Winter 12-13-2017

Comparison of Bacterial and Archaeal Communities in the Subsurface versus Surface: Implications for Nitrogen Cycling

Jason C. Kimble 5059801155
University of New Mexico

Jason C. Kimble
Candidate

Biology
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. Robert L. Sinsabaugh , Chairperson

Dr. Diana E. Northup

Dr. Clifford N. Dahm

Dr. Rebecca J. Bixby

Dr. Lisa Y. Stein

Comparison of Bacterial and Archaeal Communities in the Subsurface versus Surface: Implications for Nitrogen Cycling

by

Jason Cody Kimble

B.S. Biology, University of New Mexico, 2011

DISSERTATION

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy Biology

The University of New Mexico Albuquerque, New Mexico

Spring 2018

DEDICATION

If my late grandparents Nick and Virginia Smallridge did not have the foresight to save for my education, I may never have finished my undergraduate. Though neither of them attended college, they knew the value of an education. It is an important lesson that I take with me and will pass on to my children Kurt Thomas Kimble, Kira Lynn Kimble, and William Atticus Riker Kimble. This dissertation is dedicated in my grandparent's memory.

ACKNOWLEDGEMENTS

This journey would not have been possible without numerous emails, meetings, messaging, phone calls, and never-ending questions asked to so many. I would like to sincerely thank my committee members Dr. Diana E. Northup, Dr. Robert L. Sinsabaugh, Dr. Clifford N. Dahm, Dr. Rebecca J. Bixby, and Dr. Lisa Y. Stein. I would never have gotten as far as I did without Ara S. Winter teaching me bioinformatic methods and support from other lab members. A special thanks to the UNM IMDS program for many years of support and mentorship. I would like to thank the Fort Stanton Cave Study Project (http://fscsp.org/) and the Bureau of Land Management (BLM) for sequencing funding, sampling, and caving support.

Finally, I have much gratitude for my wife Julie's support all these years and listening to me discuss the joys and woes of graduate school.

Comparison of Bacterial and Archaeal Communities in the Subsurface versus Surface: Implications for Nitrogen Cycling

by

Jason Cody Kimble

B.S. Biology, University of New Mexico 2011

Ph.D., Biology, University of New Mexico, 2017

ABSTRACT

Arid-land caves are thought to be extremely nitrogen-limited, but almost nothing is known about how microbes in subsurface arid-land environments obtain this essential element to meet cellular demand. The depth of caves beneath the surface may represent a critical factor affecting microbial nitrogen cycling in these environments. Percolation of water and nutrients from a precipitation pulse event would affect deep arid-land carbonate caves at a much slower rate. To obtain nitrogen in deep, carbonate caves, microorganisms could use fixed N in the host rock for assimilatory biochemical pathways or for a respiratory electron acceptor. However, the latter process leads to losses of bioavailable N through production of N_2O and N_2, which can only be replaced by N_2 fixation or weathering. Fort Stanton Cave (FSC), found near the northern end of the Sacramento Mountains, is the third longest cave in New Mexico. Multicolored secondary mineral deposits of soil-like material, known as ferromanganese deposits (FMD) exist on the ceilings and walls of FSC. I hypothesized that within the FMD I would find the presence of microbial nitrogen cycling genes. Overburden and connectivity with the surface would influence archaeal and bacterial groups found in caves. As FSC is a moderately deep carbonate cave, I hypothesized that the archaeal and bacterial communities residing in the subsurface would differ from their surface counterparts, as extreme oligotrophic conditions in the cave would select for organisms with metabolisms favoring chemolithotrophy and low-nutrient adaptability. To investigate these hypotheses, Illumina shotgun metagenomics and 16S rRNA gene sequencing were used. Sequences were processed and annotated using several bioinformatic methods. Results indicate that there were genes present in the FMD related to nitrification, dissimilatory nitrate reduction to ammonium, denitrification, and assimilatory nitrate reduction pathways. Potential key players include the ammonia oxidizing archaea phylum *Thaumarchaeota* and the ammonia and/or nitrite oxidizing bacterial phylum *Nitrospirae*. Core microbiome and taxonomic results show that the archaeal and bacterial communities in surface soils are dissimilar to their cave counterparts. There were also bacterial phyla identified in the cave that were mostly absent in surface soils, suggesting

low-nutrient adaptation. Comparing the archaeal FSC dataset to Cueva Villa Luz, Tabasco, Mexico and several caves in Parashant National Monument (PARA), AZ, there were no OTUs shared across *all* samples. Our results show that cave types, host-rock geochemistry, and depth influence archaeal communities present in these subsurface environments. These results shed light on: a) how microbes in caves acquire and cycle nitrogen, b) the archaeal and bacterial diversity in these environments, and c) drivers that influence their presence of diverse archaea in these subsurface biomes.

Table of Contents

Chapter 1: Introduction

Significance and Intellectual Merit

Culture-independent studies of cave biomes, using next generation metagenomic and 16S rRNA gene sequencing, allow for taxonomic identification of bacterial and archaeal groups and functional predictions about their metabolisms. By exploring how microbial taxonomy is influenced by factors such as connectivity (depth below the surface), cave and sample types, and host rock, this research sheds light on drivers of bacterial and archaeal communities in these subsurface biomes.

Background

Caves are often considered extreme environments because of their oligotrophic nature and low productivity (Moore and Sullivan, 1997; Northup and Welbourn, 1997; Barton, 2006). However, diverse microbial communities have been described in many cave environments (Holmes *et al*., 2001; Portillo *et al*., 2008; Northup *et al*., 2008; Chen *et al*., 2009; Pasić *et al*., 2010; Tetu *et al*., 2013; Hathaway *et* al., 2014b; Tomczyk-Żak and Zielenkiewicz, 2016; Zhao *et al*., 2017). These organisms often use reduced compounds in the host rocks, such as ammonium, nitrite, sulfur, iron, and manganese (Northup and Lavoie, 2001; Barton and Northup, 2007), as sources of energy and electrons, which can be used to fix CO_2, a process known as chemolithoautotrophy. Microorganisms need sources of N and P, as well as C for growth. These sources may vary widely among caves, but the oligotrophic nature of most caves suggests that most short-term demand for nutrients is met through cycling within the microbial community. Our study focuses on the cycling of one essential nutrient: nitrogen.

Insight into Nitrogen Cycling and Key Players

One of my aims is to explain how N is cycled in a moderately deep arid-land cave. The arid-land carbonate cave that will be the primary focus of this investigation is Fort Stanton Cave (FSC), New Mexico, USA. It would seem to be an enigma that environments so low in available N would be conducive to life, and yet, evidence suggests microbial communities are present in moderate numbers in these subterranean systems (Northup *et* al., 2003; Spilde *et al*., 2005; Barton *et al*., 2007; Ortiz *et al*., 2013; Ortiz *et al*., 2014). Recent studies have provided functional and taxonomic evidence of N cycling in caves (Tetu *et* al., 2013; Ortiz *et al*., 2013; Hathaway *et* al., 2014a; Ortiz *et al*., 2014; Reitschuler *et al.*, 2015; Reitschuler *et al.*, 2016; Zhao *et* al., 2017). The project aim of chapter 2 was to study the microbial communities in these cave systems with the purpose of identifying taxonomic diversity and the presence of functional genes related to the cycling of N by Bacteria and Archaea in low-nutrient soil-like material known as ferromanganese deposits (FMD). Previous research has shown that bacterial and archaeal groups are found in FMD (Spilde *et al*., 2005; Spilde *et* al., 2006) and phylogenetic evidence suggests relatedness to ammonia oxidizers and nitrite oxidizers (Northup *et* al., 2003; Dichosa, 2008). This research contributes to our understanding of how microorganisms in low-nutrient environments acquire and cycle N, providing insight into N assimilatory and energetic pathways in caves.

Bacterial and Archaeal Surface and Cave Comparison

Research in Kartchner Caverns, AZ (Ortiz *et al*., 2013) and Lava Beds National Monument, CA (Lavoie *et al*., 2017) that compared bacterial communities present in the

caves and overlying soils, found few shared operational taxonomic units (OTUs) between the surface and caves. In shallow lava caves, such as investigated in the Lavoie *et al*. (2017) study, I hypothesize that water infiltrating from the surface after precipitation events represents the major input of organic and inorganic nitrogen into these environments. Surface water likely results in the exchange of nutrients that influence bacterial and archaeal community structure in shallow lava caves. Percolation of water and nutrients from a precipitation pulse event would affect deep arid-land carbonate caves much more slowly. Therefore, I expect to find little overlap between archaeal and bacterial surface and cave OTUs, and these environments will have dissimilar taxonomic representations. Research findings from chapter 3 were compared to previous studies (Ortiz *et al*., 2013; Lavoie *et al*., 2017) to better understand mechanisms (e.g. depth and host-rock) responsible for microbial communities in caves.

Archaeal Site Comparison

Previous investigations have identified archaeal groups in caves (Sarbu *et al*., 1996; Chen *et al*., 2009; Northup *et al*., 2003; Lehtovirta *et al*., 2009; Barton *et al*., 2014; Ortiz *et al*., 2014; Reitschuler *et al.*, 2015; Reitschuler *et al.*, 2016; Tetu *et* al., 2013; Zhao *et* al., 2017); however, we still know very little about archaeal diversity, underlying factors that influence community compositions, and similarities and/or differences of these organisms across different caves. There have been no previous cave investigations that have compared archaeal microbiomes across geographic localities and provided a comprehensive representation of archaeal groups in these biomes. Research findings from chapter 4 shed light on archaeal taxonomic differences and/or similarities among diverse cave and sample types and help determine if connectivity (e.g. depth below the surface)

and/or cave type or host-rock geochemistry are factors that determine archaeal diversity in caves.

References

Barton, H. A. (2006). Introduction to cave microbiology: a review for the non-specialist. *J. Caves Karst Stud.* 68, 43-54.

Barton, H. A., and Northup, D. E. (2007). Geomicrobiology in cave environments: Past, current and future perspectives. *J. Cave. Karst. Stud.* 69, 163-178.

Barton, H. A., Giarrizzo, J. G., Suarez, P., Robertson, C. E., Broering, M. J., and Banks, E. D. (2014). Microbial diversity in a Venezuelan orthoquartzite cave is dominated by the *Chloroflexi* (Class *Ktedonobacterales*) and *Thaumarchaeota* Group I.1c. *Front. Microbiol.* **5**:615. doi: 10.3389/fmicb.2014.00615

Barton, H. A., Taylor, N. M., Kreate, M. P., Springer, A. C., Oehrle, S. A., and Bertog, J. L. (2007). The impact of host rock geochemistry on bacterial community structure in oligotrophic cave environments. *Int. J. Speleol.* 36, 93–104. doi: 10.5038/1827-806X.36.2.5

Chen, Y., Wu, L. Q., Boden, R., Hillebrand, A., Kumaresan, D., Moussard, H. et al. (2009). Life without light: microbial diversity and evidence of sulfur- and ammonium based chemolithotrophy in Movile Cave. *ISME J.* 3, 1093-1104. doi: 10.1038/ismej.2009.57

Dichosa, A. (2008). Biogenicity and microbial community composition of desert varnish and cave ferromanganese deposits. Ph.D. dissertation: University of New Mexico.

Hathaway, J. J. M., Garcia, M. G., Balasch, M. M., Spilde, M. N., Stone, F. D., Lurdes, M. D., et al. (2014b). Comparison of bacterial diversity in Azorean and Hawai'ian lava cave microbial mats. *Geomicrobiol. J.* 31, 205-220. doi: 10.1080/01490451.2013.777491

Hathaway, J. J. M., Sinsabaugh, R. L., Dapkevicius, M. L. N. E., and Northup, D. E. (2014a). Diversity of ammonia oxidation (amoA) and nitrogen fixation (nifH) genes in lava caves of Terceira, Azores, Portugal. *Geomicrobiol. J.* 31, 221–235. https://doi.org/10.1080/01490451.2012.752424

Holmes, A. J., Tujula, N. A., Holley, M., Contos, A., James, J. M., Rogers, P., and Gillings, M. R. (2001). Phylogenetic structure of unusual aquatic microbial formations in Nullarbor caves, Australia. *Environ. Microbiol.* 74, 5934-5942. doi: 10.1046/j.1462-2920.2001.00187.x

Lavoie, K. H., Winter, A. S., Read, K. J., 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**: e0169339. doi: 10.1371/journal.pone.0169339

Lehtovirta, L. E., Prosser, J. I., and Nicol, G. W. (2009). Soil pH regulates the abundance and diversity of Group 1.1c Crenarchaeota. *FEMS Microbiol. Ecol.* 70, 367–376. https://doi.org/10.1111/j.1574-6941.2009.00748.x

Moore, G.W. and Sullivan, N. (1997). Speleology: caves and the cave environment. St. Louis, MO, USA: Cave Books.

Northup, D. E., and Lavoie, K. H. (2001). Geomicrobiology of caves: a review. *Geomicrobiol. J.* 18, 199-222. doi: 10.1080/01490450152467750

Northup, D. E. and Welbourn, W. C. (1997). Life in the twilight zone: lava tube ecology. *New Mex. Bur. Mine Miner Resour. Bull*. 156, 69-82.

Northup, D. E., Barns, S. M., Yu, L. E., Spilde, M. N., Schelble, R. T., Dano, K. E. et al. (2003). Diverse microbial communities inhabiting ferromanganese deposits in Lechuguilla and Spider Caves. *Environ. Microbiol.* 5, 1071-1086.

Northup, D. E., Connolly, C. A., Trent, A., Peck, V. M., Spilde, M. N., Welbourn, W. C., and Natvig, D. O. (2008). The nature of bacterial communities in Four Windows Cave, El Malpais National Monument, New Mexico, USA. *AMCS Bull*. 19, 119-125.

Ortiz, M., Legatzki, A., Nelson, J. W., Fryslie, B., Nelson, W. M., Wing, R. A., et al. (2014). Making a living while starving in the dark: metagenomic insights into the energy dynamics of a carbonate cave. *ISME. J.* 8, 478-491. doi:10.1038/ismej.2013.159

Ortiz, M., Neilson, J. W., Nelson, W. M., Legatzki, A., Byrne, A., Yu, Y., et al. (2013). Profiling bacterial diversity and taxonomic composition on speleothem surfaces in Kartchner Caverns, AZ. *Microb. Ecol.* 65, 371-383.

Pasić, L., Kovce, B., Sket, B., Herzog-Velikonja, B. (2010). Diversity of microbial communities colonizing the walls of a karstic cave in Slovenia. *FEMS Microbiol. Ecol.* 71, 50-60. doi: 10.1111/j.1574-6941.2009.00789.x

Portillo, M. C., Gonzalez, J. M., and Saiz-Jimenez, C. (2008). Metabolically active microbial communities of yellow and grey colonizations on the walls of Altamira Cave, Spain. *J. Appl. Microbiol.* 104, 681-691. doi: 10.1111/j.1365-2672.2007.03594.x

Reitschuler, C., Lins, P., Schwarzenauer, T., Spötl, C., Wagner. A. O., and Illmer, P. (2015). New undescribed lineages of non-extremophilic archaea form a homogeneous and dominant element within alpine moonmilk microbiomes. *Geomicrobiol. J.* 32, 890-902. doi: 10.1080/01490451.2015.1025317

Reitschuler, C., Spötl, C., Hofmann, K., Wagner, A. O., and Illmer, P. (2016). Archaeal distribution in moonmilk deposits from Alpine caves and their ecophysiological potential. *Microb. Ecol.* 71, 686-699. doi: 10.1007/s00248-015-0727-z

Sarbu, S. M., Kane, T. C., and Kinkle, B. K. (1996). A chemoautotrophically based cave ecosystem. *Science*. 272, 1953–1955. doi: 10.1126/science.272.5270.1953

Spilde, M. N., Northup, D. E., and Boston, P. J. (2006). "Ferromanganese deposits in the caves of the Guadalupe Mountains" in Caves and Karst of Southeast New Mexico, ed. L. Land L, Lueth (The New Mexico Geological Society), 161–166.

Spilde, M. N., Northup, D. E., Boston, P. J., Schelble, R. T., Kathleen, E. D., Crossey, L. J., et al. (2005). Geomicrobiology of cave ferromanganese deposits: a field and laboratory investigation. *Geomicro. J.* 22, 99-116. doi: 10.1080/01490450590945889

Tetu, S. G., Breakwell, K., Elbourne, L. D. H., Holmes, A. J., Gillings, M. R., and Paulsen, I. T. (2013). Life in the dark: metagenomic evidence that a microbial slime community is driven by inorganic nitrogen metabolism. *ISME. J.* 7, 1227–1236. doi: 10.1038/ismej.2013

Tomczyk-Żak, K. and Zielenkiewicz, U. (2016). Microbial diversity in caves. *Geomicrobiol. J.* 33, 20–38. doi: 10.1080/01490451.2014.1003341

Zhao, R., Wang, H., Yang, H., Yun, Y., and Barton, H. A. (2017). Ammonia oxidizing archaea dominant over bacterial ammonia oxidizers within alkaline cave sediments. *Geomicro. J.* 34, 511-523. doi: 10.1080/01490451.2016.1225861

Chapter 2: Nitrogen Cycling by Cave Microbes Deep in Snowy River Passage, Fort Stanton Cave, New Mexico, USA

Jason C. Kimble[1], Ara S. Winter[1], Michael N. Spilde, Robert L. Sinsabaugh[1], and Diana E. Northup[1]

[1]Department of Biology, MSC03-2020, University of New Mexico, Albuquerque, NM 87131

[2]Institute of Meteoritics, MSC03-2050, University of New Mexico, Albuquerque, NM 87131

*Corresponding Authors
Jason C. Kimble – jkimble@unm.edu

Abstract

The low biomass and productivity of arid-land caves with limited availability of nitrogen (N) presents the question of how microbes in these extreme subsurface environments acquire and cycle this essential element. Caves are ideal environments for investigating microbial functional capabilities, as they lack phototrophic activity and have near constant temperatures and high relative humidity in their Deep Zones, and caves experience little weathering, unlike the surface. From the walls of Fort Stanton Cave (FSC), multicolored secondary mineral deposits of soil-like material low in fixed N, known as ferromanganese deposits (FMD), were collected. We hypothesized that within our FMD samples we would find the presence of microbial N cycling genes and taxonomy related to N cycling microorganisms. To investigate our hypothesis, our samples were sequenced using Illumina shotgun metagenomics and 16S rRNA sequencing. Results suggest a diverse N cycle encompassing several energetic pathways including nitrification, dissimilatory nitrate reduction, and denitrification. Also, N cycling genes associated with assimilatory nitrate reduction were identified. Functional and taxonomic findings suggest several bacterial and archaeal phyla potentially play a role in nitrification pathways in FSC FMD. *Thaumarchaeota*, a deep-branching archaeal division, likely play an essential and possibly dominant role in the oxidation of ammonia. Our results provide genomic evidence for understanding how cave microbes are potentially able to acquire and cycle N in an aphotic low-nutrient subterranean environment.

Introduction

Sources of nitrogen (N) (e. g., ammonium [NH_4^+], nitrite [NO_2^-], and nitrate [NO_3^-]), phosphorus, and carbon are required by microbial communities to sustain their growth and metabolism. In subterranean biomes, these essential nutrients are difficult to acquire because they occur naturally at extremely low levels (Boston *et al*., 2001; Northup and Lavoie 2001; *Bar*ton *et al*., 2004; Spilde *et* al., 2005; Lee *et al*., 2012). The availability of these essential nutrients varies widely among caves, but the oligotrophic nature of many arid-land caves with limited surface input (Barton *et* al., 2007; Ortiz *et al*., 2013; Ortiz *et al*., 2014) suggests that most short-term demand for resources is met through cycling within the microbial community. Arid-land caves are likely to be extremely N-limited (Northup *et al*., 2003; Levy, 2007; Ortiz *et al*., 2014), but our knowledge of how microbes in subsurface environments obtain and cycle this essential element to meet cellular demand is limited.

What is known about N cycling in the subsurface comes from several investigations in subterranean environments that identified functional genes and/or taxonomy associated with the cycling of N. Spear *et al*.'s (2007) investigation of a mine adit revealed the presence of archaeal *amoA* genes, mesophilic crenarchaeal 16S rRNA sequences that shared sequence similarity with several "marine group I.1a" *Crenarchaeota*, and clones previously identified in deep South African gold mines (Takai *et al*., 2001). Mesophilic *Crenarchaeota* groups I.1a, I.1b, and I.1c are now classified as the phylum *Thaumarchaeota* (Brochier-Armanet *et al*., 2008), which includes all recognized archaeal ammonia oxidizers, as well as groups of homologous sequences with still undetermined metabolism (Pester *et al*., 2011). *Crenarchaeota* group I.1b, as well as

the presence of crenarchaeal *amoA* genes, was found in subsurface radioactive thermal spring sequencing clones (Weidler *et al*., 2007; Weidler *et al*., 2008) and evidence of bacterial ammonia and nitrite oxidation was provided from a study that looked at 16S rRNA and *amoA* gene libraries (Chen *et al*., 2009) from a surface isolated cave that is driven by sulfur and methane oxidation (Sarbu *et al*., 1996). An investigation of Azorean shallow lava caves found *amoA* genes were present and dominated by *Nitrosospira*-like sequences, though *Nitrosomonas* spp. were also found (Hathaway *et al*., 2014a). Recent metagenomic studies have also provided insight into N cycling in the subsurface. Jones *et al*. (2012) found N assimilation genes in extremely low pH biofilms that have fixed N present, and Tetu *et al*. (2013) and Ortiz *et al*. (2014) found several functional genes, as well as bacterial and archaeal taxonomy, associated with N cycling. These studies suggest the potential for an active N cycle in subsurface environments even though there has not been a comprehensive consideration of key players, pathways and genes.

Study Background

Multicolored secondary mineral deposits of soil-like material low in fixed N (Northup *et al*., 2003), known as ferromanganese deposits (FMD), exist on the ceilings and walls of many arid-land carbonate caves in New Mexico (Northup *et al*., 2000; Northup *et al*., 2003; Spilde *et al*., 2005; Spilde *et al*., 2006). This material, which is rich in Fe- and Mn-oxides, contains microbial communities driven by chemolithoautotrophic metabolic processes (Figure 1) (Northup *et al*., 2003; Spilde *et al*., 2005; Spilde *et al*., 2006). Iron and Mn-oxide deposits, thought to be of a biogenic origin as a result of microbial redox reactions, have been described in a variety of environments, including carbonate caves (Spilde *et al*., 2005; Rossi *et al*., 2010; Frierdich *et al*., 2011; Gázquez *et*

al., 2011; Carmichael *et al*., 2013), desert varnish deposits found within a semiarid region (Northup *et al*., 2010), soil sediments (He *et al*., 2008), and the caldera of the submarine Axial volcano (Kennedy *et al*., 2003). In Spider and Lechuguilla caves, New Mexico, USA, it is hypothesized that the FMD associated microbial communities are "mining" the bedrock for reduced compounds that can be used in redox pathways by chemolithoautotrophs for energetic purposes (Spilde, *et al*., 2005). The oxidation of the reduced compounds results in an oxide layer, which in some of the darker colors (e.g. brown and black crusts), was found to be enriched with elemental Fe and Mn at concentrations hundreds to thousands of times greater than the core bedrock (Spilde *et al*., 2006). It was determined that there was a relationship between the color and chemical makeup of the FMD; for example, manganese and iron oxides were found in higher concentrations in darker color deposits than in lighter colored deposits (e.g. pink FMD) (Spilde, *et al*., 2005).

Our objective was to investigate an arid-land carbonate cave for evidence of microbial N cycling using shotgun metagenomic and 16S rRNA sequencing. We identified potential key microbial players, functional genes and corresponding taxonomy, and pathways associated with N cycling, as well as potential sources of available N. Arid-land caves with limited input from the surface in particular are analogs for inferring how life could survive in a strongly nutrient limiting environment, perhaps even on other solar system bodies such as Mars. NASA's Mars Reconnaissance Orbiter has captured images of lava caves beneath the Martian surface, and it has been hypothesized that a "transient but active nitrogen cycle" exists on Mars, which has a 2.7% atmospheric composition of dinitrogen gas (Boxe *et al*., 2012). Additionally, it has been theorized that if we were to

find evidence of microbial life on Mars, it would be found in a protective, thermostable subsurface environment (Williams *et al*., 2010).

Methods

Sampling Site

Fort Stanton Cave (FSC), found near the northern end of the Sacramento Mountains, is the third longest cave in New Mexico (http://www.caverbob.com/wlong.htm). FSC formed in the lower Rio Bonito valley in Permian San Andres limestone (Kelley, 1971). Dissimilar to limestone caves found southeast of FSC within the Guadalupe Mountains, which were formed by hypogenic sulfuric acid speleogenesis (Jagnow, 1979; Hill, 1990), FSC cave is thought to be an epigenic cave that formed as the result of sinking surface waters (Davis and Land, 2006) through carbonic acid dissolution. Within FSC lies a passage known as Snowy River, named so because of the occurrence of a stream channel lined with a white-calcite continuous deposit now mapped at over 17.8 km in length, which is believed to be the longest continuous speleothem in the world (Land, 2012). Radiometric dating of the basal layer of cores taken from the Snowy River formation estimates that the passage is relatively young at only 820 ± 120 years old (Land *et al*., 2010). The cave is managed by the US Bureau of Land Management, and entrance is granted by permit only (US Bureau of Land Management, 2016). Sample collection was authorized under a US Bureau of Land Management collecting permit (Northup, P.I.). Samples were collected aseptically from the walls of Snowy River from multicolored FMD that precipitate out from the underlying mud layer (Figure 2; Figure 3). Samples were covered with sucrose lysis

buffer to break open cells and to prevent DNA decomposition (Giovannoni *et al*., 1990) and were taken to the lab, where they were stored in a freezer at -80°C and later extracted for DNA.

Molecular Processing of Samples

Genomic DNA was extracted and purified using the MoBio PowerSoil™ DNA Isolation Kit according to the manufacturer's protocol (MoBio, Carlsbad, CA). However, cells were disrupted by use of a bead beater as an alternative to vortexing and were eluted with 45 µl of elution buffer, rather than the recommended 100 µl, due to low levels of DNA found in the environmental samples. Sample extractions consisted of three replicates and a negative control. Extracted DNA was stored at -20° C at the Northup Lab and sent to Molecular Research (MR DNA) (http://www.mrdnalab.com/), Shallowater, TX in 2013 for 454 sequencing as part of a survey study to identify bacterial diversity and novelty associated with specific FMD colors found on the walls and ceiling in the Snowy River passage. From this pool, four FMD samples (Table 1; Table 2) were chosen as likely candidates to identify taxonomic and functional information associated with N cycling organisms in the FSC samples, and these samples were sequenced for metagenomic shotgun reads. Our survey showed that these samples had elevated counts associated with N cycling bacterial and archaeal groups (e.g. *Nitrospirae* and *Thaumarchaeota*) and these samples allowed us to investigate brown and black FMD, which are prevalent throughout the passage.

Sequencing

Metagenomic shotgun reads were sequenced at MR DNA using HiSeq 2x150bp Illumina HiSeq 2500 platform (www.illumina.com). Metagenome sequencing steps included the isolation and purification of genomic DNA, fragmentation, ligation to sequencing adapters and purification. Libraries were pooled and sequenced following amplification and denaturation. A total of 50 ng of DNA from each of the FSC samples was used to prepare DNA libraries using Nextera DNA library preparation kit to build individual barcode catalogs. Library insert size was determined by Experion Automated Electrophoresis Station (Bio-Rad). Pooled library (12pM) was loaded to a 600 Cycles v3 Reagent cartridge (Illumina).

Amplicon archaeal data were generated using Illumina MiSeq to target the 16S rRNA gene using the primers 349-forward (5'-GYGCASCAGKCGMGAAW-3') and 806-reverse (5'-GGACTACVSGGGTATCTAAT-3') (Takai and Horikoshi, 2000), with the barcode on the forward primer. Bacterial survey samples were also assayed using Illumina MiSeq to target the 16S rRNA gene using universal bacterial primers 46 forward (5'-GCYTAAYACATGCAAGTCG-3') and 1409 reverse (5'-GTGACGGGCRGTGTGTRCAA-3') (Northup *et al*., 2010). Amplification consisted of a 28 cycle PCR (5 cycle used on PCR products) using HotStarTaq Plus Master Mix Kit (Qiagen, Valencia, CA) under the following parameters: 94°C for 3 min, followed by 28 cycles of 94°C for 30 s, 53°C for 40 s, 72°C for 1 min, with a final elongation step at 72°C for 5 min. Amplicon PCR products were examined for sufficient amplification and band intensity using 2% agarose gel. Samples were combined in equivalent quantities based on molecular weight and DNA concentrations. Calibrated AMPure XP beads were used to purify combined samples. The Illumina DNA library was then prepared using the

combined and purified PCR product. Sequencing was performed at MR DNA following the manufacturer's guidelines.

Metagenomic Taxonomic Analyses of Snowy River FMD

A total of 36.23 GB of data were returned for the FSC FMD metagenome samples. Returned Illumina shotgun metagenomic data consisted of two FASTQ files per sample, each comprised of one forward and one reverse Illumina reads file. The forward and reverse read files were assembled into single FASTQ files using the ultrafast software tool Paired-End reAd mergeR (PAIR v0.9.6) (Zhang *et al*., 2014) with a p-value setting of 0.5, a maximum a posteriori (MAP) estimation, and a Phred score of 33. Read files were then trimmed using the software tool Sickle (v 1.33) (Joshi and Fass, 2011), which uses sliding windows accompanied by quality and length thresholds to establish which position to trim the 3'-end of reads where the quality is very poor and the position to trim the 5'-end of reads where the quality is very high. The quality threshold used to trim our reads was 20 with a length threshold of 75 bp, which is the minimum sequence length MG-RAST (Manual v3.6, revision 3) can use to determine coding regions in DNA sequences (Wilke *et al*., 2015). FASTQ files were converted to FASTA files and submitted to the Metagenome Rapid Annotation using Subsystem Technology (MG-RAST, v3.6) server (Meyer *et al*., 2008) for annotation under the project name Speleosol_Ferromanganese Deposits. MG-RAST pipeline settings included screening for *Homo sapiens* (NCBI v36) specific sequences and an ambiguous base filtering setting of 5. The data files are held on the MG-RAST server under the accession numbers 4681809.3, 4681806.3, 4681808.3, and 4681807.3.

Using the MG-RAST server (http://metagenomics.anl.gov/) (Meyer *et al*., 2008), Illumina metagenomic reads were analyzed for taxonomic composition using Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000; Kanehisa *et al*., 2016a) as the annotation source. Against the KEGG database, a maximum e value cutoff of $1e^{-5}$, a minimum identity cutoff of 60 %, and a 65 bp minimum alignment length cutoff was used. Data generated were put into tab-delimited files and analyzed for microbial community composition using R (https://www.r-project.org/) (R Core Team, 2016) and phyloseq (https://joey711.github.io/phyloseq/) (McMurdie and Holmes, 2013).

16S Taxonomic Analyses of Snowy River FMD

Returned 16S rRNA bacterial and archaeal reads were run through a Quantitative Insights Into Microbial Ecology (QIIME) (Caporaso *et al*., 2010) (qiime.org) pipeline separately to identify taxonomy at 97% or greater identity level. Reads were analyzed using the *split_libraries.py* script using the mapping, QUAL, and FASTA files. The minimum sequence length was 100 nucleotides and the maximum sequence length was 600 nucleotides, with a minimum average quality score of 30 and 0 maximum number of ambiguous bases. Chimeric sequences were identified in the *split_library/seqs.fna* output fasta files using the *identify_chimeric_seqs.py* script with USEARCH (v6.1.544) (http://www.drive5.com/usearch/) (Edgar, 2010) as the chimeric detection method against the "Gold" reference sequence database (http://drive5.com/uchime/gold.fa). Sequences were filtered from *seqs.fna* output fasta files using the *filter_fasta.py* script with the argument to negate all sequences in the sequence identifier file *chimeras.txt*. De novo operational taxonomic unit (OTU) picking was accomplished using the *pick_de_novo_otus.py* QIIME workflow

(http://qiime.org/scripts/pick_de_novo_otus.html). OTU picking was assigned with the *pick_otus.py* script with arguments indicating the input file paths as *split_library/seqs.fna* and using the SUMACLUST clustering picking method (Mercier *et al*., 2013). Taxonomy was assigned to the output reference sequences using the *assign_taxonomy.py* script with the input files *seqs_rep_set.fasta*, with a path to the QIIME-compatible SILVA_128_QIIME_release 16S only reference set *97_otus_16S.fasta* database and the 16S only ID *consensus_taxonomy_7_levels.txt* tab-delimited file (http://www.arb-silva.de/download/archive/qiime/) (Quast *et al*., 2013). OTU tables in biom format were created using the *make_otu_table.py* script and an argument indicating paths to the input SUMACLUST picked OTUs *seqs_otus.txt* files and the taxonomy assignment *seqs_rep_set_tax_assignments.txt* files. Taxonomy was summarized from the *de_novo_sumaclust.biom* files using the *summarize_taxa.py* script and an argument that allows for the representation of the absolute abundance of the lineage in each of the FSC samples. The *de_novo_sumaclust.biom* files were converted to an R (R Core Team, 2016) readable JSON biom format using the command *biom convert* and the output file *fixed.biom* (http://biom-format.org/documentation/biom_conversion.html) (Supplementary Qiime_Pipeline.txt).

Metagenomic Nitrogen Cycling Functional Analyses of Snowy River FMD

Gene calling was performed on processed FASTA files using the Prokaryotic Dynamic Programming Gene finding Algorithm (PRODIGAL, v2.6.2) (Hyatt *et al*., 2010) software package in meta optimized mode and written to a protein translation file. Protein translation files were annotated for KEGG Orthologs (KOs) (Kanehisa *et al*., 2000; Kanehisa *et al*., 2016a) identifiers using the software package HMMER (v3.0),

which uses Hidden Markov Models (HMMs) to search sequence databases (Eddy, 1998) (http://hmmer.org/). Functional gene databases screened using HMMER included the Functional Ontology Assignments for Metagenomes (FOAM) (Prestat *et al*., 2014) (http://cbb.pnnl.gov/portal/software/FOAM.html), which was used to identify all N cycling KOs, except for the identification of the nitrite oxidoreductase beta subunit (*nxrB* gene), which used the Functional Gene Pipeline and Repository (FGPR) (http://fungene.cme.msu.edu/) (Fish *et al*., 2013) in a separate HMMER search.

FOAM is a large database comprised of 73,969 HMMs, which target 2,870 individual KOs. To save computational time, a subset of FOAM was built using *hmmfetch* to retrieve only N cycling profile HMMs from the FOAM database. Only N cycling HMMs with a *single* assigned KO were retrieved from the FOAM database resulting in the retrieval of 2045 HMMs and 46 KOs (Supplementary FOAM_N_Cycle_HMMs.xlsx). The N cycling subset FOAM profile HMM database was searched against the FMD sample protein translation files individually using *hmmsearch* and the output files were in --domtblout format, which results in the *domain hits table*. The tabular files were sorted for KOs with a series of Python (https://www.python.org/) scripts, commands, and arguments (https://github.com/mmdavid/FOAM and Supplementary Functional_Pipeline.txt), which included a "Best Hits" quality control filtering step that removed all hits with a --minscore less than 25. The sorted files summarize the number of hits for each KO identified in each sample. The data were annotated by inputting the KO numbers into the search inquiry at the KO database (http://www.genome.jp/kegg/ko.html). The *nxrB* gene *hmmsearch* was accomplished by

using the *nxrB* HMM from the FGPR database (Fish *et al*., 2013) and was not assigned a KO or Enzyme Commission (EC) number.

To insure the accuracy of HMMER hits to each assigned function, additional analyses were performed. An index profile was created using *esl-sfetch* to generate an "SSI index" from the FMD sample protein translation files for each of the four FMD samples. The *domain hits table* tabular files were initially ordered using the Unix command *sort*, followed by "Best Hits" quality filtering previously discussed. The filtered HMM tabular output files were used to retrieve subsequences that correspond to hits from the HMMER search by use of their "envelope" coordinates. This was accomplished by using the Unix *grep* command and *esl-sfetch* to retrieve the subsequences from the FMD sample protein translation files (Supplementary Functional_Pipeline.txt). The output protein subsequence files were dereplicated using USEARCH (Edgar, 2010) and submitted to GhostKOALA (Kanehisa *et al*., 2016b) for functional and taxonomic annotation. The number of hits corresponding to each N cycling gene was determined by quantifying KOs associated with assigned taxonomy for each specific function.

Results

Metagenomic and 16S rRNA Gene Taxonomic Composition of Snowy River FMD

Metagenomic read counts annotated by MG-RAST (Meyer *et al*., 2008) are summarized in Table 1. The most abundant domain identified by percentage of read counts for all samples was Bacteria, followed by the Archaea (Figure 4). The average

percent of bacterial reads identified in the FMD metagenome samples was 88.7% and the average percent of archaeal reads identified in the FMD metagenome samples was 8.5% (Figure 4). Eukaryotic reads were also identified in each of the FSC samples, with an average of 0.6% in the FMD samples (Figure 4). The most abundant bacterial and archaeal phyla in the shotgun metagenomic and 16S rRNA reads are summarized in Figure 5. There were taxonomic similarities and differences with respect to reads identified amongst and between datasets. Some distinctions evident in Figure 5 are the absence of the archaeal phyla *Crenarchaeota* and *Korarchaeota* in the 16S rRNA archaeal reads, and the high occurrence of the phyla GAL15 and *Nitrospirae* in the 16S rRNA bacterial reads.

The primers used for archaeal 16S rRNA gene sequencing also amplified some bacterial groups. As such, in the 16S rRNA archaeal dataset, both bacterial and archaeal reads were identified with an average of approximately 7.7% of bacterial reads identified in FMD samples FS58 (4.6%), FS154 (5.4%), and FS155 (13.0%). Taxonomy associated with these bacterial reads are not included in our results. Sample FS56 was not investigated for 16S rRNA archaeal or bacterial reads due to insufficient quantity of available DNA to resequence this sample using Illumina MiSeq. In all samples, *Thaumarchaeota* and *Euryarchaeota* phyla were the dominant archaeal phyla identified (Figure 5); however, there were also reads associated with *Aenigmarchaeota*, *Bathyarchaeota,* Miscellaneous Euryarchaeotic Group (MEG), *Woesearchaeota* (*DHVEG-6*), and WSA2 (Supplementary Figure 1). The class *Thermoplasmata* overwhelmingly dominated the *Euryarchaeota*, representing an average of 99.9% of *Euryarchaeota* reads identified in the three FMD samples.

Metagenomic and 16S rRNA Gene Nitrification Taxonomic Results

In the FSC taxonomic results, organisms associated with ammonia and nitrite oxidation pathways were identified in both the metagenomic and 16S rRNA datasets, though differences were observed. In the metagenomic dataset, the ammonia-oxidizing bacteria (AOB) orders *Nitrosomonadales* (*Nitrosomonas* and *Nitrosospira*) and *Chromatiales* (*Nitrosococcus*) were found in each of the four FMD samples. Taxa associated with nitrite-oxidizing bacteria (NOB) found in the metagenomic reads include the order *Nitrospirales* (*Nitrospira*), which was identified in all FMD samples, and the order *Rhizobiales* (*Nitrobacter*), though this group was only found in the black FMD samples FS154 and FS155.

Ammonia-oxidizing (AO) archaeal reads identified in the shotgun metagenomic dataset were comprised of two genera associated with the marine archaeon Group I.1a, which included *Cenarchaeum* (*Cenarchaeum symbiosum*) (average of 13.7% of archaeal reads and 43.1% of *Thaumarchaeota* reads) and *Nitrosopumilus* (*Nitrosopumilus maritimus*) (average of 18.1% of archaeal reads and 56.9% of *Thaumarchaeota* reads). If the stringency of the minimum alignment length cutoff is increased to a 75 bp overlap, archaeal domain abundance in the FMD reads decreases to an average of 7.8%, with only an increase observed in sample FS56 (14.4% of archaeal reads). However, the archaeal distribution of *Thaumarchaeota* reads in the FMD metagenomes increases to an average of 70.6%, with the most dominant presence observed in sample FS154 (81.1% of archaeal reads).

In the 16S rRNA reads, some taxa similarity was found with respect to microorganisms associated with nitrification that were found in the metagenomic reads. AOB in the family *Nitrosomonadaceae* were found to comprise an average of 2.1% of bacterial reads in samples FS58 (2.0%) and FS154 (2.2%); *Nitrosomonadaceae* in sample FS155 encompassed 7.8% of reads. Reads associated with the family *Nitrosomonadaceae* could not be determined at genus level and all reads were associated with uncultured taxa. All NOB in the 16S rRNA reads were found in the family *Nitrospiracea*e (*Nitrospira*) with an average of 6.0% of bacterial reads identified in the FMD samples. There were other reads corresponding to *Nitrospiracea*e that were ascertained, though some of these reads were associated with the iron-oxidizing genus *Leptospirillum* and ambiguous and uncultured taxa.

In the 16S rRNA archaeal dataset, excluding bacterial reads, the AO phylum *Thaumarchaeota* accounted for an average of 48.9% of all reads in the three FMD samples. *Thaumarchaeota* in the FMD samples was overwhelmingly dominated by the classes South African Gold Mine Gp 1 (SAGMCG-1) (average of 64.2%) and Marine Group I (average of 25.5%), though AK31 (average of 4.2%), AK56 (average of 0.3%), FHMa11 terrestrial group (average of 4.0%), Marine Benthic Group A (average of 0.1%), and Soil Crenarchaeotic Group (SCG) (average of 1.0%) (Supplementary Figure 2) were also found. Nearly all identified 16S rRNA archaeal *Thaumarchaeota* reads were associated with uncultured taxonomy, and genus level ranking could not be discerned, not including '*Candidatus* Nitrososphaera', and a minimal number of annotated reads associated with '*Candidatus* Nitrosoarchaeum' and '*Candidatus* Nitrosopelagicus' (Supplementary Figure 2). An overwhelming majority of 16S rRNA *Thaumarchaeota*

archaeal reads were associated with subterranean and marine environments, with few reads taxonomically identifying with soils or terrestrial biomes.

Nitrogen Cycling Functional Results of Snowy River FMD

Figure 6 gives an overview of N cycle biological functional genes and metabolic pathways that were identified in the FSC FMD metagenomic samples. Information provided in Figure 7 shows that the N cycle gene counts varied between KOs and samples. There were hits associated with the nitrogenase reductase gene (*nifH*) in all the FSC samples, though KEGG GhostKOALA (Kanehisa *et al*., 2016b) functional annotation found that protein subsequences were not associated with this gene. Additional analyses using a *nifH* profile HMM from the FGPR (Fish *et al*., 2013) and a custom profile HMM built using Uniprot *nifH* protein sequences (http://www.uniprot.org/) (UniProt Consortium, 2015) yielded similar results; as such, all hits to *nifH* were disregarded. Other N fixing genes (*nifD*, n*ifK*, and *nifW*) investigated were absent from results, albeit a single hit to *nifK* in both FMD samples FS56 and FS154, neither of which subsequences were found to be associated with the *nifK* gene. N cycling FMD gene counts were prominent in N cycling pathways that are reliant on the presence of nitrite and nitrate, which included dissimilatory nitrate reductase genes (*narGH*) and denitrification NO-forming nitrite reductase (*nirK*). Although both *nrfA* and *nirBD* code dissimilatory nitrite reductase and are functionally equivalent enzymes, the gene count associated with *nirBD* was greater than that of *nrfA*. Also, while the dissimilatory NO-forming nitrite reductase gene *nirK* was identified and was associated with an elevated gene count, with respect to other identified N cycling genes in the

metagenomic dataset, the functionally comparable enzyme *nirS* was absent in all FMD samples.

Genes associated with methane/ammonia monooxygenase (AMO) subunits were found in all of the FMD samples, as was the lithotrophic nitrifying bacteria AMO intermediate hydroxylamine (NH_2OH). GhostKOALA findings showed that recovered subsequences from AMO HMMER results were dominated by archaeal reads, which accounted for 66.4% of all taxonomy identified (Figure 8). AO bacteria identified in the subsequences included *Nitrosospira*, *Nitrosomonas*, and *Nitrosococcus* (Supplementary Figure 3). Interestingly, taxonomy associated with the nitrite-oxidizing phylum *Nitrospirae* was found in annotated AMO subsequences (31 total) and their KEGG gene IDs revealed that all of these subsequences are taxonomically related to '*Candidatus* Nitrospira inopinata' (Daims *et al*., 2015; Daims *et al*., 2016). Eight of the subsequences associated with the ammonia oxidation intermediate hydroxylamine dehydrogenase (HAO) were also classified by their KEGG gene IDs as being related to '*Candidatus* Nitrospira inopinata'. The other three HAO subsequences were found to be associated with *Nitrosospira*, *Nitrosomonas*, and *Methylocystis*. An additional 29 subsequences were classified by their KEGG gene IDs as nitrite reductase (*NrfA*) related to '*Candidatus* Nitrospira inopinata'. Most bacteria AMO subsequences were found to classify as methane-oxidizing (MO) methanotrophs (Supplementary Figure 3). Bacterial *nxrB* nitrification genes were also identified in all FSC samples, and all annotated subsequences were associated with the nitrite-oxidizing bacterial phylum *Nitrospirae*.

Discussion

Overview

The results from functional N cycling gene fragment HMMER analyses suggest a potential core N cycle in the FSC samples. Genes associated with nitrification, dissimilatory nitrate reduction to ammonia, denitrification, and assimilatory nitrate reduction (ANR) were identified. In a low-nutrient environment such as FSC, the ability to cycle N through these pathways could provide energy and electrons to the microbial community residing within the FMD. Additionally, the biological cycling of N through multiple oxidation states allows for greater movement and availability of nitrogenous compounds to the microbial community. For example, the intermediate NO_2^-, formed during nitrification and denitrification, is potentially available for uptake through the assimilatory NO_2^- uptake system (Moreno-Vivián and Flores, 2007) or could be reduced to NH_4^+ via the nitrite reductase pathway.

Nitrification is the two-step biological process of oxidizing the most reduced form of N (NH_4^+) to the most oxidized species within the N cycle (NO_3^-) (Ward *et al*., 2011). The initial step of nitrification, aerobic ammonia oxidation (rate-limiting step), is carried out by either AOB or ammonia-oxidizing archaea (AOA) and the second step, aerobic nitrite oxidation, is a redox pathway associated with NOB. Nitrification is an aerobic energy yielding biological process dependent on the presence of oxygen to serve as the terminal electron acceptor. Dissimilatory nitrate reduction to ammonium (DNRA) and denitrification are anaerobic heterotrophic processes that require the availability of organic compounds while using oxidized N species as terminal respiratory electron

acceptors, yielding energy. Though not an energy yielding pathway, ANR is an ammonification process (Stein and Klotz, 2016). An important aspect to consider is that ANR and DNRA would conserve fixed N in the FMD, while denitrification would produce gaseous N loss from the microbial community. We examined functional gene and taxonomic findings associated with the metagenomic and 16S rRNA datasets, specifically examining the potential movement and cycling of this critical element.

Ammonia-Oxidizing Archaea

Thaumarchaeota classification as a phylum was proposed by Brochier-Armanet *et al*. (2008) based on deeper branching from hyperthermophilic crenarchaeota than previously hypothesized. While previously investigated in terrestrial and marine environments (Pester *et al*., 2011; Hatzenpichler *et al*., 2012), several recent studies have found this AOA phylum in subterranean ecosystems (Weidler *et al*., 2008; Tetu *et al*., 2013; Barton *et al.*, 2014, Ortiz *et al*., 2014; Reitschuler *et al.*, 2015; Reitschuler *et al.*, 2016; Zhao *et al*., 2016). A molecular survey that examined 16S rRNA bacterial and archaeal reads in soil samples from 146 sites across North America, South America and Antarctica found that relative archaeal abundance averaged just 2% across all soils and only five soil samples were over 5% (Bates *et al*., 2011). In contrast, two recent metagenomic subterranean investigations by Tetu *et al*. (2013) in a submerged passage found in Weebubbie Cave, Nullarbor Plain Australia, and Ortiz *et al*. (2014) in Kartchner Caverns, southeastern Arizona, found archaeal domain abundance encompassed 19% and 7% of all reads, respectively. *Thaumarchaeota* was disproportionately represented in both cave studies accounting for >90% of archaeal reads in Weebubbie Cave (Tetu *et al*., 2013) and 62% in Kartchner Caverns (Ortiz *et al*., 2014). Our results also show that

archaeal abundance in the metagenomes (average of 8.5%), was greater than that previously observed in soils (Bates *et al*., 2011). *Thaumarchaeota* was also determined to be one of the dominant archaeal organisms present in our metagenomic and 16S rRNA archaeal datasets (Figure 5).

We hypothesize that the oligotrophic nature of caves provides a biological niche for *Thaumarchaeota*. *Thaumarchaeota* represent approximately 20-30% of all planktonic cells in the open sea (Karner *et al*., 2001; Herndl *et al*., 2005; Martens-Habbena and Stahl, 2011), and these cells have been described in many environments, including hot springs (de la Torre *et al*., 2008; Hatzenpichler *et al*., 2008), freshwater lakes (Auguet *et al*., 2012; Hayden and Beman, 2014), open ocean (Karner *et al*., 2001; Venter *et al*., 2004; Wuchter *et al*., 2006; Herfort *et al*., 2007; Stahl and de la Torre *et al*., 2012); wetlands (Sims *et al*., 2012) and a deep soil horizon (Jung *et al*., 2014). *Thaumarchaeota* are thought to do well in low-nutrient environments due to their physiology, cellular structure, and genomic traits that have been documented in the study of the isolated strain *Nitrosopumilus maritimus* SCM1 (Könneke *et al*., 2005).

It is expected that AO bacterial and archaeal organisms compete for ammonium in the environment. Working with this hypothesis, Martens-Habbena *et al*. (2009) found that *N. maritimus* SCM1 in culture maintained exponential growth with an ammonium substrate concentration of ≤ 10 nM; 100-fold lower than required by cultivated AOB (Martens-Habbena *et al*. 2009; Pester *et al*., 2011; Urakawa *et al*., 2011). This organism had a specific affinity for ammonium >200 fold greater than that of AOB, and small substrate additions (0.2 μM ammonium) elicited a metabolic response that was not observed in cultures of *Nitrosococcus* spp. that received the same treatment (Martens-

Habbena *et al*. 2009). However, reduced activity of *N. maritimus* SCM1 was found to occur with ammonium additions ≥ 2 mM and with increased additions of substrate (Martens-Habbena *et al*. 2009), suggesting low-nutrient adaptation. Oligotrophic adaptation by *N. maritimus* SCM1 may be partially explained by the fact that this rod-shaped organism is quite small, with recorded lengths between 0.5 to 0.9 μm and 0.25 μm in width and with a cell volume around 0.023 μm^3 (Urakawa *et al*., 2011). This cell volume is 10 to 100-fold smaller than that of their bacteria AO counterparts (Urakawa *et al*., 2011; Hatzenpichler *et al*., 2012). Another oligotrophic adaptation is its relatively small 1.65 Mb genome (Walker *et al*., 2010). In culture, *N. maritimus* SCM1 has been shown to take a lengthy 15-18 h for chromosome replication and has been found to undergo cellular arrest with the depletion of substrate (Pelve *et al*., 2011). Even their method of carbon fixation, through a "modified" type of the autotrophic hydroxypropionate/hydroxybutyrate cycle associated with the *Crenarchaeota*, was found to be more energy efficient than previously described aerobic carbon-fixation pathways (Könneke *et al*., 2014).

As more than half of identified *Thaumarchaeota* reads in the metagenomic dataset were associated with *N. maritimus* SCM1, we suggest that this microbe could live in FSC FMD by potentially driving its metabolism by scavenging substrate that becomes available through ammonification or other sources discussed below. As AOB require much more substrate for ammonia oxidation to occur, in theory AOA would be anticipated to be found in a greater abundance in FMD than their bacterial AO counterparts, as has previously been observed in soils (Leininger *et al*., 2006). Archaeal dominance in subsequences associated with AMO HMMER hits (Figure 8), of which

approximately 17% was comprised of the genus *Nitrosopumilus*, provides some evidence that suggests AOA in Snowy River passage are more abundant in the FMD than are AO bacteria. This result is not completely unexpected given the limited availability of ammonium substrate for the initial step of nitrification to occur. As such, future studies to better understand AO domain dominance in FSC should use quantitative PCR (qPCR) (Leininger *et al*., 2006) to ascertain AOA presence. A recent study by Zhao *et al*. (2016) that used qPCR to investigate cave sediments found in Heshang Cave, central China, found that *Thaumarchaeota* are potentially responsible for "100% of ammonia oxidation in these sediments".

Ammonia-Oxidizing Bacteria

Based on the quantity of AOA assigned to AMO HMMER subsequences, the abundance of AOA reads in the datasets, and the presumed scarcity of available substrate for bacterial ammonia oxidation to occur, we theorize that AOB may play a decidedly limited role in aerobic ammonia oxidation in the cave. Indeed, it has been found in many soils that AOA are more abundant than AOB based on the presence of AMO gene copies (Leininger *et al*., 2006), and, as previously mentioned, AOA in cave sediments were found to be the dominant ammonia oxidizers present (Zhao *et al*., 2016). Nonetheless, taxonomy and functional genes associated with AOB were identified in both datasets. As such, we here consider their potential contribution to the N cycle in FSC.

AOB are found within the phylum *Proteobacteria*, in the orders *Nitrosomonadales* (*Nitrosomonas* and *Nitrosospira*) and *Chromatiales* (*Nitrosococcus*), and they have a wide-ranging distribution in the environment (Norton, 2011). AOB are

larger than their AOA counterparts, have larger genomes, require higher concentrations of substrate to sustain their metabolism, and have lower substrate affinity (Urakawa *et al*., 2011). All these characteristics suggest difficulty competing with their AO archaeal counterparts. Although these organisms may be weak substrate competitors in the cave, AOB have been reported in other subsurface environments (Chen *et al*., 2009; Pasić *et al*., 2010; Ortiz *et al*., 2013; Tetu *et al*., 2013; Ortiz *et al*., 2014; Hathaway *et al*., 2014a). It has been shown that representatives of *Nitrosomonas europaea*, after long-term substrate starvation, will quickly begin to oxidize augmented substrate (Wilhelm *et al*., 1998). We hypothesize that AOB are able to tolerate nutrient starvation by having a mechanism that allows for a rapid response to the presence of ammonium (Bollmann *et al*., 2005). Additionally, it has been demonstrated that nitrifying bacterial cell activity is greater when these organisms attach to charged surfaces (e.g. clays) that attract ions such as NH_4^+ (Prosser, 2007). FMD are rich in clays (Northup, 2002), which would attract NH_4^+ and potentially represent a niche that AOB could occupy and in which to derive substrate and energy. Northup (2002) determined that ammonium was present in the FMD from Spider and Lechuguilla Caves.

A noteworthy aspect associated with our AOB results was the identification of AMO and HAO genes corresponding with the recently described '*Candidatus* Nitrospira inopinata', which is metabolically capable of completing nitrification ("comammox") (Daims *et al*., 2015). An organism that could fully oxidize ammonia to nitrate in a cave would have an energetic advantage, as comammox yields more energy than oxidizing either NH_4^+ or NO_2^- in single step biochemical reactions (Daims *et al*., 2015). It has been predicted that comammox organisms would favor conditions of "slow, substrate-influx-

limited growth" that results in higher growth yield (Costa *et al*., 2006). The low availability of substrate in FSC may provide an appropriate environment for such organisms to occur. While our results provide some evidence to suggest the potential for this organism existing in the cave, further studies are warranted.

Nitrite-Oxidizing Bacteria

We found genomic evidence for the presence of the NOB key enzyme nitrite oxidoreductase (NXR) in all the investigated FSC samples, specifically the beta subunit (NxrB), which suggests the potential for the occurrence of the second step of nitrification in FSC. The *nxrB* gene has previously been found to serve as an ideal functional marker for this pathway, as well as providing "robust phylogenetic framework" for characterizing not only uncultured *Nitrospira,* but also novel diversity (Pester *et al*., 2014). The spatial distribution of *Nitrospira* populations has been found to be closely associated with AOB (Maixner, 2009), suggesting a mutualistic relationship. NOB closely associated with AOB communities would have the benefit of the presence of NO_2^-, which is a scarce nutrient in many environments (Daims *et al*., 2011). Furthermore, biological nitrite oxidation would prevent decreased AOB activity (Stein and Arp, 1998) and microbial toxicity associated with nitrite accumulation (Rowe *et al*., 1979; Bancroft *et al*., 1979; Stein and Arp, 1998). NOB activity may also be advantageous to microbes present in environmental oxygen-limited niches, such as the underlying FMD mud layer, which could provide NO_3^- as a terminal electron acceptor in the denitrification pathway and as a source of N that could be assimilated through the reduction of nitrate to ammonia for microbial metabolism.

NOB are found within the phylum *Proteobacteria*, in the orders *Rhizobiales* (*Nitrobacter*) and *Chromatiales* (*Nitrococcus* and *Nitrospina*) and the candidate genus "Nitrotoga" (Alawi *et al*., 2007), as well as the phylum *Nitrospirae* in the order *Nitrospirales* (*Nitrospira*) (Daims *et al*., 2011). In our datasets, we identified the presence of the genera *Nitrospira* and *Nitrobacter*, though the latter was only found in the metagenomic reads in two samples and were associated with fewer counts than their NO counterparts. While the presence of *Nitrospira* was anticipated, as this genus has been described in a variety of biomes (Lücker *et al*., 2010), including cave environments (Northup *et al*., 2003; Engel, 2007; Chen *et al*., 2009; Pasić *et al*., 2010; Ortiz *et al*., 2013; Tetu *et al*., 2013; Ortiz *et al*., 2014; Hathaway *et al*., 2014b), the identification of *Nitrobacter* was unexpected. Previous research suggests that *Nitrospira* (*K*-strategist) is adapted to environmental conditions that favor low nitrite and oxygen concentrations, while their NO counterpart *Nitrobacter* (*r*-strategist) is likely to be metabolically successful in surroundings that have high levels of nitrite and oxygen present (Schramm *et al*., 1999). A chemostat culture study found evidence to support the *K/r*-hypothesis that *Nitrobacter* as *r*-strategists have high growth rates and are dominant when copious amounts of substrate are available, while *Nitrospira* as *K*-strategists are slow growing organisms that flourish under conditions of nutrient limitation (Nogueira and Melo, 2006). *Nitrospira* in culture will compete with *Nitrobacter* with short-lived increased nitrite levels; however, *Nitrobacter* will become the dominant NOB if the elevated nitrite levels are sustained (Nogueira and Melo, 2006).

These findings suggest that *Nitrospira* would be better adapted to the low-nutrient conditions found in FSC than *Nitrobacter*. Based on the scarcity of *Nitrobacter* reads in

our datasets and in cave-related literature (Fliermans and Schmidt, 1977; Dichosa, 2008), we theorize that this nitrite-oxidizing genus plays a very limited, if any, role in the cycling of N in Snowy River passage. We hypothesize that '*Candidatus* Nitrospira defluvii' is the likely organism responsible for the oxidation of nitrite in the FMD. All metagenomic FMD reads were associated with this organism, as well as all taxonomy assigned to subsequences corresponding to *nxrB* HMMER hits. A study by Nowka *et al*. (2015) of several *Nitrobacter* and *Nitrospira* spp. in culture found that '*Candidatus* Nitrospira defluvii' had a generation time of 37 h and, along with *Nitrospira moscoviensis*, had the greatest affinity for nitrite of any of the other NOB investigated. Lücker *et al*. (2010) observed "constitutive expression of NXR" in culture. If the '*Candidatus* Nitrospira defluvii' identified in the FMD has similar traits, the continuous expression of NXR would be advantageous because NO_2^- that becomes available could be rapidly oxidized as a source of energy. Based on this information it is not unexpected that a putative NXR enzyme associated with '*Candidatus* Nitrospira defluvii' was identified in Kartchner Caverns, as well as taxonomic reads corresponding with a Weebubbie Cave slime curtain metagenome (Tetu *et al*., 2013).

Denitrification

In the absence of oxygen and in the presence of NO_3^- and/ or NO_2^-, denitrifiers act as facultative anaerobic chemoorganotrophs (Van Spanning *et al*., 2007). This could provide nitrifying bacteria with a source of energy from the mud layer that exists below the FMD where dissolved oxygen may be minimal. For example, AOB have been found to be associated with the production of nitric (NO) and nitrous oxide (N_2O) through the expression of genes associated with denitrification. However, very few AOB KOs

corresponding with denitrification were identified in the FMD. Also, we considered the potential for fixed N loss through denitrification by the nitrite-oxidizing bacterium '*Candidatus* Nitrospira defluvii', as results suggest the presence of this organism in the cave. Based on a genome reconstruction that found the nitrite reductase gene *nirK* to be associated with this organism, Lücker *et al*. (2010) hypothesized that '*Candidatus* Nitrospira defluvii' has the potential to use NO_2^- as a terminal electron acceptor, thus, having the ability to both oxidize and reduce nitrate depending on environmental conditions. In the FMD samples, except for FS58, there were *nirK* assigned KOs associated with '*Candidatus* Nitrospira defluvii', potentially implicating this organism in denitrification pathways in the cave.

The overwhelming majority of taxonomy assigned to *nirK* HMMER subsequences, nearly three times that of *all* identified bacterial NO-forming subsequences, were assigned to *Thaumarchaeota*, suggesting this phylum plays a key role in the loss of N in the FMD. Though *nirK* homologs have been found in nearly all *Thaumarchaeota* genomes, excluding *Cenarchaeum symbiosum*, which has a related multicopper oxidase (Hallam *et al*., 2006), any potential role these organisms play in denitrification remains unclear as gene expression of these enzymes has been observed in the presence of oxygen (Hatzenpichler *et al*., 2012). Recent research offers a likely mechanism associated with AOA that explains the expression of these genes in aerobic conditions. It is now recognized that AOA both produce and use the intermediate NH_2OH by a yet elusive gene and corresponding enzyme (Vajrala *et al*., 2013). It has been observed that NO is produced and consumed following the addition of ammonium and NH_2OH to AOA cultures; however, in the presence of an NO scavenger, the result is

inhibition of metabolic activity and NO production; suggesting this molecule is mechanistically critical to AOA (Martens-Habbena *et al*., 2015; Kozlowski, *et al*., 2016). While N_2O is produced as a byproduct of AOA metabolism, this is likely occurring "non-enzymatically", as opposed to AOB contribution of this N species through nitrifier denitrification (Kozlowski, *et al*., 2016). In summary, the potential loss of fixed N by AOA in FSC is hypothesized to be because of nitrification in oxic conditions and not through an anoxic denitrification pathway.

Potential Nitrogen Sources in Snowy River Passage

The microbial community in FSC is reliant on N present in the cave, potential N input from the surface brought in by meteoric water, atmospheric ammonia (active air exchange occurs because of barometric pressure changes on the surface), or input from N-fixing organisms. In the FSC samples, we found no evidence to suggest that diazotrophs are present. It has been proposed that the standard for identifying biological N fixation using computational extrapolation methods from sequencing data requires the presence of six genes associated with N fixation (*NifHDK* and *NifENB*) (Dos Santos *et al*., 2012).

Though FSC is now considered to be primarily hydrologically inactive (Davis and Land, 2006), surface water input following precipitation events could represent a potential nutrient input source. Flowing water was observed in Snowy River passage in 2007, 2008, and 2009 following substantial precipitation events (Land, 2012) and has currently been flowing since September 2014 as of June 2017. Though the sources of water into Snowy River passage are not entirely understood, it is theorized that a rising

water table and/or recharge through a sinkhole or losing stream are responsible for the intermittent flow of water (Land, 2012). In Midnight Creek passage, potential input from a losing surface stream has been observed in the form of particulate organic matter and a surface millipede that were found in the passage (Corcoran, personal communication).

On the surface, Little Creek is thought to be one of the potential water input sources into FSC. Two water samples collected from a presumed Little Creek insurgence on the surface had elevated levels of nitrate (sample average of 24.6 mg/L NO_3^-) (Spilde, unpublished data). Nitrate from a sample collected from Snowy River during a flow event was measured at 9.0 mg/L and a sample collected from Government Spring (Snowy River outlet into the Rio Bonito as determined by a trace dye investigation) (Land, 2012) was measured at 11.9 mg/L NO_3^- (Spilde, unpublished data). A potential anthropogenic source of the nitrate entering Snowy River passage is a local wastewater treatment pool that discharges into Little Creek and is near an assumed insurgence responsible for meteoric water input into Snowy River Passage. Madison and Burnett (1985) proposed background levels of nitrate to be 3.0 mg/L for uncontaminated groundwater. The maximum contaminant level for nitrate in drinking water is 10 mg/L as nitrate-N, as determined by federal standards (Spalding and Exner, 1993). We hypothesize that anthropogenic runoff is responsible for elevated nitrate levels detected in Snowy River waters. Support for anthropogenic input of anions and cations into Snowy River passage is evident by chloride levels that averaged 811.6 mg/L Cl^- from two samples taken from the Little Creek insurgence and another sample from the same area that was found to have a sodium level of 337.0 mg/L Na^+ (Spilde, unpublished data).

Due to the solubility and mobility of NO_3^-, the introduction of this anion to cave FMD soil-like material (Spilde *et al*., 2009) would allow for the movement of nitrate between aerobic and anaerobic zones (e.g. soil-like aggregates) (Cleemput and Samter, 1996), and hence be available to the underlying microbial communities. In addition to potential nitrate input, surface meteoric waters may carry dissolved organic carbon, phosphates, and other nitrogenous species (e.g. NH_4^+) into the cave. However, nutrient availability from meteoric waters flowing into Snowy River passage may be limited to areas in the cave where microbial communities have contact with insurgent waters. The samples from this study were collected many kilometers from areas of the passage that are associated with meteoric drip waters and are unlikely to receive fixed N from these sources.

Another possible source of N input is from the dissolution of carbonate bedrock by chemical weathering (Spilde *et al*., 2009) or from the breakdown of bedrock by microbial activity (Boston *et al*., 2001; Northup *et al*., 2003; Spilde *et al*., 2005; Spilde *et al*., 2009). Similar to what has been previously hypothesized to occur in FMD in Lechuguilla Cave (Spilde *et al*., 2006), organic acids generated from microbial metabolism break down carbonate minerals, releasing reduced metals (e.g. Fe^{2+} and Mn^{2+}) that support chemolithotrophy. This dissolution could potentially release fixed N as well (e.g. NO_3^-).

Conclusions

This study has provided functional and taxonomic evidence for biogeochemical N cycling in a deep carbonate cave. The identification of key N cycling genes also offers

information about energy metabolism, as well as loss and/or retention of N in a local microbial environment. These findings suggest that nitrification, driven by the AOA phylum *Thaumarchaeota*, represents a critical biological pathway in the cave. In the FMDs where nitrogen is limited, the AOA would act as concentrators of scarce N, which NOB could acquire and oxidize. This hypothesis is supported by the presence of the bacterial *nxrB* gene in all samples and the high number of 16S rRNA reads assigned to the nitrite-oxidizing phylum *Nitrospirae*. These phyla may exist in a syntrophic relationship where AOA scavenges and oxidizes NH_4^+ to NO_2^-, which is taken up and oxidized to NO_3^- by the NOB. The conversion of NH_4^+ into NO_3^- through this biological association could be an important nutrient and energy source for other microbial organisms present in the FMD. These results suggest a dynamic bacterial/archaeal community capable of cycling nitrogen in a moderately deep cave environment and expand our knowledge of the nitrogen cycle in caves.

Acknowledgments

We would like to thank the Fort Stanton Cave Study Project (http://fscsp.org/) and the Bureau of Land Management (BLM) for sequencing funding, sampling, and caving support. We thank John Corcoran for his mapping assistance and expert consultation on FSC. Access to the cave would not be possible without permission from the BLM and we thank them for their continued support of scientific expeditions into FSC. We thank Kenneth Ingham Photography for the cave images provided for this manuscript. We thank the NIH (GM-060201) supported UNM IMSD program for their supportive funding of Jason C. Kimble's doctoral research. We also acknowledge all members of the Northup SLiME team for their continued support and valued input.

Footnotes

1. ^http://www.caverbob.com/wlong.htm
2. ^http://www.mrdnalab.com/

3. ^www.illumina.com
4. ^http://metagenomics.anl.gov/
5. ^https://www.r-project.org/
6. ^https://joey711.github.io/phyloseq/
7. ^http://www.drive5.com/usearch/
8. ^http://drive5.com/uchime/gold.fa
8. ^http://qiime.org/scripts/pick_de_novo_otus.html
9. ^http://www.arb-silva.de/download/archive/qiime/
10. ^http://biom-format.org/documentation/biom_conversion.html
11. ^http://hmmer.org/
12. ^http://cbb.pnnl.gov/portal/software/FOAM.html
13. ^http://fungene.cme.msu.edu/
14. ^https://www.python.org/
15. ^https://github.com/mmdavid/FOAM
16. ^http://www.genome.jp/kegg/ko.html
17. ^http://www.uniprot.org/
18. ^https://www.ncbi.nlm.nih.gov/

References

Abeliovich, A. (2006). The nitrite-oxidizing bacteria. *Prokaryotes*. 5, 861-872. doi: 10.1007/0-387-30745-1_41

Alawi, M., Lipski, A., Sanders, T., Eva-Maria-Pfeiffer, and Spieck, E. (2007). Cultivation of a novel cold-adapted nitrite oxidizing betaproteobacterium from the Siberian Arctic. *ISME. J.* 1, 256–264. doi: 10.1038/ismej.2007.34

Auguet, J. C., Triadó-Margarit, X., Nomokonova, N., Camarero, L., and Casamayor, E. O. (2012). Vertical segregation and phylogenetic characterization of ammonia-oxidizing Archaea in a deep oligotrophic lake. *ISME. J.* 6, 1786-1797. doi: 10.1038/ismej.2012.33

Bancroft, K., Grant, I, F., and Alexander, M. (1979). Toxicity of NO(2): Effect of nitrite on microbial activity in an acid soil. *Appl. Environ. Microbiol.* 38, 940-944.

Barton, H. A. Taylor, M. R., and Pace, N. R. (2004). Molecular phylogenetic analysis of a bacterial community in an oligotrophic cave environment. *Geomicrobiol. J.* 21, 11-20. doi: 10.1080/01490450490253428

Barton, H. A., Taylor, N. M., Kreate, M. P., Springer, A. C., Oehrle, S. A., and Bertog, J. L. (2007). The impact of host rock geochemistry on bacterial community structure in oligotrophic cave environments. *Int. J. Speleol.* 36, 93–104. doi: 10.5038/1827-806X.36.2.5

Barton, H. A., Giarrizzo, J. G., Suarez, P., Robertson, C. E., Broering, M. J., and Banks, E. D. (2014). Microbial diversity in a Venezuelan orthoquartzite cave is dominated by the

Chloroflexi (Class *Ktedonobacterales*) and *Thaumarchaeota* Group I.1c. *Front. Microbiol.* **5**:615. doi: 10.3389/fmicb.2014.00615

Bates, S. T., Berg-Lyons, D., Caporaso, J. G., Walters, W. A., Knight, R. and Fierer, N. (2011). Examining the global distribution of dominant archaeal populations in soil. *ISME. J.* 5, 908-917. doi: 10.1038/ismej.2010.171

Bollmann, A., Schmidt, I., Saunders, A. M., and Nicolaisen, M. H. (2005). Influence of starvation on potential ammonia-oxidizing activity and amoA mRNA levels of *Nitrosospira briensis*. *Appl Environ Microbiol.* 71, 1276-1282. doi: 10.1128/AEM.71.3.1276-1282.2005

Boston, P. J., Spilde, M. N., Northup, D. E., Melim, L. A., Soroka, D. S., Kleina, L. G., et al. (2001). Cave biosignature suites: microbes, minerals, and mars. *Astrobiology*. 1, 25-55. doi: 10.1089/153110701750137413

Boxe, C. S., Hand, K.P. Nealson, K.H., Yung, Y.L. and Saiz-Lopez, A. (2012). An active nitrogen cycle on Mars sufficient to support a subsurface biosphere. *Int. J. Astrobiology*. 11, 109-115. doi:10.1017/S1473550411000401

Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008). Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* 6, 245-252. doi:10.1038/nrmicro1852

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods*. 7, 335-336. doi: 10.1038/nmeth.f.303

Carmichael, M. J., Carmichael, S. K., Santelli, C. M., Strom, A., and Bräuer, S. L. (2013). Mn(II)-oxidizing bacteria are abundant and environmentally relevant members of ferromanganese deposits in caves of the Upper Tennessee River Basin. *Geomicrobiol. J.* 30, 779-800. doi: 10.1080/01490451.2013.769651

Chen, Y., Wu, L. Q., Boden, R., Hillebrand, A., Kumaresan, D., Moussard, H. et al. (2009). Life without light: microbial diversity and evidence of sulfur- and ammonium based chemolithotrophy in Movile Cave. *ISME J.* 3: 1093-1104. doi: 10.1038/ismej.2009.57

Cleemput, O. V. and Samater, A. H. (1996). Nitrite in soils: accumulation and role in the formation of gaseous N compounds. *Fertilizer Research*. 45, 81-89. doi:10.1007/BF00749884

Costa, E., Pérez, J., and Kreft, J. U. (2006). Why is metabolic labour divided in nitrification? *Trends Microbiol.* 14, 213-219. doi: 10.1016/j.tim.2006.03.006

Daims, H., Lebedeva, E. V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., et al. (2015). Complete nitrification by Nitrospira bacteria. *Nature*. 528, 504-509. doi: 10.1038/nature16461

Daims, H., Lücker, S., Paslier, D. L., and Wagner, M. (2011). "Diversity, environmental genomics, and ecophysiology of nitrite-oxidizing bacteria," in *Nitrification*, ed. B. B. Ward (Washington, DC: ASM Press), 295-322.

Daims, H., Lücker, S., and Wagner, M. (2016). A New Perspective on Microbes Formerly Known as Nitrite-Oxidizing Bacteria. *Trends Microbiol*. 24, 699-712. doi: 10.1016/j.tim.2016.05.004

Davis D. D. and Land, L. (2006). Recently discovered passages in Fort Stanton Cave, New Mexico, and implications for speleogenesis and regional geomorphic processes in the northern Sacramento Mountains. *New Mexico Geological Society, Guidebook 57*, 219–226.

De la Torre, J. R., Walker, C. B., Ingalls, A. E., Könneke, M., and Stahl, D. A. (2008). Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environ. Microbiol.* 10, 810-818. doi: 10.1111/j.1462-2920.2007.01506.x

Dichosa, A. (2008). Biogenicity and microbial community composition of desert varnish and cave ferromanganese deposits. [dissertation thesis]. [Albuquerque (NM)]: University of New Mexico

Dos Santos, P. C., Fang, Z., Mason, S. W., Setubal, J. C. and Dixon, R. (2012). Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics.* **13**:162. doi: 10.1186/1471-2164-13-162

Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics*. 14, 755-763. doi:10.1093/bioinformatics/14.9.755

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 26, 2460-2461. doi: 10.1093/bioinformatics/btq461

Engel, A. S. (2007). Observations on the biodiversity of sulfidic karst habitats. *J. Caves Karst Stud.* 69:187–206.

Fish, J. A., Chai, B., Wang, Q., Sun, Y., Brown, C. T., Tiedje, J. M., et al. (2013). FunGene: the functional gene pipeline and repository. *Front. Microbiol.* **4**:291. doi: 10.3389/fmicb.2013.00291

Fliermans, C. B. and Schmidt, E. L. (1977). *Nitrobacter* in Mammoth Cave. *Int. J. Speleol.* 9, 1–19. doi: http://dx.doi.org/10.5038/1827-806X.9.1.1

Frierdich, A. J., Hasenmueller, E. A., and Catalano, J. G. (2011). Composition and

structure of nanocrystalline Fe and Mn oxide cave deposits: Implications for trace element mobility in karst systems. 284, 82-96. doi:10.1016/j.chemgeo.2011.02.009

Gázquez, F., Calaforra, J. M., and Forti, P. (2011) Black Mn-Fe crusts as markers of abrupt palaeoenvironmental changes in El Soplao Cave (Cantabria, Spain). *Int. J. Speleol.* 40, 163-169. doi: 10.5038/1827-806X.40.2.8

Giovannoni, S. J., Delong, E. F., Schmidt, T. M., and Pace, N.R. (1990). Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. *Appl. Environ. Microbiol.* 56, 2572-2575.

Hallam, S. J., Mincer, T. J., Schleper, C., Preston, C. M., Roberts, K., Richardson, P. M. et al. (2006). Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analyses of marine Crenarchaeota. *PLoS. Biol.* **4**:e95. doi: 10.1371/journal.pbio.0040095

Hathaway, J. J. M., Garcia, M. G., Balasch, M. M., Spilde, M. N., Stone, F. D., Lurdes, M. D., et al. (2014b). Comparison of Bacterial Diversity in Azorean and Hawai'ian Lava Cave Microbial Mats. *Geomicrobiol. J.* 31, 205-220. doi: 10.1080/01490451.2013.777491

Hathaway, J. J. M., Sinsabaugh, R. L., Dapkevicius, M. L. N. E., and Northup, D. E. (2014a). Diversity of Ammonia Oxidation (amoA) and Nitrogen Fixation (nifH) Genes in Lava Caves of Terceira, Azores, Portugal. *Geomicrobiology Journal*, *31*(3), 221–235. https://doi.org/10.1080/01490451.2012.752424

Hatzenpichler, R. (2012). Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. *Appl. Environ. Microbiol.* 78, 7501-7510. doi: 10.1128/AEM.01960-12

Hatzenpichler, R., Lebedeva, E. V., Spieck, E., Stoecker, K., Richter, A., Daims, H., et al. (2008). A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *PNAS.* 105, 2134–2139. doi: 10.1073/pnas.0708857105

Hayden, C. J. and Beman, M. J. (2014). High Abundances of Potentially Active Ammonia-Oxidizing Bacteria and Archaea in Oligotrophic, High-Altitude Lakes of the Sierra Nevada, California, USA. *PLoS. One*. **9**:e111560. doi: 10.1371/journal.pone.0111560

He, J., Zhang, L., Jin, S., and Zhu, Y. (2008). Bacterial Communities Inside and Surrounding Soil Iron-Manganese Nodules. *Geomicrobiol. J.* 25, 14-24. doi: 10.1080/01490450701829014

Herfort, L., Schouten, S., Abbas, B., Veldhuis, M. J., Coolen, M. J., Wuchter, C., et al. (2007). Variations in spatial and temporal distribution of Archaea in the North Sea in relation to environmental variables. *FEMS Microbiol. Ecol.* 62, 242-257. doi:

10.1111/j.1574-6941.2007.00397.x

Herndl, G. J., Reinthaler, T., Teira, E., van Aken, H., Veth, C., Pernthaler, A., et al. (2005). Contribution of *Archaea* to Total Prokaryotic Production in the Deep Atlantic Ocean. *Appl. Environ. Microbiol.* 71, 2303-2309. doi: 10.1128/AEM.71.5.2303-2309.2005

Hill, C.A. (1990). Sulfuric acid speleogenesis of Carlsbad Cavern and its relationship to hydrocarbons, Delaware Basin, New Mexico and Texas: Tulsa, Ok. *AAPG. Bulletin.* 74, 1685-1694.

Holmes, A. J., Costello, A., Lidstrom, M. E., and Murrell, J. C. (1995). Evidence that particulate methane monooxygenase and ammonia monooxygenase may be evolutionarily related. *FEMS Microbiol. Lett.* 132, 203-208. doi: http://dx.doi.org/10.1111/j.1574-6968.1995.tb07834.x

Hyatt, D., Chen, G., 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**:119. doi: 10.1186/1471-2105-11-119

Jagnow, D. H. (1979). Cavern development in the Guadalupe Mountains: Columbus, Ohio, Cave Research Foundation Report, 55 p.

Jones, D. S., Albrecht, H. L., Dawson, K. S., Schaperdoth, I., Freeman, K. H., Pi, Y., et al. (2012). Community genomic analysis of an extremely acidophilic sulfur-oxidizing biofilm. *ISME. J.* 6, 158-170. doi: 10.1038/ismej.2011.75

Joshi, N. A. and Fass, J. N. (2011) Sickle: A sliding-window, adaptive, quality-based trimming tool for FASTQ files. (Version 1.33) [Software]. https://github.com/najoshi/sickle

Jung, M. Y., Park, S. J., Kim, S. J., Kim, J. G., Damsté J. S. S., Jeon, C. O., et al. (2014). A mesophilic, autotrophic, ammonia-oxidizing archaeon of thaumarchaeal group I.1a cultivated from a deep oligotrophic soil horizon. *Appl. Environ. Microbiol.* 80, 3645-3655. doi: 10.1128/AEM.03730-13

Kanehisa, M. and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27-30.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016a). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, D457-D462. doi: 10.1093/nar/gkv1070

Kanehisa, M., Sato, Y., and Morishima, K. (2016b). BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *J. Mol. Biol.* 428, 726-731. doi: 10.1016/j.jmb.2015.11.006

Karner, M. B., DeLong, E. F., and Karl, D. M. (2001). Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature*. 409, 507-510. doi:10.1038/35054051

Kelley, V. C. (1971). Geology of the Pecos country, southeastern New Mexico. *Memoir*. *24*, 1-78.

Könneke, M., Bernhard, A. E., de la Torre, J., Walker, C. B., Waterbury, J. B., and Stahl, D. A. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature*. 437, 543-546. doi:10.1038/nature03911

Könneke, M., Schubert, Philip, D. M., Brown, P. C., Hügler, M., Standfest, S., and Schwander, T. (2014). Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. *PNAS.* 111, 8239-8244. doi/10.1073/pnas.1402028111

Kennedy, C.B., Scott, S. D., and Ferris, F. G. (2003). Characterization of bacteriogenic iron oxide deposits from Axial Volcano, Juan de Fuca Ridge, Northeast Pacific Ocean. *Geomicrobiol. J.* 20, 199-214. doi: 10.1080/01490450390206683

Kozlowski, J. A., Stieglmeier, M., Schleper, C., Klotz, M. G., and Stein, L. Y. (2016). Pathways and key intermediates required for obligate aerobic ammonia-dependent chemolithotrophy in bacteria and Thaumarchaeota. *ISME. J.* 10, 1836-1845. doi: 10.1038/ismej.2016.2

Land, L. (2012). Geophysical prospecting for new cave passages: Fort Stanton Cave, New Mexico, USA. *Carbonate. Evaporite.* 27, 97-102. doi: 10.1007/s13146-012-0105-6

Land L, Polyak V., and Newton B. T. (2010). The Snowy River formation in Fort Stanton Cave, New Mexico: Results from radiometric dating and hydrologic observations of the world's longest speleothem. In: Eaton M., Landres P. (eds) compilers, *Decade of Discovery in the National Landscape Conservation System*, on line report. http://www.nckri.org/research/basic_research/Snowy_River_Deposit.pdf

Lee, N. M., Meisinger, D. B., Aubrecht, R., Kovacik, L., Saiz-Jimenez, C., Baskar, S., et al. (2012). "Caves and karst environments," in *Life at Extremes: Environments, Organisms and Strategies for Survival*, ed. E. M. Bell (CAB International), 320-344.

Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G. W., et al. (2006). Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature*. 442, 806-809. doi:10.1038/nature04983

Lücker, S., Wagner, M., Maixner, F., Pelletier, E., Koch, H., Vacherie, B., et al. (2010). A Nitrospira metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria. *PNAS.* 107, 13479-13484. doi: 10.1073/pnas.1003860107

Madison, R. J., and Burnett, J. O. (1985). "Overview of the occurrence of nitrate in

groundwater of the United States," in *USGS National Water Summary 1984* (U.S. Geological Survey), 93-105.

Martens-Habbena, W. and Stahl, D. A. (2011). Nitrogen metabolism and kinetics of ammonia-oxidizing archaea. *Meth. Enzymol.* 496, 465-487. doi: 10.1016/B978-0-12-386489-5.00019-1

Martens-Habbena, W., Berube, P. M., Urakawa, H., de la Torre, J. R., and Stahl, D. A. (2009). Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature*. 461, 976-979. doi: 10.1038/nature08465

Martens-Habbena, W., Qin, W., Horak, R. E. A., Urakawa, H., Schauer, H., Moffett, J. W., et al. (2015). The production of nitric oxide by marine ammonia-oxidizing archaea and inhibition of archaeal ammonia oxidation by a nitric oxide scavenger. *Environ. Microbiol.* 17, 2261-74. doi: 10.1111/1462-2920.12677

Maixner, F. M. (2009). The ecophysiology of nitrite-oxidizing bacteria in the genus *Nitrospira*: Novel aspects and unique features. [dissertation thesis]. [Vienna (AT)]: University of Vienna

McMurdie, P. J. and Holmes, S. (2013) phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data *PLoS. One*. **8**:e61217. doi: 10.1371/journal.pone.0061217

Mercier, C., Boyer, F., Bonin, A., and Coissac, E. (2013). SUMATRA and SUMACLUST: fast and exact comparison and clustering of sequences. (Version 1.0.00) [Software]. http://metabarcoding.org/sumatra

Meyer, F., Paarmann, D., D'Souza, M. D., Glass, E. M., Kubal, M., Paczian, T., et al. (2008). The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*. **9**:386. doi:10.1186/1471-2105-9-386

Moreno-Vivián, C. and Flores, E. (2007). "Nitrate assimilation in bacteria," in *Biology of the Nitrogen Cycle*, ed. H. Bothe (The Netherlands, AMS: Elsevier), 263-282.

Nogueira, R. and Melo, L. F. (2006). Competition between Nitrospira spp. and Nitrobacter spp. in nitrite-oxidizing bioreactors. *Biotechnol. Bioeng.* 95, 169-175. doi: 10.1002/bit.21004

Northup, D. E. and Lavoie, K. H. (2001). Geomicrobiology of caves. *Geomicrobiol. J.* 18, 199-222. doi: 10.1080/01490450152467750

Northup, D. E., Barns, S. M., Yu, L. E., Spilde, M. N., Schelble, R. T., Dano, K. E., et al. (2003). Diverse microbial communities inhabiting ferromanganese deposits in Lechuguilla and Spider Caves. *Environ. Microbiol.* 5, 1071-1086. doi: 10.1046/j.1462-2920.2003.00500.x

Northup, D. E., Dahm, C. N., Melim, L. A., Spilde, M. N., Crossey, L. J., Lavoie, K. H., et al. (2000). Evidence for geomicrobiological interactions in Guadalupe caves. *J. Caves Karst Stud.* 62, 80-90.

Northup, D. E., Snider, J. R., Spilde, M. N., Porter, M. L., Van de Kamp, J. L., Boston, P. J., et al. (2010). Diversity of rock varnish bacterial communities from Black Canyon, New Mexico. *J Geophys. Res.* 115, 1-19. doi:10.1029/2009JG001107

Norton, J. M. (2011). "Diversity and environmental distribution of ammonia oxidizing bacteria," in *Nitrification*, ed. B. B. Ward (Washington, DC: ASM Press), 39-55.

Nowka, B., Daims, H., and Spieck, E. (2015). Comparison of oxidation kinetics of nitrite-oxidizing bacteria: nitrite availability as a key factor in niche differentiation. *Appl. Environ. Microbiol.* 81, 745-753. doi: 10.1128/AEM.02734-14

Ortiz, M., Legatzki, A., Nelson, J.W., Fryslie, B., Nelson, W.M., Wing, R.A., et al. (2014). Making a living while starving in the dark: metagenomic insights into the energy dynamics of a carbonate cave. *ISME. J.* 8, 478-491. doi:10.1038/ismej.2013.159

Ortiz, M., Neilson, J. W., Nelson, W. M., Legatzki, A., Byrne, A., Yu, Y., et al. (2013). Profiling bacterial diversity and taxonomic composition on speleothem surfaces in Kartchner Caverns, AZ. *Microb. Ecol.* 65, 371-383.

Pasić, L., Kovce, B., Sket, B., and Herzog-Velikonja, B. (2010). Diversity of microbial communities colonizing the walls of a Karstic cave in Slovenia. *FEMS Microbiol. Ecol.* 71, 50-60. doi: 10.1111/j.1574-6941.2009.00789.x

Pelve, E. A., Lindås, A. C., Martens-Habbena, W., de la Torre, J. R., Stahl, D. A., and Bernander, R. (2011). Cdv-based cell division and cell cycle organization in the thaumarchaeon Nitrosopumilus maritimus. *Mol. Microbiol.* 82, 555-566. doi: 10.1111/j.1365-2958.2011.07834

Pester, M., Maixner, F., Berry, D., Rattei, T., Koch, H., Lücker, S., et al. (2014). NxrB encoding the beta subunit of nitrite oxidoreductase as functional and phylogenetic marker for nitrite-oxidizing Nitrospira. *Environ. Microbiol.* 16, 3055-3071. doi: 10.1111/1462-2920.12300

Pester, M., Schleper, C., and Wagner, M. (2011). The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. *Curr. Opin. Microbiol.* 14, 300-306. doi: 10.1016/j.mib.2011.04.007

Prestat, E., David, M. M., Hultman, J., Tas, N., Lamendella, R., and Dvornik, J. (2014). FOAM (Functional Ontology Assignments for Metagenomes): A Hidden Markov Model (HMM) database with environmental focus. *Nucleic Acids Res.* **42**:e145. doi: 10.1093/nar/gku702

Prosser, J. I. (2007). "The Ecology of Nitrifying Bacteria," in *Nitrification*, ed. B. B. Ward (Washington, DC: ASM Press), 223-243.

Rossi, C., Lozano, R. P., Isanta, N., and Hellstrom, J. (2010). Manganese stromatolites in caves: El Soplao (Cantabria, Spain). *Geology*. 38, 1119-1122. doi:10.1130/G31283.1

Rowe, J. J., Yarbrough, J. M., Rake, J. B. and Eagon, R. G. (1979). *Curr. Microbiol.* 2, 51-54. doi:10.1007/BF02601735

Quast, C., Pruesse, E., Yilmaz, J., Schweer, T., Yarza, P., Peplies, J. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**:D590-D596. doi: 10.1093/nar/gks1219

R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [Software]. http://www.R-project.org

Reitschuler, C., Lins, P., Schwarzenauer, T., Spötl, C., Wagner. A. O., and Illmer, P. (2015). New undescribed lineages of non-extremophilic archaea form a homogeneous and dominant element within alpine moonmilk microbiomes. *Geomicrobiol. J.* 32, 890-902. doi: 10.1080/01490451.2015.1025317

Reitschuler, C., Spötl, C., Hofmann, K., Wagner, A. O., and Illmer, P. (2016). Archaeal Distribution in Moonmilk Deposits from Alpine Caves and Their Ecophysiological Potential. *Microb. Ecol.* 71, 686-699. doi: 10.1007/s00248-015-0727-z

Sarbu, S. M., Kane, T. C., and Kinkle, B. K., (1996). A chemoautotrophically based cave ecosystem. *Science*. 272, 1953–1955. doi: 10.1126/science.272.5270.1953

Schramm, A., De Beer, D., van den Heuvel, J. C., Ottengraf, S., and Amann, R. (1999). Microscale distribution of populations and activities of Nitrosospira and Nitrospira spp. along a macroscale gradient in a nitrifying bioreactor: quantification by in situ hybridization and the use of microsensors. *Appl. Environ. Microbiol.* 65, 3690-3696.

Sims, A., Horton, J., Gajaraj, S., McIntosh, S., Miles, R. J., Mueller, R. et al. (2012). Temporal and spatial distributions of ammonia-oxidizing archaea and bacteria and their ratio as an indicator of oligotrophic conditions in natural wetlands. *Water Res.* 46, 4121-4129. doi: 10.1016/j.watres.2012.05.007

Spalding, R. F. and Exner, M. E. (1993). Occurrence of nitrate in groundwater—a review *J. Environ. Qual.* 22, 392-402. doi: 10.2134/jeq1993.00472425002200030002x

Spear, J. R., Barton, H. A., Robertson, C. E., Francis, C. A., and Pace, N. R. (2007). Microbial community biofabrics in a geothermal mine adit. *Appl. Environ. Microbiol*. 73, 6172–6180. doi:10.1128/AEM.00393-07

Spilde, M. S., Kooser, A, Boston, P. J., and Northup, D. E. (2009). "Speleosol: A

subterranean soil," in Proceedings of the 15th International Congress of Speleology, ed. W. B. White (Huntsville, AL: National Speleological Society), 338-344.

Spilde, M. N., Northup, D. E., and Boston, P. J. (2006). "Ferromanganese deposits in the caves of the Guadalupe Mountains" in Caves and Karst of Southeast New Mexico, ed L. Land L, Lueth (The New Mexico Geological Society), 161–166.

Spilde, M. N., Northup, D. E., Boston, P. J., Schelble, R. T., Kathleen, E. D., Crossey, L. J., et al. (2005). Geomicrobiology of cave ferromanganese deposits: A field and laboratory investigation. *Geomicrobiol. J.* 22, 99-116. doi: 10.1080/01490450590945889

Stahl, D. A. and de la Torre, J. R. (2012). Physiology and diversity of ammonia-oxidizing archaea. *Annu. Rev. Microbiol.* 66, 83-101. Doi: 10.1146/annurev-micro-092611-150128

Stein, L. Y. and Arp, D. J. (1998). Loss of ammonia monooxygenase activity in *Nitrosomonas europaea* upon exposure to nitrite. *Appl. Environ. Microbiol.* 64, 4098-4102.

Stein, L. Y. and Klotz, M. G. (2016). The nitrogen cycle. *Curr Biol.* 26, 94-98. doi: 10.1016/j.cub.2015.12.021

Takai, K. and Horikoshi, K. (2000). Rapid Detection and Quantification of Members of the Archaeal Community by Quantitative PCR Using Fluorogenic Probes. *Appl. Environ. Microbiol.* 66, 5066-5072. doi: 10.1128/AEM.66.11.5066-5072.2000

Takai, K., Moser, D. P., DeFlaun, M., Onstott, T. C., and Fredrickson, J. K. (2001). Archaeal diversity in waters from deep South African gold mines. *Appl. Environ. Microbiol.* 67, 5750-5760. doi: 10.1128/AEM.67.21.5750-5760.2001

Tetu, S. G., Breakwell, K., Elbourne, L. D. H., Holmes, A. J., Gillings, M. R., and Paulsen, I. T. (2013). Life in the dark: Metagenomic evidence that a microbial slime community is driven by inorganic nitrogen metabolism. *ISME. J.* 7, 1227–1236. doi: 10.1038/ismej.2013

Vajralaa, N., Martens-Habbenab, W., Sayavedra-Sotoa, L. A., Schauerc, A., Bottomleyd, P. J., Stahl, D. A., et al. (2013). Hydroxylamine as an intermediate in ammonia oxidation by globally abundant marine archaea. *PNAS.* 110, 1006-1011. doi: 10.1073/pnas.1214272110

Venter, C. J., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., et al. (2004). Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science*. 304, 66-74. doi: 10.1126/science.1093857

UniProt Consortium. (2015). UniProt: a hub for protein information. *Nucleic Acids Res.* **43**:D202-D212. doi: 10.1093/nar/gku989

Urakawa, H., Martens-Habbena, W., and Stahl, D. A. (2011). "Physiology and genomics of ammonia-oxidizing archaea," in *Nitrification*, ed. B. B. Ward (Washington, DC: ASM Press), 117-155.

Van Spanning, R. J. M., Richardson, D. J., and Ferguson, S. J. (2007). "Introduction to the Biochemistry and Molecular Biology of Denitrification," in *Nitrification*, ed. B. B. Ward (Washington, DC: ASM Press), 3-20.

Walker, C. B., de la Torre, J. R., Klotz, M. G., Urakawa, H., Pinel, N., Arp, D. J., et al. (2010). Nitrosopumilus maritimus genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *PNAS.* 107, 8818-8823. doi: 10.1073/pnas.0913533107

Ward, B. B., Arp, D. J., and Klotz, M. G. (2011). "Nitrification: An introduction and overview of the state of the field," in *Nitrification*, ed. B. B. Ward (Washington, DC: ASM Press), 3-8.

Weidler, G. W., Dornmayr-Pfaffenhuemer, M., Gerbl, F. W., Heinen. W., and Stan-Lotter, H. (2007). Communities of *Archaea* and *Bacteria* in a subsurface radioactive thermal spring in the Austrian Central Alps, and evidence of ammonia-oxidizing *Crenarchaeota*. *Appl. Environ. Microbiol.* 73, 259–270. doi:10.1128/AEM.01570-06

Weidler, G. W., Gerbl, F. W., and Stan-Lotter, H. (2008). Crenarchaeota and their role in the nitrogen cycle in a subsurface radioactive thermal spring in the Austrian Central Alps. *Appl. Environ. Microbiol.* 74, 5934-5942. doi: 10.1128/AEM.02602-07

Wilhelm, R., Abeliovich, A., and Nejidat, A. (1998). Effect of Long-Term Ammonia Starvation on the Oxidation of Ammonia and Hydroxylamine by *Nitrosomonas Europaea*. *J. Biochem.* 124, 811-815.

Wilke, A., Glass, E. M., Bischof, J., Braithwaite, D., Gerlach, W., Harrison, T., et al. (2015) MG-RAST Manual for version 3.6, revision 3. ftp.metagenomics.anl.gov/data/manual/mg-rast-manual.pdf

Williams, K. E., McKay, C. P., Toon, O. B., and Head, J. W. (2010a). Do ice caves exist on Mars? *Icarus*. 209, 358–368. doi:10.1016/j.icarus.2010.03.039

Wuchter, C., Abbas, B., Coolen, M. J., Herfort, L., van Bleijswijk, J., Timmers, P., et al. (2006). Archaeal nitrification in the ocean. *PNAS.* 103, 12317–12322. doi: 10.1073/pnas.0600756103

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd merger. *Bioinformatics.* 30, 614-620. doi:10.1093/bioinformatics/btt593

Zhao, R., Wang, H., Yang, H., Yun, Y., and Barton, H. A. (2017). Ammonia oxidizing

archaea dominant over bacterial ammonia oxidizers within alkaline cave sediments. *Geomicrobiol. J.* http://dx.doi.org/10.1080/01490451.2016.1225861

Figures and Legends

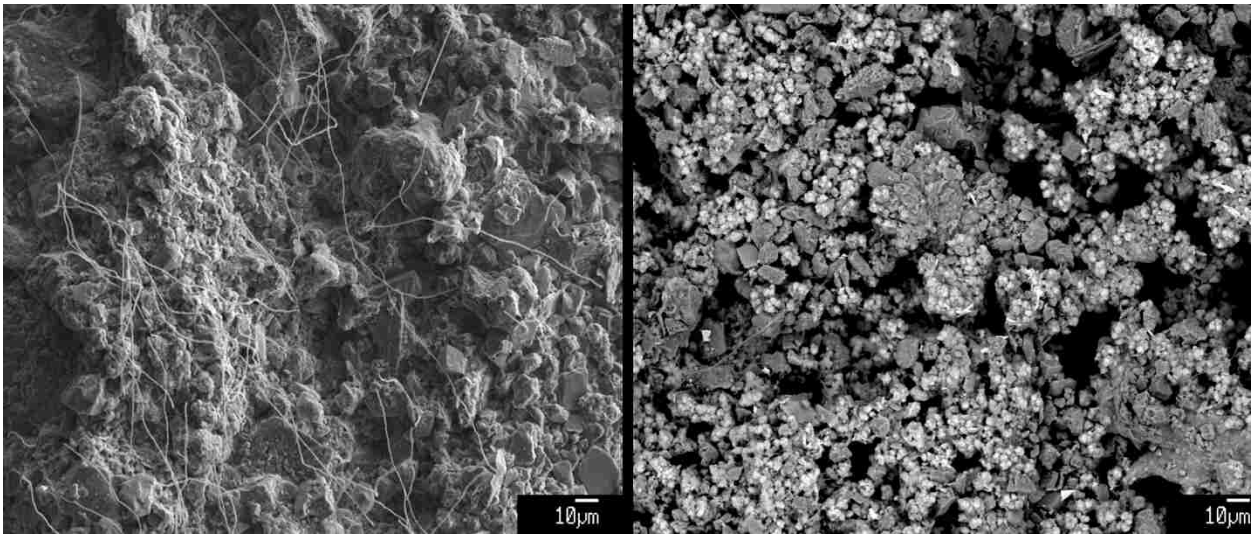


Figure 1. SEM images of black ferromanganese from a wall in Snowy River. Left: Secondary electron image of microbial filaments present on the surface. Right: Backscattered electron image of clay and fine sand (medium gray) inter-grown with manganese oxide minerals (light). Most of the manganese oxide occurs as small spherules around 2-5 μm in diameter. Scale bars are 10 μm.

Figure 2. Images of mineral deposits found on the walls and ceiling of Snowy River passage in Fort Stanton Cave. (**A**) Thin coating of manganese oxide (Black FMD) precipitating on top of mud layer. (**B**) Sample FS58 collection from manganese coating found on the wall. (**C**) A section of Snowy River passage showing the omnipresence of FMD on the walls and ceilings of the passage. (Images courtesy of Kenneth Ingham Photography)

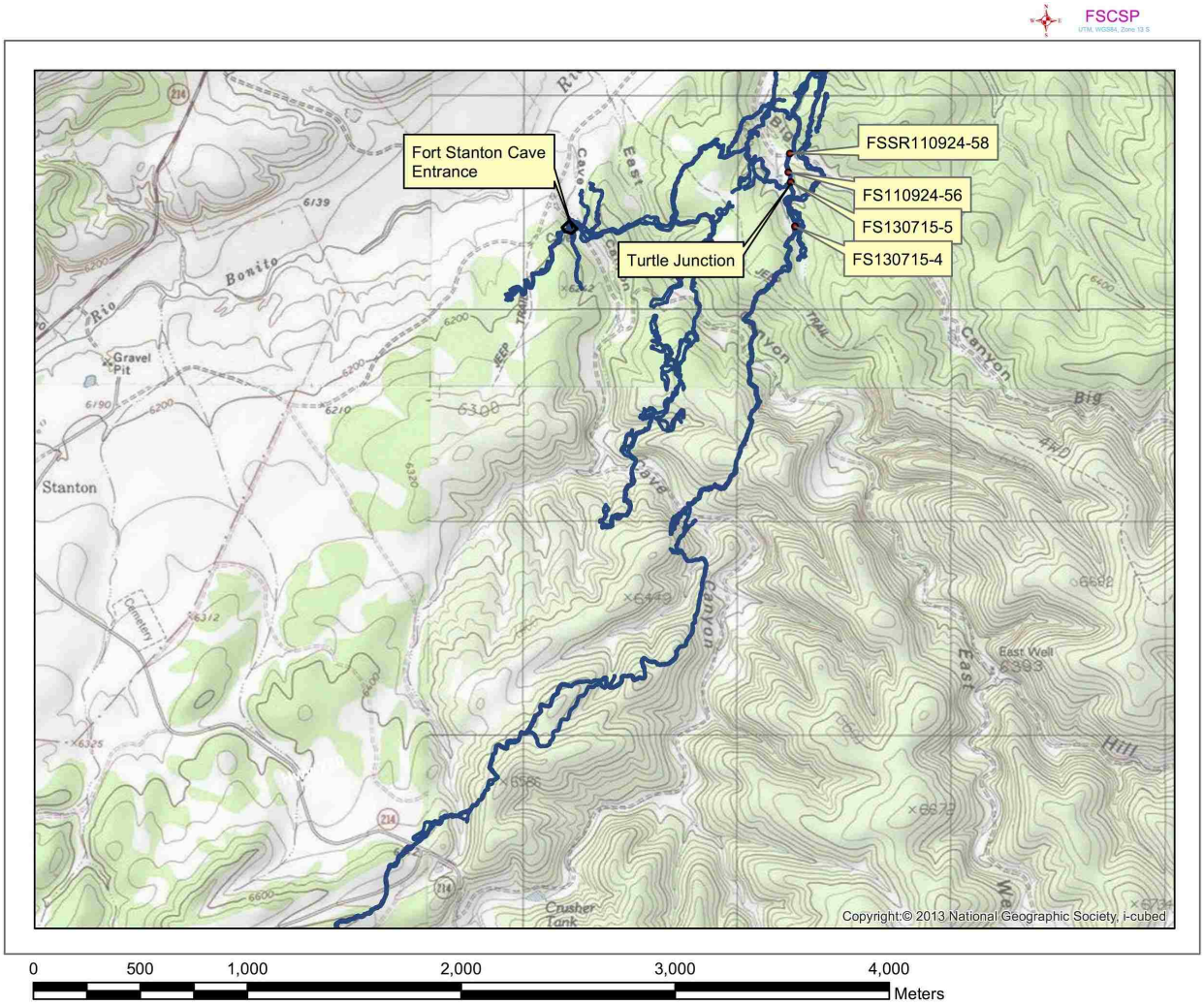


Figure 3. Topographic map showing Fort Stanton Cave passages beneath the surface. Turtle Junction indicates the access point to Snowy River, along which samples were collected. (Map courtesy of John Corcoran).

SampleID	Sample Color	Survey Station	Cave Altitude (m)	Surface Elevation (m)	Cave Depth Below Surface (m)	Distance from Turtle Junction (m)
FS110924-56 (FS56)	Brown FMD	SRS20	1835	1889	54	47
FS110924-58 (FS58)	Brown FMD	SRS14	1835	1873	38	152
FS130715-4 (FS154)	Black FMD	SRS38	1839	1897	58	234
FS130715-5 (FS155)	Black FMD	SRS23	1836	1897	51	0

Table 1. Summary of metadata associated with samples sequenced for metagenomic and 16S rRNA reads. Note: Sample ID abbreviations.

SampleID	MG-RAST ID	Raw Read Counts	Paired-End Read Counts	MGRAST Upload Counts	MG-RAST Post Read QC Counts
FS56	4681809.3	8,436,176	6,629,082	6,570,457	6,566,557
FS58	4681806.3	9,560,952	8,053,731	7,948,139	7,946,963
FS154	4681808.3	7,822,061	6,603,356	6,533,479	6,532,591
FS155	4681807.3	8,179,037	6,622,455	6,554,580	6,547,801

Table 2. Reads removed through pipeline quality control steps for each sample. In the samples, an average of 18.9% of reads were removed through quality control steps.

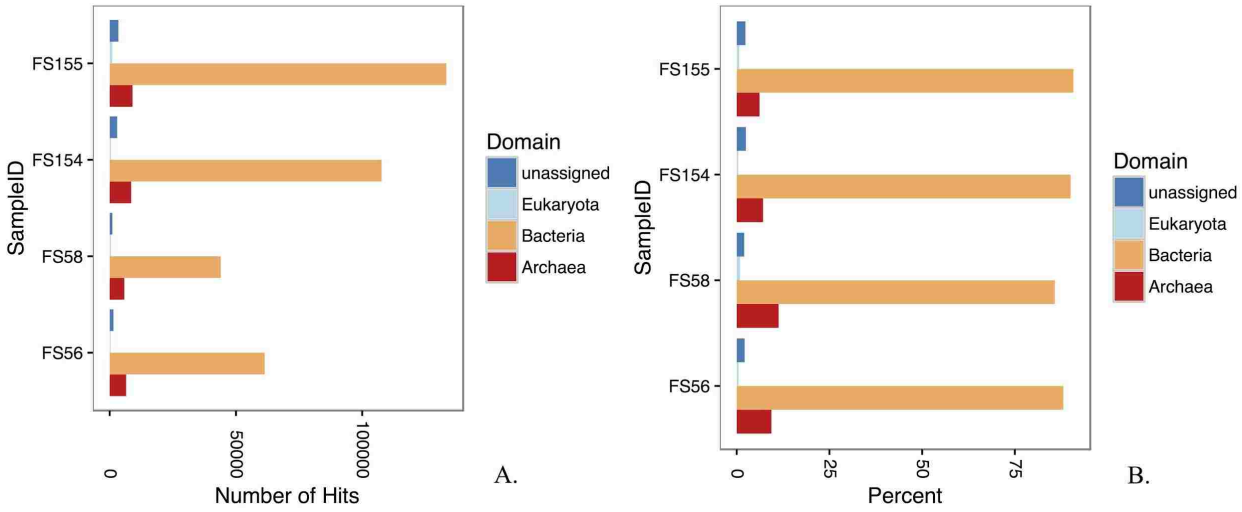


Figure 4. Domain assignments by KEGG with a maximum e value cutoff of $1e^{-5}$, a minimum identity cutoff of 60 %, and a 65 bp minimum alignment length cutoff. (**a**) The complete metagenomic dataset by sample and number of hits to each domain. (**b**) The complete metagenomic dataset representing percent by domain for each sample.

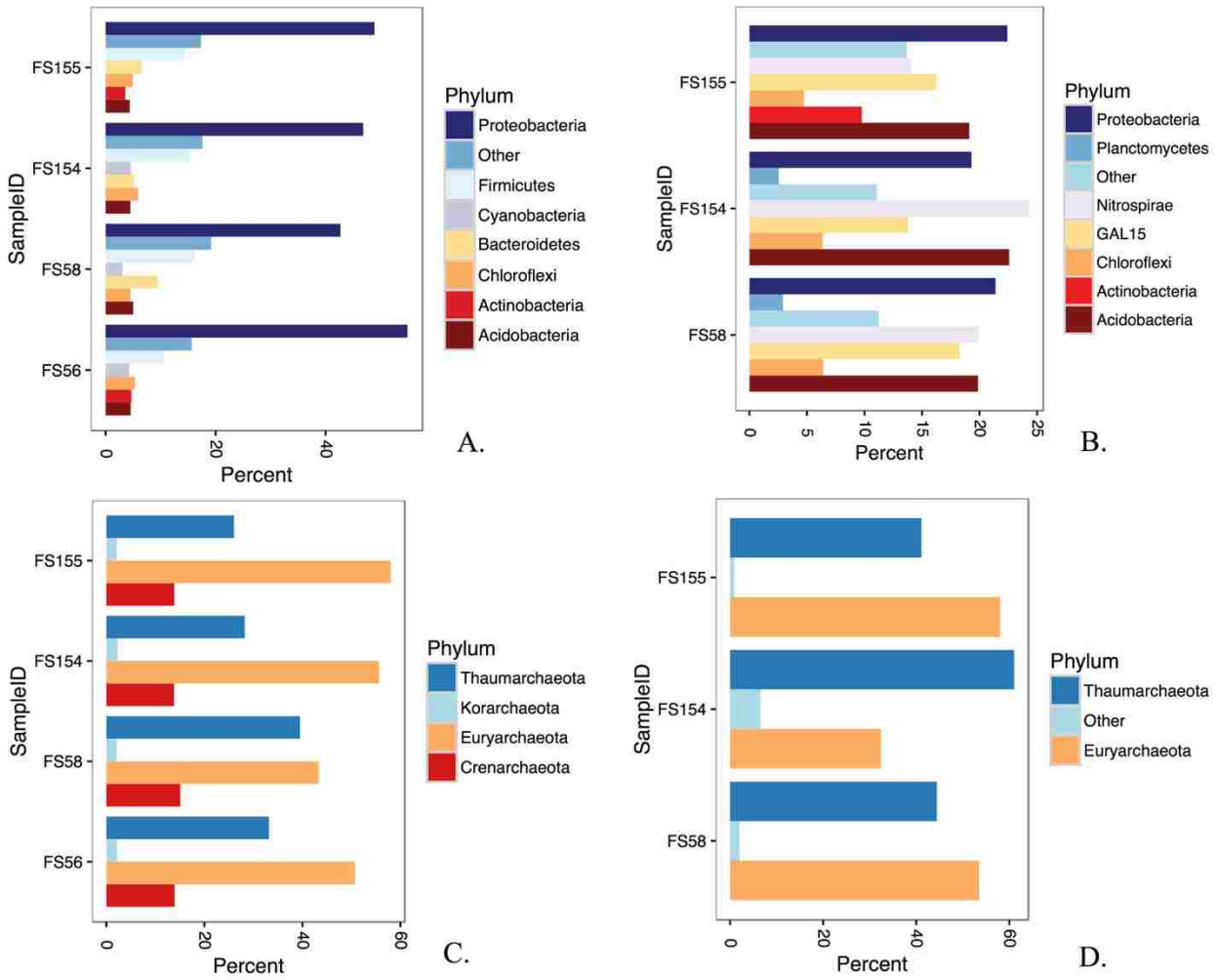


Figure 5. Taxonomic assignments by phylum for metagenomic and 16S rRNA reads. (**a**) Metagenomic top 6 bacterial phyla by sample. (**b**) 16S rRNA top 6 bacterial phyla by sample. (**c**) Metagenomic top archaeal phyla by sample. (**d**) 16S rRNA top archaeal phyla by sample. Note: Sample FS56 is absent from figure panels B and D, as there was insufficient DNA to sequence for 16S rRNA archaeal and bacterial reads.

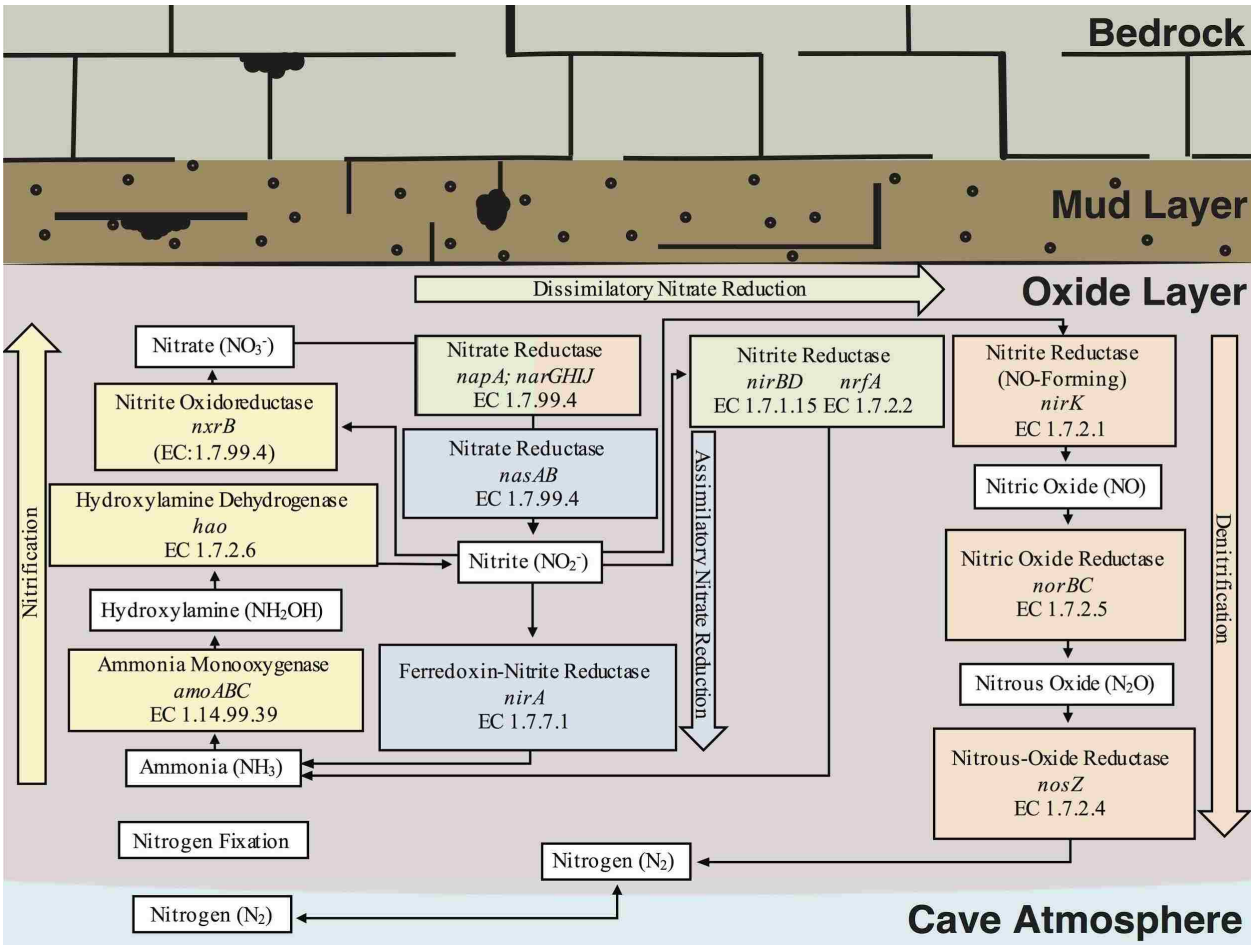


Figure 6. Overview of nitrogen cycle functional genes and metabolic pathways identified in the samples based on HMMER hits to the FOAM database.

KO	Gene	FS56	FS58	FS154	FS155
K10944	*pmoA-amoA*	129	118	175	95
K10945	*pmoB-amoB*	104	21	79	55
K10946	*pmoC-amoC*	100	65	124	58
K10535	*hao*	2	0	4	5
	nxrB	83	28	34	38
K02567	*napA*	4	2	11	9
K00370	*narG*	757	64	205	260
K00371	*narH*	276	67	81	121
K00374	*narI*	20	0	15	23
K00373	*narJ*	12	0	4	14
K00362	*nirB*	30	11	110	77
K00363	*nirD*	45	77	89	73
K03385	*nrfA*	21	3	41	44
K00368	*nirK*	165	456	362	294
K04561	*norB*	1	0	2	7
K02305	*norC*	6	10	9	2
K00376	*nosZ*	58	20	30	20
K00372	*nasA*	31	25	52	110
K00360	*nasB*	2	0	1	1
K00366	*nirA*	115	27	93	63

Figure 7. Overview of KOs and genes assigned for each of the FSC samples.

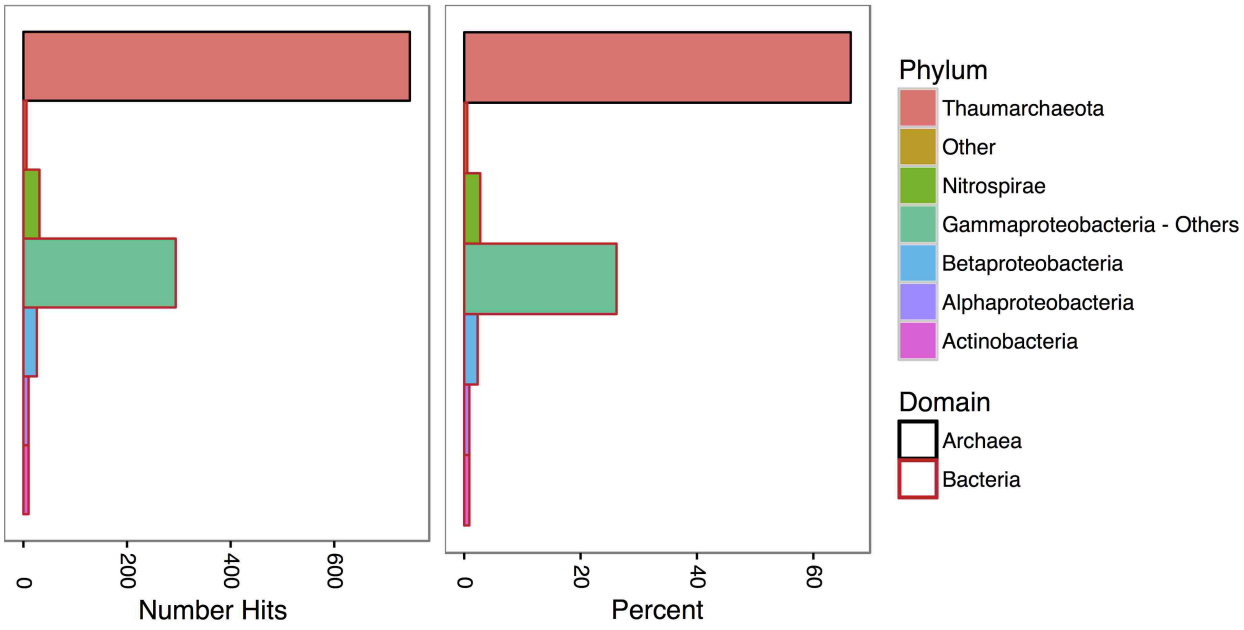
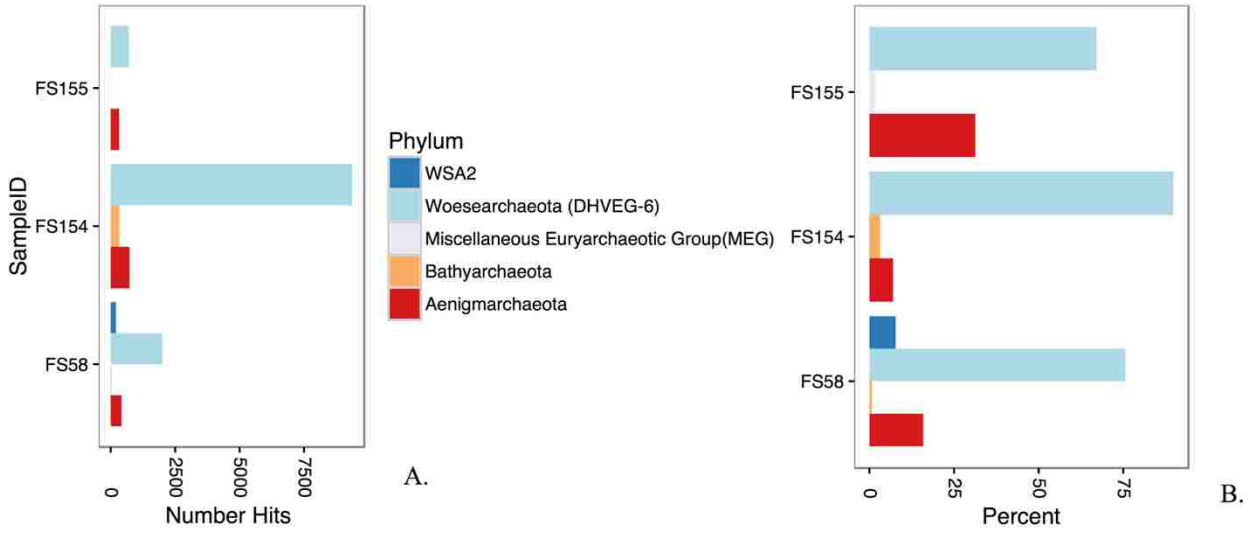
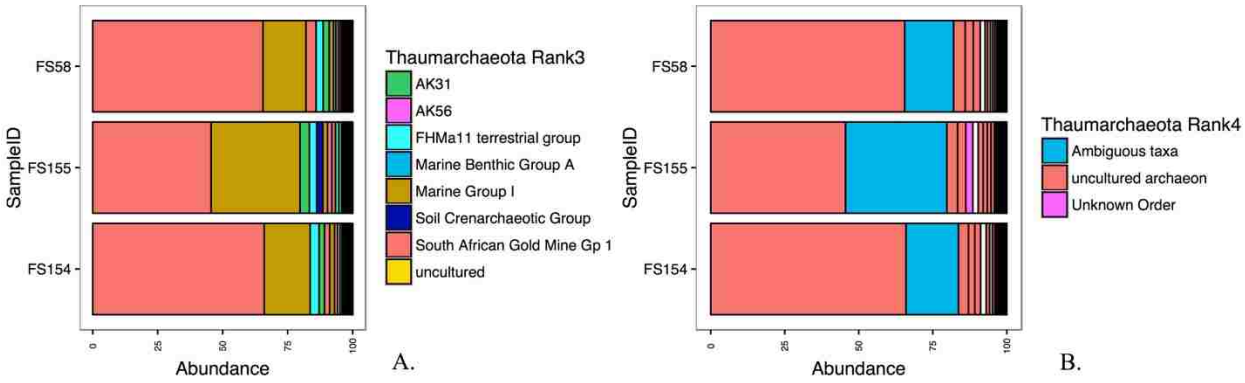


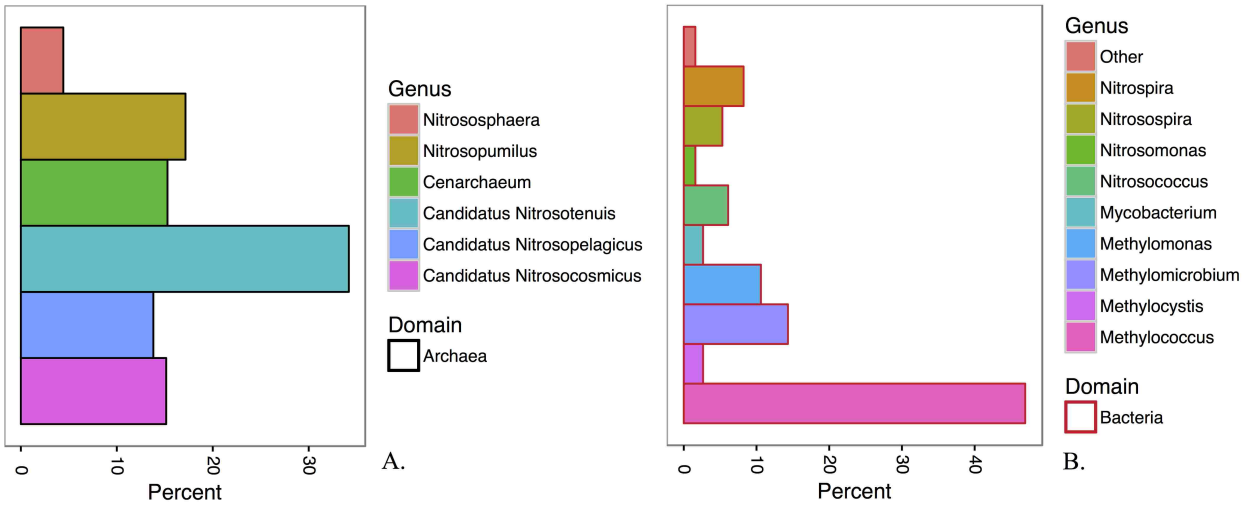
Figure 8. The bar charts represent GhostKOALA taxonomic data associated with ammonia oxidation subsequences that were identified in the 4 FMD samples using HMMER. Evident is that there were more ammonia oxidation genes identified in the archaeal domain than in Bacteria in the FMD samples.



Supplementary Figure 1. The bar charts represent a classification of "other" archaeal phyla identified in the 16S rRNA reads. (**a**) Archaeal "other" phyla by the number of hits identified. (**b**) Archaeal "other" phyla by percent of reads identified in the three FMD samples.



Supplementary Figure 2. The bar charts represent 16S rRNA taxonomic AOA archaeal data (**a**) *Thaumarchaeota* percent by rank 3 classification. (**b**) *Thaumarchaeota* percent by rank 4 classification, which indicates a majority of the reads are associated with previously uncultured organisms.



Supplementary Figure 3. The bar charts represent GhostKOALA taxonomic data associated with ammonia oxidation subsequences that were identified in the 4 FMD samples using HMMER. (**a**) Archaeal AMO subsequence associated taxonomy identified broad *Thaumarchaeota* genera. (**b**) Bacterial genera associated with ammonia oxidation were identified; however, many of the reads were associated with methane oxidizers. Ammonia monooxygenase is evolutionarily related to methane monooxygenase (Holmes *et al*., 1995).

Chapter 3: Comparison of Archaeal and Bacterial Communities Found in Surface Soils and Underlying Ferromanganese Deposits in Fort Stanton Cave, NM, USA

Jason C. Kimble[1], Ara S. Winter, Robert L. Sinsabaugh[1], and Diana E. Northup[1]

[1]Department of Biology, MSC03-2020, University of New Mexico, Albuquerque, NM 87131

*Corresponding Author
Jason C. Kimble – jkimble@unm.edu

Abstract

Deep arid-land carbonate caves receive limited surface input; hence, critical constituents (e.g. C_{org} and NH_4^+) required for microbial communities are often found in low quantities. We examined the archaeal and bacterial makeup of low-nutrient soil-like material (*speleosol*), rich in Fe- and Mn-oxides, occurring on the walls and ceilings of a carbonate cave (Fort Stanton Cave, NM, USA) and in overlying surface soils. Results of 16S rRNA gene sequences indicate that the bacterial and archaeal communities in the cave are taxonomically dissimilar to their surface counterparts. Core microbiome results, representing operational taxonomic units (OTUs) occurring in not less than 80% of all samples, determined that there were only 19 and 17 archaeal and bacterial OTUs shared between surface and cave samples, respectively. The surface archaeal community was primarily represented by the *Thaumarchaeota* class Soil Crenarchaeotic Group (SCG), while dominant archaeal groups in the subsurface included the *Euryarchaeota* class *Thermoplasmata* and the *Thaumarchaeota* classes South African Gold Mine Gp 1 (SAGMCG-1), Marine Group I (MGI), and AK31. Bacterial cave OTUs significantly different from surface bacteria included the phyla *Nitrospirae,* GAL15, *Omnitrophica*, RBG-1 (Zixibacteria), *Latescibacteria*, SBR1093, and *Ignavibacteriae*. Our results provide taxonomic evidence that Fort Stanton Cave provides a biological niche for cave-adapted oligotrophic/chemolithotrophic bacterial and archaeal groups and that the microbial community composition is influenced by depth below the surface and rock geochemistry.

Introduction

Culture-independent methods have been used over the last two decades to investigate microbial communities residing in caves and other subsurface environments (Engel, 2010). Characteristics of these environments include the absence of light within the dark zones, often extreme low-nutrient conditions, diverse host-rock geochemistry, constant temperatures and high relative humidities (95-100%) (Lee *et al*., 2012). Studies from these systems have revealed that novel bacterial and archaeal groups are found in these aphotic chemolithotrophic-driven systems (Northup and Lavoie 2001; Engel, 2010; Lee *et al*., 2012; Tetu *et al*., 2013; Ortiz *et al*., 2014). Bacterial and archaeal community composition is related to depth below the surface, which influences the influx of nutrients and host-rock substrate (Barton *et al*., 2007). A previous investigation of a deep arid-land cave provided evidence for hydrological seclusion from the surface (Levy, 2007), suggesting that at least some microbial groups would be expected to be reliant on metabolic constituents present within the cave and not from the influx of meteoric waters from the vadose zone.

Two recent studies have shown that there are bacterial groups and corresponding operational taxonomic units (OTUs) found in the subsurface that are not present in overlying surface soils (Ortiz *et al*., 2013; Lavoie *et al*., 2017). The Ortiz *et al*. (2013) investigation of a semiarid carbonate cave (Kartchner Caverns, AZ, USA) identified and compared bacterial communities present on cave speleothems, nearby rockwall samples, and surface soils. With respect to identified OTUs, the authors report only a 16% overlap between cave and surface samples. Correspondingly, there were dissimilar phylogenetic and taxonomic compositions between these environments. The authors hypothesized that

surface bacteria could have entered the cave through drip waters and adapted independently of their surface counterparts. The Lavoie *et al*. (2017) study compared bacterial communities present in visible microbial mats of different colors found on the walls and ceilings of a shallow lava cave (Lava Beds National Monument, CA, USA) and overlying surface soils. The study reported that shared OTUs between the cave and surface samples was only 11.2 %, hypothesizing that these subterranean environments select bacterial groups that can tolerate or adapt to the conditions found in the lava caves.

These previous research efforts have provided insight into bacterial dissimilarities associated with bacterial communities in caves versus their surface soil counterparts, and these studies have shown that novel bacterial and archaeal groups exist in the subsurface. However, previous studies have primarily focused on OTU-assigned bacterial taxonomy (Engel, 2010); and, until recently, there has been less emphasis on archaeal groups and associated taxonomy at and below the phylum level (Tetu *et al*., 2013, Ortiz *et al*., 2014, Reitschuler *et al.*, 2015; Reitschuler *et al.*, 2016; Zhao *et al*., 2017). The objective of our study was to characterize, compare, and contrast archaeal and bacterial communities present in surface samples found above sampling locations in Snowy River, a moderately deep (31-170 m) and long (17.8 km) carbonate cave passage found in central New Mexico. We hypothesized that the microbial communities residing in surface soils would differ from their subsurface counterparts found throughout the passage on the walls and ceilings in soil-like material (*speleosol*) that is rich in Fe- and Mn-oxides (Spilde *et al*., 2006; Spilde *et al*., 2009). Also referred to as ferromanganese deposits (FMDs) (Spilde *et al*., 2005; Spilde *et al*., 2006), these secondary mineral deposits represent an oligotrophic

environment that would likely select for organisms with metabolisms that favor local low-nutrient conditions.

Methods

Sampling Sites

Samples for the project were collected from Fort Stanton Cave (FSC) in the Fort Stanton - Snowy River National Conservation Area managed by the Bureau of Land Management. Located in the south-central region of New Mexico, FSC is the 15th longest cave in the USA, and the 63nd longest cave in the world at >50 km, and the cave is 134.4 m deep at the deepest point (http://www.caverbob.com/wlong.htm). The cave formed in the San Andreas formation in the lower Rio Bonito member (Kelley, 1971) by epigenic processes (i.e. downward waters) (Davis and Land, 2006). Access to FSC and all sample collection was allowed by a US Bureau of Land Management collecting permit (Northup, Principal Investigator). Surface and cave sampling were done using aseptic techniques to minimize the risk of human contamination. Sucrose lysis buffer, which aids in rupturing cells and acts as a genomic preservative (Giovannoni *et al*., 1990), was added immediately to all samples. Surface samples were collected in conifer forest and grassland habitats from the top 1 cm of soil directly above the corresponding survey stations in Snowy River passage that had previously been sampled (Table 1). A sterile implement was used to remove plant material prior to collection with sterile spoons. A majority of the cave samples were collected from *speleosol* (FMD) (Spilde *et al*., 2009), which is a soil-like material that forms on the ceilings and walls throughout the Snowy River passage, overtop of a mud layer. Approximately 10-20 cc of the upper

layer of *speleosol* was collect for each sample. Additionally, a sample was collected at Turtle Junction (Snowy River passage access point), by aseptically collecting small pieces of broken calcite from the upper portion of the calcium-lined channel that had visible signs of FMD residue and mud. All samples were stored at -80°C at the Northup laboratory at the University of New Mexico until extracted and sequenced.

Molecular Processing of Samples

Samples were extracted for genomic material according to the manufacturer's protocol using the MoBio PowerSoil™ DNA Isolation Kit (MoBio, Carlsbad, CA). Sample extractions consisted of three replicates and a negative control. In an effort to maximize DNA yield for sequencing, a bead beater was used to lyse cells, and samples were eluted with 40 µl of 10 mM Tris buffer, instead of the recommended 100 µl. Each sample extraction included a negative control and three replicates. A total of 15 surface soil samples and 22 cave samples was sent to Molecular Research LP (MR DNA) (http://www.mrdnalab.com/), Shallowater, TX, for sequencing.

Sequencing

The Illumina MiSeq platform (www.illumina.com) was used to render amplicon archaeal data by means of targeting the 16S rRNA gene using the primers 349-forward (5'-GYGCASCAGKCGMGAAW-3') and 806-reverse (5'-GGACTACVSGGGTATCTAAT-3') (Takai and Horikoshi, 2000). Bacterial samples were assayed using Illumina MiSeq to target the 16S rRNA gene using universal bacterial primers 46 forward (5'-GCYTAAYACATGCAAGTCG-3') and 1409 reverse (5'-GTGACGGGCRGTGTGTRCAA-3') (Northup *et al*., 2010). With the barcode on the

forward primer, HotStarTaq Plus Master Mix (Qiagen, USA) was used with a 28 cycle PCR (5 cycle used on PCR products) under the following specifications: 94°C for 3 m, followed by 28 cycles of 94°C for 30 s, 53°C for 40 s, 72°C for 1 m, with a final elongation step at 72°C for 5 m. Subsequently, amplification success and band intensity were determined by checking PCR products in a 2% agarose gel. Samples of similar molecular weights and DNA concentrations were pooled in equal amounts and purified using calibrated Ampure XP beads. DNA libraries were prepared from the pooled and purified PCR product based on the Illumina TruSeq DNA library preparation protocol. Sequencing was done at Molecular Research (MR DNA) (http://www.mrdnalab.com/), Shallowater, TX, on a MiSeq platform according to the manufacturer's guidelines.

16S rRNA Gene Taxonomic Analyses of Cave and Surface Samples

Archaeal and bacterial reads were annotated by means of a Quantitative Insights Into Microbial Ecology (QIIME) (Caporaso *et al*., 2010) (qiime.org) pipeline against the SILVA release 128 database (http://www.arb-silva.de/download/archive/qiime/) (Quast *et al*., 2013). The initial step in the pipeline involved binning reads according to each sample using the *split_libraries.py* script with the input mapping and fasta files. This script also involves a series of filtering steps with input arguments specifying direction to the qual file, a minimum sequence length of 100 and maximum of 600 nucleotides, 0 ambiguous bases, and minimum average quality score of 30. Using USEARCH (v6.1.544) (http://www.drive5.com/usearch/) (Edgar, 2010) as the detection method against the "Gold" reference sequence database (http://drive5.com/uchime/gold.fa), chimeric sequences were identified in the output *seqs.fna* files. To prevent chimeric sequences clustering with quality sequences, the script *filter_fasta.py* was used to remove

chimeras from the *seqs.fna* files prior to clustering. The output bacterial and archaeal *seqs_chimeras_filtered.fna* files were concatenated separately and run independently through the QIIME *pick_de_novo_otus.py* pipeline (http://qiime.org/scripts/pick_de_novo_otus.html), with the path to a parameter file specifying that similar sequences were clustered and assigned to OTUs by SUMACLUST (Mercier *et al*., 2013). Taxonomy was assigned to the output *seqs_rep_set.fasta* sequences using the SILVA_128_release (https://www.arb-silva.de/download/archive/qiime/) with arguments specifying the taxonomy file *consensus_taxonomy_7_levels.txt* and the reference sequence file *97_otus_16S.fasta*. The script *make_otu_table.py* was used to generate an OTU biom file from the SUMACLUST picked OTU map file *seqs_otus.txt* and an argument specifying the path to the SILVA_128 taxonomy assignment file *seqs_rep_set_tax_assignments.txt*. Using the *summarize_taxa.py script*, taxa absolute abundances were summarized from the archaeal and bacterial *de_novo_sumaclust.biom* files. Lastly, the *biom convert* command (http://biom-format.org/documentation/biom_conversion.html) was used to convert the *de_novo_sumaclust.biom* to an R (R Core Team, 2016) readable JSON biom format.

Bacterial and archaeal sequencing reads were isolated from the *de_novo_sumaclust.biom* files using the QIIME script *filter_taxa_from_otu_table.py*. The *biom convert* command (http://biom-format.org/documentation/biom_conversion.html) was used to convert the archaeal and bacterial *de_novo_sumaclust.biom* files to R (R Core Team, 2016) readable JSON biom format (*fixed_filtered.biom*). Sequencing data was analyzed and visualized using the R (https://www.r-project.org/) packages phyloseq (McMurdie and Holmes, 2013) and ggplot2 (Wickham, 2009). The *fixed_filtered.biom*

and archaeal and bacterial mapping files were used to create alpha diversity box plots (observed OTUs, Shannon and Chao1 indices). A custom script (https://github.com/joey711/phyloseq/issues/143) was used to make rarefaction curves plotted against depth. Archaeal and bacterial *fixed_filtered.biom* files were normalized using the QIIME *normalize_table.py* script using the cumulative-sum scaling (CSS) (Paulson *et al*., 2013) algorithm. The output files were used to generate archaeal and bacterial principle coordinate analyses (PCoA) plots with Brays-Curtis distance. Sorted taxonomic data was used to create bar charts using R with ggplot2. Additionally, the *filter_taxa_from_otu_table.py* script was used to isolate reads from the *de_novo_sumaclust.biom* files associated with the top five archaeal and 24 top bacterial phyla. These files were used to identify differentially abundant OTUs and corresponding taxonomy between cave and surface samples using the DESeq2 package (Love *et al*., 2014) with phyloseq (McMurdie and Holmes, 2014). DESeq2 parameters fitType = "local" and a p-value of 0.1 were used to determine log2-fold-change between cave and surface samples.

The QIIME *compute_core_microbiome.py* script was used to identify core microbiomes amongst and between cave and surface samples. The core microbiome amongst each biome and domain (e.g. *Archaea* cave only) required the *split_otu_table.py* QIIME script to isolate cave and surface reads from the archaeal and bacterial fixed*_filtered.biom* JSON formatted files. Each analysis was run independently (e.g. archaeal surface and cave) and core microbiome taxonomy related to shared OTUs found in not less than 80% of all samples was summarized from each core biome file using the *summarize_taxa.py* script.

Results

Richness and Multidimensional Scaling

Alpha diversity measurements of archaeal and bacterial differences in richness (Observed and Chao1 estimator) and abundance and evenness (Shannon index) between cave and surface samples are illustrated in Fig. 2. Box plot mediums show overall archaeal richness was greater in cave samples. The archaeal Shannon index shows that there is greater richness and corresponding relative abundances in the cave FMDs than in surface samples. Opposite relationships were observed for bacterial alpha diversity measurements, showing a much higher richness in surface soils and a more even distribution of both cave and surface OTUs (Fig. 2). Estimated sample-based rarefaction richness, plotted against depth, is included in Supplementary Fig. 1. PCoA plots (Fig. 3) indicate that archaeal and bacterial OTUs present in surface soils are tightly clustered. Surface and cave samples are dissimilar for both bacterial and archaeal-related OTUs. Bacterial and archaeal cave OTUs in black FMDs showed the most divergence.

SILVA v128 Assigned Domain

Archaeal-intended primers also used amplified 16S rDNA bacterial reads; although, amplification of bacterial reads was more prevalent in the surface soil samples. Bacterial reads in surface samples accounted for an average of 81.8% of the dataset. Archaeal reads in these samples comprised an average of 16.3% of the dataset, with unassigned reads representing an average of 1.9%. In cave samples, archaeal reads represented an average of 78.2% of the dataset; bacterial and unassigned reads accounted for an average of 8.5% and 13.3%, respectively. The mean number of archaeal reads

identified in surface samples was 17,634 in contrast to the 112,209 mean number of reads identified in cave samples.

In the bacterial dataset, only two reads were classified as archaeal. Unassigned reads in the cave accounted for an average of 6.2% of the dataset, while average unassigned surface reads were 3.9%. The average mean number of archaeal reads found across surface samples was 74,471, which was similar to the 70,376 mean number of reads found in cave samples.

Archaeal 16S rRNA Gene Taxonomic Composition of Snowy River FMD

Annotated archaeal reads were taxonomically dissimilar from their domain counterparts in overlying surface soils. Surface samples were dominated by the *Thaumarchaeota* class Soil Crenarchaeotic Group (SCG) (Fig. 4), which accounted for an average of nearly 95% of all archaeal reads in soil samples. When averaged, nearly half of all SCG reads in surface samples were identified as '*Candidatus* Nitrososphaera'. Cave SCG accounted for an average of 0.25% of all archaeal reads, and only a single core OTU was shared with surface samples. Dominant taxonomy in cave samples included archaeal groups from the phyla *Euryarchaeota* (average of 46.6% of all archaeal reads) and *Thaumarchaeota* (average of 49.6% of all archaeal reads) (Table 2). *Thaumarchaeota* cave taxa included the dominant classes South African Gold Mine Gp 1 (SAGMCG-1), Marine Group I (MGI), and AK31, which together accounted for an average of 93.3% of all *Thaumarchaeota* reads; other classes included AK56, AK59, FHMa11 terrestrial group, Marine Benthic Group A, and the SCG (Fig. 4).

Euryarchaeota cave reads were comprised of the class *Thermoplasmata*, representing an average of 99.8% of all *Euryarchaeota* reads identified. While *Euryarchaeota* reads identified in surface samples were primarily associated with the class *Thermoplasmata*, representing an average of 97.1% of *Euryarchaeota* reads, the mean number of *Thermoplasmata* reads identified in surface samples was 303, in contrast to a mean of 51,645 in cave samples. There were only nine *Thermoplasmata* OTUs shared between cave and surface samples.

Minor phyla in cave reads included *Aenigmarchaeota*, *Bathyarchaeota*, and *Woesearchaeota* (DHVEG-6); although there were a marginal number of surface reads corresponding to *Aenigmarchaeota* and *Bathyarchaeota*. The phylum *Bathyarchaeota* was also underrepresented in cave samples, evident by an average relative abundance of approximately 0.1%. The phylum *Aenigmarchaeota* was present in all cave samples; although the number of reads varied from only eight in sample FSCBLC to 17,920 in sample FSCBRB. Although *Woesearchaeota* (DHVEG-6) was in *all* samples, there were no shared OTUs between surface and cave in the core microbiome (Table 2). Other cave reads related to the archaeal phyla *Diapherotrites*, Miscellaneous Euryarchaeotic Group (MEG), *Nanohaloarchaeota*, and the WSA2 were included in the other phyla category (Fig. 4). Of these phyla, only a single read associated with the MEG was found in surface samples.

Bacterial 16S rRNA Gene Taxonomic Composition of Snowy River FMD

While there were taxonomic similarities between cave and surface samples at the phylum level (e.g. *Acidobacteria* and *Proteobacteria*) (Fig. 5), differences between the

surface and cave bacterial communities were evident when comparing taxonomy at OTU and genus level (Fig. 5). For example, *Firmicutes* and *Bacteroidetes* were dominant phyla in surface soils, while *Nitrospirae* and GAL15 relative abundances were elevated in cave FMDs (Table 2; Fig. 5). Rare bacterial phyla, such as *Ignavibacteriae*, RBG-1 (Zixibacteria), and SBR1093, were represented in the cave and, except for a single hit each, absent from surface samples.

DEseq2 Results

DESeq2 (Love *et al*., 2014) with phyloseq (McMurdie and Holmes, 2014) results, summarize archaeal and bacterial OTUs that were significantly different between surface and cave samples (Fig 6-7). Archaeal results showed that only the classes SCG, *Thermoplasmata*, and uncultured euryarchaeote had OTUs that were significantly different in surface samples. The most significant differentially abundant archaeal class between the cave and surface samples was *Thermoplasmata* (Fig. 6). *Firmicutes* was the most significant differentially abundant surface bacterial phylum, in contrast to cave-identified *Nitrospirae*. There were no significantly different surface bacterial OTUs associated with the phyla GAL15, *Omnitrophica*, RBG-1 (Zixibacteria), *Latescibacteria*, SBR1093, or *Ignavibacteriae* (Fig. 7).

16S rRNA Gene Archaeal and Bacterial Core Microbiome of Snowy River FMD

Archaeal and bacterial core microbiome results, representing OTUs occurring in not less than 80% of samples, determined minimal overlap (19 shared OTUs) between cave and surface samples (Table 3). While there were shared OTUs associated with dominant archaeal groups found in the cave, the average number of surface archaeal

reads corresponding with these OTUs was minimal (Table 3). This was in contrast to cave samples, as these shared OTUs represented a majority (average of 78.4%) of *all* archaeal reads identified (Table 3; Supplementary Fig. 2). Dominant archaeal classes representing these reads included the *Thermoplasmata* (average of 44.3%), SAGMCG-1 (average of 27.2%), FHMa11 terrestrial group (average of 19.9%), and AK31 (average of 6.53%). There were only single shared OTUs associated with *Thaumarchaeota* SCG and *Bathyarchaeota* (Table 2). There were no shared OTUs between cave and surface samples associated with *Woesearchaeota* (DHVEG-6).

When samples from each environment were compared (e.g., the core archaeal community across cave samples), taxonomy was similar between samples in both biomes, representing an average of nearly 95% of *all* reads identified from both cave and surface datasets (Table 3). Correspondingly, taxonomic similarity among samples from each biome was reflected in the high numbers of archaeal reads found in top-core OTUs. Averaging the top 10 OTUs from *each* sample revealed that a disproportionate majority of shared abundance amongst samples from the surface and cave were found in these core OTUs, 96.2% and 91.5%, respectively.

There were 17 bacterial OTUs shared between cave and surface samples, representing an average of only 6.1% of *all* reads from both environments (Table 3). Shared phyla amongst these OTUs included the *Acidobacteria*, *Actinobacteria*, *Firmicutes*, *Nitrospirae*, *Proteobacteria*, and *Verrucomicrobia* (Table 2). Approximately 80% of *all* reads were found in these OTUs from both cave and surface (Table 3). Dominant phyla from the surface included *Firmicutes* (average of 37.5%), *Acidobacteria* (average of 26.5%), *Proteobacteria* (average of 10.6%), *Acidobacteria* (average of

9.54%), and *Bacteroidetes* (average of 6.95%). *Proteobacteria* (average of 27.7%), *Nitrospirae* (average of 21.4%), *Acidobacteria* (average of 18.8%), and GAL15 (average of 11.7%) represented dominant phyla found across core-cave OTUs. Average reads associated with the top 10 core OTUs from *each* sample, amongst each biome, revealed that less than 50% of reads were found across these top surface (average of 43.7%) and cave (average of 49.8%) OTUs.

Discussion

Overview

Based on our findings showing few shared core microbiome OTUs between surface and cave samples, the high number of significantly different OTUs between the biomes, and dissimilar taxonomic compositions, we hypothesize that FSC is selective to specific archaeal and bacterial groups with metabolisms that allow these organisms to subsist in low-nutrient *speleosol*. We make several arguments below pertaining to this subject with respect to identified taxonomy and potential functional adaptability. Results also suggest that depth influences microbial communities found in caves as indicted by the near absence of SCG in FSC.

16S Archaeal Taxonomic Findings

At the phylum level (Fig. 4), significant differences exist in the archaeal communities residing in cave FMDs and overlying surface soils. *Euryarchaeota* were found in both surface and cave samples, but *Euryarchaeota* represented an average of 46.6% of annotated reads in most cave samples, primarily members of the class

Thermoplasmata. Excluding several reads associated with the methanogenic archaeon *Methanomassiliicoccus* (Dridi *et al*., 2012) in subsurface samples, *Thermoplasmata* cave and surface reads could not be classified to genus level. Most reads from both biomes were assigned as AS21 and Marine Group II uncultured archaeon, and all reads were assigned within the order *Thermoplasmatales* (Reysenbach, 2001). Lacking a cell wall, these thermoacidophilic, sulfur-respiring, facultatively anaerobic archaea have been described as being both autotrophic and heterotrophic (Huber and Stetter, 2006). Recent research has provided evidence to suggest that "uncultured *Thermoplasmatales*" are in fact methanogenic archaea (Paul *et al*., 2012).

Thermoplasmata have been reported in other cave environments; however, there is limited information concerning the ecological role these organisms play in these subsurface environments. The presence of this phylum in caves has largely been reported in sulfide-rich environments (Macalady *et al*., 2007; Jones *et al*., 2012; Jones *et al*., 2014; Gulecal-Pektas and Temel, 2017), although recent studies have reported *Thermoplasmata*-related 16S rRNA sequences in moonmilk (Reitschuler *et al*., 2014; Reitschuler *et al*., 2015) and cave sediments on a dry stream bed (Zhao *et al*., 2017). Based on 16S rRNA sequence similarities, *Thermoplasmatales* OTUs within FSC are related to sequences previously identified in subsurface environments, while *Thermoplasmatales* OTUs on the surface were mostly related to those previously found in soils. Using BLAST (Altschul *et al*. 1990) to search for highly similar sequences, the top surface *Thermoplasmatales* OTU (denovo245) had a 99% identity to a sequence reported in potato field soil (HQ269014) (Angel *et al*., 2012). Another top surface OTU (denovo214), collected and sequenced from teeth (dental calculus) exhumed from several

archeological sites, had a 99% identity (LN827541) to a *Methanomassiliicoccus* sp. (Huynh *et al*., 2016). Using similar search methods, top cave FSC *Thermoplasmatales* OTUs (denovo1, denovo2, denovo4, and denovo8) were 97-99% similar to sequences previously found in Alpine Cave moonmilk (KF964420-21) (Reitschuler *et al*., 2015), cave sediment (JX436825) Zhao *et al*., 2017), Su Bentu Cave soil (KT583765) (Barett *et al*., unpublished) and Molnar Janos thermal karst cave (LN998935) (Anda *et al*., unpublished). These OTUs were 92-97% similar to a *Methanomassiliicoccus* sp. (LN827539) (Huynh *et al*., 2016).

These findings infer a potential energy metabolism of *Thermoplasmatales* in cave FMD. At 97% similarity, we can conservatively reason that many *Thermoplasmatales* reads from the cave would share an OTU with a *Methanomassiliicoccus* sp., as denovo2 represented as high as 21% of *all* archaeal reads in an FMD (FSCBLD) sample. We hypothesize that some of the organisms classified as *Thermoplasmatales* in cave FMD derive energy through methanogenesis. A previous culture study that did complete genome sequencing found that "*Methanomassiliicoccus luminyensis*" was a methanogenic archaeon that reduced methanol in anaerobic conditions, using H_2 as an electron donor (Dridi *et al*., 2012). Found within the seventh order (*Methanoplasmatales*) of methanogens (Paul *et al*., 2012), genomic data infer that *Methanomassiliicoccus* are able to use many different types of methylated compounds (Borrel *et al*., 2013). However, sources within the FMDs to explain the high relative abundances of *Thermoplasmatales* in cave samples remain elusive.

Thaumarchaeota (Brochier-Armanet *et al*., 2008), a deep-branching ammonia-oxidizing archaea (AOA) phylum (Pester *et al*., 2011), represented most taxa and related

reads on the surface. Similar to findings reported in a recent study that investigated archaeal groups present in field and agriculture sites (Lu *et al*., 2017), *Thaumarchaeota* class Soil Crenarchaeotic Group (SCG) dominated overlying FSC surface soils. Accordingly, the top SCG OTU had a 99% identity to a sequence reported in arid-land soil crusts (EU423002) (Soule *et al*., 2009) and a 96% identity to '*Candidatus* Nitrososphaera gargensis' (CP002408) (Spang *et al*., 2012). In contrast, cave-annotated *Thaumarchaeota* reads were primarily dominated by archaeal groups previously reported in the deep subsurface (SAGMCG-1) (Takai *et al*., 2001) and marine environments (Marine Group I) (Massana *et al*., 2000). Lesser groups included *Thaumarchaeota* classes AK31, AK56, FHMa11 terrestrial group, and the surface-dominated SCG.

Thaumarchaeota have previously been reported in cave and cave-like environments. These studies include an investigation of a mine adit that identified clones related to SAGMCG-1 and Marine Group I (Group I.1a) (Spear *et al*., 2007), clones related to SCG (Group I.1b) in a subsurface thermal spring (Weidler *et al*., 2008), and *Thaumarchaeota* Group I.1c clones in sandstone Roraima Sur Cave (RSC), Venezuela (Barton *et al*., 2014). As recently identified in moonmilk deposits occurring in Austrian Alps caves (Reitschuler *et al.*, 2015; Reitschuler, *et al*., 2016), metagenomic analyses suggest that *Thaumarchaeota* is the dominant archaeal phylum present in some caves (Tetu *et al*., 2013; Ortiz *et al*., 2014). Related to all known AOA (Pester *et al*, 2011), Zhao *et al*. (2017) suggested that this phylum was the dominant ammonia-oxidizer (AO) in Heshang Cave, central China, karst sediments (Zhao *et al*., 2017). Similarly, our previous shotgun metagenomic and 16S rRNA sequencing study (Kimble *et al*., unpublished) hypothesized that *Thaumarchaeota* are key nitrifiers in FSC Snowy River

passage FMDs, predicated on functional gene and taxonomic findings. Empirical evidence indicates low substrate adaptability (Martens-Habbena *et al*. 2009; Walker *et al*., 2010; Pelve *et al*., 2011; Hatzenpichler *et al*., 2012; Könneke *et al*., 2014) of *Thaumarchaeota*; FMDs would provide a suitable biological niche. High specific affinity and ability to respond to low concentrations of substrate (Pester *et al*., 2011) would allow archaeal groups in cave FMD to compete for scarce fixed nitrogen, potentially becoming available from fixed N in the bedrock (Holloway and Dahlgren, 2002). Minimal OTU overlap and dissimilarities among identified *Thaumarchaeota* taxonomic groups are likely due to adaptations to different ammonium concentrations (Daebeler *et al*., 2014).

16S Bacterial Taxonomic Findings

The small number of shared OTUs, limited number of reads corresponding to these shared OTUs, and many discernible differences at lower taxonomic levels (e.g., at the genus level), provide evidence that the bacterial communities in each biome differ substantially. With respect to bacterial phyla on the surface, our results are similar to dominant taxa previously reported in soils, such as *Acidobacteria*, *Bacteroidetes*, *Firmicutes*, *Actinobacteria*, and *Proteobacteria* (Fierer *et al*., 2007), and related subphyla (e.g., *Alphaproteobacteria* and *Bacilli*) (Janssen, 2006). *Firmicutes* identified in surface soils had the highest relative abundance of all phyla from both biomes. Previously reported to be the dominant phyla in several land-use type soils (Kuramae *et al*., 2012), this phylum was underrepresented in all but two cave samples (FSCBLB and FSCBRF) and only accounted for a single shared OTU between surface and cave. *Bacteroidetes* and *Planctomycetes* were also found to have high relative abundances in the surface soils;

however, there were no shared OTUs found in the core microbiome that compared surface to cave samples.

Dominant bacterial phyla associated with cave FMDs were largely analogous to those reported in previous cave studies. For example, Spider Cave, in Carlsbad Caverns National Park, New Mexico, is a carbonate, oligotrophic arid-land cave that receives modest nutrient input from the surface. Within Spider Cave FMDs, similar to those found in FSC (Spilde *et al*., 2005), clones related to *Actinobacteria*, *Bacteroidetes*, *Chloroflexi*, *Nitrospirae*, and *Proteobacteria* comprised most of the relative bacterial abundances (Dichosa, 2008). These phyla, and others found in our study (e.g. *Acidobacteria*), have been reported in many karstic (Northup *et al*., 2003; Portillo *et al*., 2008; Pašić *et al*., 2010; Porca *et al*., 2012; Ortiz *et al*. 2013; Ortiz *et al*. 2014; Tomczyk-Żak and Zielenkiewicz, 2016) and lava caves (Moya *et al*., 2009; Northup *et* al., 2011; Hathaway *et al*., 2014; Lavoie *et al*., 2017). However, while these similar bacterial phyla are described in many cave environments, relative abundances described in caves can vary substantially. In RSC, an orthoquartzite (sandstone) cave, ceiling and wall samples were overwhelmingly composed of *Chloroflexi*, while *Actinobacteria* was the dominant phylum described on the surface of several karst speleothems (Ortiz *et al*., 2013) and basalt lava cave (Lavoie *et al*., 2017) samples. Many of these taxonomic differences are described in a review that downloaded and analyzed cave-related 16S rRNA gene sequences from the NCBI GenBank database (Brannen-Donnelly, 2015). For instance, *Gammaproteobacteria* reads represented the majority of abundances in a dolomite cave, while *Alphaproteobacteria* was the dominant reported class in sulfur caves.

These findings highlight fundamental differences in bacterial compositions found in caves of dissimilar geochemistry and cave types. In FSC, an observable dissimilarity to prior cave studies was the high relative abundances of the candidate phylum GAL15 in FMDs (Table 2; Fig. 5), discussed below. We hypothesize that connectivity (e.g. depth below the surface) is one of the primary factors that impacts bacterial composition in caves. Precipitation events on the surface would be expected to cause an influx of microbes and nutrient-rich waters (i.e., organic C and fixed N). Accordingly, it could be hypothesized that the number of shared OTUs and taxonomic similarities to surface microbial communities in shallow caves would increase. Using interchangeable methodologies to this study, we found that shared OTUs, relating to the core bacterial microbiome between surface and cave samples from shallow lava caves from Lava Beds National Monument (Lavoie *et al*., 2017) represented an average of 64.1% and 83.4% of *all bacterial* reads, respectively. Although their study (Lavoie *et al*., 2017) found little overlap between cave and surface substrates, it appears that a majority of reads correspond to shared OTUs. These results support our hypothesis that depth influences microbial composition in caves.

16S rRNA Rare Archaeal and Bacterial Groups

DESeq2 results (Fig. 7) associated with rare archaeal and bacterial phyla (e.g. *Aenigmarchaeota* and *Omnitrophica*, respectively) in the cave, and the limited number of surface reads associated with these phyla, suggest that some archaeal and bacterial groups exist in the cave and not in overlying soils. Referred to as 'microbial dark matter' (Marcy *et al*., 2007), partial and complete assembled genomes from metagenomics and single-cell genomics sequences (Albertsen *et al*., 2013; Rinke *et al*., 2013; Hug *et al*., 2016b)

have provided functional predictions for these primarily uncultured microorganisms (Solden *et al*., 2016). For example, genome reconstruction of the archaeal phylum *Aenigmarchaeota*, which is a member of the monophyletic superphylum DPANN (*Diapherotrites*, *Parvarchaeota*, *Aenigmarchaeota*, *Nanoarchaeota*, and *Nanohaloarchaea*) (Rinke *et al*., 2013; Hug *et al*., 2016a), offers evidence that members of this phylum use membrane-bound pyrophosphatases to pump protons for ATP generation and that there are genes for carbon fixation (RuBisCO) (Castelle *et al*., 2015). Also a member of the DPANN, *Woesearchaeota* (Castelle *et al*., 2015) (formerly *Euryarchaeota* DHVEG-6 cluster) was represented in *all* surface and cave samples. Following the trend of other identified phyla in our study, the top *Woesearchaeota* surface OTU had a 97% identity to a sequence previously identified in surface soil (FR865250) (Gan *et al*., 2012), while the top *Woesearchaeota* cave OTU had a 98% identity to sequences reported on the surface of a cave speleothem (GQ925754 and GQ925759) (Legatzki et al., 2011). Genome reconstruction suggests a potential "symbiotic or parasitic lifestyle" due to incomplete or missing core biosynthetic pathways, such as the ability to synthesize nucleotides and amino acids; suggesting it is likely an auxotroph (Castelle *et al*., 2015). Additional genomic research has indicated that *Woesearchaeota* ferment carbohydrates as part of their metabolism (Lazar *et al*., 2017).

Omnitrophica (formerly candidate division OP3), a member of the *Planctomycetes-Verrucomicrobia-Chlamydiae* (PVC) superphylum (Wagner and Horn, 2006) that have previously been found in anoxic aquatic environments (Glockner *et al*., 2010) and a limestone cave (Leuko *et al*., 2017), is predicted to use the reductive acetyl-CoA cycle and iron-only hydrogenases (hydrogen producing) (Rinke *et al*., 2013). Based

on genome and metabolic reconstruction of the deep-branching monophyletic outgroup from the *Bacteroidetes-Chlorobi* superphylum, RBG-1 (*Zixibacteria*) potentially uses several organic compounds as energy and carbon sources and was found to have genes encoding an "oxidative tricarboxylic acid (TCA) cycle, near-complete glycolysis/gluconeogenesis pathways, and an oxidative phosphorylation pathway" (Castelle *et al*., 2013). In cave FMD samples, the relative abundance of RBG-1 (*Zixibacteria*) was as high as 5.7% (FSCYLB). This may be attributed to nitrite/nitrate oxidoreductase (NXR) complexes (nitrate reductase function) that potentially oxidizes iron present in the bedrock, through a predicted nitrate-dependent nitrite oxidation pathway, under anoxic conditions and with an available organic carbon electron donor (Castelle *et al*., 2013).

Latescibacteria (formerly candidate division WS3), a member of the *Fibrobacteres-Chlorobi-Bacteroidetes* (FCB) superphylum (Rinke *et al*., 2013), has been previously reported in other cave studies (Lee *et al*., 2012; Cloutier *et al*., 2017; Gulecal-Pektas and Temel, 2017; Lavoie *et al*., 2017). Genome reconstruction has provided evidence that this phylum has "extensive transport systems for sugars" and associated pathways for metabolizing these sugars (Youssef *et al*., 2015), likely using a "saprophytic strategy" that would allow members in this phylum to degrade dead microbes in FMD (Farag *et al*., 2017). Reported in several cave studies (Porca *et al*., 2012; Ortiz *et al*., 2013; Brannen-Donnelly, 2015; De Mandal *et al*., 2017; Lavoie *et al*., 2017), a draft SBR1093 genome showed evidence for carbon fixation and the potential use of mixotrophic metabolism (Wang *et al*., 2014). Previously found in cave environments (Gulecal-Pektas and Temel, 2017), genome analyses of a *Ignavibacteriae* cultured

representative found that this *Chlorobi* related phylum is chemoorganoheterotrophic (organic carbon used as carbon and energy sources) (Podosokorskaya *et al*., 2013). These types of genomic reconstruction studies are shining light on metabolism of uncultured organisms, allowing us to infer how microbes in extreme environments subsist.

GAL15 represented a large proportion of relative abundances in many of the cave FMD samples; however, their role in FSC is poorly understood. There are no cultured representatives, and we could not find any studies that made functional predictions. We do know that GAL15 has been found in other caves (Costello *et al*., 2009; Brannen-Donnelly, 2015; Cloutier *et al*., 2017; De Mandal *et al*., 2017; Lavoie *et al*., 2017; Oliveira *et al*., 2017); however, reported relative abundances were much lower than our findings in most of our cave samples. In FMD samples, relative abundances ranged from 0.75% (FSCBLH), to 30.5% (FSCBLA). A subsurface study by Lin *et al*. (2012) that investigated archaeal and bacterial groups occurring in sediments at the Hanford Site found an average GAL15 relative abundance of 7.9% (9-18 m depths) in the oxic Ringold Formation and as high as 37% between 15.2 and 18.1 m, which was near the reduced zone (anoxic) (>18.3 m depth). While in deeper reduced sediments (18-52 m) and shallow oxic sediments (2-6 m), GAL15 relative abundances dropped sharply (Lin *et al*., 2012). These findings suggest that GAL15 may occupy a specific niche in the FMDs with respect to oxygen requirements, potentially in FMD microaerophilic zones.

Nitrifying Bacterial Groups

The *Nitrospirae* were the second-most abundant phylum found in cave FMD samples (Table 2; Fig. 5), and many reads were classified as belonging to the nitrite-

oxidizing (NO) genus *Nitrospira*. The high proportion of bacterial reads associated with this phylum suggests that these organisms occupy a niche in the FMD that provides them with a source of available nitrite (NO_2^-) required to carry out the second step of nitrification, aerobic nitrite oxidation. Nitrite is an intermediate of nitrification (Cleemput and Samater, 1996); therefore, the presumed source of nitrite in the FMD would be a result of bacterial or archaeal groups associated with this N cycling pathway. We hypothesize that the primary source of nitrite in FMDs is the result of aerobic ammonia oxidation by the phylum *Thaumarchaeota*. Their high abundance in the FMD samples and adaptability to oligotrophic conditions discussed above suggest AOA are important drivers of the first step of the nitrification pathway and are potentially responsible for the high abundance of *Nitrospirae* in FSC FMDs and other cave (Pašić *et al*., 2010; Tetu *et al*., 2013; Ivanova *et al*., 2013; Ortiz *et al*., 2014; Lavoie *et al*., 2017; Oliveira *et al*., 2017) and subsurface environments (Lin *et al*., 2010; Rempfert *et al*., 2017). Similar to spatial distributions observed with ammonia oxidizing bacteria (AOB) and nitrite oxidizing bacteria (NOB) (Maixner, 2009), we hypothesize that NOB are closely associated with AOA in FMD, as these organisms would provide a source of nitrite, which is generally scarce in the cave. Evidence for nitrogen scarcity in FMDs is provided by two samples collected from survey MK-100 with FSCBLH that had an average total nitrogen percent of 0.29 (Lewis, unpublished).

On the surface, the relative low abundance of *Nitrospirae* may be attributed to conditions in soils less favorable to AOA. Consider that more than three times as many mean number of *Thaumarchaeota* reads are present in the cave than in overlying surface surface soils. A reduction in AOA and increased substrate competition in soils could

arguably restrict NOB populations. Lavoie *et al*.'s (2017) recent investigation of bacterial communities in a shallow lava cave and overlying surface soils reported a relative abundance decrease of *Nitrospirae* in surface (3%) versus cave (7%) samples. This change was even more dramatic in FSC, which showed a reduction in *Nitrospirae* from 18.6% relative abundance in cave samples to 0.5% on the surface. This trend is explained if the ratio of AOA to NOB is considered. Normalizing (Paulson *et al*., 2013) bacterial and archaeal sequencing data using QIIME (Caporaso *et al*., 2010), *Thaumarchaeota* to *Nitrospirae* average sample ratio of 7:3 is observed in the cave versus a 35:1 ratio corresponding to surface samples. These results support the hypothesis that these chemolithotrophic organisms prefer oligotrophic conditions as K-strategists (Nogueira and Melo, 2006; Kim and Kim, 2006), and that these microbes are slow growing with low nutrient adaptations.

Lastly, reads related to the AOB in the order *Nitrosomonadaceae* (Prosser and Stein, 2014) were identified in *all* samples but with lower relative abundances found in surface samples. With respect to driving nitrification in FMD, we hypothesize they play a lesser role than their AOA counterparts. Due to the fact that AOB have a lesser substrate affinity and require higher substrate concentrations than AOA (Martens-Habbena *et al*. 2009), we previously hypothesized (Kimble *et al*., unpublished) that AOA outcompete AOB by scavenging and oxidizing ammonium (NH_4^+) that becomes available in the FMDs. Support for our hypothesis is reflected in the relative low relative abundances of *Nitrosomonadaceae* in our FMD samples. However, future studies should use quantitative PCR methods to investigate corresponding AO gene abundances (Leininger *et al*., 2006).

Conclusions

Our investigation used a 16S rRNA gene next-generation sequencing pipeline to address specific study objectives, including characterizing, comparing, and contrasting archaeal and bacterial communities existing in cave FMD and overlying surface soils. Our findings provide evidence that the bacterial and archaeal microbial communities found in Snowy River passage FMDs are in many ways taxonomically dissimilar to the microbial communities residing in overlying surface soils. Several dominant archaeal and bacterial phyla found in cave FMD are associated with taxa previously found in low-nutrient environments, which suggests metabolic adaptability specific to the low-nutrient FMDs. The limited number of shared archaeal and bacterial OTUs, and reads associated with dominant surface phyla (e.g. *Firmicutes*), suggests that meteoric waters from the vadose zone likely have limited influence on the microbial communities present in cave FMDs. Depth below the surface, which influences the influx of nutrients and host-rock substrate (Barton *et al*., 2007), likely contributes to the bacterial and archaeal community composition in the speleosols of Snowy River passage . Future investigations should address diverse caves types of dissimilar geochemistry to better understand factors that influence microbial cave compositions.

Acknowledgments

We would like to thank the Fort Stanton Cave Study Project (http://fscsp.org/) and the Bureau of Land Management (BLM) for project funding, sampling, and caving support. We thank Pete Lindsley for organizing and carrying out the surface sampling and John Corcoran for his mapping assistance and expertise consultation on FSC. Access to the cave would not be possible without permission from the BLM, and we thank them for their continued support of scientific expeditions into FSC. We thank the NIH

supported UNM IMSD program (grant number GM-060201) for their supportive funding of Jason C. Kimble's doctoral research. We also acknowledge all members of the Northup SLiME team for their continued support and valued input.

Footnotes

1. ^http://www.caverbob.com/wlong.htm
2. ^http://www.mrdnalab.com/
3. ^www.illumina.com
4. ^http://www.arb-silva.de/download/archive/qiime/
5. ^http://www.drive5.com/usearch/
6. ^http://drive5.com/uchime/gold.fa
7. ^http://qiime.org/scripts/pick_de_novo_otus.html
8. ^https://www.arb-silva.de/download/archive/qiime/
9. ^http://biom-format.org/documentation/biom_conversion.html

References

Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K. L., Tyson, G. W., and Nielsen, P. H. (2013). Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat. Biotechnol.* 31, 533–538. doi: 10.1038/nbt.2579

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–10. doi: 10.1016/S0022-2836(05)80360-2

Angel, R., Claus, P., and Conrad, R. (2012). Methanogenic archaea are globally ubiquitous in aerated soils and become active under wet anoxic conditions.*ISME. J.* 6, 847–862. doi: 10.1038/ismej.2011.141

Barton, H. A., Giarrizzo, J. G., Suarez, P., Robertson, C. E., Broering, M. J., and Banks, E. D. (2014). Microbial diversity in a Venezuelan orthoquartzite cave is dominated by the *Chloroflexi* (Class *Ktedonobacterales*) and *Thaumarchaeota* Group I.1c. *Front. Microbiol.* **5**:615. doi: 10.3389/fmicb.2014.00615

Barton, H. A., Taylor, N. M., Kreate, M. P., Springer, A. C., Oehrle, S. A., and Bertog, J. L. (2007). The impact of host rock geochemistry on bacterial community structure in oligotrophic cave environments. *Int. J. Speleol.* 36, 93–104. doi: 10.5038/1827-806X.36.2.5

Borrel, G., O'Toole, P. W., Harris, H. M. B., Peyret, P., Brugère, J. F., and Gribaldo, S. (2013). Phylogenomic data support a seventh order of methylotrophic methanogens and

provide insights into the evolution of methanogenesis. *Genome Biol. Evol*., 5, 1769–1780. doi: 10.1093/gbe/evt128

Brannen-donnelly, K. M. (2015). Bacterial diversity and function within an epigenic cave system and implications for other limestone cave systems. Ph.D. dissertation: University of Tennessee.

Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008). Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* 6, 245-252. doi:10.1038/nrmicro1852

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME. J.* 6, 1621-1624. doi: 10.1038/ismej.2012.8

Castelle, C. J., Hug, L. A., Wrighton, K. C., Thomas, B. C., Williams, K. H., Wu, D., et al. (2013). Extraordinary phylogenetic diversity and metabolic versatility in aquifer sediment. *Nat. Commun.* 4, 1–10. doi: 10.1038/ncomms3120

Castelle, C. J., Wrighton, K. C., Thomas, B. C., Hug, L. A., Brown, C. T., Wilkins, M. J. et al. (2015). Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr. Biol.* 25, 690–701. doi: 10.1016/j.cub.2015.01.014

Cleemput, O. V. and Samater, A. H. (1996). Nitrite in soils: accumulation and role in the formation of gaseous N compounds. *Fertilizer Research*. 45, 81-89. doi:10.1007/BF00749884

Cloutier, M. L. C., Carmichael, S. K., Carson, M. A., Madritch, M. D., and Bräuer, S. L. (2017). Carbon quantity and quality drives variation in cave microbial communities and regulates Mn(II) oxidation. *Biogeochemistry*, 134, 77–94. doi: 10.1007/s10533-017-0343-8

Costello, E. K., Halloy, S. R. P., Reed, S. C., Sowell, P., and Schmidt, S. K. (2009). Fumarole-supported islands of biodiversity within a hyperarid, high-elevation landscape on socompa volcano, Puna de Atacama, Andes. *Appl. Environ. Microbiol.* 75, 735–747. doi: 10.1128/AEM.01469-08

Davis D. D. and Land, L. (2006). Recently discovered passages in Fort Stanton Cave, New Mexico, and implications for speleogenesis and regional geomorphic processes in the northern Sacramento Mountains. *New Mexico Geological Society, Guidebook 57*, 219–226.

De Mandal, S., Chatterjee, R., & Kumar, N. S. (2017). Dominant bacterial phyla in caves and their predicted functional roles in C and N cycle. *BMC Microbiol*, **17**:90. doi: 10.1186/s12866-017-1002-x

Daebeler, A., Bodelier, P. L. E., Yan, Z., Hefting, M. M., Jia, Z., and Laanbroek, H. J. (2014). Interactions between Thaumarchaea, *Nitrospira* and methanotrophs modulate autotrophic nitrification in volcanic grassland soil. *ISME. J.* 8, 2397–2410. doi: 10.1038/ismej.2014.81

Dichosa, A. (2008). Biogenicity and microbial community composition of desert varnish and cave ferromanganese deposits. Ph.D. dissertation: University of New Mexico.

Dridi, B., Fardeau, M. L., Ollivier, B., Raoult, D., and Drancourt, M. (2012). Methanomassiliicoccus luminyensis gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* 62, 1902–1907. doi: 10.1099/ijs.0.033712-0

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 26, 2460-2461. doi: 10.1093/bioinformatics/btq461

Engel, A. S. (2010). "Microbial diversity of cave ecosystems," in: *Geomicrobiology: molecular and environmental perspective*, ed. L. L. Barton et al. (Dordrecht, Springer), 219-238.

Farag, I. F., Youssef, N. H., and Elshahed, M. S. (2017) Global distribution patterns and pangenomic diversity of the candidate phylum "*Latescibacteria*" (WS3). *Appl. Environ. Microbiol.* **83**:e00521-17. doi: 10.1128/AEM .00521-17

Fierer, N., Bradford, M. A., and Jackson, R. B. (2016). Toward an ecological classification of soil bacteria. *Ecology*. 88, 1354-1364. doi: 10.1890/05-1839

Giovannoni, S. J., Delong, E. F., Schmidt, T. M., and Pace, N.R. (1990). Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. *Appl. Environ. Microbiol.* 56, 2572-2575.

Glockner, J., Kube, M., Shrestha, P. M., Weber, M., Glockner, F. O., Reinhardt, R., et al. (2010). Phylogenetic diversity and metagenomics of candidate division OP3. *Environ. Microbiol.* 12, 1218–1229. doi: 10.1111/j.1462-2920.2010.02164.x

Gulecal-Pektas, Y. and Temel, M. (2017). A window to the subsurface: microbial diversity in hot springs of a sulfidic cave (Kaklik, Turkey). *Geomicrobiol. J.* 34, 374–384. doi: 10.1080/01490451.2016.1204374

Hathaway, J. J. M., Garcia, M. G., Balasch, M. M., Spilde, M. N., Stone, F. D., Lurdes, M. D., et al. (2014). Comparison of Bacterial Diversity in Azorean and Hawai'ian Lava Cave Microbial Mats. *Geomicrobiol. J.* 31, 205-220. doi: 10.1080/01490451.2013.777491

Hatzenpichler, R. (2012). Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. *Appl. Environ. Microbiol.* 78, 7501-7510. doi: 10.1128/AEM.01960-12

Holloway, J. M. and Dahlgren, R. A. (2002). Nitrogen in rock: occurrences and biogeochemical implications. *Global Biogeochem. Cycles*. 78, 7501-7510. doi:10.1029/2002GB001862

Huber, H. and Stetter, K. O. (2006). *Thermoplasmatales*. *Prokaryotes*. 3, 101-112. doi: 10.1007/0-387-30743-5_7

Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J. et al. (2016a). A new view of the tree of life. *Nat. Microbiol.* 1, 16048. doi: 10.1038/nmicrobiol.2016.48

Hug, L. A., Thomas, B. C., Sharon, I., Brown, C. T., Sharma, R., Hettich, R. L. et al. (2016b). Critical biogeochemical functions in the subsurface are associated with bacteria from new phyla and little studied lineages. *Environ. Microbiol.* 18, 159–173. doi: 10.1111/1462-2920.12930

Huynh, H. T. T., Nkamga, V. D., Signoli, M., Tzortzis, S., Pinguet, R., Audoly, et al. (2016). Restricted diversity of dental calculus methanogens over five centuries, France. *Sci. Rep.* **6**: 25775. doi: 10.1038/srep25775

Ivanova, V., Tomova, I., Kamburov, A., Tomova, A., Vasileva-Tonkova, E., and Kambourova, M. (2014). High phylogenetic diversity of bacteria in the area of prehistoric paintings in Magura Cave, Bulgaria. *J. Caves Karst Stud.* 75, 218-228. doi: 10.4311/2012MB0279

Jones, D. S., Albrecht, H. L., Dawson, K. S., Schaperdoth, I., Freeman, K. H., Pi, Y., et al. (2012). Community genomic analysis of an extremely acidophilic sulfur-oxidizing biofilm. *ISME. J.* 6, 158-170. doi: 10.1038/ismej.2011.75

Jones, D. S., Schaperdoth, I., and Macalady, J. L. (2014). Metagenomic evidence for sulfide oxidation in extremely acidic cave biofilms. *Geomicrobiol.* 31, 194–204. doi: 10.1080/01490451.2013.834008

Janssen, P. H. (2006). Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl. Environ. Microbiol.* 72, 1719–1728. doi: 10.1128/AEM.72.3.1719-1728.2006

Kelley, V. C. (1971). Geology of the Pecos country, southeastern New Mexico. *Memoir*. 24, 1-78.

Kim, D. J., and Kim, S. H. (2006). Effect of nitrite concentration on the distribution and competition of nitrite-oxidizing bacteria in nitratation reactor systems and their kinetic characteristics. *Water Res.* 40, 887–894. doi: 10.1016/j.watres.2005.12.023

Könneke, M., Schubert, Philip, D. M., Brown, P. C., Hügler, M., Standfest, S., and Schwander, T. (2014). Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. *PNAS.* 111, 8239-8244. doi/10.1073/pnas.1402028111

Kuramae, E. E., Yergeau, E., Wong, L. C., Pijl, A. S., Van Veen, J. A., and Kowalchuk, G. A. (2012). Soil characteristics more strongly influence soil bacterial communities than land-use type. *FEMS Microbiol. Ecol.* 79, 12–24. doi: 10.1111/j.1574-6941.2011.01192.x

Lavoie, K. H., Winter, A. S., Read, K. J., Hughes, E. M., Spilde, M. N., 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**: e0169339. doi: 10.1371/journal.pone.0169339

Lazar, C. S., Baker, B. J., Seitz, K. W., and Teske, A. P. (2017). Genomic reconstruction of multiple lineages of uncultured benthic archaea suggests distinct biogeochemical roles and ecological niches. *ISME. J.* 11, 1118–1129. doi: 10.1038/ismej.2016.189

Lee, N. M., Meisinger, D. B., Aubrecht, R., Kovacik, L., Saiz-Jimenez, C., Baskar, S., et al. (2012). "Caves and karst environments," in *Life at Extremes: Environments, Organisms and Strategies for Survival*, ed. E. M. Bell (CAB International), 320-344.

Legatzki, A., Ortiz, M., Neilson, J. W., Dominguez, S., Andersen, G. L., Toomey, R. S., et al. (2011). Bacterial and archaeal community structure of two adjacent calcite speleothems in Kartchner Caverns, Arizona, USA. *Geomicrobiol. J.* 28, 99–117. doi: 10.1080/01490451003738465

Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G. W., et al. (2006). Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature*. 442, 806-809. doi:10.1038/nature04983

Leuko, S., Koskinen, K., Sanna, L., D'Angeli, I. M., De Waele, J., Marcia, P., et al. (2017). The influence of human exploration on the microbial community structure and ammonia oxidizing potential of the Su Bentu limestone cave in Sardinia, Italy. *PLoS ONE*. **12**: e0180700. doi: 10.1371/journal.pone.0180700

Lin, X., Kennedy, D., Fredrickson, J., Bjornstad, B., Konopka, A. (2012). Vertical stratification of subsurface microbial community composition across geological formations at the Hanford Site. *Environ Microbiol.* 14, 414-425. doi: 10.1111/j.1462-2920.2011.02659.x

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8

Macalady, J. L., Jones, D. S., and Lyon, E. H. (2007). Extremely acidic, pendulous cave wall biofilms from the Frasassi cave system, Italy. *Environ. Microbiol.* 9, 1402–1414. doi: 10.1111/j.1462-2920.2007.01256.x

Maixner, F. M. (2009). The ecophysiology of nitrite-oxidizing bacteria in the genus *Nitrospira*: novel aspects and unique features. Ph.D. dissertation: University of Vienna.

Marcy, Y., Ouverney, C., Bik, E. M., Losekann, T., Ivanova, N., Martin, H. G., et al. (2007). Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. *Proc. Natl. Acad. Sci. U S A.* 104, 11889–11894. doi: 10.1073/pnas.0704662104

Martens-Habbena, W., Berube, P. M., Urakawa, H., de la Torre, J. R., and Stahl, D. A. (2009). Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature*. 461, 976-979. doi: 10.1038/nature08465

Massana, R., Delong, E. F., and Pedrós-Alió, C. (2000). A few cosmopolitan phylotypes dominate planktonic archaeal assemblages in widely different oceanic provinces. *Appl. Environ. Microbiol.* 66, 1777–1787. doi: 10.1128/AEM.66.5.1777-1787.2000

McMurdie, P. J. and Holmes, S. (2014) Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput. Biol.* **10**:e1003531. doi: 10.1371/journal.pcbi.1003531

Mercier, C., Boyer, F., Bonin, A., and Coissac, E. (2013). SUMATRA and SUMACLUST: fast and exact comparison and clustering of sequences. (Version 1.0.00) [Software]. http://metabarcoding.org/sumatra

Moya, M., Garcia, M. G., Spilde, M. N., and Northup, D. E. (2009). Composition of bacterial mats in El Malpais, National Monument, New Mexico, USA: comparison and contrasts with bacterial communities in Hawai'i lava tubes. *Proc.15th Congr. Speleol.* 2, 709-713.

Nogueira, R. and Melo, L. F. (2006). Competition between Nitrospira spp. and Nitrobacter spp. in nitrite-oxidizing bioreactors. *Biotechnol. Bioeng.* 95, 169-175. doi: 10.1002/bit.21004

Northup, D. E. and Lavoie, K. H. (2001). Geomicrobiology of caves. *Geomicrobiol. J.* 18, 199-222. doi: 10.1080/01490450152467750

Northup, D. E., Barns, S. M., Yu, L. E., Spilde, M. N., Schelble, R. T., Dano, K. E. et al. (2003). Diverse microbial communities inhabiting ferromanganese deposits in

Lechuguilla and Spider Caves. *Environ. Microbiol.* 5, 1071-1086.

Northup, D. E., Snider, J. R., Spilde, M. N., Porter, M. L., Van de Kamp, J. L., Boston, P. J., et al. (2010). Diversity of rock varnish bacterial communities from Black Canyon, New Mexico. *J Geophys. Res.* 115, 1-19. doi:10.1029/2009JG001107

Oliveira, C., Gunderman, L., Coles, C. A., Lochmann, J., Parks, M., Ballard, E., et al. (2017). 16S rRNA gene-based metagenomic analysis of Ozark cave bacteria. *Diversity*, 9, 1–11. doi: 10.3390/d9030031

Ortiz, M., Legatzki, A., Nelson, J.W., Fryslie, B., Nelson, W.M., Wing, R.A., et al. (2014). Making a living while starving in the dark: metagenomic insights into the energy dynamics of a carbonate cave. *ISME. J.* 8, 478-491. doi:10.1038/ismej.2013.159

Ortiz, M., Neilson, J. W., Nelson, W. M., Legatzki, A., Byrne, A., Yu, Y., et al. (2012). Profiling bacterial diversity and taxonomic composition on speleothem surfaces in Kartchner Caverns, AZ. *Microb. Ecol.* 65, 371-383.

Pasić, L., Kovce, B., Sket, B., Herzog-Velikonja, B. (2010). Diversity of microbial communities colonizing the walls of a Karstic cave in Slovenia. *FEMS Microbiol. Ecol.* 71, 50-60. doi: 10.1111/j.1574-6941.2009.00789.x

Paul, K., Nonoh, J. O., Mikulski, L., and Brune, A. (2012). "Methanoplasmatales," thermoplasmatales-related archaea in termite guts and other environments, are the seventh order of methanogens. *Appl. Environ. Microbiol.* 78, 8245–8253. doi: 10.1128/AEM.02193-12

Pelve, E. A., Lindås, A. C., Martens-Habbena, W., de la Torre, J. R., Stahl, D. A., and Bernander, R. (2011). Cdv-based cell division and cell cycle organization in the thaumarchaeon *Nitrosopumilus maritimus*. *Mol. Microbiol.* 82, 555-566. doi: 10.1111/j.1365-2958.2011.07834

Pester, M., Schleper, C., and Wagner, M. (2011). The *Thaumarchaeota*: an emerging view of their phylogeny and ecophysiology. *Curr. Opin. Microbiol.* 14, 300-306. doi: 10.1016/j.mib.2011.04.007

Podosokorskaya, O. A., Kadnikov, V. V., Gavrilov, S. N., Mardanov, A. V., Merkel, A. Y., Karnachuk, O. V., et al. (2013). Characterization of *Melioribacter roseus* gen. nov., sp. nov., a novel facultatively anaerobic thermophilic cellulolytic bacterium from the class *Ignavibacteria*, and a proposal of a novel bacterial phylum *Ignavibacteriae*. *Environ. Microbiol.* 15, 1759–1771. https://doi.org/10.1111/1462-2920.12067

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 Microbiol. Ecol.* 81, 255–266. doi: 10.1111/j.1574-6941.2012.01383.x

Portillo, M.C., Gonzalez, J.M., and Saiz-Jimenez, C. (2008). Metabolically active microbial communities of yellow and grey colonizations on the walls of Altamira Cave, Spain. *J. Appl. Microbiol.* 104, 681-691. doi: 10.1111/j.1365-2672.2007.03594.x

Prosser J. I., Head I. M., Stein L. Y. (2014). "The family *Nitrosomonadaceae*," in *The Prokaryotes*. ed. Dworkinm M., Falkow S., Rosenberg E., Schleifer K.-H., Stackebrandt E. (Berlin: Springer), 901–918.

Quast, C., Pruesse, E., Yilmaz, J., Schweer, T., Yarza, P., Peplies, J. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**:D590-D596. doi: 10.1093/nar/gks1219

R Core Team. (2016). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. [Software]. http://www.R-project.org

Rempfert, K.R., Miller, H. M., Bompard, N., Nothaft, D., Matter, J. M., Kelemen, P. (2017). Geological and geochemical controls on subsurface microbial life in the Samail Ophiolite, Oman. *Front. Microbiol.* **7**: 56. doi: 10.3389/fmicb.2017.00056

Reysenbach, A. L. (2001). "Order I. *Thermoplasmatales* ord. nov," in *Bergey's Manual of Systematic Bacteriology*, 2nd ed. Vol. 1, ed. G. Garrity (New York: Springer-Verlag), 35.

Reitschuler, C., Lins, P., Schwarzenauer, T., Spötl, C., Wagner, A. O., and Illmer, P. (2015). New undescribed lineages of non-extremophilic archaea form a homogeneous and dominant element within alpine moonmilk microbiomes. *Geomicrobiol. J.* 32, 890–902. doi: 10.1080/01490451.2015.1025317

Reitschuler, C., Lins, P., Wagner, A. O., and Illmer, P. (2014). Cultivation of moonmilk-born non-extremophilic Thaum and *Euryarchaeota* in mixed culture. *Anaerobe*. 29, 73–79. doi: 10.1016/j.anaerobe.2013.10.002

Reitschuler, C., Spötl, C., Hofmann, K., Wagner, A. O., and Illmer, P. (2016). Archaeal distribution in moonmilk deposits from alpine caves and their ecophysiological potential. *Microb. Ecol.* 71, 686–699. doi: 10.1007/s00248-015-0727-z

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J. F., et al. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature*. 499, 431–437. doi: 10.1038/nature12352

Solden, L., Lloyd, K., and Wrighton, K. (2016). The bright side of microbial dark matter: lessons learned from the uncultivated majority. *Curr. Opin. Microbiol.* 31, 217–226. doi: 10.1016/j.mib.2016.04.020

Spang, A., Poehlein, A., Offre, P., Zumbrägel, S., Haider, S., Rychlik, et al. (2012). The genome of the ammonia-oxidizing *Candidatus* Nitrososphaera Gargensis: insights into metabolic versatility and environmental adaptations. *Environ. Microbiol.* 14, 3122–3145. doi: 10.1111/j.1462-2920.2012.02893.x

Spear, J. R., Barton, H. A., Robertson, C. E., Francis, C. A., and Pace, N. R. (2007). Microbial community biofabrics in a geothermal mine adit. *Appl. Environ. Microbiol*. 73, 6172–6180. doi:10.1128/AEM.00393-07

Spilde, M. S., Kooser, A, Boston, P. J., and Northup, D. E.(2009). "Speleosol: A subterranean soil," in Proceedings of the 15th International Congress of Speleology, ed. W. B. White (Huntsville, AL: National Speleological Society), 338-344.

Spilde, M. N., Northup, D. E., and Boston, P. J. (2006). "Ferromanganese deposits in the caves of the Guadalupe Mountains" in Caves and Karst of Southeast New Mexico, ed L. Land L, Lueth (The New Mexico Geological Society), 161–166.

Spilde, M. N., Northup, D. E., Boston, P. J., Schelble, R. T., Kathleen, E. D., Crossey, L. J., et al. (2005). Geomicrobiology of cave ferromanganese deposits: a field and laboratory investigation. *Geomicro. J.* 22, 99-116. doi: 10.1080/01490450590945889

Soule, T., Anderson, I. J., Johnson, S. L., Bates, S. T., and Garcia-Pichel, F. (2009). Archaeal populations in biological soil crusts from arid lands in North America. *Soil Biol. Biochem.* 41, 2069–2074. doi: 10.1016/j.soilbio.2009.07.023

Takai, K. and Horikoshi, K. (2000). Rapid Detection and Quantification of Members of the Archaeal community by quantitative PCR using fluorogenic probes. *Appl. Environ. Microbiol.* 66, 5066-5072. doi: 10.1128/AEM.66.11.5066-5072.2000

Takai, K., Moser, D. P., DeFlaun, M., Onstott, T. C., and Fredrickson, J. K. (2001). Archaeal diversity in waters from deep South African gold mines. *Appl. Environ. Microbiol.* 67, 5750-5760. doi: 10.1128/AEM.67.21.5750-5760.2001

Tetu, S. G., Breakwell, K., Elbourne, L. D. H., Holmes, A. J., Gillings, M. R., and Paulsen, I. T. (2013). Life in the dark: metagenomic evidence that a microbial slime community is driven by inorganic nitrogen metabolism. *ISME. J.* 7, 1227–1236. doi: 10.1038/ismej.2013

Tomczyk-Żak, K., and Zielenkiewicz, U. (2016). Microbial diversity in caves. *Geomicrobiol. J.* 33, 20–38. doi: 10.1080/01490451.2014.1003341

Wagner, M., and Horn, M. (2006). The *Planctomycetes*, *Verrucomicrobia*, *Chlamydiae* and sister phyla comprise a superphylum with biotechnological and medical relevance. *Curr. Opin. Biotechnol.* 17, 241–249. doi: 10.1016/j.copbio.2006.05.005

Walker, C. B., de la Torre, J. R., Klotz, M. G., Urakawa, H., Pinel, N., Arp, D. J., et al. (2010). *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *PNAS.* 107, 8818-8823. doi: 10.1073/pnas.0913533107

Wang, Z., Guo, F., Liu, L., and Zhang, T. (2014). Evidence of carbon fixation pathway in a bacterium from candidate phylum SBR1093 revealed with genomic analysis. *PLoS ONE.* **9**: e109571. doi: 10.1371/journal.pone.0109571

Weidler, G. W., Gerbl, F. W., and Stan-Lotter, H. (2008). Crenarchaeota and their role in the nitrogen cycle in a subsurface radioactive thermal spring in the Austrian Central Alps. *Appl. Environ. Microbiol.* 74, 5934-5942. doi: 10.1128/AEM.02602-07

Youssef, N. H., Farag, I. F., Rinke, C., Hallam, S. J., Woyke, T., and Elshahed, M. S. (2015). In silico analysis of the metabolic potential and niche specialization of candidate phylum "*Latescibacteria*" (WS3). *PLoS ONE.* **10**: e0127499. doi: 10.1371/journal.pone.0127499

Zhao, R., Wang, H., Yang, H., Yun, Y., and Barton, H. A. (2016). Ammonia oxidizing archaea dominant over bacterial ammonia oxidizers within alkaline cave sediments. *Geomicro. J.* doi: 10.1080/01490451.2016.1225861

Figures and Legends

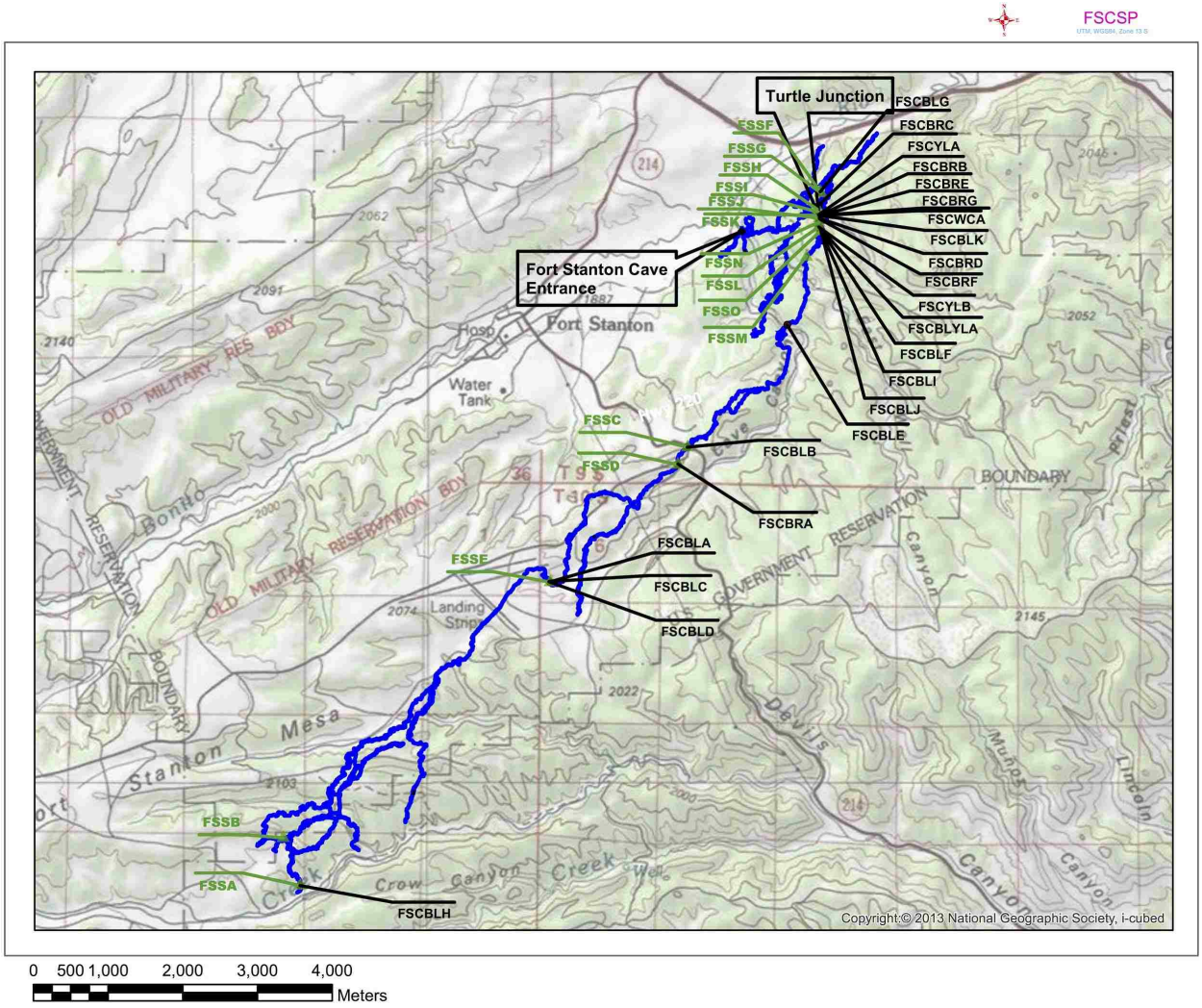


Fig. 1. Topographic map showing Snowy River passage FMD sampling locations (black) and overlying soil collection sites (green). Also indicated are the entrance to the cave and the Snowy River passage access point (Turtle Junction). (Map courtesy of John Corcoran).

Cave Sample IDs	Sample Type	Location	Depth Below Surface (m)	Distance from Turtle Junction (m)	Corresponding Surface Samples
FSCBLA	Black FMD	SRS450	170	8,761	FSSE
FSCBLB	Black FMD	SRS301	140	4,961	FSSC
FSCBLC	Black FMD	SRS450	170	8,761	FSSE
FSCBLD	Black FMD	SRS450	170	8,761	FSSE
FSCBLE	Black FMD	SRS125	69	1,891	N/A
FSCBLF	Black FMD	SRS35	31	186	N/A
FSCBLG	Black FMD	SRS1	62	374	FSSF
FSCBLH	Black FMD	MK-100	86	16,500	FSSA
FSCBLI	Black FMD	SRS38	59	234	FSSO
FSCBLJ	Black FMD	SRS38	58	234	FSSO
FSCBLK	Black FMD	SRS23	51	0	FSSJ
FSCBRA	Brown FMD	SRS315	44	5,272	FSSD
FSCBRB	Brown FMD	SRS23	51	0	FSSJ
FSCBRC	Brown FMD	SRS14	38	152	N/A
FSCBRD	Brown FMD	SRS26	50	41	FSSK
FSCBRE	Brown FMD	SRS23	51	0	FSSJ
FSCBRF	Brown FMD	SRS27	50	51	N/A
FSCBRG	Brown FMD	SRS23	51	0	FSSJ
FSCYLA	Yellow FMD	SRS22	53	14	N/A
FSCYLB	Yellow FMD	SRS34	58	172	FSSL
FSCBLYLA	Black/Yellow FMD	SRS34	58	172	FSSL
FSCWC	White Calcite	SRS23	51	0	FSSJ

Table 1. Summary of metadata associated with Snowy River passage samples and corresponding soil surface samples (where applicable).

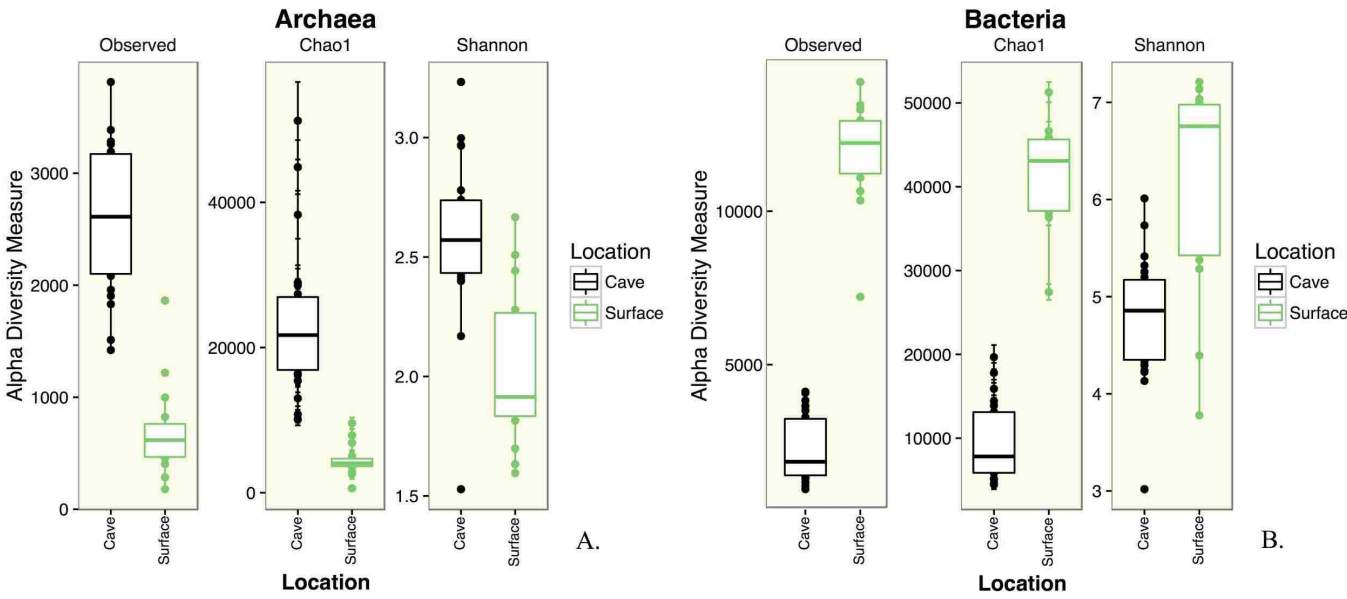


Fig. 2. Alpha diversity measurements of bacterial and archaeal groups that indicate while archaea is more diverse in the caves, the opposite relationship is observed with respect to their bacterial surface counterparts.

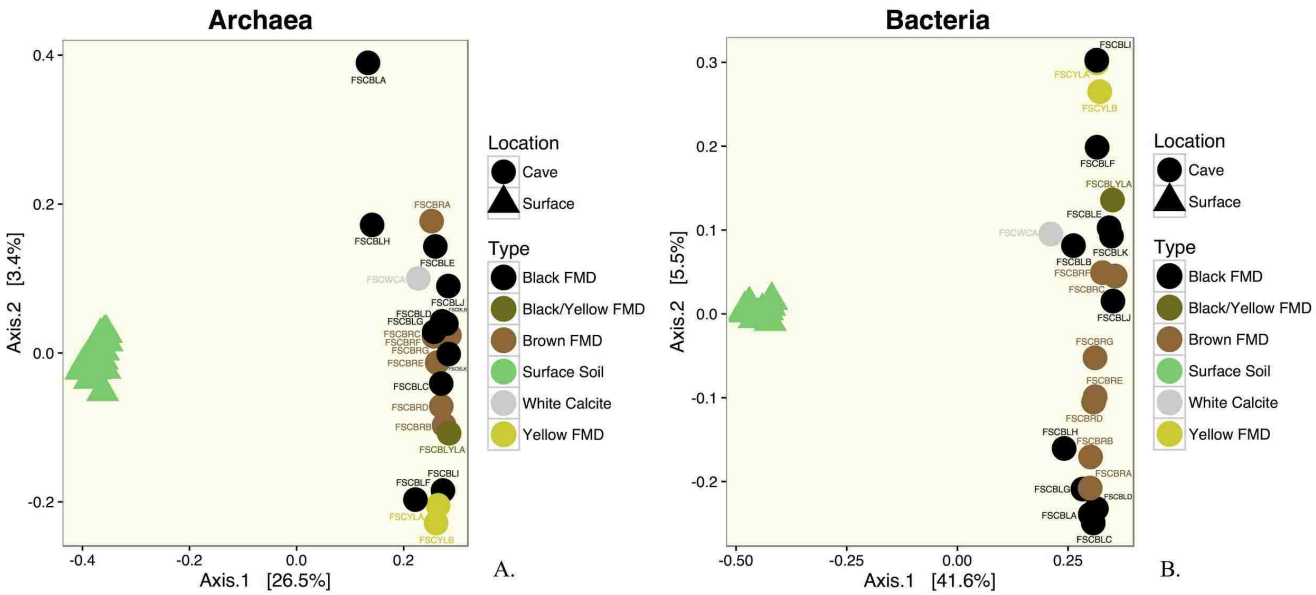


Fig. 3. Principal Coordinates Analyses plots that reveal dissimilarity between archaeal and bacterial groups in the cave and in overlying surface soils and amongst cave FMD samples.

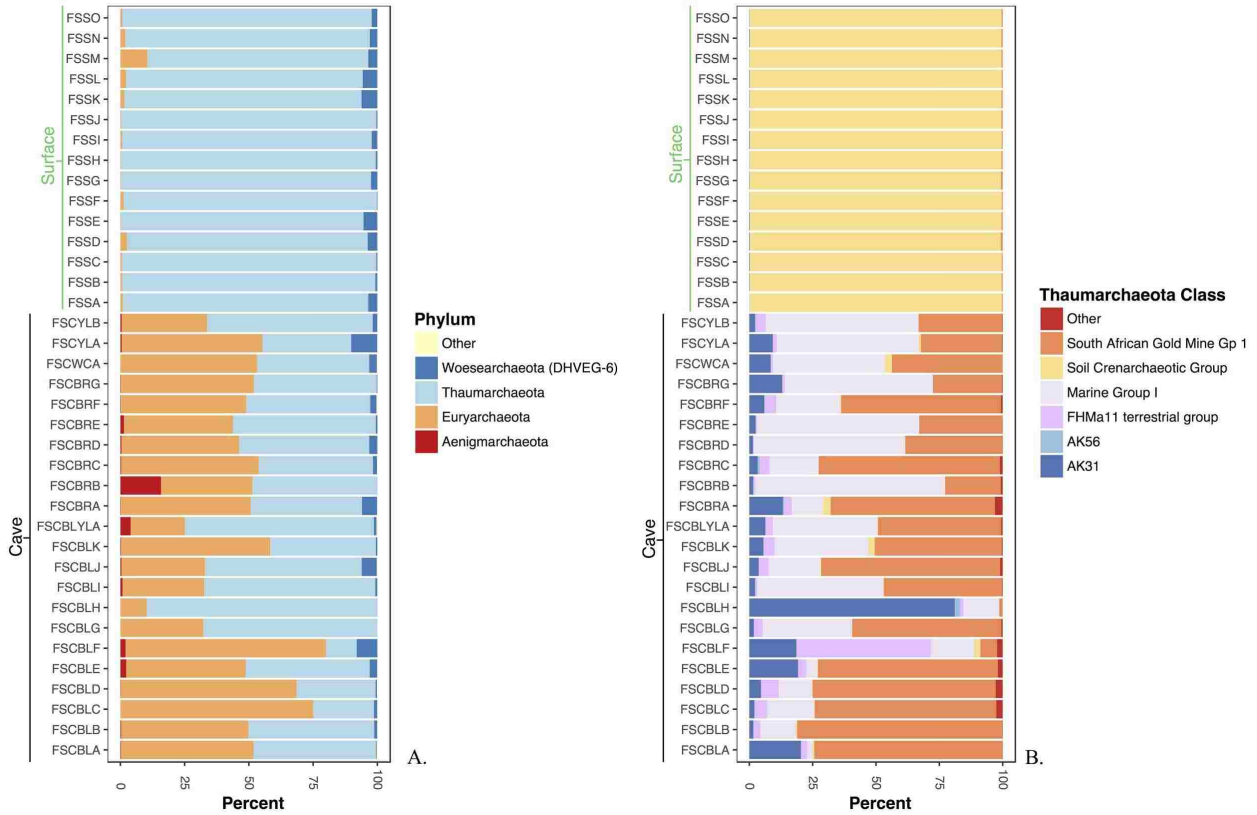


Fig. 4. Overview of 16S rRNA annotated archaeal sequences. (a) Archaeal phyla annotated from surface and cave reads. (b) Thaumarchaeota taxonomy broken down by class.

Domain/Phylum	Average Relative Abundance Cave (%)	Average Relative Abundance Surface (%)	Cave OTUs	Surface OTUs	Total OTUs	Number of shared OTUs (Core 80%)
Bacteria:						
Acidobacteria	18.2	25.2	5,079	26,148	31,057	8
Actinobacteria	4.7	10.2	1,807	14,138	15,770	2
Bacteroidetes	0.8	7.8	165	9,052	9,199	0
Chloroflexi	6.0	2.0	1,904	3,986	5,827	0
Firmicutes	1.0	32.8	479	11,726	12,133	1
GAL15	10.1	<0.1	2,723	4	2,725	0
Ignavibacteriae	1.3	<0.1	164	1	164	0
Gemmatimonadetes	2.0	1.2	866	2,402	3,250	0
Latescibacteria	0.9	<0.1	340	57	386	0
Nitrospirae	18.6	0.5	4,037	517	4,531	1
Omnitrophica	0.5	<0.1	127	238	362	0
Planctomycetes	3.1	4.4	1,146	6,286	7,368	0
Proteobacteria	27.1	11.1	7,910	14,260	21,914	4
RBG-1 (Zixibacteria)	2.6	<0.1	939	1	939	0
SBR1093	1.9	<0.1	345	1	345	0
Verrucomicrobia	0.4	2.3	153	3,534	3,662	1
Archaea:						
Aenigmarchaeota	1.4	<0.1	117	2	117	0
Bathyarchaeota	0.1	<0.1	29	4	29	1
Euryarchaeota	46.6	1.7	23,663	238	23,837	9
Thaumarchaeota	49.6	95.6	19,251	6,899	26,070	9
Woesearchaeota	2.3	2.7	576	363	918	0

Table 2. Overview of relative abundances and corresponding OTUs associated with *all* annotated bacterial and archaeal reads.

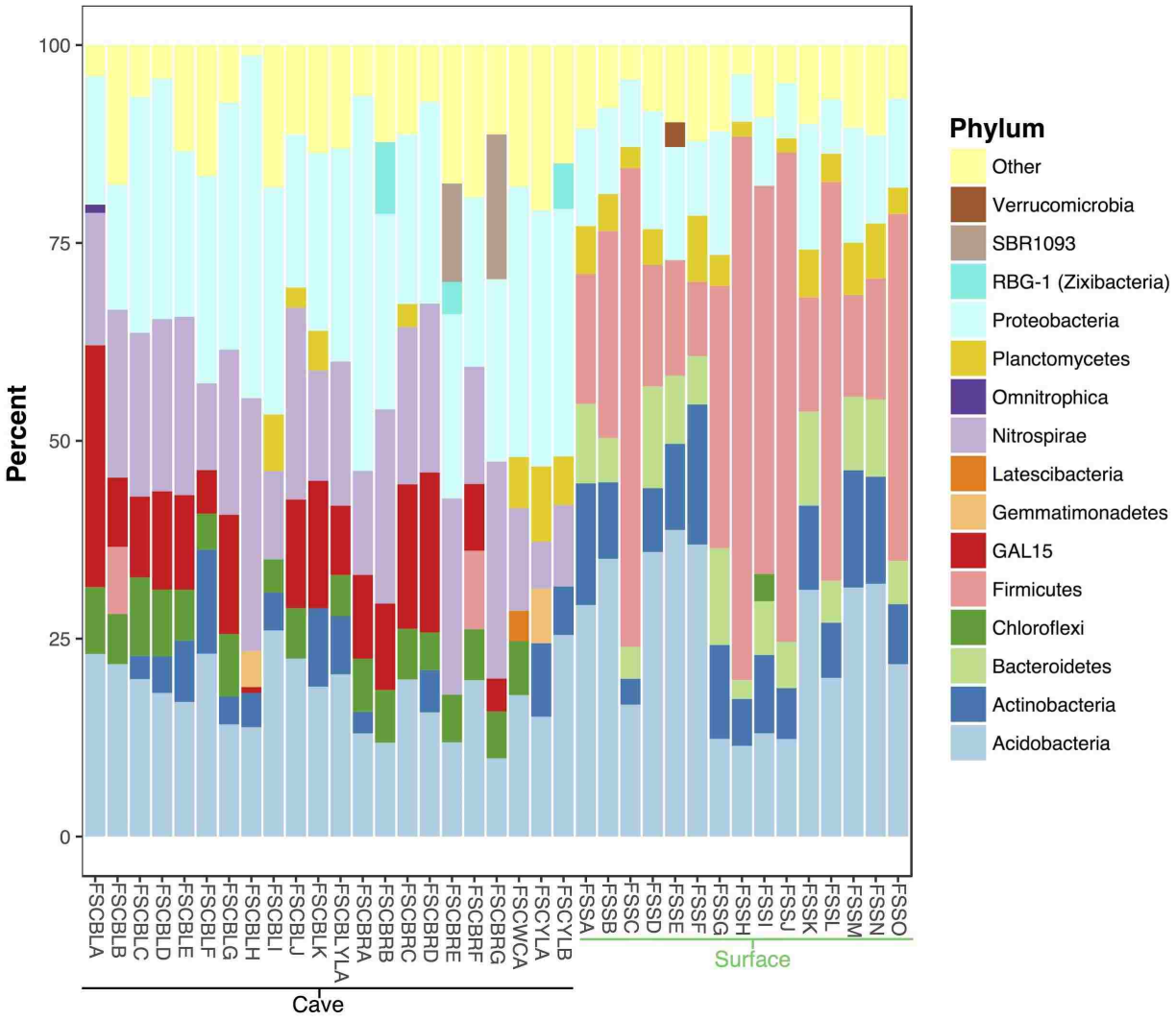


Fig. 5. Overview of 16S rRNA bacterial taxonomic findings in cave FMD and overlying surface soils.

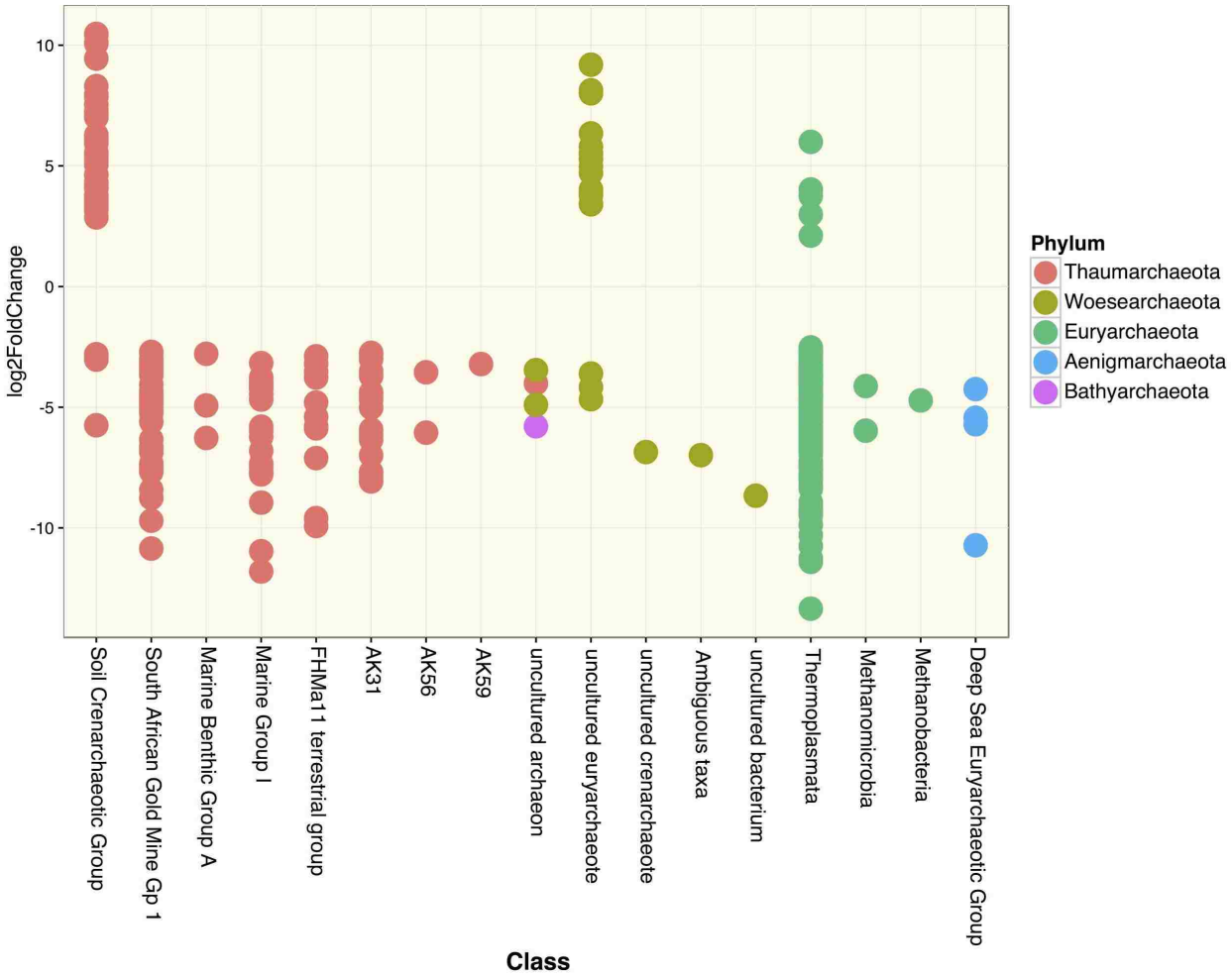


Fig. 6. DESeq2 summarizing archaeal OTUs that were significantly different between surface (top) and cave (bottom) samples.

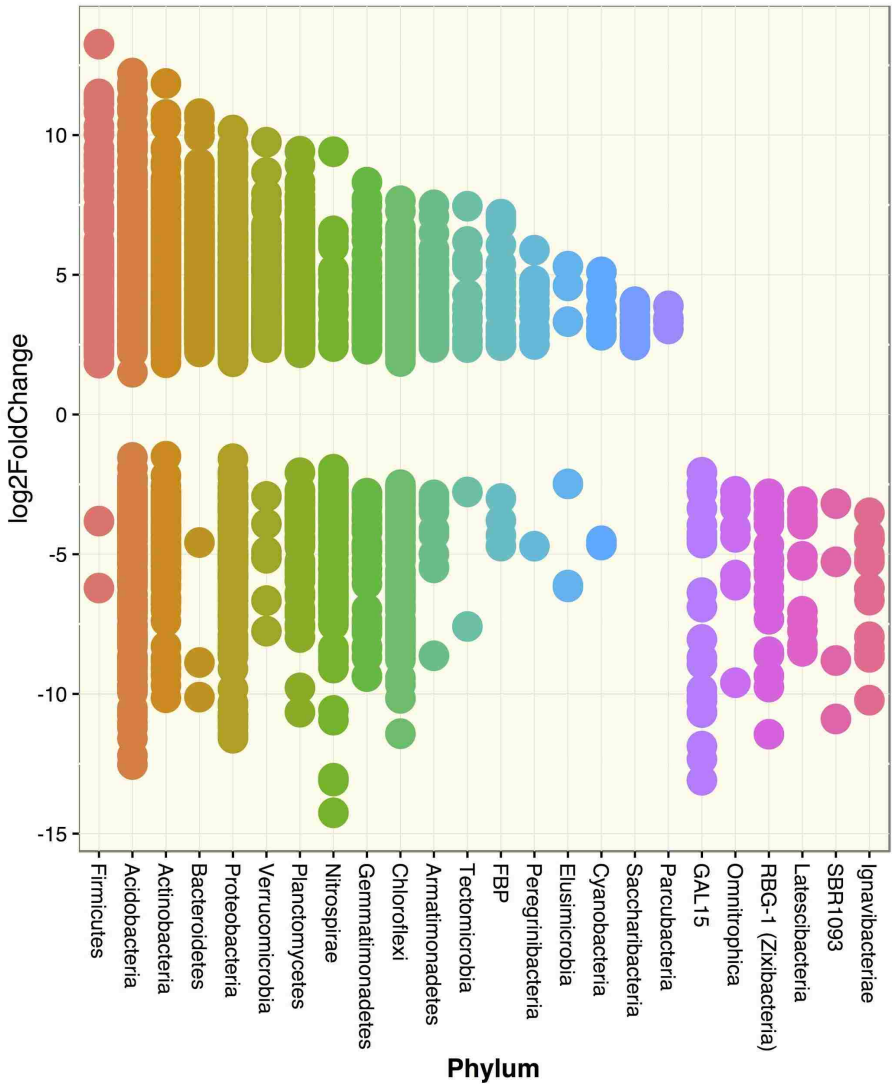
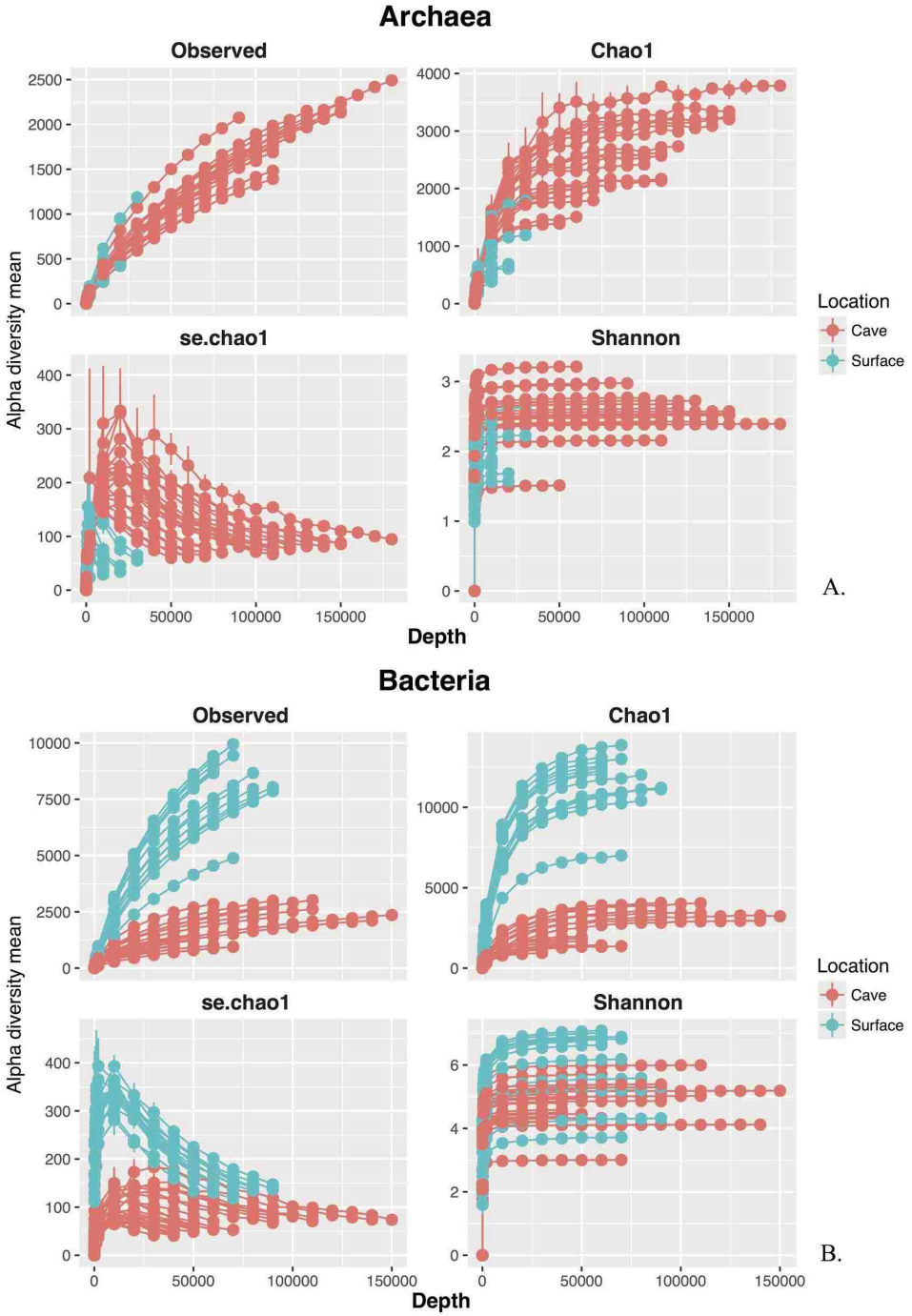


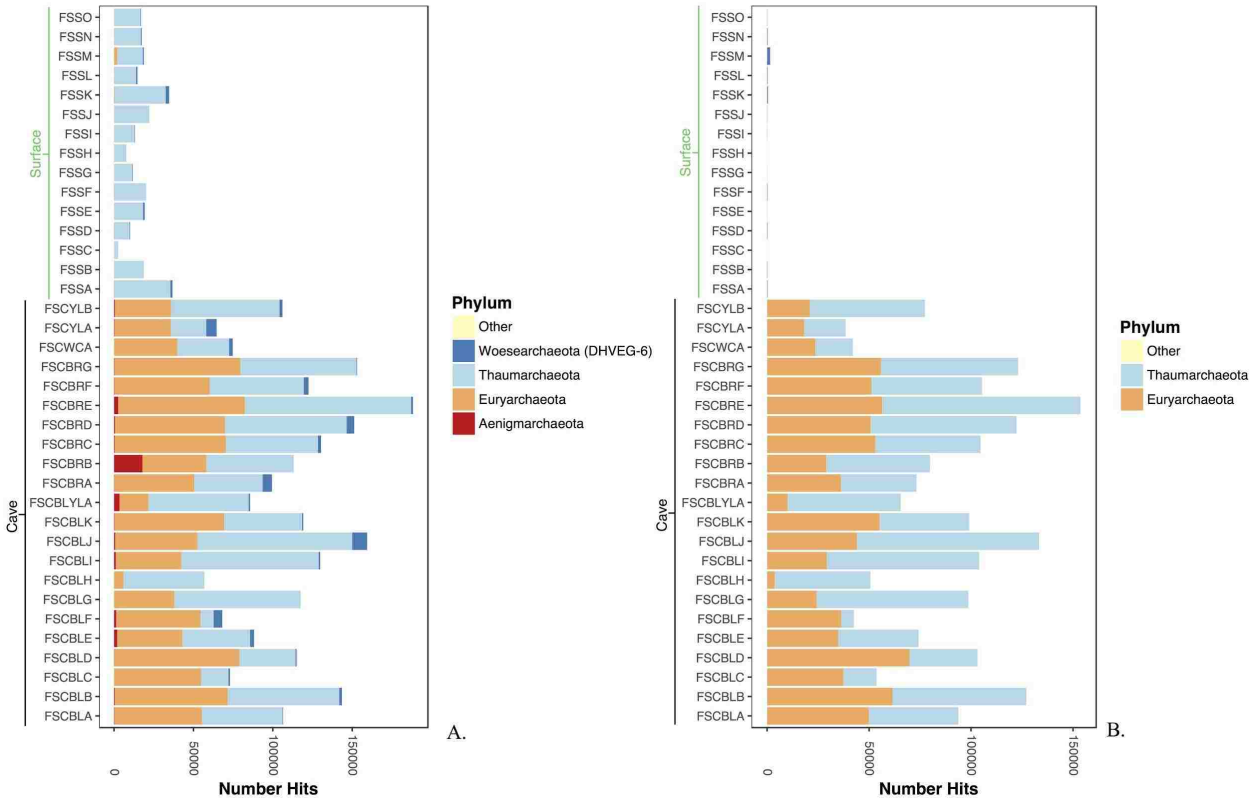
Fig. 7. DESeq2 summarizing bacterial OTUs that were significantly different between surface (top) and cave (bottom) samples. Noted are significantly different cave OTUs associated with rare bacterial phyla (e.g. GAL15 and *Ignavibacteriae*), often referred to as 'microbial dark matter' (Marcy *et al*., 2007).

Location	Average% of All Reads in Core	Average% of All Surface Reads in Core	Average% of All Cave Reads in Core	Total OTUs	Core OTUs across 80% of samples
Archaea:					
Cave	94.7	N/A	N/A	1,335	157
Surface	94.7	N/A	N/A	344	55
Cave and Surface	N/A	1.9	78.4	1,639	19
Bacteria:					
Cave	81.1	N/A	N/A	3,384	232
Surface	79.5	N/A	N/A	8,948	1,771
Cave and Surface	N/A	6.1	6.1	12,051	17

Table 3. Overview of core archaeal and bacterial microbiome OTUs and read abundances between and amongst surface and cave biomes. The total OTUs column represents only OTUs with more than 5 reads.



Supplementary Fig. 1. Alpha diversity estimated sample-based rarefaction richness, plotted against depth.



Supplementary Fig. 2. 16S rRNA raw (a) and archaeal core microbiome (b) reads found in cave and surface samples. Evident is a large majority of cave reads are associated with the 19 shared OTUs between the surface and cave. There were no shared OTUs

Chapter 4: Comparison of Archaeal Communities Found in Caves of Different Geochemistry Across Geographical Regions

Jason C. Kimble[1], Patrick A. Lewis[1], Ara S. Winter[1], Robert L. Sinsabaugh[1], and Diana E. Northup[1]

[1]Department of Biology, MSC03-2020, University of New Mexico, Albuquerque, NM 87131

*Corresponding Author
Jason C. Kimble – jkimble@unm.edu

Abstract

Recent investigations have begun to shed light on archaeal functional roles and taxonomic compositions in caves. These types of studies allow us to infer the ecological importance of these organisms in caves; however, we still know very little about other archaeal groups present, their potential metabolisms, and taxonomic similarities and/or differences among different cave types and locations. We used Illumina MiSeq to target and amplify archaeal 16S rRNA genes in samples collected from several caves of dissimilar geochemistry across three sites. There were no operational taxonomic units (OTUs) shared across the 59 cave samples; however, OTU similarities and corresponding high read abundances were observed amongst and between samples from our study sites. Non-metric dimensional scaling (NMDS) analyses demonstrated that our samples clustered by site, with the most dissimilarity observed in Parashant National Monument. Taxonomic dissimilarities were evident between sites and amongst different sample types, such as high relative abundances of the phyla *Altiarchaeales* and *Bathyarchaeota* in samples found in the hydrogen sulfide rich cave Cueva de Villa Luz. Fort Stanton Cave had the highest number of shared OTUs and taxonomic compositions were similar among samples.
We hypothesize that overburden and connectivity with the surface, evident by the high number of the *Thaumarchaeota* class Soil Crenarchaeotic Group (SCG) reads in several Parashant caves, and cave types and host-rock geochemistry, strongly influence archaeal groups present in caves. Based on our findings, we also inferred potential metabolic pathways, allowing for a better understanding of how archaeal communities vary in caves and their functional roles in these subsurface environments.

Introduction

Since Woese *et al*. (1990) used a molecular phylogenetic approach to propose that life is comprised of three domains, *Archaea* representing the third domain, our view of microbial ecology has shifted dramatically and influenced how we study these organisms. Once thought to be restricted to extreme ecological niches (Blöchl *et al*., 1997), molecular-based studies have revealed that "nonextreme" archaea are widely distributed in nature in a variety of different environments (DeLong *et al*., 1998). Some of these archaeal groups are associated with important nutrient cycling biological pathways that can influence local ecological environments (Jarrell *et al*., 2011). The autotrophic ammonia-oxidizing archaeal (AOA) phylum *Thaumarchaeota* (Brochier-Armanet *et al*., 2008) has been described in a range of environments (Erguder *et al*., 2009). Thought to comprise 20-30% of all planktonic cells in oceans (Martens-Habbena and Stahl, 2011) and found to be the dominant ammonia oxidizers in soils (Leininger *et al*., 2006), *Thaumarchaeota* are important contributors to the global nitrogen cycle (Prosser and Nicol, 2008). Other archaeal groups that participate in biological sulfidogenesis and sulfide and sulfur oxidation, by using a variety of electron donors and acceptors in reducing and oxidizing environments, are important drivers of the sulfur cycle (Offre *et al*., 2013). Recently described members of the *Thermoplasmatales*, a *Euryarchaeota* order previously associated with sulfur-reduction (Huber and Stetter, 2006), have methanogenic metabolisms (Dridi *et al*., 2012; Paul *et al*., 2012). The diverse metabolisms of *Archaea*, and their fundamental contributions to global biogeochemical cycles (Offre *et al*., 2013), could have important implications for understanding nutrient cycling and key players in caves.

Culture-independent studies have provided insight into the types of archaeal groups present in caves. In the deep carbonate Lechuguilla Cave, New Mexico, USA, Northup *et al*. (2003) sequenced clones related to *Euryarchaeota* and the *Thaumarchaeota* class South African gold mine crenarchaeotic group (SAGMCG) (Takai *et al*., 2001). Within a sulfur-based chemoautotrophically driven cave, Movile Cave, southern Romania (Sarbu *et al*., 1996), clones phylogenetically related to the Miscellaneous Crenarchaeotic Group (MCG) and Deep-sea Hydrothermal Vent (DHVE1) *Euryarchaeota* group were reported, although there was no evidence of archaeal *amoA* genes or 16S rDNA reads (Chen *et al*., 2009). *Thaumarchaeota* from group I.1b (Soil Crenarchaeotic Group (SCG)) and group I.1b (Marine Group I) were found on speleothem surfaces in Kartchner Cavern in Arizona, USA (Ortiz *et al*., 2014), while the acidophilic group I.1c (Lehtovirta *et al*., 2009) were detected in Roraima Sur Cave, an orthoquartzite cave in Venezuela (Barton *et al*., 2014), and cave sediments, central China (Zhao *et al*., 2017). Recent studies provide evidence that *Thaumarchaeota* play a critical role in the cycling of nitrogen in caves (Ortiz *et al*., 2014; Tetu *et al*., 2013; Zhao *et al*., 2017). These studies offer evidence that diverse archaeal groups are found in different cave types; however, there is a knowledge disparity with respect to bacterial investigations in caves, and we still know very little about archaeal taxonomy in caves and their functional roles (Tomczyk-Żak and Zielenkiewicz, 2016).

We lack an understanding of archaeal diversity among different cave types (e.g. carbonate and basalt) and environmental variables (e.g. depth below the surface and sulfide-rich conditions) that shape taxonomic compositions. Using 16S rRNA gene next-generation sequencing methods, our study sought to shed light on archaeal taxonomic

differences and/or similarities among diverse caves types, geochemistry, and geographic localities. We hypothesized that cave depth and underlying host-rock geochemistry would be the primary factors that determine archaeal diversity in our subsurface study sites. By providing a comprehensive representation of archaeal groups occurring in caves at our three study sites, we were able to make functional predictions based on previous research and infer ecological roles in different cave types. From our findings, we hypothesize that overburden and connectivity with the surface, which would influence the influx of nutrients and surface microbes, and host-rock geochemistry, are driving factors that influence archaeal diversity in caves.

Methods

Sampling Sites

Cueva de Villa Luz (CVL), a limestone cave in the Mexican state of Tabasco, is home to a wide range of biofilms that vary in texture, color, and microbial composition (Figure 1). The biofilms are exposed to episodic and variable inputs of reduced gases, such as hydrogen sulfide, carbon monoxide, methane, etc. Thick, mucoid biofilms line the necks of springs that emerge below the surface of the stream that runs through the cave. Other mucoid biofilms present on the cave walls include snottites (acidic, mucoid stalactites). In the dark zone of the cave, green biofilms coat the rocks in the streams and provide food for the insect life of the cave. As the hydrogen sulfide vents into the cave through 26 inlets into the stream, it is converted into elemental sulfur, sulfuric acid, and gypsum, both biologically and chemically. Many of the passages from the entrance zone to the dark zone contain raised lacey networks of biofilms, called biovermiculations.

Samples were previously obtained from areas (Table 1) with high (e.g. Phlegma Springs) and low (e.g. Ragu Passage) sulfide levels, normal (e.g. Ragu Passage) and low oxygen (e.g. Perched Pool and Yellow Roses Passage) levels, and differing pH levels. The most extreme acidic pH levels were found in snottites and gypsum crusts, while more basic levels were found in the stream deposits, where carbonate buffers any entering acid.

Parashant National Monument (PARA), AZ, samples were collected from several cave (e.g. basalt block) and sample (e.g. microbial mat) types (Table 2). Host-rock geologies include Kaibab Limestone, Red Wall Limestone, Moenkopi Mudstone, and Pleistocene Basalt. Cave formation types include sinkhole (PARA1001, PARA0901, PARA0601, and PARAMC2107), fault-driven (PARA2204 and PARA3504), basalt fissure (PARAAHF), spring void (PARA2206), and artificial (PARASVMine) caves/mine. Samples were collected from soils (top 1 cm) above several of the caves (Table 2).

Metadata and collection methods for Fort Stanton Cave (FSC) ferromanganese deposits (FMDs) and overlying surface soil samples are described in Kimble *et al*. (unpublished). The study collected samples from Snowy River passage FMD, and the samples were found throughout the passage on the walls and ceilings of the cave. FSC is the 63nd longest cave in the world at >50 km (http://www.caverbob.com/wlong.htm), and the passage is now mapped to 17.8 km in length. The 22 cave FMD samples (10-20 cc) were collected using sterile instruments and aseptic techniques to limit human-contamination. Access to the Bureau of Land Management (BLM) managed cave is by permit only (Northup, Principal Investigator). Surface samples were collected from the top 1 cm of overlying soils and corresponded to underlying survey stations in Snowy

River passage. Immediate addition of sucrose lysis buffer was made to preserve genomic material and lyse cells (Giovannoni *et al*., 1990). Samples were stored at -80°C until genomic extraction.

Molecular Processing of Samples

Samples were extracted for genomic material according to the manufacturer's protocol using the MoBio PowerSoil™ DNA Isolation Kit (MoBio, Carlsbad, CA). Sample extractions consisted of three replicates and a negative control. In an effort to maximize DNA yield for sequencing, a bead beater was used to lyse cells, and samples were eluted with 40 µl of 10 mM Tris buffer, instead of the recommended 100 µl. Each sample extraction included a negative control and three replicates. A total of 26 surface soil samples and 59 cave samples were sent to Molecular Research LP (MR DNA) (http://www.mrdnalab.com/), Shallowater, TX, for sequencing.

Sequencing

The Illumina MiSeq platform (www.illumina.com) was used to render amplicon archaeal data by means of targeting the 16S rRNA gene using the primers 349-forward (5'-GYGCASCAGKCGMGAAW-3') and 806-reverse (5'-GGACTACVSGGGTATCTAAT-3') (Takai and Horikoshi, 2000). With the barcode on the forward primer, HotStarTaq Plus Master Mix (Qiagen, USA) was used with a 28 cycle PCR (5 cycle used on PCR products) under the following specifications: 94°C for 3 min, followed by 28 cycles of 94°C for 30 s, 53°C for 40 s, 72°C for 1 min, with a final elongation step at 72°C for 5 min. Subsequently, amplification success and band intensity were determined by checking PCR products in a 2% agarose gel. Samples of similar

molecular weights and DNA concentrations were pooled in equal amounts and purified using calibrated Ampure XP beads. DNA libraries were prepared from the pooled and purified PCR product based on the Illumina TruSeq DNA library preparation protocol. Sequencing was done at Molecular Research (MR DNA) (http://www.mrdnalab.com/), Shallowater, TX, on a MiSeq platform according to the manufacturer's guidelines.

16S Taxonomic Analyses of Cave and Surface Samples

A Quantitative Insights Into Microbial Ecology (QIIME) (Caporaso *et al*., 2010) (qiime.org) pipeline and the SILVA release 128 database (http://www.arb-silva.de/download/archive/qiime/) (Quast *et al*., 2013) were used to analyze (e.g. quality controls and clustering) and annotate Illumina reads. These methods were described previously (Kimble *et al*., unpublished). Downstream analyses required the QIIME *filter_taxa_from_otu_table.py* script to remove *all* reads not classified as archaeal from our pipeline *de_novo_sumaclust.biom* file. In addition, this step also removed four PARA cave samples that had less than 1,000 reads, as determined from annotated taxonomy files. The filtered biom file (*filtered_otu_table_Archaea_Only.biom*) was converted to an R (R Core Team, 2016) readable JSON biom format using the *biom convert* command (http://biom-format.org/documentation/biom_conversion.html). The R (https://www.r-project.org/) packages phyloseq (McMurdie and Holmes, 2013) and ggplot2 (Wickham, 2009) were used to analyze and visualize sequencing data. Alpha diversity box plots (observed OTUs, Shannon and Chao1 indices) were produced by importing the *fixed_filtered_Archaea.biom* JSON formatted and mapping files. Rarefaction curves were plotted against depth using a tailored script (https://github.com/joey711/phyloseq/issues/143). Multidimensional scaling analyses

initially used the QIIME *normalize_table.py* script to normalize raw (*fixed_filtered_Archaea.biom*) archaeal filtered reads by cumulative-sum scaling (CSS) (Paulson *et al*., 2013) transformation. Ordination plots were then generated in R from the normalized archaeal biome file by non-metric dimensional scaling (NMDS) with Brays-Curtis distance. Lastly, bar chart visuals were created in R with ggplot2, using sorted taxonomy data.

Core microbiome shared OTUs amongst and between study sites were determined by using the QIIME *compute_core_microbiome.py* script. The *filter_samples_from_otu_table.py* and *split_otu_table.py* QIIME scripts were used to produce individual biom files (e.g. PARA *only* and FSC and CVL) from the raw *filtered_otu_table_Archaea_Only.biom* file. Each biome file was run independently and core microbiome taxonomy related to shared OTUs found in not less than 80% of all samples was summarized (*summarize_taxa.py*) from each core biome file.

Results

Richness and Multidimensional Scaling Site Comparison

Archaeal alpha diversity boxplot measurements from the three study sites are provided in Figure 2 and are represented as individual samples in Supplementary Figure 1. Observed and Chao1 estimator medians showed that overall archaeal richness was lowest in CVL and highest in FSC samples. Richness in PARA cave samples demonstrated the most variability. Shannon CVL indices exhibited the most unevenness, indicating some samples are dominated by few OTUs. Similar Shannon index medians were observed in FSC surface, PARA surface, and PARA cave samples. Plotted against

depth, Chao1 estimator rarefaction curves suggest that we have captured a considerable portion of the archaeal communities present in our samples (Supplementary Figure 2).

NMDS site analyses indicated that samples largely clustered by site and sample location (i.e. surface), PARA being the exception (Figure 3). Surface samples from FSC and PARA grouped together, with only slight dissimilarity. FSC and CVL samples clustered by site, while PARA cave samples exhibited the most OTU dissimilarity with some samples grouping with surface soils (Figure 3A). NMDS analyzed by cave only (Figure 3B) illustrates PARA outliers and loose clustering.

SILVA v128 Assigned Domain and Sequencing Stats

The primers that were used amplified some bacterial groups, representing an average of 28.4% and 78.6% of the dataset for all cave and surface samples, respectively. These reads were discarded and are not included in any of our analyses. Unassigned reads not classified as either archaeal or bacterial were more prevalent in cave samples with an average relative abundance of 12.7% in cave and 2.1% in surface samples. The number of unassigned reads varied between study site and sample type. The average number of unassigned cave reads in CVL, FSC, and PARA were 22.5%, 13.3%, and 7.4%, respectively. The highest unassigned relative abundances were reported in the three CVL phlegmball samples, comprising between 49.6% (CVLK) to 81.1% (CVLL) of reads. Multidimensional scaling comparison of unassigned reads (Supplementary Figure 3) showed a similar pattern to our surface and cave archaeal NMDS plot (Figure 3A).

There were 3,980,593 cave (average of 59,565 reads among 59 samples) and 483,216 (average of 17,957 reads among 26 samples) surface archaeal reads identified

and annotated. The average number of archaeal cave reads was lowest in CVL (26,729) and highest in FSC samples (112,117), while PARA had an average of 47,678 reads. Average archaeal surface reads were 17,636 at FSC and 19,880 at PARA.

Archaeal 16S rRNA Gene Taxonomic Composition of Cueva de Villa Luz (CVL)

Annotated 16S rDNA CVL reads indicate some taxonomic similarities; however, substantial differences were observed in some of the atypical samples (e.g. thick, mucoid biofilms referred to as phlegmballs) (Figure 4; Table 1). *Thermoplasmata* represented nearly all annotated *Euryarchaeota* reads (90.5-100%). This class was underrepresented in the three phlegmball and blue/black biovermiculations samples. Taxonomic classifications and relative abundances within *Thermoplasmata* varied substantially across different CVL samples. The genus *Ferroplasma* constituted 96.3% of reads in the green rock (CVLA) sample, the genus *Thermoplasma* represented approximately 98% relative abundances in the two snottite samples (CVLC and CVLI) and the brown/black biovermiculations (CVLF) sample. Other *Thermoplasmata* reads with elevated abundances were related to AMOS1A-4113-D04, ASC21, Marine Benthic Group D and DHVEG-1, and Terrestrial Miscellaneous Gp (TMEG), all uncultured Archaea.

Woesearchaeota (DHVEG-6) was identified in all samples, with elevated relative abundances in the microbial mat (white sulfur rock) (CVLB), black sediment (CVLD), and blue/black biovermiculations (CVLH) samples. The ammonia oxidizing phylum *Thaumarchaeota* was found in all samples, although the number of reads identified varied. Annotated *Thaumarchaeota* reads included the classes AK59, FHMa11 terrestrial group, Marine Group I, SCG, and South African Gold Mine Gp 1 (SAGMCG-1).

Additional *Thaumarchaeota* classes were found and classified in the other category due to the low number of reads. Elevated relative abundances of *Bathyarchaeota* and *Hadesarchaea* were only reported in the taxonomically diverse phlegmball sample CVLK. The other phlegmball samples (CVI and CVJ) were largely dominated by a single phylum, *Altiarchaeales*. OTUs related to these taxonomic findings reveal sample variations that are illustrated in an NMDS CVL only ordination plot (Figure 5).

Archaeal 16S rRNA Gene Taxonomic Composition of Parashant National Monument (PARA)

PARA16S rDNA results showed that archaeal communities in soils are primarily composed of the *Thaumarchaeota* class SCG, with elevated abundances of the class *Thermoplasmata* and *Woesearchaeota* (DHVEG-6) in several samples (Figure 6). SCG was also the dominant archaeal group in numerous cave samples with reported relative abundances >97% in cave samples PARA1001A and PARA0901A. The genus '*Candidatus* Nitrososphaera' represented a large proportion of SCG reads related to both surface and cave samples with an average surface relative abundance of 48.1% (sample average range of 6.3-80.1%) and cave sample averages between <0.1% (PARA0601B) and 85% (PARA2204E). An NMDS PARA only ordination plot showed that OTUs in several cave samples (e.g. PARA0901A) are similar to those found in surface soils (Figure 7).

Taxonomic differences amongst PARA cave samples are presented in Figure 6. Besides SCG, several PARA samples from different caves in the monument were dominated by a single phylum. A high relative abundance of *Thermoplasmata* was found

in a mineral sample PARASVMineA (Savanic Mine), while the *Euryarchaeota* class *Halobacteria* was predominant in FMD sample PARA2204A and mineral samples PARA0601A-B. Both of these samples came from carbonate caves (Table 2). The most taxonomically diverse samples, similar to archaeal groups and abundances previously found in FSC samples (Figure 8) (Kimble *et al*., unpublished), were in samples from PARA3504 cave. These included the phyla *Aenigmarchaeota* and *Woesearchaeota* (DHVEG-6) and *Thaumarchaeota* classes AK31, FHMa11 terrestrial group, Marine Group I, and South African Gold Mine Gp 1 (SAGMCG-1).

16S rRNA Gene Archaeal Core Microbiome Between and Amongst Cave Study Sites

There were no OTUs shared across 100% of the cave samples from the three sites. Only a single OTU (SAGMCG-1) was shared over 95% of the samples, absent only in two CVL samples. All other core microbiome analyses looked at OTUs occurring in not less than 80% of all samples. Using this criterion, there were four shared OTUs amongst *all* cave samples related to the *Thaumarchaeota* classes Marine Group I, SCG, and SAGMCG-1. These shared OTUs accounted for an average of 5.0%, 28.1%, and 30.6% of *all* reads in CVL, PARA, and FSC core microbiomes, respectively (Table 3). Core cave microbiome analyses between and amongst the three study sites are summarized in Table 3.

Discussion

Overview

The results of our 16S rDNA taxonomic investigation of different cave and sample types across three geographic localities expanded our knowledge of archaeal diversity and taxonomic differences among varied sample types in three geographically different subsurface environments. With respect to shared OTUs, there was little overlap between the three study sites. Within individual study sites, reads associated with shared OTUs represented a majority of taxa identified; however, taxonomic makeup varied substantially in different caves (i.e. CVL and PARAMC2107) and sample types. While taxonomic archaeal groups in FSC FMD samples of different FMD colors (Kimble *et al*., unpublished) were similar, PARA FMDs and mineral types showed dissimilarity. The abundance of SCG in several PARA soil samples and caves suggests connectivity to the surface in contrast to SCG results in samples from moderately deep FSC. Potential key players in caves include the ammonia oxidizing phylum *Thaumarchaeota* (Brochier-Armanet *et al*., 2008) and the sulfur-respiring (Huber and Stetter, 2006) and methanogenic (Dridi *et al*., 2012; Paul *et al*., 2012) phylum *Thermoplasmata*. These findings provide taxonomic evidence that archaeal groups in caves are more diverse than previously known (Macalady *et al*., 2007; Spear *et al*., 2007; Weidler *et al*., 2008; Chen *et al*., 2009; Lehtovirta *et al*., 2009; Northup *et al*., 2003; Jones *et al*., 2012; Jones *et al*., 2014; Tetu *et al*., 2013, Barton *et al*., 2014; Ortiz *et al*., 2014, Reitschuler *et al.*, 2015; Reitschuler *et al.*, 2016; Gulecal-Pektas and Temel, 2017; Zhao *et al*., 2017), and the archaeal taxa that are present are influenced by cave type, host-rock geochemistry, and surface connectivity. Multidimensional scaling analyses of unassigned reads also provides evidence that novel microbes inhabit our study sites and should be a subject for future metagenomic investigations.

16S Archaeal CVL Taxonomic Findings

CVL had high relative abundances related to shared OTUs within the site; however, taxonomic compositions varied considerably. In two CVL gelatinous biofilm phlegmball samples (CVLJ and CVLL) from different cave springs (Slot Spring and Phlegma Spring), the archaeal phylum *Altiarchaeales* comprised the most reads. However, another phlegmball sample (CVLK) from Slot Spring that contained pieces of pyrite rock was found to be more taxonomically diverse. This sample had elevated reads for the archaeal phyla *Bathyarchaeota* and *Hadesarchaea*. With respect to samples CVLJ and CVLK, most *Altiarchaeales* reads fell into a single OTU (denovo6). Using BLAST (Altschul *et al*. 1990) to search highly similar sequences, this OTU had a 98% identity to a sequence (EF444610) found in a geothermal spring in Greece (Kormas *et al*., 2009). The next closest identity was a distant 78%. Found in deep anoxic groundwaters, including cold sulfidic springs, *Altiarchaeales* form "highly-pure biofilms" (Probst and Moissl-Eichinger, 2015). Genome reconstruction suggests these at least some of these organisms fix carbon through a reductive acetyl-CoA pathway (Wood-Ljungdahl pathway) (Probst and Moissl-Eichinger, 2015). Genome reconstruction (Alti-1) from estuarine sediments indicated that some members of the *Altiarchaeales* have sulfur metabolism related genes, are autotrophic (possibly using H_2 or CO as an electron donor), and have genes related to biofilm formation (Bird *et al*., 2016). In CVL, CO levels associated with outgassing from groundwater inlets have been measured as high as 85 ppm (Hose *et al*., 2000), which would provide these organisms with an energy and carbon sources and explain their presence in the cave. Based on this information, CVL

likely provides an ecological niche for *Altiarchaeales,* and they likely participate in the formation of the gelatinous phlegmballs.

Archaeal taxonomic diversity in the phlegmball sample (CVLK) with pyrite inclusions was unexpectedly different from other phlegmball samples. *Bathyarchaeota,* which dominated the CVLK sample, were nearly absent in all other samples, including phlegmballs. Previously classified as Miscellaneous Crenarchaeota group (MCG), *Bathyarchaeota* have been described in many diverse environments and are thought to degrade aromatic compounds in anaerobic environments (Meng *et al*., 2014). A recent study, that used genome reconstruction methods, identified genes (BA1 genome) associated with the Wood-Ljungdahl and fermentation pathways and key methane metabolism genes (e.g. methanogenesis from methyl sulfides) (Evans *et al*., 2015). *Bathyarchaeota* in CVL may have diverse forms of metabolism depending on niche adaptability, as Evans *et al*. (2015) did not find key methanogenic genes (*mcrA*) in several of the metagenomes. They acknowledge that additional genomes are needed to understand their functional potential given their phylogenetic differences and presence in many environments. There were many OTUs related to *Bathyarchaeota* in CVLK that had a similar number of reads, in contrast to *Altiarchaeales*a and *Hadesarchaea*, that had one and two dominant OTUs, respectively.

Stable isotopic compositions of deep-sea subsurface sediments indicated that archaeal groups (e.g. South-African Gold Mine Miscellaneous Euryarchaeal Group (SAGMEG)) present in these sediments were heterotrophic (Biddle *et al*., 2006). Now classified as *Hadesarchaea*, these organisms with reduced genome sizes have been found to be "specialized for survival in the subsurface biosphere" with genes linked to inorganic

carbon fixation, CO and H_2 oxidation, and sulfur cycling (Baker *et al*., 2016). The two dominant *Hadesarchaea* OTUs (denovo116 and denovo172) in CVLK had 99% identities to sequences previously found in Holocene subsurface sediments in Kanto Plain, Japan (AB554234) (Takeuchi *et al*., 2011) and Lake Kivu surface sediments, East African Rift (JN853766) (Bhattarai *et al*., 2012), respectively.

Other CVL samples were primarily composed of various *Thaumarchaeota* classes and the *Euryarchaeota* class *Thermoplasmata*. The near total absence of AOA in several of the samples is explained by anoxic conditions (e.g. black sediment on stream bottom rocks) unfavorable to aerobic nitrification and pH < 2 (snottite samples), which is much lower than pH values previously reported for these organisms in acidic soils (pH 3.75 and 5.4) (Hatzenpichler *et al*., 2012). We hypothesize that elevated AOA abundances in three of the biovermiculations samples are due to their affiliation with carbonate host-rock. Dissolution of the host-rock by abiotic (Hose *et al*., 2000) and biotic (Spilde *et al*., 2006) processes releases fixed N in the form of ammonium present in the bedrock.

All *Thermoplasmata* reads in CVL were classified in the thermoacidophilic order *Thermoplasmatales* (Huber and Stetter, 2006) and were associated with four dominant OTUs that provide hints about potential metabolic pathways. The two snottite samples shared an OTU (denovo20) that had a 97% identity (KT005320) to a recently cultured extremophile that was classified as the genus *Cuniculiplasma* (Golyshina *et al*., 2016). This archaeon is a facultative anaerobic organotroph, lacks a cell wall, and is acidophilic (Golyshina *et al*., 2016). Denovo40, primarily associated with CVLF and CVLB, had a 91% identity to this genus, and had a 98% identity to a sequence described in thermal and acidophilic biofilms in Michoacan, Mexico (KJ907757) (Servin-Garciduenas,

unpublished). Nearly all reads from denovo71 corresponded to the green rock sample (CVLA), high in *Ferroplasma*. Related (97% identity) to *Ferroplasma acidiphilum* (JF891386) (Bulaev *et al*., 2011), this autotrophic acidophile "oxidizes ferrous iron as the sole energy source" (Golyshina et al., 2000). There were many *Thermoplasmatales* reads that were classified as uncultured, which recent research suggests are methanogenic archaea (Paul *et al*., 2012); however, one of the dominant OTUs (denovo17) only had an 87% identity to a *Methanomassiliicoccus* sp. (LN827539) (Huynh *et al*., 2016). These findings provide evidence that *Thermoplasmata* in CVL may have diverse metabolic capacities specific to their ecological niche.

16S Archaeal PARA Taxonomic Findings

It was evident from NMDS and core microbiome analyses that there was dissimilarity among PARA caves and sample types. PARA also had the fewest number of reads found among the core microbiome. A fundamental observation was high abundances of SCG in many subsurface samples. To investigate if SCG in the cave were related to archaeal groups in overlying soils, a core microbiome analysis was done. There were only nine shared OTUs between the 11 surface and 25 cave samples; however, these OTUs accounted for an average of 85.1% (SCG average of 99.5%) and 47.7% (SCG average of 73.3%) of *all* archaeal surface and cave reads, respectively. In individual cave samples, abundance of reads related to shared core OTUs was higher (e.g. 95.9% in PARA1001A and 78.1% in PARA0901A).

We hypothesize that connectivity to the surface is responsible for high relative abundances of SCG in cave samples. PARA1001 and PARA0901 are both sinkhole caves

with soil observed throughout the passage. PARA2204 is a fault driven cave that has evidence of animal visitors and dirt covered floors, PARAAHF is located only a few meters below the surface with cracked and porous basalt rock overburden, and PARA2107 floods regularly. The investigation by Lavoie *et al.* (2017) of shallow lava caves in Lava Beds National Monument found little overlap (11.2%) between surface and cave OTUs; however, these OTUs represented an average of 64.1% *all bacterial* reads in surface samples (Kimble *et al*., unpublished). A previous surface and cave (moderately deep FSC) archaeal and bacterial investigation found that the average percent of *all* surface reads in the core was only 1.9% and 6.1%, respectively. This shows that there are few OTUs and corresponding reads in the surface related to dominant cave taxa as evidenced by low relative abundances of SCG in FMD samples. Taxonomy identified in PARA3504 was most similar to FSC and had the most overburden (439.2 m). A core microbiome analyses of all FSC cave samples to four (PARA3504G, PARA3504B, PARA3504A, and PARA3504E) PARA samples showed that an average of 87.7% of reads in these samples were found in 117 shared OTUs with FSC. These findings show that connectivity to the surface (e.g. depth) influences the microbial composition within caves. It is unknown if SCG are only transient in PARA caves or sustained input from the surface provides appropriate ammonium concentrations for these AOA to carry out the first step of nitrification. Two of the dominant SCG OTUs (denovo8 and denovo52) prevalent in PARA cave samples had 99% and 95% identities to AOA (‘*Candidatus* Nitrosocosmicus exaquare’ and *Nitrososphaera viennensis*) adapted to higher ammonium substrate concentrations (Sauder et al., 2017; Daebeler *et al*., 2014). AOA (e.g. class AK31) in FSC and PARA3504 may be adapted to low ammonium substrate

concentration, as previously reported for *Nitrosopumilus maritimus* SCM1 (Martens-Habbena *et al*. 2009).

Conclusions

The objective of our study was to use 16S rDNA next-generation sequencing to investigate similarities and/or differences of archaeal groups inhabiting three different cave sites. Our findings suggest that cave type, geochemistry, and the amount of surface input influences archaeal composition in these environments. FMDs from the Snowy River passage in FSC are largely isolated from surface input and reside in host rock of similar composition, likely selecting for archaeal groups with similar functional roles. This is exemplified in the high number of core microbiome reads and similar taxonomic composition shared amongst samples. CVL also had a high number of core microbiome reads corresponding to shared OTUs in several samples; however, relative abundances of archaeal phyla and related classes varied considerably among cave types. This research supports the hypothesis that geochemistry is an important driver of archaeal diversity in caves because archaea phyla found in the H_2S-rich cave at CVL were unlike other sampling sites. We also determined that depth is a likely determining factor in archaeal cave taxonomy. This was suggested by a marginal number of SCG reads in moderately deep FSC, and a large proportion of SCG reads and related surface soil OTUs in several shallow PARA caves and cave types that receive large quantities of input from the surface (e.g. shallow basalt and sinkhole caves). To our knowledge, this is the first cave study that characterizes, compares, and contrasts archaeal communities in different cave types with different geochemistry and geographic locations.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–10. doi: 10.1016/S0022-2836(05)80360-2

Baker, B. J., Saw, J. H., Lind, A. E., Lazar, C. S., Hinrichs, K.-U., Teske, A. P., et al. (2016). Corrigendum: genomic inference of the metabolism of cosmopolitan subsurface Archaea, *Hadesarchaea*. *Nat. Microbiol.* 1, 16032. doi: 10.1038/nmicrobiol.2016.32

Barton, H. A., Giarrizzo, J. G., Suarez, P., Robertson, C. E., Broering, M. J., and Banks, E. D. (2014). Microbial diversity in a Venezuelan orthoquartzite cave is dominated by the *Chloroflexi* (Class *Ktedonobacterales*) and *Thaumarchaeota* Group I.1c. *Front. Microbiol.* **5**:615. doi: 10.3389/fmicb.2014.00615

Bhattarai, S., Ross, K. A., Schmid, M., Anselmetti, F. S., and Bürgmann, H. (2012). Local conditions structure unique archaeal communities in the anoxic sediments of meromictic Lake Kivu. *Microb. Ecol.* 64, 291–310. doi: 10.1007/s00248-012-0034-x

Biddle, J. F., Lipp, J. S., Lever, M. A., Lloyd, K. G., Sørensen, K. B., Anderson, R., et al. (2006). Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. *Proc. Natl. Acad. Sci. U S A.* 103, 3846–3851. doi: 10.1073/pnas.0600035103

Bird, J. T., Baker, B. J., Probst, A. J., Podar, M., and Lloyd, K. G. (2016). Culture independent genomic comparisons reveal environmental adaptations for *Altiarchaeales*. *Front. Microbiol*. 7, 1–14. doi: 10.3389/fmicb.2016.01221

Blöchl, E., Rachel, R., Burggraf, S., Hafenbradl, D., Jannasch, H. W., and Stetter, K. O. (1997). *Pyrolobus fumarii*, gen. and sp. nov., represents a novel group of archaea, extending the upper temperature limit for life to 113 degrees C. *Extremophiles*. 1, 14–21. doi: 10.1007/s007920050010

Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008). Mesophilic crenarchaeota: proposal for a third archaeal phylum, the *Thaumarchaeota*. *Nat. Rev. Microbiol.* 6, 245-252. doi:10.1038/nrmicro1852

Bulaev, A. G., Pivovarova, T. A., Melamud, V. S., Bumazhkin, B. K., Patutina, E. O., Kolganova, T. V., et al. (2011). Species composition of the association of acidophilic chemolithotrophic microorganisms participating in the oxidation of gold-arsenic ore concentrate. *Microbiology*, 80, 842–849. doi: 10.1134/S0026261711060038

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME. J.* 6, 1621-1624. doi: 10.1038/ismej.2012.8

Chen, Y., Wu, L. Q., Boden, R., Hillebrand, A., Kumaresan, D., Moussard, H. et al.

(2009). Life without light: microbial diversity and evidence of sulfur- and ammonium based chemolithotrophy in Movile Cave. *ISME J.* 3: 1093-1104. doi: 10.1038/ismej.2009.57

Daebeler, A., Bodelier, P. L. E., Yan, Z., Hefting, M. M., Jia, Z., and Laanbroek, H. J. (2014). Interactions between Thaumarchaea, *Nitrospira* and methanotrophs modulate autotrophic nitrification in volcanic grassland soil. *ISME. J.* 8, 2397–2410. doi: 10.1038/ismej.2014.81

Delong, E. F. (1998). Everything in moderation: archaea as "non-extremophiles." *Curr. Opin. Genet. Dev.* 8, 649–654. doi: 10.1016/S0959-437X(98)80032-4

Dridi, B., Fardeau, M. L., Ollivier, B., Raoult, D., and Drancourt, M. (2012). Methanomassiliicoccus luminyensis gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* 62, 1902–1907. doi: 10.1099/ijs.0.033712-0

Evans, P. N., Parks, D. H., Chadwick, G. L., Robbins, S. J., Orphan, V. J., Golding, S. D., et al. (2015). Methane metabolism in the archaeal phylum *Bathyarchaeota* revealed by genome-centric metagenomics. *Science*. 350, 434–438. doi: 10.1126/science.aac7745

Giovannoni, S. J., Delong, E. F., Schmidt, T. M., and Pace, N.R. (1990). Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. *Appl. Environ. Microbiol.* 56, 2572-2575.

Golyshina, O. V., Lünsdorf, H., Kublanov, I. V., Goldenstein, N. I., Hinrichs, K. U., and Golyshin, P. N. (2016). The novel extremely acidophilic, cell-wall-deficient archaeon *Cuniculiplasma divulgatum* gen. Nov., sp. nov. represents a new family, *Cuniculiplasmataceae* fam. nov., of the order *Thermoplasmatales*. *Int. J. Syst. Evol. Microbiol*. 66, 332–340. doi: 10.1099/ijsem.0.000725

Golyshina, O. V., Pivovarova, T. A., Karavaiko, G. I., Kondrat'eva, T. F., Moore, E. R. B., Abraham, W. R., et al. (2000). *Ferroplasma acidiphilum* gen. nov., sp. nov., an acidophilic, autotrophic, ferrous-iron-oxidizing, cell-wall-lacking, mesophilic member of the *Ferroplasmaceae* fam. nov., comprising a distinct lineage of the Archaea. *Int. J. Syst. Evol. Microbiol.* 50, 997–1006. doi: 10.1099/00207713-50-3-997

Gulecal-Pektas, Y. and Temel, M. (2017). A window to the subsurface: microbial diversity in hot springs of a sulfidic cave (Kaklik, Turkey). *Geomicrobiol. J.* 34, 374–384. doi: 10.1080/01490451.2016.1204374

Hatzenpichler, R. (2012). Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. *Appl. Environ. Microbiol.* 78, 7501-7510. doi: 10.1128/AEM.01960-12

Hose, L. D., Palmer, A. N., Palmer, M. V., Northup, D. E., Boston, P. J., and DuChene, H. R. (2000). Microbiology and geochemistry in a hydrogen-sulphide-rich karst environment. *Chem. Geol.* 169, 399–423. doi: 10.1016/S0009-2541(00)00217-5

Huber, H. and Stetter, K. O. (2006). *Thermoplasmatales*. *Prokaryotes*. 3, 101-112. doi: 10.1007/0-387-30743-5_7

Huynh, H. T. T., Nkamga, V. D., Signoli, M., Tzortzis, S., Pinguet, R., Audoly, et al. (2016). Restricted diversity of dental calculus methanogens over five centuries, France. *Sci. Rep.* **6**: 25775. doi: 10.1038/srep25775

Jarrell, K. F., Walters, A. D., Bochiwal, C., Borgia, J. M., Dickinson, T., and Chong, J. P. J. (2011). Major players on the microbial stage: Why Archaea are important. *Microbiology*, 157, 919–936. doi: 10.1099/mic.0.047837-0

Jones, D. S., Albrecht, H. L., Dawson, K. S., Schaperdoth, I., Freeman, K. H., Pi, Y., et al. (2012). Community genomic analysis of an extremely acidophilic sulfur-oxidizing biofilm. *ISME. J.* 6, 158-170. doi: 10.1038/ismej.2011.75

Jones, D. S., Schaperdoth, I., and Macalady, J. L. (2014). Metagenomic evidence for sulfide oxidation in extremely acidic cave biofilms. *Geomicrobiol.* 31, 194–204. doi: 10.1080/01490451.2013.834008

Kormas, K. A., Tamaki, H., Hanada, S., and Kamagata, Y. (2009). Apparent richness and community composition of Bacteria and Archaea in geothermal springs. *Aquat. Microb. Ecol.* 57, 113–122. doi: 10.3354/ame01333

Lehtovirta, L. E., Prosser, J. I., and Nicol, G. W. (2009). Soil pH regulates the abundance and diversity of Group 1.1c *Crenarchaeota*. *FEMS Microbiol. Ecol.* 70, 367–376. https://doi.org/10.1111/j.1574-6941.2009.00748.x

Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G. W., et al. (2006). Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature*. 442, 806-809. doi:10.1038/nature04983

Macalady, J. L., Jones, D. S., and Lyon, E. H. (2007). Extremely acidic, pendulous cave wall biofilms from the Frasassi cave system, Italy. *Environ. Microbiol.* 9, 1402–1414. doi: 10.1111/j.1462-2920.2007.01256.x

Martens-Habbena, W. and Stahl, D. A. (2011). Nitrogen metabolism and kinetics of ammonia-oxidizing archaea. *Meth. Enzymol.* 496, 465-487. doi: 10.1016/B978-0-12-386489-5.00019-1

Martens-Habbena, W., Berube, P. M., Urakawa, H., de la Torre, J. R., and Stahl, D. A. (2009). Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature*. 461, 976-979. doi: 10.1038/nature08465

McMurdie, P. J. and Holmes, S. (2013) phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data *PLoS. One*. **8**:e61217. doi: 10.1371/journal.pone.0061217

Meng, J., Xu, J., Qin, D., He, Y., Xiao, X., and Wang, F. (2014). Genetic and functional properties of uncultivated MCG archaea assessed by metagenome and gene expression analyses. *ISME. J.* 8, 650–659. doi: 10.1038/ismej.2013.174

Northup, D. E., Barns, S. M., Yu, L. E., Spilde, M. N., Schelble, R. T., Dano, K. E., et al. (2003). Diverse microbial communities inhabiting ferromanganese deposits in Lechuguilla and Spider Caves. *Environ. Microbiol.* 5, 1071-1086. doi: 10.1046/j.1462-2920.2003.00500.x

Offre, P., Spang, A., and Schleper, C. (2013). Archaea in biogeochemical cycles. *Annu. Rev. Microbiol.* 67, 437–457. doi: 10.1146/annurev-micro-092412-155614

Ortiz, M., Legatzki, A., Nelson, J.W., Fryslie, B., Nelson, W.M., Wing, R.A., et al. (2014). Making a living while starving in the dark: metagenomic insights into the energy dynamics of a carbonate cave. *ISME. J.* 8, 478-491. doi:10.1038/ismej.2013.159

Paul, K., Nonoh, J. O., Mikulski, L., and Brune, A. (2012). "*Methanoplasmatales*," *Thermoplasmatales*-related archaea in termite guts and other environments, are the seventh order of methanogens. *Appl. Environ. Microbiol.* 78, 8245–8253. doi: 10.1128/AEM.02193-12

Paulson, J. N., Stine, O. C., Bravo, H. C., and Pop, M. (2013). Differential abundance analysis for microbial marker-gene surveys. *Nat. Methods*. 10, 1200–1202. doi: 10.1038/nmeth.2658

Probst, A., & Moissl-Eichinger, C. (2015). "*Altiarchaeales*": uncultivated archaea from the subsurface. *Life*, 5, 1381–1395. doi: 10.3390/life5021381

Prosser, J. I., and Nicol, G. W. (2008). Relative contributions of archaea and bacteria to aerobic ammonia oxidation in the environment. *Environ. Microbiol*. 10, 2931–2941. doi: 10.1111/j.1462-2920.2008.01775.x

Quast, C., Pruesse, E., Yilmaz, J., Schweer, T., Yarza, P., Peplies, J. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**:D590-D596. doi: 10.1093/nar/gks1219

R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [Software]. http://www.R-project.org

Reitschuler, C., Lins, P., Schwarzenauer, T., Spötl, C., Wagner, A. O., and Illmer, P. (2015). New undescribed lineages of non-extremophilic archaea form a homogeneous and dominant element within alpine moonmilk microbiomes. *Geomicrobiol. J.* 32, 890–902. doi: 10.1080/01490451.2015.1025317

Reitschuler, C., Spötl, C., Hofmann, K., Wagner, A. O., and Illmer, P. (2016). Archaeal distribution in moonmilk deposits from alpine caves and their ecophysiological potential. *Microb. Ecol.* 71, 686–699. doi: 10.1007/s00248-015-0727-z

Sarbu, S. M., Kane, T. C., and Kinkle, B. K., (1996). A chemoautotrophically based cave ecosystem. *Science*. 272, 1953–1955. doi: 10.1126/science.272.5270.1953

Sauder, L. A., Albertsen, M., Engel, K., Schwarz, J., Nielsen, P. H., Wagner, M., et al. (2017). Cultivation and characterization of Candidatus Nitrosocosmicus exaquare, an ammonia-oxidizing archaeon from a municipal wastewater treatment system. *ISME. J.* 11, 1142–1157. doi: 10.1038/ismej.2016.192

Spear, J. R., Barton, H. A., Robertson, C. E., Francis, C. A., and Pace, N. R. (2007). Microbial community biofabrics in a geothermal mine adit. *Appl. Environ. Microbiol*. 73, 6172–6180. doi:10.1128/AEM.00393-07

Spilde, M. N., Northup, D. E., and Boston, P. J. (2006). "Ferromanganese deposits in the caves of the Guadalupe Mountains" in Caves and Karst of Southeast New Mexico, ed L. Land L, Lueth (The New Mexico Geological Society), 161–166.

Takai, K. and Horikoshi, K. (2000). Rapid Detection and Quantification of Members of the Archaeal community by quantitative PCR using fluorogenic probes. *Appl. Environ. Microbiol.* 66, 5066-5072. doi: 10.1128/AEM.66.11.5066-5072.2000

Takai, K., Moser, D. P., DeFlaun, M., Onstott, T. C., and Fredrickson, J. K. (2001). Archaeal diversity in waters from deep South African gold mines. *Appl. Environ. Microbiol.* 67, 5750-5760. doi: 10.1128/AEM.67.21.5750-5760.2001

Takeuchi, M., Yoshioka, H., Seo, Y., Tanabe, S., Tamaki, H., Kamagata, Y., et al. (2011). A distinct freshwater-adapted subgroup of ANME-1 dominates active archaeal communities in terrestrial subsurfaces in Japan. *Environ. Microbiol.* 13, 3206–3218. doi: 10.1111/j.1462-2920.2011.02517.x

Tetu, S. G., Breakwell, K., Elbourne, L. D. H., Holmes, A. J., Gillings, M. R., and Paulsen, I. T. (2013). Life in the dark: metagenomic evidence that a microbial slime community is driven by inorganic nitrogen metabolism. *ISME. J.* 7, 1227–1236. doi: 10.1038/ismej.2013

Tomczyk-Żak, K., and Zielenkiewicz, U. (2016). Microbial diversity in caves. *Geomicrobiol. J.* 33, 20–38. doi: 10.1080/01490451.2014.1003341

Weidler, G. W., Gerbl, F. W., and Stan-Lotter, H. (2008). *Crenarchaeota* and their role in the nitrogen cycle in a subsurface radioactive thermal spring in the Austrian Central Alps. *Appl. Environ. Microbiol.* 74, 5934-5942. doi: 10.1128/AEM.02602-07

Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer; 2009.

Woese, C. R., Kandlert, O., and Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains. *Proc. Natl. Acad. Sci. U S A.* 87, 4576–4579.

Zhao, R., Wang, H., Yang, H., Yun, Y., and Barton, H. A. (2017). Ammonia oxidizing archaea dominant over bacterial ammonia oxidizers within alkaline cave sediments. *Geomicro. J.* doi: 10.1080/01490451.2016.1225861

Figures and Legends

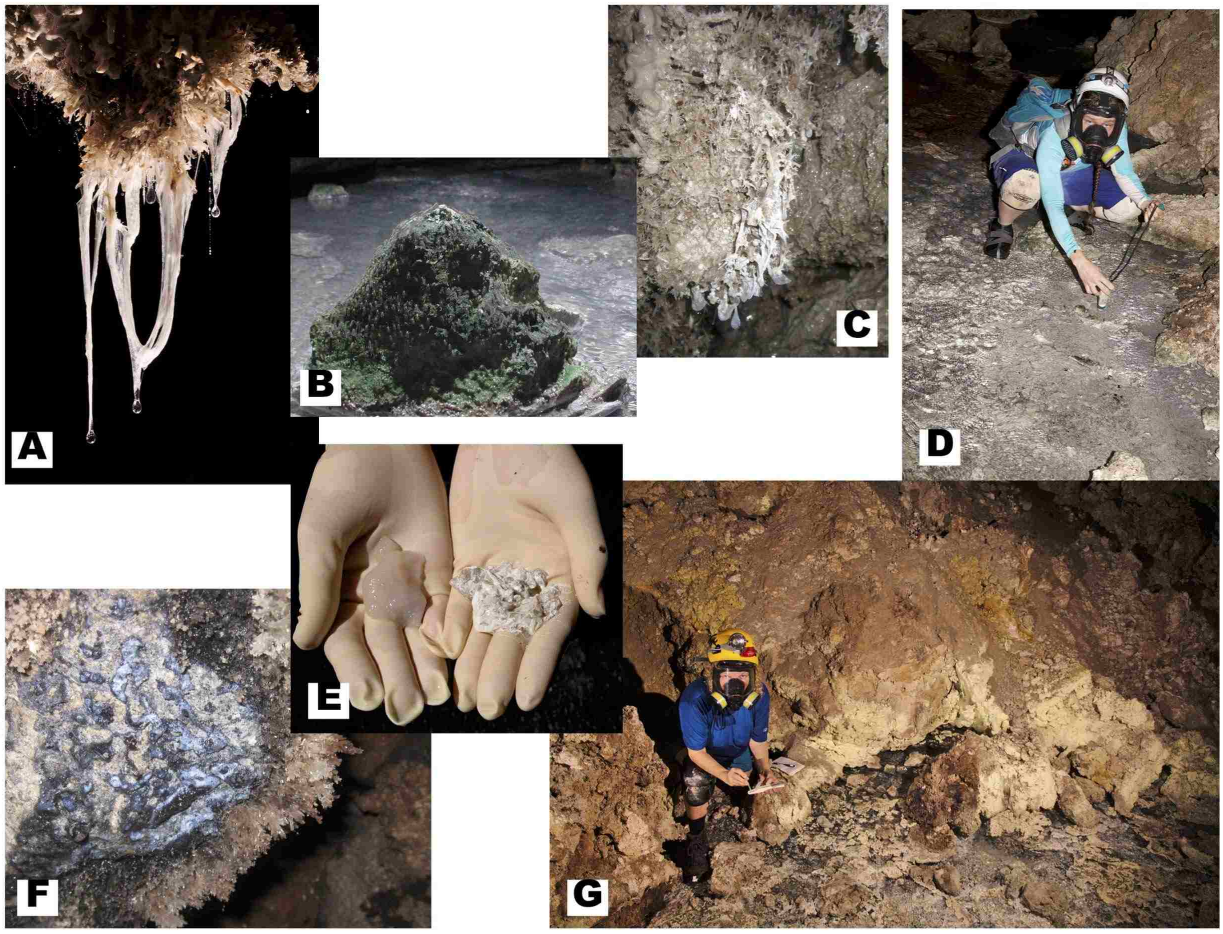


Figure 1. Microbial communities in Cueva Villa Luz. (a) Snottites, (b) green biofilms, (c), gypsum paste, (d) white microbial colonies on stream rocks, (e) phlegmballs (on gloves), (f) blue/black biovermiculations, (g) sulfur deposits (brighter yellow areas). Photos by Kenneth Ingham.

Cave Sample IDs	Sample Type	Location
CVLA	Green Rock	Phlegma Spring
CVLB	White Sulfur Rock	Phlegma Spring
CVLC	Snottite	Yellow Roses Entrance
CVLD	Black Sediment	Midway Springs
CVLE	Red Biovermiculations	Midway Springs
CVLF	Brown/Black Biovermiculations	Ragu Passage
CVLG	Red Biovermiculations	Ragu Passage
CVLH	Blue/Black Biovermiculations	Midway Springs
CVLI	Snottite	Snot Purgatory
CVLJ	Phlegmballs	Slot Spring
CVLK	Phlegmballs	Slot Spring
CVLL	Phlegmballs	Phlegma Spring

Table 1. Overview of CVL samples and corresponding metadata.

Cave Sample IDs	Sample Type	Color	Overburden (m)	Cave Type	Geology Description	Corresponding Surface Sample
PARA0601A	Mineral Crust	Yellow	12.5	Carbonate	Kaibab Limestone	PARASoilG
PARA0601B	Mineral Crust	White	12.5	Carbonate	Kaibab Limestone	PARASoilG
PARA0901A	Soil Guano	Brown	25.9	Gypsum-Carbonate	Kaibab Limestone	PARASoilB,F
PARA1001A	FMD	Pink	7.9	Gypsum-Carbonate	Kaibab Limestone	PARASoilJ
PARA2204A	FMD	Pink	107.6	Carbonate	Red Wall Limestone	N/A
PARA2204B	Microbial Mat	White	107.6	Carbonate	Red Wall Limestone	N/A
PARA2204C	FMD	Red	107.6	Carbonate	Red Wall Limestone	N/A
PARA2204D	Mineral Crust	Brown	107.6	Carbonate	Red Wall Limestone	N/A
PARA2204E	FMD	Pink	107.6	Carbonate	Red Wall Limestone	N/A
PARA2204F	FMD	Pink	107.6	Carbonate	Red Wall Limestone	N/A
PARA2204G	Mineral Crust	Yellow	107.6	Carbonate	Red Wall Limestone	N/A
PARA2206A	Montmorillonite Clay	Black	3.7	Consolidated Material	Moenkopi Mudstone	PARASoilA
PARA3504A	FMD	Red	439.2	Carbonate	Red Wall Limestone	PARASoilH-I
PARA3504B	FMD	Orange	439.2	Carbonate	Red Wall Limestone	PARASoilH-I
PARA3504C	Cricket Marshmallow	White	439.2	Carbonate	Red Wall Limestone	PARASoilH-I
PARA3504D	Microbial Mat (Mud)	White	439.2	Carbonate	Red Wall Limestone	PARASoilH-I
PARA3504E	FMD	Red	439.2	Carbonate	Red Wall Limestone	PARASoilH-I
PARA3504F	Mineral	Yellow	439.2	Carbonate	Red Wall Limestone	PARASoilH-I
PARA3504G	MoonMilk	White	439.2	Carbonate	Red Wall Limestone	PARASoilH-I
PARAAHFA	Cave Soil	Brown	16.5	Basalt Block	Pleistocene Basalt	PARASoilE
PARAAHFB	Mineral	Yellow	16.5	Basalt Block	Pleistocene Basalt	PARASoilE
PARAMC2107A	Mineral Crust	Yellow	18.9	Carbonate	Kaibab Limestone	PARASoilD
PARAMC2107B	Mineral Crust	Yellow	18.9	Carbonate	Kaibab Limestone	PARASoilD
PARAMC2107C	Microbial Mat	White	18.9	Carbonate	Kaibab Limestone	PARASoilD
PARASVMineA	Azurite Mineral	Blue	37.2	Mine	Red Wall Limestone	PARASoilK

Table 2. Summary of PARA samples and metadata. Discernable are dissimilar sample types among the different caves and surface overlying surface samples investigated above several of the caves and a mine.

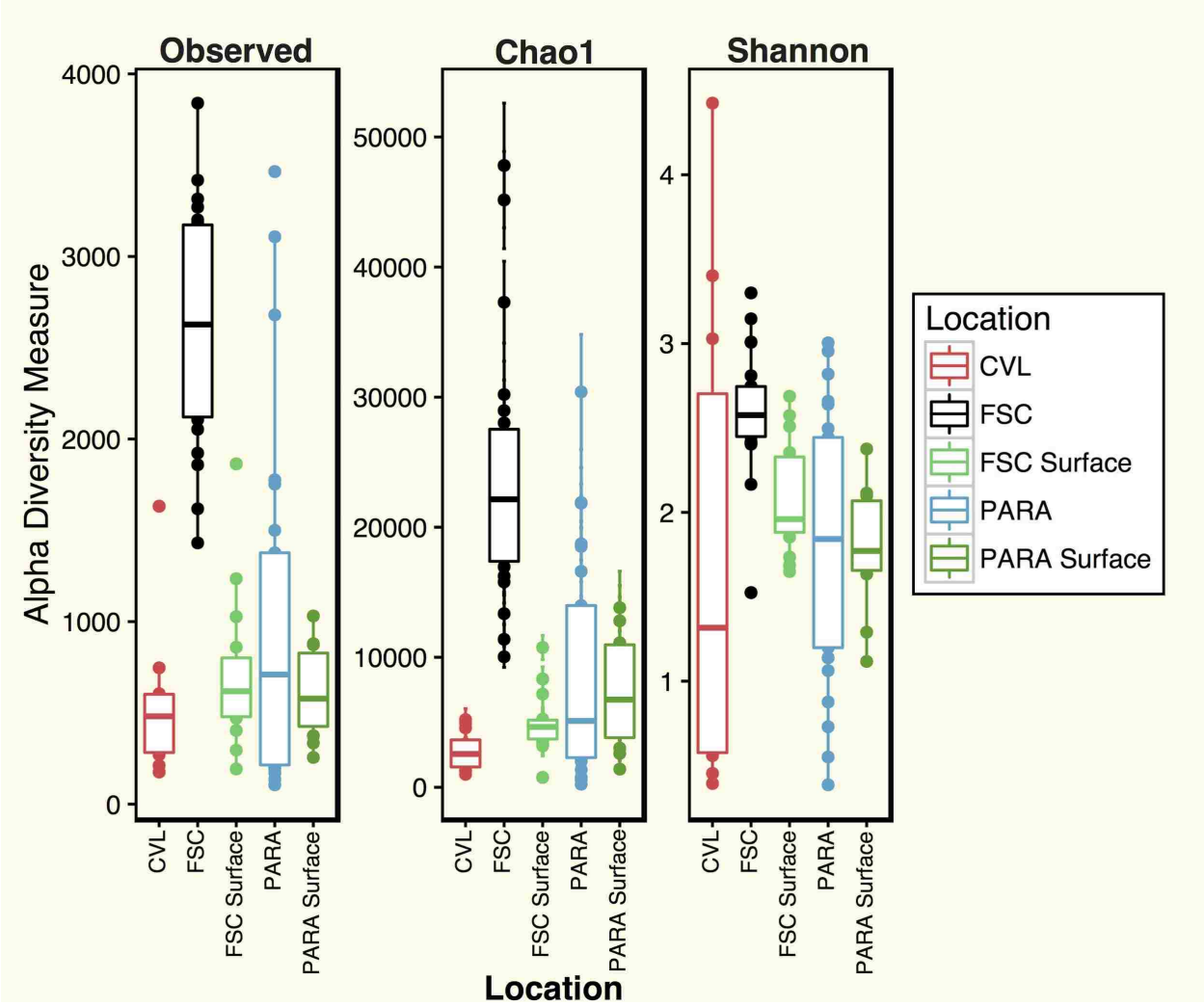


Figure 2. Alpha diversity measurements showing that overall FSC samples have the most archaeal diversity. While the CVL Shannon index boxplot medium showed the most unevenness, indicating some samples are dominated by a few OTUs.

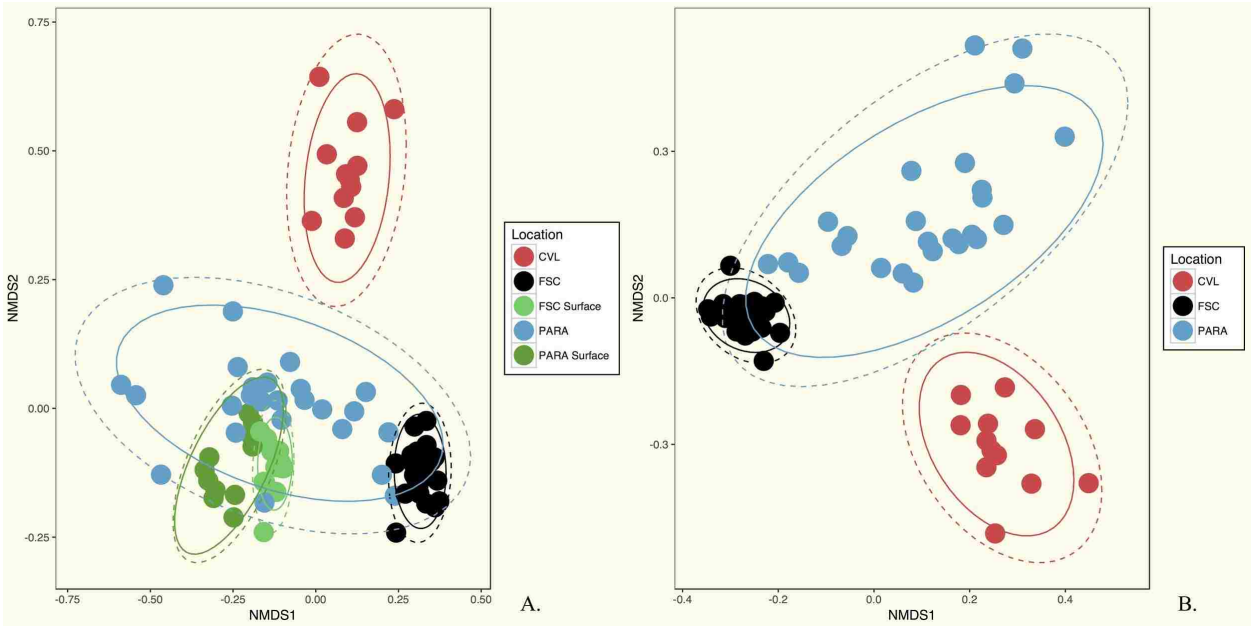


Figure 3. NMDS archaeal OTU site comparison with (a) and without (b) surface samples from FSC and PARA. Samples mostly cluster by site and sample location (i.e. surface and cave). PARA demonstrated the most dissimilarity and many samples clustered with surface soils.

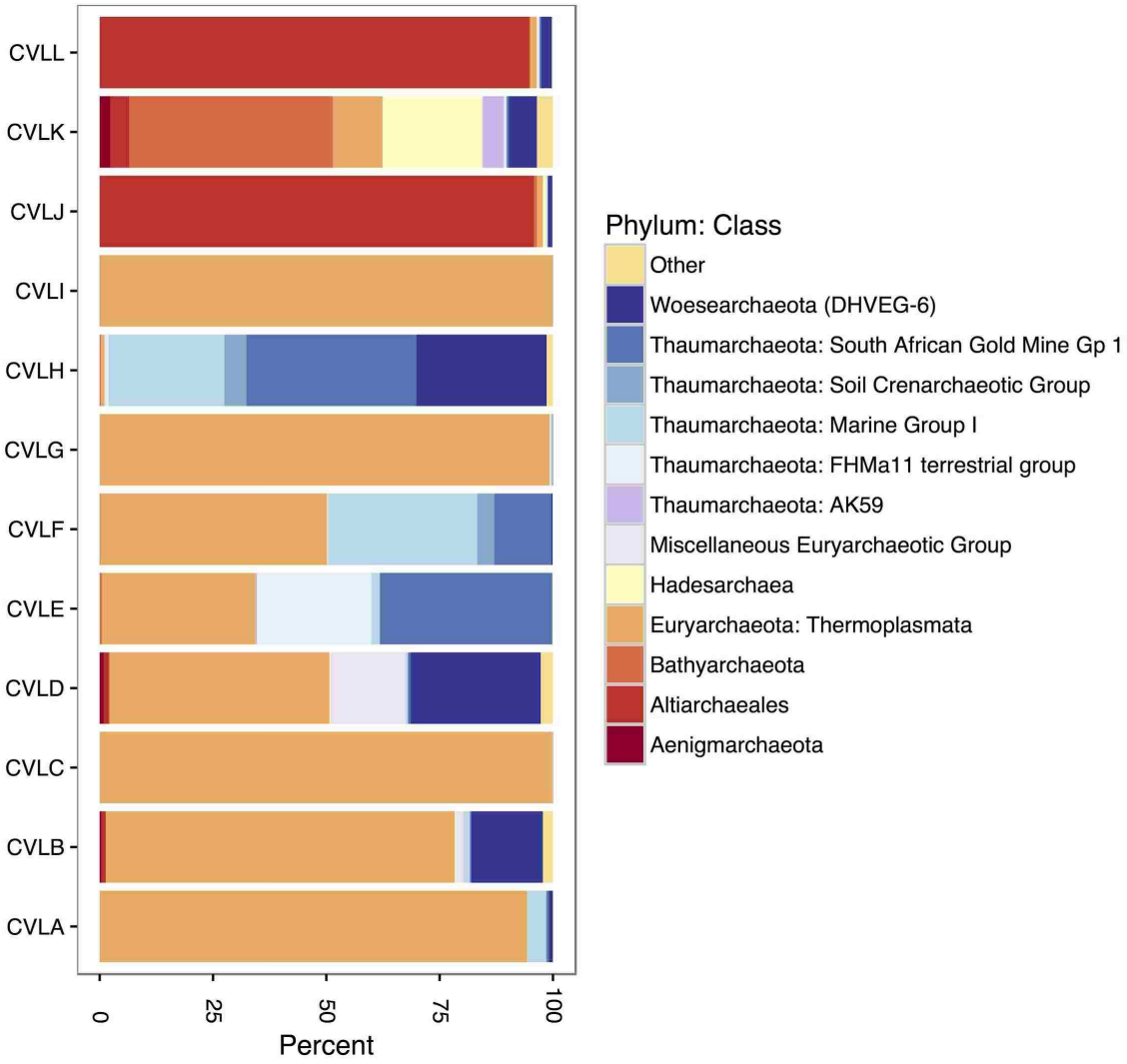


Figure 4. Overview of archaeal phyla and classes found in each sample analyzed at CVL.

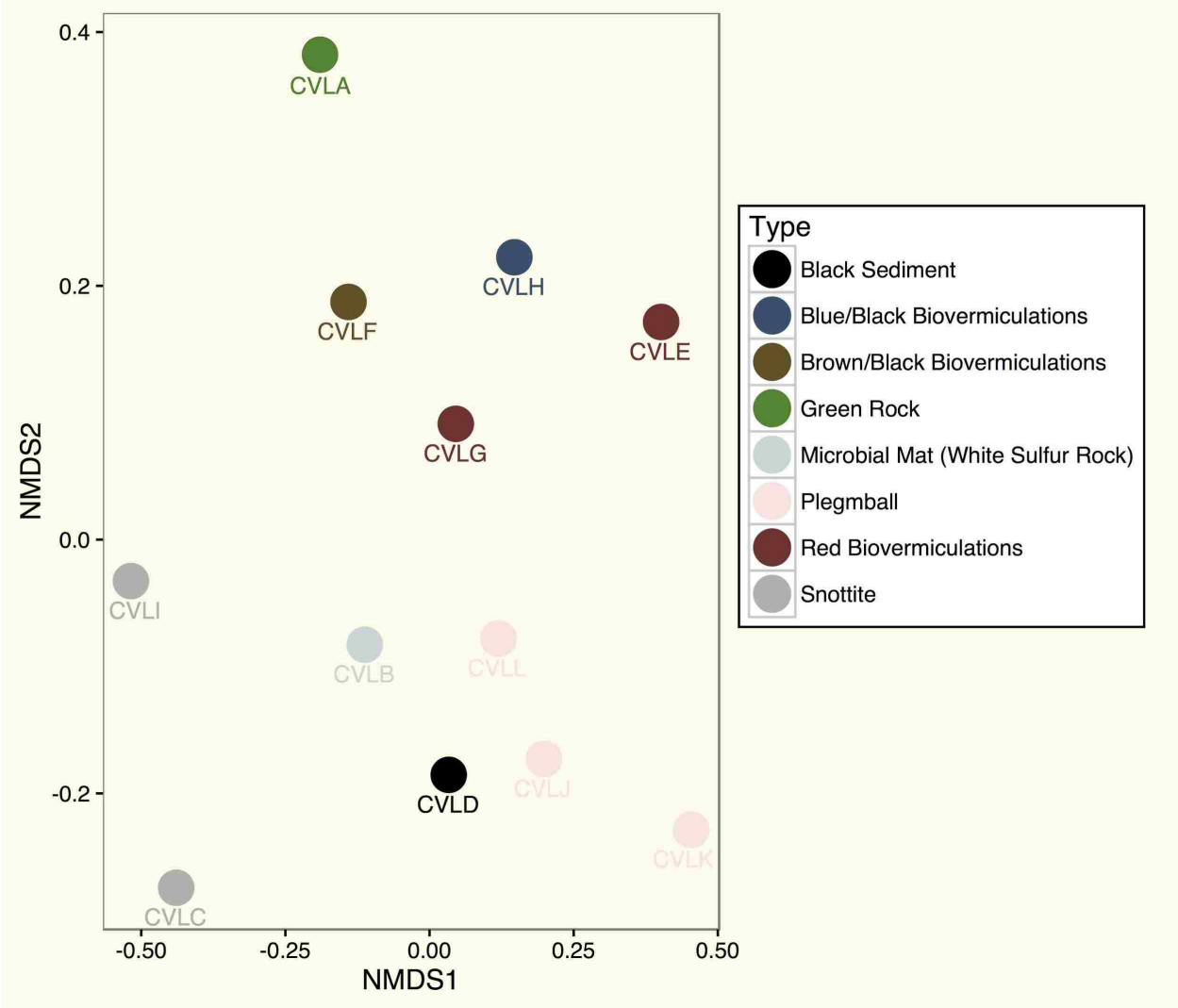


Figure 5. NMDS plot of CVL samples showing OTU dissimilarity within the site.

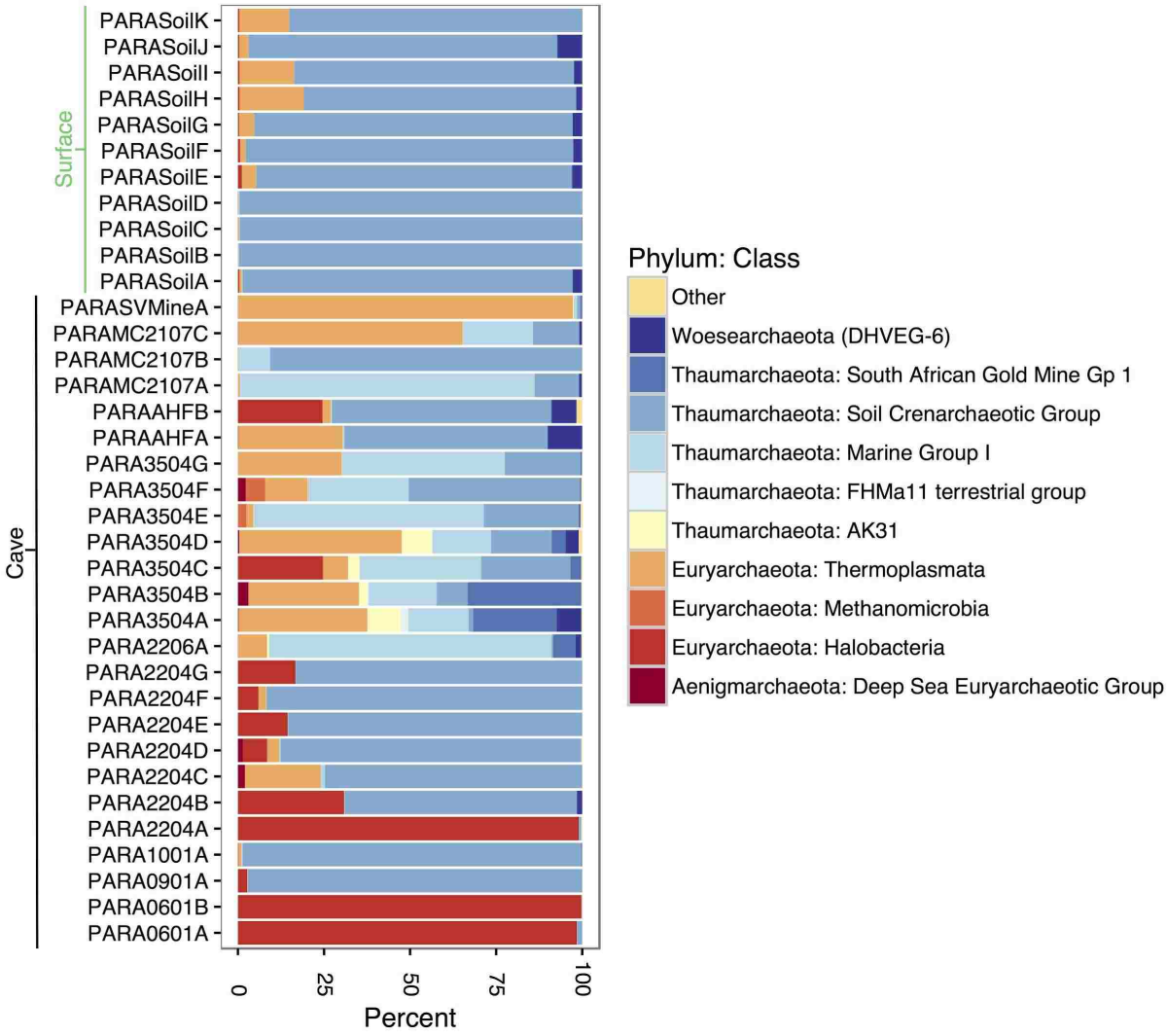


Figure 6. Taxonomy summary of archaeal phyla and classes found within PARA samples.

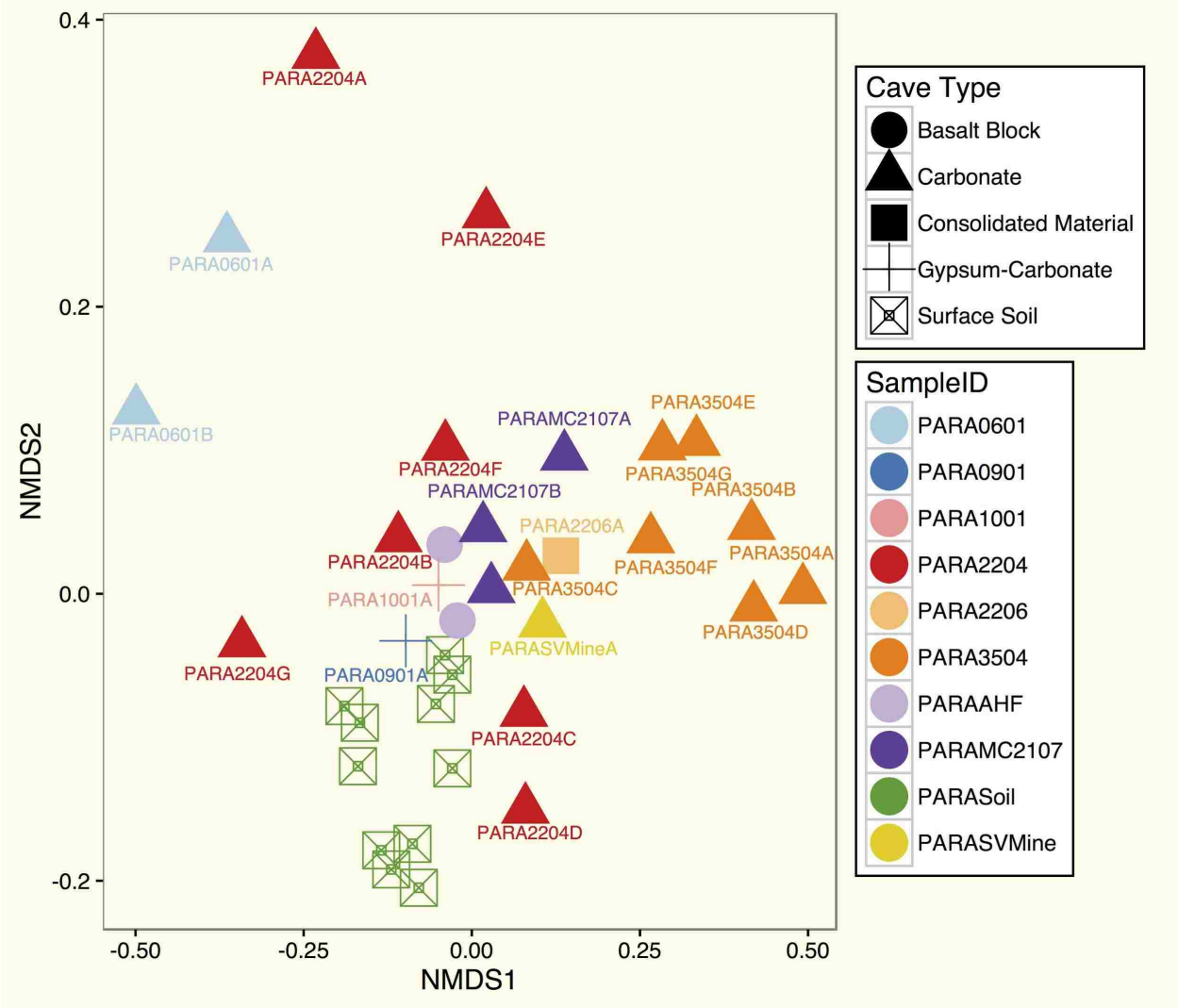


Figure 7. NMDS plot of PARA only samples illustrating OTU variability within the site and sample clustering with surface samples.

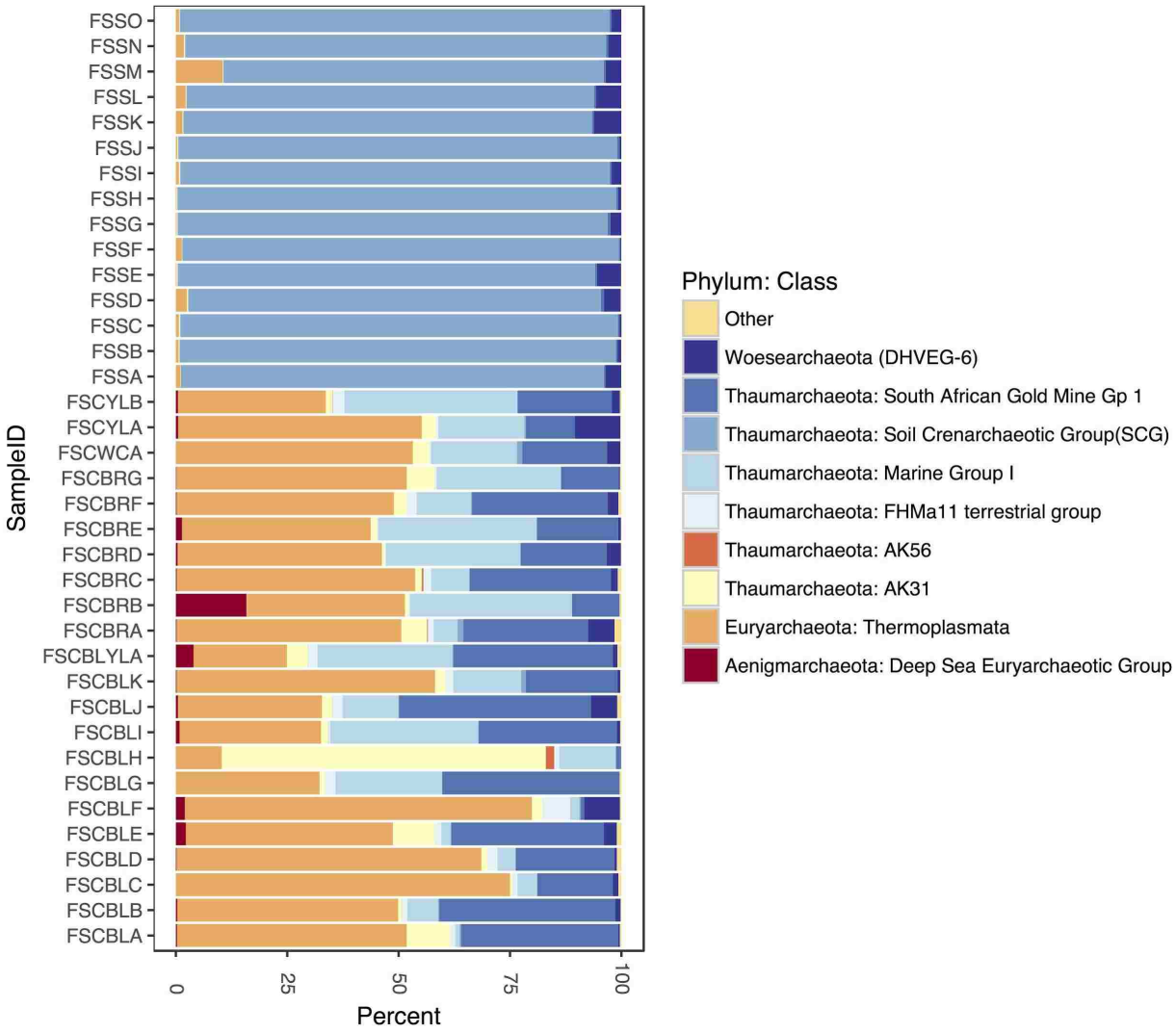
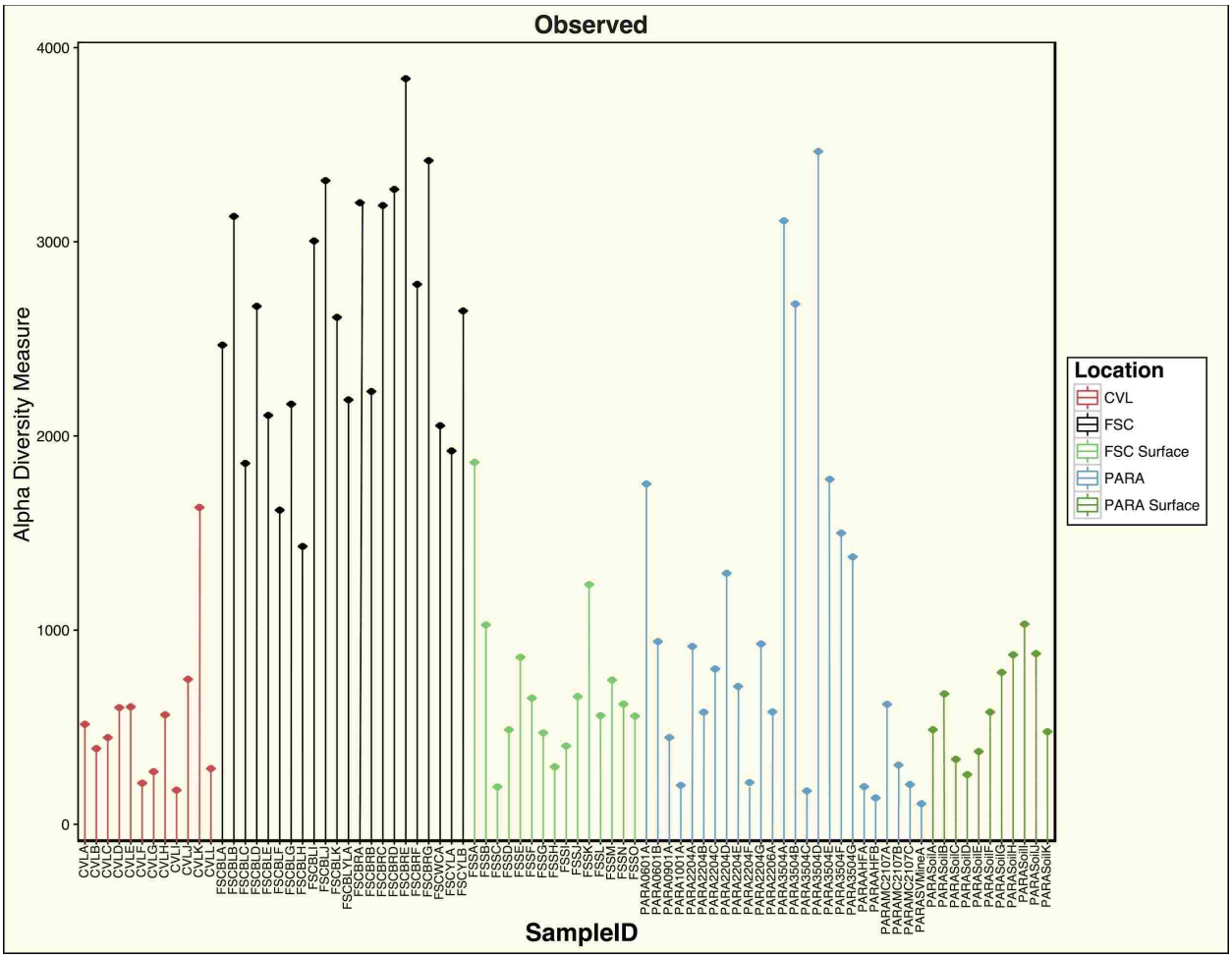


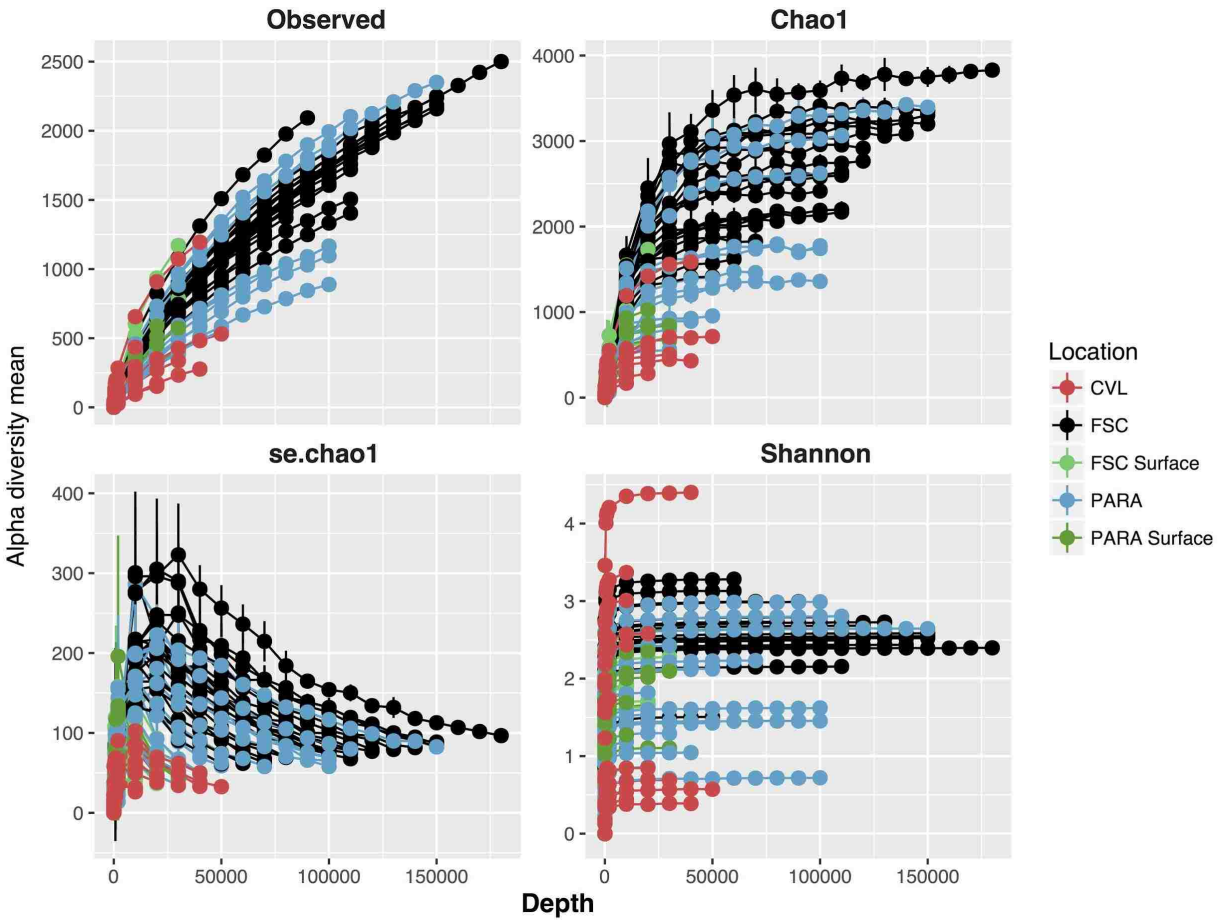
Figure 8. Summarization of archaeal taxonomy previously found in FSC (Kimble *et al.*, unpublished), broken down by phyla and classes in each sample.

Location	Core OTUs across 80% of samples	Average% of All CVL Reads in Core	Average% of All PARA Reads in Core	Average% of All FSC Reads in Core
All Sites:				
Cave	4	5.0	28.1	30.6
CVL and PARA:				
Cave	4	5.0	28.1	N/A
FSC and CVL:				
Cave	5	5.1	N/A	49.1
PARA and FSC				
Cave	5	N/A	28.2	41.1
CVL Only:				
Cave	24	80.2	N/A	N/A
PARA Only:				
Cave	14	N/A	66.4	N/A
FSC Only:				
Cave	160	N/A	N/A	94.7

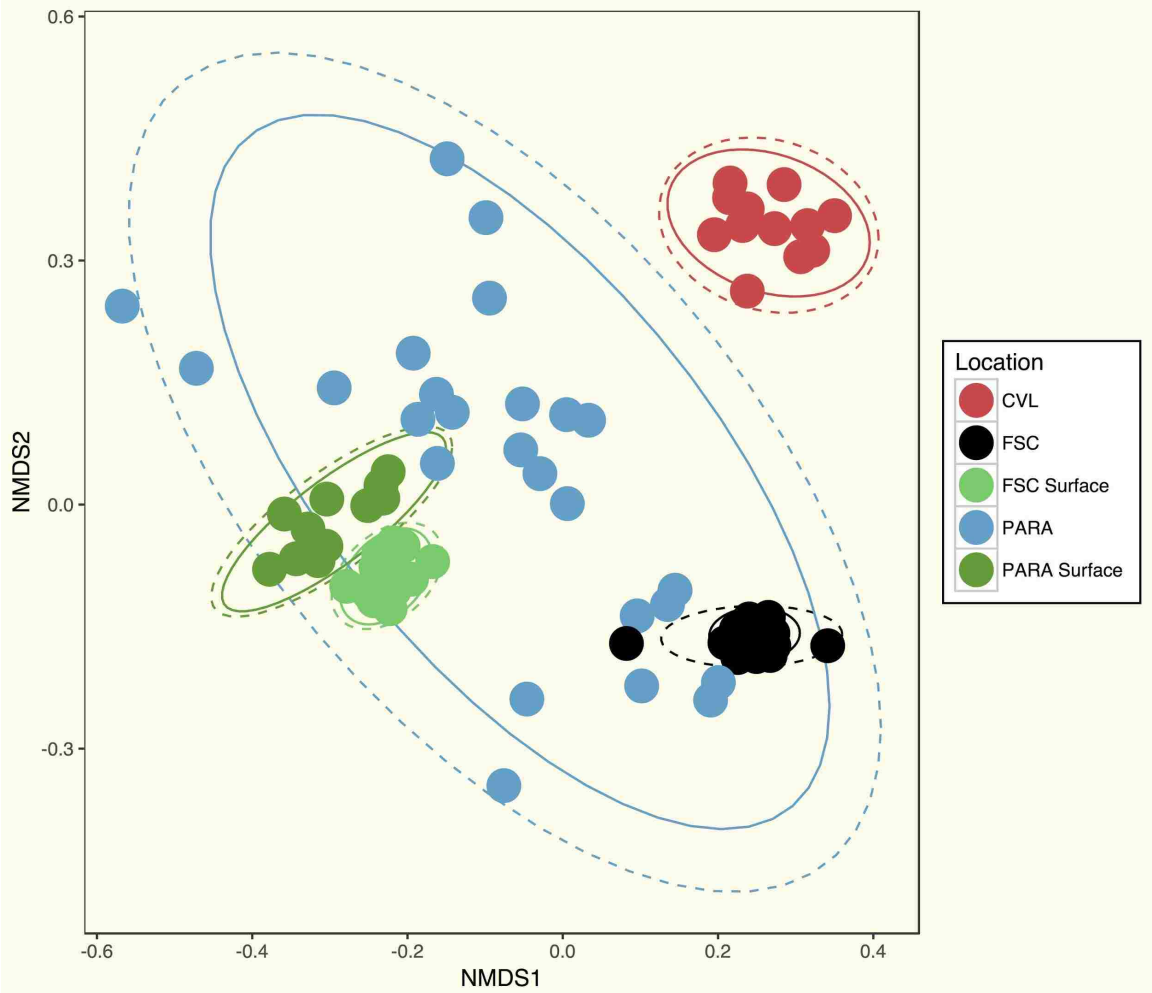
Table 3. Summary of archaeal core microbiome shared OTUs and read abundances found in not less than 80% of all samples.



Supplementary Figure 1. Observed Alpha diversity measurement given by individual samples.



Supplementary Figure 2. Alpha diversity estimated sample-based rarefaction richness, plotted by depth.



Supplementary Figure 3. NMDS plot of OTUs related to unassigned reads. Because primers used also amplified some bacterial groups, unassigned OTUs could be associated with either bacteria or archaea.

Chapter 5: Conclusions and Future Studies

This research has expanded our knowledge of N cycling in a moderately deep arid-land carbonate cave. The abundance and diversity of N cycling genes described in chapter 2 provides evidence that the microbial communities present are cycling this essential nutrient in an environment that likely receives limited input from meteoric waters from the vadose zone. An important N cycle pathway identified in FSC FMD is aerobic nitrification, an energy-producing pathway. Key players associated with nitrification were the AOA *Thaumarchaeota* and AOB/NOB *Nitrospirae*, with other AOB hypothesized to play a lesser role, with respect to the oxidation of ammonium.

These findings have broad implications for understanding how organisms in conditions of extreme oligotrophy acquire and cycle this essential element. For example, these findings may provide insights into potential nitrogen cycling on an extraterrestrial planet such as Mars. The Martian atmosphere has low levels of N_2 gas, and while ozone is present, there is not a protective ozone layer, which results in significantly higher doses of ultraviolet radiation on the terrestrial Martian surface than on Earth (Cockell *et al*., 2000). Consequently, it is probable that life, or evidence of its existence, would more likely be found in an underground terrestrial environment such as lava caves (Boston *et al*., 1992) on extraterrestrial bodies. Numerical models assessed by Williams *et al*. (2010) suggest that it is plausible that cave ice could exist in lava caves on Mars and that these areas could be of potential astrobiological importance.

Chapter 3 results from 16S rRNA gene sequencing provided evidence that archaeal and bacterial communities in FSC FMDs and overlying surface soils are

dissimilar in many ways. Core microbiome analyses showed there was little overlap between archaeal and bacterial OTUs in FMDs and overlying surface soils. Taxonomic results showed high relative abundances of *Firmicutes* and *Bacteroidetes* in soils, while *Nitrospirae*, GAL15, and *Chloroflexi* were dominant phyla in the FMDs. We hypothesize that elevated reads related to *Nitrospirae*, the second most abundant phylum found in the FMDs, is due to their adaptation to oligotrophic conditions (Nogueira and Melo, 2006; Kim and Kim, 2006) and presumed spatial association with AOA, previously observed between AOB and NOB (Maixner, 2009). This would provide these organisms with a source of nitrite for the energy yielding second step of nitrification, or NOB could potentially degrade cyanate to cross-feed AOB. There were also several bacterial groups (e.g. RBG-1 (Zixibacteria), *Latescibacteria*, and SBR1093) identified in the FMDs that were mostly absent in overlying surface soils. This was the first study to characterize, compare, and contrast both archaeal and bacterial communities present in a cave and overlying surface soils. Based on our results, we hypothesize that the low-nutrient FMDs select for organisms with metabolic pathways that favor oligotrophy in this moderately deep arid-land carbonate cave.

Our comparisons of FSC, PARA, and CVL caves in chapter 4, revealed that there were many archaeal dissimilarities (OTUs and corresponding taxonomy) between the sites. Multidimensional scaling comparison illustrated that FSC and CVL clustered by site, while PARA had loose clustering and several outliers. Although hydrogen sulfide-rich CVL (Hose *et al*., 2000) clustered by site when compared to FSC and PARA, differences among samples were apparent in a single site multidimensional scaling plot and taxonomy results. This was the only site that had elevated abundances of phyla

Altiarchaeales and *Bathyarchaeota*. Several cave and sample types in PARA were found to have high relative abundances of Soil Crenarchaeotic Group (SCG), and core microbiome analyses showed that there were many samples that shared OTUs with surface soils. Results showed that cave types, host-rock geochemistry, and surface input (e.g. depth below the surface) influence archaeal communities in caves.

Future Research

A logical next question to explore would be to ask if these genes are actively being expressed in a subsurface system such as FSC. The answer to this question could be provided by looking at the metatranscriptomics associated with the FMD found in Snowy River passage and other subterranean environments, such as a lava cave. Information provided using such technology would fill gaps in our understanding of how N is cycled in caves by providing critical knowledge, such as potential N input through the expression of nitrogenase genes and N loss through denitrification reduction pathways. Previous research in Spider and Lechuguilla caves, using epifluorescent microscopy, estimated microbial communities in FMD to be around 1.6×10^7 cells per cm^3 and found that many of these cells are metabolically active (e.g. an average of approximately 15% in Black/Brown FMD) (Northup, 2002; Spilde *et al*., 2005). Based on this information, it is reasonable to conjecture that some microbes in FSC FMD are metabolically active and that corresponding functional activity could be studied using a metatranscriptomic approach. Future investigations will consider such an approach to better understand active N cycling related gene expression in caves.

References

Boston, P.J., Ivanov, M.V., and McKay, C.P. (1992). On the possibility of chemosynthetic
ecosystems in subsurface habitats on Mars. *Icarus*. 95, 300-308.

Cockell, C.S., Catling, D.C., Davis, W.L., Snook, K., Kepner, R.L., Lee, P., and McKay, C.P. (2000). The ultraviolet environment of Mars: biological implications past, present, and future. *Icarus*. 146, 343-359.

Hose, L. D., Palmer, A. N., Palmer, M. V., Northup, D. E., Boston, P. J., and DuChene, H. R. (2000). Microbiology and geochemistry in a hydrogen-sulphide-rich karst environment. *Chem. Geol.* 169, 399-423. doi: 10.1016/S0009-2541(00)00217-5

Kim, D. J. and Kim, S. H. (2006). Effect of nitrite concentration on the distribution and competition of nitrite-oxidizing bacteria in nitratation reactor systems and their kinetic characteristics. *Water Res.* 40, 887-894. doi: 10.1016/j.watres.2005.12.023

Nogueira, R. and Melo, L. F. (2006). Competition between Nitrospira spp. and Nitrobacter spp. in nitrite-oxidizing bioreactors. *Biotechnol. Bioeng.* 95, 169-175. doi: 10.1002/bit.21004

Northup, D. E. (2002). Geomicrobiology of Caves. Ph.D. dissertation: University of New Mexico.

Spilde, M. N., Northup, D. E., Boston, P. J., Schelble, R. T., Kathleen, E. D., Crossey, L. J., et al. (2005). Geomicrobiology of cave ferromanganese deposits: a field and laboratory investigation. *Geomicrobiol. J.* 22, 99-116. doi: 10.1080/01490450590945889

Williams, K. E., McKay, C. P., Toon, O. B., and Head, J. W. (2010). Do ice caves exist in Mars? *Icarus*. 209, 358-368. doi: 10.1016/j.icarus.2010.03.039