

Metagenomic Profiling of Cave Biofilms, With a Focus on Chemolithoautotrophy and Pseudonocardiaceae

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

John Anthony Fulginiti
August 2023

ACKNOWLEDGEMENTS

Throughout my time at the University of Tennessee, I had an incredible amount of support from various organizations, faculty, friends, and family. I would like to thank the Jones Endowment and the Cave Conservancy Foundation for providing the funding that was necessary for metagenomic sequencing. As someone who had a stronger background in environmental science during my previous education, learning about metagenomics and applying it to a study was, without a doubt, the highlight of my education in graduate school. I would also like to thank my thesis committee members, Dr. Jennifer DeBruyn and Dr. Andrew Steen, for taking interest in my study and offering critical feedback that aided the direction of my thesis. I am also grateful for our lab manager, Audrey Paterson, who was there to answer many questions regarding bioinformatics pipelines. Thanks to Dr. Annette Engel for bringing me aboard the lab group, providing critical scientific writing advice and in-depth discussions on topics I found interesting in the literature, and overall, making my project possible. I would also like to thank Ethan Sweet, Hannah Rigoni, Jason Olaveson, and Allison Rogers, along with Audrey and Annette, for attending many practice presentations I gave for class projects, the GSA, and my thesis defense. Their feedback allowed me to improve my presentation skills and experience, and I am looking forward to carrying over what I learned into the professional world. Finally, I would like to thank my mom and dad for their support over the years, and giving me the opportunity to go to college and get an education in a field I had been interested in since elementary school.

ABSTRACT

The dark, humid, and generally nutrient-depleted conditions of caves set limitations on their habitability, which causes the microbial diversity and community structure of caves to differ from those in surface environments. Recent studies investigating cave microbiomes reveal that Pseudonocardiaceae, a family of bacteria belonging to the phylum Actinomycetota, are abundant in cave sediment and on rocks as subaerial biofilms. However, little work has been done to understand their ecology and geomicrobiological impact in the cave environment. This study used shotgun metagenomics to characterize the functional potential of Pseudonocardiaceae, as well as other microbes, from in Carter Cave, located in Knox County, Tennessee. From 16S rRNA gene sequence analysis, done to understand the microbial diversity at the family and genus levels, *Crossiella*, a genus belonging to the Pseudonocardiaceae, were the most abundant members of the biofilms. These findings are consistent with literature describing other cave biofilm community compositions. Two high-quality Metagenome Assembled Genomes (MAGs) with >90% completeness and <10% contamination were obtained that were taxonomically affiliated with the Pseudonocardiaceae, and revealed the functional potential for chemolithoautotrophic CO₂ fixation, denitrification, trace gas uptake, antimicrobial defense, and calcite precipitation. Two other MAGs, belonging to the classes Blastocatellia and Gammaproteobacteria were also retrieved, which contained carbon metabolic pathways indicative of heterotrophy and trace gas uptake of hydrogen and methane. All four MAGs also revealed the potential for the production of carbonic anhydrase, which has been associated with microbially induced calcite precipitation (MICP), various forms of antibiotic resistance, and the production of bacterial toxins. If metabolically active, then the Pseudonocardiaceae likely play crucial roles to the ecosystem by producing organic carbon, acting as a sink for carbon monoxide in cave air, limiting microbial interactions in biofilms, and participating in microbe–mineral interactions that can lead to calcite precipitation. This work highlights the functional potential of dominant microbial groups found in cave-wall biofilms and explores their innovative strategies that enhance their ability to thrive in the subsurface. This investigation reinforces the connections between microbial ecology and the biogeochemistry of an environment and raises interests for possible applications in biotechnology, pharmaceuticals, and astrobiology.

TABLE OF CONTENTS

SECTION ONE	1
INTRODUCTION	1
Research Objectives.....	2
Research Hypotheses	2
Thesis Organization	3
SECTION TWO.....	4
LITERATURE REVIEW	4
The Cave Environment	4
Carbon Metabolism.....	4
Nitrogen Metabolism	6
Cave Microbiomes	7
The Actinomycetota Phylum and the Family Pseudonocardiaceae	8
The Pseudomonadota	15
Geochemistry as a Tool to Trace Metabolism and Ecological Function	15
SECTION THREE.....	17
MATERIALS AND METHODS.....	17
Site Description.....	17
DNA Extraction and 16S rRNA Gene Sequence Analysis.....	17
Shotgun Metagenomic Sequencing.....	18
Metagenome Assembly, Binning, and Gene Calling.....	19
Sequence Deposition and Data Availability	20
SECTION FOUR	26
RESULTS	26
Updated 16S rRNA Microbial Community Analysis	26
Taxonomic Classification and Phylogenetic Placement of MAGs.....	26
Functional Annotations	31
Carbon Metabolism.....	31
Trace Gas Uptake.....	43
Microbe-Mineral Interactions	43
DNA Repair Strategies	43
Survival Strategies and Biofilm Formation	48
SECTION FIVE.....	53
DISCUSSION	53
Metabolic Potential for Pseudonocardiaceae in Cave-wall Biofilms	53
Pseudonocardiaceae Taxonomy	55
Metabolic Potential and Taxonomy of Other MAGs in the Cave-wall Biofilms	56
Biofilm Lifestyle – Formation, Maintenance, and Survival Implications	56
Implications for the Cave Habitat	59
SECTION SIX CONCLUSIONS	61
References	63
Vita.....	94

LIST OF TABLES

Table 2.1: Studies (not an exhaustive list and focused mostly on caves) that report Actinomycetota average relative abundances and specific genera within the phylum according to targeted 16S rRNA or other gene sequencing from cloning (C) experiments or gene surveys (GS), metagenomes (M), isolations (I) from culturing, as well as in the context of sampling location, sample type, and biofilm color (W = white, Y = yellow, G = gray, P = pink, B = blue).	9
Table 3.1: Bins, with completeness and contamination noted, along with taxonomic affiliations. Bin quality is based on Parks et al. (2015). “*” indicates exceptions that were made to the criteria of Bowers et al. (2017).	22
Table 3.2: Summary and statistics of the four highest quality bins that were generated and extracted from the Das Tool binning optimization tool.....	25
Table 4.1: Comparison of ANI values of the Pseudonocardiaceae MAGs (bin.005 and bin.008) to other Pseudonocardiaceae genera. Determinations were made using FastANI.....	42
Table 4.2: Genes involved in nitrogen metabolism in cave-wall biofilm MAGs from Carter Cave, Tennessee. Determinations were made using DRAM and KEGG Mapper.....	44
Table 4.3: Genes present in the high-quality MAGs from cave-wall biofilms, Carter Cave, Tennessee, that could be involved with the production of trace gas uptake enzyme systems. Determinations were made using DRAM and KEGG Mapper.	45
Table 4.4: Genes present in the high-quality MAGs from cave-wall biofilms, Carter Cave, Tennessee, that could be involved in carbonic anhydrase (CA). Determinations were made using DRAM and KEGG Mapper.....	46
Table 4.5: The presence (✓) or absence (X) of genes that encode for the RecF/MutHLS pathways. Determinations were made using DRAM and KEGG Mapper.	47
Table 4.6: The presence (✓) or absence (X) of genes encoding for the four subunits of RNA polymerase and various sigma factors. Determinations were made using DRAM and KEGG Mapper.	49
Table 4.7: Genes associated with biofilm formation, motility, and maintenance using <i>V. cholerae</i> , <i>P. aeruginosa</i> , <i>E. coli</i> as reference pathways. Determinations were made using DRAM and KEGG Mapper.	51
Table 4.8: Presence of genes used in antimicrobial resistance, toxin-antitoxin systems, and bacterial toxins for the high-quality MAGs from cave-wall biofilms, Carter Cave, Tennessee. Determinations were made using DRAM and KEGG Mapper.....	52

LIST OF FIGURES

Figure 3.1: CheckM quality tests of the Das Tool binning output in KBase (Arkin et al., 2018).	21
Figure 4.1: Relative abundance of taxa at the phylum level for Carter Cave sites 1 and 4, from samples collected by Frazier (2020) as swabs or wall scrapes. Only taxa with relative abundances greater than 1.0% are shown.	27
Figure 4.2: Relative abundance of taxa at the order level for Carter Cave sites 1 and 4, from samples collected by Frazier (2020) as swabs or wall scrapes. Only taxa with relative abundances greater than 1.0% are shown.	28
Figure 4.3: Relative abundance of taxa at the family level for Carter Cave sites 1 and 4, from samples collected by Frazier (2020) as swabs or wall scrapes. Only taxa with relative abundances greater than 1.0% are shown.	29
Figure 4.4: Relative abundance of taxa at the genus level for Carter Cave sites 1 and 4, from samples collected by Frazier (2020) as swabs or wall scrapes. Only taxa with relative abundances greater than 1.0% are shown.	30
Figure 4.5: A phylogenetic tree created using Species Tree v2.2.0 for the MAG resulting from Bin.005, which was most closely related to <i>Actinoalloteichus hoggarensis</i> within the Pseudonocardiaceae (Actinomycetota), and other close relatives based on a set of 49 core universal genes for COG gene families (Arkin et al., 2018).	32
Figure 4.6: A phylogenetic tree created using Species Tree v2.2.0 for the MAG resulting from Bin.008, which was most closely related to <i>Actinoalloteichus hoggarensis</i> within the Pseudonocardiaceae (Actinomycetota), like bin.005, and other close relatives based on a set of 49 core universal genes for COG gene families (Arkin et al., 2018).	33
Figure 4.7: A phylogenetic tree created using Species Tree v2.2.0 for the MAG resulting from Bin.010, which was most closely related to <i>Pyrinomonas methylaliphatogenes</i> within the Blastocatellia class (Acidobacteriota), and other close relatives based on a set of 49 core universal genes for COG gene families (Arkin et al., 2018).	34
Figure 4.8: A phylogenetic tree created using Species Tree v2.2.0 for the MAG resulting from Bin.011, which was most closely related to <i>Nitrosococcus halophilus</i> Nc_4 within the Gammaproteobacteria class (Pseudomonadota), and other close relatives based on a set of 49 core universal genes for COG gene families (Arkin et al., 2018).	35
Figure 4.9: Heat map visualization of pathways involved in carbon metabolism for the high-quality MAGs obtained from cave-wall biofilms in Carter Cave, Tennessee.	36

Figure 4.10: KEGG map of the reductive pentose phosphate (RPP) pathway in the Bin.005 Pseudonocardiaceae MAG. Green-shaded boxes refer to the presence of genes necessary for the function of the RPP pathway.	37
Figure 4.11: KEGG map of the reductive pentose phosphate (RPP) pathway in the Bin.008 Pseudonocardiaceae MAG. Green-shaded boxes refer to the presence of genes necessary for the function of the RPP pathway.	38
Figure 4.12: KEGG map of the TCA cycle in the Bin.005 Pseudonocardiaceae MAG. Green-shaded boxes refer to the presence of genes necessary for the function of the TCA cycle. Red boxes indicate the absence of genes.	39
Figure 4.13: KEGG map of the TCA cycle in the Bin.008 Pseudonocardiaceae MAG. Green-shaded boxes refer to the presence of genes necessary for the function of the TCA cycle. Red boxes indicate the absence of genes.	40
Figure 4.14: Heat map visualization using DRAM of methane metabolism pathways for high-quality MAGs from cave-wall biofilms, Carter Cave, Tennessee.	41

SECTION ONE

INTRODUCTION

Caves are dark environments with high relative humidity and generally stable temperatures, but can also have depleted concentrations of available nutrients and organic carbon that would be needed to sustain subsurface ecosystems. Such conditions set limitations on habitability and control microbial diversity, metabolism, and community structure that differ from surface environments. Yet, across a range of subsurface settings, from caves in karst landscapes, volcanic caves or lava tubes, tombs, and mines, two phyla dominate microbial communities, the Actinomycetota (the phylum previously referred to as Actinobacteria) and Pseudomonadota (the phylum previously referred to as Proteobacteria) (Ortiz et al., 2014; De Mandal., 2017; Li et al., 2018; Alonso et al., 2019; Long et al., 2019; Wiseschart et al., 2019; Zhu et al., 2019; Frazier, 2020; Turrini et al., 2020; Cimermanova et al., 2021; Gonzalez-Pimentel et al., 2021; Ma et al., 2021; Zhu et al., 2021; Martin-Pozas et al., 2022). The Pseudonocardiaceae, a family within the Actinomycetota, can have among the highest relative abundances of 16S rRNA gene sequences associated with the phylum, but many are not taxonomically related to previously cultured bacteria (e.g., Lynch et al., 2014; Turrini et al., 2020; Ghezzi et al., 2021; Gonzalez-Pimentel et al., 2021). Consequently, despite the prevalence of this group in a range of habitats, their ecology is not well understood, especially within the context of the cave environment. Put simply, we lack insight into what the Pseudonocardiaceae are doing metabolically within the cave ecosystems.

Currently, the best information about the metabolism of the Pseudonocardiaceae family comes from a combination of chemotaxonomic studies and 16S rRNA gene surveys that allude to diverse members of the family being capable of heterotrophy, meaning that they obtain their carbon and energy from organic carbon (Borsodi et al., 2012; Franco and Labeda, 2014; Cao et al., 2015; Cao et al., 2017; Fang et al., 2017; Rangseekaew and Pathom-aree, 2019). But, other metabolisms have been suggested, including chemolithoautotrophy through both aerobic and anaerobic CO₂ fixation, energy metabolism from trace gases, such as carbon monoxide (CO), as well as oxidation of different nitrogen compounds and methane (Lynch et al., 2014; Ortiz et al., 2014; Martin-Pozas et al., 2020; Martin-Pozas et al., 2022). A breakthrough came from Frazier (2020) using cave-wall biofilm samples from Carter Cave in Knox County, Tennessee (state cave

inventory number KN14). Stable carbon and nitrogen isotope ratio analyses of the cave biofilms overwhelmingly dominated by Pseudonocardiaceae suggest chemolithoautotrophic fixation of CO₂ from the cave air because the $\delta^{13}\text{C}$ [delta 13-C] values of the biofilm biomass were more negative than would be predicted if the microbes were heterotrophs and using organic carbon from within the cave. Moreover, low $\delta^{13}\text{C}$ values also point to methane-derived carbon incorporation. Nitrogen isotope analyses suggest the presence of nitrifying and nitrogen-fixing microbes within the cave biofilm samples.

Research Objectives

This study expands on the previous 16S rRNA gene sequencing and geochemical analyses from Frazier (2020) and takes a shotgun metagenomics approach to investigate the diversity and metabolic strategies of the Pseudonocardiaceae and other microbial groups from within the cave-wall biofilms, specifically focusing on the ability of these microbes to fix CO₂ and/or use CO and/or CH₄ as chemolithoautotrophs. To our knowledge, this investigation is the first to use metagenomics to tie cave-wall biofilm Pseudonocardiaceae to carbon and energy acquisition pathways, as well as among the first studies to assign specific functions to other microbes common to cave-wall biofilm communities using metagenomics.

The main research objectives are four-fold:

- Characterize and describe the microbial diversity and metabolic functions from cave-wall biofilms using 16S rRNA gene sequences and metagenome-assembled genomes (MAGs).
- Investigate the possibility of chemosynthesis via trace gas utilization.
- Investigate survival strategies, such as DNA repair mechanisms, biofilm formation, and antimicrobial production, in cave-wall biofilm MAGs.
- Determine the taxonomic relationships among cave-wall biofilms, with a specific focus on members of the Pseudonocardiaceae.

Research Hypotheses

The following hypotheses were tested:

- 1) Because cave-wall biofilms from Carter Cave contain low diversity and are dominated by members of the Pseudonocardiaceae (Frazier, 2020), it will be possible to obtain MAGs belonging to the family Pseudonocardiaceae.

- 2) Pseudonocardiaceae MAGs contain genes that encode chemolithoautotrophy enzymes involved with CO₂ fixation.
- 3) If the metagenomic assessment indicates that the cave-wall biofilm MAGs contain chemolithoautotrophy genes, then trace gas utilization genes will be present for the purpose of providing energy in the absence of sunlight.
- 4) MAGs will harbor sets of genes that encourage mutualism within biofilm communities and aid in biofilm defense against outside competitors.
- 5) Members of the Pseudonocardiaceae from Carter Cave cave-wall biofilms are novel taxa.

Thesis Organization

This thesis is organized into five additional sections following this introductory section. Section Two is a literature review of the current knowledge about cave microbiomes, focusing specifically on the Actinomycetota, to highlight potential carbon and nitrogen metabolisms that could be occurring within caves. The usefulness of geochemistry as a tool to understand cave microbiology is also included. Section Three describes the metagenome methods used in this research, with a summary of the field and other methods from Frazier (2020) that were relevant to this current study. Section Four provides the results, specifically identifying key metabolic functions in the obtained MAGs. Section Five is the main discussion and evaluation of the research hypotheses, focused on assessing the potential for chemolithoautotrophic carbon fixation by key Actinomycetota in cave-wall biofilms. Section Six concludes the thesis and provides a discussion of future research directions.

SECTION TWO

LITERATURE REVIEW

The Cave Environment

The input of allochthonous organic carbon into caves, naturally from streams, fauna, and air currents, shapes subsurface food web dynamics, biogeochemical cycling, as well as microbial diversity (Barton and Jurado, 2007). Throughout their extent, caves contain numerous microhabitats, such as in water, as subaerial or cave-wall biofilms, and within sediment. Previous studies demonstrate that cave microbial communities tend to be dominated by heterotrophic microbial activity, being dependent on allochthonous organic carbon sourced from the surface (Barton and Jurado, 2007; Gabriel and Northup, 2013; Oliveira et al., 2017). However, microbes capable of chemolithoautotrophy can also be present (Chen et al., 2009; Ortiz et al., 2014; Oliveira et al., 2017; Zhao et al., 2018; Zhu et al., 2019; Ortiz et al., 2021; Zhu et al., 2021; Ray et al., 2022), particularly when organic carbon concentrations are low, which is a condition referred to as oligotrophic. To begin to understand the importance of carbon and other nutrients to the cave microbiome, the following sections describe our current knowledge about microbial functional systems that are relevant to a cave microbiome, including carbon metabolism focused on chemolithoautotrophic carbon fixation pathways and nitrogen metabolism.

Carbon Metabolism

Despite the physical, geochemical, and biological diversity of all ecosystems on Earth, the fact remains that all living organisms need an energy source, growth substrate, and a source of carbon to sustain themselves. There are several metabolic processes responsible for fixing inorganic carbon into organic carbon and biomass, including the reductive pentose phosphate pathway (RPP) or Calvin Benson-Bassham Cycle (CBB), the reductive acetyl-coenzyme A (acetyl CoA) or Wood-Ljungdahl pathway, the reductive tricarboxylic acid (rTCA) cycle, and others. Most autotrophic CO₂ fixation occurs through the reductive pentose phosphate (RPP) pathway (Bar-Even et al., 2012; Liang et al., 2020) where CO₂ is reduced to glucose via the ribulose-1,5-bisphosphate carboxylase (RuBisCO) enzyme (Hügler and Fuchs, 2005; Hügler and Sievert, 2011). RuBisCO catalyzes the addition of one CO₂ molecule to ribulose biphosphate, and the product of this carboxylation reaction is split into two 3-phosphoglycerate molecules (Shively et al., 1998; Sharkey, 2019). The oxidative branch of this pathway, the pentose

phosphate pathway, is primarily used by heterotrophs (Bar-Even et al., 2012). The acetyl-CoA synthase enzyme serves as an excellent marker for identifying the reductive acetyl-CoA pathway in prokaryotes, as it catalyzes the final step that synthesizes acetyl CoA by combining CO and a methyl group (Wood et al., 1986; Bar-Even et al., 2012). Using this pathway, microbes can form multiple carbon compounds, like acetate, with H₂ as an electron donor (Wood et al., 1986; Bar-Even et al. 2012; Liang et al., 2020). Among aerobic bacteria, the enzyme CO dehydrogenase catalyzes the reduction of CO₂ to a carbonyl group, using CO₂ as a source of carbon and CO as a source of energy. Moreover, the 3-hydroxypropionate (3HP) cycle can also result in CO₂ fixation using acetyl-CoA and propionyl-CoA (Hügler and Fuchs, 2005; Hügler and Sievert, 2011). The 3HP pathway was previously confined to green non-sulfur bacteria that use reduced sulfur compounds as electron donors (Alber and Fuchs, 2002), but key enzymes in this pathway have also recently been detected in members within the Actinomycetota and Pseudomonadota (Garritano et al., 2022).

Carbonic anhydrase (CA) is another enzyme associated with atmospheric CO₂ uptake by microbes (Liu et al., 2020; Gonzalez-Pimentel et al., 2021). From subsurface settings, Martin-Pozas et al. (2022) link high abundances of CA genes to Actinomycetota, and specifically the genera *Crossiella* and wb1-P19. For the Actinomycetota living in caves, CA has been hypothesized to help organisms survive long periods of nutrient depletion and perhaps can be used to produce energy in carbon-depleted environments (Ortiz et al., 2014; Ortiz et al., 2021)

Seven bacterial phyla contain members capable of atmospheric gas uptake, such as H₂, CO, and methane (CH₄), using enzymes like [NiFe]-hydrogenase, [MoCu]-carbon monoxide dehydrogenase, and methane monooxygenase, respectively, including the Actinomycetota (Ray et al., 2022). These metabolisms have been identified from cave ecosystems, as well (Ortiz et al., 2014; Ortiz et al., 2021; Martin-Pozas et al., 2022; Ray et al., 2022). While CO and H₂ oxidation have not been measured in cave microbial ecology studies, microbes in soil ecosystems can oxidize both trace gases even if below atmospheric concentrations (Cordero et al., 2019; Greening et al., 2022). Methane-oxidizing (also referred to as methanotrophic) bacteria can be found in both aerobic and anaerobic environments (Oremland and Culbertson, 1992; Kou et al., 2017). Aerobic methanotrophs convert oxygen and methane into formaldehyde, which is then further processed by one of three pathways, the ribulose monophosphate pathway, the serine

pathway, or the ribulose biphosphate cyclic pathway (Stein et al. and Dunfield, 2012; Khmelenina et al., 2018). The conserved functional gene *pmoA* is necessary for the membrane-bound particulate methane monooxygenase enzyme (pMMO), which catalyzes the conversion of CH₄ to methanol (CH₃OH) (Miyaji, 2011). Methane oxidation is a biological sink for atmospheric methane (Tveit et al., 2019), which makes understanding this pathway in caves a valuable prospect for climate change mitigation.

Nitrogen Metabolism

Despite the typically low (Frazier, 2020; Paula et al., 2020; Lerch, 2011) to limited (Ortiz et al., 2014) concentrations of dissolved ammonium (NH₄⁺), nitrate (NO₃⁻), and nitrite (NO₂⁻) in cave water, a range of microbes capable of metabolizing different nitrogen compounds, including nitrogen gas (N₂), can be found in the subsurface. Ammonia oxidation is a two-step process. Under aerobic conditions, ammonia-oxidizing bacteria and archaea convert ammonia to nitrite or nitrate (Hatzenpichler, 2012), which can also provide a source of inorganic energy for chemolithoautotrophs (Hathaway et al., 2014). The major enzymes used include ammonia monooxygenase, which is encoded for by the required *amoA* gene (Gilch and Schmidt., 2009; Hathaway et al., 2014), and hydroxylamine (NH₂OH) oxidoreductase, encoded for by the *hao* gene, that converts hydroxylamine to nitrite (Lancaster and Caranto, 2018). Assimilatory nitrate reduction is carried out by assimilatory cytoplasmic nitrate reductase, nitrite reductase, and membrane-bound nitrate reductases (Feng et al., 2014). Genes belonging to the *nas* gene family encode for nitrite reductase and nitrate reductase (Lin et al., 1993). Dissimilatory nitrate reduction can also play a role in autotrophic nitrogen metabolism when oxygen is limited (Shapleigh, 2009; Feng et al., 2014). Assimilatory nitrate reduction involves the production of ammonia that is introduced into the cell to assemble proteins and amino acids (Shapleigh, 2009; Kamp et al., 2015). Nitrite can also be produced in some cases as a metabolic intermediate (Ortiz et al., 2014). Under aerobic conditions, autotrophs are typically associated with N₂ fixation and assimilatory nitrate reduction (Zehr et al., 2003; Ai et al., 2022). Some taxa within the Actinomycetota can fix N₂, such as *Mycobacterium*, *Arthrobacter*, and *Streptomyces*. Nitrogen fixation requires the genes *nifH*, *nifD*, and *nifK* that encode for different nitrogenase enzyme subunits (Zehr et al., 2003).

Anoxic environments are the typical ecological niche for microorganisms that perform dissimilatory nitrate reduction (Moreno-Vivián et al., 1999); however, competition and mutualism within microbial communities in oxic environments, such as biofilms, may encourage the use of nitrate as an electron acceptor instead of oxygen, even in the presence of oxygen (Hennecke et al., 1981; Richardson et al., 2001; Shapleigh, 2009; Feng et al., 2014). For instance, Lv et al. (2017) observe and describe members of the Actinomycetota and Pseudomonadota capable of denitrification in the presence of oxygen. Dissimilatory nitrate reduction to nitrous oxide (N₂O), N₂, or ammonium (Moreno-Vivián et al., 1999; Richardson et al., 2001; Feng et al., 2014) is a four-step process, involving the use of six different enzymes and twelve key genes (Mohan and Cole, 2007). The first step is the conversion of nitrate to nitrite via periplasmic nitrate reductase, encoded by both the *nap* and *nar* gene families (Sparacino-Watkins et al., 2014; Chen and Wang, 2015). The second step reduces nitrite to nitric oxide (NO). The principal enzyme is Cu-NIR, which allows this pathway to function more efficiently in aerobic environments (Lawton et al., 2013), and is encoded by the *nir* gene family. Although NO can serve as an energy source for some microbes, it can be toxic to others. At high concentrations, NO displays antimicrobial potential and its presence can act as a defense mechanism against competing microbes and invading viruses (Fang, 2004; Toyofuku and Yoon, 2018). Microbes that use NO do so through the nitric oxide reductase enzyme, encoded by *nor* gene family, and produce N₂O as a byproduct (Fang, 2004; Moënné-Loccoz, 2007; Borrero-de Acuña et al., 2016; Toyofuku and Yoon, 2018). The NorB subunit is responsible for accumulating two NO molecules, whereas the NorC subunit is responsible for electron transfer and receiving electrons from electron donors (Toyofuku and Yoon, 2018). The final step for N₂O to N₂ reduction, used for respiration via the N₂O reductase enzyme, is encoded for by *nos* gene family (Berks et al., 1995).

Cave Microbiomes

Studies from within the last two decades using 16S rRNA gene surveys and other methods from different types of caves, including those in karst and volcanic landscapes, show that Actinomycetota and Pseudomonadota consistently dominate microbial communities in biofilms (e.g., Barton and Jurado, 2007; Stomeo et al., 2008; Chen et al., 2009; Ortiz et al., 2014; Lavoie et al., 2017; Li et al., 2018; Long et al., 2019; Wischart et al., 2019; Zhu et al., 2019;

Frazier, 2020; Gonzalez Pimentel et al., 2021; Ma et al., 2021; Zhu et al., 2021). The Actinomycetota phylum is one of the largest and most diverse, being comprised of bacteria found in terrestrial, subterranean, aquatic, and marine ecosystems. Table 2.1 summarizes some past studies about Actinomycetota, listed with class- to genus-level classifications, if possible. The Pseudomonadota are prominently known for their association with human disease and gut microbiomes (e.g., Ravcheev and Thiele., 2017), although diverse Pseudomonadota taxa have unique survival strategies, like N₂ fixation (e.g., Tsoy et al., 2016), that likely explain why they are among the most abundant and diverse groups of bacteria on the planet, including subsurface settings like caves (Tomczyk-Żak and Zielenkiewicz, 2016; Turrini et al., 2020).

The remainder of this section is a literature review of our current understanding of the Actinomycetota and Pseudomonadota, including their metabolic diversity, notable survival strategies, and possible explanations why these phyla are prevalent in cave, subsurface, and similar environments. For instance, the prevalence of Actinomycetota and Pseudomonadota in cave-wall biofilms may be because the extreme conditions of the cave wall environment, specifically oligotrophy and changes in moisture content, select for and support the growth of distinct taxa within these phyla (Barton and Jurado, 2007; Turrini et al., 2020). Further, biosynthetic pathways and bioactive compound production, typically by many different Actinomycetota, control biofilm diversity, such that some groups are excluded from a biofilm community to enhance survival of others (Wiseschart et al., 2019; Turrini et al., 2020). Overall, however, there has been limited research done to understand the key genes associated with biofilm formation, antimicrobial resistance, bacterial toxins for cave-wall biofilm communities, which motivates exploration of these topics for this thesis.

The Actinomycetota Phylum and the Family Pseudonocardiaceae

Actinomycetota are prevalent in extreme environments, such as caves, deserts, and coastal sediments based on 16S rRNA surveys (Table 2.1) (e.g., Barton and Jurado, 2007; Gabriel and Northup, 2013; Ortiz et al., 2014; Brannen-Donnelly and Engel, 2015; Frazier, 2020; Zhu et al., 2021; Benaud et al., 2022; Martin-Pozas et al., 2022; Ray et al., 2022). Research on Actinomycetota in extreme environments can advance various industries (Qin et al., 2019) because of their biochemical properties and secondary bioactive metabolites (Barka et al., 2016;

Table 2.1: Studies (not an exhaustive list and focused mostly on caves) that report Actinomycetota average relative abundances and specific genera within the phylum according to targeted 16S rRNA or other gene sequencing from cloning (C) experiments or gene surveys (GS), metagenomes (M), isolations (I) from culturing, as well as in the context of sampling location, sample type, and biofilm color (W = white, Y = yellow, G = gray, P = pink, B = blue).

Location	Sample Type	Biofilm Color	Relative Abundance (%)	Most Abundant Taxa (% , if reported)	Method(s) Used	Study
Carlsbad Cavern, New Mexico, USA	Biofilm, weathered rock	NA	80	Genera: <i>Pseudonocardia</i> , <i>Nocardioides</i> , <i>Rhodococcus</i> , <i>Mycobacterium</i> , <i>Curtobacterium</i>	C, GS	Barton and Jurado, 2007
Ardales, Southern Spain, and Biscay, Northern Spain	Biofilm	W	44–66	Genera: <i>Pseudonocardia</i> , <i>Streptomyces</i> , <i>Rubrobacteridae</i>	C, GS	Stomeo et al., 2008
Altamira Cave, Spain	Biofilm	W, Y, G, P	50	Genera: <i>Pseudonocardinae</i>	C, GS	Porca et al., 2012
Altamira Cave, Spain	Biofilm	Y, W, G	78	Genera: <i>Pseudonocardia</i> (17)	C, GS	Cuezva et al., 2012
Kipuka Kanohina Cave Preserve, Hawai'i	Biofilm	Y, W	16	Genera: <i>Pseudonocardiaceae</i>	GS	Hathaway et al., 2014
Atacama Desert volcano debris field, Chile	Soil	NA	98	Genera: <i>Pseudonocardia</i> (56–72), <i>Saccharopolyspora</i> (9–13), <i>Actinomycetales</i> , <i>Acidimicrobiales</i> , <i>Rubrobacterales</i>	GS, M	Lynch et al., 2014
Kartchner Caverns, Arizona	Stalactites	NA	66	Orders: Actinomycetales (~75), Rubrobacterales (~60), Acidimicrobiales (~5)	GS	Ortiz et al., 2014
Mizoram, Northeast India	Sediment	NA	65	Genera: <i>Mycobacterium</i> , <i>Corynebacterium</i> , <i>Rubrobacter</i> , <i>Actinoplanes</i> , <i>Saccharothrix</i> , <i>Pseudonocardia</i>	GS	De Mandal et al., 2017
West Lake Cultural Landscape of Hangzhou, China	Stone monuments	NA	87	Genera: <i>Arthrobacter</i> , <i>Curtobacterium</i> , <i>Microbacterium</i>	GS	Li et al., 2018
Dordogne region, France	Biofilms	NA	49	Genera: <i>Arthrobacter</i> , <i>Crossiella</i> , <i>Nocardioides</i> , <i>Lentzea</i>	GS	Alonso et al., 2019
Shuanghe Cave, China	Soil, rock	NA	42–48	<i>Streptomyces</i> (52), <i>Actinoplanes</i> (13), <i>Nocardioides</i> , <i>Crossiella</i>	GS, I	Long et al., 2019

Table 2.1. Continued.

Location	Sample Type	Biofilm Color	Relative Abundance (%)	Most Abundant Taxa (% if reported)	Method(s) Used	Study
Manao-Pee Cave, Kanchanaburi province, Thailand	Soil	NA	~51	Genera: <i>Mycobacterium</i> (15), <i>Streptomyces</i> (12), <i>Nocardioides</i> (9), <i>Marmoricola</i> (5), <i>Crossiella</i> (4), <i>Amycolatopsis</i> (4), and others, including <i>Nocardia</i> , <i>Pseudonocardia</i> , <i>Saccharopolyspora</i> , <i>Actinoplanes</i> , <i>Frankia</i> , <i>Spirillospora</i> , <i>Aeromicrobium</i> , <i>Micromonospora</i> , <i>Rhodococcus</i> , <i>Salinispora</i> , <i>Sporichthya</i> , <i>Saccharothrix</i>	GS, M	Wiseschart et al., 2019
Chongqing, Sichuan, and Yunnan provinces; China	Water, sediment, weathered rock	NA	38	Family: Pseudonocardiaceae	GS	Zhu et al., 2019
Carter Cave, Tennessee, USA	Biofilm	P, W, Y, B	NA	Pseudonocardiales (44), <i>Crossiella</i> (27–52), <i>Solirubrobacter</i> (2)	GS	Frazier, 2020
New Guinea Highlands, Papua Province	Biofilm	Y, W	40	Actinomycetota, Pseudonocardiales, Streptomycetales	GS, M	Turrini et al., 2020
Slovinky, Gelnica, and Hnúšt'a mines, Slovakia	Mining waste, sludge	NA	NA	Genera: <i>Arthrobacter</i> , <i>Streptomyces</i> , <i>Crossiella</i> , <i>Actinomyces</i>	GS, I	Cimermano va et al., 2021
Heshang Cave, China	Sediment, weathered rock	NA	26	Genera: <i>Euzebya</i> , <i>Actinobacterium</i> , <i>Gaiella</i> , <i>Rubrobacter</i> , <i>Solirubrobacter</i> , <i>Crossiella</i>	GS	Ma et al., 2021
Guizhou Province, China	Water, sediment, weathered rock	NA	65	Genera: <i>Streptomyces</i> , <i>Actinomycetes</i> , <i>Crossiella</i> , <i>Arthrobacter</i> , <i>Nocardioides</i> , <i>Nocardia</i> , <i>Microbacterium</i> , <i>Janibacter</i>	GS, M, I	Zhu et al., 2021

Table 2.1. Continued.

Location	Sample Type	Biofilm Color	Relative Abundance (%)	Most Abundant Taxa (% , if reported)	Method(s) Used	Study
La Palma Island, Spain	Biofilm, mineral deposits	Y, W	13	Acidimicrobiia (5), Actinomycetota (13), Nitriliruptoria (3)	GS	Gonzalez-Pimentel et al., 2021
Imawarì Yeuta Cave, Auyan Tepui, Venezuela	Biofilm	G	50–80	Genera: <i>Crossiella</i> , <i>Acidothermus</i> , <i>Mycobacterium</i>	GS	Ghezzi et al., 2021
Mackay Glacier region, South Victoria Land, Antarctica	Soil	NA	~10–40	Order: Solirubrobacterales	M	Ortiz et al., 2021
Zhijin Cave, China	Rock wall, sediment	NA	20–40	Genera: <i>Crossiella</i> (4), <i>Pseudarthrobacter</i> (24), <i>Gaiella</i> (3)	GS	Ai et al., 2022
Nerja Cave, Spain	Drip water, carbonate precipitates	NA	12–19	Actinomycetota, Acidimicrobiia, Thermoleophilia, <i>Streptomyces</i>	GS	Jurado et al., 2022
National Park Kornati, eastern Adriatic Sea, Croatia	Water	NA	22–42	Genera: <i>Aquiluna</i> , <i>Corynebacterium</i> , <i>Cutibacterium</i> , <i>Lawsonella</i>	GS	Kajan et al., 2022
Pindal Cave, Asturias, Spain	Carbonate precipitates	NA	16–36	Genera: <i>Crossiella</i> (1–27), <i>Pseudonocardia</i> (~1), <i>Pseudarthrobacter</i> (~1), <i>Streptomyces</i> (~1)	GS	Martin-Pozas et al., 2022

Qin et al., 2019; Rangseekeaw and Pathom-aree, 2019; Wiseschart et al., 2019), such as those used as pesticides, antibiotics, and anti-tumorals (Lee et al., 2020; Silva et al., 2020), as well as because of their contributions to soil health in agriculture and forestry (Servin et al., 2008) by serving as major drivers of carbon, nitrogen, phosphorus, and potassium cycling (Hill et al., 2011; Zhang et al., 2019). Members of this phylum have diverse ways of degrading, or breaking down, organic molecules as heterotrophs (Anandan et al., 2016). For example, *Streptomyces* metabolize a range of organic compounds and, in so doing, produce secondary metabolites that act as insecticides, fungicides, and bactericides (Olanrewaju and Babalola, 2019). In surface environments and soils, Actinomycetota possess key enzymes that can help growth in nutrient-limited environments (Bao et al., 2021), like nitrogenase used for N₂ fixation and antibiotic synthase for suppressing microbial competition, as well as carbohydrate-active enzymes (CAZy) used for carbohydrate degradation. However, the function and ecological roles of Actinomycetota in extreme environments, like caves and other subsurface habitats, needs to be studied because living in the subsurface or other oligotrophic habitats likely requires even more novel metabolisms compared to if they were growing in surface habitats.

Most genomic and metagenomic studies of Actinomycetota have been from arid and cold environments (Table 2.1) (e.g., Ray et al., 2022), which calls into question whether similar metabolisms can be expected for Actinomycetota in caves. Recently, gene sequencing has revealed that some Actinomycetota can use CO, H₂, different nitrogen compounds, including atmospheric N₂, and other trace gases to obtain their energy (e.g., Ortiz et al., 2014; Liu et al., 2018; Zhao et al., 2018; Bay et al., 2021; Martin-Pozas et al., 2022; Narsing Rao and Li, 2022; Ray et al., 2022). Shivelata and Tulasī (2015) describe members from the orders Frankiales and Streptomycetales that fix atmospheric N₂ by using CO as an electron donor. Ray et al. (2022) suggest that some members of the Actinomycetota have RuBisCO and high-affinity hydrogenases, which could mean that they have the flexibility to switch between using H₂ and CO as electron donors. Other extremophilic bacteria, such as *Carboxydotherrmus hydrogenoformans*, use CO uptake to catalyze CO₂ through the reductive acetyl-CoA pathway (Wu et al., 2005; Ruickoldt, 2023).

In caves, there is evidence that trace gas uptake by microbes is responsible for the consumption of subsurface CO₂ and CH₄. For instance, Martin-Pozas et al. (2022) investigate

CO₂ fixation and CH₄ fluxes, as well as the genetic potential for these pathways in microbial communities from Pindal Cave, Spain. Methanotrophic Pseudomonadota consume CH₄, but CO₂ uptake, either from cave air or as the byproduct of methanotrophy, was tied to the genus *Crossiella* (Actinomycetota, Pseudonocardiales, Pseudonocardiaceae) that had RuBisCO and CA enzymes, as well as lived in association with moonmilk precipitation, a hydrated carbonate mineral deposit typically found on moist cave walls. Such a link between moonmilk and the presence of *Crossiella* has previously been hypothesized (Riquelme et al., 2015; Gonzalez-Pimentel et al., 2018; Maciejewska et al., 2018; Gonzalez-Pimentel et al., 2021), but the suspected association relied on predicting the presence of functional genes in the microorganisms based on only their 16S rRNA gene sequence data. Therefore, the microbes responsible for these pathways, and confirmation that these pathways are actively being performed in the caves, have not been linked genomically for any of the cave-wall biofilm taxa. Application of shotgun metagenomics could circumvent these limitations by linking the taxa to their functional potential.

The order Pseudonocardiales contains a single paraphyletic family, the Pseudonocardiaceae. All known members of the family are aerobes. There are currently 27 described species, all of which had previously been classified within the order Actinosynemmataceae (Park et al., 2008; Franco and Labeda, 2014; Nouioui et al., 2018). Nouioui et al. (2018) suggest that the Pseudonocardiales should be further expanded to include the families Actinopolysporaceae and Mzabimycetaceae that currently belong to the order Actinosporales. Franco and Labeda (2014) report that most members within the Pseudonocardiales are soil-dwelling heterotrophs that use alcohols, disaccharides, monosaccharides, and polysaccharides for growth and oxygen as a terminal electron acceptor, although some genera form symbiotic and pathogenic relationships with insects, plant roots, and animals that may influence their metabolism.

Most of what is known about members of the Pseudonocardiaceae comes from chemotaxonomic studies of species that were isolated and characterized on growth media (e.g., Borsodi et al., 2012; Franco and Labeda, 2014; Cao et al., 2015; Cao et al., 2017; Fang et al., 2017; Rangseekaew and Pathom-aree, 2019). But, even from this work, their metabolic functions are not well understood in oligotrophic environments, like caves, and despite their prevalence within microbial communities (Table 2.1). Species within the genus *Pseudonocardia* are the

most extensively studied and have been identified as a “key indicator genus” within in cave and desert environment communities, as well as from cave paintings and subsurface tombs (Table 2.1) (e.g., Lynch et al., 2014; Zhu et al., 2021). Similarly, the genera *Crossiella* and *Lentzea* are known from caves (e.g., Borsodi et al., 2012; Franco and Labeda, 2014; Cao et al., 2015; Cao et al., 2017; Fang et al., 2017; Rangseekaew and Pathom-aree, 2019), along with the genus *Kutzneria*, which has also been found in carbonate fissures in permafrost regions (Demirci and Ordu, 2020). *Kutzneria* can produce several bioactive compounds (Jakubiec-Krzesniak et al., 2018). Other Pseudonocardiaceae genera identified or described from caves include *Actinoalloteichus*, *Allokutzneria*, *Actinokineospora*, *Actinosynnema*, *Actinomyces*, *Actinophytocola*, *Alloactinosynnema*, *Amycolatopsis*, *Kibdelosporangium*, *Prauserella*, *Saccharomonospora*, *Saccharopolyspora*, *Saccharothrix*, *Sciscionella*, *Streptoalloteichus*, and *Thermocrisum* (e.g., Schabereiter-Gurtner et al., 2004; Niyomvong et al., 2012; Tomczyk-Żak and Zielenkiewicz, 2016; Fang et al., 2017; Wiseschart et al., 2019; Jurado et al., 2020; Turrini et al., 2020; Katti et al., 2021; Selensky et al., 2021).

From genome predictions according to 16S rRNA gene surveys of *Pseudonocardia* spp. sampled from the Atacama Desert, Chile, Lynch et al. (2014) predict that *Pseudonocardia* are capable of aerobic heterotrophy, nitrate and ammonia utilization, H₂ and CO oxidation, CO₂ fixation, and methanotrophy. These findings agree with other studies for different *Pseudonocardia* species. Specifically, *P. sulfidoxydans*, *P. benzenivorans*, *P. chloroethenivorans*, *P. dioxanivorans*, and *P. nitrificans* use various organic carbon sources for energy (Lee et al., 2004; Mahendra and Alvarez-Cohen, 2005; Kämpfer et al., 2006; Franco and Labeda, 2014). Other species perform CO₂ fixation by oxidizing H₂, including *P. autotrophica*, *P. petroleophila*, and *P. saturnae* (Franco and Labeda, 2014). Park et al. (2008) also describe a novel strain of *P. carboxydivorans* that oxidizes CO. One species, *P. ammonioxydans*, is a facultative mixotroph dependent on nitrogen (Franco and Labeda, 2014).

To date, *Crossiella equi* and *Crossiella cryophila* are the only cultured species within this genus. However, other 16S rRNA gene sequences with high sequence identities to *C. equi* and *C. cryophila* dominate microbial communities in caves, from weathered rock, subaerial biofilms, and paintings, as well as biofilms on weathered monuments (Table 2.1) (e.g., Li et al., 2018; Alonso et al., 2019; Zhu et al., 2019; Frazier, 2020; Cimermanova et al., 2021; Ghezzi et al.,

2021; Ma et al., 2021; Xie et al., 2021; Zhu et al., 2021). In a study investigating the relationship between microbial ecology and moonmilk, Li et al. (2018) hypothesize that *Crossiella* spp. may be directly responsible for moonmilk formation in caves because the genus comprised 84% of the moonmilk community based on relative abundances of sequence reads. Gonzalez-Pimentel et al. (2021) also identify the enzyme urease from moonmilk, which uses amine and amide groups to produce ammonia (Low, 1996), a metabolic process that increases the pH of the immediate environment and would make moonmilk deposition favorable (Maciejewksa et al., 2017). Frazier (2020) suggests that *Crossiella*-like groups might also influence carbonate precipitation and affect the carbon isotope composition of secondary carbonates (e.g., speleothems).

The Pseudomonadota

The Pseudomonadota are morphologically and metabolically diverse, with mixed tolerance to, and use of, oxygen (Kosznik-Kwaśnicka et al., 2022). This phylum is comprised of the classes Alphaproteobacteria, Betaproteobacteria, and Gammaproteobacteria. Despite their metabolic versatility, some members of the class Alphaproteobacteria are known for nitrogen and methane metabolism, and occupy cave habitats (e.g., Riquelme et al., 2015; Tsoy et al., 2016; Lavoie et al., 2017; Kumaresan et al., 2018; Zhu et al., 2019; Frazier, 2020; Jurado et al., 2020; Adesso et al., 2021). Some members within the order Hyphomicrobiales (formerly known as the Rhizobiales) fix N₂ (Carvalho et al., 2010) while others oxidize methane (Guerrero-Cruz et al., 2021). Similarly, the class Gammaproteobacteria includes taxa that are important for CO₂ fixation using the CBB cycle or the rTCA cycle with sulfur compounds or nitrate as electron acceptors, like in coastal sediments (e.g., Dykema et al., 2016; Vasquez-Cardenas et al., 2020; Vineis et al., 2021). But, other members of this class can also obtain their carbon and energy through methanotrophy, such as the family Methylococcaceae (Bowman, 2014). Within caves, Gammaproteobacteria like Methylococcaceae, as well as other taxa, can be prevalent (e.g., Chen et al., 2009; Brannen-Donnelly and Engel, 2015; Lü et al., 2018; Zhao et al., 2018; Ojeda et al., 2019; Jurado et al., 2020; Adesso et al., 2021; Zhu et al., 2021).

Geochemistry as a Tool to Trace Metabolism and Ecological Function

Not only does geochemistry play an important role in the composition of a microbial community and the relative abundances of community members (e.g., Tolar et al., 2020; Zada et al., 2021), but the geochemistry of an environment or habitat can also hint to which metabolisms

are present (e.g., Selensky et al., 2021). This is because ecological functions and specific microbial metabolisms often leave geochemical footprints. Stable carbon and nitrogen isotope ratios are among the most widely used geochemical indicators of microbial processes (e.g., Blair et al., 1985; Ryabenko, 2013; Miyamoto et al., 2022). Carbon isotopes can be used to interpret heterotrophic and autotrophic contributions to carbon pools within a system (e.g., Reichardt et al., 2011; Adrian and Marco-Urrea, 2016; Wiseschart et al., 2019). In cave and karst settings, specifically, carbon isotopes have been used to evaluate food webs and key autotrophic functions, such as CO₂ fixation and CH₄ oxidation, from distinct isotope discrimination against the heavy carbon isotope, ¹³C (e.g., Engel et al., 2004; Hutchins et al., 2016; Frazier, 2020; Martin-Pozas et al., 2020; Selensky et al., 2021; Nicolosi et al., 2022). In standard delta (δ) notation, CO₂ fixation through the CBB cycle can produce biomass with a δ¹³C value that is up to 20 to 30 per mil (‰) lighter (i.e., having more negative values) than the initial carbon source, whereas carbon use from methanotrophy can result in biomass isotope compositions that are up to 80‰ lighter (e.g., Penger et al., 2012; Selensky et al., 2021). Heterotrophic carbon metabolisms result in δ¹³C values are generally within 3‰ of the dissolved organic carbon (DOC) source used (Hayes, 2001; Boschker and Middelburg, 2002). In contrast, use of nitrogen isotopes as tracers is challenging because there may be low concentrations of nitrogen species to measure in an environment and because various biotic and abiotic processes affect the isotopic composition of nitrogen species differently (e.g., Ryabenko, 2013). Nevertheless, in the previous of study of Carter Cave, Tennessee, cave-wall biofilms dominated by Pseudonocardiaceae (Frazier, 2020), bulk δ¹³C analyses from the biofilms indicate chemolithoautotrophic CO₂ fixation, despite the cooccurrence of other heterotrophic microbial groups. However, although provocative, the methods used in the Frazier (2020) study were based on 16S rRNA gene surveys from the same samples used for the geochemical analyses, and metabolism was inferred based on putative functions from the literature. A deeper way of assessing metabolic function within the biofilms was needed and serves as the foundation for the metagenomic approach taken in thesis research.

SECTION THREE

MATERIALS AND METHODS

Site Description

This study used samples previously collected from Carter Cave (state cave KN14), Tennessee, in 2020 (Frazier, 2020). The cave formed in interbedded limestone and dolomite of the Kingsport Formation (Hardeman et al., 1966). A stream flows through the cave that can introduce organic carbon to the cave walls during floods. During the Frazier (2020) study, the cave was well ventilated and had overall air temperatures that averaged 11.46 °C at nearly 100% relative humidity. The percent moisture of cave-wall surfaces ranged from 19% to 84%. Biofilms near the stream and around cave drips had higher moisture content. Samples collected from site 1 were approximately 0.9 m above the cave stream and from a biofilm of pink, white, and blue colonies, whereas samples from site 4, which was approximately 2.15 m above the stream and near site 1, had white and yellow colonies. Percent moisture varied, with site 1 biofilms having a reading of 44% and site 4 of 19%. The temperatures at both locations were 9.7°C and 10.5°C, respectively. The $\delta^{13}\text{C}$ of carbonates from sites 1 and 4 site were -9.03‰ and -8.88‰, values indicative of microbial induced calcite precipitation (MICP; Lü et al., 2018). Overall, $\delta^{13}\text{C}_{\text{CO}_2}$ values for Carter Cave air was -14‰. The samples from these sites were chosen because of their isotopic compositions and dominance of Pseudonocardiales (i.e., having more than 50%, based on relative abundances, of retrieved 16S rRNA gene sequences). The site 1 biofilms had $\delta^{13}\text{C}$ values that averaged $-42.45 \pm 0.58\text{‰}$ ($n = 3$) and a $\delta^{15}\text{N}$ value of -1.02 ‰, whereas site 4 biofilms had $\delta^{13}\text{C}$ values of $-38.30 \pm 1.18\text{‰}$ ($n = 3$) and a $\delta^{15}\text{N}$ value of -2.55‰ (Frazier, 2020).

DNA Extraction and 16S rRNA Gene Sequence Analysis

DNA was extracted from the site 1 and site 4 cave-wall biofilm samples of Frazier (2020). For Carter Cave, DNA was collected from biofilms using a combination of three swabs and one scrape for each site. Samples that were obtained through scrapes were derived from multiple biofilm features, including the underlying substrate (e.g., mud, rock, etc.), while a swab was used to sample a particular biofilm feature from the top surface. DNA was extracted from swabs and scrapes separately. Briefly, DNeasy PowerSoil DNA Isolation kits were used, following manufacturer instructions (Qiagen, Hilden, Germany), with modifications to lyse Gram-positive cells (i.e., Actinomycetota) by incubating lysate at 70°C for 10 minutes prior to

the bead beating. The Illumina MiSeq platform was used to sequence the V4 region of the 16S rRNA gene with the primers 515F and 806R (Molecular Research LP, MRDNA, Shallowater, TX). The resulting paired-end raw sequences were unbinned, and primers and linker sequences were removed using FASTqProcessor software (MRDNA, Shallowater, TX).

Frazier (2020) used the SILVA v.138 database, created by the SILVA ribosomal RNA gene database project, to classify the taxonomy of 16S rRNA gene sequences (Quast et al., 2013), but this reference file no longer reflects recent changes and updates to microbial taxonomy (Oren et al., 2022). Therefore, the 16S rRNA gene sequences from Frazier (2020) from sites 1 and 4 (i.e., 8 total samples) were reprocessed using the updated (as of 27 Aug 2020) SILVA reference file v.138.1. Sequences were acquired from the NCBI (<http://www.ncbi.nlm.nih.gov/bioproject/PRJNA649551>). The computer program MOTHUR v.1.44 was used, following the standard operating procedure (Schloss et al., 2009; Schloss, 2020) specific for MiSeq technology. Briefly, R1 and R2 files were combined into contigs, then sequences that were the length errors and/or had ambiguous bases were removed using the pre.cluster command, followed by the chimera.vsearch command to check for chimera, which were also removed. To assign taxonomy, the classify.seqs command used the SILVA v.138.1 reference file. The dist.seqs command was used to calculate pairwise distances between aligned sequences, followed by the cluster command to assign each operational taxonomic unit (OTU) to a taxonomic rank at 97% sequence identity using the classify.otu command. Amplicon sequence variants (ASVs) were assigned to a taxonomic rank using the same command, by replacing the 0.03 cutoff with the ASV label, as described by Schloss et al. (2020), to identify exact sequences. Microbial abundance plots for phylum, order, family, and genus ranks of the Carter Cave samples were created using the taxonomy and OTU files. The command rarefaction.single was used to calculate Shannon and Simpson indices for alpha-level diversity.

Shotgun Metagenomic Sequencing

Frazier (2020) found no significant differences between microbial community composition and biofilm colors for either the swabs and scrapes. Therefore, metagenomes were sequenced from a composite sample from site 1 (i.e., three swabs and one scrape) DNA and from site 4 DNA (i.e., three swabs and one scrape) at MRDNA using the same DNA extracted from Carter Cave (Frazier, 2020). These samples also had low alpha diversity, with Shannon's

Diversity indices of 2.73 to 2.83, and Simpson's indices of about 0.72. Collectively, the results suggested that the samples were good candidates to recover Pseudonocardiales MAGs.

Metagenome shotgun sequencing libraries were prepared using a Tagmentation library preparation kit (Illumina), following the manufacturer's user guide. A Qubit® dsDNA HS Assay Kit (Life Technologies) was used to measure the initial concentrations of DNA, after which the evaluation was made that concentrations were too low (i.e., 14.10 ng/μl for site 1 and 12.10 ng/μl for site 4). Whole genome amplification was carried out by using Qiagen a REPLI-g Midi kit, following manufacturer instructions, and the resulting linear amplified DNA was cleaned using a Qiagen DNEasy PowerClean Pro Cleanup Kit, following manufacturer instructions, before sequencing using the Illumina Novaseq 6000 platform to generate up to thirty million reads per sample with paired-end read lengths of 2x150bp.

Metagenome Assembly, Binning, and Gene Calling

Metagenome libraries were analyzed using the Department of Energy's data science platform Systems Biology Knowledgebase (KBase) and following the "Genome Extraction from Shotgun Metagenome Sequence Data" narrative (Arkin et al., 2018). Reads were cropped of adapters to purify the FASTQ files using Trimmomatic v0.36 (Bolger et al., 2014), which also removed reads that had PHRED scores of 15 or lower. The quality of each file was assessed using FastQC v0.11.9 (Andrews, 2010). Because Pseudonocardiales dominated both site 1 and site 4 samples, Merge Reads Libraries v1.0.1 (Arkin et al., 2018) was used to combine all of the libraries to increase coverage.

Reads were co-assembled separately using IDBA-UD v1.1.3 (Peng et al., 2012), metaSPAdes v3.15.3 (Prjibelski et al., 2020), and MEGAHIT v1.2.9 (Li et al., 2015), where each output was individually evaluated for contig length using the Compare Assembled Contig Distributions v1.1.2 application (Arkin et al., 2018). metaSPAdes generated the longest contig, at 195,228 bp, and from 19 contigs also produced the largest assembly of 2,442,036 bp in length. Therefore, from this metaSPAdes assembly, contigs were then clustered into bins using three separate methods, MaxBin2 v2.2.4 (Wu et al., 2015), MetaBAT2 Contig Binning v1.7 (Kang et al., 2015) and CONCOCT v1.1 (Alneberg et al., 2014). The results of each binning method were run through DAS Tool v1.2 (Sieber et al., 2018) to optimize and integrate the results of each binning method into a final binning output. Average Nucleotide Identity (ANI) values were

determined by making pairwise comparisons of orthologous gene pairs between MAGs and other members of the Pseudonocardiaceae using fastANI (Jain et al., 2018).

The completion and contamination of each binning method, including DAS Tool, was assessed via CheckM v1.0.18 (Parks et al., 2015). The Das Tool binning optimization resulted in a total of 6,629 contigs grouped into 11 bins (Figure 3.1), which were filtered to select for nearly complete bins, using the criteria of Parks et al. (2015). This selection yielded four bins with >90% completeness: Bin.005, Bin.008, Bin.010, and Bin.011 (Tables 3.1 and 3.2). Regarding the Minimum Information about a Metagenome-Assembled Genome (MIMAG) standard of Bowers et al. (2017), Bin.008 and Bin.011 were recognized as high quality. Although Bin.005 and Bin.010 would normally be recognized as medium quality because contamination was slightly above 5%, but because completeness was greater than 90%, both bins were used because they could provide important information for this study.

Initial taxonomic classifications were done for the four high-quality MAGs using GTDB-Tk Classify v1.7.0 (Chaumeil et al., 2019) in KBase using single-copy phylogenetic markers and prior to reannotation following methods outlined by the Annotate Microbial Genome application in RASTtk to make gene predictions via Prodigal (Hyatt et al., 2010) and adding SEED (Overbeek et al., 2014) functions and annotation to protein-encoding genes. After determining taxonomic affiliations, SpeciesTree v2.2.0 was used to visualize phylogenetic relatives to each MAG using a set of 49 core universal genes defined by clusters of orthologous groups (COG) gene families (Arkin et al., 2018).

The annotated MAGs were evaluated using the Annotate and Distill Assemblies with DRAM application (Shaffer et al., 2020) to view metabolic summaries and functions of interest. MAGs that had complete sets of genes for multiple carbon pathways identified through DRAM were further evaluated in KEGG to visualize complete metabolic pathways as well as Enzyme Commission (E.C.) numbers belonging to each pathway (Kanehisa et al., 2023).

Sequence Deposition and Data Availability

The 16S rRNA gene sequences obtained by Frazier (2020) and reanalyzed in this study were deposited under the NCBI BioProject accession number PRJNA649551 (<http://www.ncbi.nlm.nih.gov/bioproject/PRJNA649551>). MAGs resulting from this study, using the same BioSamples, are deposited under the same BioProject number.

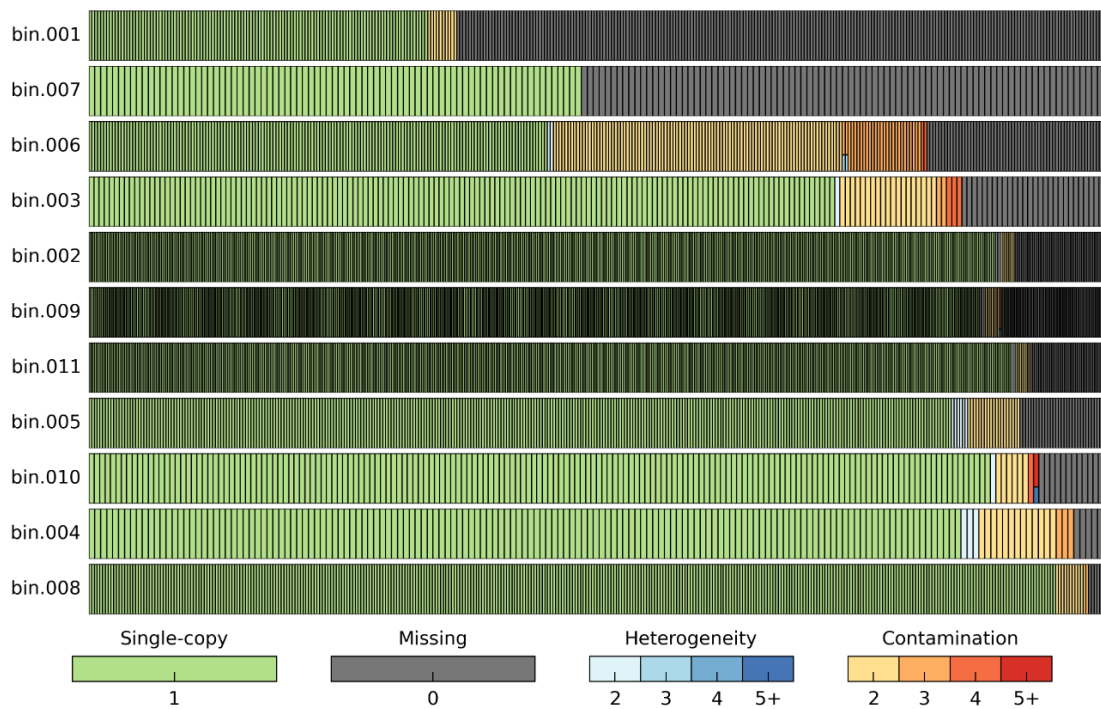


Figure 3.1: CheckM quality tests of the Das Tool binning output in KBase (Arkin et al., 2018).

Table 3.1: Bins, with completeness and contamination noted, along with taxonomic affiliations. Bin quality is based on Parks et al. (2015). “*” indicates exceptions that were made to the criteria of Bowers et al. (2017).

Bin	Taxonomy	N50	N75	Completeness	Contamination	Quality
Bin.001	Phylum: Pseudomonadota Class: Alphaproteobacteria Order: Dongiales Family: - Genus: -	4,164	2,391	30.60%	2.77%	Low Quality
Bin.002	Phylum: Pseudomonadota Class: Gammaproteobacteria Order: JACCXJ01 Family: JACCXJ01 Genus: USC γ -Taylor	20,429	9,347	88.34%	1.23%	Medium Quality
Bin.003	Phylum: Actinomycetota Class: Thermoleophilia Order: Solirubrobacterales Family: Solirubrobacteraceae Genus: -	8,178	3,996	85.99%	17.34%	Low Quality
Bin.004	Phylum: Pseudomonadota Class: Gammaproteobacteria Order: Xanthomonadales Family: Xanthomonadaceae Genus: -	18,334	9,134	86.58%	17.09%	Low Quality

Table 3.1, Continued.

Bin	Taxonomy	N50	N75	Completeness	Contamination	Quality
Bin.005	Phylum: Actinomycetota Class: Actinomycetia Order: Pseudonocardiales Family: Pseudonocardiaceae Genus: GCA-003244245	36,243	17,110	91.89%	6.39%	High Quality*
Bin.006	Phylum: Pseudomonadota Class: Alphaproteobacteria Order: DSM-16000 Family: Inquilinaceae Genus: <i>Inquilinus</i>	4,450	3,253	82.71%	51.07%	Low Quality
Bin.007	Phylum: Pseudomonadota Class: Alphaproteobacteria Order: Rhodospirillales Family: Reyranellaceae Genus: -	6,800	4,280	44.07%	0.00%	Low Quality
Bin.008	Phylum: Actinomycetota Class: Actinomycetia Order: Pseudonocardiales Family: Pseudonocardiaceae Genus: GCA-003244245	58,690	30,353	98.22%	3.83%	High Quality
Bin.009	Phylum: Pseudomonadota Class: Gammaproteobacteria Order: Xanthomonadales Family: Xanthomonadaceae Genus: -	11,167	7080	68.84%	1.35%	Medium Quality
Bin.010	Phylum: Acidobacteriota Class: Blastocatellia Order: RBC074 Family: - Genus: -	24,995	12,772	91.98%	8.12%	High Quality*

Table 3.1, Continued.

Bin	Taxonomy	N50	N75	Completion	Contamination	Quality
Bin.011	Phylum: Pseudomonadota Class: Gammaproteobacteria Order: JACCXJ01 Family: - Genus: -	29,652	16,929	90.18%	2.16%	High Quality

Table 3.2: Summary and statistics of the four highest quality bins that were generated and extracted from the Das Tool binning optimization tool.

Nearly Complete Bins	Length (bp)	GC content	Number of Contigs	Number of Features
Bin.005	5,078,116	66.86%	161	5,651
Bin.008	7,098,898	68.62%	337	8,534
Bin.010	10,790,763	59.22%	730	10,356
Bin.011	4,349,206	63.69%	252	4,537

SECTION FOUR

RESULTS

Updated 16S rRNA Microbial Community Analysis

A total of 10 phyla from the Carter Cave biofilms had relative abundances above 1%, (Figure 4.1). The top four most abundant phyla included Actinomycetota (ranging from 46% to 63% relative abundance), Pseudomonadota (19–30%), Acidobacteriota (1–11%), Chloroflexota (1–5%), and Planctomycetota (1–4%). At the order level, the Pseudonocardiales were the most abundant taxa (47–55%), followed by Hyphomicrobiales (6–8%) and Steroidobacterales (1–8%), as well as unidentified members of Blastocatellales, Steroidobacteraceae, and Inquilinaceae (Figure 4.2). Pseudonocardiaceae (Figure 4.3) and *Crossiella* (Figure 4.4) were the most abundant family and genus, respectively, within their taxonomic ranks. Other abundant members at the genus level included unidentified members of the Blastocatellales, *Inquilinus*, *Steroidobacter*, and unclassified bacteria (Figure 4.4).

Because of their prevalence, *Crossiella* sequences were analyzed separately to determine ASV-level (i.e., identical sequences) relationships. In total, there were 142 ASVs with more than a single read that belonged to *Crossiella*. Of these, ASV001 made up 98.2% of all *Crossiella* sequences. The number of *Crossiella* ASVs indicated that there is strain level diversity, which may contribute to the retrieval of similar but functionally distinct Pseudonocardiaceae MAGs.

Taxonomic Classification and Phylogenetic Placement of MAGs

The taxonomy of the four high-quality MAGs was evaluated according to GTDB-Tk classification and phylogenetic tree assignments. Two MAGs (Bin.005 and Bin.008) belonged to the phylum Actinomycetota, order Pseudonocardiales, and family Pseudonocardiaceae (Table 3.1). Although both MAGs belonged to the same family, Bin.008 contained approximately two million more base pairs than Bin.005 and had less contamination. One MAG (Bin.010) was affiliated with the Acidobacteriota phylum, class Blastocatellia, and was identified as an uncultured isolate (RBC074) from a municipal wastewater treatment plant in Guelph, Ontario, Canada (Spasov et al., 2020) (Table 3.1). Lastly, the fourth high-quality MAG (Bin.011) was identified as a member of the Pseudomonadota phylum, class Gammaproteobacteria, and was most closely related to the novel putative methanotrophic order JACCXJ01 that was observed in Antarctic desert soils by Ortiz et al. (2021) (Table 3.1).

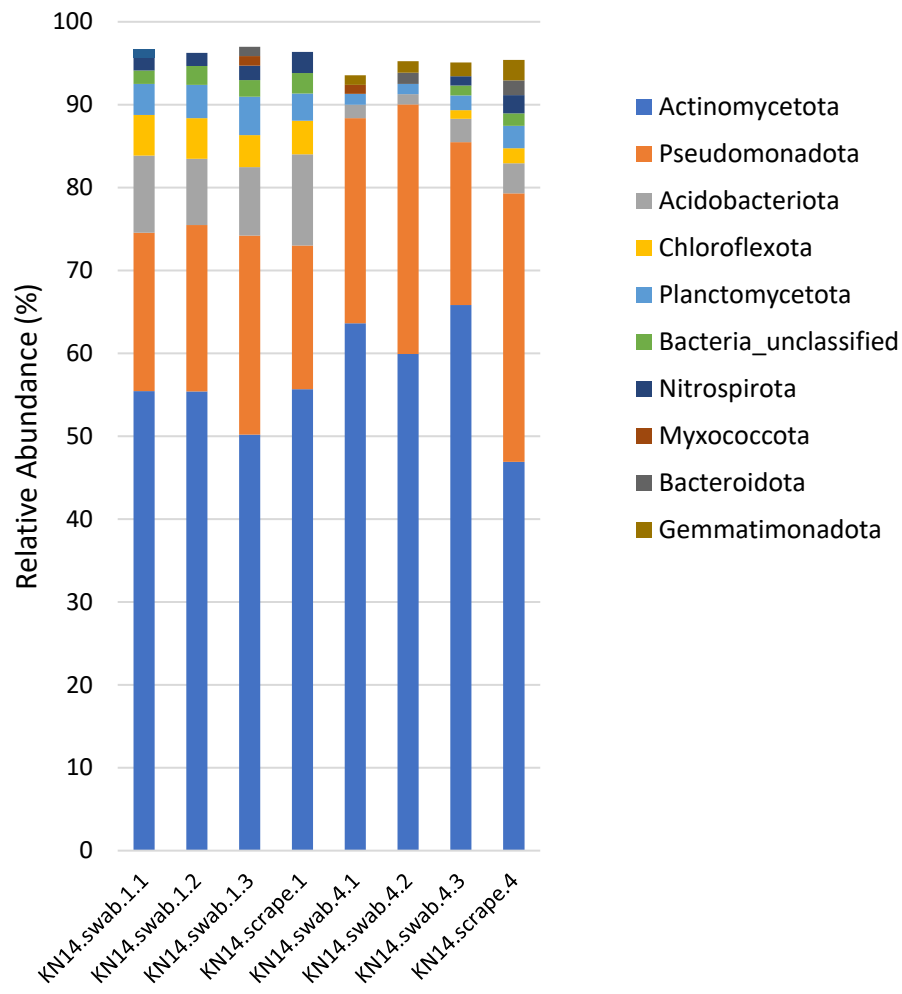


Figure 4.1: Relative abundance of taxa at the phylum level for Carter Cave sites 1 and 4, from samples collected by Frazier (2020) as swabs or wall scrapes. Only taxa with relative abundances greater than 1.0% are shown.

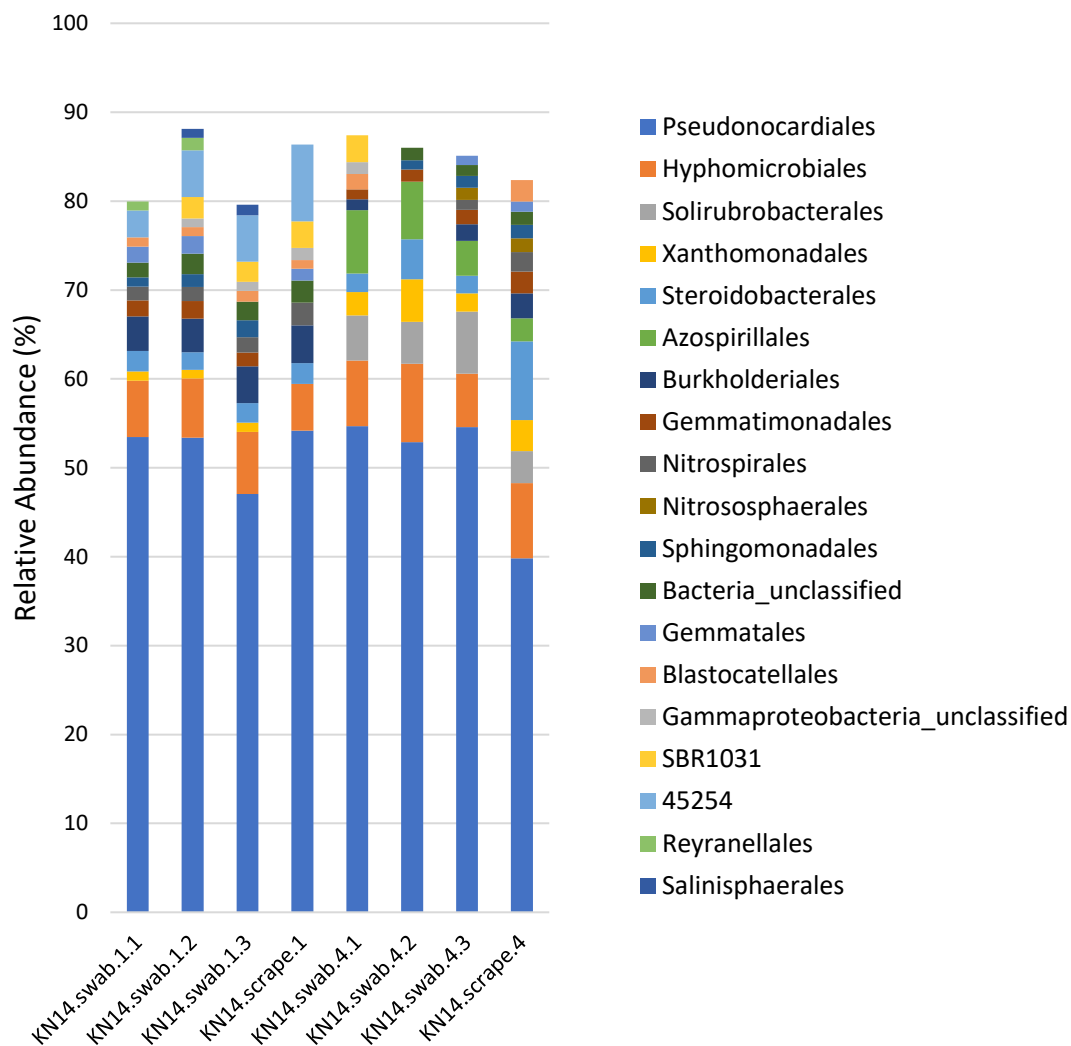


Figure 4.2: Relative abundance of taxa at the order level for Carter Cave sites 1 and 4, from samples collected by Frazier (2020) as swabs or wall scrapes. Only taxa with relative abundances greater than 1.0% are shown.

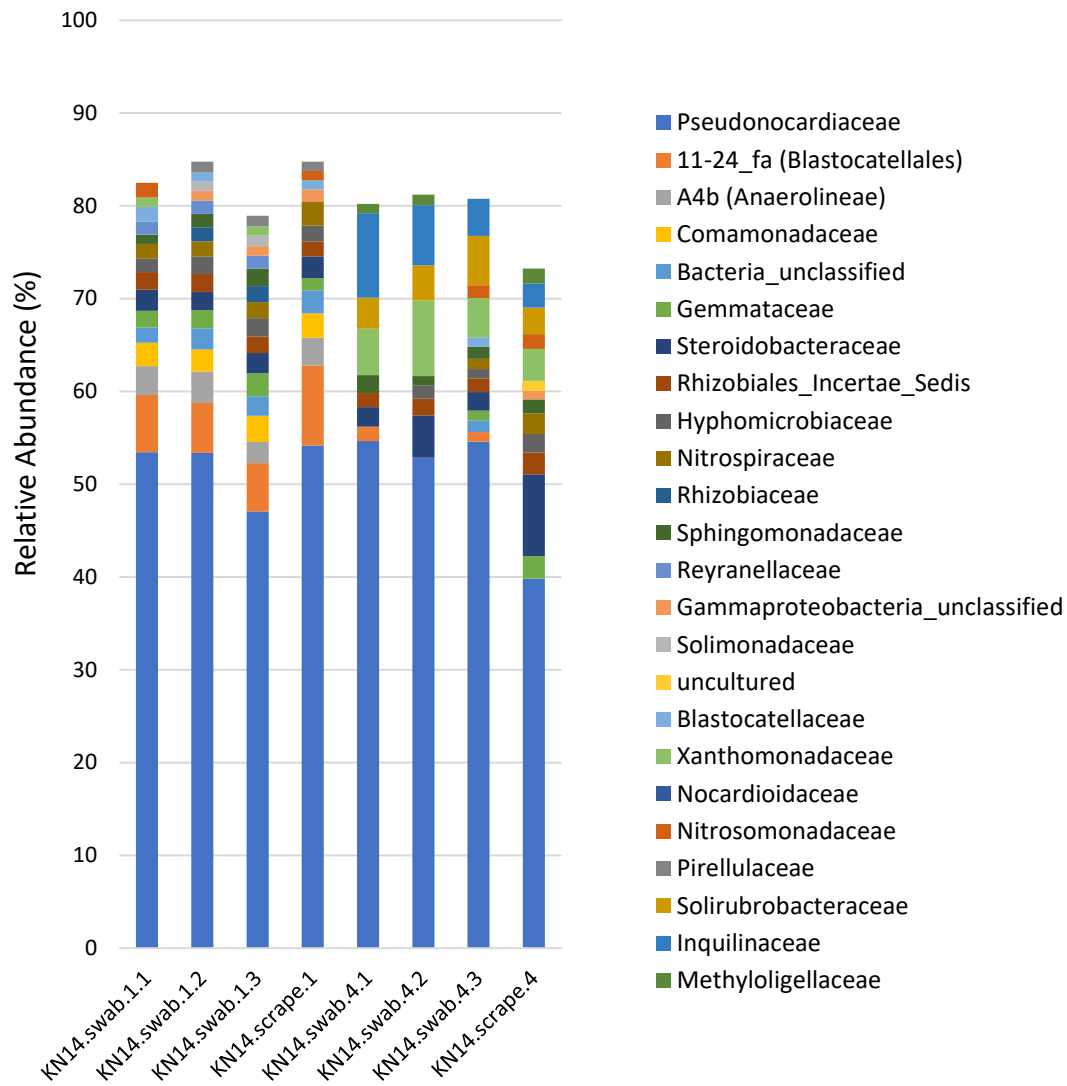


Figure 4.3: Relative abundance of taxa at the family level for Carter Cave sites 1 and 4, from samples collected by Frazier (2020) as swabs or wall scrapes. Only taxa with relative abundances greater than 1.0% are shown.

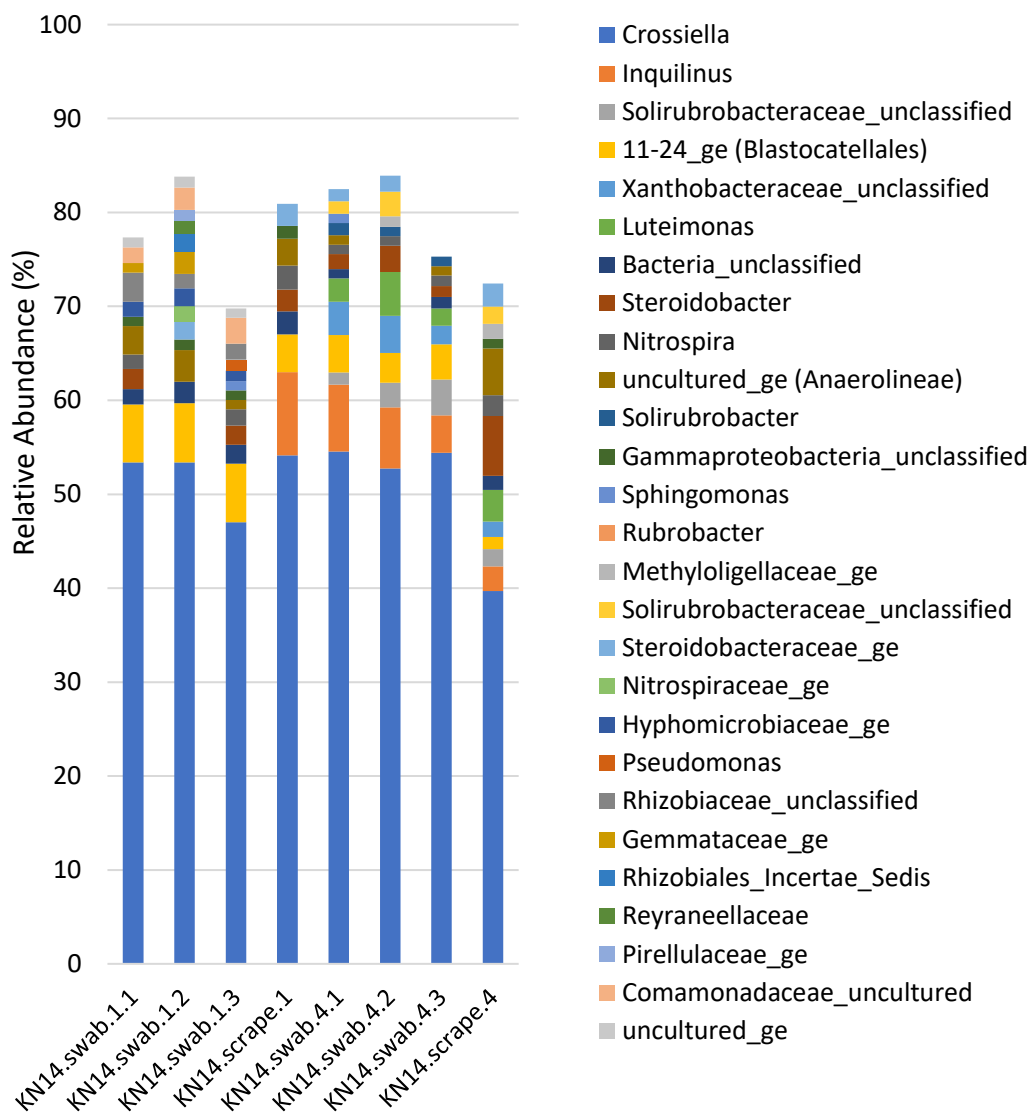


Figure 4.4: Relative abundance of taxa at the genus level for Carter Cave sites 1 and 4, from samples collected by Frazier (2020) as swabs or wall scrapes. Only taxa with relative abundances greater than 1.0% are shown.

Phylogenetic trees for both Pseudonocardiaceae MAGs were made based on a set of 49 core universal genes for COG gene families (Figures 4.5 and 4.6). *Actinoalloteichus hoggarensis* was the closest relative to both MAGs. However, ANI values revealed that the closest relative to Bin.005 and Bin.008 MAGs as *Gandjariella thermophila* at 77.2% and 77.8%, respectively, whereas the ANI values compared to *Actinoalloteichus hoggarensis* were 76.4% and 76.2%, respectively (Table 4.1). Using the cutoff criteria of Xu et al. (2021), the values suggested that both MAGs showed enough genetic relatedness to be categorized within the same family. The Bin.010 Blastocatellia MAG was closely related to *Pyrinomonas methylaliphatogenes* (Figure 4.7), with an ANI value of 75.51%. The Bin.011 Gammaproteobacteria MAG's closest relative was *Nitrosococcus halophilus* (Figure 4.8); however, the ANI value between the MAG and *N. halophilus* genome was 54.41%, indicated order-level differences.

Functional Annotations

Carbon Metabolism

Metabolic functions were annotated for each of the high-quality MAGs using the RastTK application and visualized using DRAM (Figure 4.9) and KEGG Mapper (Figures 4.10–4.13). Both Pseudonocardiaceae MAGs (Bin.005 and Bin.008) had similar potential functions. According to DRAM, both MAGs contained all sets of genes necessary for the RPP (or CBB) pathway and rTCA cycles, six out of seven genes necessary to perform the pentose phosphate cycle for heterotrophy, and seven out of nine genes necessary for glycolysis. Bin.005 also had 50% of the genes needed for the 3HP cycle. KEGG Mapper indicated that both bins only had a complete RPP pathway (Figures 4.10 and 4.11), and that Bin.005 also had multiple steps missing during the conversion of 2-oxoglutarate to succinyl-CoA in the TCA cycle (Figure 4.12), compared to Bin.008 that had a nearly complete rTCA cycle (Figure 4.13). The Blastocatellia MAG (Bin.010) had no complete carbon pathways (Figure 4.9). The Gammaproteobacteria MAG (Bin.011) had genes for the pentose phosphate pathway and the TCA cycle, as well as three-quarters of the genes necessary for the glyoxylate cycle, and rare genes needed for the 3HP cycle and the reductive acetyl-CoA pathway (two of seven genes). All four MAGs had genes that encoded for acetyl-CoA synthetase, an enzyme that facilitates the conversion of acetate to methane during methanogenesis (Figure 4.14) (Kurade et al., 2019). The Pseudonocardiaceae Bin.008 and Gammaproteobacteria Bin.011 had genes encoding for acetate kinase, typically

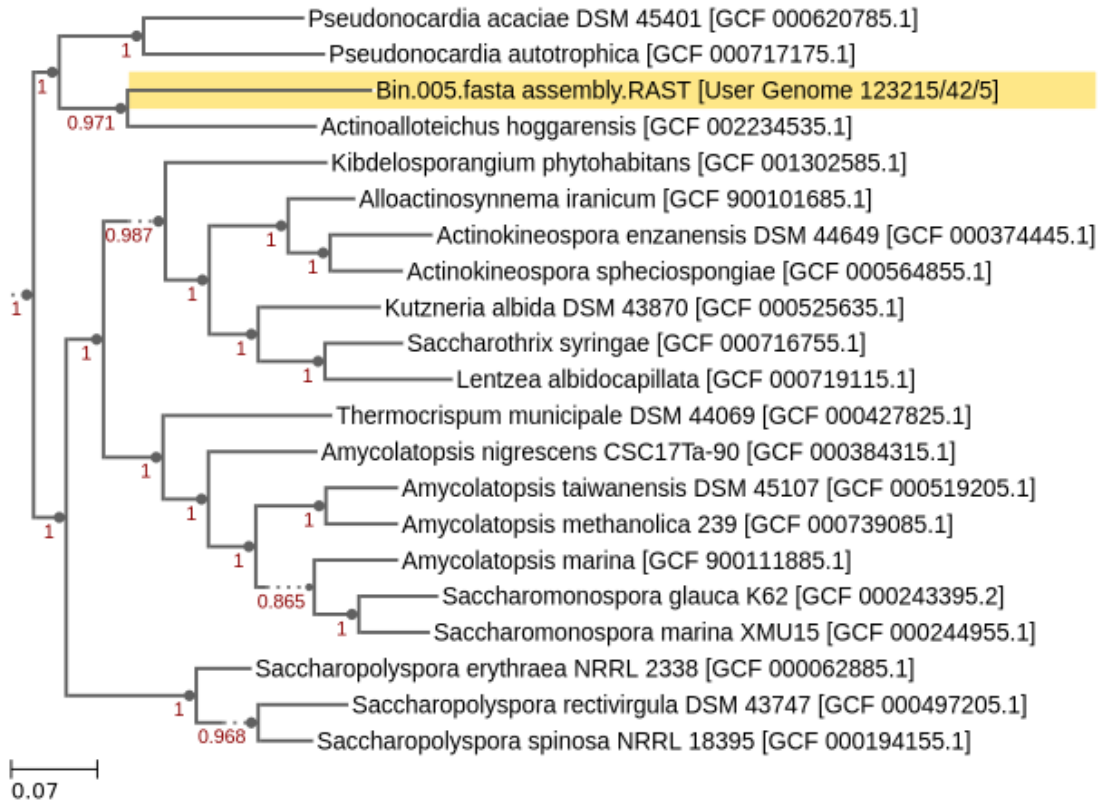


Figure 4.5: A phylogenetic tree created using Species Tree v2.2.0 for the MAG resulting from Bin.005, which was most closely related to *Actinoalloteichus hoggarensis* within the Pseudonocardiaceae (Actinomycetota), and other close relatives based on a set of 49 core universal genes for COG gene families (Arkin et al., 2018).

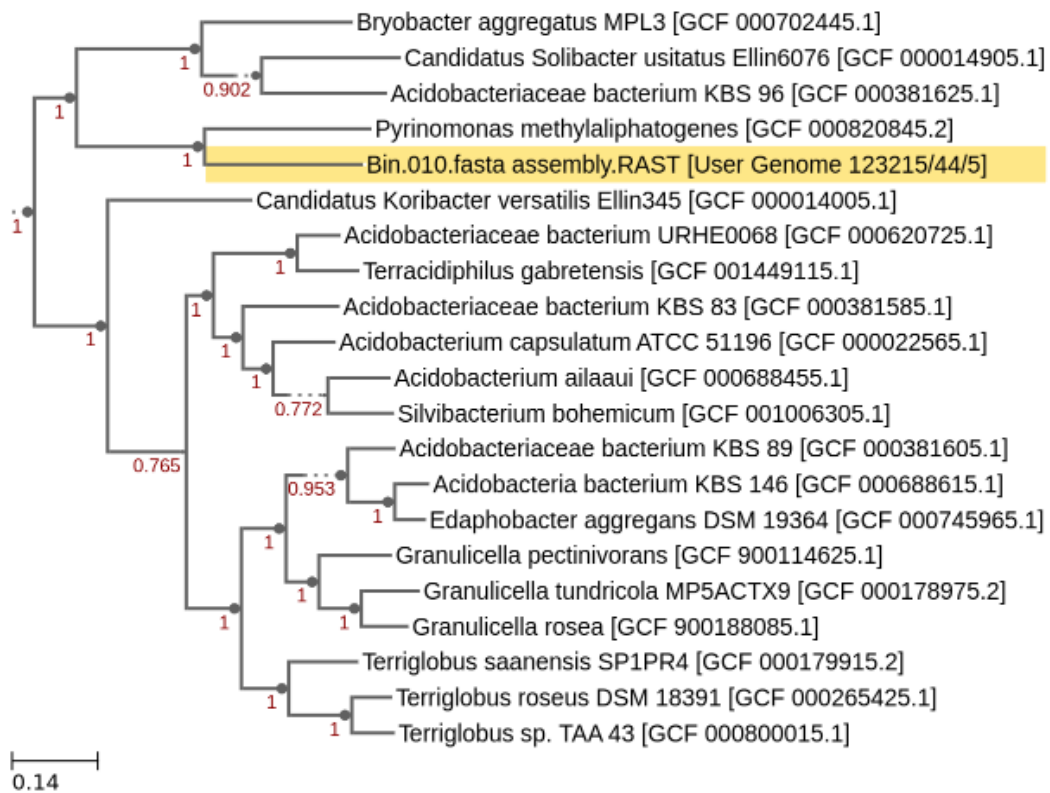


Figure 4.7: A phylogenetic tree created using Species Tree v2.2.0 for the MAG resulting from Bin.010, which was most closely related to *Pyrinomonas methylaliphatogenes* within the Blastocatellia class (Acidobacteriota), and other close relatives based on a set of 49 core universal genes for COG gene families (Arkin et al., 2018).

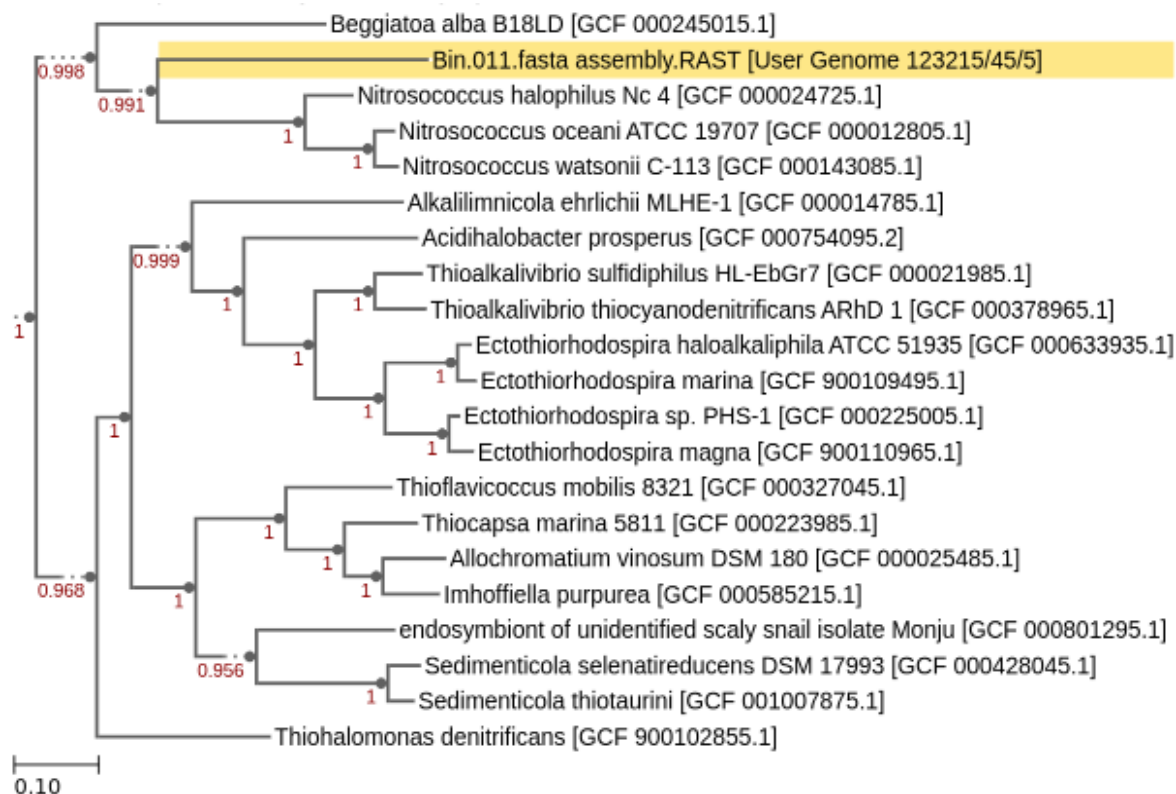


Figure 4.8: A phylogenetic tree created using Species Tree v2.2.0 for the MAG resulting from Bin.011, which was most closely related to *Nitrosococcus halophilus* Nc_4 within the Gammaproteobacteria class (Pseudomonadota), and other close relatives based on a set of 49 core universal genes for COG gene families (Arkin et al., 2018).

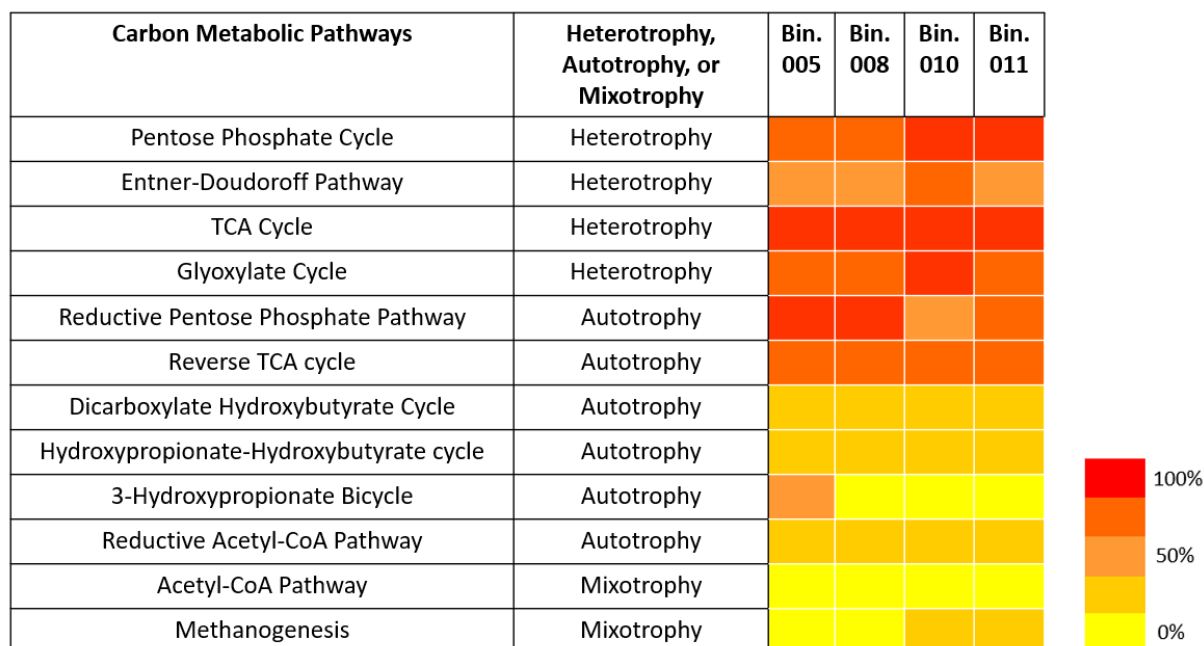
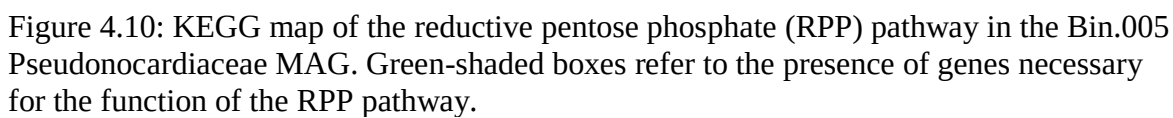


Figure 4.9: Heat map visualization of pathways involved in carbon metabolism for the high-quality MAGs obtained from cave-wall biofilms in Carter Cave, Tennessee.



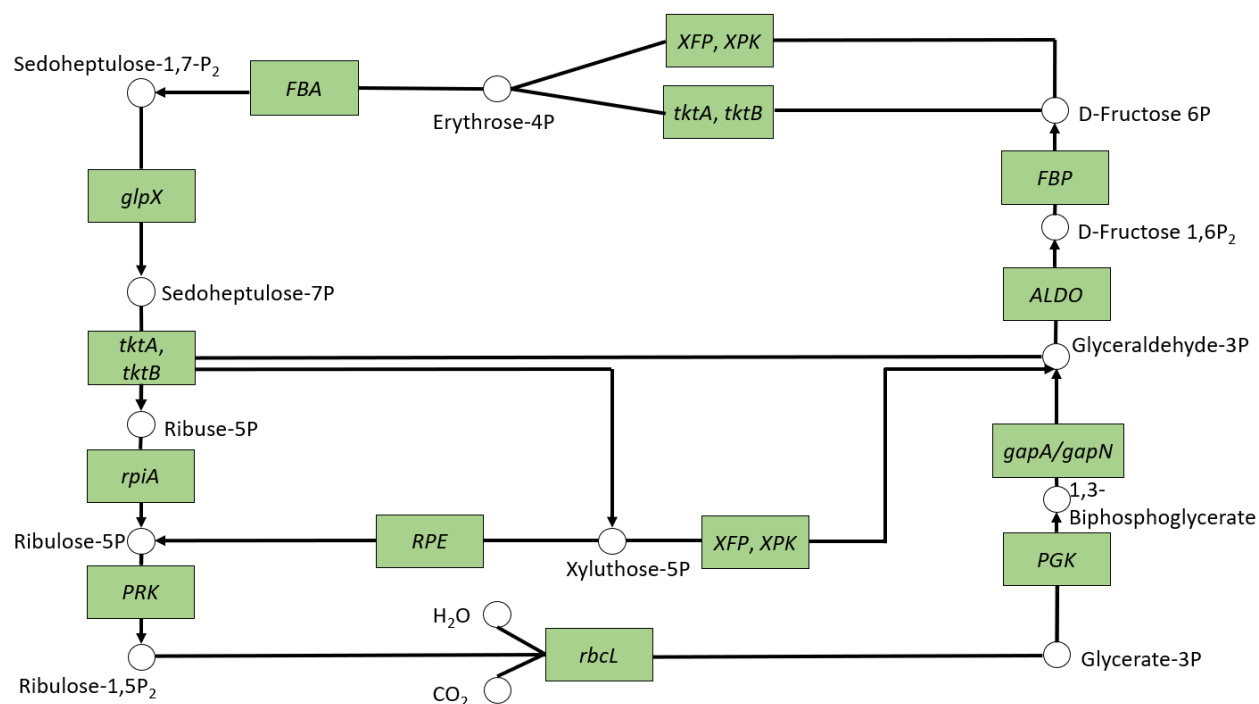


Figure 4.11: KEGG map of the reductive pentose phosphate (RPP) pathway in the Bin.008 Pseudonocardia MAG. Green-shaded boxes refer to the presence of genes necessary for the function of the RPP pathway.

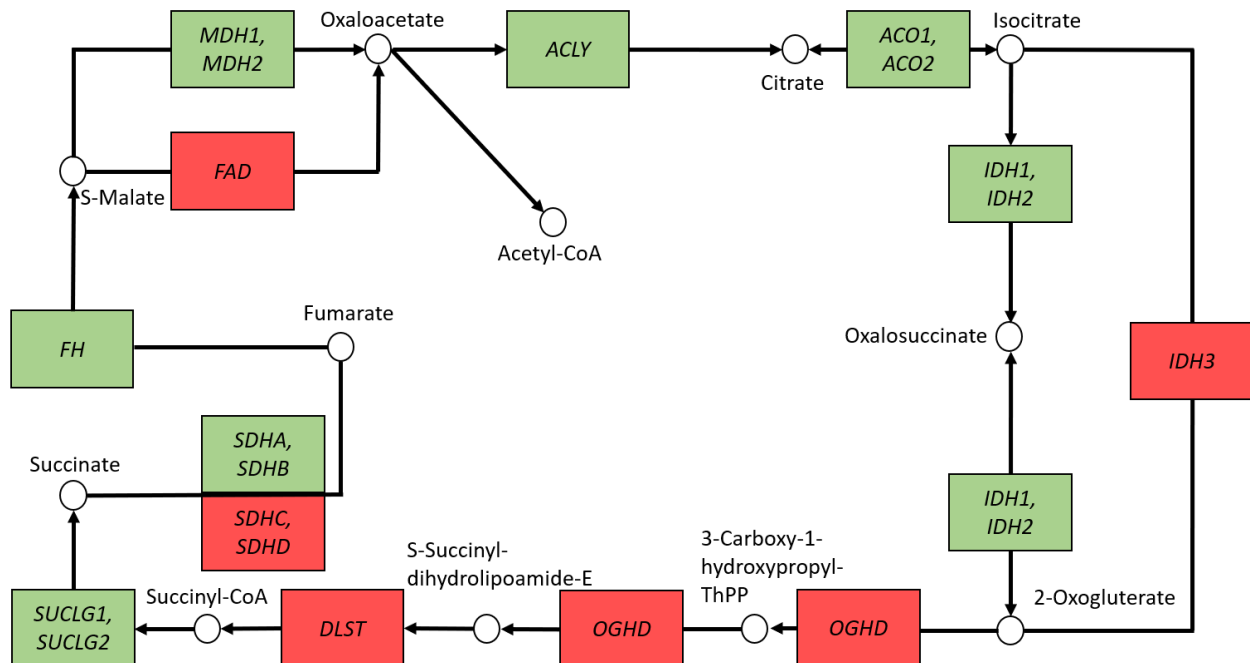


Figure 4.12: KEGG map of the TCA cycle in the Bin.005 Pseudonocardiaceae MAG. Green-shaded boxes refer to the presence of genes necessary for the function of the TCA cycle. Red boxes indicate the absence of genes.

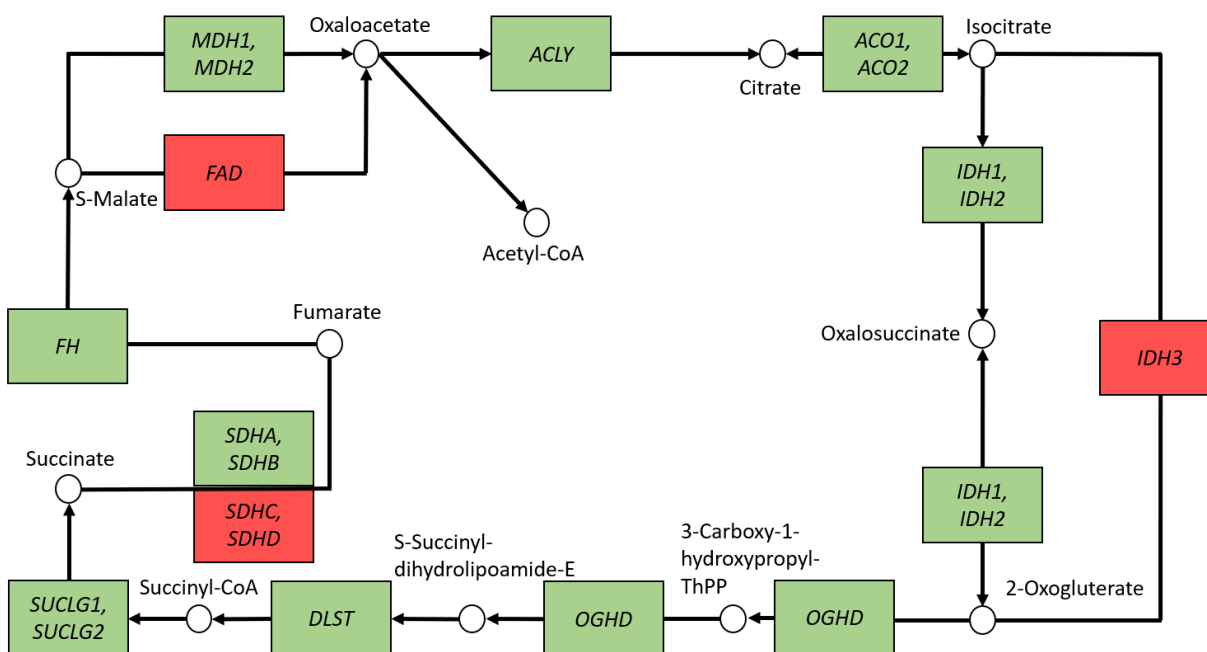


Figure 4.13: KEGG map of the TCA cycle in the Bin.008 Pseudonocardiaceae MAG. Green-shaded boxes refer to the presence of genes necessary for the function of the TCA cycle. Red boxes indicate the absence of genes.

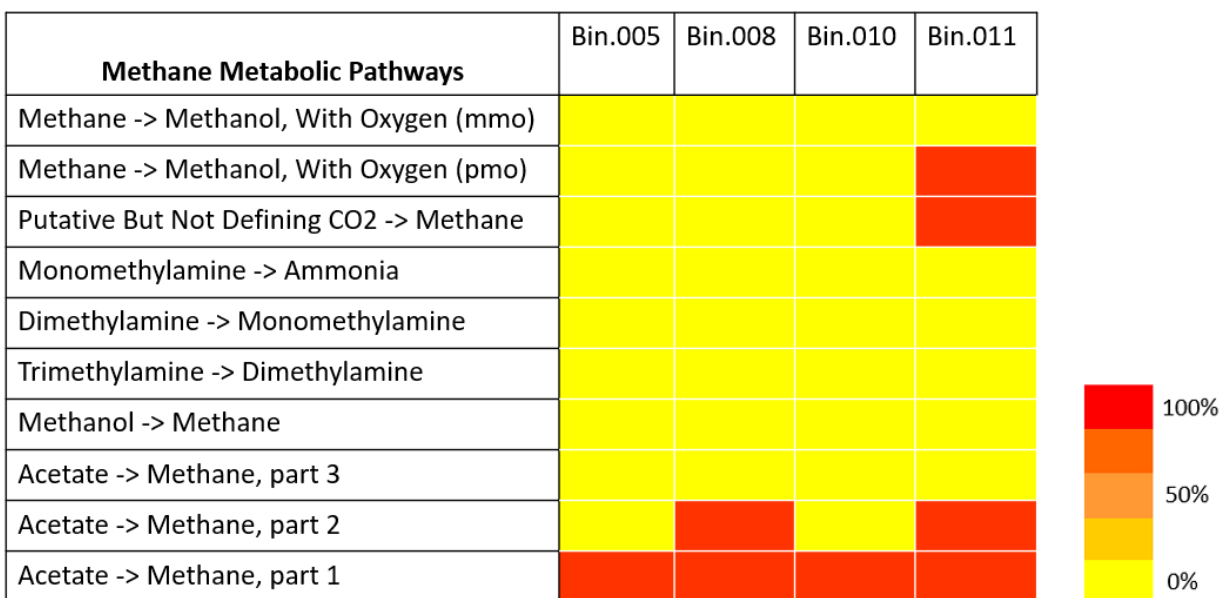


Figure 4.14: Heat map visualization using DRAM of methane metabolism pathways for high-quality MAGs from cave-wall biofilms, Carter Cave, Tennessee.

Table 4.1: Comparison of ANI values of the Pseusonocardiaceae MAGs (Bin.005 and Bin.008) to other Pseudonocardiaceae genera. Determinations were made using FastANI (Jain et al., 2018).

Pseudonocardiaceae Genera	Bin.005 ANI (%) Comparison	Bin.008 ANI (%) Comparison
<i>Actinoalloteichus hoggarensis</i>	76.4	76.2
<i>Actinokineospora alba</i>	76.8	77
<i>Actinomycespora chiangmaiensis</i>	76.5	76.5
<i>Actinophytocola algerensis</i>	76.6	76.7
<i>Actinosynnema mirum</i>	76.6	76.8
<i>Allokutzneria albata</i>	76.9	77.2
<i>Amycolatopsis acidicola</i>	76.7	76.9
<i>Crossiella equi</i>	77.1	77.2
<i>Gandjariella thermophila</i>	77.1	77.8
<i>Goodfellowiella sp.</i>	76.9	76.9
<i>Haloechothrix alba</i>	76.4	76.6
<i>Halopolyspora algerensis</i>	76.7	76.9
<i>Herbihabitans rhizosphaerae</i>	77.0	77.1
<i>Kibdelosporangium aridum</i>	76.6	76.7
<i>Kutzneria albida</i>	76.9	77
<i>Labedeae rhizosphaerae</i>	76.7	76.8
<i>Lechevaliera xinjiangensis</i>	76.3	76.4
<i>Lentzea albida</i>	76.1	76.4
<i>Prauserella coralliicola</i>	76.5	76.2
<i>Pseudonocardia autotrophica</i>	76.7	76.8
<i>Saccharomonospora amisosensis</i>	76.6	76.7
<i>Saccharothrix algeriensis</i>	76.8	77.1
<i>Sciscionella marina</i>	76.3	76.5
<i>Tamarichihabitans halophyticus</i>	76.8	76.7
<i>Thermocrisum agreste</i>	76.8	77.2
<i>Umezawaea tangerina</i>	76.5	76.5
<i>Yuhushiella deserti</i>	76.9	77.3

found in hydrogenotrophic methanogens that use CO₂ and H₂ as a substrate in anoxic environments (Wagner et al., 2018). The potential for complete methanotrophy was unique to Bin.011, which had genes that encoded for methane oxidation to methanol via pMMO.

Nitrogen Metabolism

All MAGs, except for the Blastocatellia MAG (Bin.011), had genes related to nitrogen metabolism (Table 4.2). Pseudonocardiaceae (Bin.005) contained complete sets of genes for both assimilatory and dissimilatory nitrate reduction. Pseudonocardiaceae (Bin.008) and Gammaproteobacteria (Bin.010) contained complete sets of genes for dissimilatory nitrate reduction, as well. The Blastocatellia MAG did not possess genes for the oxidation or reduction of nitrogen species for energy (Table 4.2).

Trace Gas Uptake

Genes associated with H₂ uptake were present in Bin.005, Bin.008, and Bin.010 (Table 4.3). All three MAGs contained structural genes necessary to encode [Ni-Fe] hydrogenase, along with multiple accessory genes (Lacasse and Zamble, 2016). Structural genes that are responsible for encoding all necessary subunits (CoxL, CoxM, CoxS) for CO dehydrogenase were also present in both Pseudonocardiaceae MAGs, as well as Bin.010. Accessory proteins were present in both Pseudonocardiaceae MAGs, with only Bin.008 containing *coxD* and *coxF*. The accessory gene *coxD* is crucial for the maturation of the CO dehydrogenase active site and plays a key role for growth in the presence of CO, whereas *coxF* is used in the electron transport chain (Pelzmann et al., 2009; Pelzmann et al., 2014). Genes for H₂ and CO uptake were absent in Bin.011.

Microbe-Mineral Interactions

Because of the association of carbonic anhydrase (CA) and putative calcite precipitation, genes encoding for CA were investigated (Table 4.4). All MAGs contained genes that encode for variants of CA, with beta-class CA being the most shared among the MAGs. Both Pseudonocardiaceae MAGs also contained genes necessary for the expression of BicA, although Bin.005 was the only MAG to contain genes associated with alpha-class CA.

DNA Repair Strategies

Genes for DNA repair strategies were present in the high-quality MAGs (Table 4.5). Both Bin.005 and Bin.008 had all necessary genes required for the RecF DNA repair pathway, a common DNA repair mechanism associated with single-strand DNA breaks (Handa et al., 2009).

Table 4.2: Genes involved in nitrogen metabolism in cave-wall biofilm MAGs from Carter Cave, Tennessee. Determinations were made using DRAM and KEGG Mapper.

MAG	Metabolism	Genes Present
Pseudonocardiaceae (Bin.005)	Assimilatory nitrate reduction	<i>nasC, nasD, nasE</i>
	Dissimilatory nitrate reduction	<i>narG, narH, narI, nirB, nirD</i>
Pseudonocardiaceae (Bin.008)	Dissimilatory nitrate reduction	<i>narG, narH, narI, nirB, nirD</i>
Blastocatellia (Bin.010)	N/A	N/A
Gammaproteobacteria (Bin.011)	Dissimilatory nitrate reduction	<i>narG, narH, narI, nirB, nirD</i>

Table 4.3: Genes present in the high-quality MAGs from cave-wall biofilms, Carter Cave, Tennessee, that could be involved with the production of trace gas uptake enzyme systems. Determinations were made using DRAM and KEGG Mapper.

MAG	[Ni-Fe]-hydrogenase		Carbon monoxide dehydrogenase	
	structural genes	accessory genes	structural genes	accessory genes
Pseudonocardiaceae (Bin.005)	<i>hydA, hydB</i>	<i>hypA, hypB, hypC, hypD, hypE, hypF</i>	<i>coxL, coxM, coxS</i>	<i>coxD</i>
Pseudonocardiaceae (Bin.008)	<i>hydA, hydB</i>	<i>hypA, hypB, hypC, hypD, hypE, hypF</i>	<i>coxL, coxM, coxS</i>	<i>coxD, coxF</i>
Blastocatellia (Bin.010)	<i>hydA, hydB</i>	<i>hypA, hypB, hypC, hypD, hypE, hypF</i>	<i>coxL, coxM, coxS</i>	N/A
Gammaproteobacteria (Bin.011)	N/A	N/A	N/A	N/A

Table 4.4: Genes present in the high-quality MAGs from cave-wall biofilms, Carter Cave, Tennessee, that could be involved in carbonic anhydrase (CA). Determinations were made using DRAM and KEGG Mapper.

MAG	Potential Carbonic Anhydrase Enzymes
Pseudonocardiaceae (Bin.005)	• Carbonic anhydrase, beta class • Na⁺ dependent bicarbonate transporter, Carbonic anhydrase, beta class (<i>bicA</i>)
Pseudonocardiaceae (Bin.008)	• Carbonic anhydrase, alpha class • Carbonic anhydrase, beta class • Na⁺ dependent bicarbonate transporter Carbonic anhydrase, beta class (<i>bicA</i>)
Blastocatellia (Bin.010)	• Carbonic anhydrase, family 3 (<i>yrdA</i>) • Carbonic anhydrase, beta class
Gammaproteobacteria (Bin.011)	• Carbonic anhydrase • Carbonic anhydrase, beta class

Table 4.5: The presence (✓) or absence (X) of genes that encode for the RecF/MutHLS pathways. Determinations were made using DRAM and KEGG Mapper.

MAG	<i>recF</i>	<i>recO</i>	<i>recR</i>	<i>recJ</i>	<i>recN</i>	<i>recG</i>	<i>recQ</i>	<i>recA</i>	<i>lexA</i>	<i>mutH</i>	<i>mutL</i>	<i>mutS</i>
Pseudonocardiaceae (Bin.005)	✓	✓	✓	X	✓	✓	✓	✓	✓	X	X	X
Pseudonocardiaceae (Bin.008)	✓	✓	✓	X	✓	✓	✓	✓	✓	X	X	X
Blastocatellia (Bin.010)	✓	X	✓	✓	X	✓	✓	✓	X	✓	✓	✓
Gammaproteobacteria (Bin.011)	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

In this process, LexA controls the expression of genes that participate in the RecF pathway. RecN and RecR are responsible for recognizing damaged single-stranded DNA and directing Rec proteins. RecF then binds to the damaged DNA strand and promotes the activation of RecA, which is stabilized by RecO to separate the damaged DNA strand and conduct repairs. Genes encoding enzymes used to fix damaged double-stranded DNA were absent. Although Bin.010 and Bin.011 did not have genes encoding for the RecF repair pathway, both had genes that encoded for DNA mismatch repairing mechanisms and genes that encoded the MutHLS pathway, where the MutS protein is responsible for recognizing and binding to mismatched base pairs and recruits the MutL protein. Moreover, MutH is responsible for marking a newly synthesized DNA strand that will be used after the damaged strand degrades (Marti et al., 2002).

Survival Strategies and Biofilm Formation

Genes related to physical, biological, and environmental stressors were also investigated. All four MAGs had genes that encoded cold-shock proteins, although Bin.010 and Bin.011 were unspecified. Bin.005 and Bin.008 had CspB through the *cspB* gene. RNA polymerase is also associated with responses to harsh conditions. RNA polymerase has an alpha subunit (RpoA), a beta subunit (RpoB), and a beta prime subunit (RpoC), along with multiple sigma factors that recognize specific promotor regions of DNA during stress (Sutherland and Murakami, 2018). All four MAGs had all RNA polymerase subunits (Table 4.6), as well as *rpoD*, *rpoE*, *rpoN*, and *rpoS* that encode for Sigma D, Sigma E, Sigma N, and Sigma S, respectively. Sigma D regulates genes involved with cell growth and stress response (Wade and Struhl, 2004) and is typically expressed during the transition from exponential growth to stationary growth when nutrients are limited (Pletnev et al., 2020). Sigma E regulates genes involved with cold stress and entry into the stationary phase (Miticka et al., 2003). Sigma N is responsible for activating genes involved with environmental nitrogen limitation response, and typically allows for alternative nitrogen compound use for energy (Dong and Schellhorn, 2011). It has also recently been associated with motility and biofilm formation (Liu et al., 2021; Zhang et al., 2021). Sigma S is associated with identifying and expressing genes that respond to dehydration during stationary growth (Chen and Goulian, 2018), as well as biofilm formation (Keenan, 2016). The only MAG that contained *rpoH* was Bin.011. This gene encodes for Sigma H, which regulates the expression of genes

Table 4.6: The presence (✓) or absence (X) of genes encoding for the four subunits of RNA polymerase and various sigma factors. Determinations were made using DRAM and KEGG Mapper.

MAG	<i>rpoA</i>	<i>rpoB</i>	<i>rpoC</i>	<i>rpoD</i>	<i>rpoE</i>	<i>rpoH</i>	<i>rpoN</i>	<i>rpoS</i>	<i>rpoZ</i>
Pseudonocardiaceae (Bin.005)	✓	✓	✓	✓	✓	X	✓	✓	✓
Pseudonocardiaceae (Bin.008)	✓	✓	✓	✓	✓	X	✓	✓	✓
Blastocatellia (Bin.010)	✓	✓	✓	✓	✓	X	✓	✓	X
Gammaproteobacteria (Bin.011)	✓	✓	✓	✓	✓	✓	✓	✓	✓

involved in response to higher temperatures, detoxification, and responses to oxidative stress (Zhao et al., 2005). Biofilm formation serves multiple purposes, including combating harsh conditions, nutrient acquisition, and communication. Although KEGG Mapper contains only references for biofilm formation from *Pseudomonas aeruginosa*, *Vibrio cholerae*, and *Escherichia coli*, all four cave-wall biofilm MAGs contained genes associated with putative biofilm formation, motility, and maintenance, including *crp*, *rpoS*, *rpoN*, *fliA*, *vpsU*, *trpE*, and *pelF* (Table 4.7). However, none of the MAGs contained a complete set of genes for any specific pathway. The Blastocatellia MAG and Gammaproteobacteria MAG contained a broader variety of biofilm formation genes compared to both Pseudonocardiaceae MAGs.

Defense against microbial competitors may also play a large role in the dominance of major taxa in cave-wall biofilms. Antimicrobial resistance and bacterial toxins are important mechanisms that can prevent foreign microbes from interfering with microbial processes within biofilms, and toxin-antitoxin systems can serve the same role by regulating gene expression, cell growth, and inducing a dormant state during unfavorable conditions (Unterholzner et al., 2013). Defensive mechanisms in the MAGs included antimicrobial resistance genes/pathways, toxin/antitoxin systems, and bacterial toxins (Table 4.8), including TlyC pore-forming toxins and resistance to beta-lactam antibiotics. Both Pseudonocardiaceae MAGs contained the potential for resistance to beta-lactam antibiotics, TlyC pore-forming toxins, and the VapC-VapB toxin/antitoxin system. The Blastocatellia and Gammaproteobacteria MAGs also had a variety of multi-drug resistance modules.

Table 4.7: Genes associated with biofilm formation, motility, and maintenance using *V. cholerae*, *P. aeruginosa*, *E. coli* as reference pathways. Determinations were made using DRAM and KEGG Mapper.

Pseudonocardiaceae (Bin.005)	Pseudonocardiaceae (Bin.008)	Blastocatellia (Bin.010)	Gammaproteobacteria (Bin.011)
<i>crp</i> <i>rpoS</i> <i>rpoN</i> <i>fliA</i> <i>vpsU</i> <i>trpE</i> <i>pelF</i> <i>glgA</i> <i>glgC</i> <i>glgP</i>	<i>crp</i> <i>rpoS</i> <i>rpoN</i> <i>fliA</i> <i>vpsU</i> <i>trpE</i> <i>trpG</i> <i>glgA</i> <i>glgC</i> <i>glgP</i>	<i>cpr</i> <i>rpoS</i> <i>rpoN</i> <i>uvrY</i> <i>hfq</i> <i>cyaA</i> <i>vspU</i> <i>gspC</i> <i>pilJ</i> <i>pilH</i> <i>chpA</i> <i>cyaB</i> <i>trpE</i> <i>trpG</i> <i>uvrY</i> <i>pslA</i> <i>cyaA</i> <i>dksA</i> <i>bcsA</i> <i>wza</i> <i>glgA</i> <i>glgC</i> <i>glgP</i>	<i>cpr</i> <i>rpoS</i> <i>rpoN</i> <i>uvrY</i> <i>csrA</i> <i>cyaA</i> <i>cyaB</i> <i>fis</i> <i>hfq</i> <i>fliA</i> <i>fliA</i> <i>vpsU</i> <i>gspC</i> <i>pilJ</i> <i>pilI</i> <i>pilH</i> <i>pilG</i> <i>chpA</i> <i>cpdA</i> <i>trpE</i> <i>rhlB</i> <i>ppkA</i> <i>pslA</i> <i>dksA</i> <i>wza</i> <i>pgaC</i>

Table 4.8: Presence of genes used in antimicrobial resistance, toxin-antitoxin systems, and bacterial toxins for the high-quality MAGs from cave-wall biofilms, Carter Cave, Tennessee. Determinations were made using DRAM and KEGG Mapper.

MAG	Antimicrobial Resistance Genes/Gene Sets	Toxin-Antitoxin Systems	Bacterial Toxins
Pseudonocardiaceae (Bin.005)	Beta-lactam resistance • penP; beta-lactamase class A Vancomycin resistance • vanX; zinc D-Ala-D-Ala dipeptidase	YoeB-YefM VapC-VapB FitB-FitA	<i>tlyC</i> ; magnesium and cobalt exporter
Pseudonocardiaceae (Bin.008)	Imipenem resistance • cusR, copR, silR	YoeB-YefM VapC-VapB HigB-HigA	<i>tlyC</i> ; magnesium and cobalt exporter
Blastocatellia (Bin.010)	Beta-lactam resistance • penP; beta-lactamase class A Imipenem resistance • cusR, copR, silR Vancomycin resistance • vanX; zinc D-Ala-D-Ala dipeptidase Multi-drug resistance modules • MexAB-OprM • MexJK-OprM • MexXY-OprM • MexPQ-OpmE • AcrEF-TolC • MdtEF-TolC	YoeB-YefM ParE-ParD HigB-HigA	<i>tlyC</i> ; magnesium and cobalt exporter
Gammaproteobacteria (Bin.011)	Beta-lactam resistance • penP; beta-lactamase class A Vancomycin resistance • vanX; zinc D-Ala-D-Ala dipeptidase Multi-drug resistance modules • MexAB-OprM • MexJK-OprM • MexXY-OprM • AcrEF-TolC • MdtEF-TolC • BpeEF-OprC	MazF-MazE YoeB-YefM HigB-HigA VapC-VapB FitB-FitA	<i>tlyC</i> ; magnesium and cobalt exporter

SECTION FIVE

DISCUSSION

Microbial metabolism in cave ecosystems, especially of the most abundant taxonomic groups, is responsible for carbon and nutrient cycling. Prior investigations of the Carter Cave cave-wall biofilms (Frazier, 2020) suggest that microbes capable of chemolithoautotrophy using CO₂ and possibly CH₄ as carbon sources were responsible for imparting negative bulk biofilm carbon isotope values and that members of the Pseudonocardiaceae were the most abundant groups according to 16S rRNA surveys. Despite Pseudonocardiaceae being repeatedly observed in high relative abundances from caves (Table 2.1) (e.g., Cuezva et al., 2012; Porca et al., 2012; Ortiz et al., 2014; Lavoie et al., 2017; Li et al., 2018; Long et al., 2019; Wiseschart et al., 2019; Zhu et al., 2019; Frazier, 2020; Gonzalez-Pimentel et al., 2021; Ma et al., 2021; Zhu et al., 2021), their metabolic functions have remained a mystery. Consequently, the purpose of this investigation was to focus on cave-wall Pseudonocardiaceae to understand potential carbon and energy acquisition pathways, along with mechanisms involved in biofilm formation, mutualism, and competition. Because Pseudonocardiaceae made up over half of the Carter Cave biofilm community for some samples, in addition to their low overall diversity, these samples were ideal for metagenome sequencing. To our knowledge, this is the first known study to turn to shotgun metagenomics to describe cave-wall biofilm Pseudonocardiaceae and associated members of the microbial community, including to the Blastocatellia and Gammaproteobacteria, also common in cave-wall biofilms but with limited knowledge about their metabolic potential (Spasov et al., 2020; Ortiz et al., 2021).

Metabolic Potential for Pseudonocardiaceae in Cave-wall Biofilms

The two high-quality Pseudonocardiaceae MAGs contained all steps for the RPP pathway and the TCA cycle, nitrate reduction, and H₂ and CO uptake (Figures 4.9–4.12, Tables 4.2 and 4.3). Coupled with $\delta^{13}\text{C}$ values from Frazier (2020) for the same samples, there is strong evidence that these Pseudonocardiaceae are capable of autotrophy. MAG annotation using DRAM and KEGG indicated a complete RPP pathway in both MAGs, but the rTCA cycle was complete according to DRAM and incomplete with KEGG. The RPP pathway is the most widely used mechanism for CO₂ fixation in photoautotrophs and chemolithoautotrophs with suitable

electron donors, which could include H₂ and CO based on the complete suite of structural and accessory *hyd* and *cox* genes in the MAGs (Table 4.3).

The use of CO and H₂ is more widespread than previously thought (Cordero et al., 2019; Greening et al., 2022). The average CO concentration in the atmosphere at ground level varies daily and seasonally, from about 0.05 to 0.3 ppm or higher. Hydrogen gas is also present at about 0.5–0.6 ppm in the atmosphere at ground level. Carboxydotrophs metabolize CO to supplement energy and carbon requirements in environments with elevated CO concentrations, but chemolithoautotrophic growth can be supported at levels as low as 0.05 ppm (King et al., 2007; Cordero et al., 2019). Cordero et al. (2019) also report that carboxydotrophs can oxidize CO to support the RPP pathway. Chemolithoautotrophs can also use H₂ at atmospheric concentrations with the aid of high affinity Ni-Fe hydrogenases (Greening et al., 2022). Microbial degradation of organic matter (e.g., fermentation) can supply high concentrations of H₂. Greening et al. (2022) note that H₂ oxidation in aerobic bacteria is well studied from soil ecosystems, but is understudied in subterranean environments, although CO and H₂ concentrations in caves are virtually unknown. Future cave microbial ecology studies should consider the use of transcriptomics in combination with H₂ and CO measurements to verify that there are microbes with cave-wall communities capable of these metabolic functions.

The TCA cycle is considered a heterotrophic metabolism that generates ATP and NADH by combining acetyl-CoA with oxaloacetate to produce CO₂ during respiration (Kanehisa et al., 2023), which could also include nitrate use as an alternative terminal electron acceptor. Both MAGs indicated dissimilatory nitrate reduction, although Pseudonocardiaceae (Bin.005) also had a complete set of genes for assimilatory nitrate reduction (Table 4.2). Although neither MAG was 100% complete, and some genes could be missing from the MAGs, the noted differences among these gene presences/absences could indicate that there is potentially strain-level diversity between the MAGs. Even so, MAGs are not considered to be identical to the actual genomes of the organisms in question, requiring greater consideration for the absence of genes in high quality MAGs and placing further emphasis on future attempts for cultivation and whole-genome sequencing.

Pseudonocardiaceae Taxonomy

Determining the taxonomic identity of the two *Pseudonocardiaceae* MAGs could have been straightforward if the reference databases contained representatives for all genera in this family, but the SILVA v.138.1 reference database used for 16S rRNA gene phylogenies and the GTDB-Tk database used for metagenome taxonomic determination are incomplete. Associating a taxonomic identity to both MAGs was further complicated because both MAGs lacked 16S rRNA genes, a common dilemma associated with short read MAGs (Orellana et al., 2023). The 16S rRNA survey revealed that the genus *Crossiella* is the dominant *Pseudonocardiaceae* genus within the biofilm samples (Figure 4.4). This genus has also been reported as the most prevalent from other cave biofilms studies (Table 2.1), but this could also be an artifact of others using the SILVA database. However, the GTDB-Tk database has only two cultured *Crossiella* species genomes for reference. Based on phylogenetic analysis a set of 49 core universal genes for COG gene families from the two cave MAGs, gene homologs were not more closely related to those from *Crossiella* reference genomes but to *Actinoalloteichus hoggarensis*, yet ANI values for both MAGs were highest when compared to *G. thermophila*, at 77.1% and 77.8% (Xu et al., 2021) (Figures 4.5 and 4.6, Table 4.1). Neither *A. hoggarensis* nor *G. thermophila* were retrieved from the 16S rRNA gene surveys from the cave. Representative sequences exist in the SILVA database for *A. hoggarensis* from which phylogenetic analyses could be made, but representative sequences for *G. thermophila* are absent in the SILVA database yet present in the GTDB-tk database. The 16S rRNA gene was extracted from the *Gandjariella* MAG in the GTDB-tk database and ran through BLAST to determine its sequence similarity to both *Crossiella* and *Actinoalloteichus*, revealing inter-genus relatedness (95.32% and 94.66% sequence identities, respectively). Moreover, previous work reports that *A. hoggarensis* was isolated from Saharan soils and it metabolizes a range of organic compounds (Boudjelal et al., 2015), which is inconsistent with the reconstructed metabolism of the two cave MAGs. *G. thermophila* was also recently identified in forest soils near a geothermal environment (Ningsa et al., 2019). It is currently the only species belonging to this genus and genomic information regarding carbon acquisition is inconclusive. Consequently, it is still possible that the Carter Cave MAGs belong to an novel species of *Crossiella*, as the 16S rRNA analysis may suggest. The sequences belonging to ASV001 (to which 98.2% of all *Crossiella* sequence reads belonged) were ran

through BLAST. Resulting sequence similarities above 98.65% belonged to unclassified Pseudonocardiales. Future work should address this issue and perhaps generate metagenomes using long sequence reads that could yield 16S rRNA genes and perhaps more complete MAGs.

Metabolic Potential and Taxonomy of Other MAGs in the Cave-wall Biofilms

Ecologically, Blastocatellia are slow-growing, aerobic heterotrophs capable of surviving desiccation and oligotrophy (Huber et al., 2014; Pascual et al., 2015; Wüst et al., 2016). The Blastocatellia MAG from this study had carbon metabolism pathways indicative of heterotrophy, including the PP and TCA pathways, as well as the glyoxylate cycle. The MAG also had genes for H₂ and CO uptake as an alternative source of energy, but did not recover genes related to nitrogen metabolism (Figure 4.9, Table 4.3). Wüst et al. (2016) describe members within this class as K-strategists, and Adesso et al. (2021) hypothesize that Blastocatellia are opportunists in caves by oxidizing metabolic compounds from nearby Gammaproteobacteria, which may explain why they have an array of biofilm formation, antimicrobial resistance, and toxin-antitoxin genes. The Blastocatellia MAG was taxonomically unresolved, which likely supports that the Carter Cave species is novel.

The recovered genes from the Gammaproteobacteria MAG suggested a heterotroph because of having complete PP and TCA pathways using oxygen (Figure 4.9), as well as genes that indicated nitrate use for energy as a terminal electron acceptor (Table 4.2). No genes were present for H₂ and/or CO use, but there were genes for methanotrophy, specifically the *pmoA*, that encodes for the methane monooxygenase subunit A (Figure 4.14). Taxonomically, the most closely related species to this MAG was taxonomy *N. halophilus*, but the ANI value was low (Figure 4.8). Based on taxonomic affinity for the 16S rRNA genes, *Steroidobacter* was one of the most abundant gammaproteobacterial groups (Figure 4.4). More research is needed to resolve the identity of this MAG, which is likely a novel organism.

Biofilm Lifestyle – Formation, Maintenance, and Survival Implications

Lack of light, carbon, and nutrients, low temperatures, high humidity, and microbial competition are all factors that affect microbial growth in cave-wall biofilms. The formation of biofilms serves multiple purposes, including mutualism, protection from environmental conditions, nutrient acquisition, nutrient source, and protection from foreign microbes. All MAGs contained the CRP/FNR family transcriptional regulator gene, *crp* (Table 4.7). CRP has

been associated with promoting biofilm formation in *E. coli* and *K. pneumoniae* by inhibiting the production of fimbriae to adhere to surfaces (Liu et al., 2020). Similarly, *fliA* encodes for the transcription factor FliA, which is associated with the development of type 1 pili and chemotaxis that are important for colonizing surfaces (Buck et al., 2021). Only *trpE*, responsible for catalyzing the first step of the tryptophan pathway and early stages of biofilm production, was present in all MAGs (Hamilton et al., 2009). Moreover, the Gammaproteobacteria MAG also had genes to produce type 4 pili, which aids in adhesion, motility, and chemotaxis (Leong et al., 2017). All MAGs contained *vspU*, a gene responsible for maintaining biofilm morphology through VspU, although a multitude of *vps* genes (e.g., *vpsU*, *vpsA-K*, *vpsL-Q*) are required for exopolysaccharide production during biofilm building (Gao et al., 2020). All MAGs had *pelF*, which encodes glycosyltransferase and is responsible for the production of a glucose-rich polysaccharide biofilm matrix (Rastgar et al., 2022). A glucose-rich biofilm can be beneficial because it forms a more adhesive and rigid structure that can act as a barrier against environmental stressors (Whitfield et al., 2020). Both Pseudonocardiaceae MAGs and the Blastocatellia MAG contained genes associated with glycogen biosynthesis and biofilm formation (Chen et al., 2019). The accumulation of glycogen increases the size of biofilms and can also act as a source of carbon and energy (Bonafonte et al., 2000).

Therefore, biofilm production affects the metabolic flexibility of community members but also can protect against environmental stress. All four MAGs had genes encoding for RNA polymerase (Table 4.6) that regulate gene expression in response to changing environmental conditions, such as cold or changing temperatures and oligotrophy, through sigma factors (Burgess, 2001; Miticka et al., 2003; Wade and Struhl, 2004; Zhao et al., 2005; Lovely et al., 2011; Pletnev et al., 2020). All MAGs also contained cold-shock genes that encode proteins typically expressed at colder temperatures. Both Pseudonocardiaceae MAGs contained genes of unspecified cold-shock proteins, but the Blastocatellia and Gammaproteobacteria MAGs had *cspB*, a cold-shock protein expressed during stationary growth phase (Kram et al., 2020). The ability to maintain stationary growth while conserving energy during environmental stress, is encoded for by *rpoS* and results in the formation of smaller biofilms (Sheldon et al., 2012). This may explain why biofilms on cave walls, even if dominated by microbes with large cell sizes in high relative abundance, are small and generally do not cover all available surfaces.

Moreover, oligotrophy encourages competition. Bacteria deal with competition using a plethora of strategies that include antimicrobial resistance and toxin production. All MAGs had genes necessary for various forms of antimicrobial resistance, toxin-antitoxin systems, and bacterial toxins that protect against antibiotics and cellular damage, as well as the production of pore-forming toxins that are meant to degrade cellular structures (Table 4.8). Specifically, one of the Pseudonocardiaceae MAGs (Bin.005), and the Blastocatellia and Gammaproteobacteria MAGs, had *penB*, a gene responsible for encoding class A beta-lactamase, an enzyme associated with beta-lactam resistance (Somprasong et al., 2020). Beta-lactam is a class of antibiotics that include compounds like penicillin, amoxicillin, and ampicillin (Holton and Onusko, 2000). Pseudomonadota, Bacillota, and other members of the Actinomycetota can produce beta-lactam antibiotics (Adesoji and Ogunjobi, 2016; Lazarus et al., 2019), and are all members found repeatedly in cave-wall biofilm communities (Ortiz et al., 2014; Lavoie et al., 2017; Li et al., 2018; Long et al., 2019; Wischart et al., 2019; Zhu et al., 2019; Frazier, 2020; Gonzalez Pimentel et al., 2021; Ma et al., 2021; Zhu et al., 2021). Vancomycin resistance genes were also found in the Bin.005 Pseudonocardiaceae MAG, and the Blastocatellia and Gammaproteobacteria MAGs. Vancomycin is an antibiotic that is commonly associated with the Pseudonocardiaceae taxa, *Amycolatopsis orientalis*, and is used to treat infections caused by *Staphylococcus aureus* (Applbaum and Bozdogan, 2004; Lee et al., 2016).

Collectively, living with active competitors also requires antimicrobial/antibiotic resistance, and some bacteria likely evolved mechanisms to combat antimicrobial compounds and maintain a critical advantage when competing with other microbes on surfaces. For instance, the Gammaproteobacteria and Blastocatellia MAGs had Mex-Opr complexes that are found in Gram-negative bacteria that pump antimicrobial compounds out of the cell, thereby reducing compound potency (Aeschlimann, 2003; Tsutsumi et al., 2019). This system has been well studied in *Pseudomonas aeruginosa* (Fujiwara et al., 2022). Similarly, toxin-antitoxin systems aid in survival during environmental disturbance, antibiotic exposure, genetic damage, and nutrient starvation. All four MAGs contained genes for the YoeB-YefM toxin-antitoxin system, as well as VapC-VapB (Pseudonocardiaceae Bin.005, Pseudonocardiaceae Bin.008, Gammaproteobacteria) or HigB-HigA (Pseudonocardiaceae Bin.008, Blastocatellia, Gammaproteobacteria), which have been linked to stress response, persister cell production, and

biofilm formation (Maisonneuve et al., 2011; Shavit et al., 2016; Wood and Wood, 2016; Yang and Walsh, 2017; Eroshenko et al., 2020; Xue et al., 2020; Ma et al., 2021; Patel et al., 2022). Moreover, all MAGs contained the genes to produce the pore-forming toxin TlyC, which lyses bacterial cell walls and degrades peptidoglycan (Vigneron et al., 2022). The *tlyC* gene was the only gene in the Carter Cave MAGs associated with bacterial toxicity. To defend against microbial competitors and their toxins that cause DNA mismatches and direct DNA damage, microbes need DNA repair mechanisms. All MAGs possessed DNA repair genes. Both Pseudonocardiaceae MAGs had genes for the RecF pathway used to repair single-strand DNA breaks. Only one other study has observed and described the RecF pathway in cave microbiomes (Bendia et al., 2022). Also, the Blastocatellia and Gammaproteobacteria MAGs encoded for the MutHLS complex to correct base pair mismatches. Further work would need to be done to pinpoint specific survival mechanisms under the influence of these genes, should they be expressed.

Implications for the Cave Habitat

To understand potential microbial roles in the karst environment, genes encoding for carbonic anhydrase (CA) were investigated (Table 4.4). Along with functions regarding metabolism and homeostasis, CA is one of the main enzymes responsible for microbe-induced calcite precipitation (MICP) (Riquelme et al., 2015; Gonzalez-Pimentel et al., 2018; Maciejewska et al., 2018; Liu et al., 2020; Gonzalez-Pimentel et al., 2021). All four MAGs contained genes to produce CA variants, and all MAGs contained beta-class CA genes and both Pseudonocardiaceae MAGs also had genes for BicA, an Na⁺-dependent bicarbonate transporter that transports bicarbonate ions across the cell membrane using Na⁺ as an electrochemical gradient for the main purpose of regulating cellular pH (Bernardo et al., 2006). The only MAG that had genes for alpha-class CA was the Bin.008, a Pseudonocardiaceae MAG. Both classes of CA, along with BicA, perform similar functions for CO₂ uptake and storage, pH regulation, and have comparable influence on calcite precipitation (Frost, 2014; Price and Howitt, 2014; Murray and McKenna, 2019; Wang et al., 2019; Martin-Pozas et al. 2022). Precipitates resulting from MICP typically are nanocrystal specks that accumulate at the base of the biofilm and grow outward (Li et al., 2015; Zambare et al., 2020). Li et al. (2015) also describe how the development of calcite precipitation can act to regulate the structure and composition of biofilms,

by deforming biofilm morphology and detaching bacterial cells. The presence of CA genes in the MAGs also supports that microbes have different ways of bringing inorganic carbon into the cell, which can regulate the pH of the surrounding environment and influence mineral precipitation. Furthermore, the potential for MICP can impact biofilm dynamics, including biofilm stability and community control. Frazier (2020) notes secondary carbonate speleothems in the vicinity of the sampled biofilms, which also had lighter carbon isotope compositions than expected if precipitated from the dissolved inorganic carbon of dripwater coming into the cave. The hypothesis was that the speleothems could be MICP, and future work should investigate these speleothems and associated biofilms in more detail given the metabolic potential from recovered MAGs.

SECTION SIX

CONCLUSIONS

Four high-quality MAGs were generated using samples from Frazier (2020) and analyzed for genes related to carbon metabolism, nitrogen metabolism, biofilm formation, and survival strategies. Two of the MAGs belonged to Pseudonocardiaceae, and the other two belonged to Blastocatellia and Gammaproteobacteria. The two high-quality Pseudonocardiaceae MAGs contained all steps for the RPP pathway and the TCA cycle, nitrate reduction, and H₂ and CO uptake. Coupled with $\delta^{13}\text{C}$ values from Frazier (2020) for the same samples, there is strong evidence that these Pseudonocardiaceae members are capable of autotrophy. Due to the array of results regarding their taxonomic identity, both MAGs likely represent novel members of the Pseudonocardiaceae. The Blastocatellia MAG contained all of the necessary genes for the pentose phosphate pathway, the TCA cycle, and the glyoxylate cycle, which suggests this microbe is a heterotroph, but there is also the potential to uptake H₂ and CO. The Gammaproteobacteria MAG possessed evidence of heterotrophy and mixotrophy, based on a suite of genes for the pentose phosphate pathway, the TCA cycle, nitrate reduction, and possibly methanotrophy. All four MAGs contained the genetic potential for unique survival strategies, including response to cold conditions, nutrient limitation, biofilm production and maintenance, antimicrobial resistance, and virulence, all traits that improve their ability to survive on the extreme cave-wall environment.

Although shotgun metagenomics is a useful tool that can provide insight for the metabolic functions of a microorganism, this approach cannot guarantee that functions of interest are expressed by members of the microbial community. Therefore, future studies regarding the metabolism of the Pseudonocardiaceae should utilize 1) metatranscriptomics to understand which carbon and energy acquisition genes are being expressed, along with the quantity of genes that are actively expressed, and 2) metaproteomics to understand the production of proteins that are used for carbon and energy metabolism. Given the evidence that the cave-wall biofilm microbes are likely novel species, it is possible that their proteins are also novel. Furthermore, steps should be taken to quantify active members of the Pseudonocardiaceae in biofilms via quantitative PCR. Apart from the Pseudonocardiaceae, the Blastocatellia had also been recovered from other cave microbiome studies, albeit in lower relative abundances (Frazier, 2020; Jurado et

al., 2020; Adesso et al., 2021; Gonzalez-Pimentel et al., 2021; Jurado et al., 2022). Wüst et al. (2016) provide discussion regarding the survival of Blastocatellia in harsh environments, like caves, and Adesso et al. (2021) note a possible mutualistic relationship among various members of Blastocatellia and Gammaproteobacteria in caves. This relationship should be explored to understand the survival of K-strategists in oligotrophic environments. Finally, mechanisms involved in cave biofilm formation, maintenance, and defense are understudied. Understanding enzymes of cave biofilms may partly explain why certain taxa appear in higher relative abundances and elaborate on microbe-microbe interactions in the cave environment, as these interactions are ultimately what attracts attention to the pharmaceutical and biotechnology industries to cave microbiomes (Radhakrishnan et al., 2014; Pawlowski et al., 2016; Pipite et al., 2022).

REFERENCES

- Addesso, R., Gonzalez-Pimentel, J.L., D'Angeli, I.M., De Waele, J., Saiz-Jimenez, C., Jurado, V., Miller, A.Z., Cubero, B., Vigliotta, G. and Baldantoni, D., 2021. Microbial community characterizing vermiculations from karst caves and its role in their formation. *Microbial Ecology*, 81(4): 884-896.
- Adesoji, A.T. and Ogunjobi, A.A., 2016. Detection of extended spectrum beta-lactamases resistance genes among bacteria isolated from selected drinking water distribution channels in southwestern Nigeria. *BioMed Research International*, 2016(9): 7149295.
- Adrian, L. and Marco-Urrea, E., 2016. Isotopes in geobiochemistry: tracing metabolic pathways in microorganisms of environmental relevance with stable isotopes. *Current Opinion in Biotechnology*, 41: 19-25.
- Aeschlimann, J.R., 2003. The role of multidrug efflux pumps in the antibiotic resistance of *Pseudomonas aeruginosa* and other gram-negative bacteria: insights from the Society of Infectious Diseases Pharmacists. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, 23(7): 916-924.
- Ai, J., Guo, J., Li, Y., Zhong, X., Lv, Y., Li, J. and Yang, A., 2022. The diversity of microbes and prediction of their functions in karst caves under the influence of human tourism activities—a case study of Zhijin Cave in Southwest China. *Environmental Science and Pollution Research*, 29(17): 25858-25868.
- Alber, B.E. and Fuchs, G., 2002. Propionyl-coenzyme A synthase from *Chloroflexus aurantiacus*, a key enzyme of the 3-hydroxypropionate cycle for autotrophic CO₂ fixation. *Journal of Biological Chemistry*, 277(14): 12137-12143.

Alneberg, J., Bjarnason, B.S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., Lahti, L., Loman, N.J., Andersson, A.F. and Quince, C., 2014. Binning metagenomic contigs by coverage and composition. *Nature Methods*, 11(11): 1144-1146.

Alonso, L., Pommier, T., Kaufmann, B., Dubost, A., Chapulliot, D., Doré, J., Douady, C.J. and Moënné-Loccoz, Y., 2019. Anthropization level of Lascaux Cave microbiome shown by regional-scale comparisons of pristine and anthropized caves. *Molecular Ecology*, 28(14): 3383-3394.

Anandan, R., Dharumadurai, D. and Manogaran, G.P., 2016. An introduction to Actinobacteria, Actinobacteria-basics and biotechnological applications. IntechOpen, Rijeka, pp. Ch. 1.

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

Appelbaum, P.C. and Bozdogan, B., 2004. Vancomycin resistance in *Staphylococcus aureus*. *Clinics in Laboratory Medicine*, 24(2): 381-402.

Arkin, A.P., Cottingham, R.W., Henry, C.S., Harris, N.L., Stevens, R.L., Maslov, S., Dehal, P., Ware, D., Perez, F. and Canon, S., 2018. KBase: the United States Department of Energy Systems Biology Knowledgebase. *Nature Biotechnology*, 36(7): 566-569.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M. and Kubal, M., 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, 9(1): 1-15.

Bao, Y., Dolfing, J., Guo, Z., Chen, R., Wu, M., Li, Z., Lin, X. and Feng, Y., 2021. Important ecophysiological roles of non-dominant Actinobacteria in plant residue decomposition, especially in less fertile soils. *Microbiome*, 9: 1-17.

Bar-Even, A., Noor, E. and Milo, R., 2012. A survey of carbon fixation pathways through a quantitative lens. *Journal of Experimental Botany*, 63(6): 2325-2342.

Barka, E.A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.-P., Clément, C., Ouhdouch, Y. and van Wezel, G.P., 2016. Taxonomy, physiology, and natural products of Actinobacteria. *Microbiology and Molecular Biology Reviews*, 80(1): 1-43.

Barton, H.A. and Jurado, V., 2007. What's up down there? Microbial diversity in caves. *American Society for Microbiology. Microbe* 2:132-38

Bay, S.K., Waite, D.W., Dong, X., Gillor, O., Chown, S.L., Hugenholtz, P. and Greening, C., 2021. Chemosynthetic and photosynthetic bacteria contribute differentially to primary production across a steep desert aridity gradient. *The ISME Journal*, 15(11): 3339-3356.

Benaud, N., Chelliah, D.S., Wong, S.Y. and Ferrari, B.C., 2022. Soil substrate culturing approaches recover diverse members of Actinomycetota from desert soils of Herring Island, East Antarctica. *Extremophiles*, 26(2): 24.

Bendia, A.G., Callefo, F., Araújo, M.N., Sanchez, E., Teixeira, V.C., Vasconcelos, A., Battilani, G., Pellizari, V.H., Rodrigues, F. and Galante, D., 2022. Metagenome-assembled genomes from Monte Cristo Cave (Diamantina, Brazil) reveal prokaryotic lineages as functional models for life on Mars. *Astrobiology*, 22(3): 293-312.

Berks, B.C., Ferguson, S.J., Moir, J.W. and Richardson, D.J., 1995. Enzymes and associated electron transport systems that catalyse the respiratory reduction of nitrogen oxides and oxyanions. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1232(3): 97-173.

Bernardo, A.A., Bernardo, C.M., Espiritu, D.J. and Arruda, J.A., 2006. The sodium bicarbonate cotransporter: structure, function, and regulation, *Seminars in nephrology*. Elsevier, pp. 352-360.

Blair, N., Leu, A., Olsen, J., Kwong, E. and Des Marais, D., 1985. Carbon isotopic fractionation in heterotrophic microbial metabolism. *Applied and Environmental Microbiology*, 50(4): 996-1001.

Bolger, A.M., Lohse, M. and Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15): 2114-20.

Bonafonte, M.A., Solano, C., Sesma, B., Alvarez, M., Montuenga, L., García-Ros, D. and Gamazo, C., 2000. The relationship between glycogen synthesis, biofilm formation and virulence in *Salmonella enteritidis*. *FEMS Microbiology Letters*, 191(1): 31-36.

Borrero-de Acuña, J.M., Rohde, M., Wissing, J., Jänsch, L., Schobert, M., Molinari, G., Timmis, K.N., Jahn, M. and Jahn, D., 2016. Protein network of the *Pseudomonas aeruginosa* denitrification apparatus. *Journal of Bacteriology*, 198(9): 1401-1413.

Borsodi, A.K., Knáb, M., Krett, G., Makk, J., Márialigeti, K., Erőss, A. and Mádl-Szőnyi, J., 2012. Biofilm bacterial communities inhabiting the cave walls of the Buda Thermal Karst System, Hungary. *Geomicrobiology Journal*, 29(7): 611-627.

Boschker, H. and Middelburg, J., 2002. Stable isotopes and biomarkers in microbial ecology. *FEMS Microbiology Ecology*, 40(2): 85-95.

Boudjelal, F., Zitouni, A., Bouras, N., Schumann, P., Spröer, C., Sabaou, N. and Klenk, H.-P., 2015. *Actinoalloteichus hoggarensis* sp. nov., an actinomycete isolated from Saharan soil. *International Journal of Systematic and Evolutionary Microbiology*, 65(Pt_6): 2006-2010.

Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K., Schulz, F., Jarett, J., Rivers, A.R., Eloie-Fadrosch, E.A., Tringe, S.G., Ivanova, N.N., Copeland, A., Clum, A., Becraft, E.D., Malmstrom, R.R., Birren, B., Podar, M., Bork, P., Weinstock, G.M., Garrity, G.M., Dodsworth, J.A., Yooseph, S., Sutton, G., Glöckner, F.O., Gilbert, J.A., Nelson,

W.C., Hallam, S.J., Jungbluth, S.P., Ettema, T.J.G., Tighe, S., Konstantinidis, K.T., Liu, W.-T., Baker, B.J., Rattei, T., Eisen, J.A., Hedlund, B., McMahon, K.D., Fierer, N., Knight, R., Finn, R., Cochrane, G., Karsch-Mizrachi, I., Tyson, G.W., Rinke, C., Kyrpides, N.C., Schriml, L., Garrity, G.M., Hugenholtz, P., Sutton, G., Yilmaz, P., Meyer, F., Glöckner, F.O., Gilbert, J.A., Knight, R., Finn, R., Cochrane, G., Karsch-Mizrachi, I., Lapidus, A., Meyer, F., Yilmaz, P., Parks, D.H., Murat Eren, A., Schriml, L., Banfield, J.F., Hugenholtz, P., Woyke, T. and The Genome Standards, C., 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature Biotechnology*, 35(8): 725-731.

Bowman, J.P., 2014. The Family Methylococcaceae. In: E. Rosenberg, E.F. DeLong, S. Lory, E. Stackebrandt and F. Thompson (Editors), *The Prokaryotes: Gammaproteobacteria*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 411-440.

Brannen-Donnelly, K. and Engel, A.S., 2015. Bacterial diversity differences along an epigenic cave stream reveal evidence of community dynamics, succession, and stability. *Frontiers in Microbiology*, 6: 729.

Buck, L.D., Paladino, M.M., Nagashima, K., Brezel, E.R., Holtzman, J.S., Urso, S.J. and Ryno, L.M., 2021. Temperature-dependent influence of FliA overexpression on PHL628 *E. coli* biofilm growth and composition. *Frontiers in Cellular and Infection Microbiology*, 11: 1282.

Burgess, R.R., 2001. Sigma factors. In: S. Brenner and J.H. Miller (Editors), *Encyclopedia of Genetics*. Academic Press, New York, pp. 1831-1834.

Cao, C.-L., Zhou, X.-Q., Qin, S., Tao, F.-X., Jiang, J.-H. and Lian, B., 2015. *Lentzea guizhouensis* sp. nov., a novel lithophilous actinobacterium isolated from limestone from the Karst area, Guizhou, China. *Antonie van Leeuwenhoek*, 108: 1365-1372.

Cao, C., Yuan, B., Qin, S., Jiang, J., Tao, F. and Lian, B., 2017. *Lentzea pudingi* sp. nov., isolated from a weathered limestone sample in a karst area. International Journal of Systematic and Evolutionary Microbiology, 67(11): 4873-4878.

Carvalho, F.M., Souza, R.C., Barcellos, F.G., Hungria, M. and Vasconcelos, A.T.R., 2010. Genomic and evolutionary comparisons of diazotrophic and pathogenic bacteria of the order Rhizobiales. BMC Microbiology, 10(1): 1-15.

Chen, H. and Wang, Z. (2015). Insights on nitrate respiration by *Shewanella*. Marine Molecular Biology and Ecology, 428, 73-80.

Chen, Y., Wu, L., Boden, R., Hillebrand, A., Kumaresan, D., Moussard, H., Baciú, M., Lu, Y. and Colin Murrell, J., 2009. Life without light: microbial diversity and evidence of sulfur-and ammonium-based chemolithotrophy in Movile Cave. The ISME Journal, 3(9): 1093-1104.

Chen, A.I. and Goulian, M., 2018. A network of regulators promotes dehydration tolerance in *Escherichia coli*. Environmental Microbiology, 20(3): 1283-1295.

Chen, T., Liu, N., Ren, P., Xi, X., Yang, L., Sun, W., Yu, B., Ying, H., Ouyang, P. and Liu, D., 2019. Efficient biofilm-based fermentation strategies for L-threonine production by *Escherichia coli*. Frontiers in Microbiology, 10: 1773.

Cimermanova, M., Pristas, P. and Piknova, M., 2021. Biodiversity of Actinomycetes from heavy metal contaminated technosols. Microorganisms, 9(8): 1635.

Cordero, P.R.F., Bayly, K., Man Leung, P., Huang, C., Islam, Z.F., Schittenhelm, R.B., King, G.M. and Greening, C., 2019. Atmospheric carbon monoxide oxidation is a widespread mechanism supporting microbial survival. The ISME Journal, 13(11): 2868-2881.

Cuezva, S., Fernandez-Cortes, A., Porca, E., Pašić, L., Jurado, V., Hernandez-Marine, M., Serrano-Ortiz, P., Hermosin, B., Cañaveras, J.C. and Sanchez-Moral, S., 2012. The biogeochemical role of Actinobacteria in Altamira cave, Spain. *FEMS Microbiology Ecology*, 81(1): 281-290.

De Mandal, S., Chatterjee, R. and Kumar, N.S., 2017. Dominant bacterial phyla in caves and their predicted functional roles in C and N cycle. *BMC Microbiology*, 17(1).

Demirci, H. and Emel, O., 2020. Metagenomic analysis of the microbial community in Çal Cave soil to elucidate biotechnological potential. *Turkish Journal of Forestry*, 21(3): 254-259.

Dong, T., Yu, R. and Schellhorn, H., 2011. Antagonistic regulation of motility and transcriptome expression by RpoN and RpoS in *Escherichia coli*. *Molecular Microbiology*, 79(2): 375-386.

Dyksma, S., Bischof, K., Fuchs, B.M., Hoffmann, K., Meier, D., Meyerdierks, A., Pjevac, P., Probandt, D., Richter, M. and Stepanauskas, R., 2016. Ubiquitous Gammaproteobacteria dominate dark carbon fixation in coastal sediments. *The ISME Journal*, 10(8): 1939-1953.

Engel, A.S., Porter, M.L., Stern, L.A., Quinlan, S., Bennett, P.C., 2004. Bacterial diversity and ecosystem function of filamentous microbial mats from aphotic (cave) sulfidic springs dominated by chemolithoautotrophic “Epsilonproteobacteria.” *FEMS Microbiology Ecology* 51, 31–53.

Eroshenko, D.V., Polyudova, T.V. and Pyankova, A.A., 2020. VapBC and MazEF toxin/antitoxin systems in the regulation of biofilm formation and antibiotic tolerance in nontuberculous mycobacteria. *International Journal of Mycobacteriology*, 9(2): 156.

Fang, F.C., 2004. Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. *Nature Reviews Microbiology*, 2(10): 820-832.

Fang, B.-Z., Han, M.-X., Liu, L., Zhang, Z.-T., Liu, W.-L., Shen, J.-T., Wang, Y., Zhang, W.-Q., Wei, D.-Q. and Li, W.-J., 2017. *Lentzea cavernae* sp. nov., an actinobacterium isolated from a karst cave sample, and emended description of the genus *Lentzea*. International Journal of Systematic and Evolutionary Microbiology, 67(7): 2357-2362.

Feng, H., Sun, Y., Zhi, Y., Wei, X., Luo, Y., Mao, L. and Zhou, P., 2014. Identification and characterization of the nitrate assimilation genes in the isolate of *Streptomyces griseorubens* JSD-1. Microbial Cell Factories, 13(1):174.

Franco, C.M.M. and Labeda, D.P., 2014. The Order Pseudonocardiales. Springer Berlin Heidelberg, pp. 743-860.

Frazier, V.E., Carbon Metabolism in Cave Subaerial Biofilms. Master's Thesis, University of Tennessee, 2020. https://trace.tennessee.edu/utk_gradthes/5867.

Fujiwara, M., Yamasaki, S., Morita, Y. and Nishino, K., 2022. Evaluation of efflux pump inhibitors of MexAB-or MexXY-OprM in *Pseudomonas aeruginosa* using nucleic acid dyes. Journal of Infection and Chemotherapy, 28(5): 595-601.

Gabriel, C.R. and Northup, D.E., 2013. Microbial ecology: caves as an extreme habitat. Cave microbiomes: a novel resource for drug discovery. SpringerBriefs in Microbiology, vol 1. Springer, New York, NY: 85-108.

Gao, H., Ma, L., Qin, Q., Qiu, Y., Zhang, J., Li, J., Lou, J., Diao, B., Zhao, H. and Shi, Q., 2020. Fur represses *Vibrio cholerae* biofilm formation via direct regulation of *vieSAB*, *cdgD*, *vpsU*, and *vpsA-K* transcription. Frontiers in Microbiology, 11: 587159.

Garritano, A.N., Song, W. and Thomas, T., 2022. Carbon fixation pathways across the bacterial and archaeal tree of life. PNAS Nexus, 1(5): pgac226.

Ghezzi, D., Sauro, F., Columbu, A., Carbone, C., Hong, P.-Y., Vergara, F., De Waele, J. and Cappelletti, M., 2021. Transition from unclassified Ktedonobacterales to Actinobacteria during amorphous silica precipitation in a quartzite cave environment. *Scientific Reports*, 11(1): 1-18.

Gilch, S., Meyer, O. and Schmidt, I., 2009. A soluble form of ammonia monooxygenase in *Nitrosomonas europaea*. *De Gruyter*. vol. 390, no. 9, 2009, pp. 863-873.

Gonzalez-Pimentel, J.L., Miller, A.Z., Jurado, V., Laiz, L., Pereira, M.F. and Saiz-Jimenez, C., 2018. Yellow coloured mats from lava tubes of La Palma (Canary Islands, Spain) are dominated by metabolically active Actinobacteria. *Scientific Reports*, 8(1): 1944.

Gonzalez-Pimentel, J.L., Martin-Pozas, T., Jurado, V., Miller, A.Z., Caldeira, A.T., Fernandez-Lorenzo, O., Sanchez-Moral, S. and Saiz-Jimenez, C., 2021. Prokaryotic communities from a lava tube cave in La Palma Island (Spain) are involved in the biogeochemical cycle of major elements. *PeerJ*, 9: e11386.

Greening, C., Islam, Z.F. and Bay, S.K., 2022. Hydrogen is a major lifeline for aerobic bacteria. *Trends in Microbiology*, 30(4): 330-337.

Guerrero-Cruz, S., Vaksmaa, A., Horn, M.A., Niemann, H., Pijuan, M. and Ho, A., 2021. Methanotrophs: discoveries, environmental relevance, and a perspective on current and future applications. *Frontiers in Microbiology*, 12: 678057.

Hamilton, S., Bongaerts, R.J., Mulholland, F., Cochrane, B., Porter, J., Lucchini, S., Lappin-Scott, H.M. and Hinton, J.C., 2009. The transcriptional programme of *Salmonella enterica* serovar *Typhimurium* reveals a key role for tryptophan metabolism in biofilms. *BMC genomics*, 10: 1-21.

Handa, N., Morimatsu, K., Lovett, S.T. and Kowalczykowski, S.C., 2009. Reconstitution of initial steps of dsDNA break repair by the RecF pathway of *E. coli*. *Genes & Development*, 23(10): 1234-1245.

Hardeman, W.D., Miller, R.A., & Swingle, G.D. (1966) Geologic Map of Tennessee: Division of Geology, Tennessee Department of Environment and Conservation.

Hathaway, J.J.M., Sinsabaugh, R.L., Dapkevicius, M.D.L.N. and Northup, D.E., 2014. Diversity of ammonia oxidation (*amoA*) and nitrogen fixation (*nifH*) genes in lava caves of Terceira, Azores, Portugal. *Geomicrobiology Journal*, 31(3): 221-235.

Hatzenpichler, R., 2012. Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. *Applied and Environmental Microbiology*, 78(21): 7501-7510.

Hayes, J., 2001. Fractionation of carbon and hydrogen isotopes in biosynthetic processes. *Reviews in Mineralogy and Geochemistry*, 43: 225-277.

Hennecke, H., Bothe, H. and Trebst, A., 1981. Biology of inorganic nitrogen and sulfur. *Proceedings in Life Sciences*. Springer, Berlin, Heidelberg, pp. 64–77.

Hill, P., Krišťufek, V., Dijkhuizen, L., Boddy, C., Kroetsch, D. and van Elsas, J.D., 2011. Land use intensity controls actinobacterial community structure. *Microbial Ecology*, 61: 286-302.

Holten, K.B. and Onusko, E.M., 2000. Appropriate prescribing of oral beta-lactam antibiotics. *American Family Physician*, 62(3): 611-620.

Huber, K.J., Wüst, P.K., Rohde, M., Overmann, J. and Foesel, B.U., 2014. *Aridibacter famidurans* gen. nov., sp. nov. and *Aridibacter kavangonensis* sp. nov., two novel members of subdivision 4 of the Acidobacteria isolated from semiarid savannah soil. *International Journal of Systematic and Evolutionary Microbiology*, 64(Pt_6): 1866-1875.

Hügler, M. and Fuchs, G., 2005. Assaying for the 3-hydroxypropionate cycle of carbon fixation. *Methods in Enzymology*, 397: 212-221.

Hügler, M. and Sievert, S.M., 2011. Beyond the Calvin cycle: autotrophic carbon fixation in the ocean. *Annual Review of Marine Science*, 3: 261-289.

Hutchins, B.T., Engel, A.S., Nowlin, W.H. and Schwartz, B.F., 2016. Chemolithoautotrophy supports macroinvertebrate food webs and affects diversity and stability in groundwater communities. *Ecology*, 97(6): 1530-1542.

Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W. and Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11(1): 119.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T. and Aluru, S., 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature communications*, 9(1): 5114.

Jakubiec-Krzesniak, K., Rajnisz-Mateusiak, A., Guspiel, A., Ziemska, J. and Solecka, J., 2018. Secondary metabolites of actinomycetes and their antibacterial, antifungal and antiviral properties. *Polish Journal of Microbiology*, 67(3): 259-272.

Jurado, V., Gonzalez-Pimentel, J.L., Miller, A.Z., Hermosin, B., D'Angeli, I.M., Tognini, P., De Waele, J. and Saiz-Jimenez, C., 2020. Microbial communities in vermiculation deposits from an Alpine cave. *Frontiers in Earth Science*, 8: 586248.

Jurado, V., Del Rosal, Y., de Cisneros, C.J., Liñan, C., Martin-Pozas, T., Gonzalez-Pimentel, J.L., Hermosin, B. and Saiz-Jimenez, C., 2022. Microbial communities in carbonate precipitates from drip waters in Nerja Cave, Spain. *PeerJ*, 10: e13399.

Kajan, K., Cukrov, N., Cukrov, N., Bishop-Pierce, R. and Orlić, S., 2021. Microeukaryotic and prokaryotic diversity of anchialine caves from eastern Adriatic Sea Islands. *Microbial Ecology*: 1-14.

Kämpfer, P., Kohlweyer, U., Thiemer, B. and Andreesen, J.R., 2006. *Pseudonocardia tetrahydrofuranoxydans* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*, 56(7): 1535-1538.

Kamp, A., Høgslund, S., Risgaard-Petersen, N. and Stief, P., 2015. Nitrate storage and dissimilatory nitrate reduction by eukaryotic microbes. *Frontiers in Microbiology*: 1492.

Kang, D.D., Froula, J., Egan, R. and Wang, Z., 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ*, 3: e1165.

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. and Ishiguro-Watanabe, M., 2023. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research*, 51(D1): D587-d592.

Katti, A.K.S., AK, S. and Mudgulkar, S.B., 2022. Diversity and classification of rare Actinomycetes, Actinobacteria: Ecology, Diversity, Classification and Extensive Applications. Springer, pp. 117-142.

Keenan, D., 2016. *Encyclopedia of Food and Health*. Academic Press: Cambridge, MA, USA. P.401-416

Khmelenina, V.N., Colin Murrell, J., Smith, T.J. and Trotsenko, Y.A., 2018. Physiology and Biochemistry of the Aerobic Methanotrophs. In: F. Rojo (Editor), *Aerobic Utilization of Hydrocarbons, Oils and Lipids*. Springer International Publishing, Cham, pp. 1-25.

King, G.M. and Weber, C.F., 2007. Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nature Reviews Microbiology*, 5(2): 107-118.

Kosznik-Kwaśnicka, K., Golec, P., Jaroszewicz, W., Lubomska, D. and Piechowicz, L., 2022. Into the unknown: Microbial communities in caves, their role, and potential use. *Microorganisms*, 10(2): 222.

Kou, Y., Li, J., Wang, Y., Li, C., Tu, B., Yao, M. and Li, X., 2017. Scale-dependent key drivers controlling methane oxidation potential in Chinese grassland soils. *Soil Biology and Biochemistry*, 111: 104-114.

Kurade, M.B., Saha, S., Salama, E.-S., Patil, S.M., Govindwar, S.P. and Jeon, B.-H., 2019. Acetoclastic methanogenesis led by *Methanosarcina* in anaerobic co-digestion of fats, oil and grease for enhanced production of methane. *Bioresource Technology*, 272: 351-359.

Lacasse, M.J. and Zamble, D.B., 2016. [NiFe]-hydrogenase maturation. *Biochemistry*, 55(12): 1689-701.

Lancaster, K.M., Caranto, J.D., Majer, S.H. and Smith, M.A., 2018. Alternative bioenergy: updates to and challenges in nitrification metalloenzymology. *Joule*, 2(3): 421-441.

Lavoie, K.H., Winter, A.S., Read, K.J.H., Hughes, E.M., Spilde, M.N. and Northup, D.E., 2017. Comparison of bacterial communities from lava cave microbial mats to overlying surface soils from Lava Beds National Monument, USA. *PLOS ONE*, 12(2): e0169339.

Lawton, T.J., Bowen, K.E., Sayavedra-Soto, L.A., Arp, D.J. and Rosenzweig, A.C., 2013. Characterization of a nitrite reductase involved in nitrifier denitrification. *Journal of Biological Chemistry*, 288(35): 25575-25583.

- Lazarus, J.E., Warr, A.R., Kuehl, C.J., Giorgio, R.T., Davis, B.M. and Waldor, M.K., 2019. A new suite of allelic-exchange vectors for the scarless modification of proteobacterial genomes. *Applied and Environmental Microbiology*, 85(16): e00990-19.
- Lee, J.Y., Moon, S.S., Yun, B.S., Yoo, I.D. and Hwang, B.K., 2004. Thiobutacin, a novel antifungal and antioomycete antibiotic from *Lechevalieria aerocolonigenes*. *Journal of Natural Products*, 67(12): 2076-2078.
- Lee, K., Lee, B., Ryu, J., Kim, D., Kim, Y. and Lim, S.K., 2016. Increased vancomycin production by overexpression of MbtH-like protein in *Amycolatopsis orientalis* KFCC10990P. *Letters in Applied Microbiology*, 63(3): 222-228.
- Lee, L.-H., Goh, B.-H. and Chan, K.-G., 2020. Actinobacteria: Prolific producers of bioactive metabolites. *Frontiers Media SA*, pp. 1612.
- Leong, C.G., Bloomfield, R.A., Boyd, C.A., Dornbusch, A.J., Lieber, L., Liu, F., Owen, A., Slay, E., Lang, K.M. and Lostroh, C.P., 2017. The role of core and accessory type IV pilus genes in natural transformation and twitching motility in the bacterium *Acinetobacter baylyi*. *PloS one*, 12(8): e0182139.
- Lerch, R.N., 2011. Contaminant transport in two central Missouri karst recharge areas. *Journal of Cave and Karst Studies*, 2011 v.73 no.2: pp. 99-113.
- Li, X., Chopp, D.L., Russin, W.A., Brannon, P.T., Parsek, M.R. and Packman, A.I., 2015. Spatial patterns of carbonate biomineralization in biofilms. *Applied and Environmental Microbiology*, 81(21): 7403-7410.
- Li, D., Liu, C.M., Luo, R., Sadakane, K. and Lam, T.W., 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, 31(10): 1674-6.

Li, Q., Zhang, B., Yang, X. and Ge, Q., 2018. Deterioration-associated microbiome of stone monuments: Structure, variation, and assembly. *Applied and Environmental Microbiology*, 84(7): e02680-17.

Liang, B., Zhao, Y. and Yang, J., 2020. Recent advances in developing artificial autotrophic microorganism for reinforcing CO₂ fixation. *Frontiers in Microbiology*, 11: 592631.

Lin, J., Goldman, B. and Stewart, V., 1993. Structures of genes *nasA* and *nasB*, encoding assimilatory nitrate and nitrite reductases in *Klebsiella pneumoniae* M5al. *Journal of Bacteriology*, 175(8): 2370-2378.

Liu, Z., Zhang, Y., Fa, K., Zhao, H., Qin, S., Yan, R. and Wu, B., 2018. Desert soil bacteria deposit atmospheric carbon dioxide in carbonate precipitates. *Catena*, 170: 64-72.

Liu, D., Ma, Q., Valiela, I., Anderson, D.M., Keesing, J.K., Gao, K., Zhen, Y., Sun, X. and Wang, Y., 2020. Role of C4 carbon fixation in *Ulva prolifera*, the macroalga responsible for the world's largest green tides. *Communications Biology*, 3(1): 494.

Liu, C., Sun, D., Zhu, J., Liu, J. and Liu, W., 2020. The regulation of bacterial biofilm formation by cAMP-CRP: a mini-review. *Frontiers in Microbiology*, 11: 802.

Liu, X., Ye, Y., Zhu, Y., Wang, L., Yuan, L., Zhu, J. and Sun, A., 2021. Involvement of RpoN in regulating motility, biofilm, resistance, and spoilage potential of *Pseudomonas fluorescens*. *Frontiers in Microbiology*, 12: 641844.

Long, Y., Jiang, J., Hu, X., Zhou, J., Hu, J. and Zhou, S., 2019. Actinobacterial community in Shuanghe Cave using culture-dependent and -independent approaches. *World Journal of Microbiology and Biotechnology*, 35(10): 153.

Lovely, D.R., Ueki, T., Zhang, T., Malvankar, N.S., Shrestha, P.M., Flanagan, K.A., Aklujkar, M., Butler, J.E., Giloteaux, L., Rotaru, A.-E., Holmes, D.E., Franks, A.E., Orellana, R., Risso, C. and Nevin, K.P., 2011. *Geobacter*: The microbe electric's physiology, ecology, and practical applications. In: R.K. Poole (Editor), *Advances in Microbial Physiology*. Academic Press, pp. 1-100.

Lü, X.-F., He, Q.-F., Wang, F.-K., Zhao, R.-Y. and Zhang, H., 2018. Impact of tourism on bacterial communities of karst underground river: A case study from two caves in Fengdu, Chongqing. *Huan Jing Ke Xue*, 39(5): 2389-2399.

Lv, P., Luo, J., Zhuang, X., Zhang, D., Huang, Z. and Bai, Z., 2017. Diversity of culturable aerobic denitrifying bacteria in the sediment, water and biofilms in Liangshui River of Beijing, China. *Scientific Reports*, 7(1): 10032.

Lynch, R.C., Darcy, J.L., Kane, N.C., Nemergut, D.R. and Schmidt, S.K., 2014. Metagenomic evidence for metabolism of trace atmospheric gases by high-elevation desert Actinobacteria. *Frontiers in Microbiology*, 5: 698.

Ma, D., Gu, H., Shi, Y., Huang, H., Sun, D. and Hu, Y., 2021. *Edwardsiella piscicida* YefM-YoeB: a type II Toxin-Antitoxin system that is related to antibiotic resistance, biofilm formation, serum survival, and host infection. *Frontiers in Microbiology*, 12: 646299.

Ma, L., Huang, X., Wang, H., Yun, Y., Cheng, X., Liu, D., Lu, X. and Qiu, X., 2021. Microbial interactions drive distinct taxonomic and potential metabolic responses to habitats in karst cave ecosystem. *Microbiology Spectrum*, 9(2): e01152-21.

Maciejewska, M., Całusińska, M., Cornet, L., Adam, D., Pessi, I.S., Malchair, S., Delfosse, P., Baurain, D., Barton, H.A. and Carnol, M., 2018. High-throughput sequencing analysis of the actinobacterial spatial diversity in moonmilk deposits. *Antibiotics*, 7(2): 27.

- Marti, T.M., Kunz, C. and Fleck, O., 2002. DNA mismatch repair and mutation avoidance pathways. *Journal of Cellular Physiology*, 191(1): 28-41.
- Mahendra, S. and Alvarez-Cohen, L., 2005. *Pseudonocardia dioxanivorans* sp. nov., a novel actinomycete that grows on 1, 4-dioxane. *International Journal of Systematic and Evolutionary Microbiology*, 55(2): 593-598.
- Maisonneuve, E., Shakespeare, L.J., Jørgensen, M.G. and Gerdes, K., 2011. Bacterial persistence by RNA endonucleases. *Proceedings of the National Academy of Sciences*, 108(32): 13206-13211.
- Martin-Pozas, T., Gonzalez-Pimentel, J.L., Jurado, V., Cuezva, S., Dominguez-Moñino, I., Fernandez-Cortes, A., Cañaveras, J.C., Sanchez-Moral, S. and Saiz-Jimenez, C., 2020. Microbial activity in subterranean ecosystems: Recent Advances. *Applied Sciences*, 10(22): 8130.
- Martin-Pozas, T., Cuezva, S., Fernandez-Cortes, A., Cañaveras, J.C., Benavente, D., Jurado, V., Saiz-Jimenez, C., Janssens, I., Seijas, N. and Sanchez-Moral, S., 2022. Role of subterranean microbiota in the carbon cycle and greenhouse gas dynamics. *Science of the Total Environment*, 831: 154921.
- Miticka, H., Rowley, G., Rezuchova, B., Homerova, D., Humphreys, S., Farn, J., Roberts, M. and Kormanec, J., 2003. Transcriptional analysis of the *rpoE* gene encoding extracytoplasmic stress response sigma factor σE in *Salmonella enterica* serovar Typhimurium. *FEMS Microbiology Letters*, 226(2): 307-314.
- Miyaji, A., 2011. Particulate methane monooxygenase from *Methylosinus trichosporium* OB3b, *Methods in Enzymology*. Elsevier, pp. 211-225.
- Miyamoto, H., Asano, F., Ishizawa, K., Suda, W., Miyamoto, H., Tsuji, N., Matsuura, M., Tsuboi, A., Ishii, C. and Nakaguma, T., 2022. A potential network structure of symbiotic

bacteria involved in carbon and nitrogen metabolism of wood-utilizing insect larvae. *Science of The Total Environment*, 836: 155520.

Moënne-Loccoz, P., 2007. Spectroscopic characterization of heme iron–nitrosyl species and their role in NO reductase mechanisms in diiron proteins. *Natural Product Reports*, 24(3): 610-620.

Mohan, S.B. and Cole, J.A., 2007. The dissimilatory reduction of nitrate to ammonia by anaerobic bacteria, *Biology of the nitrogen cycle*. Elsevier, pp. 93-106.

Moreno-Vivián, C., Cabello, P.n., Martínez-Luque, M., Blasco, R. and Castillo, F., 1999. Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases. *Journal of Bacteriology*, 181(21): 6573-6584.

Murray, A.B. and McKenna, R., 2019. β -Carbonic anhydrases, *Carbonic Anhydrases*. Elsevier, pp. 55-77.

Narsing Rao, M.P. and Li, W.-J., 2022. Diversity of actinobacteria in various habitats, *Actinobacteria: Microbiology to Synthetic Biology*. Springer, pp. 37-58.

Nicolosi, G., Galdenzi, S., Messina, M.A., Miller, A.Z., Petralia, S., Sarbu, S.M. and Isaia, M., 2022. Sulfidic habitats in the gypsum karst system of Monte Conca (Italy) host a chemoautotrophically supported invertebrate community. *International Journal of Environmental Research and Public Health*, 19(5): 2671.

Niyomvong, N., Pathom-Aree, W., Thamchaipenet, A. and Duangmal, K., 2012. Actinomycetes from tropical limestone caves. *Chiang Mai Journal of Science*, 39(3): 373-388.

Nouioui, I., Carro, L., García-López, M., Meier-Kolthoff, J.P., Woyke, T., Kyrpides, N.C., Pukall, R., Klenk, H.-P., Goodfellow, M. and Göker, M., 2018. Genome-based taxonomic classification of the phylum Actinobacteria. *Frontiers in Microbiology*, 9: 2007.

Ojeda, L., Vadillo, I., Etiope, G., Benavente, J., Liñán, C., del Rosal, Y., Tapia, S.T., Moríñigo, M.Á. and Carrasco, F., 2019. Methane sources and sinks in karst systems: the Nerja cave and its vadose environment (Spain). *Geochimica et Cosmochimica Acta*, 259: 302-315.

Olanrewaju, O.S. and Babalola, O.O., 2019. *Streptomyces*: implications and interactions in plant growth promotion. *Applied Microbiology and Biotechnology*, 103: 1179-1188.

Oliveira, C., Gunderman, L., Coles, C.A., Lochmann, J., Parks, M., Ballard, E., Glazko, G., Rahmatallah, Y., Tackett, A.J. and Thomas, D.J., 2017. 16S rRNA gene-based metagenomic analysis of Ozark cave bacteria. *Diversity*, 9(3): 31.

Olm, M.R., Brown, C.T., Brooks, B. and Banfield, J.F., 2017. dRep: A tool for fast and accurate genome de-replication that enables tracking of microbial genotypes and improved genome recovery from metagenomes. Cold Spring Harbor Laboratory.

Orellana, L.H., Krüger, K., Sidhu, C., Amann, R., 2023. Comparing genomes recovered from time-series metagenomes using long- and short-read sequencing technologies. *Microbiome*, 11(105). <https://doi.org/10.1186/s40168-023-01557-3>

Oremland, R.S. and Culbertson, C.W., 1992. Importance of methane-oxidizing bacteria in the methane budget as revealed by the use of a specific inhibitor. *Nature*, 356(6368): 421-423.

Ortiz, M., Legatzki, A., Neilson, J.W., Fryslie, B., Nelson, W.M., Wing, R.A., Soderlund, C.A., Pryor, B.M. and Maier, R.M., 2014. Making a living while starving in the dark: metagenomic insights into the energy dynamics of a carbonate cave. *The ISME Journal*, 8(2): 478-491.

Ortiz, M., Leung, P.M., Shelley, G., Jirapanjawat, T., Nauer, P.A., Van Goethem, M.W., Bay, S.K., Islam, Z.F., Jordaan, K., Vikram, S., Chown, S.L., Hogg, I.D., Makhalanyane, T.P., Grinter, R., Cowan, D.A. and Greening, C., 2021. Multiple energy sources and metabolic

strategies sustain microbial diversity in Antarctic desert soils. *Proceedings of the National Academy of Sciences*, 118(45): e2025322118.

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A.R., Xia, F. and Stevens, R., 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res*, 42(Database issue): D206-14.

Pascual, J., Wüst, P.K., Geppert, A., Foesel, B.U., Huber, K.J. and Overmann, J., 2015. Novel isolates double the number of chemotrophic species and allow the first description of higher taxa in Acidobacteria subdivision 4. *Systematic and Applied Microbiology*, 38(8): 534-544.

Park, S.W., Park, S.T., Lee, J.E. and Kim, Y.M., 2008. *Pseudonocardia carboxydivorans* sp. nov., a carbon monoxide-oxidizing actinomycete, and an emended description of the genus *Pseudonocardia*. *International journal of Systematic and Evolutionary Microbiology*, 58(11): 2475-2478.

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

Patel, H., Buchad, H. and Gajjar, D., 2022. *Pseudomonas aeruginosa* persister cell formation upon antibiotic exposure in planktonic and biofilm state. *Scientific Reports*, 12(1): 16151.

Paula, C. Bichuette, M.E. and Selegim, M.H.R., 2020. Nutrient availability in tropical caves influences the dynamics of microbial biomass. *MicrobiologyOpen*, 9(7): e1044.

Pawlowski, A.C., Wang, W., Koteva, K., Barton, H.A., McArthur, A.G. and Wright, G.D., 2016. A diverse intrinsic antibiotic resistome from a cave bacterium. *Nature Communications*, 7(1): 13803.

Pelzmann, A., Ferner, M., Gnida, M., Meyer-Klaucke, W., Maisel, T., Meyer, O.. 2009. The CoxD protein of *Oligotropha carboxidovorans* is a predicted AAA+ ATPase chaperone involved in the biogenesis of the CO dehydrogenase [CuSMoO₂] cluster. *Journal of Biological Inorganic Chemistry*, 284(14):9578-86.

Pelzmann, A., Mickoleit, F., Meyer, O., 2014. Insights into the posttranslational assembly of the Mo-, S- and Cu-containing cluster in the active site of CO dehydrogenase of *Oligotropha carboxidovorans*. *Journal of Biological Inorganic Chemistry*, 19(8):1399-414.

Peng, Y., Leung, H.C.M., Yiu, S.M. and Chin, F.Y.L., 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*, 28(11): 1420-1428.

Penger, J., Conrad, R. and Blaser, M., 2012. Stable carbon isotope fractionation by methylotrophic methanogenic archaea. *Applied and Environmental Microbiology*, 78(21): 7596-7602.

Pipite, A., Lockhart, P.J., McLenachan, P.A., Christi, K., Kumar, D., Prasad, S. and Subramani, R., 2022. Isolation, antibacterial screening, and identification of bioactive cave dwelling bacteria in Fiji. *Frontiers in Microbiology*, 13: 1012867.

Pletnev, P., Pupov, D., Pshanichnaya, L., Esyunina, D., Petushkov, I., Nesterchuk, M., Osterman, I., Rubtsova, M., Mardanov, A., Ravin, N., Sergiev, P., Kulbachinskiy, A. and Dontsova, O., 2020. Rewiring of growth-dependent transcription regulation by a point mutation in region 1.1 of the housekeeping σ factor. *Nucleic Acids Research*, 48(19): 10802-10819.

Porca, E., Jurado, V., Žgur-Bertok, D., Saiz-Jimenez, C. and Pašić, L., 2012. Comparative analysis of yellow microbial communities growing on the walls of geographically distinct caves

indicates a common core of microorganisms involved in their formation. FEMS Microbiology Ecology, 81(1): 255-266.

Price, G.D. and Howitt, S.M., 2014. Topology mapping to characterize cyanobacterial bicarbonate transporters: BicA (SulP/SLC26 family) and SbtA. Molecular Membrane Biology, 31(6): 177-182.

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A. and Korobeynikov, A., 2020. Using SPAdes de novo assembler. Current Protocols in Bioinformatics, 70(1): e102.

Qin, S., Li, W.-J., Klenk, H.-P., Hozzein, W.N. and Ahmed, I., 2019. Actinobacteria in special and extreme habitats: diversity, function roles and environmental adaptations. Frontiers Media SA, pp. 944.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research, 41(D1): D590-D596.

Radhakrishnan, M., Raman, V.A., Bharathi, S., Balagurunathan, R. and Kumar, V., 2014. Anti MRSA and antitubercular activity of phenoxazinone containing molecule from Borra Caves Streptomyces sp. BCA1. International Journal of Pharmaceutical Sciences and Research, 5(12): 5342.

Rangseekaew, P. and Pathom-Aree, W., 2019. Cave actinobacteria as producers of bioactive metabolites. Frontiers in Microbiology, 10: 387.

Rastgar, M.G., Rasti, B. and Zamani, H., 2022. Ibuprofen involves with the reduced expression of pelD and pelF in pathogenic *Pseudomonas aeruginosa* strains. Archives of Microbiology, 204(6): 329.

Ravcheev, D.A. and Thiele, I., 2017. Comparative genomic analysis of the human gut microbiome reveals a broad distribution of metabolic pathways for the degradation of host-synthesized mucin glycans and utilization of mucin-derived monosaccharides. *Frontiers in genetics*, 8: 111.

Ray, A.E., Zaugg, J., Benaud, N., Chelliah, D.S., Bay, S., Wong, H.L., Leung, P.M., Ji, M., Terauds, A. and Montgomery, K., 2022. Atmospheric chemosynthesis is phylogenetically and geographically widespread and contributes significantly to carbon fixation throughout cold deserts. *The ISME Journal*, 16(11): 2547-2560.

Reichardt, N., Barclay, A.R., Weaver, L.T. and Morrison, D.J., 2011. Use of stable isotopes to measure the metabolic activity of the human intestinal microbiota. *Applied Environmental Microbiology*, 77(22): 8009-14.

Richardson, D.J., Berks, B.C., Russell, D.A., Spiro, S. and Taylor, C.J., 2001. Functional, biochemical and genetic diversity of prokaryotic nitrate reductases. *Cellular and Molecular Life Sciences*, 58(2): 165-178.

Riquelme, C., Rigal, F., Hathaway, J.J., Northup, D.E., Spilde, M.N., Borges, P.A., Gabriel, R., Amorim, I.R. and Dapkevicius, M.d.L.N., 2015. Cave microbial community composition in oceanic islands: disentangling the effect of different colored mats in diversity patterns of Azorean lava caves. *FEMS Microbiology Ecology*, 91(12).

Ruickoldt, J., 2023. On the coupling of the catalytical activities of the CODH/ACS complex from *Carboxydotherrmus hydrogenoformans*, Humboldt-Universität zu Berlin.

Ryabenko, E., 2013. Stable isotope methods for the study of the nitrogen cycle. *Topics in Oceanography*: 1-40.

Schabereiter-Gurtner, C., Saiz-Jimenez, C., Piñar, G., Lubitz, W. and Rölleke, S., 2004. Phylogenetic diversity of bacteria associated with Paleolithic paintings and surrounding rock walls in two Spanish caves (Llonin and La Garma). *FEMS Microbiology Ecology*, 47(2): 235-247.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H. and Robinson, C.J., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, 75(23): 7537-7541.

Schloss, P.D., 2020. Reintroducing mothur: 10 years later. *Applied and Environmental Microbiology*, 86(2): e02343-19.

Selensky, M.J., Masterson, A.L., Blank, J.G., Lee, S.C. and Osburn, M.R., 2021. Stable carbon isotope depletions in lipid biomarkers suggest subsurface carbon fixation in lava caves. *Journal of Geophysical Research: Biogeosciences*, 126(7): e2021JG006430.

Servin, J.A., Herbold, C.W., Skophammer, R.G. and Lake, J.A., 2008. Evidence excluding the root of the tree of life from the actinobacteria. *Molecular Biology and Evolution*, 25(1): 1-4.

Shaffer, M., Borton, M.A., McGivern, B.B., Zayed, A.A., La Rosa, Sabina L., Solden, L.M., Liu, P., Narrowe, A.B., Rodríguez-Ramos, J., Bolduc, B., Gazitúa, M.C., Daly, R.A., Smith, G.J., Vik, D.R., Pope, P.B., Sullivan, M.B., Roux, S. and Wrighton, Kelly C., 2020. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Research*, 48(16): 8883-8900.

Shapleigh, J.P., 2009. Dissimilatory and assimilatory nitrate reduction in the purple photosynthetic bacteria. Springer Netherlands, pp. 623-642.

Sharkey, T.D., 2019. Discovery of the canonical Calvin–Benson cycle. *Photosynthesis Research*, 140(2): 235-252.

Shavit, R., Lebendiker, M., Pasternak, Z., Burdman, S. and Helman, Y., 2016. The vapb–vapc operon of *Acidovorax citrulli* functions as a bona-fide toxin–antitoxin module. *Frontiers in Microbiology*, 6: 1499.

Sheldon, J.R., Yim, M.-S., Saliba, J.H., Chung, W.-H., Wong, K.-Y. and Leung, K.T., 2012. Role of rpoS in *Escherichia coli* O157: H7 strain H32 biofilm development and survival. *Applied and Environmental Microbiology*, 78(23): 8331-8339.

Shively, J.M., Van Keulen, G. and Meijer, W.G., 1998. Something from almost nothing: carbon dioxide fixation in chemoautotrophs. *Annual Review of Microbiology*, 52(1): 191-230.

Shivlata, L. and Satyanarayana, T., 2015. Thermophilic and alkaliphilic Actinobacteria: biology and potential applications. *Frontiers in Microbiology*, 6: 1014.

Sieber, C.M.K., Probst, A.J., Sharrar, A., Thomas, B.C., Hess, M., Tringe, S.G. and Banfield, J.F., 2018. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nature Microbiology*, 3(7): 836-843.

Silva, L.J., Crevelin, E.J., Souza, D.T., Lacerda-Júnior, G.V., de Oliveira, V.M., Ruiz, A.L.T.G., Rosa, L.H., Moraes, L.A.B. and Melo, I.S., 2020. Actinobacteria from Antarctica as a source for anticancer discovery. *Scientific Reports*, 10(1): 1-15.

Somprasong, N., Hall, C.M., Webb, J.R., Sahl, J.W., Wagner, D.M., Keim, P., Currie, B.J. and Schweizer, H.P., 2020. *Burkholderia ubonensis* meropenem resistance: insights into distinct properties of class A β -lactamases in *Burkholderia cepacia* complex and *Burkholderia pseudomallei* complex bacteria. *Mbio*, 11(2): e00592-20.

Sparacino-Watkins, C., Stolz, J.F. and Basu, P., 2014. Nitrate and periplasmic nitrate reductases. *Chemical Society Reviews*, 43(2): 676-706.

Spasov, E., Tsuji, J.M., Hug, L.A., Doxey, A.C., Sauder, L.A., Parker, W.J. and Neufeld, J.D., 2020. High functional diversity among *Nitrospira* populations that dominate rotating biological contactor microbial communities in a municipal wastewater treatment plant. *The ISME Journal*, 14(7): 1857-1872.

Stein, L.Y., Roy, R. and Dunfield, P.F., 2012. Aerobic methanotrophy and nitrification: processes and connections. *Encyclopedia of Life Sciences*, <https://doi.org/10.1002/9780470015902.a0022213>.

Stomeo, F., Portillo, M., Gonzalez, J., Laiz, L. and Sáiz-Jiménez, C., 2008. *Pseudonocardia* in white colonizations in two caves with Paleolithic paintings. *International Biodeterioration & Biodegradation*, 62(4): 483-486.

Sutherland, C. and Murakami, K.S., 2018. An introduction to the structure and function of the catalytic core enzyme of *Escherichia coli* RNA polymerase. *EcoSal Plus*, 8(1).

Tolar, B.B., Boye, K., Bobb, C., Maher, K., Bargar, J.R. and Francis, C.A., 2020. Stability of floodplain subsurface microbial communities through seasonal hydrological and geochemical cycles. *Frontiers in Earth Science*, 8: 338.

Tomczyk-Żak, K. and Zielenkiewicz, U., 2016. Microbial diversity in caves. *Geomicrobiology Journal*, 33(1): 20-38.

Toyofuku, M. and Yoon, S.-S., 2018. Nitric oxide, an old molecule with noble functions in *Pseudomonas aeruginosa* biology. *Advances in Microbial Physiology*, 72: 117-145.

Tsoy, O.V., Ravcheev, D.A., Čuklina, J. and Gelfand, M.S., 2016. Nitrogen fixation and molecular oxygen: comparative genomic reconstruction of transcription regulation in Alphaproteobacteria. *Frontiers in Microbiology*, 7: 1343.

Tsutsumi, K., Yonehara, R., Ishizaka-Ikeda, E., Miyazaki, N., Maeda, S., Iwasaki, K., Nakagawa, A. and Yamashita, E., 2019. Structures of the wild-type MexAB–OprM tripartite pump reveal its complex formation and drug efflux mechanism. *Nature Communications*, 10(1): 1520.

Turrini, P., Tescari, M., Visaggio, D., Pirolo, M., Lugli, G.A., Ventura, M., Frangipani, E. and Visca, P., 2020. The microbial community of a biofilm lining the wall of a pristine cave in Western New Guinea. *Microbiological Research*, 241: 126584.

Tveit, A.T., Hestnes, A.G., Robinson, S.L., Schintlmeister, A., Dedysh, S.N., Jehmlich, N., von Bergen, M., Herbold, C., Wagner, M. and Richter, A., 2019. Widespread soil bacterium that oxidizes atmospheric methane. *Proceedings of the National Academy of Sciences*, 116(17): 8515-8524.

Unterholzner, S.J., Poppenberger, B., Rozhon, W., 2013. Toxin-antitoxin systems: Biology, identification, and application. *Mob Genet Elements*, 3(5):e26219

Vasquez-Cardenas, D., Meysman, F.J. and Boschker, H.T., 2020. A cross-system comparison of dark carbon fixation in coastal sediments. *Global Biogeochemical Cycles*, 34(2): e2019GB006298.

Vigneron, A., Cruaud, P., Lovejoy, C. and Vincent, W.F., 2022. Genomic evidence of functional diversity in DPANN archaea, from oxic species to anoxic vampiristic consortia. *ISME Communications*, 2(1): 4.

- Vineis, J.H., Bulseco, A.N. and Bowen, J.L., 2021. Microbial dark carbon fixation fueled by nitrate enrichment. *bioRxiv*: 2021.08. 24.457596.
- Wade, J.T. and Struhl, K., 2004. Association of RNA polymerase with transcribed regions in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 101(51): 17777-17782.
- Wagner, T., Watanabe, T. and Shima, S., 2018. Hydrogenotrophic methanogenesis, *Handbook of Hydrocarbon and Lipid Microbiology Series. Biogenesis of Hydrocarbons*. Springer. 1-29
- Wang, C., Sun, B., Zhang, X., Huang, X., Zhang, M., Guo, H., Chen, X., Huang, F., Chen, T. and Mi, H., 2019. Structural mechanism of the active bicarbonate transporter from cyanobacteria. *Nature Plants*, 5(11): 1184-1193.
- Whitfield, G.B., Marmont, L.S., Ostaszewski, A., Rich, J.D., Whitney, J.C., Parsek, M.R., Harrison, J.J. and Howell, P.L., 2020. Pel polysaccharide biosynthesis requires an inner membrane complex comprised of PelD, PelE, PelF, and PelG. *Journal of Bacteriology*, 202(8): e00684-19.
- Wiseschart, A., Mhuantong, W., Tangphatsornruang, S., Chantasingh, D. and Pootanakit, K., 2019. Shotgun metagenomic sequencing from Manao-Pee cave, Thailand, reveals insight into the microbial community structure and its metabolic potential. *BMC Microbiology*, 19(1).
- Wood, H.G., Ragsdale, S.W. and Pezacka, E., 1986. The acetyl-CoA pathway of autotrophic growth. *FEMS Microbiology Reviews*, 2(4): 345-362.
- Wood, T.L. and Wood, T.K., 2016. The HigB/HigA toxin/antitoxin system of *Pseudomonas aeruginosa* influences the virulence factors pyochelin, pyocyanin, and biofilm formation. *Microbiologyopen*, 5(3): 499-511.

Wu, M., Ren, Q., Durkin, A.S., Daugherty, S.C., Brinkac, L.M., Dodson, R.J., Madupu, R., Sullivan, S.A., Kolonay, J.F. and Nelson, W.C., 2005. Life in hot carbon monoxide: the complete genome sequence of *Carboxydotherrnus hydrogenofornans* Z-2901. PLoS Genetics, 1(5): e65.

Wu, Y.-W., Simmons, B.A. and Singer, S.W., 2015. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics, 32(4): 605-607.

Wüst, P.K., Foesel, B.U., Geppert, A., Huber, K.J., Luckner, M., Wanner, G. and Overmann, J., 2016. *Brevitalea aridisoli*, *B. deliciosa* and *Arenimicrobium luteum*, three novel species of Acidobacteria subdivision 4 (class Blastocatellia) isolated from savanna soil and description of the novel family Pyrinomonadaceae. International Journal of Systematic and Evolutionary Microbiology, 66(9): 3355-3366.

Xie, J., Wu, Z., Zhang, X., Peng, T., Yang, C., Zhang, J. and Liang, J., 2021. Diversity and structural characteristics of soil microbial communities in different habitats of wild *Lilium regale* Wilson in Wenchuan area. Bioengineered, 12(2): 10457-10469.

Xu, Z., Masuda, Y., Wang, X., Ushijima, N., Shiratori, Y., Senoo, K. and Itoh, H., 2021. Genome-Based taxonomic rearrangement of the order Geobacterales including the description of *Geomonas azotofigens* sp. nov. and *Geomonas diazotrophica* sp. nov. Frontiers in Microbiology, 12: 737531.

Xue, L., Yue, J., Ke, J., Khan, M.H., Wen, W., Sun, B., Zhu, Z. and Niu, L., 2020. Distinct oligomeric structures of the YoeB–YefM complex provide insights into the conditional cooperativity of type II toxin–antitoxin system. Nucleic Acids Research, 48(18): 10527-10541.

Yang, Q.E. and Walsh, T.R., 2017. Toxin–antitoxin systems and their role in disseminating and maintaining antimicrobial resistance. FEMS Microbiology Reviews, 41(3): 343-353.

- Zada, S., Xie, J., Yang, M., Yang, X., Sajjad, W., Rafiq, M., Hasan, F., Hu, Z. and Wang, H., 2021. Composition and functional profiles of microbial communities in two geochemically and mineralogically different caves. *Applied Microbiology and Biotechnology*, 105: 8921-8936.
- Zambare, N.M., Naser, N.Y., Gerlach, R. and Chang, C.B., 2020. Mineralogy of microbially induced calcium carbonate precipitates formed using single cell drop-based microfluidics. *Scientific Reports*, 10(1): 1-11.
- Zehr, J.P., Jenkins, B.D., Short, S.M. and Steward, G.F., 2003. Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environmental Microbiology*, 5(7): 539-554.
- Zhang, N., Zhang, S., Ren, W., Gong, X., Long, H., Zhang, X., Cai, X., Huang, A. and Xie, Z., 2021. Roles of rpoN in biofilm formation of *Vibrio alginolyticus* HN08155 at different cell densities. *Microbiological Research*, 247: 126728.
- Zhang, B., Wu, X., Tai, X., Sun, L., Wu, M., Zhang, W., Chen, X., Zhang, G., Chen, T. and Liu, G., 2019. Variation in actinobacterial community composition and potential function in different soil ecosystems belonging to the arid Heihe River Basin of Northwest China. *Frontiers in Microbiology*, 10: 2209.
- Zhao, K., Liu, M. and Burgess, R.R., 2005. The global transcriptional response of *Escherichia coli* to induced σ_{32} protein involves σ_{32} regulon activation followed by inactivation and degradation of σ_{32} in vivo. *Journal of Biological Chemistry*, 280(18): 17758-17768.
- Zhao, R., Wang, H., Cheng, X., Yun, Y. and Qiu, X., 2018. Upland soil cluster γ dominates the methanotroph communities in the karst Heshang Cave. *FEMS Microbiology Ecology*, 94(12): 192.

Zhu, H.-Z., Zhang, Z.-F., Zhou, N., Jiang, C.-Y., Wang, B.-J., Cai, L. and Liu, S.-J., 2019. Diversity, distribution and co-occurrence patterns of bacterial communities in a karst cave system. *Frontiers in Microbiology*, 10: 1726.

Zhu, H.-Z., Zhang, Z.-F., Zhou, N., Jiang, C.-Y., Wang, B.-J., Cai, L., Wang, H.-M. and Liu, S.-J., 2021. Bacteria and metabolic potential in karst caves revealed by intensive bacterial cultivation and genome assembly. *Applied and Environmental Microbiology*, 87(6): e02440-20.

VITA

John Anthony Fulginiti was born in Wilkes-Barre, Pennsylvania, a region of the state known for its coal mining history. John's curiosity in Earth Science came at a young age and was influenced by the history of Pennsylvania's natural resources, as well as the beautiful state parks around northeast Pennsylvania. John studied Geobiology at Penn State University from 2016 – 2020, and held numerous undergraduate lab technician positions throughout his studies relating to geology, oceanography, botany, and entomology. John graduated from Penn State during the COVID-19 pandemic, and immediately began to pursue opportunities for graduate studies despite open positions being harder to come by during the country's shutdown. In the fall of 2020, he applied to the University of Tennessee, Knoxville, and joined an interdisciplinary geobiology lab to pursue a M.S. in Geology.