

Microbial Taxonomic Variation Across Hydrothermal Mats at Lake Kleifarvatn, Iceland

Laidlaw Research Summer Report

Felecia Selgrade

Research done with Dr. Sarah Stewart Johnson, Dr. Zachary K. Garvin, Dr. Shauna Bennett, At
the Johnson Biosignatures Lab in Georgetown University

August 14, 2026



GEORGETOWN UNIVERSITY
Center for Research & Fellowships



Abstract:

Hydrothermal systems serve as analogs for planetary environments offering insight into extraterrestrial habitability. Two microbial mat samples were collected near Lake Kleifarvatn, Iceland, by NASA's Goddard Instrument Field Team (GIFT), at differing distances from an inactive hydrothermal vent. DNA was extracted and sequenced for downstream metagenomic analyses. Taxonomic classification of the sequenced samples revealed differences associated with distance from the hydrothermal vent source. The proximal sample contained less diverse tax and a higher abundance of chemolithotrophic organisms reflecting reliance on vent-derived inorganic chemistry. Conversely, the distal sample contained a higher abundance of certain phyla and classes that are associated with carbon and nitrogen cycling. These results indicate large selective pressure on microbial communities depending on proximity to hydrothermal activity at just meters-scale, even within a contiguous microbial mat. Future work will incorporate metagenome-assembled genomes to confirm functional sulfur cycling in the mats alongside additional samples to enable statistical comparisons of community diversity across sites.

Introduction:

Through the support of the Laidlaw Foundation and the Johnson Biosignatures Lab at Georgetown I was able to continue research with postdoctoral fellow Zachary Gavin. When I joined the Johnson lab in my second year at Georgetown, I had no background in astrobiology or geochemistry, but I was eager to explore other areas of biology. During my time in the lab, I worked with Giovanni Calvillo Castrejon and Gisselle Gastelum in learning about different microbial communities in evolving geochemical conditions. We worked together on soil microbial mats from Lake Kleifarvatn in southwest Iceland. This lake is a hydrothermally fed, closed-basin system in which acidic (pH ~5), ~100°C vent fluids mix with cooler (~17°C),

alkaline lake water, resulting in an overall acid-neutral environment. This mixing creates a favorable habitat for a wide range of microorganisms adapted to mineral-rich conditions, including exposure to sulfates, Fe-oxides, and Ti-oxides (Markusson, Stefansson, 2011).

Iceland is located on the Mid-Atlantic Ridge (a divergent plate boundary) above a mantle hotspot, serving as one of the only locations on Earth where the ridge is above sea level (subaerial). This results in a unique geological setting of basaltic rock, extensive volcanic activity, and hydrothermal systems. The environments hosted on the island are widely used as planetary analogs due to the geology, chemistry, and environmental conditions resembling those of other planets and moons. Lake Kleifarvatn is of particular interest due to the diversity of hydrothermal systems present in and around the lake itself (Table 1), which could be representative of both modern (inactive) and ancient (active) Mars. Research by Dr. Garvin and a team of NASA scientists is focused on how organic molecules and specific mineral patterns, which only form in these hydrothermal settings in the presence of water and heat, can shape microbial composition as well as nutrient cycling within this extreme environment.

Soil samples were collected as part of a field expedition organized by the NASA Goddard Instrument Field Team (GIFT) in different activity sections of Lake Kleifarvatn. Samples were collected, by Dr. Garvin, from both actively venting and relict (inactive) sections of the lake to capture a series of environments representing different stages of evolution and a range of geochemical conditions across the lake. Sites were classified by the “activity” of the environment, indicating whether or not hydrothermal fluids are still flowing through the system. Critically, certain microbial communities could rely on the mineralogical and geochemical conditions provided by the different activity levels for their energy metabolisms. This means that

differing activity levels, shown in Figure 1, affect the microbial communities that can be supported.



Figure 1. Vent and Deposit Activity Gradient at Lake Kleifarvatn. Comparison of hydrothermal vent deposits across the activity gradients at Lake Kleifarvatn, where vent fluids are active within the shallow lake, recently inactive and exposed on the shoreline, inactive after longer term dormancy further inland of the shore.

During the school year we worked with two mat samples, collected from different sites within the same exposed sandbar area. Both had recently been submerged but were uncovered as the water level declined, classifying them as “recently inactive”. The first sample (“Proximal”) was located closer to the presumed hydrothermal source (i.e., subaqueous vent); thus, it likely experienced greater, more direct exposure to the hydrothermal fluid and its chemistry. The second sample (“Distal”) was collected further from this source, and it was expected to have been less directly influenced by the vent activity. Since the environment plays a key role in microbial community composition, we analyzed the two specific mat samples and performed DNA extraction and shotgun metagenomics analyses. Looking at the microbial communities that make up Lake Kleifarvatn would provide valuable insight into the types of microbes and chemical process that could have existed on Mars (Arroyo & Barry, 2026). I was interested in

analyzing the microbial communities for metabolisms associated with geothermal chemistries, such as methanogenesis and sulfur cycling, while also comparing microbial composition between the two samples to determine the effects of vent proximity on the communities. Knowing more about these extremophiles that tolerate and thrive under intense conditions can further open new opportunities for research of biotechnological applications in the areas of wastewater treatment and agriculture (Bolger et al., 2014; Thakur et al., 2022).

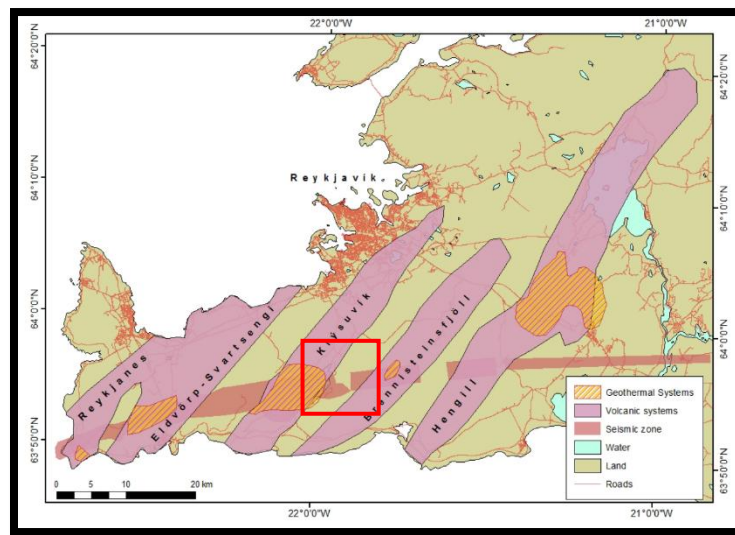


Figure 2. Geothermal and Volcanic Systems of the Reykjanes Peninsula, Iceland. Map of activate volcanic systems, geothermal areas and seismic zones across the Reykjanes Peninsula, Iceland. Lake Kleifarvatn sits within the volcanic system highlighted by the red box.

Research Methodology:

Collection of Samples

As described previously, microbial mat samples were collected near Lake Kleifarvatn, Iceland during the NASA GIFT field expedition by Dr. Garvin (Fig.3a and 3b). The mat samples described here are referred to as Proximal and Distal based on their relative radial positions around the presumed hydrothermal vent (1.7 m and 2.8 m, respectively). Samples were handled

with sterile tools and vials, immediately placed on dry ice on site, and carried back to the US via cryoshipper in liquid nitrogen. Samples were promptly stored in the lab in a -80°C freezer prior to analysis.

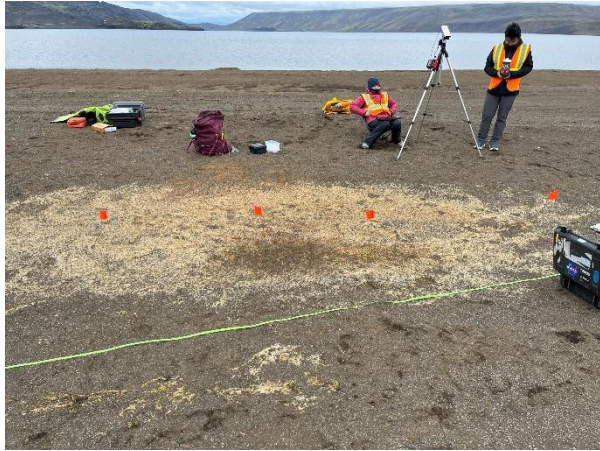


Fig 3a. (left)

**Fig 3b.
(right)**



Figure 3a. Sample sites were located on a sandbar recently exposed due to falling lake levels.

Figure 3b. Close up of mat's orange color reflecting potential hydrothermal chemistry and iron/sulfur metabolisms.

DNA Extraction and Sequencing

During the school year our group worked on DNA extraction of these two samples using the DNeasy PowerSoil Kit (Qiagen, Carlsbad, CA). We followed the manufacturer's procedure provided by the kit. Due to high yields, the extracted DNA was diluted prior to measuring their concentrations via a Qubit Fluorometer and dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA). After dilution of both samples, Proximal sample had a DNA concentration of 15.6 ng/uL and Distal sample had a DNA concentration of 24.7 ng/uL.

Following extraction, the DNA from both samples were sent to the Rush University Genomics and Microbiome Core Facility for Illumina sequencing. Shotgun metagenomics were performed by sequencing the DNA (2 x 150 base pairs) on a single flow cell lane with a NovaSeq X System (Illumina). The Illumina sequencing generated short paired-end reads (150 base pairs), which were provided to us in a FASTQ file format.

Quality-control of the reads were performed via trimming and filtering using the Trimmomatic program to remove low-quality bases and adapter sequences, which resulted in a cleaner set of high-quality reads (Bolger et al., 2014). During the school year our lab worked closely with Professor Bennett and focused on the viruses in the mats, using a suite of viral analysis programs : Virsorter2 (Guo et al., 2021), GeNomad (Camargo et al., 2023), iPHoP (Roux et al., 2023). This data analysis work started to prepare me for the future work I continued in my Research summer with Laidlaw.

Analysis of Microbial Community

During my Laidlaw Research summer, I decided to shift my focus from viruses to bacteria, aiming to characterize the Metagenome-Assembled Genomes or MAGs recovered from the sequenced DNA. MAGs are near-complete bacterial genomes reconstructed by grouping together (“binning”) contigs, which are longer contiguous DNA sequences constructed from the short 150 bp reads. Metagenomics in general is a valuable technique for studying extreme environment interactions by looking at communities without the need for culturing. In environments like Lake Kleifarvatn, where most microbes cannot be grown in a lab, this allows us to identify species and understand their metabolic capabilities directly from their genomic information. MAGs in particular are beneficial because they enable us to analyze the encoded

metabolic capabilities within individual organisms, which can help to understand dominant processes involved in nutrient cycling and other microbe-environment interactions.

With this new focus, I started by reassembling contigs from the trimmed reads. MetaSPAdes identifies patterns across reads using de Bruijn graph assembly to overlap sequences of DNA, allowing for the generation of contigs (Nurk et al., 2017). Contigs can provide immense insight into bacterial taxonomy, gene presence and abundance, and downstream generation of MAGs.

The contigs were filtered and reformatted using Anvi'o v-9, retaining only sequences >300 bp (Eren et al., 2015). Filtering out contigs shorter than 300 bp helps improve the quality and reliability of downstream analysis. Very short contigs are often assembly fragments with too little sequence information to be confidently used for taxonomic classification or binning into MAGs. For the proximal sample this process reduced the assembly from 4,039,082 contigs to 2,597,043 contigs, removing about 35.7% of contigs.

During this point of the research, Postdoc Zack and I decided to characterize the composition and diversity of the mat communities by taxonomic classification of the bacterial and comparison between samples. For taxonomy, I used Kaiju, which classifies short reads by comparing them against a reference database of protein sequences from microbial and viral genomes, assigning it to a taxonomic identity (Menzel et al., 2016). The *nr_euk* was used as the reference database, comprising Archaea, Bacteria, and viral sequences sourced from the nonredundant NCBI BLAST database. Within the database, taxon identifiers are contained in *nodes.dmp* and *names.dmp* files, which are needed to create the classification summary (Menzel et al., 2016).

Kaiju was run on the trimmed, quality-filtered sequences (forward and reverse) to produce Kaiju outputs files (*kaiju.out*) for both proximal and distal samples. These kaiju outputs were used along with *nodes.dmp* and *names.dmp* files to make a classification summary table for a given taxonomic rank (class and phylum). Counting unclassified reads towards the total number of reads when calculating relative abundance percentages was disabled in order to provide a more accurate picture of the composition of the identifiable microbial community.

The tab-separated output files were converted to excel workbooks for manual manipulation and assessment. The data was imported into R for data analysis and visualization, including the generation of bar plots of relative abundance at the class and phylum levels for bacteria and archaea. This produces a *kaiju.out* file containing the taxonomic classification for each read, which can then be summarized into class- and phylum-level abundance table. Krona plots were also generated based on the Kaiju output. I generated a Krona plot for each sample for a holistic comparison of community taxa.

Results:

Short read classification from Kaiju revealed differences in microbial community composition between the two samples (Fig.4). The Proximal sample primarily consisted of the phyla Pseudomonadota, Chloroflexota, Bacteroidota, and Actinomycetota. The Distal sample exhibited greater diversity at the phylum level, sharing all five of the major phyla within the Proximal sample while also containing a large abundance of Cyanobacteriota (6.1%) and Bacillota (7.5%), which had a much lower abundance in the Proximal sample. The percent of the total reads that were unclassified was 47% for the proximal sample and 61% for the distal sample for both class and phylum taxa. The presence of Cyanobacteriota is notable as members of this phylum can perform oxygenic photosynthesis and nitrogen fixation, which means they have

important roles in carbon and nitrogen biogeochemical cycles. Bacillota bacteria have diverse metabolic capabilities and range from aerobes, facultative anaerobes and strict anaerobes. Most members of the phylum are chemoorganotrophs or chemolithotrophs (Mehdizadeh Allaf & Peerhossaini, 2022). Together, the abundance of these two phyla at greater distance from the vent source could indicate that bacteria further away from the hydrothermal vents encode the ability to perform a distinct set of metabolisms not present in communities closer to the vents, such as photosynthesis and nitrogen fixation. These metabolisms may be less favorable under an environment that receives more heat and chemical influence from hydrothermal fluid. For instance, photosynthesis has been shown to have a temperature limit in hydrothermal systems. Conversely, communities closer to the vent could be more dominated by a less diverse set of chemolithotrophs reliant on the reduced gases, mineral phases or ions supplied by the hydrothermal fluid.

At the class level, the Proximal sample primarily consisted of Gammaproteobacteria, Flavobacteriia, Cytophagia, Betaproteobacteria, Anaerolineae, Alphaproteobacteria, and Actinomycetes. Once again, there was significant overlap in the Proximal and Distal communities, as the Distal samples also contained all eight of these classes at similar abundances (Fig.5). The most notable changes were an 8% drop in Gammaproteobacteria and the addition of Cyanophyceae, Chitinophagia and Bacilli. Cyanophyceae are a class under Cyanobacteria phylum and perform photosynthesis and nitrogen fixation (Mehdizadeh Allaf & Peerhossaini, 2022). Chitinophagia are within the Bacteroidota phylum which breaks down chitin and help recycle dead plant biomass (McKee et al. 2019). Bacilli is a narrower taxon within Bacillota and has similar metabolic functions of nutrient cycling (Argentino et al., 2025). These results further establish the distal samples as containing organisms with metabolisms less dependent on

hydrothermal activity. Photosynthesis and heterotrophic degradation of plant material are both unlikely to be favorable processes to the Proximal communities, which have access to a more diverse array of inorganic energy sources provided by the hydrothermal fluid and its water-rock interactions.

The observed changes in taxonomic composition and diversity from Proximal to Distal are interpreted as a spatial relationship. The bacterial communities closer to hydrothermal vents are constrained to a more thermally and geochemically influenced environment. On the other hand, communities further away have adapted to milder conditions that support greater metabolic diversity and allow less specialized taxa to thrive. This is particularly evident based on the abundance of Cyano, Phyceae and Chitinophagia in only the Distal sample, suggesting that there is more phototrophic and organoheterotrophic activity compared to the proximal sample. Together, these taxa show that the bacterial communities contain organisms that are reliant on different primary energy sources, with Proximal communities more dependent on inorganic compounds supplied by the vent. The Proximal community points towards more vent and chemistry driven resources, with less bacteria for nitrogen fixation or photosynthesis that are likely inhibited by the high temperatures and less favorable in an environment replete with inorganic energy source.

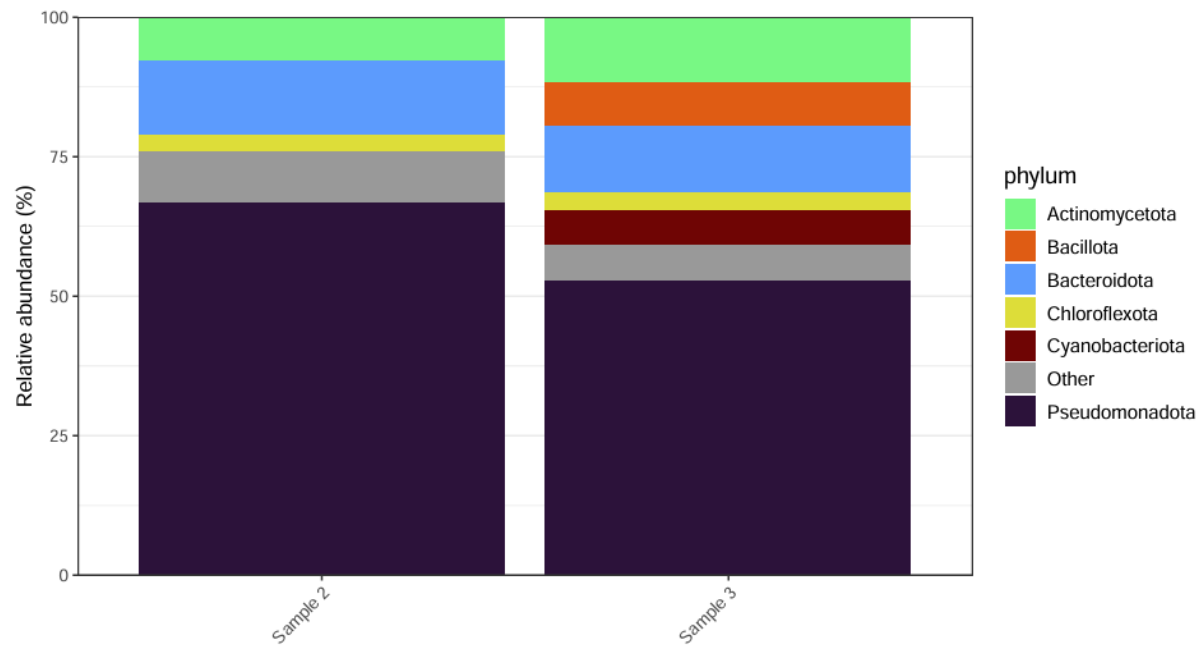


Figure 4. Distal microbial community exhibits greater phylum-level diversity, seen by the increase of Cyanobacteriota and Bacillota. All listed phyla are >2% relative abundance classified by Kaiju.

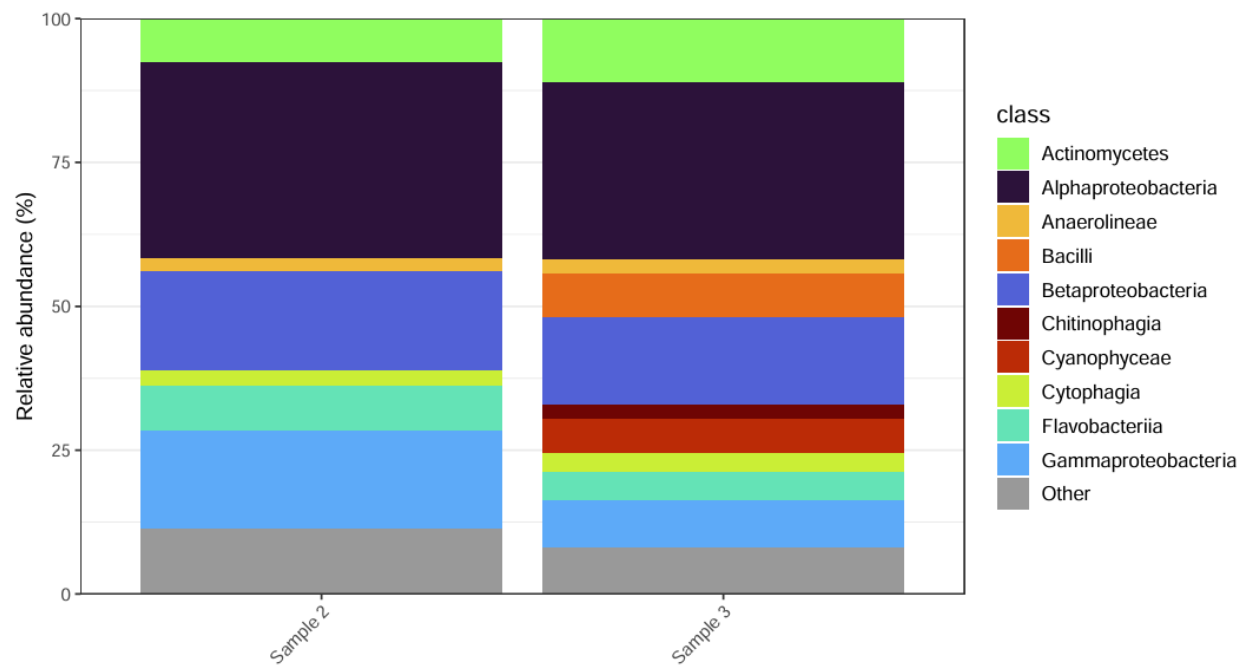


Figure 5. Class-level composition shows an increase of relative abundance of bacteria Cyanophyceae, Chitinophagia and Bacilli in the Distal sample. All listed classes are >2% relative abundance classified by Kaiju.

Krona plots were also made alongside the relative abundance bar plots which provide insight into taxonomic classification as well as percentages of those classifications. The bar plots showed that the Distal sample contained bacteria that do not solely rely on inorganic compounds and the Krona plots support this. (Fig. 6 and 7). In the proximal sample *Sphingomonas* make up 15% of bacteria while in the distal sample they make up only 4% of bacteria. *Sphingomonas* are strictly aerobic and oligotrophic, meaning they can survive with little oxygen and low nutrients (White et al., 1996). They are capable of sulfur-oxidation, pollution breakdown, and organic breakdown. Sulfur-oxidation is a particularly important metabolism in hydrothermal systems as reduced sulfur compounds are abundant in vent fluid due to water-rock interactions. Bacteria capable of oxidizing these compounds can use them as an energy source instead of using photosynthesis, due to the lack of sunlight reaching the bacteria.

Another chemolithotrophic metabolism seen in the Proximal sample is the consuming of methane and using the carbon as an energy source. The Proximal sample had a small percent of *Methylmonas* (1% of bacteria), and one function is to oxidize methane, but this type of bacteria was not found in the Distal sample. This supports that the bacteria in the Proximal sample are influenced by the chemistry of the hydrothermal systems, as methane is commonly emitted from the vents. Instead of the bacteria relying on sunlight or organic matter the *Methylomonas* can use more available resources like vent-driven methane as a carbon source (Tikhonova et al., 2023).

Lastly, another large difference was the decrease of Terrabacteria in the Proximal sample (14% of bacteria) compared to the Distal sample (29% of bacteria). Terrabacteria is a large biological clade made up of Actinobacteria, Firmicutes, Cyanobacteria and Chloroflexi which are highly driven by carbon and nitrogen cycling to provide energy (Poyntner et al., 2026). This pattern is consistent with the phylum and class bar plots discussed earlier and shows that the bacterial communities in the Distal sample are capable of carbon and nitrogen cycling rather than relying on inorganic energy sources supplied by the hydrothermal fluid.

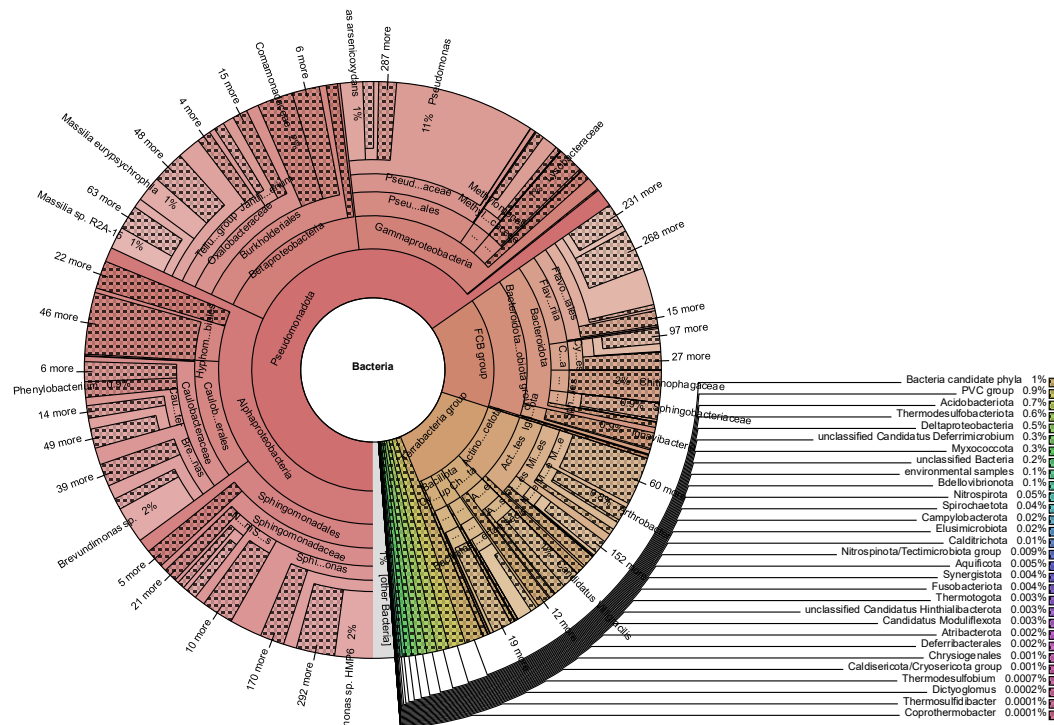


Figure 6. Krona plot of Proximal sample shows higher percentage of chemolithotroph bacteria supporting a vent-driven bacterial community. Plot displays full taxonomic hierarchy and relative abundance (%) of bacterial classifications based on Kaiju read classification.

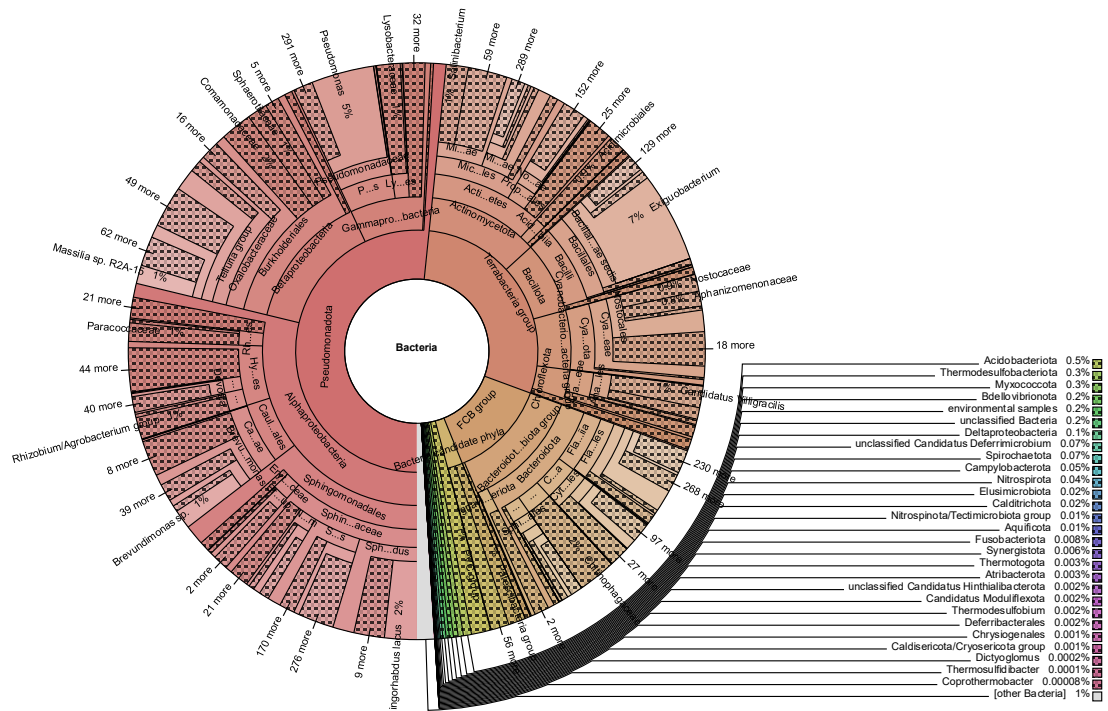


Figure 7. Krona plot of the Distal sample shows higher percentage of bacteria driven by carbon and nitrogen cycling for energy. Plot displays full taxonomic hierarchy and relative abundance (%) of bacterial classifications based on Kaiju read classification.

Future Directions:

Future research will continue with MAGs and analysis of MAGs to determine the true influence of chemolithotrophy near the hydrothermal vent. Reconstructing the genomes would directly identify functional genes involved in sulfur cycling, confirming whether sulfur metabolism is present and active rather than inference from taxonomy.

Alongside MAGs analysis, we would also like to continue research with other samples to provide a larger data set for analysis. A larger data set of samples would allow for Bray-curtis dissimilarity and beta/alpha diversity calculations.

Acknowledgements:

This experience would not have been possible without the mentorship and support of Dr. Johnson, Dr. Garvin, and Dr. Bennett. I am deeply grateful and thankful to the Laidlaw Scholars Foundation and the Georgetown University Laidlaw Scholars Programme for their support during my experience. Finally, I would like to thank the NASA Goddard Instrument Field Team (GIFT) who organized and facilitated the expedition to Iceland. Special thanks to Team ATOMIC members Dr. Dina Bower, Dr. Amy McAdam, and Dr. Casey Honniball.

References

- Argentino, I. C. V., Godoy, M. G., Seldin, L., & Diogo Jurelevicius. (2025). Distribution of bacillota in water and sediments from aquatic environments. *Microbial Ecology*, 88(1). <https://doi.org/10.1007/s00248-024-02482-0>
- Arroyo, Y., & Barry, C. (2026, February 9). *NASA fieldwork spotlight: Iceland*. NASA Science. <https://science.nasa.gov/solar-system/planetary-analogs/nasas-goddard-instrument-field-team-iceland-2025/>
- Bhaya, D., Birzu, G., & Rocha, E. P. C. (2025). Horizontal gene transfer and recombination in cyanobacteriota. *Annual Review of Microbiology*, 79(1), 685–711. <https://doi.org/10.1146/annurev-micro-041522-100420>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.
- Camargo, A. P., Roux, S., Schulz, F., Babinski, M., Xu, Y., Hu, B., Chain, P. S. G., Nayfach, S., & Kyrpides, N. C. (2023). Identification of mobile genetic elements with geNomad. *Nature Biotechnology*, 1–10. <https://doi.org/10.1038/s41587-023-01953-y>
- Eren, A. M., Esen, Ö. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., & Delmont, T. O. (2015). Anvi'o: An advanced analysis and visualization platform for 'omics data. *PeerJ*, 3, e1319. <https://doi.org/10.7717/peerj.1319>
- Guo, J., Bolduc, B., Zayed, A. A., Varsani, A., Dominguez-Huerta, G., Delmont, T. O., Pratama, A. A., Gazitúa, M. C., Vik, D., Sullivan, M. B., & Roux, S. (2021). VirSorter2: A multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome*, 9(1). <https://doi.org/10.1186/s40168-020-00990-y>
- Markusson, Sigurdur, Stefansson, Andri. (2011). Geothermal surface alteration of basalts, Krýsuvík Iceland—Alteration mineralogy, water chemistry and the effects of acid supply on the alteration process. *ScienceDirect*,

- McKee, L. S., Martínez-Abad, A., Ruthes, A. C., Vilaplana, F., & Brumer, H. (2019). Focused Metabolism of β -Glucans by the Soil *Bacteroidetes* Species *Chitinophaga pinensis*. *Applied and environmental microbiology*, 85(2), e02231-18. <https://doi.org/10.1128/AEM.02231-18>
- Mehdizadeh Allaf, M., & Peerhossaini, H. (2022). Cyanobacteria: Model Microorganisms and Beyond. *Microorganisms*, 10(4), 696. <https://doi.org/10.3390/microorganisms10040696>
- Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications*, 7(1). <https://doi.org/10.1038/ncomms11257>
- Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017). metaSPAdes: A new versatile metagenomic assembler. *Genome Research*, 27(5), 824–834. <https://doi.org/10.1101/gr.213959.116>
- Poyntner, C., Täumer, J., Dahl, M. B., Jensen, C., Kuss, A. W., Margesin, R., & Urich, T. (2026). Changes in the (micro-)biome structure and function of two forest sites in the Alps at different elevations and seasons using metatranscriptomics. *Applied Soil Ecology*, 223, 107050. <https://doi.org/10.1016/j.apsoil.2026.107050>
- Roux, S., Camargo, A. P., Coutinho, F. H., Dabdoub, S. M., Dutilh, B. E., Nayfach, S., & Tritt, A. (2023). iPHoP: An integrated machine learning framework to maximize host prediction for metagenome-derived viruses of archaea and bacteria. *PLOS Biology*, 21(4), e3002083. <https://doi.org/10.1371/journal.pbio.3002083>
- Thakur, N., Singh, S. P., & Zhang, C. (2022). Microorganisms under extreme environments and their applications. *Current Research in Microbial Sciences*, 3, 100141. <https://doi.org/10.1016/j.crmicr.2022.100141>
- Tikhonova, E. N., Suleimanov, R. Z., Miroshnikov, K. K., Oshkin, I. Y., Belova, S. E., Danilova, O. V., Ashikhmin, A. A., Konopkin, A. A., But, S. Y., Khmelenina, V. N., Pimenov, N. V., & Dedysh, S. N. (2023). *Methylobacterium rapida* sp. nov., a novel species of fast-growing, carotenoid-producing obligate methanotrophs with high biotechnological potential. *Systematic and Applied Microbiology*, 46(2), 126398. <https://doi.org/10.1016/j.syapm.2023.126398>
- White, D. C., Sutton, S. D., & Ringelberg, D. B. (1996). The genus *sphingomonas*: Physiology and ecology. *Current Opinion in Biotechnology*, 7(3), 301–306. [https://doi.org/10.1016/s0958-1669\(96\)80034-6](https://doi.org/10.1016/s0958-1669(96)80034-6)