Characterizing and Evaluating the Diverse Microbial Communities and Their Mercury Resistance Potential from Hot Spring Sites Representing Gradients in Temperature and pH in Yellowstone National Park

Yelizaveta Rassadkina, Spencer Roth, and Tamar Barkay
Department of Biochemistry and Microbiology, Rutgers University

Abstract

Yellowstone National Park is home to many different hot springs, lakes, geysers, pools, and basins that range in pH, chemical composition, and temperature. These different environmental variations provide a broad range of conditions that select and grow diverse communities of microorganisms. In this study, we collected samples from geochemically diverse lakes and springs to characterize the microbial communities present through 16S rRNA metagenomic analysis. This information was then used to observe how various microorganisms survive in high mercury environments. The results show the presence of microorganisms that have been studied in previous literature. The results also depict gradients of microorganisms including thermophilic bacteria and archaea that exist in these extreme environments. In addition, beta diversity analyses of the sequence data showed site clustering based primarily on temperature instead of pH or sample site, suggesting that while pH, temperature, and sample site were all shown to be significant, temperature is the strongest factor driving microorganism community development. While it is important to characterize the microorganism community present, it is also important to understand how this community functions as a result of its selection. Along with looking at community composition, genomic material was tested to see if it contained mercury methylating (*hgcA*) or mercury reducing (*merA*) genes. Out of 22 samples, three of them were observed to have *merA* genes, while no samples had *hgcA* genes. These results indicate that microorganisms in Mustard and Nymph Springs may use mercury reduction. Understanding how microorganisms survive in environments with high concentrations of toxic pollutants is crucial because it can be used as a model to better understand mechanisms of resistance and the biogeochemical cycle, as well as for bioremediation and other solutions to anthropogenic problems.

Key terms
Chemolithotrophs, mercury methylation, mercury reduction, thermophiles.

Introduction

Yellowstone National Park is well known for its unique aquatic geothermal habitats, which exhibit a wide array of extreme temperatures and pH levels. Extreme conditions select for lithotrophic extremophiles, organisms that obtain reducing equivalents from inorganic substrates (Cuhel et al., 2002). Chemolithotrophs are a specific group of lithotrophic organisms that use inorganic substrates for ATP production, as well as inorganic material for carbon compound biosynthesis such as CO_2 fixation (Cuhel et al., 2002). These groups of organisms belong to the archaeal and bacterial domains and are dominant in the extreme environments common in Yellowstone National Park (Cuhel et al., 2002).

Yellowstone's geothermal activity leads to the presence of hydrothermal springs, which emit compounds such as sulfide, elemental sulfur, arsenic, and iron (Shock et al., 2010). These compounds can then be used as a source of energy for local microorganisms. Since most of the microorganisms in hot springs are chemolithotrophic thermophiles, they require specific nutrients and extreme conditions that are not easily replicated in a laboratory setting. Therefore, the ecology of microbial communities within Yellowstone is not as well-characterized as communities from many mesophilic environments (Reysenbach et al., 2000). This is why we used a metagenomic sequencing approach—rather than cultures—to better understand the target microorganism community. It is important to understand which microbial guilds dominate these environments, how they live and interact with each other, and how selective environmental pressures influence these interactions.

The environmental pressures of hot springs include a range of moderate to high temperatures, extreme pH ranges, and several toxic compounds. Mercury (Hg) is one such compound that has been observed at measurable concentrations in ~20% of geothermal sites in western North America (Geesey et al., 2016). Hg is a pollutant and a toxic metal present throughout many natural environments, such as hot springs and sediments. Due to its toxicity, Hg applies selective pressure on local organisms, especially in high concentrations. Mercury-resistant organisms can be found in some geothermal sites in North America, such as Yellowstone National Park (Geesey et al., 2016). Microbes that can tolerate high concentrations of Hg often have genes that allow them to reduce Hg(II) to its volatile form, Hg(0). Some prokaryotes have the *mer* operon, associated with mercury reduction, which encodes for mercury transport and transformations (Barkay and Wagner-Dobler, 2005). There are many different *mer* genes that participate in the uptake of Hg(II) into the cell. After being reduced, Hg(0) is released into the atmosphere in its gaseous state (Barkay and Wagner-Dobler, 2005). It is crucial to understand how organisms withstand toxic pollutants in extreme conditions within a properly functioning ecosystem.

In this study, we aim to characterize the microbial community in Yellowstone National Park and identify the mechanisms of mercury resistance that allow them to survive in these geothermally

diverse hot springs. For this purpose, DNA sequencing was used to identify the mercury reducing organisms or mercury methylating organisms present in the community. In addition, diversity metrics were used to determine if the physical and chemical properties of the spring environment correlate with the structure and function of their microbial communities. Our study differs from previous literature in that we have a pH and temperature range as well as a matrix range. This allows us to study the differences between different types of samples taken from the same site, i.e. scum vs. sediment from the sample site. We used DNA sequencing and diversity metrics to broaden our understanding of the structure and function of the microbial community in the hot springs of Yellowstone National Park. This information can help us better understand the impact of microorganism communities and their diversity in nature.

Methods

Geochemistry

The conditions of the hot springs (i.e. temperature and pH) from which the samples were extracted were recorded by Dr. JoAnn Holloway and team.

Sample Description

The samples were collected in Yellowstone National Park in September 2017 by Dr. JoAnn Holloway (USGS). Samples were taken from Upper Geyser Basin, Nymph basin, Norris Geyser Basin, Mud Volcano, Artist Paint Pots Basin, Gibbon Canyon, and Midway Geyser Basin. Samples were then stored in Lifeguard solution (Qiagen). For each sample, approximately one gram was collected and placed in a tube with Lifeguard solution. Scum and water samples were each collected with a sterile syringe and then filtered. Filters were placed into a Lifeguard solution and then stored on ice. Samples were shipped to Rutgers University and stored at −80°C until further analysis.

Extraction

DNA was extracted using the MoBio Powerlyzer PowerSoil DNA Isolation kit (Qiagen) following the manufacturer's protocol. This kit is intended to be used for the isolation of microbial nucleic acids from soil, environmental samples, and tough microbes. It uses bead-based homogenizers to eliminate humic substances and other inhibitors, which allows for downstream application. DNA extracts were stored at −20°C following extraction.

PCR

PCR (Polymerase Chain Reaction) was performed after every extraction to amplify the 16S rRNA gene from the extracts to determine the presence of bacterial and/or archaeal DNA. For PCR products obtained with GoTaq Green Master Mix (Promega), a 515 forward primer and 806

reverse primer 16S V4 primer set was used (Parada et al., 2016; Apprill et al., 2015). Products of PCR were separated by 1% agarose gel electrophoresis.

DNA Sequencing

The V4 region of the 16S rRNA gene was sequenced using the Illumina MiSeq platform following the Earth Microbiome Project at the Rutgers University Genome Cooperative (UL-Hasan et al., 2019).

hgcA Detection

PCR was performed on every sample to see if *hgcA* genes were present within the extracted DNA. *hgcA* was targeted using a 261 forward primer and a 912 reverse primer set as previously published (Schaefer et al, 2014). Products of PCR were separated by 2% agarose gel electrophoresis.

merA Detection

PCR was performed on every sample to see if *merA* genes were present within the extracted DNA. The PCR products were obtained with GoTaq Green Master Mix, a merA2 forward primer and a merA2 reverse primer set was used (Poulain et al., 2015). Products of PCR were separated by 2% agarose gel electrophoresis.

Microbial Ecology Analysis (Figure 1)

The bioinformatics platform for Quantitative Insights Into Microbial Ecology 2 (QIIME2) was used because it includes sequence alignments, quality filtering, phylogeny building, and taxonomic classification (Hall et al., 2018). QIIME2, as opposed to QIIME, allows for microbial marker gene analysis to better quantify and compare the extracted genomic material (Hall et al., 2018). For data analysis, the Divisive Amplicon Denoising Algorithm 2 (DADA2) pipeline was used to detect and correct the processed sequences (Callahan et al., 2016). Alpha rarefaction and beta diversity metrics were all performed and analyzed (Anderson, 2001). Taxonomy was assigned using a naïve-Bayesian classifier trained on the V4 region of the Silva-132 99% 16S database (Pedregosa et al., 2011) Then, a feature table and a feature data summary were made after sequence filtering. Lastly, a taxonomic bar plot was created along with beta diversity metrics such as Weighted UniFrac, which qualitatively measures community dissimilarity (Bolyen et al., 2019; Anderson, 2001).
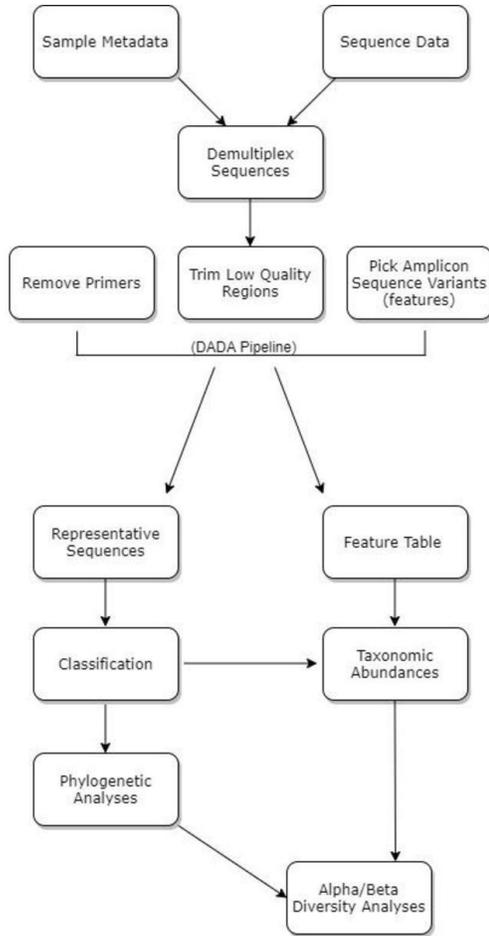
Figure 1. Pipeline depicting the 16S rRNA gene bioinformatics analysis. This pipeline was performed on QIIME2 using the extracted genomic material that was sent for sequencing of the V4 region of the 16S rRNA gene. Twenty-two sample sequences were used to start and then after the DADA function was applied, only 11 samples were kept for further analyses because they rarified at 500 sequences. The remaining sequences were then taxonomically classified and used to perform beta diversity analysis (weighted UniFrac).

Results

Samples were taken from various geothermally diverse sediments, waters, mats, and scum sites. A total of twenty-two samples were taken from distinct environments that ranged in pH from 1.71 – 6.65 and in temperature from 54.2°C – 87.8°C. These twenty-two samples came from various springs, geysers, pools, and basins located in Yellowstone National Park (Figure 2). Several sample matrices were taken in duplicate.
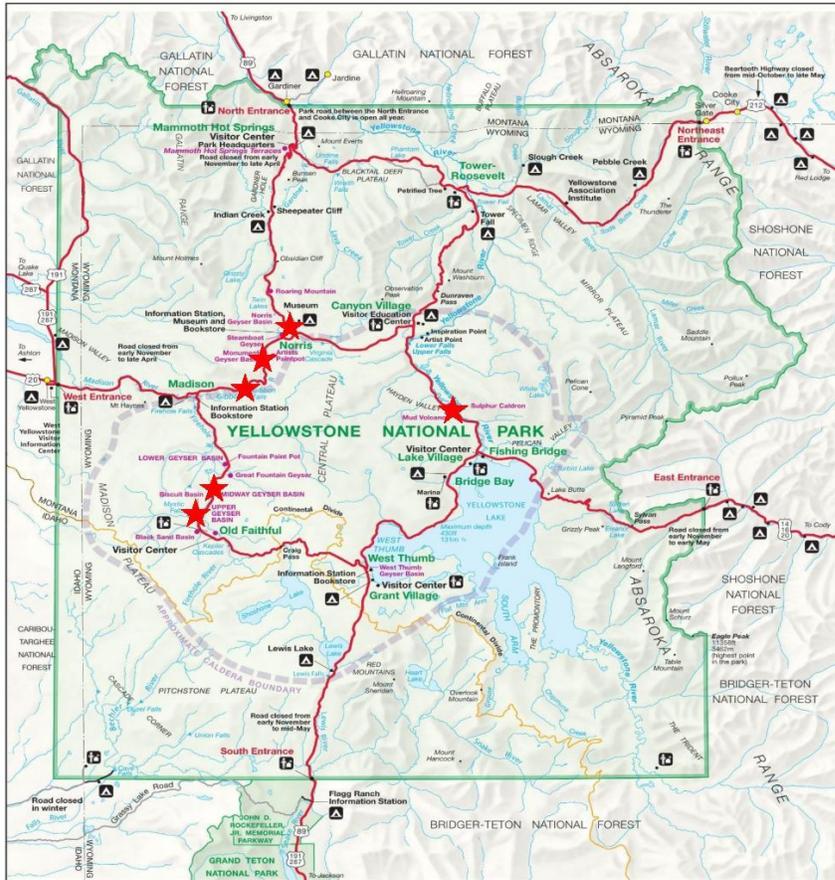
Figure 2. Map of Yellowstone National Park depicting the main hot springs and the landmarks near them. The red stars indicate the location of sites from which the samples were taken. Twenty-two different samples were taken from six different locations. The image of the map was taken from Yellowstone Maps website https://yellowstone.net/maps/yellowstone-park-map/.

After collection, the twenty-two samples were sent for sequencing analysis after completion of the genomic extractions. Eleven out of twenty-two samples were retained after rarefication at 500 or more sequence reads (Table 1). The other eleven samples had very small amounts of biomass and thus resulted in a low number of reads (<500), so these samples were excluded from further analyses. The samples that had a low number of reads consisted mainly of water and scum samples from sample matrices.

Table 1. Samples from which DNA extracts were obtained, their source spring, water pH, temperature, and the number of reads.

Sampled Spring	Sample Matrix	Sample ID	pH	Temperature (°C)	Number of Sequence Reads
Mud Volcano	Sediment	MV-S	1.71	71.5	2737
Mud Volcano	Mat	MV-M	2.07	54.2	1185
Mud Volcano	Scum	MV-SC	2.07	54.2	534
Mud Volcano	Pool Edge Sediment	MV-PS	2.07	54.2	2507
Mud Volcano	Pool Edge Sediment I	MV-PS1	2.07	54.2	2610
Mud Volcano	Water	MV-W	2.07	54.2	1451
Crystal Sister W	Sediment	CSW-S	2.19	81.3	18044
Crystal Sister W	Sediment I	CSW-S1	2.19	81.3	635
Nymph	Sediment	N-S	2.32	59.5	503
Crystal Sister E	Sediment	CSE-S	4.22	81.2	5927
Gibbon Canyon	Outlet Sediment	GC-O	6.65	87.8	777

Weighted UniFrac analysis was performed to observe how the various sample matrices cluster together based on several environmental factors that drive community selection. The three main selection factors that were investigated were temperature, pH, and sample matrix. These environmental conditions were tested to isolate the primary factor that influenced sample clustering. The Principle Coordinate Analysis (PCA) plot showed that the samples exhibited clustering based on sample matrix (Figure 3a). One cluster was composed of sediment and outlet sediment samples, while the other was composed of sediment and water samples. The mat and scum samples, both from Mud Volcano, did not cluster together with the other samples. In addition, the pool edge sediment from the Crystal Sister West sample was shown to be distinct from the other sample matrices.

The Weighted UniFrac beta diversity analysis showed that the microbial communities in Yellowstone National Park clustered primarily based on temperature, rather than pH, which had no clustering pattern (Figures 3b and 3c). Microbial communities from samples that had a temperature of 50–60℃ tended to cluster together, while communities that were taken from sampling sites with a temperature of 70–90℃ tended to cluster together, with the exception of Crystal Sister West sediment. Beta diversity dissimilarity results showed that pH accounts for 13.9% (p = 0.006) of variance, while temperature accounts for 22.7% (p < 0.001) and sample matrix accounts for 38.1% (p = 0.009).
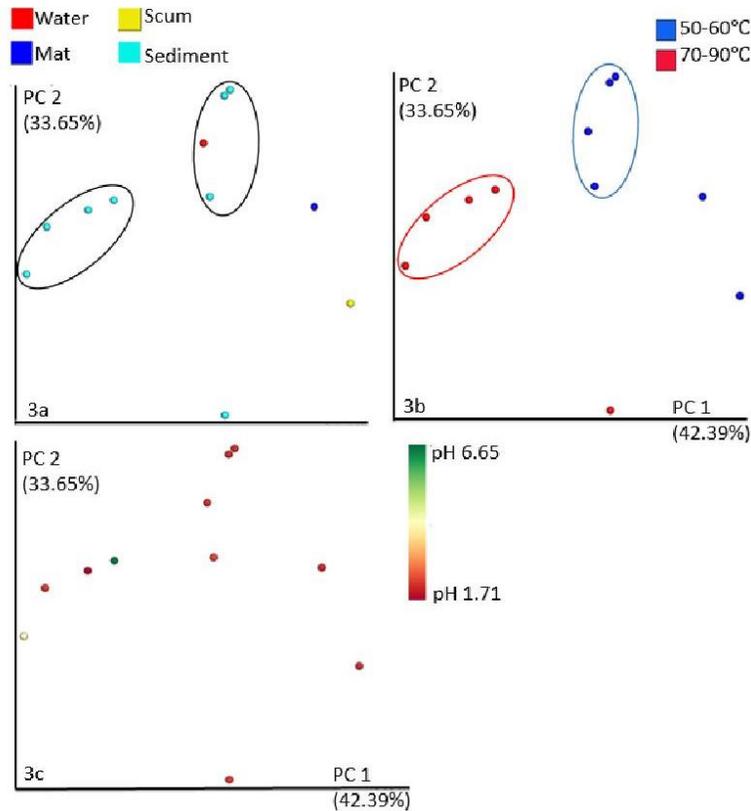
Figure 3a-3c. PCA plots of Weighted UniFrac distances showing the beta dissimilarity of the samples based on the sample type (left, 3a), spring water temperature (right, 3b), and pH (bottom, 3c) of the different samples. Figure 3a shows that sediment samples cluster together, while figure 3b demonstrates that there is a clustering of samples based on the temperature of the spring. Figure 3c shows that there is no clustering of the samples based on the pH of the hot spring. Principal Component 1 (PC 1) represents the most variation within the samples, 42.39%, and separates the groups based on temperatures except for three sample outliers: Crystal Sister West sediment, Mud Volcano scum, and Mud Volcano mat. Principal Component 2 (PC 2) accounts for the second most variation present within the samples which is 33%. The beta-diversity dissimilarity was significantly associated with pH, temperature, and sample type ($p < 0.05$).

Sequence taxonomy was assigned to determine the major taxa present in the samples. The archaeal order Thermoproteales was the most abundant throughout the samples, except for at the Mud Volcano mat and Nymph water sites, where it was not detected (Figure 4). The second most dominant order was Candidatus Micrarchaeum, which was present in seven of the samples. In addition, the most dominant bacterial taxon was Acetobacterales, while the most dominant archaeal taxon was Desulfurococcales. Overall, there were more archaea than bacteria present in physiologically extreme conditions.
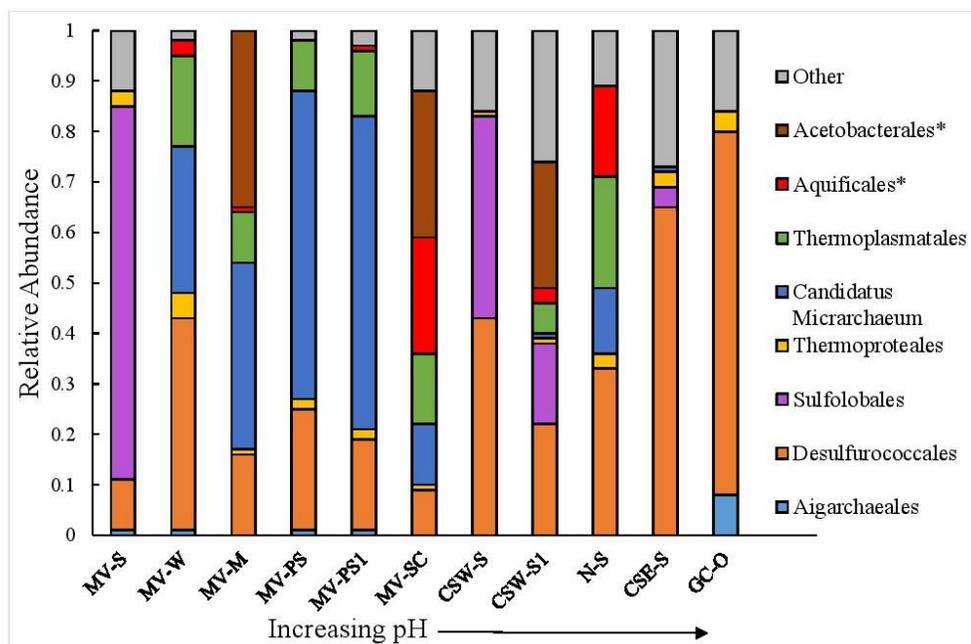
Figure 4. Most common archaeal and bacterial orders present in the samples. All samples returned >500 sequence reads. Sites are organized along the x-axis from the lowest pH to the highest. The pH ranges from 1.71 in Mud Volcano sediment to a pH of 6.65 in sample Gibbon Canyon sediment. The stars indicate orders that belong to the bacterial domain while all other orders belong to the archaeal domain. The "Other" category represents sequenced reads that could not be taxonomically classified. Samples for which <500 reads were obtained are not included in the presentation. (MV-Mud Volcano, CSW- Crystal Sister W, N- Nymph, CSE- Crystal Sister E, GC- Gibbon Canyon, S- Sediment, W- Water, M- Mat, PS- Pool Edge Sediment, SC-Scum, O- Outlet Sediment)

Within the eleven samples, those that had more than a 50% relative frequency of archaea included Mud Volcano mat, scum, pool edge sediment, and water; sediments from site samples Crystal Sister E, Nymph, Crystal Sister W, Mud Volcano; and outlet sediments from Gibbon Canyon. The samples dominated by bacteria (more than 50% relative frequency) were Mud Volcano scum and mat. Within Mud Volcano, samples collected from the sediment were dominated by archaea, while the mat and scum samples were dominated by bacteria. Candidatus Micrarchaeum was abundant in samples with a pH range of 1.0–2.32. Desulfurococcales and Thermoproteales were present in the pH range of 1.71–6.65. Archaeal 16S rRNA reads were detected in all samples with pH at or below 6.65.

Lastly, all of the samples were also tested for mercury reduction or mercury methylation genes. The gene encoding a mercuric reductase, *merA*, was detected in two samples from Mustard Spring and one sample from Nymph. Mustard spring has a pH of 8.19 with a temperature of 86.2 °C, while Nymph Spring has a pH of 2.32 and a temperature of 59.5 °C. No copies of *hgcA*, a gene responsible for Hg methylation, were detected in the samples.

Discussion/Conclusion

Sequences of 16 rRNA gene PCR products were obtained from most of the samples that were extracted. Samples that had over one thousand sequence reads were mostly sediment samples. Fewer reads, above one hundred but below one thousand, were obtained from the mats, water, and scum samples. In addition, the samples excluded from the PCA plot had very low number of reads (<500 reads) such as mat and water samples. This is not that surprising because (i) samples were very small—yielding little biomass in all but the sediments and extreme biomass scarcity in water samples—and (ii) the DNA extraction kit was most suited for soil and sediment samples.

The beta diversity analysis showed that the microbial communities in Yellowstone National Park cluster primarily by temperature rather than pH. This matches previous literature that states that the main selection factor in hot springs tends to be temperature (Uribe-Lorio et al., 2019). Beta diversity dissimilarity results showed statistical significance of sample separation based on pH, temperature, and sample type. The Weighted UniFrac PCA plot, which measures community dissimilarity while incorporating the relative abundance of the organisms, showed two distinct groups clustered based on the temperatures of the sites (Lozupone et al., 2011). In addition to these two grouping clusters, the results also showed a dominance of archaeal species. This is likely due to the extreme thermophilic and pH conditions which allow archaea to outcompete bacteria in these niches (Reed et al., 2013). Archaea are known extremophiles that have evolved specific proteins that allow them to proliferate in extreme environments and tolerate normally toxic compounds, such as arsenic and mercury (Reed et al., 2013). It is not surprising that they dominate these communities in the extreme conditions of some Yellowstone springs.

In addition to temperature, the microbial taxa also clustered based on sampling matrix. Clusters for sediments and outlet sediments, pool edge sediments, water, and scum samples were all observed. The Weighted UniFrac plot showed two distinct clustering patterns between the different sample matrices. Environmental conditions in the different matrices are expected to vary. Sediments are expected to contain less dissolved oxygen, less UV radiation, more sulfide, more arsenic, and higher mercury concentrations compared to water, mat, and scum samples, which were taken from the top of the hot spring (Jiang et al., 2016).

Sequence reads were taxonomically classified at the order level for all sites where more than 500 sequence reads were available. Around 10% of the obtained reads were unidentified, possibly due to many microorganisms being unculturable and thus not being included in databases. Results showed that the archaeal order Thermoproteales was present in all but one, Crystal Sister West, of the sampling sites, indicating that this order thrives in these extreme thermophilic conditions. Since Thermoproteales was detected in most of the sequenced libraries, it can be inferred that this order is selected for in these niches, and that its growth is not limited to a singular matrix in the hot spring. This aligns with previous literature that

depicts diverse metabolic capabilities of Thermoproteales, which includes sulfur oxidation/reduction, oxygen respiration, H_2 respiration, and CO_2 fixation (Jay et al., 2016).

Several samples from the Mud Volcano site were collected, extracted, and sequenced. These included three mat samples, one water sample, two pool edge sediment samples, and two scum samples. This allowed us to compare communities from different matrices and locations within one site. The Mud Volcano sediments (central and pool edge sediment) were dominated by archaea while the mats and scum showed the presence of both bacteria and archaea. Mud Volcano has moderately hot and acidic water (Table 1), which has magnesium, calcium, and sodium along with high concentrations of both organic carbon and sulfur (Sherman et al., 2009). These are all necessary trace minerals that can be used for energy and growth, so it is not surprising that we found an abundance of both acidophilic and thermophilic bacteria and archaea.

Observing the abundance of taxa in different sample matrices and springs suggests a potential gradient. Most of the samples dominated by archaea are the sediment samples, while the mats, water, and scum samples are dominated by bacteria (excluding Mud Volcano). The anoxic sediments of the hot springs, which have a higher concentration of Hg and sulfide (Krabbenhoft et al., 1999), show an abundance of archaeal species. On the other hand, the mat, water, and scum samples were taken from the water above the sediment and thus exposed to oxygen, sunlight, less mercury, and less sulfide has a higher abundance of bacterial taxa. This is likely due to the innate characteristics of archaea, such as a lipid monolayer and having specific proteins that have evolved to remain stable and active in various harsh conditions that allow them to thrive and proliferate in extreme environments (Reed et al., 2013). In addition, such gradients could be due to the competition between the archaea and bacteria. The archaea outnumber the bacteria at the bottom of the springs, thus out-competing them for resources and nutrients. Archaea are better adapted to survive in oligotrophic, low nutrient environments. They are able to live in sediments that are energy stressed due to having lower cell permeability than bacteria, thus reducing energy loss (Vuillemin et al., 2019). In addition, archaea are also able to outcompete bacteria for urea, which is used as a source of carbon for the cell (Seyler et al., 2019). This shows that archaea have developed several adaptations that allow them to utilize nutrients that often cannot be used by bacteria, thus allowing them to grow and survive in more environments and outcompete bacteria for the limited nutrients.

At low temperatures, bacterial and archaeal species were found in similar abundances, while archaea were much more dominant at high temperatures. This indicates that higher temperatures allow archaeal dominance within acidic to neutral pH ranges (1.71–6.65). This might be because most alkaliphilic archaea are also considered halophilic due to sharing similar genomic characteristics (Reed et al., 2013). The sampled alkaline springs, like Octopus Spring, did not have a high salt concentration, thus possibly selecting hyperthermophilic and alkaliphilic bacteria instead of archaea.

In Yellowstone National Park, as well as other geothermal springs, there are often moderate to high levels of toxic compounds, such as Hg, or those that may serve as growth substrates, such as sulfur. Consequently, organisms often carry genes for resistance to these compounds or may utilize these compounds for better survival and growth. In our study, genes necessary for Hg methylation, *hgcA*, were not detected in any of the hot spring samples, and no known Hg methylating genera were detected by 16S sequencing. This, however, does not indicate that there is no presence of Hg methylation, it just means that the samples that were collected did not contain microorganisms with this gene or that organisms with this gene were detected but not characterized due to the limited amount of information that is known about uncultured extremophiles.

When looking at Hg resistance, the taxonomy derived from the 16S sequences showed the presence of the archaeon *Sulfolobus* spp., which has a *mer* operon (Barkay and Wagner-Dobler, 2005), in the Infant Geyser outlet sediment sample. This suggests that it is possible that Infant Geyser, which has a pH of 3.05, ~5.5 x 10^4 ng/L of total mercury, and a temperature of 82.7℃, has thermophilic mercury resistant species (Barkay, Personal Communication). Sequence reads in the order Aquificales, which also contains the *mer* operon, were detected in several sampling sites including Nymph (~1,000 ng/L total mercury), Crystal Sister West (>10,000 ng/L total mercury), Mud Volcano (1 x 10^4 – 1 x 10^5 ng/L total mercury), and Gibbon Canyon (>1,000 ng/L total mercury) (Geesey et al., 2016; Barkay, Personal Communication). Aquificales bacteria were most abundant in Nymph and Mud Volcano. This suggests a potential of mercury reduction occurring within those environments, and this makes sense because these four sites all have a pH of around two. We detected the *merA* gene in DNA from three samples from Mustard Spring and Nymph. Mustard spring has a pH of 8.19 with a temperature of 86.2 °C while Nymph Spring has a pH of 2.32 and a temperature of 59.5 °C. This data demonstrates that in the springs where the total mercury concentration in unfiltered water is around 1,000 ng/L and above, there are taxa that have the potential for mercury resistance (Barkay, Personal Communication). In hot springs that have a low pH, there is often a higher concentration of total mercury, thus selecting for mercury resistant organisms. While the genes needed for mercuric reduction were detected within the samples, detection of the gene does not indicate activity. Further testing is needed to determine if mercury reduction genes are active within these species.

It is important to understand Hg reduction because it plays a critical role in the global cycling of Hg. The reduction of mercury is the only transformation of a toxic metal that can be done on a large scale (Barkay et al., 2003). Microorganisms that are capable of mercuric reduction are able to transform both organic and inorganic Hg during biotic and abiotic processes (Barkay et al., 2003). This transformation helps to prevent Hg bioaccumulation, which is a serious problem in some parts of the world. For many decades now, Hg reducing bacteria have been utilized by humans to remove toxins from human products such as wastewater (Barkay et al., 2003). This means that Hg reduction is not only necessary for ecosystem homeostasis, but is also fundamental in the reduction of Hg through bioremediation, which is the degradation of

environmental contaminants through the use of organisms, mainly microorganisms (Vidali, 2001). The different taxa shown here to live in hot springs with high mercury concentration, such as *Sulfolobus* spp. and the Aquificales bacterial order, can be used for further research to support the bioremediation process.

Our results correspond with previous literature about community structure within hot springs in Yellowstone National Park. Extracting metagenomic information from several springs demonstrated that the microbial communities cluster together based on the temperature of the spring and sample matrix. As one selection factor often dominates in extreme environments (Uribe-Lorio et al., 2019), it is important to know that the main selection factor for these communities in Yellowstone is temperature. A potential gradient based on the abundance of archaeal and bacterial species was observed in samples from different compartments within the same spring. Archaeal species were more abundant in the sediment samples and their presence decreased at the surface of the water, while bacterial species exhibited a reverse gradient where they were most abundant in the water and mats at the top of the spring and less abundant down towards the sediment. This shows that there is a negative correlation between the presence of bacterial and archaeal growth within the hot springs. This is likely due to the different environmental conditions and compounds that are present at the top and bottom of the hot spring. This information gives us a better understanding of how microbial communities form, function, and interact with the surrounding environment.

Even though the data obtained in this study matches the previous literature, there are still several limitations that need to be addressed. First, the sample size of this study is small, so the data presented in this paper cannot be widely generalized. There were originally twenty-two samples, and only eleven of them had over 500 sequences and were kept for further analyses. Five different basins were sampled. All sites had at least one water sample, however, only six sites had mat samples and only five sites had sediment samples. It was hard to extract genetic material from the water samples, as they had very low biomass. Furthermore, in those water samples, even if material was extracted, the sequence number was low. Therefore, the samples could not be considered. This issue caused the number of usable samples for metagenomic analysis to decrease by half. Out of the five basins originally sampled, only four (Mud Volcano, Nymph, Gibbon Canyon, and Norris Geyser Basin) had samples that were used for analysis and data extrapolation. Another limitation is that some of the basins used for data analysis are similar in their pH and temperature range. This means that this data cannot be applied broadly; more samples from other basins with varying pH levels must be analyzed before results can be generalized. Additionally, this paper used a metagenomic approach to show that mercury reduction genes are present, while mercury methylation genes were not. This, however, does not indicate if the *merA* genes are active within the community. Rather, it only shows that these genes are present. Regarding the absence of mercury methylation genes, this might not be representative of all of the sites due to the small sample size that was analyzed. In addition, the Hg concentration was not measured at the time of sampling but rather at another time in the same year. Overall, one of the more considerable limitations to this study is that many of the

microorganisms tested were unculturable. This means that there is often little information on these organisms, so it would be impossible to classify genes even if they were detected. This happened during the taxonomic analysis of the eleven samples, which were often categorized with only domain, family classification, or just "uncultured." These samples therefore did not have enough information for us to determine a bigger picture of the community structure.

Due to the various limitations that are present in this study, there is more that can be done for future research and analysis. First, more samples need to be taken, specifically more sediment samples, and they should be taken from a wider range of basins located throughout Yellowstone National Park. This will not only help to increase the sample size but also to search for other mercury reducers and mercury methylators within the different environments. Second, other *mer* operon genes can be tested instead of just mercuric reductase (*merA*). This can improve understanding of the different capabilities of the microorganisms inhabiting hot springs. Lastly, proteomics analyses can be done to see if the *merA* genes are active within the sampled microorganisms. Other research that can further this study is to look at other ways these microorganisms resist toxic mercury concentrations. While some microbes do not have the *mer* operon, they could have other ways to resist Hg. These methods could include influx, accumulation, efflux, and metallothionein, a protein that protects the cell against the toxicity of Hg (Irawati et al., 2012). It is important to know the many diverse ways that microorganisms resist toxic metals. This new information can give us a more holistic view of the unique environment of the microorganisms in the hot springs of Yellowstone National Park.

Acknowledgments

References

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, *26*(1), 32-46. doi: 10.1111/j.1442-9993.2001.01070.pp.x

Apprill, A., Mcnally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, *75*(2), 129-137. doi: 10.3354/ame01753

Barkay, T., Miller, S. M., & Summers, A. O. (2003). Bacterial mercury resistance from atoms to ecosystems. *FEMS Microbiology Reviews*, *27*(2-3), 355-384. doi: 10.1016/s0168-6445(03)00046-9

Barkay, T., & Wagner-Döbler, I. (2005). Microbial Transformations of Mercury: Potentials, Challenges, and Achievements in Controlling Mercury Toxicity in the Environment. *Advances in Applied Microbiology Advances in Applied Microbiology Volume 57*, 1-52. doi: 10.1016/s0065-2164(05)57001-1

Bolyen, E., Rideout, J. R., Dillon, M. R. *et al.* (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol 37,* 852-857 doi:10.1038/s41587-019-0209-9

Callahan, B. J., Mcmurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., & Holmes, S. P. (2016). DADA2: High resolution sample inference from amplicon data. *Nature Methods*, *13*(7), 581 doi: 10.1101/024034

Cuhel R. L., Aguilar C., Anderson P. D., Maki J. S., Paddock R. W., Remsen C. C., et al. (2002). *Underwater domains in Yellowstone Lake hydrothermal vent geochemistry and bacterial chemosynthesis* In: Anderson, R.J., and D. Harmon, eds. 2002. *Yellowstone Lake: Hotbed of Chaos or Reservoir of Resilience? Proceedings of the 6th Biennial Scientific Conference on the Greater Yellowstone Ecosystem.* Mammoth Hot Springs Hotel, Yellowstone National Park. Yellowstone National Park, Wyo., and Hancock, Mich. Yellowstone Center for Resources and The George Wright Society.

Geesey, G. G., Barkay, T., & King, S. (2016). Microbes in mercury-enriched geothermal springs in western North America. *Science of The Total Environment*, *569-570*, 321-331. doi: 10.1016/j.scitotenv.2016.06.080

Hall, M., & Beiko, R. G. (2018). 16S rRNA Gene Analysis with QIIME2. *Methods in Molecular Biology Microbiome Analysis*, 113-129. doi: 10.1007/978-1-4939-8728-3_8

Irawati, W., P., Soraya, Y., & Baskoro, A. H. (2012). A Study on Mercury-Resistant Bacteria Isolated from a Gold Mine in Pongkor Village, Bogor, Indonesia. *HAYATI Journal of Biosciences, 19*(4), 197-200. doi:10.4308/hjb.19.4.197

Jay, Z. J., Beam, J. P., Kozubal, M. A., Jennings, R. D., Rusch, D. B., & Inskeep, W. P. (2016). The distribution, diversity and function of predominant Thermoproteales in high-temperature environments of Yellowstone National Park. *Environmental Microbiology*, *18*(12), 4755-4769. doi: 10.1111/1462-2920.13366

Jiang, Z., Li, P., Nostrand, J. D. V., Zhang, P., Zhou, J., Wang, Y., … Wang, Y. (2016). Microbial communities and arsenic biogeochemistry at the outflow of an alkaline sulfide-rich hot spring. Scientific Reports, 6(1). doi: 10.1038/srep25262

Krabbenhoft, D.P., Wiener, J.G., Brumbaugh, W.G., Olson, M.L., Dewild, J.F., & Sabin, T.J. (1999). A National Pilot Study of Mercury Contamination of Aquatic Ecosystems Along Multiple Gradients: Bioaccumulation in Fish. *Environmental Science.*

Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., & Knight, R. (2011). UniFrac: An effective distance metric for microbial community comparison. *The ISME Journal, 5*(2), 169-172. doi:10.1038/ismej.2010.133

Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, *18*(5), 1403-1414. doi: 10.1111/1462-2920.13023

Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., and Duchesnay E. (2011). Scikit-learn: machine learning in python. *Journal of machine learning research*, 12, 2825-2830.

Poulain, A. J., Aris-Brosou, S., Blais, J. M., Brazeau, M., Keller, W., & Paterson, A. M. (2015). Microbial DNA records historical delivery of anthropogenic mercury. *The ISME Journal*, *9*(12), 2541-2550. doi: 10.1038/ismej.2015.86

Reed, C. J., Lewis, H., Trejo, E., Winston, V., & Evilia, C. (2013). Protein Adaptations in Archaeal Extremophiles. *Archaea*, 1-14. doi: 10.1155/2013/373275

Reysenbach A-L., Ehringer M, Hershberger K. (2000). Microbial diversity at 83°C in Calcite Springs, Yellowstone National Park: another environment where the Aquificales and "Korarchaeota" coexist. *Extremophiles* 4:61-67.

Schaefer, J. K., Kronberg, R.-M., Morel, F. M. M., & Skyllberg, U. (2014). Detection of a key Hg methylation gene,hgcA, in wetland soils. *Environmental Microbiology Reports*, *6*(5), 441-447. doi: 10.1111/1758-2229.12136

Seyler, L. M., Tuorto, S., Mcguinness, L. R., Gong, D., & Kerkhof, L. J. (2019). Bacterial and Archaeal Specific-Predation in the North Atlantic Basin. Frontiers in Marine Science, 6. doi: 10.3389/fmars.2019.00555

Sherman, L., Blum, J., Nordstrom, D., Mccleskey, R., Barkay, T., & Vetriani, C. (2009). Mercury isotopic composition of hydrothermal systems in the Yellowstone Plateau volcanic field and Guaymas Basin sea-floor rift. *Earth and Planetary Science Letters*, *279*(1-2), 86-96. doi: 10.1016/j.epsl.2008.12.032

Shock, E. L., Holland, M., Meyer-Dombard, D. A., Amend, J. P., Osburn, G., & Fischer, T. P. (2010). Quantifying inorganic sources of geochemical energy in hydrothermal ecosystems, Yellowstone National Park, USA. *Geochimica Et Cosmochimica Acta*, *74*(14), 4005-4043. doi: 10.1016/j.gca.2009.08.036

Ul-Hasan S, Bowers RM, Figueroa-Montiel A, Licea-Navarro AF, Beman JM, Woyke T, Nobile CJ. (2019). Community ecology across bacteria, archaea and microbial eukaryotes in the sediment and seawater of coastal Puerto Nuevo, Baja California. *PLoS ONE* 14(2): e0212355. doi: 10.1371/journal.pone.0212355

Uribe-Lorío, L., Brenes-Guillén, L., Hernández-Ascencio, W., Mora-Amador, R., González, G., Ramírez-Umaña, C. J., … Pedrós-Alió, C. (2019). The influence of temperature and pH on bacterial community composition of microbial mats in hot springs from Costa Rica. *MicrobiologyOpen*, *8*(10). doi: 10.1002/mbo3.893

Vidali, M. P. (2001). Bioremediation. An overview. *Pure and Applied Chemistry, 73*(7), 1163-1172. doi:https://doi.org/10.1351/pac200173071163

Vuillemin, A., Wankel, S. D., Coskun, Ö K., Magritsch, T., Vargas, S., Estes, E. R., . . . Orsi, W. D. (2019). Archaea dominate oxic subseafloor communities over multimillion-year time scales. *Science Advances, 5*(6). doi:10.1126/sciadv.aaw4108

Yellowstone National Park – Official Park Map. (n.d.). In *Yellowstone Maps*. Retrieved from https://yellowstone.net/maps/yellowstone-park-map/