DE).

2.6 Metatranscriptomic Analysis

Metatranscriptomic reads were quality controlled using the previously described pipeline (www.githib.com/mooreryan/qc). Ribosomal RNA transcripts were removed using SortMeRNA v2.1 (Kopylova et al., 2012). Non-rRNA transcripts were then mapped to the thirty representative MAGs using Bowtie2 v2.3.2 (Langmead et al., 2012). Bedtools v2.26.0 was used to identify the raw counts of reads that mapped to each ORF (Quinlan & Hall, 2010). These raw counts were normalized to transcripts per million (TPM), using the following equation:

$$TPM = \frac{R_g \times rl \times 10^6}{fl \times T}$$

$$T = \sum_{g \in G} \frac{R_g \times rl}{fl_g}$$

Where R_g is the number of reads mapping to an ORF, *rl* is the read length, and *fl* is length of ORF in bases (Wagner et al., 2012).

2.7 Building the *Sulfurovum* Pangenomes and Defining Functional Clusters

Identification of genetic diversity within the *Sulfurovum* was done by building a concatenated single copy gene tree using the protein sequence of genes found in all chosen MAGs/genomes (Table A.1). These sequences were then aligned using

MUSCLE (Edgar, 2004), and then curated to eliminate large gaps and areas that did not align well, which resulted in a final alignment of 1,165 amino acids. RAxML (Stamatakis, 2006) was used to cluster this final alignment based on the maximum likelihood method, and ultimately used to infer phylogeny of MAGs. Genomes for all *Sulfurovum* MAGs were downloaded from NCBI (https://www.ncbi.nlm.nih.gov/genome/genomes/13561).

Pangenomic analysis was performed using the Anvi'o pangenomic workflow (http://merenlab.org/2016/11/08/pangenomics-v2/, Delmont & Eren, 2018) to compute the Pozzo dei Cristalli *Sulfurovum* pangenome. To strengthen the validity of gene annotation all *Sulfurovum* MAGs were annotated using the command 'anvi-run-ncbi-cogs', with in Anvi'o (Eren et al., 2015). After annotation the command "anvi-pan-genome" using the parameters '--minbit 0.5', and '--mcl-inflation 1.5', was used to perform the pangenomic analysis. Anvi'o was also used to interactively bin gene clusters into groups based on presence within a MAG/genome. Using an in-house script, genes were binned based on the functional annotation identified from the COG database.

Chapter 3

RESULTS AND DISCUSSION

3.1 Geochemistry of Pozzo di Cristalli Stream and Colonization Observations

Sulfurovum streamers abundant in sulfur are common in the sulfide rich waters of the Pozzo dei Cristalli (Hamilton et al., 2015). Ecological mechanisms by which these organisms produce and consume S(0) is still poorly understood. To survey transcriptional changes and genetic diversity of sulfur oxidizers within the Pozzo dei Cristalli, we set up colonization nets near the surface of the stream to provide the environmental niche for *Sulfurovum* streamer growth, high sulfide to oxygen ratios (Macalady et al., 2008) (Figure 2A). During colonization sulfide and oxygen concentrations remained stable, as well as, other measured geochemical parameters (Table 1).

Growth of streamers emerged on the nets by three days after installment. During the study, we observed increase in biomass of these streamers (Figure 2B). The streamer morphology on the nets is consistent with previously observed *Sulfurovum* mats (Hamilton et al., 2015). Samples for molecular analysis were collected over seventeen days from the three net locations (Table 2). By day 17 we noticed two distinct biofilm morphologies on net 3, both fluffy streamers and a crusty biofilm

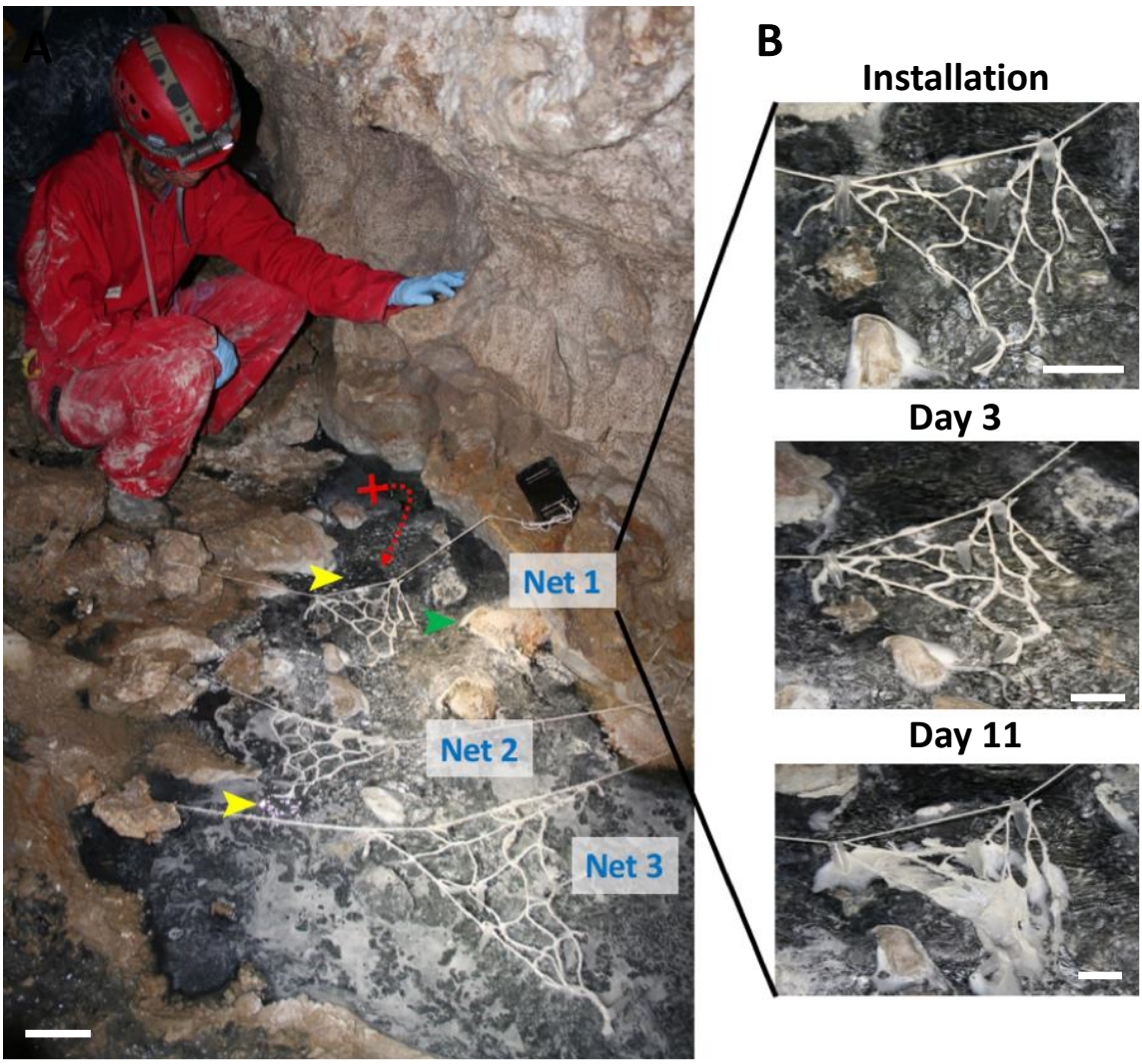

Figure 2 (A) Installation of colonization nets within the Pozzo di Cristalli stream. Red X indicates the stream source, and the dashed line represents the flow of the stream. Yellow arrows indicate the locations of *in situ* geochemistry measurements. Green arrow indicates location of natural streamer collected for comparison to biofilms that colonized nets. (B) Macroscopic observation of biofilm colonization of Net 1 over time. All scales bars are 8cm.

Table 1 *In situ* sulfide and oxygen concentrations during colonization. Measurements were made at stream source (Up) and between Net 2 and Net 3 (Down). Exact locations are indicated on Figure 1. At all sampling times pH and percent salinity was 7.1-7.2 and 1.3, respectively. Temperature remained stable throughout colonization at 13.7°C

Day	**0**	**3**		**15**		**17**	
Measurement Location	Up	Up	Down	Up	Down	Up	Down
$[S^{2-}]$ mM	NA	421	426	487	518	460	450
$[O_2]$ mM	12	10	16	13	18	17	19

(Figure 3). The location of these nets within the high sulfide microaerophilic waters of the Pozzo dei Cristalli supported growth of microbial communities whose morphology is suggestive of *Sulfurovum*, while also providing a spatially different niche for growth of a morphologically distinct crusty biofilm.

Table 2 Microbial community morphologies of samples collected for molecular analysis

Sample Name	**Net**	**Time Point**	**Morphology**	**Date**
PC16-4a	Net 1	Day 3	Streamer	11-July-2016
PC16-12	Net 1	Day 11	Streamer	19-July-2016
PC16-40a	Net 1	Day 15	Streamer	23-July-2016
PC16-51a	Net 1	Day 17	Streamer	25-July-2016
PC16-19	Net 2	Day 11	Streamer	19-July-2016
PC16-41a	Net 2.1	Day 4	Streamer	23-July-2016
PC16-55a	Net 2.1	Day 6	Streamer	25-July-2016
PC16-59a	Net 3	Fluffy	Streamer	25-July-2016
PC16-60b	Net 3	Crusty	Biofilm	25-July-2016
PC16-3a	Natural Streamer	Natural Streamer	Streamer	25-July-2016

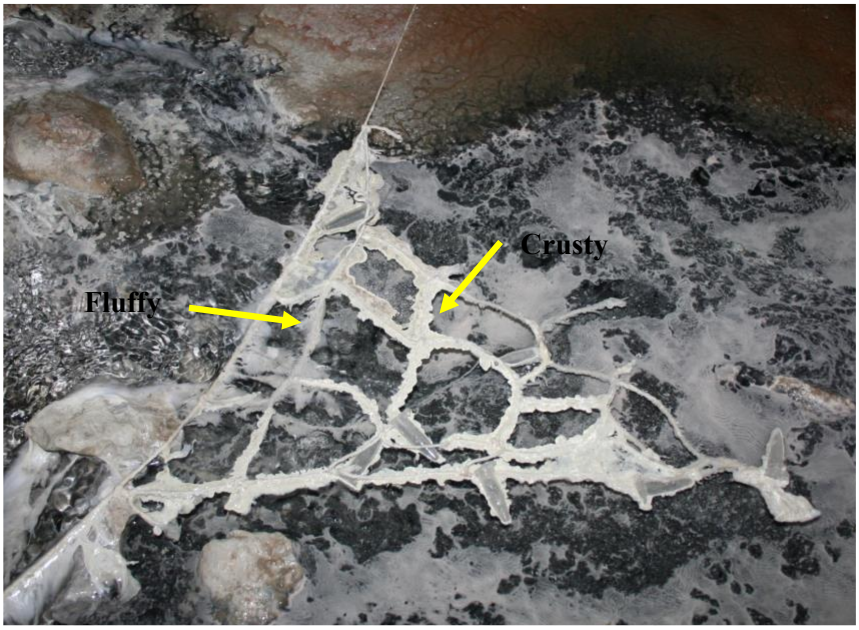

Figure 3 Biofilm morphologies observed on Net 3 Day 17.

3.2 Metagenomic Sample Selection

To get an initial view of the community structure and identify samples for metagenomics, we sequenced the 16S rRNA gene and then clustered samples based on OTU abundance (Figure 4A). The streamer community was simple, dominated by two OTUs, OTU 1 Gammaproteobacteria *Halothiobacillaceae* and OTU 2 Campylobacterota *Sulfurovum* (Figure 4B). The earliest samples collected at Net 1 Day 3 and Net 2.1 Day 4 clustered together, while the later samples for the fluffy morphology clustered together. The relative abundance of OTUs within the natural streamer was similar to that of samples from the later time points, suggesting that this streamer represents a more mature microbial community. The morphologically distinct

crusty biofilm and sediment samples were the most dissimilar, with higher abundance of OTU3 Gammaproteobacteria Halothiobacillaceae.

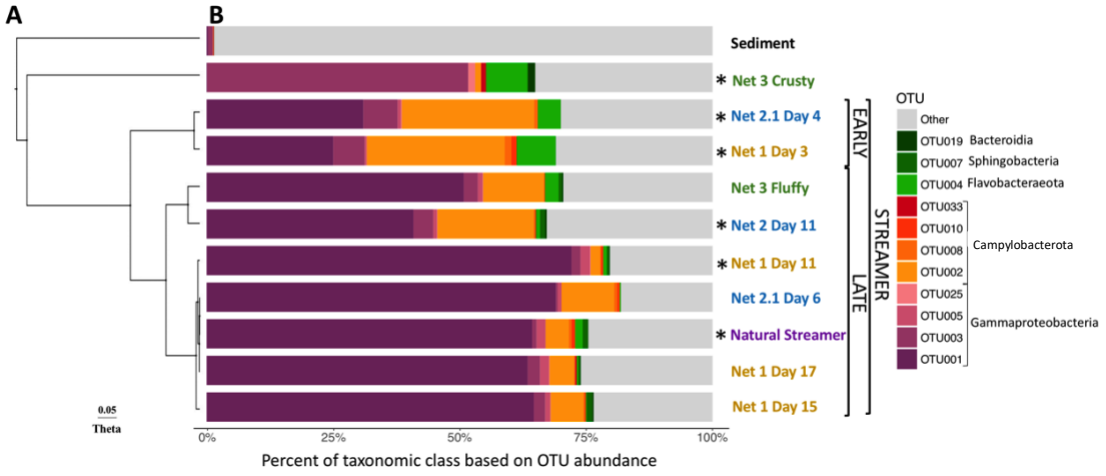

Figure 4 Microbial community structure based on the 16S sequences. (A) Tree based on the Yue & Clayton measure of dissimilatory between the structures of the biofilm communities. (B) Distribution of OTU and their taxonomic class. (*) indicates samples chosen for metagenomics.

Due to our interest in the *Sulfurovum*, we chose samples with the highest abundance to achieve adequate coverage for genome reconstruction (Figure 4B). To compare the natural community with the net colonizers, we chose to sequence the natural streamer. The crusty biofilm sample was also chosen to investigate the sulfur oxidizers, which form a distinct morphological community. A total of six samples were chosen for metagenomic investigation, including genome reconstruction for identification of taxon specific expression profiles and genetic diversity within this sulfur oxidizing community.

address issues that come from strain level heterogeneity commonly found in high abundant organisms (Dick et al., 2009). These assembly methods resulted in 754 – 56,378 contigs >2000bp and were binned by metagenomic sample (Table 4). After dereplication of MAGs, these contigs produced thirty high quality MAGs (>90% complete and <10% redundant) and were used as references for metatranscriptome analysis (Table 5) and the ones identified as *Sulfurovum* were used in the pangenome analysis.

Table 3 Metagenome sequencing statistics.

Timepoint	**Raw Reads**	**Quality Controlled Reads**
Net 3 Crusty	23,412,281	23,362,077
Net 1 Day 3	21,447,100	21,404,047
Net 2.1 Day 4	22,653,430	22,610,871
Natural Streamer	23,600,090	23,532,177
Net 2 Day 11	24,741,228	24,687,584
Net 1 Day 11	23,660,336	23,585,752
Total	139,514,465	139,182,508

Table 4 Statistics from the three assembly methods

Assembly Method	Time Point	Assembled QC Contigs	Max Contig Length (bp)	Contigs >2000bp	Percent of Reads Mappting to Contigs >2000bp
Single Library					
	Net 3 Crusty	304,236	175,452	13,559	78
	Net 1 Day 3	94,767	171,424	13,838	68
	Net 2.1 Day 4	147,784	238,274	11,184	70
	Natural Streamer	127,430	454,090	13,778	67
	Net 2 Day 11	153,537	187,108	6,822	60
	Net 1 Day 11	139,1974	155,360	30,162	69
Co-assembly					
	All	452,928	56,378	56,378	69
Sub-Assembly					
	Net 2 Day 11	4,077	42,742	754	62
Total		1,564,096	NA	146,475	NA

Across the ten metatranscriptome libraries, an average of 143 million quality-controlled transcripts per library were produced (Table 6). Of these reads, 18% to 52% were identified as non-rRNA, which suggests that rRNA reduction was successful as it is reported that 80% of the RNA within a microbial cell is ribosomal (Rosenow et al., 2001). The 30 MAGs chosen for investigation recruited 52% to 85% of these non-rRNA reads, suggesting adequate capture of expression within the microbial community.

Table 5 MAGs used for metatranscriptomic mapping. SL- single library SA – sub-assembly

Genome ID	Phylogenetic Assignment	Percent Complete	Percent Redundant	Percent G+C	Size (Mbp)	Assembly (Sample Name)	# of CDS
PC16-C1	Halothiobacillaceae	100	2.1	65.4	2.4	SL (PC16_4a)	2249
PC16-C2	Sulfurovum	96	2.5	33.9	1.9	SA (PC16_19)	1871
PC16-C3	Flavobacterium	99	1.3	31.2	2.6	SL (PC16_4a)	2412
PC16-C5	Fibrobacteres	99	3.3	45.0	3.8	SL (PC16_3a)	3505
PC16-C6	Sulfurovum	97	4.5	35.5	2.8	SL (PC16_41a)	2633
PC16-C7	Sulfurovum	99	2.1	39.9	1.9	SL (PC16_4a)	2113
PC16-C8	Sulfurovum	97	3.9	39.4	2.3	SL (PC16_19)	2214
PC16-C9	Sulfurovum	94	2.9	41.3	1.6	SL (PC16_41a)	1612
PC16-C10	Arcobacter	96	6.5	30.9	2.2	SL (PC16_4a)	2156
PC16-C11	Beggiatoa	92	0.6	46.9	2.1	SL (PC16_3a)	2003
PC16-C12	Piscirickettsiaceae	100	0.9	45.9	2.1	SL (PC16_41a)	2003
PC16-C13	Sulfurovum	98	0.4	37.8	2.0	SL (PC16_60b)	2027
PC16-C14	Sulfurovum	95	1.1	34.7	1.5	SL (PC16_3a)	1564
PC16-C15	Bacteroidia	99	0.5	38.7	2.8	SL (PC16_60b)	2344
PC16-C16	Sulfuricurvum	97	3.9	43.9	2.0	SL (PC16_3a)	2035
PC16-C17	Bacteroidia	93	1.7	37.5	3.2	SL (PC16_60b)	2500
PC16-C18	Halothiobacillaceae	93	8.9	46.8	2.1	SL (PC16_4a)	2068
PC16-C19	Thiothrix	99	0.7	51.3	3.5	SL (PC16_60b)	3464
PC16-C20	Sulfuricurvum	92	2.1	41.0	1.5	SL (PC16_12)	1632
PC16-C22	Paludibacter	96	0.5	37.7	2.5	SL (PC16_60b)	2092
PC16-C23	Sulfuricurvum	99	1.8	41.9	2.4	SL (PC16_12)	2565
PC16-C28	Cyanobacteria	91	0.9	27.7	3.7	SL (PC16_60b)	3344
PC16-C29	Helicobacteraceae	96	3.0	41.7	2.7	SL (PC16_19)	2922
PC16-C32	Bacteroidia	97	0.0	30.1	2.8	SL (PC16_60b)	2245
PC16-C33	Bacteroidia	99	0.5	37.3	3.0	SL (PC16_60b)	2449
PC16-C35	Halothiobacillaceae	96	0.6	53.0	2.1	SL (PC16_60b)	2078
PC16-C37	Piscirickettsiaceae	98	3.2	43.6	2.3	SL (PC16_60b)	2302
PC16-C39		91	2.3	66.7	3.6	SL (PC16_60b)	3360
PC16-C43	Myxococcales	92	2.5	65.9	4.5	SL (PC16_60b)	3758
PC16-C59	Flavobacteria	93	3.4	33.4	2.7	Co-Assembly	2528

Table 6 Statistics from metatranscriptome and mapping to MAGs.

Net and Day	Total Raw Reads in Millions	Total QC Reads in Millions	Non-rRNA Reads in Millions (% of Total QC Reads)	Non-rRNA Reads Mapping to MAGs in Millions (% of Non-rRNA Reads)
Natural Streamer	14.62	14.58	2.64 (18)	1.66 (67)
Net 3 Fluffy	12.98	12.95	4.99 (39)	3.67 (74)
Net 3 Crusty	16.25	16.21	5.36 (33)	2.79 (52)
Net 2.1 Day 4	13.48	13.45	6.97 (52)	5.86 (85)
Net 2.1 Day 6	14.45	14.41	4.38 (30)	2.96 (68)
Net 2 Day 11	14.73	14.69	4.03 (27)	2.52 (64)
Net 1 Day 3	13.76	13.73	4.49 (33)	3.23 (72)
Net 1 Day 11	12.69	12.66	3.38 (27)	2.07 (62)
Net 1 Day 15	14.26	14.23	5.27 (37)	3.54 (68)
Net 1 Day 17	15.83	15.79	4.92 (31)	3.38 (69)
Mean	14.305	14.27	4.64 (33)	3.17 (68)

3.2.1 Phylogeny of Community Composition and Metatranscriptome

To identify the abundance of particular organisms (represented by MAGs) during colonization, the community percentage based on coverage was calculated for each MAG in each sample's metagenome (Figure 5A). The streamer community was dominated by *Halothiobacillaceae* PC16-C1 and the *Sulfurovum* PC16-C2, which made up 77% to 89% of the community composition. Other sulfur oxidizing members of the community were *Campylobacterota*, including six other *Sulfurovum*, *Arcobacter* PC16-10, and *Sulfuricurvum* PC16-C16. The remaining portion of the streamer community, 2% to 3% consisted of heterotrophic organisms *Flavobacterium* PC16-C3 and *Fibrobacteres* PC16-C5 (Ransom-Jones et al., 2012; McBride & Nakane, 2015), which could indicate an important link between the primary producers and heterotrophic consumption. The community of the crusty biofilm was dominated by *Sulfurovum*, making up 54% of its composition. However, the distribution of *Sulfurovum* MAGs was different from the streamers

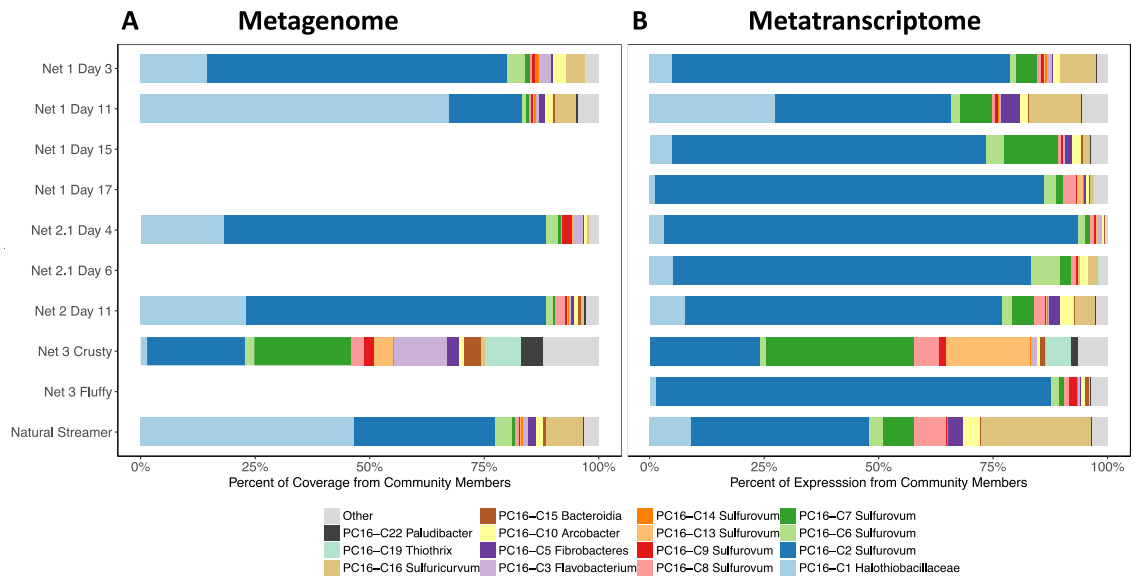

Figure 5 Coverage and community expression of the MAGs within the streamer samples. (A) Metagenomic coverage (B) Total expression of MAGs

with two other *Sulfurovum* (PC16-C7 and PC16-C13) at a higher abundance in the crusty biofilm.

To get a proxy for relative activity of organisms in the streamers, we recruited non-rRNA transcripts to the MAGs. The percent of transcripts mapping to the representative MAGs within the metatranscriptomes is generally similar to the taxonomic distribution that was observed in the metagenome. (Figure 5B). Activity in the streamer communities was dominated by *Sulfurovum* PC16-C2, while other major contributors were *Halothiobacillaceae* PC16-C1, *Sulfurovum* PC16-C6, and *Sulfuricurvum* PC16-C16. The natural streamer was distinguished by the high expression of *Sulfuricurvum* PC16-C16. Again, the morphologically distinct crusty

biofilm was dominated by the *Sulfurovum*, but there were three active *Sulfurovum* populations (PC16-C2, PC16-C7, and PC16-C13).

3.3 Metabolic Functions of the Community

In order to determine the biogeochemical roles of the entire microbial community, we analyzed the abundance of key genes within the metagenomes and metatranscriptomes (Figure 6). Genes involved in sulfur oxidation (*sqr*, *fccAB*, and *sox*) were among the highest expressed genes in the metatranscriptome. On average across all samples 70.6 % transcripts mapping to the *sqr* gene, responsible for sulfide oxidation to S(0), were affiliated to *Sulfurovum* PC16-C2 (Table 7). Based on the abundance of *Sulfurovum* PC16-C2 and the high expression of *sqr* by this population, we propose that *Sulfurovum* PC16-C2 is a major contributor of S(0) deposition within the streamer.

In contrast, genes involved in thiosulfate oxidation (*sox*) and *fccB* (the catalytic subunit) were found to be predominately affiliated with PC16-C1 *Halothiobacillaceae* PC16-C1 (Table 7). The reactions carried out by *sqr* and *fccAB* differ by providing electrons to quinone pool or to cytochrome c, respectively (Friedrich, 1998; Meyer & Cusanovich, 2003). Thus, oxidation of sulfide by *sqr*,

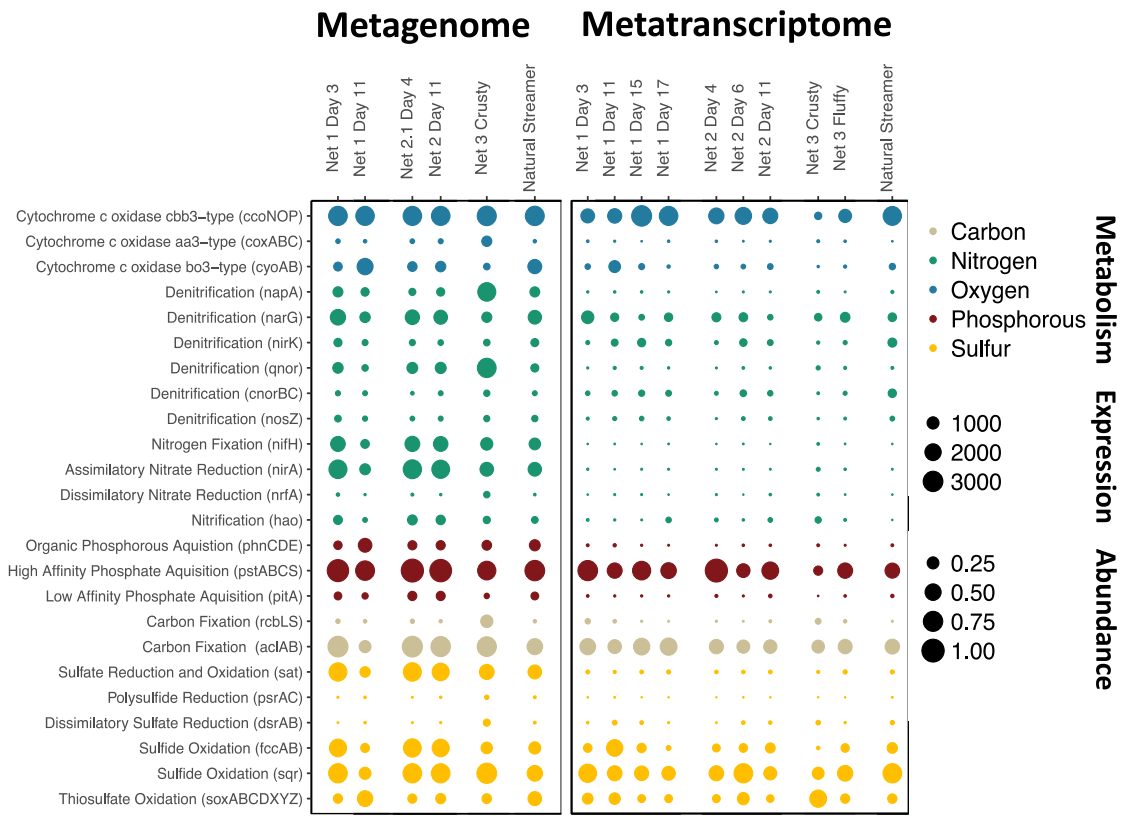

Figure 6 Normalized abundance and expression of pathways found within the streamer community for oxygen, nitrogen, phosphorous, carbon and sulfur metabolisms. Community is made up of the 30 MAGs. (A) Metagenomes were normalized by the single copy gene, DNA-directed RNA polymerase beta subunit gene (*rpoB*). Pathways were normalized to number of genes, and a minmax normalized abundance of this subset is plotted. (B) Metatranscriptome is plotted as expression (TPM), and pathways were normalized to number of genes. *ccoNOP* was divided by 10.

should in theory generate more energy than *fccAB* because proton motive force is generated at the b/c1 complex level through the reduction of quinone pool (Friedrich, 1998; Shutz et al., 1999; Friedrich et al, 2005). Expression of *fccA* in PC16-C1

Table 7 Percent of TPM coming from the dominant expressing organisms for sulfur oxidation genes

Gene	MAG	Net 1 Day 3	Net 1 Day 11	Net 1 Day 15	Net 1 Day 17	Net 2.1 Day 4	Net 2.1 Day 6	Net 2 Day 11	Crusty	Fluffy	Natural Streamer	Average
sqr												
	PC16-C2 Sulfurovum	77.9	40.9	83.1	84.5	96.0	88.0	65.6	39.4	94.7	36.3	70.6
soxA												
	PC16-C1 Halothiobacillaceae	24.2	71.2	91.4	32.1	84.4	84.7	78.6	0.1	59.5	68.5	59.5
soxB												
	PC16-C1 Halothiobacillaceae	42.9	76.0	93.3	38.8	90.7	83.8	90.5	21.6	51.3	75.4	66.4
soxC												
	PC16-C1 Halothiobacillaceae	58.7	97.0	80.7	8.3	62.2	82.4	59.6	1.3	40.4	87.5	57.8
soxD												
	PC16-C1 Halothiobacillaceae	33.3	91.5	78.8	72.9	88.9	91.4	77.0	1.7	32.0	87.8	65.5
sozX												
	PC16-C1 Halothiobacillaceae	6.2	85.2	92.3	57.1	77.1	81.8	82.6	0.0	32.6	84.4	59.9
soxY												
	PC16-C1 Halothiobacillaceae	6.7	68.6	63.1	8.2	46.1	86.8	29.7	0.6	14.9	53.5	37.8
soxZ												
	PC16-C1 Halothiobacillaceae	8.4	63.8	61.4	5.4	35.3	82.1	27.1	0.1	13.5	58.4	35.6
fccB												
	PC16-C1 Halothiobacillaceae	99.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.4	100.0	99.9

Halothiobacillaceae suggests this population is also depositing S(0). However, *fccA* is found in many organisms that contain the *sox* pathway and has been suggested as a sole route of sulfide detoxification, as opposed to solely energy production (Friedrich, 1998; Meyer & Cusanovich, 2003). However, the ability to obtain energy and detoxify, as well as its ability to catalyze a six-electron transfer (*soxCD*) during thiosulfate oxidation may contribute to its overall abundance within these streamer samples.

Table 8 Percent of TPM coming from the dominant expressing organisms for terminal oxidases

Gene	MAG	Net 1 Day 3	Net 1 Day 11	Net 1 Day 15	Net 1 Day 17	Net 2.1 Day 4	Net 2.1 Day 6	Net 2 Day 11	Crusty	Fluffy	Natural Streamer	Average
ccoN												
	PC16-C2 Sulfurovum	94.0	87.3	93.8	97.1	98.6	95.4	89.8	44.3	97.5	79.5	87.7
ccoO												
	PC16-C2 Sulfurovum	76.2	58.8	85.2	92.2	96.5	88.1	73.7	20.0	93.2	55.4	73.9
ccoP												
	PC16-C2 Sulfurovum	84.1	69.9	86.7	91.6	95.5	87.6	81.2	38.9	93.9	66.1	79.6
narG												
	PC16-C2 Sulfurovum	89.1	35.4	80.1	86.5	99.2	89.0	58.8	87.6	97.8	22.2	74.6

The oxidation of reduced sulfur compounds and transport of electrons to the respiratory chain link sulfur oxidation to aerobic respiration. In connection, genes encoding for the high-affinity terminal oxidase *ccoNOP* were consistently found to be in the top 10 highest expressed genes in all metatranscriptomes (Table A.2). Expression of *ccoNOP* was predominately affiliated with that of *Sulfurovum* PC16-C2 (Table 8), which suggest that this population is functioning under low-oxygen conditions. When oxygen is limiting, organisms are known to couple denitrification to sulfur oxidation (Cardoso et al., 2006). Genes for the complete denitrification pathway were found within all metagenomes. Compared to the other nitrogen genes, only expression of *narG* was prominent across all samples with 74.6% of transcripts affiliated to *Sulfurovum* PC16-C2. A member of the large family of molybdopterin oxidoreductases, *narG*, catalyzes the first step in denitrification, nitrate to nitrite, pumping two-protons from the quinol pool to the periplasm. The denitrification gene *narG* is only used by denitrifying bacteria that depend on the process for conservation

of energy (Chen & Strous, 2013). Expression of energy producing enzymes, *ccoNOP* and *narG,* by *Sulfurovum* PC16-C2 suggest that this population is carrying out sulfide oxidation coupled to both oxygen and nitrate. The metabolic versatility of *Sulfurovum* PC16-C2 allow this population to inhabit both microaerophilic and anaerobic niches and could be a reason for its dominance with the streamer biofilms.

Reducing equivalents generated from the respiratory chain can be used for carbon fixation pathways. Genes indicative of carbon assimilation by the reductive tricarboxylic acid (rTCA) and Calvin-Benson-Bassham (CBB) cycles were found in the metagenomes and expressed in the metatranscriptomes (Figure 6). ATP citrate lyase (*aclAB*) is indicative of the rTCA being used to fix carbon. All transcripts were affiliated to the Campylobacterota (Table 9). RuBisCo, *rbcL,* a key enzyme in the CBB cycle was found to be mainly expressed by the Halothiobacillaceae populations (PC16-C1 and PC16-C35). Campylobacterota are considered to be primary producers in sulfidic environments such as hydrothermal vents and sulfidic springs (Campbell et al., 2006). While the Campylobacterota may be playing a large role in primary production, the presence of RuBisCo expression by *Halothiobacillaceae* PC16-C1, *Piscirickettsiaceae* PC16-C37, and *Halothiobacillaceae* PC16-C35 suggest that populations other than the Campylobacterota are contributing to primary production within the streamers.

Table 9 Percent of TPM coming from the top three expressing MAGs for gene involved in carbon fixation.

Gene	MAG	Net 1 Day 3	Net 1 Day 11	Net 1 Day 15	Net 1 Day 17	Net 2.1 Day 4	Net 2.1 Day 6	Net 2 Day 11	Crusty	Fluffy	Natural Streamer	Average
aclA												
	PC16-C2 Sulfurovum	79.4	52.9	69.6	83.7	94.0	74.8	66.0	14.4	91.1	31.9	65.8
	PC16-C7 Sulfurovum	7.2	14.9	18.9	2.8	1.7	4.8	11.9	49.2	2.2	15.5	12.9
	PC16-C16 Sulfuricurvum	9.4	21.6	2.4	1.8	0.3	4.0	10.6	0.2	0.1	41.7	9.2
aclB												
	PC16-C2 Sulfurovum	78.6	59.0	78.3	86.3	94.4	75.8	77.0	23.1	90.7	47.9	71.1
	PC16-C7 Sulfurovum	5.6	13.7	11.0	2.2	1.6	5.0	6.3	38.0	1.6	8.4	9.3
	PC16-C16 Sulfuricurvum	12.0	18.3	2.2	1.1	0.2	5.5	8.1	0.1	0.1	34.8	8.2
rbcL												
	PC16-C1 Halothiobacillaceae	46.7	86.9	59.8	89.5	79.9	76.5	76.9	70.3	19.3	5.6	61.1
	PC16-C37 Piscirickettsiaceae	17.0	6.7	29.4	0.0	2.9	20.6	10.8	2.7	8.3	66.7	16.5
	PC16-C35 Halothiobacillaceae	6.9	2.0	7.6	8.8	6.4	0.0	3.9	25.6	35.6	5.5	10.2

We identified three phosphorous acquisition methods (*phnCDE, pstABCS,* and *pitA*) within the metagenome, suggesting a need to survive in a phosphorous limiting environment. The most highly expressed of those was the high affinity phosphate transporter, *pstABCS.* However, stability studies on *pst*S show mRNA stability ranging from 1.8 mins to 4.5 mins in *Escherichia coli* and *Bacillus subtilis,* respectively (Bernstein et al., 2002; Aguena et al., 2009). This may account for the relatively high expression of this transcript. Yet the expression of subunits *pst*ABC were also highly expressed (Figure A.1) and may suggests that phosphorous could be a limiting nutrient within the cave system, and a limiting factor in the rate at which S(0) is deposited.

3.4 Pangenomic Analysis

To determine the phylogeny of the seven *Sulfurovum* MAGs, a concatenated single copy gene tree with taxa representatives was built (Figure 7). The most abundant *Sulfurovum* (PC16-C2, PC16-C6, and PC16-C7) are each closely related to *Sulfurovum* MAGs from the Hamilton et al. metagenomes. Interestingly, at least one PC16 *Sulfurovum* was closely related to a *Sulfurovum* from a different environment. For example, *Sulfurovum* PC16-C8 and PC16-C9 were more closely related to *Sulfurovum* found within Mine wastewater than the *Sulfurovum* from the Pozzo dei Cristalli streamers. Our MAGs span the full spectra of phylogenetic diversity within the *Sulfurovum*, which may allow future work to determine if they are representative of their phylogenetic group.

The use of pangenomes in studying closely related organisms provides an understanding of core genes shared within the group, as well as strain specific genes that may lead to niche determination. A pangenomic analysis of the seven PC16 *Sulfurovum* resulted in 6917 gene clusters using the Anvi'o pangenomic workflow, with an mcl-inflation setting of 1.5 (Eren et al., 2015) (Figure 8A). In this study, a 'gene cluster' represents sequences of one or more predicted ORFs that grouped together based on their homology. There was a total of 544 (21% - 34% of a *Sulfurovum* genome) gene clusters that were found in the core genome (shared by all *Sulfurovum*), and a total of 747 (28% - 47% of a *Sulfurovum* genome) gene clusters shared by all but one *Sulfurovum.* The low percentage of core gene clusters within the

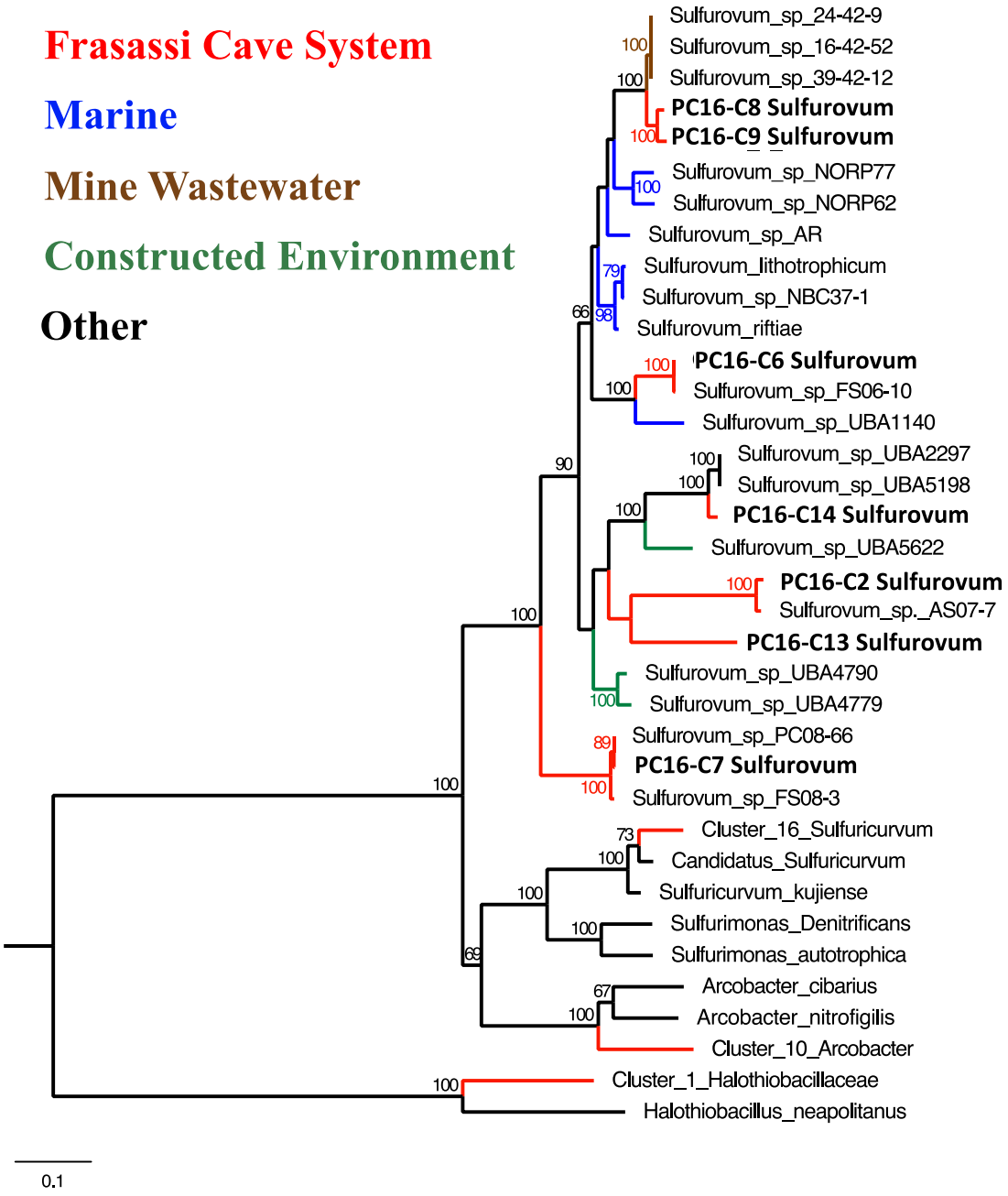

Figure 7 Maximum likelihood phylogenetic tree of 11 concatenated single-copy proteins showing the taxonomic placement of the *Campylobacterota* MAGs. Branch colors indicate environment form where genome or MAG was isolated.

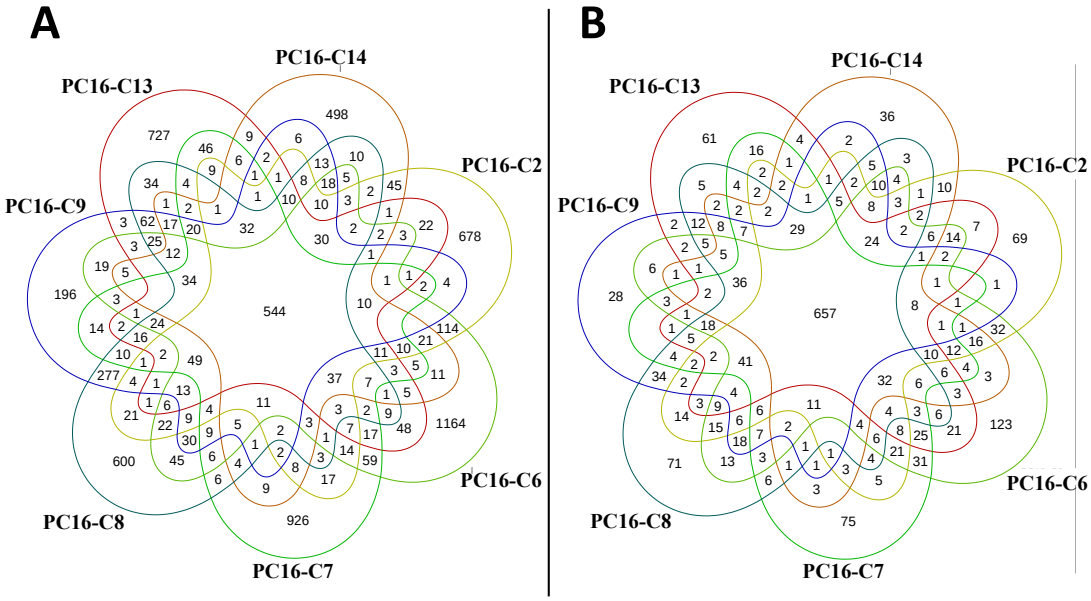

Figure 8 Venn diagram of gene clusters and functional annotations shared by the seven Sulfurovum Pozzo dei Cristalli Metagenomes. (A) Unique gene clusters (9703) identified in the Anvi'o pangenomic pipeline. (B) Functional Pangenome based on 1906 unique COG annotations.

PC16 *Sulfurovum*, is consistent with previous work on *Prochlorococcus* (10.4% of total gene clusters) and in *Pseudomonas* (15% of total gene clusters) (Mosquera-Rendon et al., 2016; Delmont & Eren, 2018). However, the low percentage of core gene clusters may be related to wide diversity of *Sulfurovum* we found within our streamers. The genetic variation within the *Sulfurovum* is high, but we expected that the functional annotation of these protein sequences would have less variation. would expect

To understand the functional diversity and adaptations within the PC16 *Sulfurovum*, genes were clustered based on unique gene functional annotations (COG-based), which we have termed a functional pangenome (Figure 8B). Out of 1904 unique functions 657 (30.5%) were found in the core functional genome, while 512 (26.9%) were MAG specific. However, some COGs did not correspond to the RAST annotation, which were used to look at community wide expression in section 3.4. For example, ORFs annotated by RAST as being *soxYZ* were placed in COGs as either "Uncharacterized Protein" or "Predicted secreted protein." The presence of *sox* genes within the *Sulfurovum* MAGs was confirmed by blasting these sequences (Table 10). Additionally, *sqr* was placed into the "Lipid transport and metabolism" category. It is known that annotation methods can have a significant impact of the results, which limited conclusions to be made from the functional pangenome.

Table 10 RAST verification of mis-annotated *sox* genes

RAST Annotation	COG	MAG	E-value	Percent Identifty	Organism
soxZ	Uncharacterized protein	PC16-C8	2.00E-53	84	Sulfurovum sp. 16-42-52
soxY	Predicted secreted protein	PC16-C8	1.00E-94	93	Sulfurovum sp. 35-42-20
soxY	Predicted secreted protein	PC16-C9	3.00E-92	91	Sulfurovum sp. 35-42-20
soxY	Predicted secreted protein	PC16-C13	6.00E-63	66	Sulfurovum sp. UBA12169
soxZ	Uncharacterized protein	PC16-C13	4.00E-43	67	Sulfurovum sp. AR

The incongruent annotations led to a lower resolution of genetic differences (functional annotation vs. genetic sequence) but more stringent analysis at the level of COG categories. Functional differences between the core genome and genes unique to the three dominant *Sulfurovum* (PC16-C2, PC16-C7, and PC16-C6), were reflected in the proportion of genes within COG categories (Table 11). Functional characterization of the *Sulfurovum* core genome indicated that while translation and cell wall biosynthesis categories are enriched in the core genome, a large proportion were annotated as Function unknown. The two COG categories cell wall biosynthesis and intracellular trafficking were found in higher proportion in *Sulfurovum* PC16-C2. Interestingly, the proportion of genes in carbohydrate transport and metabolism was much higher in *Sulfurovum* PC16-C6. We also observed a large proportion of *Sulfurovum* PC16-C7 strain specific genes to be involved in signal transduction and inorganic ion transport. The gene cluster and functional pangenomes could not alone provide enough information to identify the functional diversity within the PC16 *Sulfurovum*. Future work will expand the pangenomic analysis to include more *Sulfurovum* genomes and identify environmental or phylogenetic factors that affect the functional diversity of the *Sulfurovum*.

Table 11 Percent of COG Category found within each pangenome MAG

COG Category	All	Core	C2	C6	C7
[C] Energy production and conversion	5.42	5.09	5.19	2.27	7.87
[D] Cell cycle control, cell division, chromosome partitioning	1.12	0.92	3.90	2.27	0.00
[G] Carbohydrate transport and metabolism	4.86	1.74	2.60	13.64	4.49
[H] Coenzyme transport and metabolism	2.45	5.14	1.30	3.03	0.00
[I] Lipid transport and metabolism	2.37	1.95	0.00	3.79	1.12
[J] Translation, ribosomal structure and biogenesis	4.56	9.83	1.30	2.27	3.37
[K] Transcription	2.10	2.03	1.30	5.30	2.25
[L] Replication, recombination and repair	8.29	3.50	1.30	2.27	5.62
[M] Cell wall/membrane/envelope biogenesis	2.22	7.14	15.58	6.06	6.74
[N] Cell motility	4.39	0.40	1.30	2.27	0.00
[O] Post-translational modification, protein turnover, and chaperones	7.46	4.76	1.30	3.79	3.37
[P] Inorganic ion transport and metabolism	0.60	2.57	2.60	4.55	7.87
[Q] Secondary metabolites biosynthesis, transport, and catabolism	4.53	0.19	0.00	0.76	1.12
[R] General function prediction only	4.17	5.47	28.57	17.42	15.73
[S] Function unknown	0.55	33.99	6.49	10.61	7.87
[T] Signal transduction mechanisms	7.60	3.39	10.39	9.85	21.35
[V] Defense mechanisms	26.51	1.29	5.19	2.27	3.37
[X] Mobilome: prophages, transposons	4.28	0.34	3.90	7.58	3.37
[E] Amino acid transport and metabolism	2.00	5.43	1.30	0.00	2.25
[F] Nucleotide transport and metabolism	2.89	2.60	0.00	0.00	2.25
[U] Intracellular trafficking, secretion, and vesicular transport	0.21	1.95	5.19	0.00	0.00
[W] Extracellular structures	1.35	0.19	1.30	0.00	0.00
[Z] Cytoskeleton	0.06	0.09	0.00	0.00	0.00

3.5 Biogeochemical Niches of *Sulfurovum* and Other Dominant MAGs

To get a better understanding of the biogeochemical interplay between the *Sulfurovum* and other dominant organisms we did a targeted comparison of metabolisms involved in biogeochemical cycling. While the absence of key metabolic genes within our MAGs does not necessarily equate to their absences in the actual genome, general predictions can be made about the geochemical niche occupied by

each MAG. The metabolic diversity within dominant MAGs, was apparent when we compared carbon fixation pathways, nitrogen cycling, and sulfur oxidation genes (Table 12).

3.5.1 Carbon

The cycling of carbon in the Pozzo dei Cristalli streamers provides an important aspect of where these chemolithoautotrophic obtain carbon for biomass. Carbon fixation genes indicative of the rTCA and CBB cycle were found in the dominant classes of the streamer, which suggests that organic carbon levels are low, and majority of the community needs to fix CO_2 (Table 12). The rTCA cycle (~0.6 mol ATP/mol CO_2 for pyruvate) is energetically less expensive than the CBB cycle (~2.3 mol ATP/mol CO_2 for pyruvate), and the payoff greater with only 3-10 molecules of CO_2 being fixed per second by *rbcLS* (Ellis, 2010; Berg, 2011; Boyle & Morgan, 2011). The enzymatic efficiency of the rTCA may help to facilitate the success of the *Campylobacterota* within these streamers. Moreover, the rate at which the *Campylobacterota* are fixing carbon may be faster than that of *Halothiobacillaceae* PC16-C1. While, *Halothiobacillaceae* PC16-C1was seen at the early samples *Sulfurovum* PC16-C2 in theory is equipped to more quickly colonize the net and play a larger role in the

Table 12 Presence and absence of key metabolic genes in dominant MAGs. Bolded letters indicate presence of only one subunit.

Genes	Annotation	Metabolism	PC16 - C1	PC16 - C10	PC16 - C16	PC16 - C2	PC16 - C6	PC16 - C7	PC16 - C8	PC16 - C9	PC16 - C13	PC16 - C14	NBC37 - 1	AR	Litotrophicum
korABCD	2-oxoglutarate oxidoreductase	Carbon												**B**	
sdhAB	succinate dehydrogenase	Carbon													
aclAB	ATP-citrate lyase	Carbon													
rbcLS	Ribulose bisphotphate carboylase	Carbon													
actP	Cation/acetate symporter	Carbon													
acsA	Acetyl-coenzyme A synthetase	Carbon													
ccoNOP	Cytochrome c oxidase cbb3	Oxygen													
coxABC	Cytochrome c oxidase aa3	Oxygen													
cycAB	Cytochrome c oxidase bo3	Oxygen													
nifH	Nitrogenase reductase	Nitrogen													
napA	Periplasmic nitrate reductase	Nitrogen													
narGH	Respiratory nitrate reductase	Nitrogen													
nirK	Cytochrome cd1 nitrite reductase	Nitrogen													
cnorBC	Nitric-oxide reductase	Nitrogen					**B**								
qnor	Nitric-oxide reductase quinol-dependent	Nitrogen													
nirA	Ferredoxin--nitrite reductase	Nitrogen													
nozZ	Nitrous-oxide reductase	Nitrogen													
hao	Cytochrome p460	Nitrogen													
phnCDE	Phosphonate ABC transporter	Phosphorous													
pstABCS	Phosphate transport system	Phosphorous													
fccAB	sulfide dehydrogenase flavocytochrome C	Sulfur				**B**									
sqr	Sulfide:quinone oxidoreductase	Sulfur													
soxAX	c-type cytochrome	Sulfur												**X**	
soxYZ	sulfur-binding protein	Sulfur													
soxB	Thiol sulfate esterase	Sulfur													
soxCD	Sulfur dehydrogenase	Sulfur													

primary production of organic carbon within the biofilm.

However, the reducing carboxylases in the rTCA are oxygen sensitive because they contain iron-sulfur clusters, while *rbcLS* has been proposed to be oxygen insensitive (Kletzin & Adams, 1996; Berg, 2011). As a result, the oxygen sensitivity of these carbon fixation genes could be influencing the niches for primary production. PC16-C1 *Halothiobacillaceae* PC16-C1, may be acting as a primary producer in more aerobic regions of the next while *Campylobacterota* in more microaerophilic or anaerobic regions.

While all MAGs chosen for targeted exploration contain potential to fix their own carbon, we did fine the genetic potential for acetate assimilation in three of the *Campylobacterota* (PC16-C6, PC16-C14, and PC16-C16). The ability to directly assimilate acetate requires both *actP,* an acetate import system, and *acsA,* responsible for acetate activation (Gimenez et al., 2003). The potential for direct acetate assimilation in only a portion of these *Campylobacterota*, suggests that these flanking community members are equipped for a mixotrophic lifestyle, which has been proposed in *Sulfurovum* by Campbell et al., (2006).

3.5.2 Oxygen

The presence and expression of terminal oxidases can be used to determine the geochemical niche for which they are adapted. All MAGs contain the genes for *ccoNOP,* which encode for subunits of the cbb3 cytochrome c oxidase (Table 12). The

gene products of *ccoNOP* are components of the high affinity cytochrome c oxidase and used when oxygen concentrations are very low (Xie et al., 2014). The prevalence of *ccoNOP* within all the major MAGs is consistent with the location of our nets (air-water interface) and low oxygen concentrations for S(0) deposition.

However, *Halothiobacillaceae* PC16-C1, and two *Sulfurovum* (PC16-C8 and PC16-C9) have the genetic potential for other types of cytochrome c oxidases, *cyoAB* and *coxABC*, respectively.

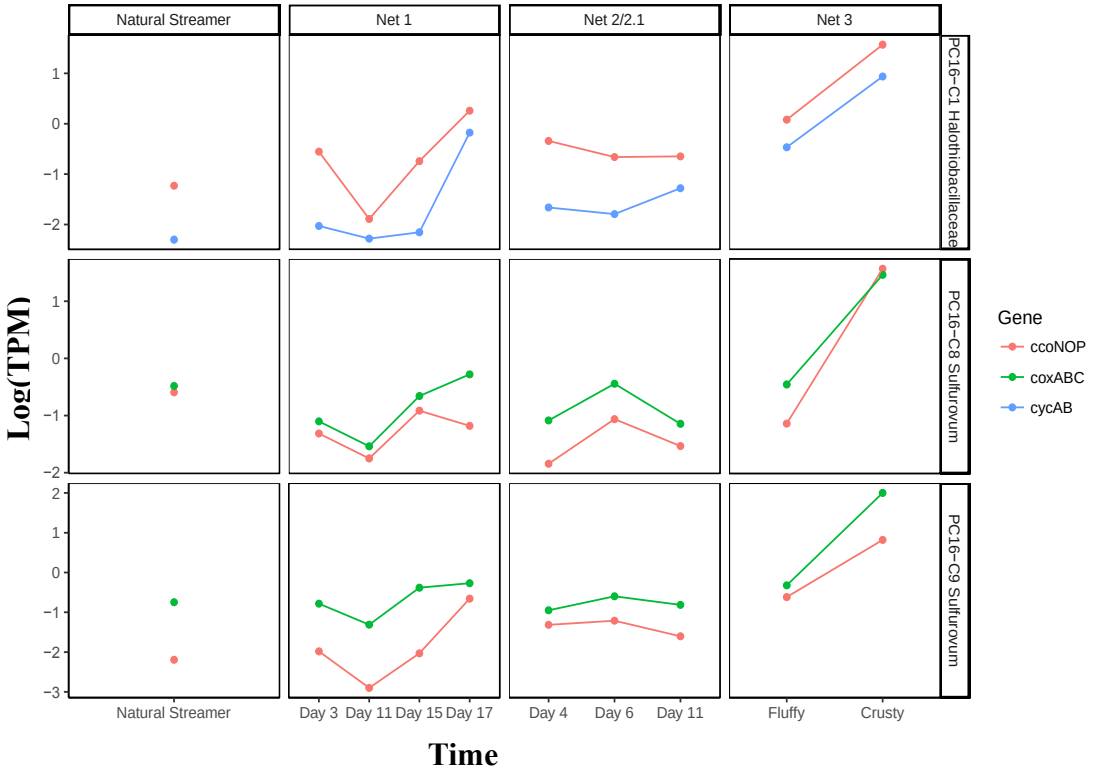

Figure 9 Log transformed TPM of terminal cytochrome c oxidases within the Halothiobacillaceae PC16-C1 and the two Sulfurovum (PC16-C8 and PC16-C9)

The metatratranscriptome allowed us to identify a preference in expression for the cytochrome c oxidase within these three microbial populations (Figure 9). These two *Sulfurovum* populations are expressing *coxABC* at a higher level, while *Halothiobacillaceae* PC16-C1 is expressing *ccoNOP* at a higher level. The aa3-type cytochrome (*cox*) has a low affinity for oxygen, but maximizes proton translocation, as supported by work done in *Bradyrhizhombium japonicum, Rhodobacter sphaeroides,* and *Bacillus subtilis* (Flory & Donohue, 1997; Winstedt & von Wachenfeldt, 2000; Arai et al., 2008; Kawakami et al., 2010). The oxidase *cyoAB* also has low affinity for oxygen and shows resistance to nitric-oxides (D'Mello et al., 1995; Kawakami et al., 2010). In conjunction, with the expression of genes encoding for these cytochrome oxidases of *Sulfurovum* PC16-C8 and PC16-C9 may be inhabiting regions of the biofilm where levels of oxygen are higher, and levels of sulfide are lower, which result in these two populations playing a smaller role in S(0) deposition.

3.5.3 Nitrogen

The coupling of nitrate to sulfide oxidation with production of S(0) is a proposed metabolism within the Frasassi cave system (Hamilton et al., 2015). To understand the community members that have potential to reduce nitrate, denitrify the biofilm, and fix nitrogen, we invested these genes within the dominant organisms (Table 12). The genetic potential for nitrogen cycling was found mainly in *Campylobacterota,* while there were no genes identified in *Halothiobacillaceae* PC16-

C1. To understand if the *Campylobacterota* are dominating cycling of nitrogen within the biofilm we compared expression to that of other MAGs at the class level (Table 13). For each transformation of nitrogen, the *Campylobacterota* are the dominant expressers of genes that have potential to carry out these reactions. We hypothesize that the *Campylobacterota* are major contributors to denitrification within the streamers. This further suggests that anaerobic S(0) deposition within the streamers may primarily be carried out by the *Campylobacterota*.

Table 13 Percent of expression by class expressing genes indicative of nitrogen cycling.

Gene	Reaction	Metabolism	Class	Percent of average TPM contribution
narG	$NO_3 \rightarrow NO_2$	Nitrate Reduction		
			Campylobacterota	100.00
napA	$NO_3 \rightarrow NO_2$	Nitrate Reduction		
			Campylobacterota	99.53
			Gammaproteobacteria	0.47
nirK	$NO_2 \rightarrow NO$	Denitrification		
			Campylobacterota	96.02
			Gammaproteobacteria	3.98
norBC/qnor	$NO \rightarrow N_2O$	Denitrification		
			Campylobacterota	96.24
			Gammaproteobacteria	3.76
nosZ	$N_2O \rightarrow N_2$	Denitrification		
			Campylobacterota	87.16
			Gammaproteobacteria	12.74
			Flavobacteriia	0.10
nifH	$N_2 \rightarrow NH_3$	Nitrogen Assimilation		
			Campylobacterota	95.59
			Gammaproteobacteria	0.14
			Fibrobacteres	4.26

Nitrate levels in the Pozzo dei Cristalli can be found at very low concentrations (Macalady et al., 2006; Jones et al., 2010; Hamilton et al., 2015). Although nitrate levels may be low, expression of *narGH* by *Sulfurovum* PC16-C2 suggests that it may still be using nitrate as a source of energy. The high abundance of *Sulfurovum* PC16-C2 and low levels of nitrate, may suggest that this population can rapidly convert nitrate to nitrite. The production of nitrite by *Sulfurovum* PC16-C2 may be providing niches for further denitrification that are occupied by the other *Campylobacterota*. During denitrification oxidation of nitrous oxide can contribute to energy production when *qnor* is present (Chen & Strous, 2013). Expression of *qnor* in *Sulfurovum* PC16-C7 suggests that this population could be occupying this energy generating niche while also participating in denitrification within the streamer. To summarize, nitrogen affiliated niches within the steamers are occupied by *Campylobacterota* and may be a key driver in the source of diversity found within the *Sulfurovum.*

3.5.4 Sulfur

The streamer community as a whole is primarily using reduced sulfur compounds as electron donors, yet the distribution of sulfur oxidizing genes within the dominant members provides an insight into the players of S(0) accumulation and consumption (Table 12). Based on the presence of the complete *sox* pathway found in the MAGs of *Halothiobacillaceae* PC16-C1 and *Sulfuricurvum* PC16-C16 we proposed that these two populations can carry out the complete oxidation of thiosulfate to sulfate (Figure 10). The remaining *Sulfurovum* and *Arcobacter* PC16-

C10 all contain the genetic potential to deposit S(0), indicated by presence of *sqr*, *fccB*, and incomplete *sox* pathways.

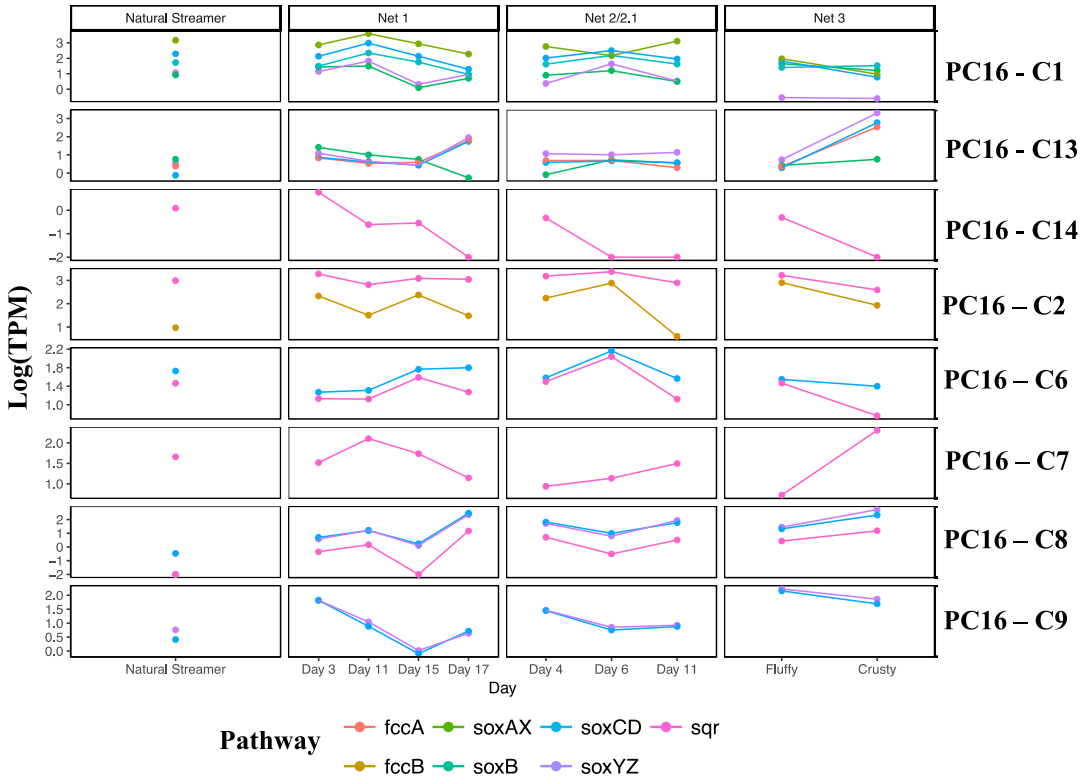

Figure 10 Log transformed TPM of sulfur oxidation genes with the *Sulfurovum* and *Halothiobacillaceae.*

Distribution of *sox* genes within the *Sulfurovum* prompted further exploration of their expression during colonization (Figure 10). In particular, *Sulfurovum* PC16-C6, PC16-C8, and PC16-C13 are expressing *sqr* at the same level or less than that of *sox* genes. From current literature, *in situ* assays have shown that sulfide or S(0) can be oxidized if either *soxAX* or *soxB* is missing, but at a slower rate (Rother et al., 2001). It has also been reported that *C. tepidum soxB* mutant does not secrete S(0) globules and was able

to completely oxidize sulfide to sulfate (Azai et al., 2009). This may suggest that the *Sulfurovum* expressing components of the *sox* pathway may not be playing a role in the deposition of S(0) within the streamers. The significance of which suggests that *Sulfurovum* PC16-C2 may be dominating S(0) within these streamers.

The dominance of S(0) by *Sulfurovum* PC16-C2 could also be playing a role in the diversity of *Sulfurovum* observed. Previous work by, Pjevac et al., (2014), showed that sulfur oxidizers *Sulfurimonas denitrificans* and *Allochromatium vinosum* with the same genetic potential, *soxABXYZ,* had very different growth patterns when cyclooctasulfur was the sole electron source. Further, microaerophilic waters are known to have different confirmations of S(0) (cyclooctasulfur, polymeric sulfur or different length polysulfides) (Holmvist et al., 2011; Lichtschlag et al., 2013). While our *Sulfurovum* lack the presence of *soxB*, a key enzyme for thiosulfate oxidation, the ability to access different polysulfides as electron donors could be a source of niche partitioning within the *Sulfurovum*.

3.6 Significance and Conclusion

The conspicuous white streamers within the Pozzo dei Cristalli streamer host a complex microbial community, that is sustained by an ecologically successful population of *Sulfurovum* PC16-C2. As the predominant microbial population within the streamer, they are biogeochemically important within the Pozzo dei Cristalli as well as, the drivers of the ecosystem.

As the streamer architects they appear to be the most dominant primary producers fixing carbon through the rTCA cycle, as indicated by expression of *aclAB*. With a bulk of the organic carbon within the streamer produced by *Sulfurovum* PC16-C2, they provide niches that are inhabited by heterotrophic microbes that can survive in this extreme environment. This syntrophic relationship with *Sulfurovum* PC16-C2 and the heterotrophs links back to three other populations of the *Campylobacterota* that have potential to assimilate acetate. The functional role of *Sulfurovum* PC16-C2 within the carbon cycle of the Pozzo dei Cristalli appears to drive both heterotrophic and mixotrophic lifestyles within the streamers.

Nitrogen cycling within microbial communities has the potential to structure them, as well as control productivity on the ecosystem scale. *Sulfurovum* PC16-C2 again functions as a conductor for the reduction of nitrate into nitrite, the first step in denitrification. Expression of genes involved in denitrification, and eventually nitrogen assimilation, are almost exclusively affiliated to the *Campylobacterota.* It appears denitrification is playing an important role in the diversity we see among the *Campylobacterota.* This phylum is driving nitrogen detoxification and production of biologically accessible nitrogen within the streamers and could be a source dependent microbial partners within the streamer. This beneficial microbial interaction between Sulfurovum PC16-C2, and the denitrifying and assimilating members of the *Campylobacterota*, provide an insight into the mutualistic metabolism if the dominant organism and the flanking community members.

Sulfide and oxygen concentrations within the waters of the Pozzo dei Cristalli have previously been shown to act as drivers of niche differentiation among these sulfur oxidizing microbes (Macalday et al., 2008). We have been able to provide further evidence that forms of S(0) within these streamers is also playing a role in the ability of organisms to colonize. With the highest expression of genes indicative of S(0) as a metabolic intermediate, *Sulfurovum* PC16-C2 again plays a functional role in both the deposition of S(0) within the streamer, as well as, potentially playing a role in who can consume S(0). Within the *Sulfurovum* there was a lot of diversity within the metabolic potential and expression of genes involved in the sox pathway. To date the exact substrate of the sox pathway has not been well characterized, yet there is evidence that supports its ability to oxidize sulfide, thiosulfate, S(0), and sulfite *in vitro* (Rother et al., 2001). The metabolic potential and actual sulfur substrates within these flanking *Sulfurovum* members are still unknown, yet their ability to access and utilize different sulfur compounds may be playing a role in their diversity. Further work should focus on biochemically characterizing the sox pathway in a broader range of organism, especially in the presences of extracellular S(0). Overall, we have shown that *Sulfurovum* PC16-C2 population drives the deposition of S(0) and possibly ecosystem with regards to sulfur accessible within the streamer. within the cave, in all areas of carbon, nitrogen, and in sulfur.

Our data indicate Pozzo de Cristalli Sulfurovum PC16-C2 population are potentially catalyzing S(0) precipitation using both oxygen and nitrate as a terminal electron acceptor. They are coupling this energy production to a lithoautotrophic

lifestyle that sustains streamer community in organic carbon production, nitrite concentrations, and possible accessibility of S(0).

We were able to capture a wide range of genetic diversity within the *Sulfurovum* population. In an evolutionary context this diversity is indicative of *Sulfurovum* found in geochemically diverse environments. Further pangenomic work should be done to explore the genetic similarities between the Pozzo dei Cristalli *Sulfurovum* and *Sulfurovum* from marine and other terrestrial environments. Identifying the common genetic traits within the *Sulfurovum,* as well as, environment specific genetic traits, will valuable insight into how the *Sulfurovum* are ecologically success and the importance they in biogeochemical cycling on planet earth.

REFERENCES

Aguena, M., Ferreira, G.M., & Spira, B. (2009). Stability of the *pstS* transcript of *Escherichia coli.* Arch. Microbiol. 191. 105-112.

Arai, H., Roh, J.H., & Kaplan, S. (2008). Transcriptome dynamics during the transition from anaerobic photosynthesis to aerobic respiration in *Rhodobacter sphaeroides* 2.4.1. J. Bacteriol. 190. 286–299

Azai, C., Tsukatani, Y., Harada, J., & Oh-oka, H. (2009). Sulfur oxidation in mutants of the photosynthetic green sulfur bacterium *Chlorobium tepidum* devoid of cytochrome c-554 and SoxB. Photosynth. Res. 100. 57-65.

Berg, I.A. (2011). Ecological aspects of the distribution of different autotrophic CO2 fixation pathways. Appl Environ Microbiol. 77:6. 1925-1936.

Boyle, N.R. & Morgan, J.A. (2011). Computation of metabolic fluxes and efficiencies for biological carbon dioxide fixation. Metab Eng. 13:2. 150-158.

Campbell, B. J., Engel, A. S., Porter, M. L., & Takai, K. (2006). The versatile ε-proteobacteria: key players in sulphidic habitats. Nature Reviews Microbiology, 4(6), 458–468. http://doi.org/10.1038/nrmicro1414

Campbell, B. J., Stein, J. L., & Cary, S. C. (2003). Evidence of Chemolithoautotrophy in the Bacterial Community Associated with Alvinella pompejana, a Hydrothermal Vent Polychaete. Applied and Environmental Microbiology, 69(9), 5070–5078. http://doi.org/10.1128/AEM.69.9.5070-5078.2003

Campbell, B.J., Engel, A.S., Porter, M.L. & Takai, K. (2006). The versatile epsilon-proteobacteria: key players in sulphidic habitats. Nat Rev Microbiol 4: 458–468.

Canfield, D. E. (2004). The evolution of the Earth surface sulfur reservoir. American Journal of Science, 304(10), 839–861. http://doi.org/10.2475/ajs.304.10.839

Cardoso, R.B., Sierra-Alvares, R., Rowlette, P., Flores, E.R., Gomes, J., & Field, J.A. (2006). Sulfide oxidation under chemolithoautotrophic denitrifying conditions. Biotechnol Bioeng. 95:6. 1148-1157.

Chen, J., & Strous, M. (2013). Denitrification and aerobic respiration, hybrid electron transport chains and co-evolution. Biochimica et Biophysica Acta. 1827. 136-144.

D'Mello, R., Hill, S., Poole, R.K. (1995). The oxygen affinity of cytochrome bo' in *Escherichia coli* determined by the deoxygenation of oxyleghemoglobin and oxymyoglobin: Km values for oxygen are in the submicromolar range. J. Bacteriol. 177 867–870.

Dahl, C. (2017). Sulfur Metabolism in Phototrophic Bacteria. In Modern Topics in the Phototrophic Prokaryotes (Vol. 183, pp. 27–66). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-51365-2_2

Darling, A.E., Jospin, G., Lowe E., Matsen IV, F.A., Bik, H.M., & Eisen, J.A. (2014). PhyloSift: phylogenetic analysis of genomes and metagenomes. PeerJ. 2:e243.

Delmont, T.O. & Eren, A.M. (2018). Linking pangenomes and metagenomes: the *Prochlorococcus* metapangenome. PeerJ 6:e4320.

Dick, G.J., Andersson, A.F., Baker, B.J., Simmons, S.L., Thomas, B.C., Yelton, A.P., & Banfield, J.F. (2009). Community-wide analysis of microbial genome sequence signatures. Genome Biol. 10:8. R85

Eddie, B. J., & Hanson, T. E. (2013). Chlorobaculum tepidum TLS displays a complex transcriptional response to sulfide addition. Journal of Bacteriology, 195(2), 399–408.

Edgar, C., R. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:6. 1792-1797.

Ehrlich, H. L. (2002). Geomicrobiology 4th Edition. Marcel Dekker Inc. New York.

Ellis, R.J. (2010). Biochemistry: Tracking unintelligent design. Nature. 463. 164-165.

Embley, R. W., Baker, E. T., Butterfield, D. A., Chadwick, W. W., Lupton, J. E., Resing, J. A., et al. (2007). Exploring the Submarine Ring of Fire. Oceanography, 20(4), 68–70.

Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., Delmont, T.O. (2015) Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ,* 3:e131

Flores, G. E., Campbell, J. H., Kirshtein, J. D., Meneghin, J., Podar, M., Steinberg, J. I., et al. (2011). Microbial community structure of hydrothermal deposits from geochemically different vent fields along the Mid-Atlantic Ridge. Environ. Microbiol. 13, 2158–2171.

Flory, J.E., Donohue T.J. (1997). Transcriptional control of several aerobically induced cytochrome structural genes in Rhodobacter sphaeroides. Microbiology 143. 3101–3110.

Fox, G.E., Wisotzkey, J.D., & Jurtshuck, P. (1992). How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. Int. J. Syst. Bacteriol, 42, 166-170.

Friedrich, C. G. (1998). Physiology and genetics of sulfur-oxidizing bacteria. Advances in Microbial Physiology, 39. 235-289.

Friedrich, C. G., Bardischewsky, F., Rother, D., Quentmeier, A., & Fischer, J. (2005). Prokaryotic sulfur oxidation. Current Opinion in Microbiology, 8(3), 253–259.

Frigaard, N.-U., & Dahl, C. (2008). Sulfur Metabolism in Phototrophic Sulfur Bacteria (Vol. 54, pp. 103–200). Elsevier. http://doi.org/10.1016/S0065-2911(08)00002-7

Fuchs, T., Huber, H., Burggraf, S., & Stetter, K. O. (1996). 16S rDNA-based Phylogeny of the Archaeal Order Sulfolobales and Reclassification of Desulfurolobus ambivalens as Acidianus ambivalens comb. nov. Systematic and Applied Microbiology, 19(1), 56–60. http://doi.org/10.1016/S0723-2020(96)80009-9

Galdenzi, S., Cocchioni, M., Morichetti, L., Amici, V., & Scuri, S. (2008). Sulfidic ground-water chemistry in the Frasassi Caves, Italy. Journal of Cave and Karst Studies, 7(2), 94–107.

Ghosh, W., & Dam, B. (2009). Biochemistry and molecular biology of lithotrophic sulfur oxidation by taxonomically and ecologically diverse bacteria and archaea. FEMS Microbiology Reviews, 33(6), 999–1043. http://doi.org/10.1111/j.1574-6976.2009.00187.x

Gimenez, R., Nuñez, M.F., Badia, J., Aguilar, J., & Baldoma, L. (2003). The gene yjcG, co-transcribed with the gene acs, encodes an acetate permease in Escherichia coli. 185:21. 6448-6455.

Hamilton, T. L., Jones, D. S., Schaperdoth, I., & Macalady, J. (2015). Metagenomic insights into S(0) precipitation in a terrestrial subsurface lithoautotrophic ecosystem. Frontiers in Microbiology, 1–16. http://doi.org/10.3389/fmicb.2014.00756/abstract

Hayes, J. M., & Waldbauer, J. R. (2006). The carbon cycle and associated redox processes through time. Philosophical Transactions of the Royal Society B: Biological Sciences, 361(1470), 931–950. http://doi.org/10.1098/rstb.2006.1840

Holmkvist, L., Ferdelman, T.G., and Jørgensen, B.B. (2011). A cryptic sulfur cycle driven by iron in the methane zone of marine sediment (Aarhus Bay, Denmark). Geochim Cosmochim Acta 75. 3581–3599.

Howarth, R.W., & Jorgensen, B.B. (2013) Formation of 35S-lablled elemental sulfur and pyrite in coastal marine sediments during short term sulfate reduction measurements. Geochim Cosmochim Acta. 48. 1807-1818.

Huber, J. A., Cantin, H. V., Huse, S. M., Mark Welch, D. B., Sogin, M. L., and Butterfield, D. A. (2010). Isolated communities of Epsilonproteobacteria in hydrothermal vent fluids of the Mariana Arc seamounts. FEMS Microbiol. Ecol. 73, 538–549.

Hyatt, D., Chen, G., LoCscio, P.F., Land, M.L., Larimer, F.W., & Hauser, L.J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 11:119.

Imhoff, J. F. (2001) The anoxigenic phototrophic purple bacteria. Bergey's Manual of Systematic Bacteriology, Vol. 1 (Boone DR, Castenholz RW & Garrity GM, eds), pp. 631–637. Springer, New York.

Imhoff, J.F. (1995) Taxonomy and physiology of phototrophic purple bacteria and green sulfur bacteria. Anoxygenic Photosynthetic Bacteria (Blankenship RE, Madigan MT & Bauer CE, eds), pp. 1–15. Kluwer Academic Publishers, Dordrecht.

Ivanov, M. V. (1968) Microbial Processes in the Formation of Sulfur Deposits.

Jannasch, H. W., Wirsen, C. O., & Molyneaux, S. J. (1991). Chemoautotrophic sulfur-oxidizing bacteria from the Black Sea. Deep Sea Research Part a, Oceanographic Research Papers, 38(Supplement 2), S1105–S1120.

Johnson, D. B., Joulian, C., d'Hugues, P., & Hallberg, K. B. (2008). Sulfobacillus benefaciens sp. nov., an acidophilic facultative anaerobic Firmicute isolated from mineral bioleaching operations. Extremophiles, 12(6), 789–798. http://doi.org/10.1007/s00792-008-0184-4

Jones, D. S., Lyon, E. H., & Macalady, J. L. (2008). Geomicrobiology of biovermiculations from the Frasassi Cave System, Italy. Journal of Cave and Karst Studies, 7(2), 78–93.

Kawakami, T., Kuroki, M., Ishii, M., Igarashi, Y., & Arai H. (2010). Differential expression of multiple terminal oxidases for aerobic respiration in *Pseudomonas aeruginosa*. Environ. Microbiol. 12. 1399–1412

Kletzin, A. & Adams, M.W. (1996). Molecular and phylogenetic characterization of pyruvate and 2-ketoisovalerate ferredoxin oxidoreductases from pyrococcus furiosus and pyruvate ferredoxin oxidoreductases from Thermatoga maritima. J Bacteriol. 178:1. 248-257.

Klingenberg, H., & Meinicke, P. (2017). How to normalize metatranscriptomic count data for differential expression analysis. PeerJ, 5(10), e3859–29. http://doi.org/10.7717/peerj.3859

Kopylova, E., Noé, L., Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data, *Bioinformatics*. 28:24, 3211–3217

Kostanjevecki, V., Brige, A., Meyer, T. E., Cusanovich, M. A., Guisez, Y., & Van Beeumen, J. (2000). A Membrane-Bound Flavocytochrome c-Sulfide Dehydrogenase from the Purple Phototrophic Sulfur Bacterium Ectothiorhodospira vacuolata. Journal of Bacteriology, 182(11), 3097–3103. http://doi.org/10.1128/JB.182.11.3097-3103.2000

Lavik, G., Stührmann, T., Brüchert, V., Van der Plas, A., Mohrholz, V., Lam, P., et al. (2008). Detoxification of sulphidic African shelf waters by blooming chemolithotrophs. Nature, 457(7229), 581–584.

Lee, I., Kim, Y.O., Park, S.C., Chun ,J. (2016). OrthoANI: an improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol. 66:1100-1103.

Lichtschlag, A., Kamyshny, A., Jr, Ferdelman, T.G., and de Beer, D. (2013) Intermediate sulfur oxidation state compounds in the euxinic surface sediments of the Dvurechenskii mud volcano (Black Sea). Geochim Cosmochim Acta 105. 130–145.

Luther, G. W., III, Church, T. M., & Powell, D. (1991). Sulfur speciation and sulfide oxidation in the water column of the Black Sea. Deep Sea Research Part a, Oceanographic Research Papers, 38(Supplement 2), S1121–S1137. http://doi.org/10.1016/S0198-0149(10)80027-5

Luz, D.E., Nepomuceno, R.S.L., Spira, B., & Ferreira, R.C.C. (2012). The Pst system of *Streptococcus mutans* is important for phosphate transport and adhesion to abiotic surfaces. Molecular Oral Microbiology. 27. 172-181.

m, F.F., Dahl, C. (2008) Thiosulfate and sulfur oxidation in purple sulfur bacteria. Microbial Sulfur Metabolism. (Dahl C & Friedrich CG, eds), pp 101-116. Springer Verlag, Berlin.

Macalady, J. L., Dattagupta, S., Schaperdoth, I., Jones, D. S., Druschel, G. K., & Eastman, D. (2008). Niche differentiation among sulfur-oxidizing bacterial populations in cave waters. The ISME Journal, 2(6), 590–601.

Macalady, J. L., Jones, D. S., & Lyon, E. H. (2007). Extremely acidic, pendulous cave wall biofilms from the Frasassi cave system, Italy. Environmental Microbiology, 9(6), 1402–1414

Macalady, J. L., Lyon, E. H., Koffman, B., Albertson, L. K., Meyer, K., Galdenzi, S., & Mariani, S. (2006). Dominant Microbial Populations in Limestone-Corroding Stream Biofilms, Frasassi Cave System, Italy. Applied and Environmental Microbiology, 72(8), 5596–5609. http://doi.org/10.1128/AEM.00715-06

Madsen, E. L. (2005). Identifying microorganisms responsible for ecologically significant biogeochemical processes. Nature Reviews Microbiology, 1–8.

Marcia, M., Ermler, U., Peng, G., & Michel, H. (2009). The structure of Aquifex aeolicus sulfide:quinone oxidoreductase, a basis to understand sulfide detoxification and respiration. Proceedings of the National Academy of Sciences, 106(24), 9625–9630. http://doi.org/10.1073/pnas.0904165106

McBride, M.J., & Nakane, D. (2015). Flavobacterium gliding motility and the type IX secretion system. Curr. Opin. Microbiol. 28. 72-77.

Meyer, T. E., & Cusanovich, M. A. (2003). Discovery and characterization of electron transfer proteins in the photosynthetic bacteria. Photosynthesis Research, 76, 111–126.

Middleburg, J. (2000). The geochemical sulfur cycle. In: Lens P, Hulshoff Pol W (eds) Enviormental technologies to treat sulfur pollution. IWA Publishing, London, 33-46.

Monds, R.D., Silby, M.W., & Mahanty, K.H. (2001). Expression of the Pho regulon negatively regulates biofilm formation by *Pseudomonas aureofaciens* PA147-2. Mol Microbiol 42. 415-426.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A., and Kanehisa, M.; KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 35, W182-W185 (2007).

Mosquera-Rendon, J., Rada-Bravo, A.M., Cardenas-Brito, S., Corredor, M., Restrepo-Pineda, E., & Paez, A. (2016). Pangenome-wide and molecular evolution analyses of the Pseudomonas aeruginosa species. BMC Genomics. 17:45

Nguyen, N., Warnow, T., Pop, M. & White, B. (2016). A perspective on 16S rRNA operational taxonomic unit clustering using sequence similarity. Npj Biofilms and Microbioms, 2, 1-8.

Nordstrom, D. K., Ball, J. W., & McCleskey, R. B. (2005). Ground Water to Surface Water: Chemistry of Thermal Outflows in Yellowstone National Park. Geothermal Biology and Geochemistry in Yellowstone National Park, 76–94.

O'May, G.A., Jacobsen, S.M., Longwell, M., Stoodley, P., Mobley, H.L. and Shirtliff, M.E. (2009) The high-affinity phosphate transporter Pst in Proteus mirabilis HI4320and its importance in biofilm formation. Microbiology. 155. 1523–1535.

Ochsner, U.A., Wilderman, P.J., Vasil, A.I., & Vasil, M.L. (2002). GeneChip expression analysis of the iron starvation response in *Pseudomonas aeruginosa*: identification of novel pyoverdine biosynthesis genes. Mol. Microbiol. 45. 1277–1287

Ogawa, T., Furusawa, T., Nomura, R., Seo, D., Hosoya-Matsuda, N., Sakurai, H., & Inoue, K. (2008). SoxAX Binding Protein, a Novel Component of the Thiosulfate-Oxidizing Multienzyme System in the Green Sulfur Bacterium Chlorobium tepidum. Journal of Bacteriology, 190(18), 6097–6110. http://doi.org/10.1128/JB.00634-08

Pasteris, J. D., Freeman, J. J., Goffredi, S. K., & Buck, K. R. (2001). Raman spectroscopic and laser scanning confocal microscopic analysis of sulfur in living sulfur-precipitating marine bacteria. Chemical Geology, 180(1-4), 3–18. http://doi.org/10.1016/S0009-2541(01)00302-3

Quentmeier, A., & Friedrich, C. G. (2001). The cysteine residue of the SoxY protein as the active site of protein-bound sulfur oxidation of Paracoccus pantotrophusGB17. FEBS Letters, 503(2-3), 168–172. http://doi.org/10.1016/S0014-5793(01)02727-2

Quinlan, A.R., & Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26, 6, pp. 841–842.

Ransom-Jones, E., Jones, D.L., McCarthy, A.J., & McDonald, J.E. (2012). The Fibrobacteres: an important phylum of cellulose-degrading bacteria. Micrb. Ecol. 63:2. 267-281.

Robertson, L. A., & Kuenen, J. G. (1983). Thioshaera pantotropha gen. nov. sp. nov., a Facultatively Anaerobic, Facultatively Autotrophic Sulphur Bacteria. Journal of General Microbiology, 129, 2847–2855.

Roelsen, H., & Jorgensen, B. B. (1982). Seasonal Dynamics of Elemental Sulfur in Two Coastal Sediments. Estuarine, Costal and Shelf Science, 255–266.

Rother, D., Henrich, H.J., Quentmeir, A., Bardischewsky, F., & Friedrich, C.G. (2001). Novel genes of the sox gene cluster, mutagenesis of the flavoprotein soxF, and evidence for a general sulfur-oxidizing system in Paracoccus pantotrophus. GB17. J. Bacteriol. 183:15. 4499-4508.

Roy, A.B. & Trudinger, P.A. (1970) The Biochemistry of Inorganic Compounds of Sulfur. Cambridge, U.K.: Cambridge University Press.

Sauve, V., Bruno, S., Berks, B. C., & Hemmings, A. M. (2007). The SoxYZ Complex Carries Sulfur Cycle Intermediates on a Peptide Swinging Arm. Journal of Biological Chemistry, 282(32), 23194–23204. http://doi.org/10.1074/jbc.M701602200

Shutz, M., Maldener, I., Griesbeck, C., & Hauska, G. (1999). Sulfide-Quinone Reductase from Rhodobacter capsulatus: Requirement for Growth, Periplasmic Localization, and Extension of Gene Sequence Analysis. Journal of Bacteriology, 181(20), 6516–6523.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:21. 2688-2690.

Stockdreher, Y., Venceslau, S.S., Josten, M., Sahl, H-G., Pereira. I.A.C., & Dahl, C. (2012). Cytoplasmic Sulfurtransferases in the Purple Sulfur Bacterium Allochromatium vinosum: Evidence for Sulfur Transfer from DsrEFH to DsrC. PLoS ONE 7(7): e40785.

Taylor, S. R, (1964). Abundance of chemical elements in the continental crust: a new table. Geochimica Acta. 28, 1273-1285.

Waite, D.W., Vanwonterghem, I., Rinke, C., Parks, D.H., Zhang, Y., Takai, K., Sievert, S.M., Simon, J., Campbell, B.J., Hanson, T.E., Woyke, T, Klotz, M.G., & Hugenholtz, P. (2017). Comparative Genomic Analysis of the Class Epsilonproteobacteria and Proposed Reclassification to Campylobacterota (phyl. nov.). Front. Microbiol. 8:682.

Weissgerber, T., Watanabe, M., Hoefgen, R., & Dahl, C. (2014). Metabolomic profiling of the purple sulfur bacterium Allochromatium vinosum during growth on different reduced sulfur compounds and malate. Metabolomics, 10(6), 1094–1112.

Winstedt, L., & von Wachenfeldt, C. (2000). Terminal oxidases of *Bacillus subtilis* strain 168: one quinol oxidase, cytochrome aa3 or cytochrome bd, is required for aerobic growth. J. Bacteriol.182. 6557–6564

Wright, K. E., Williamson, C., Grasby, S. E., Spear, J. R., & Templeton, A. S. (2013). Metagenomic evidence for sulfur lithotrophy by Epsilonproteobacteria as the major energy source for primary productivity in a sub-aerial arctic glacial deposit, Borup Fiord Pass. Frontiers in Microbiology, 1–20. http://doi.org/10.3389/fmicb.2013.00063/abstract

Xie, H., Buschmann, S., Langer, L.D., Bernd, L., & Michel, H. (2014). Biochemical and Biophysical Characterization of the Two Isoforms of the cbb3-Type Cytochrome c Oxidation from *Pseudomans stutzeri.* J Bacteriol. 196:2. 472-482.

Yamanaka, T., Fukumori, Y., & Okunuki, K. (1979). Preparation of Subunits of Flavocytochromes c Derived from Chlorobium limicola f. thiosulfatophilum and Chromatium vinosum. Analytical Biochemistry, 95, 209–213.

Yue, J.C. & Clayton, M.K., (2005). A similarity measure based on species proportions. Communications in Statistics-Theory and Methods. 34:11, 2123-2131.

Zhang, J.-Z., & Millero, F. J. (1993). The products from the oxidation of H2S in seawater. Geochimica Et Cosmochimica Acta, 57, 1705–1718.

Appendix A

SUPPLEMENTARY MATERIAL

Table A.1 Protein sequences used for the concatenated single copy gene tree

Protein Name
Initiation Factor 2
Ribosomal Protein L22
Ribosomal Protein L23
Ribosomal Protein L2-C
Ribosomal Protein L3
Ribosomal Protein L4
Ribosomal Protein S11
Ribosomal Protein S15
Ribosomal Protein S19
RNA Polymerase Rpb6
Alanine—tRNA ligase

Table A.2 The Rank abundance of genes in the cbb3 terminal oxidase. These ranks do not include hypothetical and structural ribosome genes.

Gene	Net 1 Day 3	Net 1 Day 11	Net 1 Day 15	Net 1 Day 17	Net 2.1 Day 4	Net 2.1 Day 6	Net 2 Day 11	Crusty	Fluffy	Natural Streamer
ccoN	1	1	1	1	1	1	1	35	1	1
ccoO	4	2	3	2	2	2	4	19	5	2
ccoP	5	3	2	3	4	3	2	16	4	3

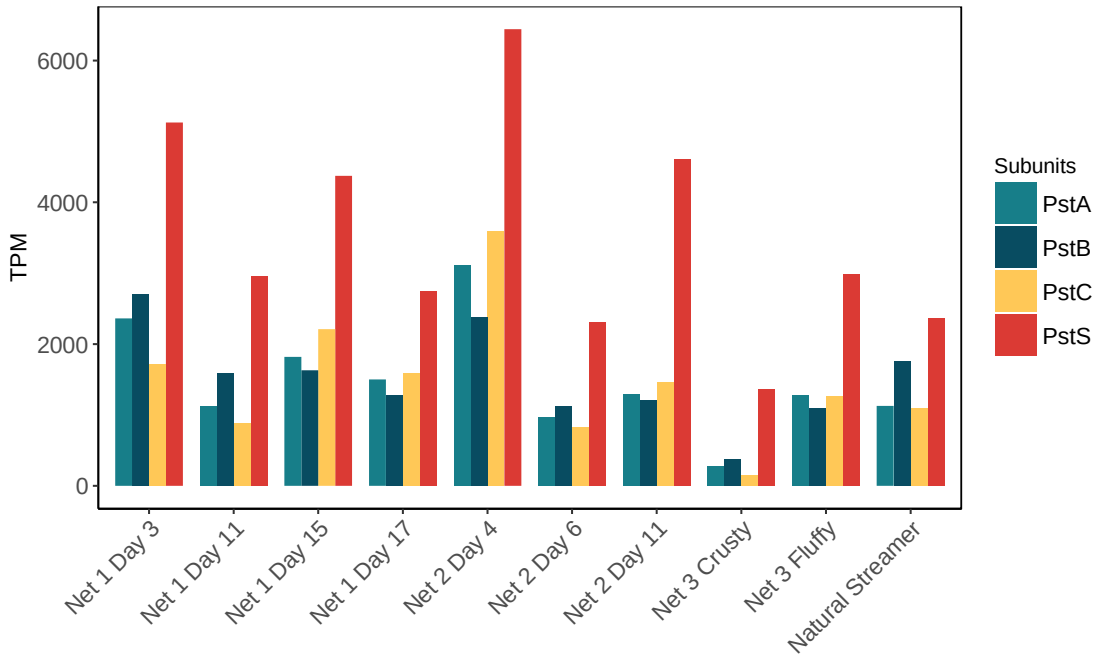

Figure A.1 Community wide expression of the subunits involved in the pst system

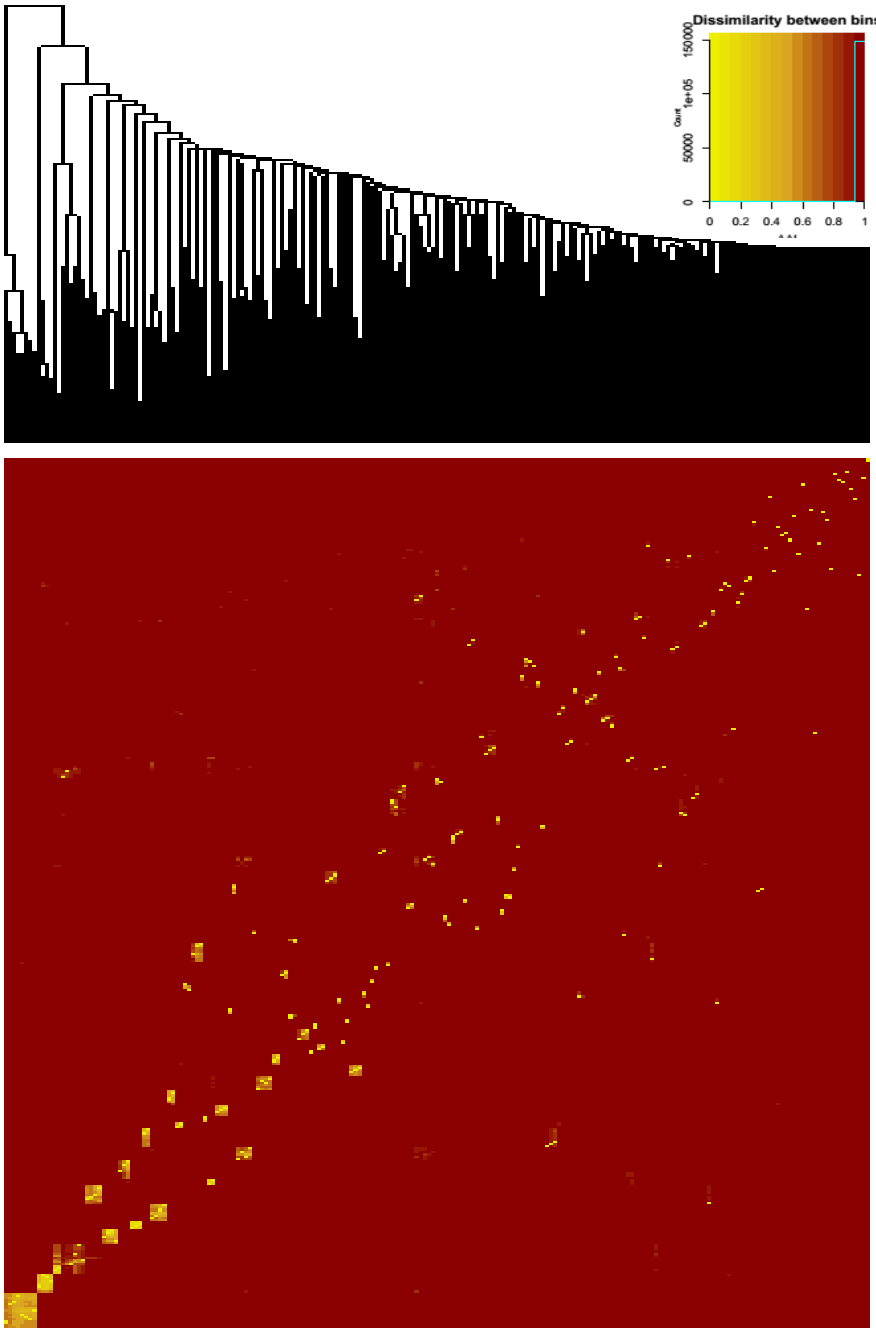

Figure A.2 Pairwise comparison of Amino Acid Identity (AAI) within 281 MAGs generated from the three assembly methods. Each leaf on the dendrogram represents a different MAG comparison. Clusters of MAGs with >95% AAI identity are in yellow.