# EFFECTS OF PERMIAN BASIN OIL AND GAS WASTEWATER DUMPING ON SOIL MICROBIAL COMMUNITIES

by

Mitra Kashani

A Thesis
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Master of Science
Biology

Committee:

| | |
|---|---|
| _____ | Dr. Denise M. Akob, Thesis Director |
| _____ | Dr. Ancha Baranova, Thesis Chair |
| _____ | Dr. Jennifer Salerno, Committee Member |
| _____ | Dr. Patrick Gillevet, Committee Member |
| _____ | Dr. Iosif Vaisman, Director, School of Systems Biology |
| _____ | Dr. Donna M. Fox, Associate Dean, Office of Student Affairs & Special Programs, College of Science |
| _____ | Dr. Fernando Miralles-Wilhelm, Dean, College of Science |

Date: _____7/27/2020_____

Summer Semester 2020
George Mason University
Fairfax, VA

# Effects of Permian Basin Oil and Gas Wastewater Dumping on Soil Microbial Communities

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at George Mason University

by

Mitra Kashani
Bachelor of Science, Biology
George Mason University, 2018

Director: Denise M. Akob, Research Microbiologist
U.S. Geological Survey

Summer Semester 2021
George Mason University
Fairfax, VA

## DEDICATION

I would like to dedicate this thesis to my family – Mohsen, Faranak, and Ashkaan Kashani – who motivated me to pursue my Masters and supported me unconditionally throughout my work.

## ACKNOWLEDGEMENTS

I would like to thank my team at the Reston Microbiology Laboratory at the U.S. Geological Survey (Reston, Virginia) for all of their support in helping me craft this research and thesis. Thank you to my advisor Dr. Denise M. Akob for her dedication to mentoring me and her encouragement and support, and thank you to the rest of my team at the U.S. Geological Survey for all of their support and mentorship in helping me craft this research and thesis, namely Dr. Adam Mumford, Robert Andrews, Amy Springfield, Jeanne Jaeschke, Allison Leninger, Cassandra Harris, Marisa Mihori, Timothy Bushman, Bridgette Polite. Also, thank you to the collaborators on this project Dr. Mark A. Engle, Dr. Douglas B. Kent, Terry Gregston, Dr. Aaron M. Jubb, Dr. Isabelle Cozzarelli, Dr. Aïda Farag, and Dr. William B. Orem for their work and insights. Thank you to my committee members Dr. Jennifer Salerno and Dr. Patrick M. Gillevet for their guidance, and to my committee chair Dr. Ancha Baranova for her support and encouragement. Finally, thank you to the U.S. Geological Survey's Toxic Substances Hydrology Program for funding this research.

## TABLE OF CONTENTS

## LIST OF TABLES

**Table** **Page**

## LIST OF FIGURES

**Figure** **Page**

## LIST OF SUPPLEMENTAL MATERIALS

## LIST OF ABBREVIATIONS AND SYMBOLS

Advanced Research Computing……ARC
Aktiengesellschaft……AG
Bromide……..Br
Bureau of Land Management……BLM
California……..CA
Carbon……C
Centimeter……cm
Chloride…….Cl
Deoxyribonucleic acid…….……DNA
Inductively coupled plasma-mass spectrometry……...ICP-MS
Inductively coupled plasma-optical emission spectroscopy……...ICP-OES
Ion chromatography……...IC
Massachusetts……..…MA
Micrometer……..µm
Millisiemens per centimeter……...……mS/cm
New Jersey…….....NJ
Nitric acid……...……...$HNO_3$
Nitrogen……...……...…N
Non-metric multidimensional scaling……NMDS
Ohio……..OH
Oil and gas ……..……OG
Operational taxonomic unit……..…OTU
Percent……%
Permutational Multivariate Analysis of Variance ……...……PERMANOVA
Polymerase Chain Reaction……..……...…PCR
Radium……..……Ra
Real Time Analysis……RTA
Silver……...……...…Ag
Sodium …….Na
Strontium……Sr
Texas……TX
Tin……..……..……..…Sn
Total dissolved solids……...……TDS
U.S. Geological Survey……...USGS
Underground Injection Control……...UIC
United States of America…….…….……USA
United States ……..……U.S.

## ABSTRACT

EFFECTS OF PERMIAN BASIN OIL AND GAS WASTEWATER DUMPING ON SOIL MICROBIAL COMMUNITIES

Mitra Kashani, M.S.

George Mason University, 2021

Thesis Director: Dr. Denise M. Akob

The Permian Basin, straddling New Mexico and Texas, is the largest oil and gas (OG) producing region in the United States. OG production yields large volumes of wastewaters that have complex chemistries, which pose unknown environmental health risks, particularly in the case of accidental or intentional releases. Starting in November 2017, at least 39 illegal dumps of OG wastewater were identified in southeastern New Mexico, which resulted in ~4,000 barrels of fluids being released onto desert soils. To evaluate the impacts of these releases, we analyzed changes in soil geochemistry and microbial community structure at 5 sites by comparing soils from within OG wastewater dump-affected zones to unaffected (control) soils collected from the same site. Soil geochemistry across all dump sites differed from unaffected soils, reflecting the residual salts and hydrocarbons left behind by the OG wastewater release (e.g., enriched in sodium (Na), chloride (Cl), and bromide (Br)). Microbial community analysis, using Illumina 16S iTag sequencing, revealed that communities in dump zones had lower alpha diversity

compared to control sites. Subjecting a weighted UniFrac distance matrix to a permutational multivariate analysis of variance (PERMANOVA) revealed significant ($p<0.01$) variation between dump and control soil community structure. Shifts in the phylogenetic composition and potential function of microorganisms in dump-affected soils were also observed. Dump-affected soils showed an overwhelming increase in OTUs with halophilic capabilities, such as those affiliated with the Marinobacteraceae, Halomonadaceae, and Halobacteroidaceae, suggesting that the brine composition of the OG wastewater dumps were a strong selective pressure for microbial communities. Taxa with capabilities to metabolize hydrocarbons were also detected in the dump-affected soil communities. Overall, this study elucidates changes in arid soil geochemistry and microbial community dynamics due to OG wastewater exposure. Ultimately, we elucidate the potential role for microorganisms to serve as biological markers for, and their potential to naturally attenuate, OG wastewaters.

# INTRODUCTION

## Oil and Gas Development in the United States

In the United States, the demand for fossil fuels has increased and resulted in increased production from onshore resources. Both conventional and unconventional oil and gas (OG) reservoirs exist within the contiguous U.S. with the potential to be extracted economically (King, 2010). Conventional OG formations refer to resources that are generally more accessible for extraction due to the rock formation's high porosity and permeability. However, decades of depletion of conventional reservoirs has driven industry towards unconventional OG resources (King, 2010). Hydrocarbons in unconventional geologic formations are more difficult to extract, as resources are trapped in low permeability and porosity reservoirs. Specialized technologies such as horizontal drilling and hydraulic fracturing allow production of unconventional resources that were previously inaccessible or uneconomical for extraction (Arthur et al., 2008; Gallegos & Varela, 2015; King, 2012). The most productive geologic provinces of both conventional and unconventional OG reservoirs in the continental U.S. are the Appalachian, Permian, and Bakken regions (Figure 1) (U.S. Energy Information Administration, 2016).

Hydraulic fracturing is a technique used to increase OG production by injecting hydraulic fracturing fluids at high pressure into hydrocarbon reservoirs, fracturing the rock

formation to allow for greater hydrocarbon flow rates (Frithsen, 2015; Gallegos & Varela, 2015). Directional drilling, combined with hydraulic fracturing, allows producers to access a greater variety of OG reserves. While vertical drilling targets conventional OG formations, horizontal drilling, which can turn a vertically oriented wellbore as much as 90º, is used to access low-permeability, unconventional OG reservoirs, such as shale gas, coalbed methane, and tight sands (Arthur et al., 2008; King, 2012).

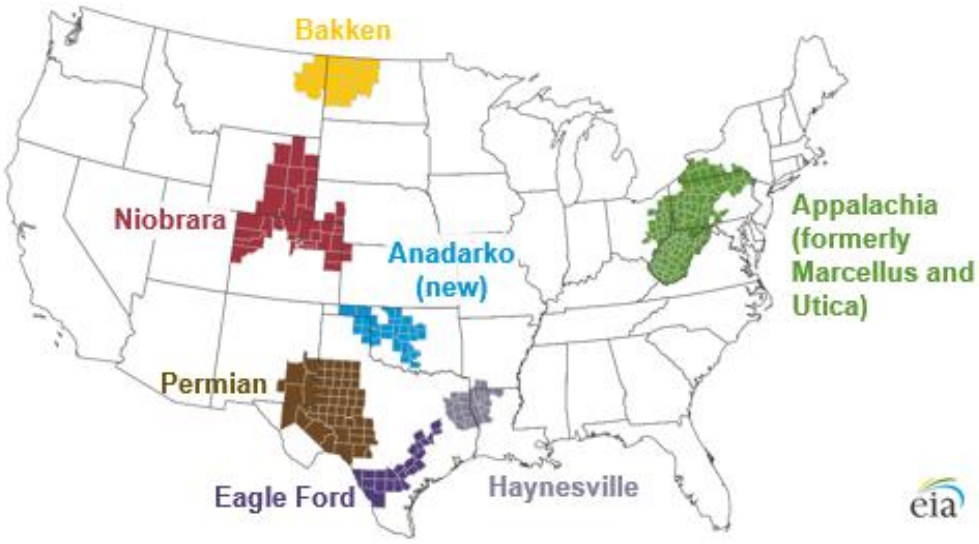



**Figure 1.** Map of shale gas plays in the continental Unites States *(Source: Energy Information Administration).*

Fluids used for hydraulic fracturing are composed of a mixture of water, proppants, and chemical modifiers (Arthur et al., 2008; King, 2012). Proppants comprise ~1 – 2 % of hydraulic fracturing fluids and are small particles, typically sand or ceramics, that hold fractures open to allow hydrocarbons to flow (King, 2012). Chemical modifiers constitute

less than 1% of hydraulic fracturing fluids, and include biocides, friction reducers, and corrosion inhibitors that are used to enhance production (King, 2012). An estimated one to four million gallons of water may be required to hydraulically fracture a single OG well (Engle et al., 2014). These large volumes of water are commonly derived from local surface waters, pumped from a municipal water source (Gregory et al., 2011), or reused for additional well completions. Of the injected fluid volume, 10 - 40 % returns to the surface as "flowback" for up to two weeks and is comprised of a mixture of components from the injected fluids and those naturally present in the geologic formation (Arthur et al., 2008; Engle et al., 2014). Following the initial period of flowback, waters continually come to the surface along with hydrocarbons and these waters are referred to as produced waters (Engle et al., 2014). These produced waters are a combination of flowback water, injected fluids for production, and/or formation waters from the reservoir. As both flowback and produced waters are operational definitions, they are referred to collectively here as OG wastewaters.

The composition of OG wastewaters changes over time, initially reflecting the chemistry of the injected fluids, then shifting to reflect the chemistry of the formation along with residual fracturing fluids (Engle et al., 2014; King, 2012; Rowan et al., 2015). OG wastewaters are typically brines with elevated levels of total dissolved solids (TDS). OG wastewaters from the Appalachian basin exhibit TDS ranging from a few thousand to around 400,000 mg/L (Barbot et al., 2013; Gregory et al., 2011; Rowan et al., 2015) (for perspective, freshwater has a TDS of <1,000 mg/L while seawater has a TDS of roughly 35,000 mg/L). OG wastewaters may also include heavy and light chain hydrocarbons from

the OG formation and naturally occurring radioactive materials (Gregory et al., 2011; Orem et al., 2017; Rowan et al., 2015; Akob et al., 2016).

## Oil and Gas Wastewater Management and Disposal

In major OG production regions, production and water use has increased for hydraulic fracturing and subsequently, OG wastewater production has increased from 2011 to 2016 (Kondash et al., 2018). The complex, variable chemistry of these wastewaters, and their large volumes (~21 billion barrels of OG wastewaters produced annually) (Clark & Veil, 2009; King, 2012; Kondash et al., 2018), pose a unique challenge regarding their proper management and disposal. Produced waters are collected from an OG well on-site, then transported by trucks or pipelines for subsequent disposal, treatment, and/or reused for further hydraulic fracturing operations (U.S. EPA, 2013). Methods for onshore produced water management include injection back into an OG well for further hydrocarbon production, or into a non-producing formation for disposal (Clark & Veil, 2009; Gregory et al., 2011). Designated Class II Underground Injection Control (UIC) wells, which make up 20% of all Class II wells in the U.S. (U.S. EPA, 2015), are often used for deep underground injection of OG wastewaters as a means of disposal (Clark & Veil, 2009; Gregory et al., 2011). However, there is limited availability of UIC disposal wells throughout the contiguous U.S. (Clark & Veil, 2009; U.S. EPA, 2015). Proper OG wastewater disposal and treatment practices may be constrained by local policies, geographic location, and available infrastructure and technology (Gregory et al., 2011).

## Environmental Impacts of Oil and Gas Wastewaters

Unintended releases of OG wastewaters, including spills and pipeline leaks, or intentional releases, e.g. illegal dumping, are of environmental concern given the water's complex chemistry. The potential environmental health effects of such unintended releases are not fully understood; however, studies have shown that accidental releases can impact environmental health by altering stream biogeochemistry, microbiology, and the health of other aquatic life (Akob et al., 2016; Cozzarelli et al., 2017). Akob et al. (2016) studied the environmental health effects of activities at an UIC OG wastewater disposal well in West Virginia, whereas Cozzarelli et al. (2017) assessed similar effects from a large pipeline leak in North Dakota. These studies showed that OG wastewater releases resulted in streams having elevated concentrations of Cl, Br, strontium (Sr), radium (Ra), and hydrocarbons downstream from a UIC well or the spill site. Surface waters elevated in these constituents also exhibited increased endocrine disruption activity in laboratory bioassays using mammalian and yeast reporter cell lines (Cozzarelli et al., 2017; Kassotis et al., 2016). These bioassays indicated potential disruption to development and/or reproduction for aquatic and human life. Hydrocarbons were detected at higher concentrations downstream as compared to upstream (Cozzarelli et al., 2017; Orem et al., 2017). At the North Dakota spill site, the changes in stream biogeochemistry were also observed in conjunction with reduced fish survival (Cozzarelli et al., 2017). Furthermore, in North Dakota, elements of the OG wastewater spill persisted for months after remediation (Cozzarelli et al., 2017).

At the West Virginia site, the microbial ecology was investigated to understand the ecological impacts from wastewater releases. Microbial communities in sediments downstream from the OG wastewater disposal facility had lower diversity as well as shifts in community composition and function (Akob et al., 2016; Fahrenfeld et al., 2017; Mumford et al., 2018). Community composition was altered in sediments downstream from the OG wastewater disposal site in West Virginia, with observed increases in members of the Deltaproteobacteria and Methanomicrobia, and decreases in Thaumarchaeota (Akob et al., 2016; Fahrenfeld et al., 2017). Metagenomic sequencing indicated potential changes in respiration and nutrient cycling, and further identified increased markers of stress in impacted stream sediments (Fahrenfeld et al., 2017). Specific gene changes included those involved in dormancy, sporulation, methanogenic respiration, and antibiotic resistance, which were reported at higher concentrations downstream from the spill (Fahrenfeld et al., 2017). Changes in the microbial community structure and function were also observed during laboratory studies of sediment microbial activity in the presence of hydraulic fracturing fluid additives (Mumford et al., 2018). When enriched in the presence of biocides, guar gum, and/or ethylene glycol, microbial communities from sediments both downstream (impacted) and upstream (background) of the OG wastewater disposal facility displayed shifts in community composition. Changes in differential abundance of microbial taxa during the experiment were accompanied by alterations in potential microbial function, including varied iron- and sulfate-reducing activity, and increased dehalogenation potential in the face of highly saline OG wastewater releases and halogenated biocides (Mumford et al., 2018). These results highlight the

sensitivity of microbial community structure and function to even small amounts of hydraulic fracturing fluid additives, many of which are found residually in OG wastewaters. Microorganisms have the potential to act as bioindicators of OG wastewater releases in the environment, and may aid in the remediation of organic compounds from the releases in impacted ecosystems.

## Illegal Dumping of OG Wastewaters in the Permian Basin

The Permian Basin is the largest technically recoverable OG reserve in the continental United States (Gaswirth et al., 2018). The geological province spans the northern reaches of the Chihuahuan Desert ecoregion in New Mexico and Texas, and lies under semi-arid, desert habitat (Figure 1). The Chihuahuan Desert, which is the largest desert in North America, exhibits high biodiversity, and provides key ecosystem services through the contributions of native plants, raw materials, and climate regulation, as well as cultural and commercial services (Taylor et al., 2017). Local ranchers also rely on the region's plant life as a food source for grazing livestock. The recent energy boom in the Permian Basin has raised local concern over the environmental health impacts of OG development on the desert ecosystem and the services it provides (*Chihuahuan Desert Ecoregion (U.S. National Park Service)*, 2018).

Hydrocarbon production in the region began in the early 1900's and initially focused on conventional OG. Starting in the 2000's, production in the Permian Basin has expanded to unconventional resources, primarily to tight oil reservoirs including the Bone

Spring and Wolfcamp formations which, since 2015, have become the more common source of oil produced (U.S. Energy Information Administration, 2019). The production of tight oil, a type of hydrocarbon trapped in impermeable shale and limestone rock deposits, has resulted in an increase in OG production in the Permian Basin from 8200 $m^3$ per well in 2011 to 16,700 $m^3$ per well in 2016, a 130% increase (Kondash et al., 2018). This spike in production has been accompanied by a 770% increase in water use for OG recovery (more than 42,500 $m^3$ of water per well), and an increase in wastewater discharge (Kondash et al., 2018). In southeast New Mexico, 1 billion barrels (or 42 billion gallons) of produced water was generated by the OG industry in 2018 alone (*New Mexico Environment Department New Mexico Produced Water*, 2019).

A large portion of OG leases in the Permian Basin are on public lands managed by the Bureau of Land Management (BLM), which has over 45 million acres of mineral estate under its jurisdiction (Bureau of Land Management, 2016). Starting in November 2017, the BLM identified at least 39 illegal OG wastewater dumps in southeastern New Mexico, which released approximately 4,000 barrels of OG drilling and production wastewaters onto arid, desert lands. Despite anecdotal reports of livestock and wildlife kills from exposure to spilled OG wastewater, little is known about the environmental impacts of OG wastewater spills on arid lands. A team of U.S. Geological Survey (USGS), BLM, and university researchers are seeking to 1) identify the sources of spilled OG wastewaters, e.g., the specific reservoir for production, and 2) evaluate the impacts resulting from these OG wastewater spills by analyzing changes in soil geochemistry and microbial community structure.

***Research Questions and Hypothesis***

Approximately 4 months after the dump sites were initially identified and mapped, a total of 5 dump sites were sampled for geochemistry and microbial community analysis (Figure 2). These OG wastewater dump sites were characterized by visible salt crusts and hydrocarbons on foliage and soil surfaces. We investigated the responses of arid soil microbial communities to inputs of OG wastewater from the illegal dumping activities that occurred on BLM lands in the Permian Basin. This study hypothesizes that microbial communities in OG wastewater dump-affected soils will be impacted by OG wastewater geochemistry and will take on a composition distinct from that of unaffected arid soils. We test this hypothesis by comparing the geochemistry and microbial community structure of dump-affected soils to that of unaffected control soils at 5 OG wastewater dump sites. Specifically, we addressed the following questions:

**1)** What is the composition of microbial communities in control vs. dump-affected soils at illegal OG wastewater dump locations?

**2)** How are potential microbial community functions, e.g., hydrocarbon degradation, affected by OG wastewater inputs?

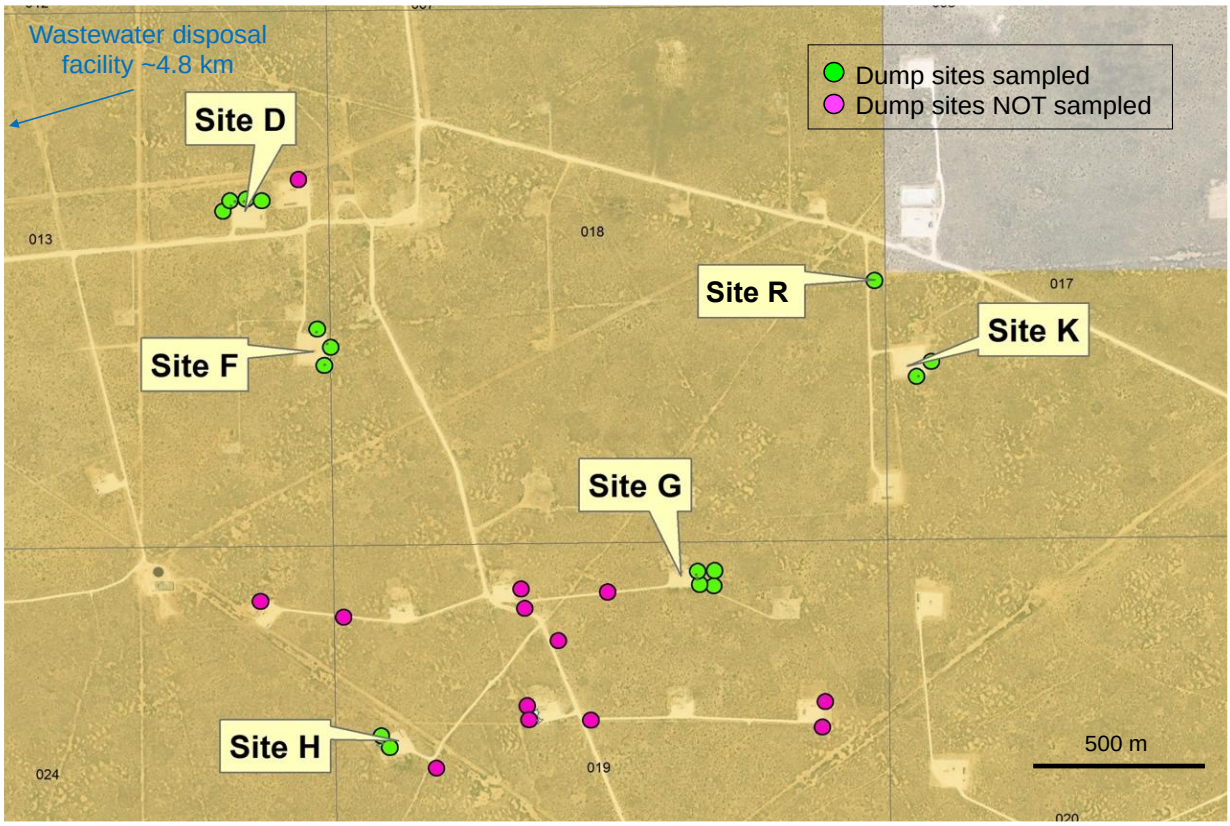



**Figure 2.** Map of OG wastewater dump sites in Permian Basin study area. Green dots indicate locations where sample were collected, while pink dots represent sites which were not sampled. Site R was excluded from future analysis due to chemistries inconsistent with an OG wastewater dump site.

# MATERIALS AND METHODS

## Sample Collection

In March of 2018, five sites were sampled by a USGS and BLM field team which had been previously identified as dump sites and mapped by Terry Gregston (BLM) (Figure 2). At each site, the areas impacted by the dump (dump zone) and areas outside of the dump (control zone) were identified visually by the presence or absence of salt crusts and hydrocarbons. Within each zone, a minimum of 3 replicate soil samples were collected using aseptic techniques for geochemistry and microbiology analysis. Soils were collected from the upper 0 – 5 cm using sterile scoops for surface grab samples. Additional samples were collected from select sites using a soil corer to characterize microbial communities and geochemistry at 0-10 cm below surface. Samples for geochemical analysis were stored in iChem jars (Thomas Scientific, Swedesboro, NJ) on ice in the field. Samples for microbial community analysis were stored in sterile Whirl-Pak® bags and frozen with dry ice in the field. In the lab, geochemistry samples were stored at 4ºC and microbial samples were stored at -80ºC.

**Soil Geochemical Analysis**

Soils were processed to measure percent (%) moisture, % carbon (C) and nitrogen (N), and water extractable anions and cations. For % moisture, C, and N analysis, easily visible leaf & twig parts (>3 mm) were removed from homogenized soil samples. From each field replicate, triplicate soil subsamples were weighed, and dried at 50ºC until reaching a steady dry weight to determine the % water content. For % C and N, approximately 8-13 mg of the dried soil samples were weighed into silver (Ag) cups, which were then inserted into a tin (Sn) cup and sealed. Duplicates of four samples (D-S-06, K-S-04, K-S-05, and D-S-10) were analyzed to check for reproducibility. The % C and N were analyzed using a ThermoScientific Flash 2000 Elemental Analyzer.

For the analysis of water extractable anions and cations, duplicate soil subsamples were extracted (1:1 weight per weight) in ultra-pure water by incubating for 1 hour at room temperature while shaking. After incubation, the soil:water solutions were centrifuged for 10 min at 5,000 x g and supernatants were 0.45 µm filtered. Conductivity and pH were measured from soil:water extracts using an InLab™ conductivity probe and Micro pH electrode, respectively on a SevenCompact™ Duo S213 pH/Conductivity Meter (Mettler Toledo™, Columbus, OH). Water extract samples for anion measurements were stored at 4ºC, whereas, samples for cations were preserved at pH < 2 with Optima-grade nitric acid ($HNO_3$) prior to storage at 4ºC. Major cation concentrations in water extracts were determined using inductively coupled plasma-optical emission spectroscopy (ICP-OES) (Perkin-Elmer Optima 4300 ICP-OES, Perkin-Elmer, Waltham, MA) and inductively

coupled plasma-mass spectrometry (ICP-MS) (Perkin-Elmer ELAN 9000 ICP-MS, Perkin-Elmer, Waltham, MA). Anion concentrations in extracts were determined using ion chromatography (Dionex$^{TM}$ ICS 1000 IC) with an electrochemical detector and AS14 column.

**Microbial Community Characterization**

*DNA Extraction & Illumina Sequencing.* The soil microbial community composition was analyzed using Illumina 16S iTag sequencing to capture the phylogenetic composition of the whole community. DNA was extracted using the PowerSoil® DNA Isolation kit (Qiagen) (MO BIO Laboratories, Carlsbad, CA), according to the manufacturer's instructions. Extracted DNA was quantified using Quant-iT PicoGreen® dsDNA assay kit (Invitrogen, Carlsbad, CA), following the procedure described in Blotta et al. (2005) with slight modifications. In brief, samples were quantified with PicoGreen in triplicate in a 96-well plate format using a CFX Connect Real-Time PCR Detection System (Bio-Rad Laboratories, Inc., Hercules, CA) with ROX as internal standard. Genomic DNA was diluted to a final concentration of 10 ng/µL using Ambion® nuclease-free water (Ambion, Inc., Austin, TX) in preparation for amplicon sequencing. Diluted samples were sent to the Michigan State University Genomics Core Facility for Illumina 16S iTag sequencing (Illumina, Inc., San Diego, CA). Libraries were prepared using the dual-indexed 515f (5'- GTGCCAGCMGCCGCGGTAA-3') and 806r (5'- GGACTACHVGGGTWTCTAAT-3') Illumina compatible primers (Caporaso et al.,

2011) following the Schloss lab protocol (Kozich et al., 2013). PCR products were batch normalized using a SequalPrep™ Normalization Plate Kit (Thermo Fisher Scientific, Waltham, MA) and product recovered from the plate pooled. The pool was quality controlled and quantified using a combination of Qubit™ dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA), Agilent 4200 TapeStation System HS DNA1000 (Agilent Technologies, Santa Clara, CA) and Kapa Library Quantification Kit qPCR assays (Roche Holding AG, Basel, Switzerland). This pool was then loaded onto an Illumina MiSeq v2 flow cell and sequencing was performed in a 2x250bp paired end format using a MiSeq Reagent Kit v2 (500-cycles) (Illumina, Inc., San Diego, CA). Custom sequencing and index primers complementary to the 515f and 806r target specific sequences were added to appropriate wells of the reagent cartridge. Base calling was done by Illumina Real Time Analysis (RTA) v1.18.54 and output of RTA was demultiplexed and converted to FASTQ format with Illumina bcl2fastq v2.19.1.

*Sequence Processing.* Initial quality control, alignment, and taxonomic assignment of microbial sequence data were performed using mothur v1.42.3 and 1.43.0 (Schloss et al., 2009) according to the mothur MiSeq standard operating procedure (Kozich et al., 2013) and using the USGS Advanced Research Computing (ARC) Yeti high-performance computing facility. Reads from paired-end fastq files were merged into contiguous sequences with the 'make.contigs' command using 64 processors. Sequences from the fasta file were then aligned with the Silva 132 non-redundant database (Quast et al., 2013) using 'align.seqs'. Chimeras were checked using VSEARCH (Rognes et al., 2016) in mothur (command: 'chimera.vsearch'). Operational taxonomic units (OTUs) were generated based

on a 97% similarity cutoff against the Silva 132 non-redundant database (Quast et al., 2013) using the VSEARCH clustering algorithm (Rognes et al., 2016) (command: cluster.split); consensus taxonomies were assigned using the command 'classify.otu'. Unwanted lineages (mitochondrial, eukaryotic, chloroplast, and unknown) were removed using the command 'remove.lineage'. Representative sequences for each OTU were found (mothur; 'get.oturep'), and singleton sequences were removed from all samples using 'remove.rare'. Samples were rarefied to the smallest library size (109,370 sequences) using the phyloseq package (McMurdie & Holmes, 2013) in R (R Core Team, 2014).

Statistical and diversity analyses were performed in R using the vegan (version 2.5.6), phyloseq (1.26.1), and DESeq2 packages (Love et al., 2014; McMurdie & Holmes, 2013; Oksanen et al., 2019). Alpha diversity measures (number of OTUs, Inverse Simpson index, and Shannon diversity index) were calculated for the subsampled libraries without singleton OTUs using the 'estimate_richness' command. Alpha diversity metrics were visualized and subjected to non-parametric Mann-Whitney U tests to determine statistically significant differences between dump-affected and control communities using Prism version 8.4.2 (GraphPad Software, La Jolla, CA).

To determine if microbial community structure in dump soils differed from that of control soils, a UniFrac distance matrix (Lozupone & Knight, 2005) was constructed and used for multivariate analysis. An unrooted maximum likelihood tree was created using the FastTree algorithm (Price et al., 2010) in Geneious Prime version 2019.2.1 (https://www.geneious.com) which resulted in 38,216 tips and 38,214 internal nodes across 33 samples. The maximum likelihood tree was randomly rooted in phyloseq (rngseed =

711) and the weighted UniFrac distance matrix was calculate using the 'phyloseq:distance' command to consider the relatedness of microbial taxa within dump and control soils. To elucidate whether OG wastewater dump-affected soil communities differed significantly from control soil communities, the weighted Unifrac distance matrix was subjected to a Permutational Multivariate Analysis of Variance (PERMANOVA) with condition as a factor, and 999 permutations using the 'adonis' function in vegan. A non-metric multidimensional scaling (NMDS) ordination was used to visualize the distance matrix. To determine which geochemical factors significantly co-varied with microbial community structure, geochemical vectors were fitted onto the NMDS ordination using 'envfit' in vegan, with 999 permutations. The R package ggplot2 version 3.2.1 (Wickham, 2016) was used to visualize the ordination, with confidence interval ellipses set at 95%.

To calculate relative abundance of various taxa the number of reads for each taxa (after rarefaction) was divided by by the total number of reads per sample (i.e. 109,370 reads) for each individual sample (n=33). Then, the resulting percent abundances were averaged across all dump (n=18) and all control (n=15) samples. Percent relative abundance was visualized at the phyla level using Excel. To determine significant differential abundance in microbial taxa between controls and dumps, DESeq2 was run on the unrarefied dataset in R, and the $\log_2$fold changes were visualized in Prism version 8.4.2 (GraphPad Software, La Jolla, CA).

# RESULTS

## Differences in Geochemistry Between Dump and Control Soils

Soil geochemistry differed significantly between dump and control soils (Figures 3 and 4). Percent moisture and % C in bulk soil measurements were significantly elevated in dump affected soils (1.96 ± 1.70 % moisture and 1.56 % C ± 1.39), while % N did not significantly differ between dump and control soils (Figure 3).

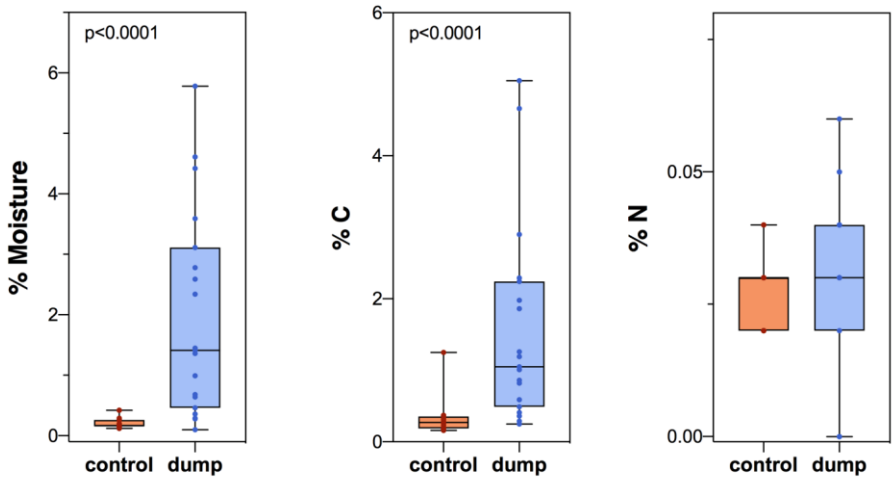



**Figure 3.** Comparisons of soil properties between dump and control soils across all sampling sites. The lower and upper hinges correspond to the 25th and 75th percentiles, respectively, and the whiskers correspond to minimum and maximum values with the middle line being the median. The percent moisture, and percent carbon were significantly elevated in dump affected soils, while the percent nitrogen between dump and control soils was not significantly different.

Soil conductivity, measured on soil:water extracts, was elevated in dump soils ranging from 0.517 to 30.8 mS/cm compared to control soils with a range of 0.008 to 1.96 mS/cm (Figure 4). The water extractable anions and cations that had significantly elevated concentrations (measured in mg/kg) in dump soils included sodium (Na), chloride (Cl), sulfate ($SO_4^{2-}$), bromide (Br), strontium (Sr), lithium (Li), boron (B), and barium (Ba). These elements were previously identified as useful tracers of OG wastewater releases into streams (Akob et al., 2016; Cozzarelli et al., 2017; Orem et al., 2017). Sodium and Cl were present at orders of magnitude greater in dump soils as compared to controls (Figure 4), with dump soils having an average of 3,524 ± 2,124 mg/kg Na and 7,217 ± 4,424 mg/kg Cl (control soils had an average of 1.03 ± 1.86 mg/kg of Na and 2.84 ± 4.44 mg/kg of Cl).

Dump soils had an average sulfate concentration of 224 ± 141 mg/kg while control soils had a significantly ($p<0.0001$) lower concentration with an average 2.88 ± 0.91 mg/kg. Bromide was also elevated in dump soils with a concentration range of 11.8 to 127 mg/kg, while concentrations in control samples were all below the detection limit of 0.160 mg/kg. Strontium concentrations were significantly ($p<0.0001$) higher in dump soils as well, with an average of 13.7 ± 10.8 mg/kg, compared to control concentrations which averaged 0.08 ± 0.04 mg/kg. Low concentrations of lithium (Li) were also detected in dump soils ranging from 0.11 to 1.37 mg/kg, while control Li concentrations were all below the detection limit of 0.02 mg/kg except for a replicate in Site H which had a Li concentration of 0.03 mg/kg. Low concentrations of boron (B) were also detected, averaging at 3.43 ± 1.53 mg/kg in dump soils.

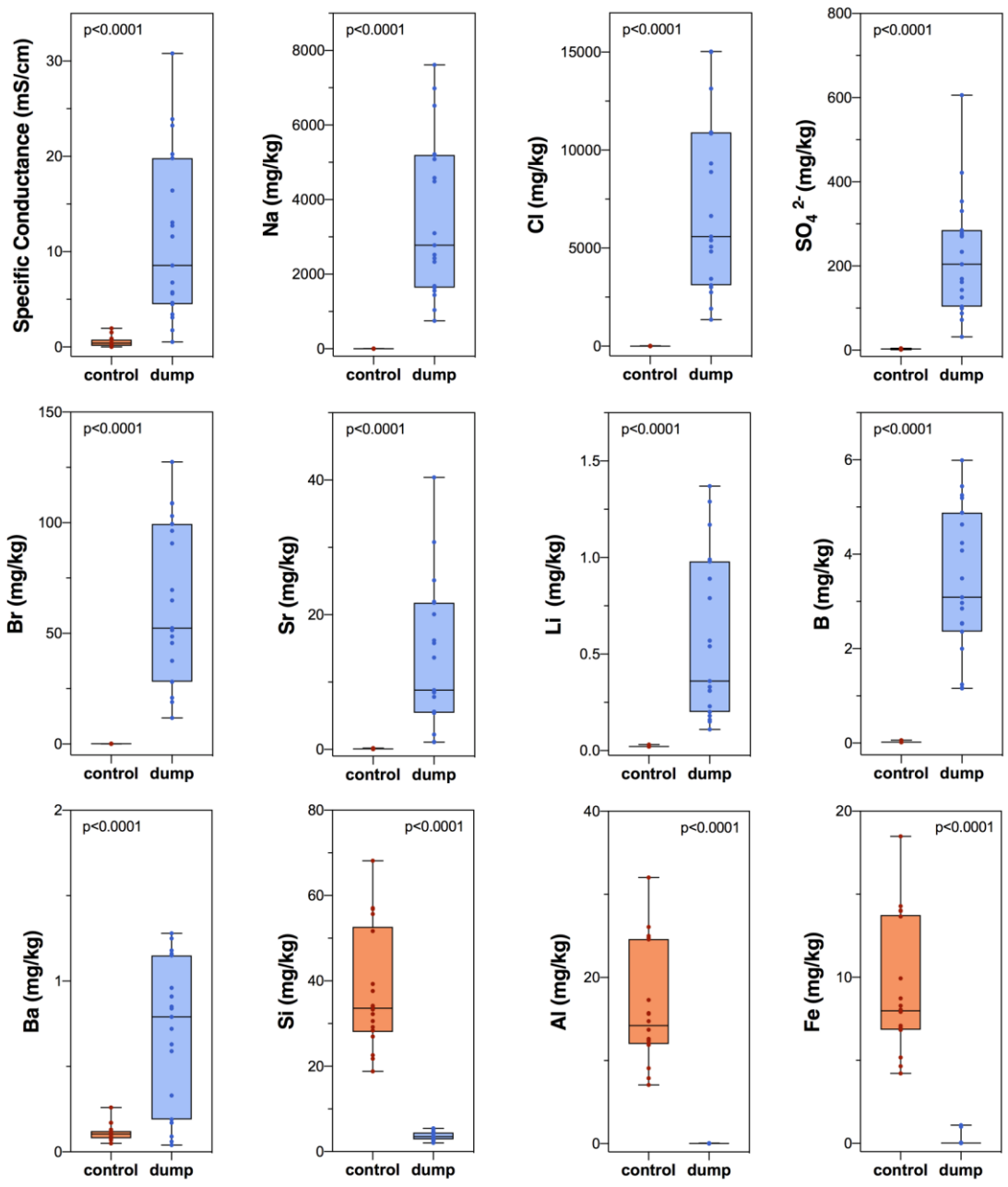



**Figure 4.** Comparisons of anions and cations in soil water extracts between dump and control soils across all sampling sites. The lower and upper hinges correspond to the 25th and 75th percentile, respectively, and the whiskers correspond to minimum and maximum values with the middle line being the median.

Almost all control soils had B concentrations below the detection limit of 0.02 mg/kg with the exception of samples D-C-03, and D-S-01, which had concentrations of 0.06 and 0.02 mg/kg, respectively (Table 1). Barium was also elevated in dump soils (average of 0.69 ± 0.43 mg/kg) compared to control soils (average of 0.11 ± 0.05 mg/kg).

Additional constituents that were elevated significantly in OG wastewater dump-affected soils included calcium (Ca), potassium (K), magnesium (Mg), and zinc (Zn) (Table 1). Calcium, which was present in control soils at low concentrations ranging from 8.27 to 20.3 mg/kg, was found at significantly higher concentrations in dump soils, ranging from 43.6 to 1002 mg/kg. Potassium was also observed at minimal concentrations in control soils (average of 7.38 ± 1.77 mg/kg) but was elevated significantly in dump soils (average 98.2 ± 49.3 mg/kg). Similarly, Mg and Zn, both present at very low concentrations in control soils (averages of 4.01 ± 1.06 and 0.04 ± 0.01 mg/kg respectively), were detected at elevated concentrations in dump soils (averages of 67.1 ± 44.7 and 0.11 ± 0.05 mg/kg). Manganese (Mn) ranged from 0.08 to 0.38 mg/kg in control soils and 0.02 to 2.01 mg/kg in dump soils and was not significantly elevated in dump soils.

Control soils were marked by significantly elevated silicon (Si), aluminum (Al), and iron (Fe) concentrations. Si, Al, and Fe averaged 3.66 ± 0.97, 0.05± 0.02, and 0.40 ± 0.61 mg/kg in control soils, respectively. In dump-affected soils, Si, Al, and Fe had concentrations of 37.7 ± 14.2, 16.4 ± 7.12, and 9.22 ± 3.99, respectively (Table 1). In addition to differences in geochemistry on an element-by-element basis, overall soil geochemical composition in control samples were more similar to each other, while greater

variability was observed in geochemical composition across all dump samples, as seen in Figure 4 and Table 1.

**Table 1.** Geochemistry of control and dump-affected soils with average ± standard deviations (SD) of water extractable specific conductance, anion and cation concentrations (mg/kg) and % moisture, carbon, and nitrogen. Average concentrations below detection limits denoted with "<". Variations in detection limits within an analyte are due to varying dilution factors as a result of potentially elevated concentrations in dump samples. Conductivity, pH, anions, and cations were measured on soil:water extracts.

| | **Control (average ± SD)** | **Dump (average ± SD)** |
|---|---|---|
| **pH** | 7.57 ± 0.31 | 7.71 ± 0.20 |
| **Specific conductance (mS/cm)** | 0.51 ± 0.53 | 11.4 ± 8.77 |
| **$F^-$** | 0.72 ± 0.53 | <24.8 |
| **$Cl^-$** | 2.84 ± 4.44 | 7217 ± 4424 |
| **$Br^-$** | <0.160 | 63.8 ± 36.0 |
| **$NO_3^-$** | 2.93 ± 2.42 | <11.3 |
| **$PO_4^{3-}$** | 1.78 ± 1.57 | <83.8 |
| **$SO_4^{2-}$** | 2.85 ± 0.91 | 223 ± 141 |
| **Ca** | 14.4 ± 3.56 | 484 ± 317 |
| **Na** | 1.03 ± 1.86 | 3524 ± 2124 |
| **Mg** | 4.01 ± 1.06 | 67.1 ± 44.7 |
| **K** | 7.38 ± 1.77 | 98.2 ± 49.3 |
| **Si** | 37.7 ± 14.2 | 3.66 ± 0.97 |
| **Sr** | 0.076 ± 0.04 | 13.7 ± 10.8 |
| **Al** | 16.4 ± 7.12 | 0.05 ± 0.02 |
| **Fe** | 9.22 ± 3.99 | 0.40 ± 0.61 |
| **Mn** | 0.22 ± 0.10 | 0.50 ± 0.51 |
| **B** | 0.035 ± 0.02 | 3.43 ± 1.53 |
| **Ba** | 0.11 ± 0.05 | 0.69 ± 0.43 |
| **Li** | 0.02 ± 0.00 | 0.58 ± 0.42 |
| **Cu** | <0.02 | 0.03 ± 0.01 |
| **Zn** | 0.04 ± 0.01 | 0.11 ± 0.05 |
| **% Moisture** | 0.21 ± 0.09 | 1.96 ± 1.70 |
| **% C** | 0.32 ± 0.24 | 1.56 ± 1.39 |
| **% N** | 0.03 ± 0.01 | 0.03 ± 0.02 |

Phosphate and nitrate were both detected at concentrations below the detection limit across most dump samples (Table 1). In the OG wastewater dump-affected samples, nitrate concentrations were below the detection limit of 11.3 mg/kg, with the exception of sample G-S-06 (557 mg/kg nitrate). Control samples had an average of 2.96 ± 2.42 mg/kg of nitrate. The majority of the control samples were observed to have phosphate values below the detection limit of 1.68 mg/kg, with the exception of sample K-S-05 with 2.90 mg/kg of phosphate detected. Similarly, all of the dump samples had phosphate concentrations below the detection limit of 83.8 mg/kg. It is important to note that differences in the detection limits between the control and dump sites are due to differences in dilutions between control and dump soil:water extracts.

## Microbial Community Changes Due to OG Wastewater Dumping

### *Changes in Diversity and Community Structure Due to OG Wastewater Dumping*

Soil microbial communities were characterized using Illumina 16S iTag sequencing and yielded between 109,370 to 185,009 reads (Table 2). There were a total of 4,477,573 reads across all 33 samples prior to rarefying which binned into 43,203 unique OTUs. After rarefying to an even sequencing depth of 109,370 sequences per sample (the smallest library size) the data set contained a total of 3,609,210 total reads and 38,216 unique OTUs. Significant variation in alpha diversity measures were observed between dump and control samples (Figure 5).

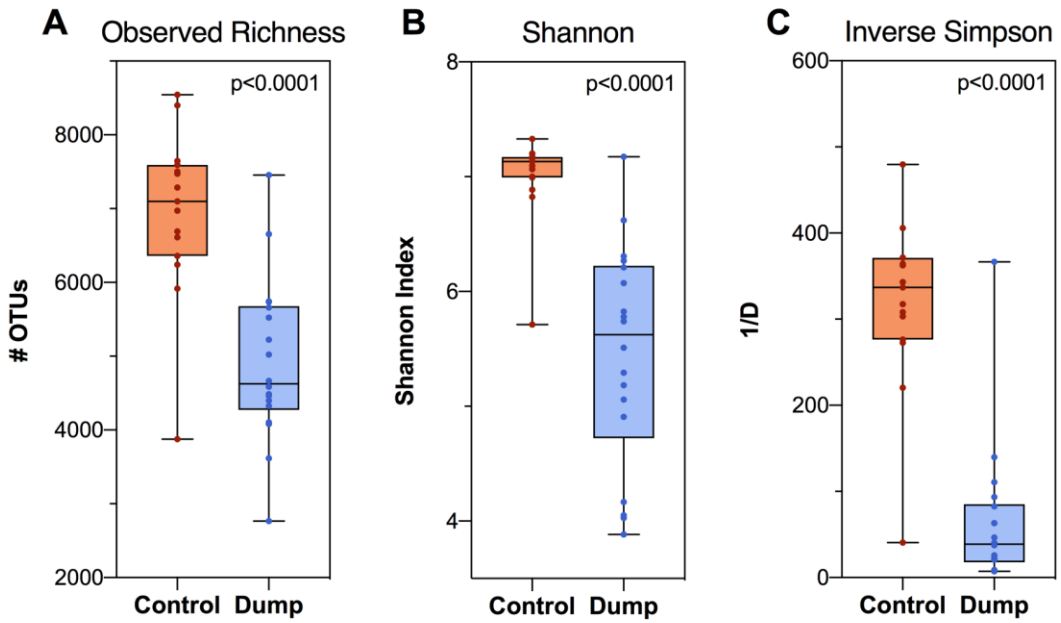



**Figure 5.** Alpha diversity measures (A) observed OTU richness, (B) Shannon diversity index, and (C) Inverse Simpson diversity index comparing dump and control soil microbial communities. Each sample is represented as a dot on the box and whisker plot. P-values correspond to non-parametric pairwise Mann-Whitney rank-sum test. The lower and upper hinges correspond to the 25th and 75th percentile, respectively, and the whiskers correspond to minimum and maximum value. Across all four measures, a significant decrease in alpha diversity was observed in dump soil communities. OTU richness was determined using raw counts of rarefied dataset.

Pairwise comparisons using the non-parametric pairwise Mann-Whitney rank-sum test demonstrated that observed richness, and Shannon and Inverse Simpson diversity indices between control and dump communities were all significantly different ($p<0.0001$) (Figure 5). Control soils displayed an average of 6,947 observed OTUs across 15 samples, while dump-affected soils had significantly less observed OTU richness, with an average of about 4,917 OTUs across 18 samples (Figure 5). Shannon and Inverse Simpson alpha diversity metrics also showed significant differences in control versus dump microbial communities.

**Table 2.** Sequence summary for control and dump soils. Number of reads and coverage were calculated before rarefaction whereas observed OTUs and diversity were calculated post-rarefaction.

| Condition | Sample Name | Number of Reads | Coverage | Observed OTUs | Shannon Diversity | Inverse Simpson Diversity |
|---|---|---|---|---|---|---|
| Control | D-C-03 | 129,020 | 0.98 | 8,543 | 7.33 | 478 |
| | D-S-01 | 113,845 | 0.98 | 6,971 | 7.17 | 406 |
| | D-S-02 | 153,891 | 0.98 | 7,472 | 7.16 | 318 |
| | F-S-01 | 124,002 | 0.98 | 6,360 | 6.99 | 303 |
| | F-S-02 | 134,891 | 0.99 | 6,690 | 7.06 | 371 |
| | F-S-03 | 117,997 | 0.98 | 7,593 | 7.13 | 364 |
| | G-S-07 | 116,912 | 0.98 | 6,611 | 7.00 | 276 |
| | G-S-08 | 127,885 | 0.98 | 6,240 | 6.83 | 221 |
| | H-S-01 | 129,337 | 0.99 | 7,098 | 7.20 | 337 |
| | H-S-02 | 139,099 | 0.99 | 7,503 | 7.17 | 363 |
| | H-S-03 | 125,000 | 0.98 | 7,286 | 7.09 | 273 |
| | K-S-01 | 134,178 | 0.98 | 7,645 | 7.14 | 343 |
| | K-S-03 | 157,047 | 0.98 | 8,400 | 7.20 | 371 |
| | K-S-04 | 131,832 | 0.98 | 5,918 | 6.89 | 308 |
| | K-S-05 | 118,805 | 0.99 | 3,875 | 5.71 | 40.5 |
| Dump | D-C-08 | 130,881 | 0.99 | 4,328 | 4.17 | 7.50 |
| | D-S-06 | 112,653 | 0.98 | 6,655 | 6.62 | 140 |
| | D-S-07 | 148,351 | 0.99 | 5,022 | 5.74 | 40.2 |
| | D-S-11 | 161,727 | 0.99 | 4,082 | 4.05 | 8.80 |
| | D-S-12 | 160,342 | 0.99 | 5,224 | 6.21 | 111 |
| | D-S-13 | 152,381 | 0.99 | 4,667 | 4.91 | 22.4 |
| | F-S-04 | 127,482 | 0.99 | 4,400 | 5.18 | 37.5 |
| | F-S-05 | 141,933 | 0.99 | 4,484 | 5.06 | 21.8 |
| | F-S-06 | 122,182 | 0.98 | 5,522 | 5.82 | 63.2 |
| | G-C-03 | 116,112 | 0.99 | 3,618 | 5.78 | 46.6 |
| | G-S-01 | 109,370 | 0.99 | 4,587 | 6.27 | 82.5 |
| | G-S-02 | 125,993 | 0.99 | 4,463 | 5.51 | 20.6 |
| | G-S-06 | 126,756 | 0.99 | 5,742 | 6.07 | 63.3 |
| | H-S-04 | 140,737 | 0.99 | 4,103 | 4.03 | 7.20 |
| | H-S-05 | 185,009 | 0.99 | 5,736 | 6.31 | 93.5 |
| | H-S-06 | 171,670 | 0.99 | 2,767 | 3.88 | 8.20 |
| | K-S-02 | 153,911 | 0.99 | 5,659 | 5.29 | 25.7 |
| | K-S-06 | 136,342 | 0.98 | 7,457 | 7.17 | 367 |

The Shannon diversity index was significantly lower for the dump soil microbial community (average of 5.45 ± 0.93) compared to control soil communities (average of 7.01 ± 0.37) (Figure 5) A similar trend was observed with the Inverse Simpson diversity measure, with a decrease in dump affected (mean of 64.76 ± 82.21) compared to control (mean 318.24 ± 94.96) microbial community diversity (Figure 5)

In addition to significant differences in alpha diversity between dump-affected and control soil microbial communities, beta diversity analyses indicated changes in community structure. A PERMANOVA test indicated that a significant variation in microbial community composition existed between soil conditions (i.e. OG wastewater dump-affected soils and controls; $p<0.001$, $r^2=0.362$). The variation in microbial community structure between control and dump-affected soils was visually represented using a non-metric multidimensional scaling (NMDS) ordination. Control communities showed distinct clustering along the NMDS axis 1, with some variation along the NMDS axis 2 (Figure 6A). Dump-affected soil microbial communities displayed greater heterogeneity across dump sites, with a wider distribution along both NMDS axes 1 and 2 (Figure 6A). When mapped onto the NMDS ordination, several geochemical constituents that were shown to be tracers of OG wastewaters (Akob et al., 2015; Engle et al., 2014, 2016; Rowan et al., 2015) were found to co-occur with significant shifts in dump-affected soil microbial community structure (Figure 6B, Table S5). This included significant co-occurrence with elevated concentrations of Na, Cl, K, Ba, Br, and B ($p<0.01$, $r^2 = 0.509$). Meanwhile higher concentrations of Al, Fe, and Si co-occurred together with control soil microbial community structure.

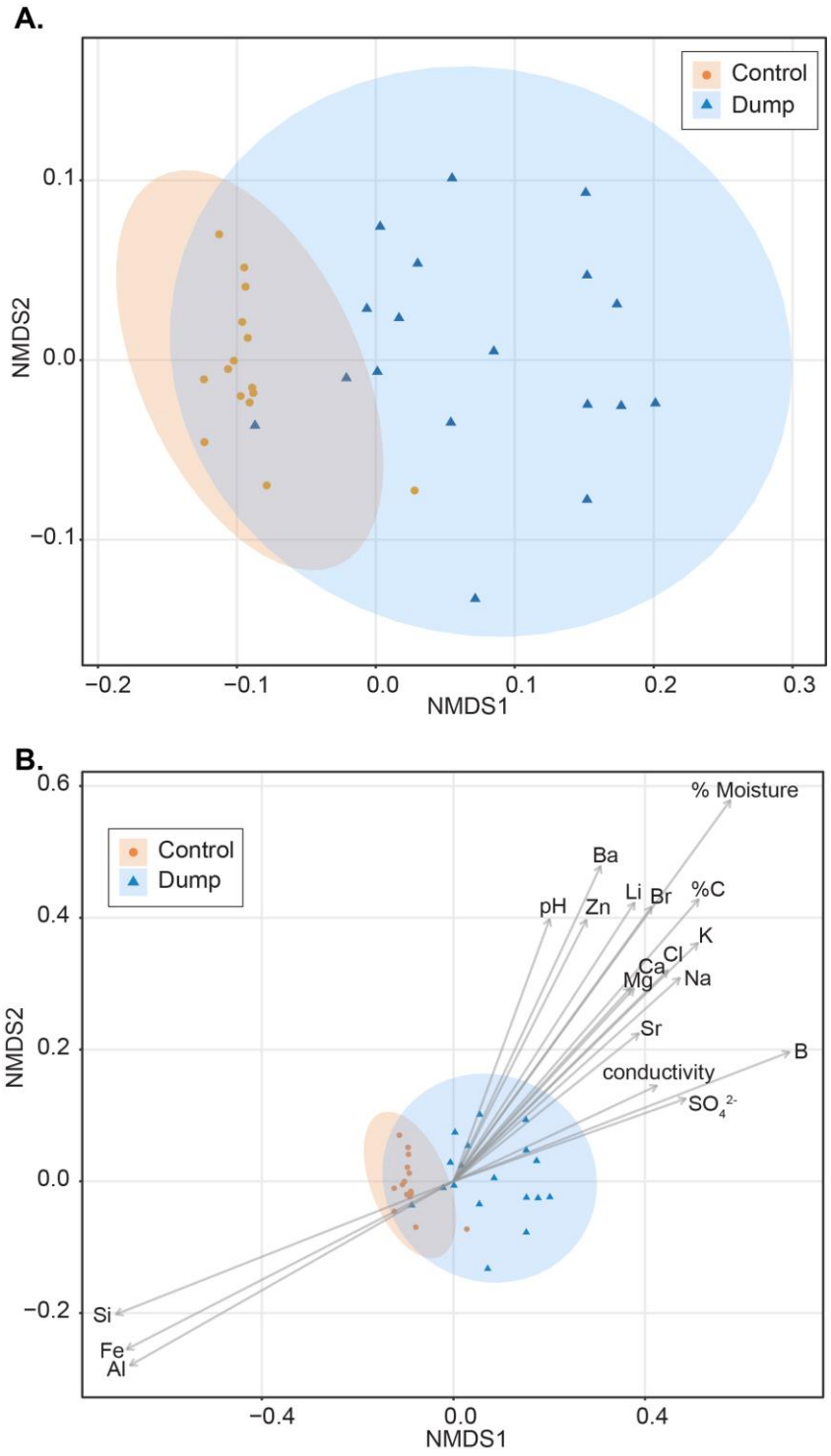



**Figure 6.** NMDS of A) weighted UniFrac distances among control and dump-affected soil microbial communities and B) with significant (ENVFIT, $p < 0.05$) co-occurring geochemical vectors fitted. Ellipses represent 95% confidence intervals. Significant differences in community structure were determined using PERMANOVA with soil condition as a factor, and 999 permutations. Vectors are the geochemical constituents plotted in Figures 3 and 4.

We found that B, % moisture, and %C were the strongest predictors of shifting microbial communities in dump soils ($p<0.001$), followed by Na, Cl, Br, Ba, and Li ($p<0.01$). These trends were also reflected by the length of the vectors (arrows) mapped onto the NMDS (Figure 6B), with longer arrows indicating stronger predictors of a shifting microbial community. The relationship of $NO_3^-$, Mn, and % N to microbial community structure was not significant, suggesting these constituents did not play a significant role in affecting microbial community shifts in dump soils.

### *Phylum and Class Level Differences Between Dump and Control Soils*

Reflecting the changes in community structure, the phylogenetic composition of microbial communities in dump-affected soils differed from that of unimpacted control soils. Microbial communities were dominated by members of the domain Bacteria (average 98 ± 0.01 % of total reads in dump soils, and 96 ± 0.01 % in control soils) while only ~ 1-4% of reads across both conditions were affiliated with the domain Archaea (Figure 7, Table 3). Archaea were predominantly comprised of the phylum Euryarchaeota, with 0.41 ± 0.002 % relative abundance in control soils and 0.20 ± 0.001 % in dump soils. Fourteen Bacteria phyla were detected at greater than 1 % abundance across all dump and control samples: Proteobacteria, Actinobacteria, Acidobacteria, Bacteroidetes, Chloroflexi, Planctomycetes, Verrucomicrobia, Thaumarchaeota, Gemmatimonadetes, an unclassified Bacterial phylum, Firmicutes, Cyanobacteria, Epsilonbacteraeota, and Patescibacteria (Figure 8). Proteobacteria, Firmicutes, Epsilonbacteraeota, and Patescibacteria relative

abundances were elevated in dump soils relative to controls. In contrast, the phyla Actinobacteria, Acidobacteria, Bacteroidetes, Chloroflexi, Planctomycetes, Verrucomicrobiota, Thaumarchaeota, Gemmatimonadetes, and Cyanobacteria all had elevated relative abundances in control soils compared to dump soils (Figure 7, Table 3).

The most abundant phyla across both dump-affected and control soils were Proteobacteria, with relative abundances of 50.6 ± 71.5 % in dump-affected soils and 28.7 ± 29.8 % in control samples (Figure 7). Three Proteobacteria classes, Alphaproteobacteria, Deltaproteobacteria, and Gammaproteobacteria, in addition to unclassified Proteobacteria were present in both conditions (Table 3). Gammaproteobacteria dominated dump-affected soils with a relative abundance of 35.3 ± 48.2% (compared to 6.65 ± 8.88% in

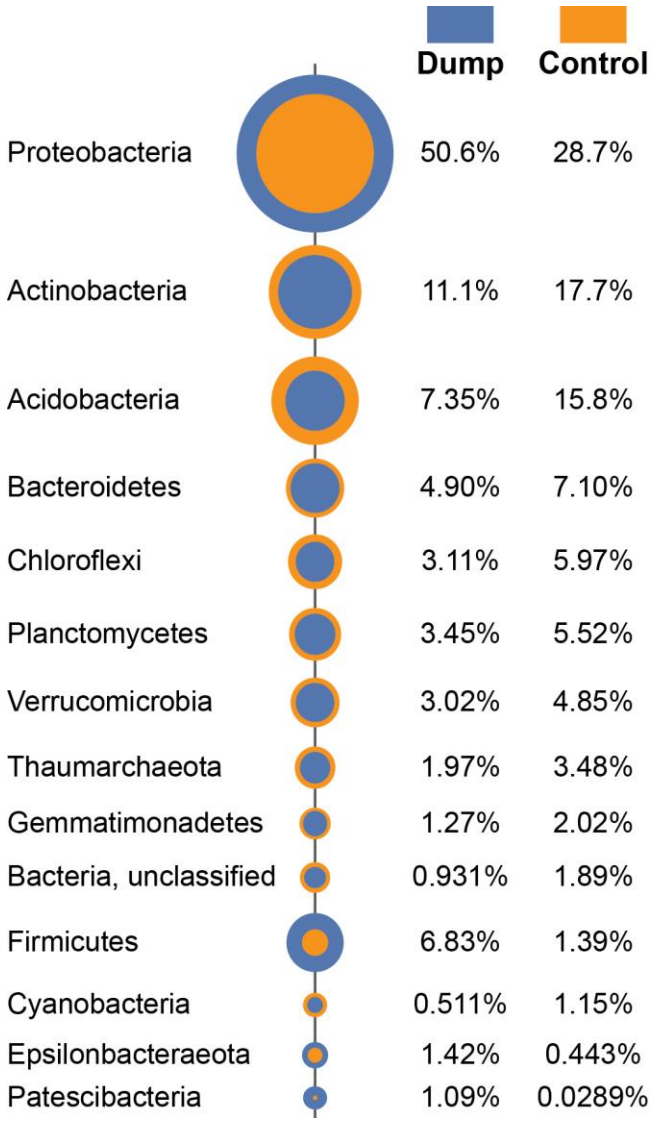



**Figure 7.** Relative abundance of OTUs at the phylum level in all dump soil samples compared to all control samples. Only phyla with relative abundances of greater than 1% in at least one control and/or dump sample are shown. Size of circles correspond to relative abundance of phyla, with orange circles representing control samples contrasted over blue circles which represent dump samples.

controls). In contrast, members of the Alpha- and Deltaproteobacteria were in greater abundance in control soil samples (Table 3). Alphaproteobacteria represented 16.2 ± 13.7% of reads in control soils compared to 11.9 ± 17.0% of reads in dump soils, Deltaproteobacteria were observed as 5.70 ± 6.92% in control soils and 3.33 ± 5.68% of total reads dump soils, respectively (Table 3).

Actinobacteria was the second most abundant phylum in both dump and control soils, but was lower in abundance in the dump-affected samples (11.1 ± 15.4% of reads) compared to control samples (17.7 ± 16.1% of reads) (Figure 7, Table 3). Specifically, the classes Actinobacteria and Thermoleophilia were in higher abundance in control soils (7.36 ± 6.6% and 5.93 ± 5.3%, respectively) compared to dump soils (4.89 ± 6.5% and 3.81 ± 5.3%, respectively) (Table 3). Actinobacteria are prevalent in arid and desert soils (Bachar et al., 2010; Crits-Christoph et al., 2013; Maza et al., 2019; Torres-Cortés et al., 2012) including the Chihuahuan Desert (J. S. Clark et al., 2009). Members of this phylum are capable of growing in low soil moisture conditions (Zvyagintsev et al., 2007), have multiple UV repair adaptations (Gao & Garcia-Pichel, 2011), and capacity for sporulation (Ensign, 1978). Acidobacteria, being the third most abundant phyla across control and dump samples, were also found at higher abundance in control (15.8 ± 14.1%) compared to dump-affected soils (7.35 ± 10.2%). A number of Subgroup classifications for Acidobacteria were detected in both dump and control soils, with the class Blastocatellia (Subgroup 4) having the highest relative abundance in control samples at 5.45 ± 5.2% compared to 2.00 ± 3.1% in dump samples (Table 3). Acidobacteria are among the most abundant and diverse phyla in soil ecosystems and all cultured members are heterotrophic,

although they are strongly underrepresented in culture collections (Elena P. Ivanova et al., 2014; Kielak et al., 2016).

**Table 3.** Relative abundance of phyla, and classes within phyla, across all control and dump soils as averages ± standard deviation (SD). Only phyla and classes with a relative abundance >0.1% are presented.

| **Phylum** | **Class** | **Average Relative Abundance (% ± SD)** | |
|---|---|---|---|
| | | **Dump** | **Control** |
| Acidobacteria | | 7.35 ± 10.11 | 15.8 ± 14.1 |
| | Acidobacteriia | 0.55 ± 1.13 | 1.44 ± 1.47 |
| | Blastocatellia_(Subgroup_4) | 2.00 ± 3.10 | 5.45 ± 5.15 |
| | Holophagae | 0.74 ± 1.10 | 1.08 ± 1.24 |
| | Subgroup_17 | 0.12 ± 0.15 | 0.13 ± 0.12 |
| | Subgroup_6 | 3.08 ± 3.16 | 5.65 ± 4.00 |
| | Thermoanaerobaculia | 0.67 ± 1.01 | 1.84 ± 1.71 |
| Actinobacteria | | 11.1 ± 15.4 | 17.7 ± 16.1 |
| | 0319-7L14 | 0.13 ± 0.24 | 0.24 ± 0.33 |
| | Actinobacteria | 4.89 ± 6.45 | 7.36 ± 6.60 |
| | Unclassified | 0.86 ± 1.23 | 1.26 ± 1.24 |
| | MB-A2-108 | 0.57 ± 1.16 | 0.67 ± 0.88 |
| | Nitriliruptoria | 0.06 ± 0.08 | 0.12 ± 0.14 |
| | Rubrobacteria | 0.76 ± 0.83 | 2.04 ± 1.17 |
| | Thermoleophilia | 3.81 ± 5.31 | 5.93 ± 5.71 |
| Armatimonadetes | | 0.22 ± 0.44 | 0.54 ± 0.70 |
| | uncultured | 0.19 ± 0.36 | 0.48 ± 0.60 |
| Bacteria_unclassified | | 0.93 ± 2.19 | 1.89 ± 3.03 |
| Bacteroidetes | | 4.90 ± 8.97 | 7.10 ± 8.33 |
| | Bacteroidia | 4.72 ± 8.50 | 6.95 ± 8.09 |
| | Rhodothermia | 0.11 ± 0.31 | 0.04 ± 0.06 |
| Chloroflexi | | 3.11 ± 5.11 | 5.97 ± 7.33 |
| | Anaerolineae | 0.64 ± 1.36 | 0.91 ± 1.51 |
| | Chloroflexi_unclassified | 0.08 ± 0.21 | 0.20 ± 0.36 |
| | Chloroflexia | 0.82 ± 1.56 | 1.80 ± 2.54 |
| | Dehalococcoidia | 0.10 ± 0.14 | 0.12 ± 0.11 |
| | Gitt-GS-136 | 0.22 ± 0.22 | 0.17 ± 0.20 |

| Phylum | Class | Average Relative Abundance (% ± SD) | |
|---|---|---|---|
| | | Dump | Control |
| | KD4-96 | 0.74 ± 0.73 | 1.42 ± 1.08 |
| | Ktedonobacteria | 0.09 ± 0.18 | 0.49 ± 0.73 |
| | TK10 | 0.32 ± 0.45 | 0.79 ± 0.71 |

***Table 3****. cont.*

| Phylum | Class | Average Relative Abundance (%) ± SD | |
|---|---|---|---|
| | | Dump | Control |
| Cyanobacteria | | 0.51 ± 1.17 | 1.15 ± 2.20 |
| | Melainabacteria | 0.10 ± 0.28 | 0.09 ± 0.17 |
| | Oxyphotobacteria | 0.35 ± 0.77 | 0.92 ± 1.79 |
| | Sericytochromatia | 0.06 ± 0.12 | 0.14 ± 0.22 |
| Elusimicrobia | | 0.07 ± 0.17 | 0.14 ± 0.25 |
| Entotheonellaeota | Entotheonellia | 0.10 ± 0.22 | 0.50 ± 0.53 |
| Epsilonbacteraeota | Campylobacteria | 1.42 ± 3.16 | 0.44 ± 1.71 |
| Euryarchaeota | | 0.20 ± 0.47 | 0.41 ± 0.50 |
| | Unclassified | 0.10 ± 0.21 | 0.26 ± 0.29 |
| | Thermoplasmata | 0.10 ± 0.26 | 0.15 ± 0.21 |
| FBP | | 0.22 ± 0.36 | 0.61 ± 0.60 |
| Firmicutes | | 6.83 ± 11.9 | 1.39 ± 2.97 |
| | Bacilli | 6.81 ± 11.8 | 1.38 ± 2.94 |
| Gemmatimonadetes | | 1.27 ± 2.17 | 2.03 ± 2.09 |
| | Gemmatimonadetes | 1.01 ± 1.69 | 1.83 ± 1.81 |
| | S0134_terrestrial_group | 0.12 ± 0.28 | 0.11 ± 0.15 |
| Halanaerobiaeota | Halanaerobiia | 0.91 ± 2.65 | 0.56 ± 2.15 |
| Nitrospirae | Nitrospira | 0.18 ± 0.30 | 0.43 ± 0.41 |
| Patescibacteria | | 1.09 ± 2.72 | 0.03 ± 0.10 |
| | Gracilibacteria | 1.07 ± 2.62 | 0.00 ± 0.01 |
| Planctomycetes | | 3.45 ± 5.87 | 5.52 ± 7.10 |
| | Phycisphaerae | 1.04 ± 1.66 | 1.48 ± 1.76 |
| | Planctomycetacia | 2.25 ± 3.86 | 3.88 ± 5.05 |
| Proteobacteria | | 50.6 ± 71.5 | 28.7 ± 29.8 |

| Phylum | Class | Average Relative Abundance (%) ± SD | |
|---|---|---|---|
| | | Dump | Control |
| | Alphaproteobacteria | 11.9 ± 17.0 | 16.2 ± 13.7 |
| | Deltaproteobacteria | 3.33 ± 6.13 | 5.68 ± 6.93 |
| | Gammaproteobacteria | 35.3 ± 48.2 | 6.65 ± 8.88 |
| | Proteobacteria_unclassified | 0.08 ± 0.23 | 0.15 ± 0.31 |
| Rokubacteria | NC10 | 0.26 ± 0.45 | 0.41 ± 0.41 |
| Thaumarchaeota | | 1.97 ± 2.48 | 3.48 ± 2.60 |
| | Nitrososphaeria | 1.97 ± 2.48 | 3.48 ± 2.60 |
| Verrucomicrobia | Verrucomicrobiae | 3.02 ± 5.30 | 4.85 ± 5.98 |

Members of the Firmicutes were more abundant in dump-affected soils (6.83 ± 11.9% of total reads) compared to control soils (1.39 ± 3.97%) (Figure 7) and have been found in high abundance in previous studies of OG produced waters (Struchtemeyer & Elshahed, 2012). Two additional phyla that had increased relative abundance in dump-affected soils compared to controls were Epsilonbacteraeota and Patescibacteria, each comprising less than 1.5% of total reads in dump-affected soils (Figure 7).

Bacteroidetes, Chloroflexi, Planctomycetes, Verrucomicrobiota, Thaumarchaeota, Gemmatimonadetes, and Cyanobacteria all had elevated relative abundances in control soils, and each group represented between 1.15 to 7.10% of total reads in control samples (Figure 7, Table 3). These phyla are among some of those found commonly in arid, desert soil ecosystems (Connon et al., 2007; Crits-Christoph et al., 2013; Kuske et al., 1997; Makhalanyane et al., 2015; Maza et al., 2019; Skujinš, 1984), although groups like Planctomycetes and Verrucomicrobiota are underrepresented in cultivation studies (Bergmann et al., 2011; Kuske et al., 1997). Members of Chloroflexi are known to be

thermophilic and, along with Acidobacteria and Actinobacteria, are able to oxidize atmospheric $H_2$, an important process under nutrient deprivation encountered in arid soil ecosystems (Islam et al., 2019). Thaumarchaeota was the sole Archaea group represented in high abundance among control soil samples (Table 8). When growing in arid soils, these Archaea may receive nutrients and photosynthetic capabilities from other unrelated groups, and Proteobacterium isolates with photosynthetic capabilities were shown to pass photosynthetic capacity to Gemmatimonadetes (Zeng et al., 2014). This may be an important process in supporting microbial populations in nutrient limited (i.e. in arid soil) ecosystems which may be disturbed due to OG wastewater dumping.

***Most Abundant OTUs Observed in Dump-Affected and Control Soils***

In dump-affected soils, the most abundant OTUs were members of Gammaproteobacteria, Firmicutes, Epsilonbacteraeota, Patescibacteria, Bacteroidetes, Halanaerobiaeota, and Acidobacteria, which represented 36% of the reads across all samples (Table 4A). The most abundant OTUs affiliated with the Gammaproteobacteria class were members of the families Marinobacteraceae, Halomonadaceae, Idiomarinaceae, Cellvibrionaceae, and Alcanivoracaceae (Table 4A). In the family Marinobacteraceae, two *Marinobacter* OTUs had a combined relative abundance of 15.6% in dump-affected soils (compared to 0.47% across all control samples; Table 4A). Marinobacter species have previously been isolated from hydrocarbon-polluted sediments (Gauthier et al., 1992) and have the capacity to use a variety of hydrocarbon as their sole source of carbon and energy

(Duran, 2010; Gauthier et al., 1992). *Halomonas* and *Chromohalobacter,* both members of the halophilic Halomonadaceae family, were also found in greater abundance in dump-affected soils (Table 3A). An OTU belonging to Idiomarinaceae, a heterotrophic and halophilic aerobe (Duran, 2010; Gauthier et al., 1992), was present at an average relative abundance of 4.24 ± 2.9% across all dump samples, while it comprised only 0.02 ± 0.08% of all control samples. *Marinimicrobium* (Cellvibrionaceae family), was observed at a greater relative abundance across all dump soils (1.81 ± 4.2 %) compared to control soils (0.01 ± 0.03 %).

**Table 4.** Relative abundance of 15 OTUs with highest abundance across all (A) dump and (B) control samples, represented as average (%) relative abundance ± standard deviation. Relative abundance of the OTUs in the opposite condition are provided in the right-most column for comparison.

A.

| OTU # | Phylum | Class | Order | Family | Genus | Dump | Control |
|---|---|---|---|---|---|---|---|
| Otu00001 | Proteobacteria | Gammaproteobacteria | Alteromonadales | Marinobacteraceae | *Marinobacter* | 10.2 ± 11.1 | 0.16 ± 0.59 |
| Otu00002 | Proteobacteria | Gammaproteobacteria | Alteromonadales | Marinobacteraceae | *Marinobacter* | 5.43 ± 6.98 | 0.30 ± 1.16 |
| Otu00003 | Proteobacteria | Gammaproteobacteria | Alteromonadales | Idiomarinaceae | *Idiomarina* | 4.24 ± 2.94 | 0.02 ± 0.079 |
| Otu00005 | Proteobacteria | Gammaproteobacteria | Oceanospirillales | Halomonadaceae | *Halomonas* | 2.75 ± 2.10 | 0.43 ± 1.66 |
| Otu00006 | Proteobacteria | Gammaproteobacteria | Oceanospirillales | Halomonadaceae | *Halomonas* | 2.67 ± 5.55 | <0.001 |
| Otu00004 | Firmicutes | Bacilli | Bacillales | Bacillaceae | *Halolactibacillus* | 2.04 ± 2.46 | 0.52 ± 2.01 |
| Otu00008 | Proteobacteria | Gammaproteobacteria | Cellvibrionales | Cellvibrionaceae | *Marinimicrobium* | 1.81 ± 4.24 | 0.01 ± 0.033 |
| Otu00009 | Epsilonbacteraeota | Campylobacteria | Campylobacterales | Arcobacteraceae | *Arcobacter* | 1.39 ± 3.09 | 0.43 ± 1.67 |
| Otu00022 | Patescibacteria | Gracilibacteria | JGI_0000069-P22 | JGI_0000069-P22_fa | JGI_0000069-P22_ge | 1.06 ± 2.59 | <0.001 |
| Otu00015 | Bacteroidetes | Bacteroidia | Flavobacteriales | Flavobacteriaceae | *Salegentibacter* | 0.93 ± 1.67 | <0.001 |
| Otu00014 | Halanaerobiaeota | Halanaerobiia | Halanaerobiales | Halobacteroidaceae | *Orenia* | 0.83 ± 2.52 | 0.55 ± 2.11 |
| Otu00011 | Proteobacteria | Gammaproteobacteria | Oceanospirillales | Halomonadaceae | *Chromohalobacter* | 0.82 ± 1.23 | <0.001 |
| Otu00025 | Proteobacteria | Gammaproteobacteria | Oceanospirillales | Alcanivoracaceae | *Alcanivorax* | 0.72 ± 1.40 | <0.001 |
| Otu00035 | Firmicutes | Bacilli | Bacillales | Bacillaceae | Unclassified *Bacillaceae* | 0.71 ± 1.07 | <0.001 |
| Otu00007 | Acidobacteria | Blastocatellia (Subgroup 4) | Pyrinomonadales | Pyrinomonadaceae | RB41 | 0.62 ± 0.69 | 1.89 ± 0.94 |

B.

| OTU # | Phylum | Class | Order | Family | Genus | Control | Dump |
|---|---|---|---|---|---|---|---|
| Otu00007 | Acidobacteria | Blastocatellia (Subgroup 4) | Pyrinomonadales | Pyrinomonadaceae | RB41 | 1.89 ± 0.94 | 0.62 ± 0.69 |
| Otu00010 | Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae | Ellin6055 | 1.01 ± 0.41 | 0.58 ± 0.54 |
| Otu00013 | Proteobacteria | Alphaproteobacteria | Rhizobiales | Beijerinckiaceae | *Microvirga* | 0.96 ± 0.51 | 0.53 ± 0.48 |
| Otu00032 | Actinobacteria | Rubrobacteria | Rubrobacterales | Rubrobacteriaceae | *Rubrobacter* | 0.81 ± 0.36 | 0.22 ± 0.22 |
| Otu00020 | Actinobacteria | Actinobacteria | Micromonosporales | Micromonosporaceae | Unclassified | 0.77 ± 0.55 | 0.41 ± 0.23 |
| Otu00019 | Proteobacteria | Alphaproteobacteria | Rhizobiales | Beijerinckiaceae | *Microvirga* | 0.73 ± 0.41 | 0.36 ± 0.34 |
| Otu00036 | Thaumarchaeota | Nitrososphaeria | Nitrososphaerales | Nitrososphaeraceae | Unclassified | 0.73 ± 0.47 | 0.21 ± 0.29 |
| Otu00023 | Proteobacteria | Alphaproteobacteria | Rhizobiales | Beijerinckiaceae | Unclassified | 0.62 ± 0.42 | 0.32 ± 0.30 |
| Otu00024 | Cyanobacteria | Oxyphotobacteria | Nostocales | Phormidiaceae | *Tychonema* (CCAP_1459-11B) | 0.60 ± 1.14 | 0.11 ± 0.16 |
| Otu00028 | Proteobacteria | Deltaproteobacteria | Myxococcales | Unclassified | Unclassified | 0.59 ± 0.24 | 0.33 ± 0.29 |
| Otu00012 | Actinobacteria | Unclassified | Unclassified | Unclassified | Unclassified | 0.59 ± 0.54 | 0.50 ± 0.65 |
| Otu00029 | Proteobacteria | Deltaproteobacteria | Unclassified | Unclassified | Unclassified | 0.58 ± 0.42 | 0.23 ± 0.23 |
| Otu00017 | Actinobacteria | Actinobacteria | Micrococcales | Micrococcaceae | *Pseudarthrobacter* | 0.58 ± 0.39 | 0.49 ± 0.54 |
| Otu00014 | Halanaerobiaeota | Halanaerobiia | Halanaerobiales | Halobacteroidaceae | *Orenia* | 0.55 ± 2.11 | 0.83 ± 2.52 |
| Otu00018 | Acidobacteria | Subgroup 6 | Unclassified | Unclassified | Unclassified | 0.54 ± 0.31 | 0.35 ± 0.29 |

*Marinimicrobium* species have been shown to be moderately halotolerant and have previously been isolated from tidal flat sediments (Lim et al., 2006) and soil (Song et al., 2019). *Alcanivorax,* a marine microorganism ubiquitously detected in hydrocarbon-contaminated marine environments (Naether et al., 2013), was also higher in abundance in dump soils with 0.72 ± 1.4 % relative abundance (with less than 0.001% abundance in control samples) (Table 4A).

Within the Flavobacteriaceae family, one *Salegentibacter*-related OTU was elevated in abundance in dump samples (0.93 ± 1.67%) (Table 4A). Members of *Salegentibacter* are known to be strictly aerobic, are slightly to moderately halophilic, and have previously been isolated from marine sand (Bowman, 2016; Siamphan & Kim, 2014). The lactic acid bacterium *Halolactobacillus* (Firmicutes) was also among the top OTUs found in dump samples with a higher relative abundance of 2.04 ± 2.46% in dump affected soils compared to 0.52 ± 2.01% in controls. Previous isolates of *Halolactobacillus* from marine samples in Japan were found to be slightly halophilic, highly halotolerant, and alkaliphilic (Bowman, 2016; Siamphan & Kim, 2014). *Orenia,* a member of the family Halobacteroidaceae, was observed to be more abundant in all dump soil samples than in control samples, with relative abundances of 0.83 ± 2.5 % in dump soils and 0.55 ± 2.11 % in control soils (Table 4A). The elevated relative abundance of *Orenia* in controls (Table 4A) was primarily due to a single control sample from Site K (K-S-05) which had a relative abundance of 0.55%. Halanaerobiaeota members are known to be obligatory anaerobic, moderate halophiles that require NaCl concentrations between 0.5 and 3.4 Molar (M) and are sulfidogenic (Oren, 2014; Rainey et al., 1995). Within the family Arcobacteraceae,

*Arcobacter* was also among the OTUs with highest abundance in dump soils (Table 4A), and has previously been detected in OG wastewaters (Evans et al., 2018).

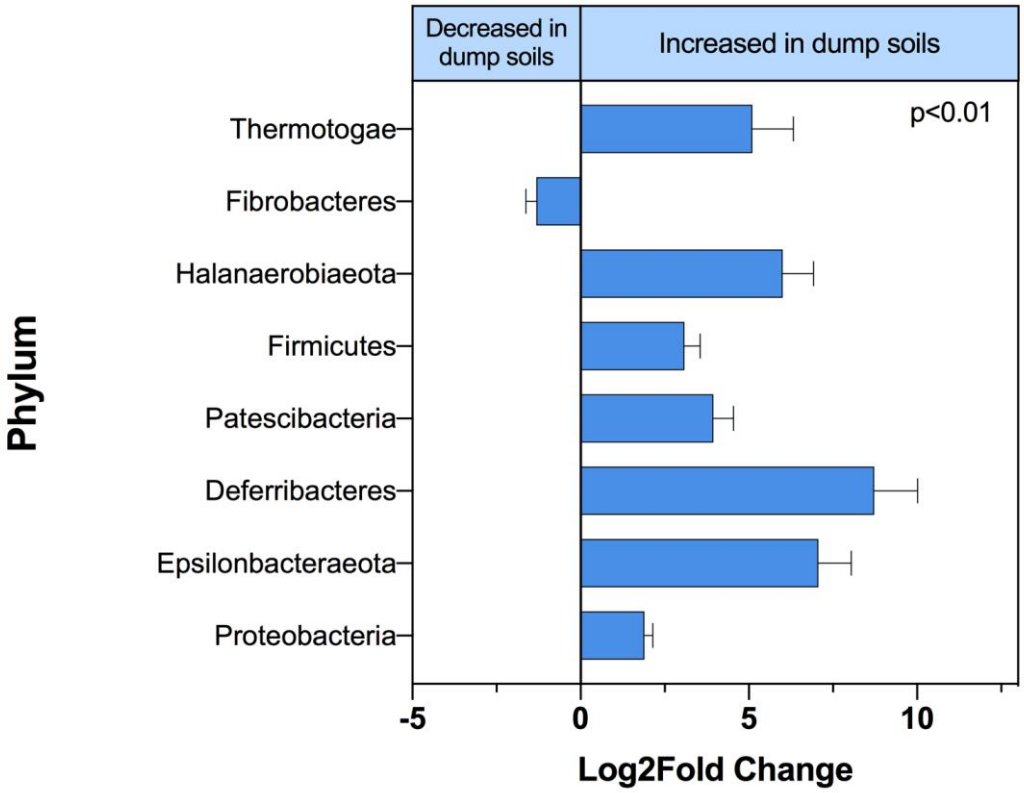
In control soils, the top OTUs were members of the Acidobacteria, Proteobacteria (Alpha- and Deltaproteobacteria classes), Actinobacteria, Cyanobacteria, and Thaumarchaeota phyla (Table 4B). The most abundant OTUs in control samples only comprised 11.5% of all control samples and were found at relatively low abundances (ranging from 0.54% to 1.89%). The most abundant OTU in control samples belonged to the family of aerobic chemoorganoheterotrophs Pyrinomonadaceae (Pascual et al., 2018) at 1.89 ± 0.94%, which was also among the most abundant OTUs in dump samples at 0.62 ± 0.69% abundance (Table 4B). An OTU related to the *Microvirga*, part of the Beijerinckiaceae family, had a relative abundance of 1.69% in control soils (compared to a combined abundance of 0.89% in dump soils) (Table 4B). Members of Beijerinckiaceae are free-living, nitrogen fixing aerobes, which enables them to survive in nitrogen-limited ecosystems (Marín & Arahal, 2014). An OTU belonging to the family Sphingomonadaceae, which are commonly isolated from soils (Glaeser & Kämpfer, 2014), was also among the most abundant microorganisms in control samples, with an average relative abundance of 1.01 ± 0.41% (Table 4B). The remaining top OTUs in control samples were found at abundances of less than 1% included those belonging to the families Rubrobacteriaceae, Micromonosporaceae, Nitrososphaeraceae, Phormidiaceae, Micrococcaceae, Halobacteroidaceae, an Acidobacteria subgroup (Subgroup 6), and unclassified members of Deltaproteobacteria, Actinobacteria, and Acidobacteria (Table 4B). Among them, Rubrobacteriaceae (Actinobacteria) have been detected in high

abundance in other arid and semi-arid soils (Bachar et al., 2010; Torres-Cortés et al., 2012). Additionally, Nitrososphaeraceae (Thaumarchaeota) are known to be important ammonia oxidizers and involved in nitrification (Tourna et al., 2011).

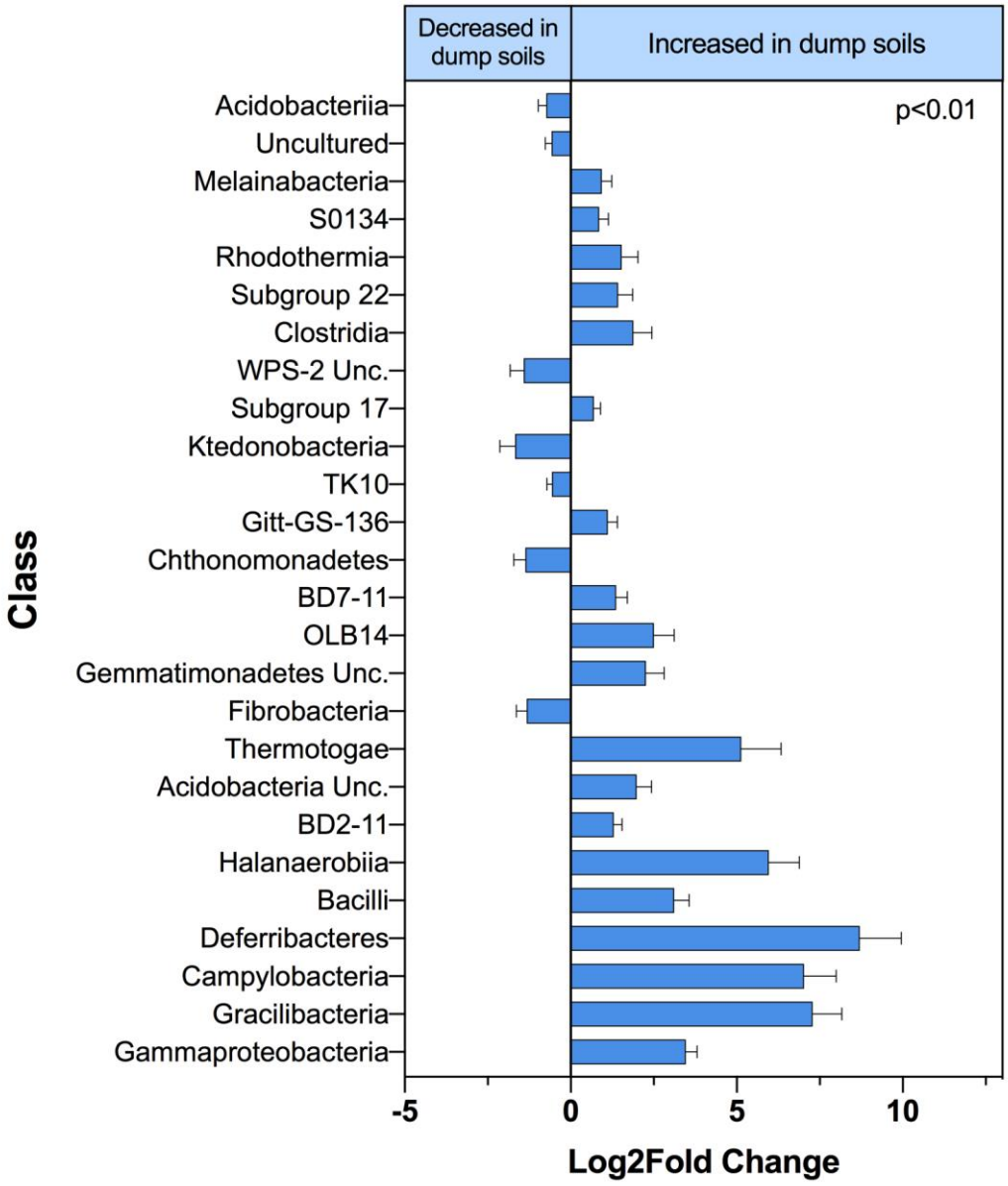
### *Differential Abundance of Microorganisms from Control to Dump Soil Conditions*

We used DESeq2 to determine which taxa were significantly ($p<0.01$) differentially abundant in OG wastewater dump soils. Differential abundance, as $\log_2$ fold change, was assessed at different taxonomic levels, comparing changes from the control to the OG wastewater dump condition and only taxa within the domain Bacteria had significant changes. The phyla Proteobacteria, Epsilonbacteraeota, Deferribacteres, Patescibacteria, Firmicutes, Halanaerobiaeota, and Thermotogae were all significantly ($p<0.01$) differentially abundant in the dump condition compared to the control condition (Figure 8A, Table 10), and all but Halanaerobiaeota and Deferribacteres were detected at relative abundances of >1%. We observed a $\log_2$ fold increase of ~2 in Proteobacteria abundance, and ~7 in Epsilonbacteraeota from control to dump soil conditions (Figure 8A). Patescibacteria and Firmicutes were observed to have 3.95 and 3.08 $\log_2$ fold increases in abundance, respectively (Table 10A). Members of the Fibrobacteres phylum were significantly less abundant in dump conditions (Figure 8A and Table 10A), and was detected at less than 1% relative abundance overall.

**A. Phylum Level**



**B. Class Level**



**C. Family Level**

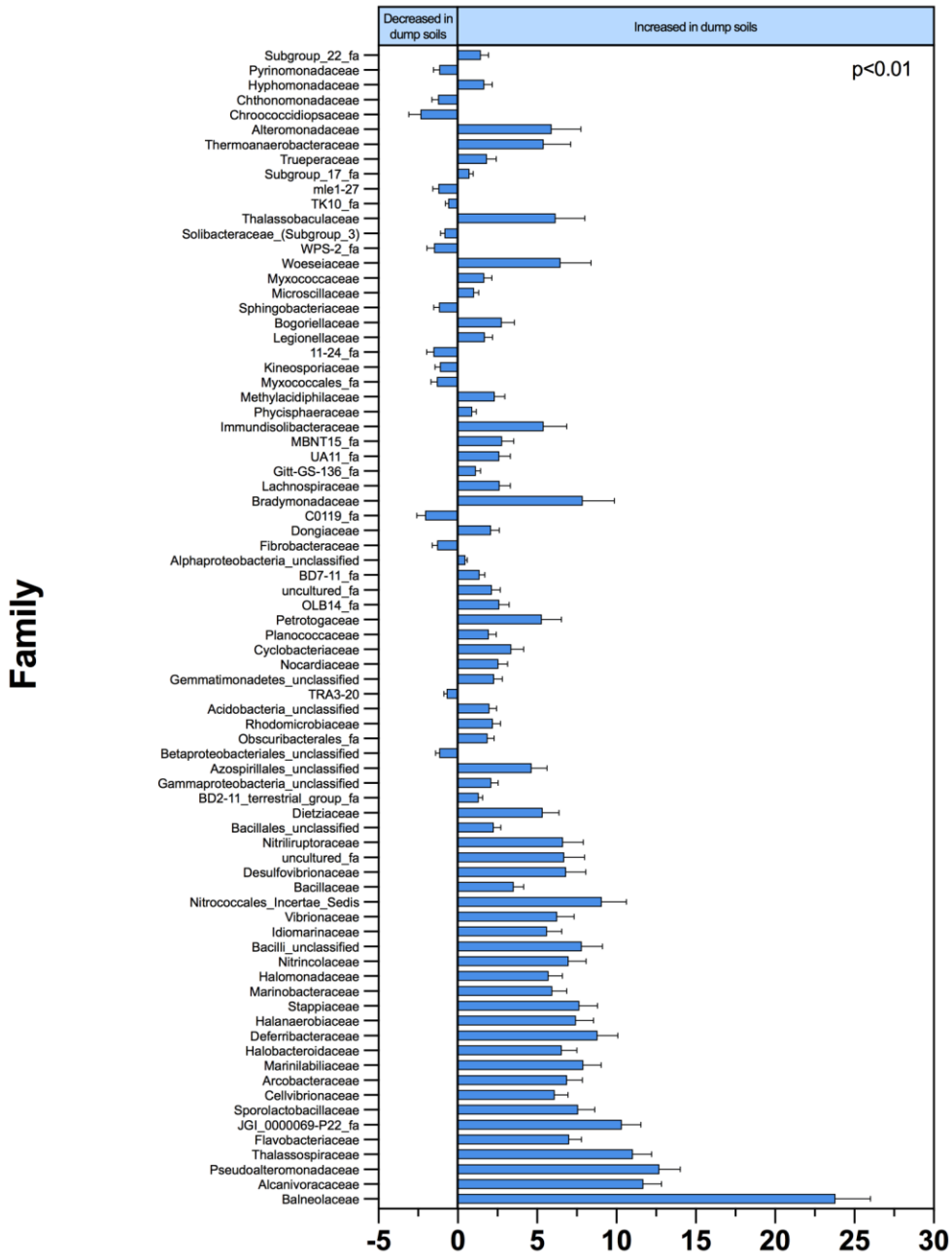



**Figure 8.** Differential abundance of microbial taxa shifting in composition from control to dump conditions at the (A) phylum, (B) class, and (C) family levels. Values are $\log_2$ fold change (doubling) ± standard error. Bars on the left indicate a significant decrease in abundance in dump soils while bars on the right indicate an increase in dump soils.

The families Marinobacteraceae, Idiomarinaceae, Arcobacteraceae, Cellvibrionaceae, Halobacteroidaceae, and Bacillaceae, which were members of the OTUs with the highest relative abundance in dump soil conditions (Figure 8C), had $\log_2$ fold changes of 3.5 to 6.9 from control to dump conditions (Figure 8C). In contrast, the family Pyrinomonadaceae, which was a member of the most abundant OTU in control soils (Figure 8C), significantly decreased in abundance ($\log_2$ fold change of -1.17) in dump soil conditions (Figure 8C).

Despite a relative abundance of less than 1% (Table 3), the phyla Halanaerobiaeota, Thermotogae, and Deferribacteres were all significantly differentially abundant in dump soil conditions with a $\log_2$ fold increase of >5. We observed a $\log_2$ fold increase of >8 in Deferribacteres abundance, >6 increase in Halanaerobiaeota, and >5 in Thermotogae (Figure 8A, Table 10A). All three of these phyla include organisms that are either halophilic and/or have thermophilic capabilities (Alauzet & Jumas-Bilak, 2014; Bhandari & Gupta, 2014; Oren, 2014). In addition to halophilic capacity, Deferribacteres are anaerobic and are typically found in thermophilic, sulfide-rich conditions (Alauzet & Jumas-Bilak, 2014).

At the family level, taxa with halotolerant and halophilic members, including Balneolaceae (Xia et al., 2017) , Alcanivoracaceae (Yakimov et al., 2019), Cellvibrionaceae (Lim et al., 2006), Halobacteroidaceae and Halanaerobiaceae (Oren, 2014; Rainey et al., 1995), and Idiomarinaceae (Ivanova et al., 2000) had a significant $\log_2$ fold increase from control to dump soil conditions (by more than 5-fold) (Figure 8B & C). Five families were found to have $\log_2$ fold changes of >10 in dump soil samples

($p<0.01$; Figure 8C, Table 4C). These families included Balneolaceae, Pseudoalteromonadaceae, Alcanivoracaceae, and Thalassospiraceae. Balneolaceae exhibited the greatest $\log_2$ fold increase from control to dump conditions ($p<0.01$) and has exhibited thermophilic and halophilic characteristics (Xia et al., 2017). The family Alcanivoracaceae include hydrocarbon degraders (Hara et al., 2003; Sabirova et al., 2011). The Thalassospiraceae family also had a significant $\log_2$ fold increase from control to dump conditions (Figure 8C, Table 10C), and have been previously found in petroleum-contaminated seawater and oil-polluted saline soils (Johnsen et al., 2019). Taxa that were significantly ($p<0.01$) decreased in abundance in dump soils with a $\log_2$ fold decrease of >2 were members of the family Chroococcidiopsaceae (Cyanobacteria), some of which are unicellular photosynthetic bacteria (Büdel, 2011), and the class Ktedonobacteria (Chloroflexi), which are found ubiquitously in terrestrial environments (Zheng et al., 2019).

# DISCUSSION

***Oil and Gas Wastewater Dumping Altered the Arid Soil Environment and Microbiome.***

This study revealed that OG wastewater dumping on arid soils in the Permian Basin led to shifts in soil microbial community structure and altered the geochemistry of impacted soils. Dumping resulted in a significant ($p<0.01$) increase in major anions and cations (Na, Cl, $SO_4^{2-}$, Br, Sr, Li, B, Ba, Ca, K, Mg, and Zn) in the soil that are commonly used as tracers of OG wastewater. OG wastewaters are characterized by high TDS (Engle et al., 2019) and specific geochemical tracers including Na, Ca, Cl, Li, B, Ba, Sr, and Br (Akob et al., 2016; Cozzarelli et al., 2017; Lauer et al., 2016; Murali Mohan et al., 2013; Struchtemeyer & Elshahed, 2012). These elements were elevated in the dump soils, which confirmed that the visual observations in the field were due to intentional releases of OG wastewaters. Dump-impacted samples also saw an increase in % moisture and % C, suggesting retention and persistence of the wastewaters and potential hydrocarbon residues four months after the dumps occurred, which has also been observed in previous releases of OG wastewaters (Cozzarelli et al., 2017).

The altered geochemistry observed in the New Mexico dump-affected soils reflects the chemistry of local OG wastewaters. In the Wolfcamp shale formation (part of the Permian Basin), formation waters were marked by elevated Na (between 38,101 to 45,096 mg/L) and Cl (between 63,052 to 75,370 mg/L) concentrations (Engle et al., 2016).

Elevated Br, B, $SO_4^{2-}$, and Ca concentrations in the New Mexico dump soils is consistent with the chemistry of Permian Basin formation waters which comprise a portion of OG wastewaters (Engle et al., 2016). In the dump soils, Fe, Al, and Si were found at lower concentrations compared to the unimpacted control soils, which differs from observations of previous releases. In Akob et al. (2016), Fe concentrations were elevated in sediments downstream from an OG wastewater disposal facility in West Virginia compared to upstream. Fe was similarly elevated in brine spills in North Dakota streams (Lauer et al., 2016) compared to background waters. Work is ongoing by USGS and university team members to understand what is driving changes in Al, Fe, and Si in the New Mexico dump soils.

The complex soil chemistry generated by OG wastewater dumps was accompanied by shifts in soil microbial community structure. Microbial communities in soils impacted by OG wastewater dumps had significantly lower alpha diversity (observed number of OTUs, Shannon diversity index, and Inverse Simpsons diversity index) compared to the unimpacted soils. They also differed in overall structure, with more heterogeneity among dump-affected samples compared to control soil communities (Figure 6A). The control soil microbial communities clustered more closely together, indicating a higher level of similarity between those samples. This may indicate a shift in dump soils towards more niche microbial communities that are better adapted for and reflect the complex chemistries that accompany OG wastewaters. This also supports the 'Anna Karenina principle' for microbiomes in which disturbed or stressed communities display more heterogeneity than communities under an ecosystem's 'normal' conditions (Zaneveld et al., 2017).

In addition to having lower diversity, we saw shifts in the taxonomic composition and potential function of microorganisms in OG wastewater dump-affected soils compared to unaffected soils. Arid desert soils typically harbor a number of ubiquitous taxa such as Actinobacteria, Acidobacteria, Bacteroidetes, and Proteobacteria (Connon et al., 2007; Crits-Christoph et al., 2013; Makhalanyane et al., 2015; Skujinš, 1984). Actinobacteria and Acidobacteria often prevail in phylogenetic studies of arid and semi-arid soil microbial communities (Crits-Christoph et al., 2013) and were some of the most abundant OTUs in control soil communities. Members of the *Rubrobacter* genus (Actinobacteria), also among the most abundant OTUs in control soils, have been observed in high abundance within arid and semi-arid soils globally (Bachar et al., 2010; Crits-Christoph et al., 2013; Torres-Cortés et al., 2012). While arid soil microbial communities are adapted to low levels of soil moisture for extended periods of time, they have been shown to be vulnerable to changes in precipitation frequency and timing in the Chihuahuan Desert (J. S. Clark et al., 2009). The microbial community shifts that we observed may be attributed to the increased soil moisture resulting from OG wastewater dumping. Surprisingly, this increased soil moisture was still measurable 4 months after dumping occurred. Another factor that may have driven changes in the community was the input of ions which increased the salinity of soils. Soil salinization, either through anthropogenic or natural causes, has been shown to decrease alpha diversity and impact community structure of arid soil microorganisms (Zhang et al., 2019) as we observed in this study.

### *Potential Microbial Functions Shifted as a Result of OG Wastewater Dumping*

Changes in microbial populations in dump soils may also indicate changes in ecological processes in arid soils. With negligible rainfall, much of the soil in arid and semi-arid ecosystems lacks decaying organic matter and less abundant microbial populations (Wiesman, 2009) but the dumps resulted in elevated carbon in soils. Microbial metabolisms in these ecosystems are typically adapted to low carbon and nitrogen availability, high UV radiation, and high temperatures (Leung et al., 2020; Makhalanyane et al., 2015). Dump soil communities were enriched in taxa with functions not typical for arid soil ecosystems (e.g. halophiles and hydrocarbon degraders). While this study was not designed to assess shifts in ecological processes, we suggest this may be an important avenue to explore further. The changes observed for dump-affected microbial communities suggest that microbial function and the ecosystem changed due to differences in nutrient availability. We predict that OG wastewater dumping may alter biogeochemical processes as halophilic and hydrocarbon-degrading taxa rise in abundance in arid soils.

An increase in microbial populations that possessed halophilic capabilities were observed in OG wastewater dump-affected soils. The rise in salt-tolerant microorganisms in dump soils suggests that Na and Cl from OG wastewater dumps were major selective pressures in determining community composition. OG wastewater dump soils had a high abundance of halotolerant and halophilic populations affiliated with Gammaproteobacteria, Epsilonproteobacteria, Bacteroidia, Bacilli, Campylobacteria, and Gracilibacteria (Figure 8C). The most abundant genera in dump soils (Table 3),

*Marinobacter, Idiomarina, Halomonas, Halolactobacillus, Marinimicrobium, Arcobacter, Salegentibacter, Orenia, Chromohalobacter, and Alcanivorax* all possess strains known for their halophilic or halotolerant capacities. *Marinobacter, Halomonas,* and members of Halanaerobiia have also previously been detected in OG wastewaters (Akob et al., 2015; Lipus et al., 2017; Murali Mohan, Hartsock, Hammack, et al., 2013). *Marinobacter* can grow in NaCl concentrations of 0.5 to 20.0% (Duran, 2010), *Idiomarina* can grow between 0.6 to 15.0% (Ivanova et al., 2000), and members of *Chromohalobacter* can grow in a range between 3.0 to 25.0% NaCl (Aguilera et al., 2007; Prado et al., 2006). The moderately halophilic obligate anaerobe *Orenia* (Halanaerobiia class) is capable of growing in 2.9 to 20.0% NaCl (Oren, 2014).

In addition to being enriched in salt-tolerating microbial populations, dump soil communities may have been influenced by elevated carbon and sulfate concentrations. Elevated % carbon in dump-affected soils (Table 1) supports the field observations that hydrocarbons were released onto the landscape with OG wastewater dumping. While characterization of the soil hydrocarbons is ongoing by the USGS team, we infer that the elevated carbon is from the residual product and other carbon sources used to enhance OG production, e.g., complex polysaccharides (like guar gum) (King, 2012). This may help explain the rise of microorganisms capable of hydrocarbon degradation and those found in previous studies of OG wastewaters and hydrocarbon-contaminated environments. Several of these microorganisms belonged to Gammaproteobacteria, including *Marinobacter*, which are known hydrocarbon degraders (88), and *Alcanivorax*, which have been found ubiquitously in hydrocarbon-contaminated environments (Hara et al., 2003; Schneiker et

al., 2006; Yakimov et al., 2019). Members of *Halomonas* are also capable of metabolizing a variety of organic, aromatic compounds (García et al., 2004) and were detected in significantly higher (>1%) abundance in dump soils compared to control soils (Table 3). Halanaerobiia, which were also detected in significantly higher abundance in OG dump soils, have been detected ubiquitously in OG wastewaters and are known hydrocarbon degraders (Oren, 2014). Sulfate may have also influenced the microbial community composition, as organisms with sulfur-cycling metabolisms like Thermotogae (Bhandari & Gupta, 2014) and the previously mentioned Halanaerobiia emerged in greater abundance in dump-affected soils contrasted to control soils.

Microorganisms that are well adapted to the unique chemistries of OG wastewaters, such as halophiles and those than can metabolize hydrocarbons, may have utility as natural attenuators of OG wastewater releases into the environment. These populations may also have utility in engineered remediation strategies. Microbial communities consisting of *Marinobacter*, *Halomonas*, *Idiomarinaceae*, and *Arcobacter* have previously been utilized in engineered microbial mats to remediate constituents of OG wastewaters (Akyon et al., 2015). Akyon et al. (2015) found that microorganisms in the engineered mats were able to metabolize complex carbohydrates and other constituents of produced waters from southwest Pennsylvania shales with TDS concentrations of up to 100,000 mg/L. Given that these microbial community members were also detected in high abundance in dump-affected soils, they may be useful in managed remediation strategies for OG wastewater releases in the Permian Basin.

In addition to having bioremediation potential, the microbial taxa detected in the OG wastewater dump soils may serve as indicators of OG wastewater releases into the environment. Halophilic and hydrocarbon-degrading taxa that typically do not comprise significant portions of arid soil ecosystems (Makhalanyane et al., 2015) may signify changes in the environment reflective of OG wastewater releases. Specific community members, like members of Halanaerobiia, may also serve as indicator species for OG wastewater releases as they have been observed in high abundance in produced waters of the Marcellus Shale (Akob et al., 2016; Cluff et al., 2014; Lipus et al., 2017; Murali Mohan et al., 2013) and Utica Shale (Booker et al., 2017). As these organisms were observed to be significantly more abundant in dump affected soils it provides an indication of the effect of OG wastewater release. However, the source of Halanaerobiales in dump-affected soils is unknown; they could be rare indigenous soil taxa that increased in abundance due to changing geochemistry or released into the environment with dumped OG wastewaters. To assess the source, additional studies are needed. Our work suggests that Halanaerobiales may serve as a microbial "fingerprint" for future OG wastewater releases due to its near ubiquitous detection across OG wastewaters and produced waters across U. S.

# CONCLUSIONS

Our data demonstrated that OG wastewater dumping in the Permian Basin did in fact alter the geochemistry of arid soils and led to a change in soil microbiology, supporting our hypothesis. This included a significant decrease in the microbial community's alpha diversity, and a distinct community structure compared to the unimpacted soil communities. The brine composition of the OG wastewater dumps as well as elevated % C and $SO_4^{2-}$ may have allowed for enrichment of certain microbial populations, namely those with halophilic, sulfidogenic, and/or hydrocarbon metabolizing abilities.

The Permian Basin is currently experiencing a boom in OG production (U.S. Energy Information Administration, 2018). The management and proper disposal of large volumes of OG wastewaters – typically brines with varying compositions of ions, hydrocarbons, and other constituents, as we saw in this study – is a growing challenge. The long-term environmental health impacts of illegal OG wastewater dumps still remains understudied. By characterizing environmental changes due to OG wastewater dumps in the Permian Basin, we gain further insights into how OG wastewaters alter arid environments and soil microbial functions. Thus far, geochemical indicators of OG wastewaters – such as Na, Ca, Cl, B, Br, Ba, and Sr – have been useful in identifying the presence of OG wastewaters in the environment. Microbial community members that were in detected at elevated abundance in the OG wastewater dump-affected soils could also serve as evidence of such releases. They may also serve in remediating contamination by

OG wastewater dump constituents. Our results highlight the importance of understanding the role microorganisms play in OG wastewater dumps.

# APPENDIX

## Supplemental Methods

### *Sequence processing script*

```
# Processed sequences using mothur v.1.42.3 and v.1.43.0

# to assemble contigs, alignment, taxonomic assignment, and OTU calling
make.contigs(file=BLM2_file.txt, processors=64)
summary.seqs()

#screen sequences, parameters here are for v4 region
screen.seqs(fasta=current, group=current, maxambig=1, optimize=start-end-minlength-
maxlength, criteria=90, maxhomop=8)
summary.seqs()
unique.seqs(fasta=current)
count.seqs(name=current, group=current)
summary.seqs(count=current)

# Aligned sequences to Silva v132, using customized Silva v132 truncated to the v4 region
using pcr.seqs to save time and memory.
align.seqs(fasta=current,
reference=/home/amumford/Working_Silva_Files/silva.nr_v132.align, flip=T)
summary.seqs(count=current)

#screened sequences after alignment, parameters here are for v4, and #summarized
sequences
screen.seqs(fasta=current, count=current, optimize=start-end-minlength-maxlength,
criteria=95)
summary.seqs(count=current)

# Pre-processing sequences prior to OUT clustering
filter.seqs(fasta=current, vertical=T, trump=.)
summary.seqs(count=current)
unique.seqs(fasta=current, count=current)

#pre-cluster sequences, parameters for v4
pre.cluster(fasta=current, count=current, diffs=3)
summary.seqs(count=current)

# Checked sequences for chimeras
```

chimera.vsearch(fasta=current, count=current, dereplicate=t)

# Removed chimeric sequences from fasta file
remove.seqs(fasta=current, accnos=current)
summary.seqs(count=current)

# Classified sequences against Silva v132
classify.seqs(fasta=current, count= current,
reference=/home/amumford/Working_Silva_Files/silva.nr_v132.align,
taxonomy=/home/amumford/Working_Silva_Files/silva.nr_v132.tax, cutoff=80, probs=F)

# Removed unknown and non-bacterial/archaeal sequences
remove.lineage(fasta=current, count=current, taxonomy=current,taxon=Chloroplast-Mitochondria-unknown-Eukaryota)
summary.seqs(count=current)

# Called and classified OTUs based on 97% sequence similarity
cluster.split(fasta=
BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.fasta,
count=BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.pick.count_table,
taxonomy=BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.pick.nr_v132.wang.pick.taxonomy, splitmethod=classify, taxlevel=5, cutoff=0.03)
make.shared(list=current, count=current, label=0.03)
classify.otu(list=current, count=current, taxonomy=current, label=0.03)

#basic pre-analysis of diversity
count.groups(shared=current)
collect.single(shared=current, calc=chao-invsimpson, freq=100)
rarefaction.single(shared=current, calc=sobs, freq=100)
summary.single(shared=current, calc=nseqs-coverage-sobs-invsimpson, subsample=T)

# remove singletons
mothur>
remove.rare(list=BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.opti_mcc.list,
count=BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.pick.count_table, nseqs=1)

# generated a FASTA file of representative sequences and mothur shared file for downstream analysis

mothur >
list.seqs(list=BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.opti_mcc.0.03.pick.list)

```
mothur>
get.seqs(fasta=BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.fast
a,
accnos=BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.opti_mcc.
0.03.pick.accnos)

mothur>
get.oturep(list=BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.opti
_mcc.0.03.pick.list,
column=BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.pick.dist,
count=BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.
pick.pick.count_table, fasta=
BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.pick.fasta)

mothur>
make.shared(list=current,
count=BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.
pick.pick.count_table)

# changed file names for brevity in mothur

mothur > system(cp
BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.opti_mcc.0.03.pick
.0.03.rep.fasta BLM2_repotus_NoSingle.fasta)

mothur > system(cp
BLM2_file.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.opti_mcc.0.03.pick
.shared BLM2_NoSingle.shared)
```

***Cleaning up FASTA names to format for FastTree***

# used ‘sed’ text editor to remove header information from FASTA file to make FastTree compatible file

```
sed 's/\M*\b.*\t//' BLM2_repotus_NoSingle.fasta |sed 's/|.*//' > BLM_allclean3.fasta
```

# uploaded cleaned up version of FASTA file into Geneious Prime version 2019.2.1 (https://www.geneious.com) to create a maximum likelihood tree using the FastTree function

***R analysis scripts***

```
#libraries required for importing mothur files, processing OTU tables,
#performing statistical analyses, and making figures
library(DESeq2)
library(phyloseq)
library(ggplot2)
library(vegan)
packageVersion("phyloseq") #‘1.26.1’
packageVersion("ggplot2") #‘3.2.1’
packageVersion("vegan") #‘2.5.6’

# Tree file name: BLM2_repotus_NoSingle_FastTree.newick
# Shared file name: BLM2_NoSingle.shared
# Taxonomy file name: BLM2_0.03.cons.taxonomy
# Set working directory >setwd("~/Documents/BLM/phyloseq/NMBLM_20191021")

# Imported mothur shared and taxonomy, and FastTree's newick formatted phylogenetic tree file
# Note: singletons removed in mothur
NMBLM_NS_raw <- import_mothur(mothur_shared_file ="BLM2_NoSingle.shared",
                              mothur_constaxonomy_file ="BLM2_0.03.cons.taxonomy",
                              mothur_tree_file = "BLM2_repotus_NoSingle_FastTree.newick")
# Loaded in NMBLM metadata file
NMBLM_NS_metadata<- data.frame(read.csv("NMBLM_Metadata_20200416.csv", header = TRUE, row.names = 1))

# Formatted metadata for phyloseq
NMBLM_NS_metadata_phyloseq <-
sample_data(data.frame(Site=NMBLM_NS_metadata["Site"],
                       Indv=NMBLM_NS_metadata["Indv"],
                       Condition=NMBLM_NS_metadata["Condition"],
Type=NMBLM_NS_metadata["Type"],
                       Depth_cm=NMBLM_NS_metadata["Depth_cm"],
All_Params=NMBLM_NS_metadata["All_Params"],
                       Caliche=NMBLM_NS_metadata["Caliche"],
Site_K=NMBLM_NS_metadata["Site_K"]))
```

```
# Merged metadata and mothur sequence data
NMBLM_NS_full_phyloseq <- merge_phyloseq(NMBLM_NS_raw,
NMBLM_NS_metadata_phyloseq)

# Changing column names in tax table from ranks to taxon
colnames(tax_table(NMBLM_NS_full_phyloseq))[1] <-"Domain"
colnames(tax_table(NMBLM_NS_full_phyloseq))[2] <-"Phylum"
colnames(tax_table(NMBLM_NS_full_phyloseq))[3] <-"Class"
colnames(tax_table(NMBLM_NS_full_phyloseq))[4] <-"Order"
colnames(tax_table(NMBLM_NS_full_phyloseq))[5] <-"Family"
colnames(tax_table(NMBLM_NS_full_phyloseq))[6] <-"Genus"

# Checked new tax table column names
rank_names(NMBLM_NS_full_phyloseq)

# Reviewed otu and tax table summary
tax_table(NMBLM_NS_full_phyloseq)
otu_table(NMBLM_NS_full_phyloseq) # 43,203 taxa and 46 samples. Taxa are in rows.
sum(taxa_sums(NMBLM_NS_full_phyloseq)) #6,170,399
min(sample_sums(NMBLM_NS_full_phyloseq)) #107,495
max(sample_sums(NMBLM_NS_full_phyloseq)) #185,009
# subset samples of interest
# removed Site C (caliche pit)
NMBLM_NS_withoutC<-subset_samples(NMBLM_NS_full_phyloseq, Caliche=="no")
sample_data(NMBLM_NS_withoutC)

# subset surface soils (0-5cm) and surface cores (0-10 cm)
NMBLM_NS_surface<-subset_samples(NMBLM_NS_withoutC, Type=="surface_grab")
sample_data(NMBLM_NS_surface)

# summary
otu_table(NMBLM_NS_surface) #43,203 taxa by 33 samples
sum(taxa_sums(NMBLM_NS_surface)) #4,477,573
min(sample_sums(NMBLM_NS_surface)) #109,370
max(sample_sums(NMBLM_NS_surface)) #185,009

# estimated richness prior to rarefying
NMBLM_NS_surface_NOTRARE_rich<-estimate_richness(NMBLM_NS_surface)
write.csv(NMBLM_NS_surface_NOTRARE_rich,
"NMBLM_NS_surface_NOTRARE_rich.csv")

# Observed richness summary
min(NMBLM_NS_surface_NOTRARE_rich$Observed) #3,300
max(NMBLM_NS_surface_NOTRARE_rich$Observed) #9,026
mean(NMBLM_NS_surface_NOTRARE_rich$Observed) #6,240.061
sd(NMBLM_NS_surface_NOTRARE_rich$Observed) #1,580.519

# rarefied to smallest library size
# set.seed to 711 for reproducibility
```

```
NMBLM_NS_rare_surface <- rarefy_even_depth(NMBLM_NS_surface,
sample.size=min(sample_sums(NMBLM_NS_surface)),
                                rngseed=711, replace=FALSE, trimOTUs=TRUE)
# 4,987 OTUs were removed because they are no longer present in any sample after
random subsampling
sample_sums(NMBLM_NS_rare_surface) # 109,370 seqs.

# summary
tax_table(NMBLM_NS_rare_surface)  #38216
otu_table(NMBLM_NS_rare_surface) #38,216 taxa by 33 samples
sum(taxa_sums(NMBLM_NS_rare_surface)) #3,609,210 -- total # of individuals observed
from each species/taxa/OTU
min(sample_sums(NMBLM_NS_rare_surface)) #109,370 (total # individuals observed from
each sample, min. observed)
max(sample_sums(NMBLM_NS_rare_surface)) #109,370

# estimate_richness
# diversity metrics were then plotted and statistical tests run in Prism
NMBLM_NS_rare_surface_rich<-estimate_richness(NMBLM_NS_rare_surface)
write.csv(NMBLM_NS_rare_surface_rich, "NMBLM_NS_rare_surface_rich.csv")

# creating unifrac distance matrix using newick formatted phylogenetic tree
# randomly rooted UniFrac tree in phyloseq
# Used code described by:http://john-quensen.com/r/unifrac-and-tree-roots/
phy_tree(NMBLM_NS_rare_surface) # phylogenetic tree with 38,216 tips and 38,214 internal
nodes. Unrooted; includes branch lengths.
set.seed(711)
wunifracdist.NMBLM_NS_rare_surface<-phyloseq::distance(NMBLM_NS_rare_surface,
method="unifrac", weighted=T)
#Randomly assigned root as -- Otu07710-- in the phylogenetic tree in the data you provided.

# NMDS orindation with weighted UniFrac distance matrix of rarefied dataset
NMDSord.wunifrac_NMBLM<-ordinate(NMBLM_NS_rare_surface, method = "NMDS",
distance = wunifracdist.NMBLM_NS_rare_surface)
NMDSord.wunifrac_NMBLM

# plot ordination with ggplot2
theme_set(theme_bw())
NMDSord.wunifrac_NMBLM.plot3<-plot_ordination(NMBLM_NS_rare_surface,
NMDSord.wunifrac_NMBLM, color = "Condition", shape="Condition", label="Site") +
  geom_point(aes(colour=Condition, shape=Condition, fill=Condition), size = 2) +
labs(x="NMDS1", y="NMDS2", cex=4) +
  theme(axis.text = element_text(size=12), axis.title = element_text(size=14, face="bold"))+
  stat_ellipse(geom = "polygon",type = "norm", level=0.95, alpha = 0.25, aes(fill =
Condition))+
  scale_colour_manual(values = c("#F79D0B","#185892"),
                      name = "Condition")+
  scale_fill_manual(values = c("#ed8a47","#69B4F7"))
NMDSord.wunifrac_NMBLM.plot3
```

```
# permanova on weighted unifrac distance matrix
set.seed(711)
wunifrac_adonis <- adonis(wunifracdist.NMBLM_NS_rare_surface ~
sample_data(NMBLM_NS_rare_surface)$Condition)
wunifrac_adonis # between all dump and all control, p=0.001 ***, R2 = 0.36162

# loaded in geochemistry metadata (all analytes included except for
# F-, PO4_3-, and Cu which had concentrations below detection limits in    # both control and
dump samples)

geochem <- data.frame(read.csv("geochem_for_vegan_minusFPo4Cu.csv", header = TRUE,
row.names = 1))

#ordinate UniFrac distance matrix in vegan using metaMDS
NMBLM_wunifrac_metaMDSord<-metaMDS(wunifracdist.NMBLM_NS_rare_surface, k=2,
try=1000, trymax = 10000)

#used envfit function in vegan to fit geochemistry vectors
set.seed(711)
NMBLM_envfit_NMDS<-envfit(NMBLM_wunifrac_metaMDSord, geochem_v2, perm=999,
choices = c(1,2))
View(NMBLM_envfit_NMDS)
#Site - p=0.495 , R2 = 0.1131
#Condition - p<0.001 ** , R2 = 0.5093
#No. permutations = 999

### retrieved scores from oridination
NMBLM_wunifrac_scores<-scores(NMBLM_wunifrac_metaMDSord)
NMBLM_wunifrac_scores

#Merged community data with geochem factors
microbe_geochem<-merge(NMBLM_wunifrac_scores, geochem_v2, by="row.names")

#Ensured sample names were correctly formatted as row names
row.names(microbe_geochem)<-microbe_geochem$Row.names
View(microbe_geochem)

microbe_geochem$Condition <- factor(microbe_geochem$Condition, levels =
c("Control","Dump"), labels = c("Control","Dump"))


#formatted vectors
NMBLM_as_list<-as.list(NMBLM_envfit_NMDS$vectors)
NMBLM_pvals<-as.data.frame(NMBLM_as_list$pvals)
NMBLM_arrows<-as.data.frame(NMBLM_as_list$arrows*sqrt(NMBLM_as_list$r))
NMBLM_cbind<-cbind(NMBLM_arrows, NMBLM_pvals)
NMBLM_cbind_sig<-subset(NMBLM_cbind, NMBLM_as_list$pvals<0.05)
NMBLM_cbind_sig<-cbind(NMBLM_cbind_sig, Factors = rownames(NMBLM_cbind_sig))
```

```
#plotted NMDS with vectors fitted

NMBLM_micro_with_geochemvectors.plot<-ggplot(microbe_geochem, aes(x=NMDS1,
y=NMDS2)) +
  theme_bw()+
  geom_point(aes(colour=Condition, shape=Condition, fill=Condition), size = 2) +
labs(x="NMDS1", y="NMDS2", cex=4) +
  theme(axis.text = element_text(size=12), axis.title = element_text(size=14, face="bold"),
        plot.title = element_text(size=16, hjust = 0.5)) +
  scale_color_manual(values = c("#ed8a47","#0072B2"),
                     name = "Condition") +
   stat_ellipse(geom = "polygon",type = "norm", level=0.95, alpha = 0.25, aes(fill =
Condition))+
  scale_fill_manual(values = c("#ed8a47","#69ACF7"))+
  geom_segment(data=NMBLM_cbind_sig,
               aes(x=0, xend = NMDS1, y=0, yend = NMDS2),
               arrow = arrow(length = unit(0.25, "cm")), colour = "grey", alpha=0.5) +
  geom_text(data=NMBLM_cbind_sig, aes(x = NMDS1, y = NMDS2, label = Factors),
            size = 4, position=position_jitter(width=.02, height=.02))
NMBLM_micro_with_geochemvectors.plot

# exported raw count data and taxa table of rarefied dataset for relative abundance analysis
NMBLM_NS_countdata<-as.data.frame(otu_table(NMBLM_NS_rare_surface))
write.csv(NMBLM_NS_countdata, "NMBLM_NS_countdata.csv")
NMBLM_NS_taxatable<-as.data.frame(tax_table(NMBLM_NS_rare_surface))
write.csv(NMBLM_NS_taxatable, "NMBLM_NS_taxatable.csv")

# observed relative abundances at Domain level to determine percentage of Archae and
Bacteria across all samples
Domain<-tax_glom(NMBLM_NS_rare_surface,
taxrank=rank_names(NMBLM_NS_rare_surface)[1])
Domain_NS_RA<-transform_sample_counts(Domain, function(OTU)(OTU/sum(OTU))*100)
Domain_otutable<-as.data.frame(otu_table(Domain_NS_RA))
write.csv(Domain_otutable, "Domain_OTU_Tabale.csv")

#Summary - Domain level
sum(taxa_sums(Domain_NS_RA)) #Total of 3300 taxa
taxa_sums(Domain_NS_RA)
#Archaea (Otu00036) = 97.91
#Bacteria (Otu00001) = 3,202.09

## ran DESeq2 to determine differential abundance of taxa
# binned OTUs to Phylum level from dataset that was not rarefied -->
sample_data(NMBLM_NS_surface)

#Phylum level

#Collapse OTUs into phylum classifications
NMBLM_NS_surface_phylum<-tax_glom(NMBLM_NS_surface, taxrank = "Phylum")
```

```
sample_data(NMBLM_NS_surface_phylum)$Condition <-
as.factor(sample_data(NMBLM_NS_surface_phylum)$Condition)
#Made DESeq objects
DESeq.phylum_allsites <- phyloseq_to_deseq2(NMBLM_NS_surface_phylum, ~Condition)
DESeq.phylum_allsites<-DESeq(DESeq.phylum_allsites)
#Assigned siginificance for determining differerntial abundance
alpha=0.01
results_phylum<- results(DESeq.phylum_allsites, contrast = c("Condition", "Dump",
"Control"), alpha=alpha)
results_phylum<- results_phylum[order(results_phylum$padj, na.last = NA), ]
results_phylum_sig<-results_phylum[results_phylum$padj<alpha, ]
#Added taxa names from phyloseq object "tax_table"
results_phylum_sig.data<-cbind(as(results_phylum_sig, "data.frame"),

as(tax_table(NMBLM_NS_surface_phylum)[rownames(results_phylum_sig), ],
                              "matrix"))

#Export significant phyla data as .csv to be plotted in Prism
write.csv(results_phylum_sig.data, "NMBLM_NS_surface_phylum_allsites_v2.csv")

#Repeated for Class level with same steps as Phylum level

NMBLM_NS_surface_Class<-tax_glom(NMBLM_NS_surface, taxrank = "Class")
sample_data(NMBLM_NS_surface_Class)$Condition <-
as.factor(sample_data(NMBLM_NS_surface_Class)$Condition)
DESeq.class_allsites <- phyloseq_to_deseq2(NMBLM_NS_surface_Class, ~Condition)
DESeq.class_allsites<-DESeq(DESeq.class_allsites)
alpha=0.01
results_Class<- results(DESeq.class_allsites, contrast = c("Condition", "Dump", "Control"),
alpha=alpha)
results_Class<- results_Class[order(results_Class$padj, na.last = NA), ]
results_Class_sig<-results_Class[results_Class$padj<alpha, ]
results_Class_sig.data = cbind (as(results_Class_sig, "data.frame"),
                            as(tax_table(NMBLM_NS_surface_Class)[rownames(results_Class_sig),
],
                               "matrix"))
write.csv(results_Class_sig.data, "NMBLM_NS_surface_Class_allsites_v2.csv")

#Repeated for Family level with same steps as Phylum and Class level

NMBLM_NS_surface_Family<-tax_glom(NMBLM_NS_surface, taxrank = "Family")
sample_data(NMBLM_NS_surface_Family)$Condition <-
as.factor(sample_data(NMBLM_NS_surface_Family)$Condition)
DESeq.family_allsites <- phyloseq_to_deseq2(NMBLM_NS_surface_Family, ~Condition)
DESeq.family_allsites<-DESeq(DESeq.family_allsites)
alpha=0.01
results_Family<- results(DESeq.family_allsites, contrast = c("Condition", "Dump", "Control"),
alpha=alpha)
results_Family<- results_Family[order(results_Family$padj, na.last = NA), ]
results_Family_sig<-results_Family[results_Family$padj<alpha, ]
```

```
results_Family_sig.data = cbind (as(results_Family_sig, "data.frame"),

as(tax_table(NMBLM_NS_surface_Family)[rownames(results_Family_sig), ],
                                    "matrix"))

write.csv(results_Family_sig.data, "NMBLM_NS_surface_Family_allsites.csv")
```

## Supplemental Tables

**Table 5.** Geochemistry of (A) control and (B) dump-affected soil samples. Sample names labeled "Dup" are duplicates. Columns with "<" indicate values below the detection limit. Variations in detection limits within an analyte are due to varying dilution factors as a result of potentially elevated concentrations in dump samples. Conductivity, pH, anions, and cations were measured on soil:water extracts.

### A. Control Geochemistry

| Sample Name | Site | pH | Specific Conductance (mS/cm) | F (mg/kg) | Cl (mg/kg) | Br (mg/kg) | $NO_3^-$ (mg/kg) | $PO_4^{3-}$ (mg/kg) | $SO_4^{2-}$ (mg/kg) |
|---|---|---|---|---|---|---|---|---|---|
| D-C-03, 0-10 cm | D | 7.80 | 0.083 | 0.284 | 0.517 | <0.064 | 0.769 | <0.67 | 1.00 |
| D-S-01 | D | 7.57 | 0.134 | <0.495 | 0.670 | <0.160 | 1.42 | <1.68 | 1.81 |
| D-S-01-Dup | D | 6.85 | 0.008 | <0.495 | 0.817 | <0.160 | 2.31 | <1.68 | 2.06 |
| D-S-02 | D | 7.30 | 0.010 | <0.495 | 0.861 | <0.160 | 2.36 | <1.68 | 2.58 |
| F-S-01 | F | 7.92 | 0.047 | 0.243 | 0.764 | <0.064 | 1.20 | 0.672 | 1.66 |
| F-S-02 | F | 7.61 | 1.53 | <0.495 | 1.18 | <0.160 | 1.91 | <1.68 | 2.56 |
| F-S-03 | F | 7.68 | 0.730 | <0.495 | 0.733 | <0.160 | 2.37 | <1.68 | 2.32 |
| G-S-07 | G | 7.81 | 0.328 | <0.495 | 1.52 | <0.160 | 3.55 | <1.68 | 2.90 |
| G-S-08 | G | 7.82 | 0.460 | <0.495 | 2.30 | <0.160 | 3.07 | <1.68 | 3.21 |
| H-S-01 | H | 7.35 | 0.391 | <0.495 | 2.57 | <0.160 | 2.48 | <1.68 | 3.68 |
| H-S-02 | H | 7.26 | 0.187 | <0.495 | 0.930 | <0.160 | 1.52 | <1.68 | 3.19 |
| H-S-03 | H | 7.14 | 0.147 | <0.495 | 0.824 | <0.160 | 1.13 | <1.68 | 3.37 |
| K-S-01 | K | 7.93 | 0.398 | <0.495 | 0.618 | <0.160 | 3.82 | <1.68 | 2.17 |
| K-S-03 | K | 7.69 | 0.768 | <0.495 | 2.998 | <0.160 | 7.09 | <1.68 | 3.57 |
| K-S-05 | K | 7.41 | 0.390 | <0.495 | 2.576 | <0.160 | 9.59 | 2.90 | 4.56 |
| K-S-05-Dup | K | 7.44 | 0.809 | <0.495 | 1.676 | <0.160 | 6.29 | <1.68 | 3.31 |
| K-S-04 | K | 7.88 | 1.96 | 1.13 | 16.2 | <0.160 | 0.93 | <1.68 | 4.15 |
| K-S-04-Dup | K | 7.85 | 0.887 | 1.22 | 13.3 | <0.160 | 0.99 | <1.68 | 3.17 |

***Table 5A.*** *cont.*

| Sample Name | Site | Ca (mg/kg) | Na (mg/kg) | Mg (mg/kg) | K (mg/kg) | Si (mg/kg) | Sr (mg/kg) | Al (mg/kg) | Fe (mg/kg) |
|---|---|---|---|---|---|---|---|---|---|
| D-C-03, 0-10 cm | D | 19.4 | 1.09 | 5.20 | 8.93 | 51.7 | 0.081 | 24.58 | 13.66 |
| D-S-01 | D | 11.6 | 0.26 | 3.22 | 6.27 | 33.3 | 0.046 | 14.74 | 8.29 |
| D-S-01-Dup | D | 12.3 | 0.28 | 4.31 | 7.66 | 56.8 | 0.048 | 26.08 | 14.28 |
| D-S-02 | D | 14.8 | 0.28 | 3.10 | 6.01 | 39.3 | 0.055 | 17.30 | 9.93 |
| F-S-01 | F | 14.7 | 1.10 | 3.68 | 5.72 | 30.6 | 0.065 | 12.10 | 6.84 |
| F-S-02 | F | 11.2 | 0.62 | 2.12 | 4.66 | 18.8 | 0.053 | 7.87 | 4.65 |
| F-S-03 | F | 16.6 | 0.28 | 4.04 | 6.51 | 37.7 | 0.070 | 15.70 | 8.73 |
| G-S-07 | G | 17.0 | 0.36 | 3.48 | 6.16 | 29.2 | 0.078 | 12.04 | 6.91 |
| G-S-08 | G | 17.4 | 0.69 | 3.86 | 7.28 | 34.2 | 0.088 | 13.70 | 7.92 |
| H-S-01 | H | 10.4 | 0.13 | 3.11 | 7.12 | 28.4 | 0.043 | 12.61 | 7.08 |
| H-S-02 | H | 11.1 | 0.08 | 3.70 | 6.47 | 32.2 | 0.053 | 15.63 | 8.05 |
| H-S-03 | H | 11.0 | 0.31 | 5.61 | 9.39 | 68.2 | 0.058 | 32.03 | 18.49 |
| K-S-01 | K | 17.7 | 0.11 | 5.53 | 8.34 | 57.1 | 0.086 | 24.76 | 14.02 |
| K-S-03 | K | 12.8 | 0.58 | 4.95 | 10.81 | 55.7 | 0.069 | 25.00 | 13.99 |
| K-S-05 | K | 13.2 | 0.16 | 3.66 | 11.19 | 27.0 | 0.073 | 12.39 | 6.83 |
| K-S-05-Dup | K | 8.27 | 0.17 | 2.42 | 7.85 | 21.8 | 0.045 | 9.09 | 5.17 |
| K-S-04 | K | 20.3 | 6.59 | 4.64 | 5.93 | 22.6 | 0.184 | 7.08 | 4.21 |
| K-S-04-Dup | K | 19.1 | 5.54 | 5.52 | 6.48 | 33.9 | 0.169 | 11.88 | 6.96 |

***Table 5A.*** *cont.*

| Sample Name | Site | Mn (mg/kg) | B (mg/kg) | Ba (mg/kg) | Li (mg/kg) | Cu (mg/kg) | Zn (mg/kg) | % Moisture | % C | % N |
|---|---|---|---|---|---|---|---|---|---|---|
| D-C-03, 0-10 cm | D | 0.355 | 0.063 | 0.255 | <0.02 | <0.02 | 0.055 | 0.419 | 0.171 | 0.017 |
| D-S-01 | D | 0.206 | 0.021 | 0.114 | <0.02 | <0.02 | 0.039 | 0.425 | 0.159 | 0.015 |
| D-S-01-Dup | D | 0.234 | 0.020 | 0.117 | 0.020 | <0.02 | 0.063 | 0.294 | 0.284 | 0.031 |
| D-S-02 | D | 0.122 | <0.02 | 0.073 | <0.02 | <0.02 | 0.035 | 0.196 | 0.267 | 0.019 |
| F-S-01 | F | 0.143 | <0.02 | 0.082 | <0.02 | <0.02 | 0.038 | 0.174 | 0.271 | 0.018 |
| F-S-02 | F | 0.099 | <0.02 | 0.049 | <0.02 | <0.02 | 0.021 | 0.256 | 0.178 | 0.016 |
| F-S-03 | F | 0.201 | <0.02 | 0.096 | <0.02 | <0.02 | 0.048 | 0.159 | 0.351 | 0.029 |
| G-S-07 | G | 0.199 | <0.02 | 0.089 | <0.02 | <0.02 | 0.041 | 0.158 | 1.246 | 0.044 |
| G-S-08 | G | 0.174 | <0.02 | 0.082 | <0.02 | <0.02 | 0.041 | 0.124 | 0.201 | 0.022 |
| H-S-01 | H | 0.302 | <0.02 | 0.124 | <0.02 | <0.02 | 0.052 | 0.183 | 0.331 | 0.034 |
| H-S-02 | H | 0.305 | <0.02 | 0.171 | <0.02 | <0.02 | 0.040 | 0.164 | 0.176 | 0.020 |
| H-S-03 | H | 0.357 | <0.02 | 0.174 | 0.025 | <0.02 | 0.052 | 0.165 | 0.187 | 0.021 |
| K-S-01 | K | 0.379 | <0.02 | 0.129 | <0.02 | <0.02 | 0.068 | 0.118 | 0.371 | 0.040 |
| K-S-03 | K | 0.326 | <0.02 | 0.120 | 0.020 | <0.02 | 0.064 | 0.165 | 0.259 | 0.033 |
| K-S-05 | K | 0.298 | <0.02 | 0.108 | <0.02 | <0.02 | 0.032 | 0.261 | 0.372 | 0.035 |
| K-S-05-Dup | K | 0.115 | <0.02 | 0.051 | <0.02 | <0.02 | 0.020 | 0.149 | 0.227 | 0.029 |
| K-S-04 | K | 0.078 | <0.02 | 0.077 | <0.02 | <0.02 | 0.021 | 0.146 | 0.370 | 0.032 |
| K-S-04-Dup | K | 0.116 | <0.02 | 0.084 | <0.02 | <0.02 | 0.028 | 0.188 | 0.307 | 0.025 |

**B. Dump Soil Chemistry**

| Sample Name | Site | pH | Specific Conductance (mS/cm) | F mg/kg | Cl mg/kg | Br mg/kg | $NO_3^-$ mg/kg | $PO_4^{3-}$ mg/kg | $SO_4^{2-}$ mg/kg |
|---|---|---|---|---|---|---|---|---|---|
| D-C-08, 0-10 cm | D | 7.54 | 5.72 | <24.8 | 3006 | 28.1 | <11.3 | <83.8 | 103 |
| D-S-06 | D | 7.51 | 4.61 | <24.8 | 3432 | 37.6 | <11.3 | <83.8 | 99.5 |
| D-S-06-Dup | D | 7.66 | 1.74 | <24.8 | 5071 | 48.6 | <11.3 | <83.8 | 143 |
| D-S-07 | D | 7.57 | 4.47 | <24.8 | 15033 | 109 | <11.3 | <83.8 | 606 |
| D-S-11 | D | 8.37 | 0.517 | 0.546 | 1355 | 11.8 | <0.90 | <1.675 | 31.8 |
| D-S-12 | D | 7.61 | 3.08 | <24.8 | 10900 | 103 | <11.3 | <83.8 | 279 |
| D-S-13 | D | 7.70 | 3.41 | <24.8 | 13138 | 109 | <11.3 | <83.8 | 330 |
| F-S-04 | F | 7.79 | 19.8 | <24.8 | 8889 | 69.6 | <11.3 | <83.8 | 271 |
| F-S-05 | F | 7.74 | 20.2 | <24.8 | 9314 | 96.3 | <11.3 | <83.8 | 234 |
| F-S-06 | F | 7.66 | 23.2 | <24.8 | 10900 | 99.4 | <11.3 | <83.8 | 354 |
| G-C-03, 0-10 cm | G | 7.80 | 13.0 | <24.8 | 5402 | 52.4 | <11.3 | <83.8 | 170 |
| G-S-01 | G | 7.64 | 23.9 | <24.8 | 10853 | 90.7 | <11.3 | <83.8 | 285 |
| G-S-02 | G | 7.73 | 12.7 | <24.8 | 5587 | 64.9 | 557 | <83.8 | 204 |
| G-S-06 | G | 7.44 | 30.8 | <24.8 | 15006 | 127 | <11.3 | <83.8 | 422 |
| H-S-04 | H | 7.84 | 5.59 | <24.8 | 1914 | 18.9 | <11.3 | <83.8 | 72.2 |
| H-S-05 | H | 7.56 | 11.6 | <24.8 | 4830 | 45.7 | <11.3 | <83.8 | 162 |
| H-S-06 | H | 7.86 | 6.74 | <24.8 | 2748 | 20.9 | <11.3 | <83.8 | 87.6 |
| K-S-02 | K | 7.59 | 16.4 | <24.8 | 6638 | 51.5 | <11.3 | <83.8 | 273 |
| K-S-06 | K | 7.88 | 8.53 | <24.8 | 3108 | 28.2 | <11.3 | <83.8 | 125 |

***Table 5B.*** *cont.*

| Sample Name | Site | Ca mg/kg | Na mg/kg | Mg mg/kg | K mg/kg | Si mg/kg | Sr mg/kg | Al mg/kg | Fe mg/kg |
|---|---|---|---|---|---|---|---|---|---|
| D-C-08, 0-10 cm | D | 141 | 1642 | 21.4 | 53.9 | 2.59 | 5.44 | 0.048 | 0.041 |
| D-S-06 | D | 329 | 1684 | 37.4 | 53.3 | 3.47 | 5.60 | <0.02 | <0.02 |
| D-S-06-Dup | D | 471 | 2426 | 56.5 | 70.7 | 3.60 | 8.80 | <0.02 | <0.02 |
| D-S-07 | D | 1000 | 7612 | 135 | 168 | 2.05 | 40.4 | <0.02 | <0.02 |
| D-S-11 | D | 43.6 | 747.9 | 5.50 | 30.9 | 2.89 | 1.15 | 0.073 | 0.047 |
| D-S-12 | D | 911 | 5087 | 126 | 131 | 4.67 | 21.9 | <0.02 | <0.02 |
| D-S-13 | D | 738 | 6519 | 106 | 187 | 4.46 | 25.1 | <0.02 | <0.02 |
| F-S-04 | F | 590 | 4581 | 83.0 | 115 | 4.41 | 16.2 | <0.02 | <0.02 |
| F-S-05 | F | 655 | 4488 | 94.5 | 132 | 5.19 | 15.8 | <0.02 | <0.02 |
| F-S-06 | F | 810 | 5197 | 113 | 151 | 4.14 | 21.8 | <0.02 | <0.02 |
| G-C-03, 0-10 cm | G | 335 | 2777 | 47.6 | 78.7 | 5.46 | 8.47 | <0.02 | 1.10 |
| G-S-01 | G | 675 | 5218 | 96.3 | 143 | 4.20 | 20.1 | 0.034 | <0.02 |
| G-S-02 | G | 298 | 2520 | 41.1 | 75.7 | 2.85 | 8.61 | <0.02 | <0.02 |
| G-S-06 | G | 1002 | 6986 | 143 | 172 | 3.51 | 30.8 | <0.02 | <0.02 |
| H-S-04 | H | 47.1 | 1038 | 6.11 | 39.6 | 3.14 | 1.10 | <0.02 | <0.02 |
| H-S-05 | H | 309 | 2333 | 40.6 | 72.5 | 2.42 | 7.82 | <0.02 | <0.02 |
| H-S-06 | H | 95.4 | 1440 | 13.8 | 61.3 | 3.25 | 2.23 | <0.02 | <0.02 |
| K-S-02 | K | 533 | 3100 | 77.4 | 81.2 | 2.73 | 13.6 | <0.02 | <0.02 |
| K-S-06 | K | 210 | 1556 | 30.9 | 48.7 | 4.47 | 5.43 | <0.02 | <0.02 |

***Table 5B.*** *cont.*

| Sample Name | Site | Mn mg/kg | B mg/kg | Ba mg/kg | Li mg/kg | Cu mg/kg | Zn mg/kg | % Moisture | % C | % N |
|---|---|---|---|---|---|---|---|---|---|---|
| D-C-08, 0-10 cm | D | 0.23 | 4.24 | 0.09 | 0.33 | <0.02 | 0.0978 | 2.59 | 0.589 | 0.0233 |
| D-S-06 | D | 0.11 | 1.17 | 0.85 | 0.18 | <0.02 | 0.0993 | 1.41 | 0.863 | 0.0430 |
| D-S-06-Dup | D | 0.25 | 1.24 | 1.16 | 0.31 | <0.02 | 0.122 | 1.45 | 1.19 | 0.0000 |
| D-S-07 | D | 2.01 | 5.25 | 0.59 | 1.37 | <0.02 | 0.180 | 0.99 | 0.247 | 0.0000 |
| D-S-11 | D | <0.02 | 1.16 | 0.17 | 0.16 | <0.02 | <0.02 | 3.59 | 2.29 | 0.0000 |
| D-S-12 | D | 0.42 | 2.97 | 1.15 | 0.98 | <0.02 | 0.162 | 4.42 | 2.90 | 0.0386 |
| D-S-13 | D | 1.07 | 5.99 | 1.18 | 1.29 | <0.02 | 0.137 | 3.11 | 4.66 | 0.0615 |
| F-S-04 | F | 0.52 | 4.88 | 0.91 | 0.57 | <0.02 | 0.124 | 0.29 | 1.01 | 0.0273 |
| F-S-05 | F | <0.02 | 4.63 | 1.28 | 0.79 | <0.02 | 0.164 | 4.61 | 0.288 | 0.0254 |
| F-S-06 | F | 0.54 | 5.44 | 1.25 | 0.89 | <0.02 | 0.183 | 1.36 | 1.26 | 0.0438 |
| G-C-03, 0-10 cm | G | 0.08 | 2.85 | 0.96 | 0.54 | <0.02 | 0.0995 | 2.78 | 1.98 | 0.0000 |
| G-S-01 | G | 0.44 | 4.08 | 0.84 | 0.99 | <0.02 | 0.122 | 0.68 | 1.86 | 0.0379 |
| G-S-02 | G | 0.21 | 3.09 | 0.72 | 0.36 | <0.02 | 0.0899 | 0.36 | 0.821 | 0.0271 |
| G-S-06 | G | 1.15 | 5.19 | 0.79 | 1.17 | <0.02 | 0.178 | 0.28 | 1.05 | 0.0362 |
| H-S-04 | H | <0.02 | 2.00 | 0.06 | 0.11 | <0.02 | 0.0215 | 2.34 | 2.24 | 0.0451 |
| H-S-05 | H | 0.45 | 2.53 | 0.19 | 0.23 | 0.02 | 0.0951 | 0.64 | 0.361 | 0.0290 |
| H-S-06 | H | 0.02 | 2.54 | 0.04 | 0.20 | 0.02 | 0.0376 | 5.78 | 5.05 | 0.0514 |
| K-S-02 | K | 0.47 | 3.49 | 0.63 | 0.31 | 0.05 | 0.0730 | 0.46 | 0.414 | 0.0281 |
| K-S-06 | K | 0.10 | 2.36 | 0.33 | 0.15 | <0.02 | 0.0203 | 0.10 | 0.494 | 0.0290 |

**Table 6.** Summary of non-parametric Mann-Whitney test statistics on soil geochemistry between all control and all dump soil samples. All P values are two-tailed, and are exact (as opposed to approximate).

| **Mann-Whitney U Results** | **pH** | **Specific Conductance (mS/cm)** | **F** | **Cl** | **Br** | **$NO_3^-$** | **$PO_4^{2-}$** | **$SO_4^{2-}$** | **Ca** | **Na** |
|---|---|---|---|---|---|---|---|---|---|---|
| P value | 0.4565 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| Significantly different (P < 0.05)? | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Sum of ranks in Controls, Dump-Affected | 317, 386 | 178, 525 | 173, 530 | 171, 532 | 171, 532 | 205, 498 | 179.5, 523.5 | 171, 532 | 171, 532 | 171, 532 |
| Mann-Whitney U | 146 | 7 | 2 | 0 | 0 | 34 | 8.5 | 0 | 0 | 0 |
| Median of Controls | 7.645, n=18 | 0.3906, n=18 | 0.4950, n=18 | 1.055, n=18 | 0.1600, n=18 | 2.335, n=18 | 1.675, n=18 | 3.035, n=18 | 13.93, n=18 | 0.2950, n=18 |
| Median of Dump-Affected | 7.660, n=19 | 8.533, n=19 | 24.80, n=19 | 5587, n=19 | 52.38, n=19 | 11.30, n=19 | 83.80, n=19 | 204.2, n=19 | 470.7, n=19 | 2777, n=19 |
| Difference between Medians: Actual | 0.015 | 8.142 | 24.31 | 5586 | 52.22 | 8.965 | 82.13 | 201.1 | 456.8 | 2776 |
| Difference between Medians: Hodges-Lehmann | 0.0725 | 8.142 | 24.31 | 5586 | 52.22 | 8.94 | 82.13 | 201.1 | 456.8 | 2776 |

***Table 6.*** *cont.*

| **Mann-Whitney U Results** | **Mg** | **K** | **Si** | **Sr** | **Al** | **Fe** | **Mn** | **B** | **Ba** | **Li** |
|---|---|---|---|---|---|---|---|---|---|---|
| P value | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | 0.4379 | <0.0001 | <0.0001 | <0.0001 |
| Significantly different (P < 0.05)? | Yes | Yes | Yes | Yes | Yes | Yes | No | Yes | Yes | Yes |
| Sum of ranks in Controls, Dump-Affected | 174, 529 | 171, 532 | 513, 190 | 171, 532 | 513, 190 | 513, 190 | 316, 387 | 171, 532 | 218.5, 484.5 | 171, 532 |
| Mann-Whitney U | 3 | 0 | 0 | 0 | 0 | 0 | 145 | 0 | 47.5 | 0 |
| Median of Controls | 3.780, n=18 | 6.815, n=18 | 33.60, n=18 | 0.06500, n=18 | 14.22, n=18 | 7.985, n=18 | 0.2050, n=18 | 0.02000, n=18 | 0.1050, n=18 | 0.02000, n=18 |
| Median of Dump-Affected | 56.45, n=19 | 78.66, n=19 | 3.510, n=19 | 8.800, n=19 | 0.02000, n=19 | 0.02000, n=19 | 0.2500, n=19 | 3.090, n=19 | 0.7900, n=19 | 0.3600, n=19 |
| Difference between Medians: Actual | 52.67 | 71.85 | -30.09 | 8.735 | -14.2 | -7.965 | 0.045 | 3.07 | 0.685 | 0.34 |
| Difference between Medians: Hodges-Lehmann | 52.67 | 71.21 | -29.74 | 8.735 | -14.18 | -7.9 | 0.09 | 3.07 | 0.67 | 0.34 |

***Table 6.*** *cont.*

| **Mann-Whitney U Results** | **Cu** | **Zn** | **% Moisture** | **% Carbon** | **% Nitrogen** |
|---|---|---|---|---|---|
| P value | >0.9999 | 0.0004 | <0.0001 | <0.0001 | 0.2299 |
| Significantly different (P < 0.05)? | No | Yes | Yes | Yes | No |
| Sum of ranks in Controls, Dump-Affected | 333, 370 | 231, 472 | 196.5, 506.5 | 201, 502 | 303.5, 399.5 |
| Mann-Whitney U | 162 | 60 | 25.5 | 30 | 132.5 |
| Median of Controls | 0.02000, n=18 | 0.04000, n=18 | 0.1650, n=18 | 0.2700, n=18 | 0.03000, n=18 |
| Median of Dump-Affected | 0.02000, n=19 | 0.1000, n=19 | 1.410, n=19 | 1.050, n=19 | 0.03000, n=19 |
| Difference between Medians: Actual | 0 | 0.06 | 1.245 | 0.78 | 0 |
| Difference between Medians: Hodges-Lehmann | 0 | 0.07 | 1.215 | 0.8 | 0.01 |

**Table 7.** Summary of significance testing for alpha diversity metrics determined using non-parametric Mann-Whitney test comparing all dump sample alpha diversity to all control sample alpha diversity. All P-values are exact, and two-tailed.

| **Mann Whitney Test** | **Observed OTUs** | **Shannon Diversity** | **Inverse Simpsons Diversity** |
|---|---|---|---|
| **P value** | <0.0001 | <0.0001 | <0.0001 |
| **Significantly different (P < 0.05)?** | Yes | Yes | Yes |
| **Sum of ranks in Control, Dump** | 362, 199 | 370, 191 | 372, 189 |
| **Mann-Whitney U** | 28 | 20 | 18 |
| **Median of Control** | 7098, n=15 | 7.131, n=15 | 337.0, n=15 |
| **Median of Dump** | 4627, n=18 | 5.625, n=18 | 38.84, n=18 |
| **Difference Between Medians: Actual** | -2471 | -1.506 | -298.2 |
| **Difference Between Medians: Hodges-Lehmann** | -2143 | -1.43 | -279.9 |

**Table 8**. Scores from non-metric multidimensional scaling (NMDS) ordination of microbial community structure, i.e. Figure 6A.

| Sample Name | Condition | NMDS1 | NMDS2 |
|---|---|---|---|
| D_C_03 | Control | -0.088181 | -0.018265 |
| D_C_08 | Dump | 0.176664 | -0.025490 |
| D_S_01 | Control | -0.090862 | -0.023579 |
| D_S_02 | Control | -0.102170 | -0.000354 |
| D_S_06 | Dump | 0.001148 | -0.006449 |
| D_S_07 | Dump | 0.054015 | -0.034724 |
| D_S_11 | Dump | 0.173614 | 0.031057 |
| D_S_12 | Dump | 0.003134 | 0.074253 |
| D_S_13 | Dump | 0.152073 | 0.047151 |
| F_S_01 | Control | -0.094700 | 0.051497 |
| F_S_02 | Control | -0.078522 | -0.069715 |
| F_S_03 | Control | -0.088986 | -0.015293 |
| F_S_04 | Dump | 0.152095 | -0.077642 |
| F_S_05 | Dump | 0.151096 | 0.093154 |
| F_S_06 | Dump | 0.084967 | 0.004902 |
| G_C_03 | Dump | 0.054903 | 0.101171 |
| G_S_01 | Dump | 0.016691 | 0.023436 |
| G_S_02 | Dump | 0.030137 | 0.053746 |
| G_S_06 | Dump | -0.006439 | 0.028591 |
| G_S_07 | Control | -0.096141 | 0.021243 |
| G_S_08 | Control | -0.112724 | 0.069961 |
| H_S_01 | Control | -0.123719 | -0.010706 |
| H_S_02 | Control | -0.106358 | -0.004973 |
| H_S_03 | Control | -0.123397 | -0.045632 |
| H_S_04 | Dump | 0.152388 | -0.024756 |
| H_S_05 | Dump | -0.021171 | -0.009994 |
| H_S_06 | Dump | 0.201256 | -0.024004 |
| K_S_01 | Control | -0.092206 | 0.012353 |
| K_S_02 | Dump | 0.071551 | -0.132814 |
| K_S_03 | Control | -0.097304 | -0.019990 |
| K_S_04 | Control | -0.093786 | 0.040833 |
| K_S_05 | Control | 0.027911 | -0.072552 |
| K_S_06 | Dump | -0.086976 | -0.036417 |

**Table 9.** ENVFIT scores for geochemical parameters for NMDS of soil microbial communities dump-affected and control soils, i.e. Figure 6B.

| | **NMDS1** | **NMDS2** | **$r^2$** | **Pr (>r)** | **Significant?** |
|---|---|---|---|---|---|
| pH | 0.450 | 0.893 | 0.199 | 0.038 | Yes |
| Specific Conductance | 0.946 | 0.323 | 0.202 | 0.035 | Yes |
| Cl | 0.813 | 0.582 | 0.303 | 0.002 | Yes |
| Br | 0.704 | 0.710 | 0.345 | 0.002 | Yes |
| $NO_3^-$ | 0.208 | 0.978 | 0.042 | 0.523 | No |
| $SO_4^{2-}$ | 0.968 | 0.250 | 0.251 | 0.011 | Yes |
| Ca | 0.790 | 0.613 | 0.227 | 0.019 | Yes |
| Na | 0.837 | 0.547 | 0.319 | 0.002 | Yes |
| Mg | 0.783 | 0.622 | 0.224 | 0.024 | Yes |
| K | 0.816 | 0.578 | 0.391 | 0.002 | Yes |
| Si | -0.961 | -0.276 | 0.537 | 0.001 | Yes |
| Sr | 0.865 | 0.502 | 0.200 | 0.032 | Yes |
| Al | -0.924 | -0.383 | 0.534 | 0.001 | Yes |
| Fe | -0.937 | -0.350 | 0.531 | 0.001 | Yes |
| Mn | 0.418 | -0.909 | 0.028 | 0.663 | No |
| B | 0.963 | 0.269 | 0.531 | 0.001 | Yes |
| Ba | 0.540 | 0.842 | 0.323 | 0.004 | Yes |
| Li | 0.666 | 0.746 | 0.322 | 0.002 | Yes |
| Zn | 0.573 | 0.820 | 0.235 | 0.019 | Yes |
| % Moisture | 0.707 | 0.707 | 0.669 | 0.001 | Yes |
| % C | 0.767 | 0.641 | 0.445 | 0.001 | Yes |
| % N | 0.903 | -0.430 | 0.037 | 0.586 | No |

**Table 10.** Differential abundance of OTUs at A) phylum, B) class, and C) family taxonomic levels that observed a significant ($p<0.01$) shift from control to dump conditions. $Log_2$ fold change between control and dump conditions shows degree of "doubling" of the taxa ("2" indicates 4-fold increase or decrease in observed taxa), “Std Error” indicates standard error of the $log_2$ fold change. P-values determined using the Wald test, and adjusted (“p-adjusted”) using the Benjamini-Hochberg method. Positive $log_2$ fold change indicates an increase in abundance in dump soil, while negative $log_2$ fold change indicates a decrease in abundance in dump soil.

**A. Phylum Level**

| **Phylum** | **$log_2$ fold change** | **Std. Error** | **p-adjusted** |
|---|---|---|---|
| Proteobacteria | 1.90 | 0.25 | 8.02E-13 |
| Epsilonbacteraeota | 7.07 | 0.98 | 1.06E-11 |
| Deferribacteres | 8.72 | 1.29 | 1.87E-10 |
| Patescibacteria | 3.95 | 0.60 | 4.71E-10 |
| Firmicutes | 3.08 | 0.47 | 5.20E-10 |
| Halanaerobiaeota | 6.00 | 0.92 | 5.20E-10 |
| Fibrobacteres | -1.32 | 0.31 | 0.00010272 |
| Thermotogae | 5.10 | 1.22 | 0.000147972 |

## B. Class Level

| Phylum | Class | $log_2$ fold change | Std. Error | p-adjusted |
|---|---|---|---|---|
| Proteobacteria | Gammaproteobacteria | 3.5 | 0.34 | 4.3E-22 |
| Patescibacteria | Gracilibacteria | 7.3 | 0.88 | 5.5E-15 |
| Epsilonbacteraeota | Campylobacteria | 7.0 | 0.97 | 1.5E-11 |
| Deferribacteres | Deferribacteres | 8.7 | 1.2 | 7.3E-11 |
| Firmicutes | Bacilli | 3.1 | 0.46 | 2.1E-10 |
| Halanaerobiaeota | Halanaerobiia | 6.0 | 0.91 | 1.3E-09 |
| Gemmatimonadetes | BD2-11 (Terrestrial Group) | 1.3 | 0.25 | 3.5E-06 |
| Acidobacteria | Unclassified Acidobacteria | 2.0 | 0.45 | 1.6E-04 |
| Thermotogae | Thermotogae | 5.1 | 1.2 | 2.6E-04 |
| Fibrobacteres | Fibrobacteria | -1.3 | 0.32 | 2.9E-04 |
| Gemmatimonadetes | Unclassified Gemmatimonadetes | 2.3 | 0.54 | 2.9E-04 |
| Chloroflexi | OLB14 | 2.5 | 0.61 | 3.8E-04 |
| Planctomycetes | BD7-11 | 1.4 | 0.34 | 4.8E-04 |
| Armatimonadetes | Chthonomonadetes | -1.4 | 0.35 | 8.0E-04 |
| Chloroflexi | Gitt-GS-136 | 1.1 | 0.29 | 1.1E-03 |
| Chloroflexi | TK10 | -0.56 | 0.15 | 1.5E-03 |
| Chloroflexi | Ktedonobacteria | -1.7 | 0.46 | 1.6E-03 |
| Acidobacteria | Subgroup 17 | 0.70 | 0.20 | 2.4E-03 |
| WPS-2 | WPS-2 | -1.4 | 0.41 | 3.2E-03 |
| Firmicutes | Clostridia | 1.9 | 0.57 | 4.8E-03 |
| Acidobacteria | Subgroup 22 | 1.4 | 0.44 | 6.2E-03 |
| Bacteroidetes | Rhodothermia | 1.5 | 0.49 | 8.6E-03 |
| Gemmatimonadetes | S0134 (Terrestrial Group) | 0.86 | 0.27 | 8.6E-03 |
| Cyanobacteria | Melainabacteria | 0.93 | 0.30 | 9.3E-03 |
| Armatimonadetes | uncultured | -0.57 | 0.19 | 9.8E-03 |
| Acidobacteria | Acidobacteriia | -0.74 | 0.24 | 9.8E-03 |

## C. Family Level

| Phylum | Class | Order | Family | $log_2$ fold change | Std. Error | p-adjusted |
|---|---|---|---|---|---|---|
| Bacteroidetes | Rhodothermia | Balneolales | Balneolaceae | 24 | 2.2 | 9.8E-25 |
| Proteobacteria | Gammaproteobacteria | Oceanospirillales | Alcanivoracaceae | 12 | 1.1 | 3.0E-22 |
| Proteobacteria | Gammaproteobacteria | Alteromonadales | Pseudoalteromonadaceae | 13 | 1.3 | 5.3E-20 |
| Proteobacteria | Alphaproteobacteria | Rhodospirillales | Thalassospiraceae | 11 | 1.2 | 6.4E-19 |
| Bacteroidetes | Bacteroidia | Flavobacteriales | Flavobacteriaceae | 7.0 | 0.77 | 8.2E-18 |
| Patescibacteria | Gracilibacteria | JGI_0000069-P22 | JGI_0000069-P22 | 10 | 1.2 | 7.6E-16 |
| Firmicutes | Bacilli | Bacillales | Sporolactobacillaceae | 7.6 | 1.0 | 2.0E-11 |
| Proteobacteria | Gammaproteobacteria | Cellvibrionales | Cellvibrionaceae | 6.1 | 0.85 | 3.8E-11 |
| Epsilonbacteraeota | Campylobacteria | Campylobacterales | Arcobacteraceae | 6.9 | 1.0 | 7.2E-11 |
| Bacteroidetes | Bacteroidia | Bacteroidales | Marinilabiliaceae | 7.9 | 1.1 | 7.8E-11 |
| Halanaerobiaeota | Halanaerobiia | Halanaerobiales | Halobacteroidaceae | 6.6 | 0.9 | 1.4E-10 |
| Deferribacteres | Deferribacteres | Deferribacterales | Deferribacteraceae | 8.8 | 1.3 | 1.4E-10 |
| Halanaerobiaeota | Halanaerobiia | Halanaerobiales | Halanaerobiaceae | 7.4 | 1.1 | 4.4E-10 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Stappiaceae | 7.7 | 1.2 | 1.0E-09 |
| Proteobacteria | Gammaproteobacteria | Alteromonadales | Marinobacteraceae | 6.0 | 0.9 | 2.2E-09 |
| Proteobacteria | Gammaproteobacteria | Oceanospirillales | Halomonadaceae | 5.7 | 0.9 | 3.0E-09 |
| Proteobacteria | Gammaproteobacteria | Oceanospirillales | Nitrincolaceae | 7.0 | 1.1 | 8.3E-09 |
| Firmicutes | Bacilli | Unclassified Bacilli | Unclassified Bacilli | 7.8 | 1.3 | 3.5E-08 |
| Proteobacteria | Gammaproteobacteria | Alteromonadales | Idiomarinaceae | 5.6 | 0.9 | 4.0E-08 |
| Proteobacteria | Gammaproteobacteria | Vibrionales | Vibrionaceae | 6.3 | 1.1 | 9.4E-08 |
| Proteobacteria | Gammaproteobacteria | Nitrococcales | Nitrococcales (Incertae Sedis) | 9.1 | 1.6 | 1.2E-07 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | 3.5 | 0.62 | 2.7E-07 |

***Table 10C. Family*** *cont.*

| **Phylum** | **Class** | **Order** | **Family** | **$log_2$ fold change** | **Std. Error** | **p-adjusted** |
|---|---|---|---|---|---|---|
| Proteobacteria | Deltaproteobacteria | Desulfovibrionales | Desulfovibrionaceae | 6.8 | 1.2 | 7.8E-07 |
| Firmicutes | Bacilli | Uncultured | Uncultured | 6.7 | 1.3 | 2.8E-06 |
| Actinobacteria | Nitriliruptoria | Nitriliruptorales | Nitriliruptoraceae | 6.6 | 1.3 | 3.6E-06 |
| Firmicutes | Bacilli | Bacillales | Unclassified | 2.3 | 0.44 | 3.6E-06 |
| Actinobacteria | Actinobacteria | Corynebacteriales | Dietziaceae | 5.4 | 1.0 | 3.9E-06 |
| Gemmatimonadetes | BD2-11 (Terrestrial Group) | BD2-11 | BD2-11 | 1.3 | 0.25 | 4.0E-06 |
| Proteobacteria | Gammaproteobacteria | Unclassified | Unclassified | 2.1 | 0.43 | 1.7E-05 |
| Proteobacteria | Alphaproteobacteria | Azospirillales | Unclassified | 4.7 | 1.0 | 2.7E-05 |
| Proteobacteria | Gammaproteobacteria | Betaproteobacteriales | Unclassified | -1.2 | 0.25 | 3.3E-05 |
| Cyanobacteria | Melainabacteria | Obscuribacterales | Obscuribacterales | 1.9 | 0.41 | 6.7E-05 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhodomicrobiaceae | 2.2 | 0.49 | 1.0E-04 |
| Acidobacteria | Unclassified | Unclassified | Unclassified | 2.0 | 0.45 | 1.0E-04 |
| Proteobacteria | Gammaproteobacteria | Betaproteobacteriales | TRA3-20 | -0.71 | 0.16 | 1.1E-04 |
| Gemmatimonadetes | Unclassified | Unclassified | Unclassified | 2.3 | 0.52 | 1.1E-04 |
| Actinobacteria | Actinobacteria | Corynebacteriales | Nocardiaceae | 2.6 | 0.58 | 1.1E-04 |
| Bacteroidetes | Bacteroidia | Cytophagales | Cyclobacteriaceae | 3.4 | 0.78 | 1.6E-04 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | 2.0 | 0.46 | 1.9E-04 |
| Thermotogae | Thermotogae | Petrotogales | Petrotogaceae | 5.3 | 1.2 | 2.0E-04 |
| Chloroflexi | OLB14 | OLB14 | OLB14 | 2.6 | 0.61 | 2.0E-04 |
| Proteobacteria | Alphaproteobacteria | uncultured | uncultured | 2.1 | 0.52 | 3.8E-04 |
| Planctomycetes | BD7-11 | BD7-11 | BD7-11 | 1.4 | 0.34 | 4.3E-04 |
| Proteobacteria | Alphaproteobacteria | Unclassified | Unclassified | 0.48 | 0.12 | 5.4E-04 |

***Table 10C. Family*** *cont.*

| Phylum | Class | Order | Family | $log_2$ fold change | Std. Error | p-adjusted |
|---|---|---|---|---|---|---|
| Fibrobacteres | Fibrobacteria | Fibrobacterales | Fibrobacteraceae | -1.3 | 0.33 | 6.3E-04 |
| Proteobacteria | Alphaproteobacteria | Dongiales | Dongiaceae | 2.1 | 0.52 | 6.6E-04 |
| Chloroflexi | Ktedonobacteria | C0119 | C0119 | -2.1 | 0.52 | 6.6E-04 |
| Proteobacteria | Deltaproteobacteria | Bradymonadales | Bradymonadaceae | 7.9 | 2.0 | 8.0E-04 |
| Firmicutes | Clostridia | Clostridiales | Lachnospiraceae | 2.6 | 0.68 | 8.7E-04 |
| Chloroflexi | Gitt-GS-136 | Gitt-GS-136 | Gitt-GS-136 | 1.1 | 0.29 | 9.7E-04 |
| Verrucomicrobia | Verrucomicrobiae | UA11 | UA11 | 2.6 | 0.69 | 1.3E-03 |
| Proteobacteria | Deltaproteobacteria | MBNT15 | MBNT15 | 2.8 | 0.74 | 1.3E-03 |
| Proteobacteria | Gammaproteobacteria | Immundisolibacterales | Immundisolibacteraceae | 5.4 | 1.5 | 1.6E-03 |
| Planctomycetes | Phycisphaerae | Phycisphaerales | Phycisphaeraceae | 0.9 | 0.25 | 2.2E-03 |
| Verrucomicrobia | Verrucomicrobiae | Methylacidiphilales | Methylacidiphilaceae | 2.3 | 0.64 | 2.5E-03 |
| Proteobacteria | Deltaproteobacteria | Myxococcales | Myxococcales | -1.3 | 0.37 | 2.8E-03 |
| Actinobacteria | Actinobacteria | Kineosporiales | Kineosporiaceae | -1.1 | 0.32 | 2.9E-03 |
| Acidobacteria | Blastocatellia (Subgroup 4) | 11-24 | 11-24 | -1.5 | 0.43 | 3.2E-03 |
| Proteobacteria | Gammaproteobacteria | Legionellales | Legionellaceae | 1.7 | 0.49 | 3.5E-03 |
| Actinobacteria | Actinobacteria | Micrococcales | Bogoriellaceae | 2.8 | 0.80 | 4.0E-03 |
| Bacteroidetes | Bacteroidia | Sphingobacteriales | Sphingobacteriaceae | -1.2 | 0.34 | 4.2E-03 |
| Bacteroidetes | Bacteroidia | Cytophagales | Microscillaceae | 1.0 | 0.30 | 4.3E-03 |
| Proteobacteria | Deltaproteobacteria | Myxococcales | Myxococcaceae | 1.7 | 0.48 | 4.3E-03 |
| Proteobacteria | Gammaproteobacteria | Steroidobacterales | Woeseiaceae | 6.5 | 1.9 | 4.6E-03 |
| WPS-2 | WPS-2 | WPS-2 | WPS-2 | -1.5 | 0.44 | 4.6E-03 |

***Table 10C. Family*** *cont.*

| **Phylum** | **Class** | **Order** | **Family** | **$log_2$ fold change** | **Std. Error** | **p-adjusted** |
|---|---|---|---|---|---|---|
| Acidobacteria | Acidobacteriia | Solibacterales | Solibacteraceae (Subgroup 3) | -0.8 | 0.25 | 4.8E-03 |
| Proteobacteria | Alphaproteobacteria | Thalassobaculales | Thalassobaculaceae | 6.2 | 1.8 | 5.2E-03 |
| Chloroflexi | TK10 | TK10 | TK10 | -0.60 | 0.18 | 5.2E-03 |
| Proteobacteria | Deltaproteobacteria | Myxococcales | mle1-27 | -1.2 | 0.36 | 5.2E-03 |
| Acidobacteria | Subgroup 17 | Subgroup 17 | Subgroup 17 | 0.74 | 0.22 | 5.9E-03 |
| Deinococcus-Thermus | Deinococci | Deinococcales | Trueperaceae | 1.8 | 0.57 | 8.0E-03 |
| Firmicutes | Clostridia | Thermoanaerobacterales | Thermoanaerobacteraceae | 5.4 | 1.7 | 8.2E-03 |
| Proteobacteria | Gammaproteobacteria | Alteromonadales | Alteromonadaceae | 5.9 | 1.8 | 8.2E-03 |
| Cyanobacteria | Oxyphotobacteria | Nostocales | Chroococcidiopsaceae | -2.3 | 0.74 | 8.6E-03 |
| Armatimonadetes | Chthonomonadetes | Chthonomonadales | Chthonomonadaceae | -1.2 | 0.39 | 8.9E-03 |
| Proteobacteria | Alphaproteobacteria | Caulobacterales | Hyphomonadaceae | 1.7 | 0.52 | 8.9E-03 |
| Acidobacteria | Blastocatellia (Subgroup 4) | Pyrinomonadales | Pyrinomonadaceae | -1.2 | 0.37 | 9.5E-03 |
| Acidobacteria | Subgroup 22 | Subgroup 22 | Subgroup 22 | 1.5 | 0.47 | 1.0E-02 |

## REFERENCES


Aguilera, M., Cabrera, A., Incerti, C., Fuentes, S., Russell, N. J., Ramos-Cormenzana, A., & Monteoliva-Sánchez, M. (2007). Chromohalobacter salarius sp. Nov., a moderately halophilic bacterium isolated from a solar saltern in Cabo de Gata, Almeria, southern Spain. *International Journal of Systematic and Evolutionary Microbiology*, *57*(Pt 6), 1238–1242. https://doi.org/10.1099/ijs.0.64953-0

Akob, D. M., Cozzarelli, I. M., Dunlap, D. S., Rowan, E. L., & Lorah, M. M. (2015). Organic and inorganic composition and microbiology of produced waters from Pennsylvania shale gas wells. *Applied Geochemistry*, *60*, 116–125. https://doi.org/10.1016/j.apgeochem.2015.04.011

Akob, D. M., Mumford, A. C., Orem, W., Engle, M. A., Klinges, J. G., Kent, D. B., & Cozzarelli, I. M. (2016). Wastewater Disposal from Unconventional Oil and Gas Development Degrades Stream Quality at a West Virginia Injection Facility. *Environmental Science & Technology*, *50*(11), 5517–5525. https://doi.org/10.1021/acs.est.6b00428

Akyon, B., Stachler, E., Wei, N., & Bibby, K. (2015). Microbial Mats as a Biological Treatment Approach for Saline Wastewaters: The Case of Produced Water from Hydraulic Fracturing. *Environmental Science & Technology*, *49*(10), 6172–6180. https://doi.org/10.1021/es505142t

Alauzet, C., & Jumas-Bilak, E. (2014). The Phylum Deferribacteres and the Genus Caldithrix. In E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, & F. Thompson (Eds.), *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea* (pp. 595–611). Springer. https://doi.org/10.1007/978-3-642-38954-2_162

Arthur, J. D., Layne, M., & Bohm, B. (2008). *Hydraulic Fracturing Considerations for Natural Gas Wells of the Marcellus Shale*. 17.

Bachar, A., Al-ashhab, A., Soares, M. I., M, Sklarz, M. Y., Angel, R., Ungar, E. D., & Gillor, O. (2010). Soil Microbial Abundance and Diversity Along a Low Precipitation Gradient. *Microbial Ecology; Heidelberg*, *60*(2), 453–461. http://dx.doi.org.mutex.gmu.edu/10.1007/s00248-010-9727-1

Barbot, E., Vidic, N. S., Gregory, K. B., & Vidic, R. D. (2013). Spatial and Temporal Correlation of Water Quality Parameters of Produced Waters from Devonian-Age Shale following Hydraulic Fracturing. *Environmental Science & Technology*, *47*(6), 2562–2569. https://doi.org/10.1021/es304638h

Bergmann, G. T., Bates, S. T., Eilers, K. G., Lauber, C. L., Caporaso, J. G., Walters, W. A., Knight, R., & Fierer, N. (2011). The under-recognized dominance of Verrucomicrobia in soil bacterial communities. *Soil Biology & Biochemistry*, *43*(7), 1450–1455. https://doi.org/10.1016/j.soilbio.2011.03.012

Bhandari, V., & Gupta, R. S. (2014). The Phylum Thermotogae. In E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, & F. Thompson (Eds.), *The Prokaryotes:*

*Other Major Lineages of Bacteria and The Archaea* (pp. 989–1015). Springer. https://doi.org/10.1007/978-3-642-38954-2_118

Blotta, I., Prestinaci, F., Mirante, S., & Cantàfora, A. (2005). Quantitative assay of total dsDNA with PicoGreen reagent and real-time fluorescent detection. *Annali Dell'Istituto Superiore Di Sanità*, *41*, 119–123.

Booker, A. E., Borton, M. A., Daly, R. A., Welch, S. A., Nicora, C. D., Hoyt, D. W., Wilson, T., Purvine, S. O., Wolfe, R. A., Sharma, S., Mouser, P. J., Cole, D. R., Lipton, M. S., Wrighton, K. C., & Wilkins, M. J. (2017). Sulfide Generation by Dominant Halanaerobium Microorganisms in Hydraulically Fractured Shales. *MSphere*, *2*(4). https://doi.org/10.1128/mSphereDirect.00257-17

Bowman, J. P. (2016). Salegentibacter. In *Bergey's Manual of Systematics of Archaea and Bacteria* (pp. 1–11). American Cancer Society. https://doi.org/10.1002/9781118960608.gbm00338.pub2

Büdel, B. (2011). Chroococcidiopsis. In J. Reitner & V. Thiel (Eds.), *Encyclopedia of Geobiology* (pp. 273–274). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9212-1_56

Bureau of Land Management. (2016, August 17). *Bureau of Land Management Programs: Energy and Minerals: Oil and Gas* [Government]. Www.Blm.Gov. https://www.blm.gov/programs/energy-and-minerals/oil-and-gas

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N., & Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the*

*National Academy of Sciences*, *108*(Supplement 1), 4516–4522. https://doi.org/10.1073/pnas.1000080107

*Chihuahuan Desert Ecoregion (U.S. National Park Service)*. (2018, September 20). [Government]. National Park Service. https://www.nps.gov/im/chdn/ecoregion.htm

Clark, C. E., & Veil, J. A. (Environmental S. D. (2009, September 1). *Produced water volumes and management practices in the United States.* [Report]. Digital Library. https://doi.org/10.2172/1007397

Clark, J. S., Campbell, J. H., Grizzle, H., Acosta-Martìnez, V., & Zak, J. C. (2009). Soil Microbial Community Response to Drought and Precipitation Variability in the Chihuahuan Desert. *Microbial Ecology*, *57*(2), 248–260. https://doi.org/10.1007/s00248-008-9475-7

Cluff, M. A., Hartsock, A., MacRae, J. D., Carter, K., & Mouser, P. J. (2014). Temporal Changes in Microbial Ecology and Geochemistry in Produced Water from Hydraulically Fractured Marcellus Shale Gas Wells. *Environmental Science & Technology*, *48*(11), 6508–6517. https://doi.org/10.1021/es501173p

Connon, S. A., Lester, E. D., Shafaat, H. S., Obenhuber, D. C., & Ponce, A. (2007). Bacterial diversity in hyperarid Atacama Desert soils. *Journal of Geophysical Research: Biogeosciences*, *112*(G4). https://doi.org/10.1029/2006JG000311

Cozzarelli, I. M., Skalak, K. J., Kent, D. B., Engle, M. A., Benthem, A., Mumford, A. C., Haase, K., Farag, A., Harper, D., Nagel, S. C., Iwanowicz, L. R., Orem, W. H., Akob, D. M., Jaeschke, J. B., Galloway, J., Kohler, M., Stoliker, D. L., & Jolly,

G. D. (2017). Environmental signatures and effects of an oil and gas wastewater spill in the Williston Basin, North Dakota. *Science of The Total Environment*, *579*, 1781–1793. https://doi.org/10.1016/j.scitotenv.2016.11.157

Crits-Christoph, A., Robinson, C. K., Barnum, T., Fricke, W. F., Davila, A. F., Jedynak, B., McKay, C. P., & DiRuggiero, J. (2013). Colonization patterns of soil microbial communities in the Atacama Desert. *Microbiome*, *1*(1), 28. https://doi.org/10.1186/2049-2618-1-28

Duran, R. (2010). Marinobacter. In K. N. Timmis (Ed.), *Handbook of Hydrocarbon and Lipid Microbiology* (pp. 1725–1735). Springer. https://doi.org/10.1007/978-3-540-77587-4_122

Engle, M. A., Cozzarelli, I. M., & Smith, B. D. (2014). *USGS Investigations of Water Produced During Hydrocarbon Reservoir Development* (p. 4) [Fact Sheet]. U.S. Geological Survey. https://dx.doi.org/10.3133/fs20143104

Engle, M. A., Reyes, F. R., Varonka, M. S., Orem, W. H., Ma, L., Ianno, A. J., Schell, T. M., Xu, P., & Carroll, K. C. (2016). Geochemistry of formation waters from the Wolfcamp and "Cline" shales: Insights into brine origin, reservoir connectivity, and fluid flow in the Permian Basin, USA. *Chemical Geology*, *425*, 76–92. https://doi.org/10.1016/j.chemgeo.2016.01.025

Engle, M. A., Saraswathula, V., Thordsen, J. J., Morrissey, E. A., Gans, K. D., Blondes, M. S., Kharaka, Y. K., Rowan, E. L., & Reidy, M. E. (2019). *U.S. Geological Survey National Produced Waters Geochemical Database v2.3* [Data set]. U.S. Geological Survey. https://doi.org/10.5066/F7J964W8

Ensign, J. C. (1978). Formation, Properties, and Germination of Actinomycete Spores. *Annual Review of Microbiology*, *32*(1), 185–219. https://doi.org/10.1146/annurev.mi.32.100178.001153

Evans, M. V., Panescu, J., Hanson, A. J., Welch, S. A., Sheets, J. M., Nastasi, N., Daly, R. A., Cole, D. R., Darrah, T. H., Wilkins, M. J., Wrighton, K. C., & Mouser, P. J. (2018). Members of Marinobacter and Arcobacter Influence System Biogeochemistry During Early Production of Hydraulically Fractured Natural Gas Wells in the Appalachian Basin. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.02646

Fahrenfeld, N. L., Delos Reyes, H., Eramo, A., Akob, D. M., Mumford, A. C., & Cozzarelli, I. M. (2017). Shifts in microbial community structure and function in surface waters impacted by unconventional oil and gas wastewater revealed by metagenomics. *Science of The Total Environment*, *580*, 1205–1213. https://doi.org/10.1016/j.scitotenv.2016.12.079

Frithsen, J. (2015, June). *Assessment of the Potential Impacts of Hydraulic Fracturing for Oil and Gas on Drinking Water Resources (External Review Draft)* [Reports & Assessments]. Www.Epa.Gov. https://cfpub.epa.gov/ncea/hfstudy/recordisplay.cfm?deid=244651

Gallegos, T. J., & Varela, B. A. (2015). *Trends in Hydraulic Fracturing Distributions and Treatment Fluids, Additives, Proppants, and Water Volumes Applied to Wells Drilled in the United States from 1947 through 2010—Data Analysis and Comparison to the Literature* (Scientific Investigations Report 2014, p. 15)

[Scientific Investigations Report]. U.S. Geological Survey. http://dx.doi.org/10.3133/sir20145131

Gao, Q., & Garcia-Pichel, F. (2011). Microbial ultraviolet sunscreens. *Nature Reviews Microbiology*, *9*(11), 791–802. https://doi.org/10.1038/nrmicro2649

García, M. T., Mellado, E., Ostos, J. C., & Ventosa, A. (2004). Halomonas organivorans sp. Nov., a moderate halophile able to degrade aromatic compounds. *International Journal of Systematic and Evolutionary Microbiology*, *54*(Pt 5), 1723–1728. https://doi.org/10.1099/ijs.0.63114-0

Gaswirth, S. B., French, K. L., Pitman, J. K., Marra, K. R., Mercier, T. J., Leathers-Miller, H. M., Schenk, C. J., Tennyson, M. E., Woodall, C. A., Brownfield, M. E., Finn, T. M., & Le, P. A. (2018). *Assessment of undiscovered continuous oil and gas resources in the Wolfcamp Shale and Bone Spring Formation of the Delaware Basin, Permian Basin Province, New Mexico and Texas, 2018* (Fact Sheet) [USGS Numbered Series]. U.S. Geological Survey. http://pubs.er.usgs.gov/publication/fs20183073

Gauthier, M. J., Lafay, B., Christen, R., Fernandez, L., Acquaviva, M., Bonin, P., & Bertrand, J.-C. (1992). Marinobacter hydrocarbonoclasticus gen. Nov., sp. Nov., a New, Extremely Halotolerant, Hydrocarbon-Degrading Marine Bacterium. *International Journal of Systematic and Evolutionary Microbiology*, *42*(4), 568–576. https://doi.org/10.1099/00207713-42-4-568

Glaeser, S. P., & Kämpfer, P. (2014). The Family Sphingomonadaceae. In E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, & F. Thompson (Eds.), *The Prokaryotes:*

*Alphaproteobacteria and Betaproteobacteria* (pp. 641–707). Springer. https://doi.org/10.1007/978-3-642-30197-1_302

Gregory, K. B., Vidic, R. D., & Dzombak, D. A. (2011). Water Management Challenges Associated with the Production of Shale Gas by Hydraulic Fracturing. *Elements*, *7*(3), 181–186. https://doi.org/10.2113/gselements.7.3.181

Hara, A., Syutsubo, K., & Harayama, S. (2003). Alcanivorax which prevails in oil-contaminated seawater exhibits broad substrate specificity for alkane degradation. *Environmental Microbiology*, *5*(9), 746–753. https://doi.org/10.1046/j.1468-2920.2003.00468.x

Islam, Z. F., Cordero, P. R. F., Feng, J., Chen, Y.-J., Bay, S. K., Jirapanjawat, T., Gleadow, R. M., Carere, C. R., Stott, M. B., Chiri, E., & Greening, C. (2019). Two Chloroflexi classes independently evolved the ability to persist on atmospheric hydrogen and carbon monoxide. *The ISME Journal*, *13*(7), 1801–1813. https://doi.org/10.1038/s41396-019-0393-0

Ivanova, E. P., Romanenko, L. A., Chun, J., Matte, M. H., Matte, G. R., Mikhailov, V. V., Svetashev, V. I., Huq, A., Maugel, T., & Colwell, R. R. (2000). Idiomarina gen. Nov., comprising novel indigenous deep-sea bacteria from the Pacific Ocean, including descriptions of two species, Idiomarina abyssalis sp. Nov. And Idiomarina zobellii sp. Nov. *International Journal of Systematic and Evolutionary Microbiology*, *50 Pt 2*, 901–907. https://doi.org/10.1099/00207713-50-2-901

Ivanova, Elena P., Ng, H. J., & Webb, H. K. (2014). The Family Pseudoalteromonadaceae. In E. Rosenberg, E. F. DeLong, S. Lory, E.

Stackebrandt, & F. Thompson (Eds.), *The Prokaryotes: Gammaproteobacteria* (pp. 575–582). Springer. https://doi.org/10.1007/978-3-642-38922-1_229

Johnsen, A. R., Christensen, J. H., Poulsen, K. G., Hansen, L. H., Ellegaard-Jensen, L., & Fritt-Rasmussen, J. (2019). A study of microbial crude oil degradation in water and sediment from the Greenland Sea. : *: Aarhus University, DCE – Danish Centre for Environment and Energy ©*, *No. 347*, 92.

Kassotis, C. D., Iwanowicz, L. R., Akob, D. M., Cozzarelli, I. M., Mumford, A. C., Orem, W. H., & Nagel, S. C. (2016). Endocrine disrupting activities of surface water associated with a West Virginia oil and gas industry wastewater disposal site. *Science of The Total Environment*, *557–558*, 901–910. https://doi.org/10.1016/j.scitotenv.2016.03.113

Kielak, A. M., Barreto, C. C., Kowalchuk, G. A., van Veen, J. A., & Kuramae, E. E. (2016). The Ecology of Acidobacteria: Moving beyond Genes and Genomes. *Frontiers in Microbiology*, *7*. https://doi.org/10.3389/fmicb.2016.00744

King, G. E. (2010). *Thirty Years of Gas Shale Fracturing: What Have We Learned?* 50.

King, G. E. (2012). Hydraulic Fracturing 101: What Every Representative, Environmentalist, Regulator, Reporter, Investor, University Researcher, Neighbor and Engineer Should Know About Estimating Frac Risk and Improving Frac Performance in Unconventional Gas and Oil Wells. *SPE Hydraulic Fracturing Technology Conference*. SPE Hydraulic Fracturing Technology Conference, The Woodlands, Texas, USA. https://doi.org/10.2118/152596-MS

Kondash, A. J., Lauer, N. E., & Vengosh, A. (2018). The intensification of the water footprint of hydraulic fracturing. *Science Advances*, *4*(8). https://doi.org/10.1126/sciadv.aar5982

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Appl. Environ. Microbiol.*, *79*(17), 5112–5120. https://doi.org/10.1128/AEM.01043-13

Kuske, C. R., Barns, S. M., & Busch, J. D. (1997). Diverse uncultivated bacterial groups from soils of the arid southwestern United States that are present in many geographic regions. *Applied and Environmental Microbiology*, *63*(9), 3614–3621.

Lauer, N. E., Harkness, J. S., & Vengosh, A. (2016). Brine Spills Associated with Unconventional Oil Development in North Dakota. *Environmental Science & Technology*, *50*(10), 5389–5397. https://doi.org/10.1021/acs.est.5b06349

Leung, P. M., Bay, S. K., Meier, D. V., Chiri, E., Cowan, D. A., Gillor, O., Woebken, D., & Greening, C. (2020). Energetic Basis of Microbial Growth and Persistence in Desert Ecosystems. *MSystems*, *5*(2). https://doi.org/10.1128/mSystems.00495-19

Lim, J.-M., Jeon, C. O., Lee, J.-C., Song, S.-M., Kim, K.-Y., & Kim, C.-J. (2006). Marinimicrobium koreense gen. Nov., sp. Nov. And Marinimicrobium agarilyticum sp. Nov., novel moderately halotolerant bacteria isolated from tidal flat sediment in Korea. *International Journal of Systematic and Evolutionary Microbiology*, *56*(3), 653–657. https://doi.org/10.1099/ijs.0.64075-0

Lipus, D., Vikram, A., Ross, D., Bain, D., Gulliver, D., Hammack, R., & Bibby, K. (2017). Predominance and Metabolic Potential of Halanaerobium spp. In Produced Water from Hydraulically Fractured Marcellus Shale Wells. *Applied and Environmental Microbiology*, *83*(8). https://doi.org/10.1128/AEM.02659-16

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550–550. PubMed. https://doi.org/10.1186/s13059-014-0550-8

Lozupone, C., & Knight, R. (2005). UniFrac: A New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology*, *71*(12), 8228–8235. https://doi.org/10.1128/AEM.71.12.8228-8235.2005

Makhalanyane, T. P., Valverde, A., Gunnigle, E., Frossard, A., Ramond, J.-B., & Cowan, D. A. (2015). Microbial ecology of hot desert edaphic systems. *FEMS Microbiology Reviews*, *39*(2), 203–221. https://doi.org/10.1093/femsre/fuu011

Marín, I., & Arahal, D. R. (2014). The Family Beijerinckiaceae. In E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, & F. Thompson (Eds.), *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* (pp. 115–133). Springer. https://doi.org/10.1007/978-3-642-30197-1_255

Maza, F., Maldonado, J., Vásquez-Dean, J., Mandakovic, D., Gaete, A., Cambiazo, V., & González, M. (2019). Soil Bacterial Communities From the Chilean Andean Highlands: Taxonomic Composition and Culturability. *Frontiers in Bioengineering and Biotechnology*, *7*. https://doi.org/10.3389/fbioe.2019.00010

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE*, *8*(4), 1–11. https://doi.org/10.1371/journal.pone.0061217

Mumford, A. C., Akob, D. M., Klinges, J. G., & Cozzarelli, I. M. (2018). Common Hydraulic Fracturing Fluid Additives Alter the Structure and Function of Anaerobic Microbial Communities. *Applied and Environmental Microbiology*, *84*(8), e02729-17. https://doi.org/10.1128/AEM.02729-17

Murali Mohan, A., Hartsock, A., Bibby, K. J., Hammack, R. W., Vidic, R. D., & Gregory, K. B. (2013). Microbial Community Changes in Hydraulic Fracturing Fluids and Produced Water from Shale Gas Extraction. *Environmental Science & Technology*, *47*(22), 13141–13150. https://doi.org/10.1021/es402928b

Murali Mohan, A., Hartsock, A., Hammack, R. W., Vidic, R. D., & Gregory, K. B. (2013). Microbial communities in flowback water impoundments from hydraulic fracturing for recovery of shale gas. *FEMS Microbiology Ecology*, *86*(3), 567–580. https://doi.org/10.1111/1574-6941.12183

Naether, D. J., Slawtschew, S., Stasik, S., Engel, M., Olzog, M., Wick, L. Y., Timmis, K. N., & Heipieper, H. J. (2013). Adaptation of the Hydrocarbonoclastic Bacterium Alcanivorax borkumensis SK2 to Alkanes and Toxic Organic Compounds: A Physiological and Transcriptomic Approach. *Applied and Environmental Microbiology*, *79*(14), 4282–4293. https://doi.org/10.1128/AEM.00694-13

*New Mexico Environment Department New Mexico Produced Water*. (n.d.). Retrieved June 26, 2020, from https://www.env.nm.gov/new-mexico-produced-water/

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community Ecology Package* (2.5-5) [Computer software]. https://CRAN.R-project.org/package=vegan

Orem, W., Varonka, M., Crosby, L., Haase, K., Loftin, K., Hladik, M., Akob, D. M., Tatu, C., Mumford, A., Jaeschke, J., Bates, A., Schell, T., & Cozzarelli, I. (2017). Organic geochemistry and toxicology of a stream impacted by unconventional oil and gas wastewater disposal operations. *Applied Geochemistry*, *80*, 155–167. https://doi.org/10.1016/j.apgeochem.2017.02.016

Oren, A. (2014). The Order Halanaerobiales, and the Families Halanaerobiaceae and Halobacteroidaceae. In E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, & F. Thompson (Eds.), *The Prokaryotes: Firmicutes and Tenericutes* (pp. 153–177). Springer. https://doi.org/10.1007/978-3-642-30120-9_218

Pascual, J., Huber, K. J., & Overmann, J. (2018). Pyrinomonadaceae. In *Bergey's Manual of Systematics of Archaea and Bacteria* (pp. 1–4). American Cancer Society. https://doi.org/10.1002/9781118960608.fbm00310

*Permian region is expected to drive U.S. crude oil production growth through 2019—Today in Energy—U.S. Energy Information Administration (EIA)*. (n.d.). Retrieved July 17, 2020, from https://www.eia.gov/todayinenergy/detail.php?id=36936

Prado, B., Lizama, C., Aguilera, M., Ramos-Cormenzana, A., Fuentes, S., Campos, V., & Monteoliva-Sánchez, M. (2006). Chromohalobacter nigrandesensis sp. Nov., a

moderately halophilic, Gram-negative bacterium isolated from Lake Tebenquiche on the Atacama Saltern, Chile. *International Journal of Systematic and Evolutionary Microbiology*, *56*(3), 647–651. https://doi.org/10.1099/ijs.0.63983-0

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLOS ONE*, *5*(3), e9490. https://doi.org/10.1371/journal.pone.0009490

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Gloeckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

R Core Team. (2014). *R: The R Project for Statistical Computing*. https://www.r-project.org/

Rainey, F. A., Zhilina, T. N., Boulygina, E. S., Stackebrandt, E., Tourova, T. P., & Zavarzin, G. A. (1995). The Taxonomic Status of the Fermentative Halophilic Anaerobic Bacteria: Description of Haloanaerobiales ord. nov., Halobacteroidaceae fam. nov., Orenia gen. nov. and further Taxonomic Rearrangements at the Genus and Species Level. *Anaerobe*, *1*(4), 185–199. https://doi.org/10.1006/anae.1995.1018

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *4*. https://doi.org/10.7717/peerj.2584

Rowan, E. L., Engle, M. A., Kraemer, T. F., Schroeder, K. T., Hammack, R. W., & Doughten, M. W. (2015). Geochemical and isotopic evolution of water produced from Middle Devonian Marcellus shale gas wells, Appalachian basin, Pennsylvania. *Aapg Bulletin*, *99*(2), 181–206. https://doi.org/10.1306/07071413146

Sabirova, J. S., Becker, A., Lünsdorf, H., Nicaud, J.-M., Timmis, K. N., & Golyshin, P. N. (2011). Transcriptional profiling of the marine oil-degrading bacterium Alcanivorax borkumensis during growth on n-alkanes: Transcriptomic responses of Alcanivorax borkumensis. *FEMS Microbiology Letters*, *319*(2), 160–168. https://doi.org/10.1111/j.1574-6968.2011.02279.x

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Horn, D. J. V., & Weber, C. F. (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology*, *75*(23), 7537–7541. https://doi.org/10.1128/AEM.01541-09

Schneiker, S., dos Santos, V. A. M., Bartels, D., Bekel, T., Brecht, M., Buhrmester, J., Chernikova, T. N., Denaro, R., Ferrer, M., Gertler, C., Goesmann, A., Golyshina, O. V., Kaminski, F., Khachane, A. N., Lang, S., Linke, B., McHardy, A. C., Meyer, F., Nechitaylo, T., Pühler, A., Regenhardt, D., Rupp, O., Sabirova, J.S., Selbitschka, W., Yakimov, M.M., Timmis, K.N., Vorhölter, F.J., Weidner, S., Kaiser, O., Golyshin, P. N. (2006). Genome sequence of the ubiquitous

hydrocarbon-degrading marine bacterium Alcanivorax borkumensis. *Nature Biotechnology*, *24*(8), 997–1004. https://doi.org/10.1038/nbt1232

Siamphan, C., & Kim, W. (2014). Salegentibacter chungangensis sp. Nov., isolated from a sea sand and emended description of the genus Salegentibacter. *International Journal of Systematic and Evolutionary Microbiology*, *64*(Pt 5), 1514–1519. https://doi.org/10.1099/ijs.0.059170-0

Skujinš, J. (1984). Microbial Ecology of Desert Soils. In K. C. Marshall (Ed.), *Advances in Microbial Ecology* (pp. 49–91). Springer US. https://doi.org/10.1007/978-1-4684-8989-7_2

Song, L., Liu, H., Cai, S., & Zhou, Y. (2019). Marinimicrobium alkaliphilum sp. Nov., an alkaliphilic bacterium isolated from soil and emended description of the genus Marinimicrobium. *International Journal of Systematic and Evolutionary Microbiology*, *69*(2), 481–485. https://doi.org/10.1099/ijsem.0.003183

Struchtemeyer, C. G., & Elshahed, M. S. (2012). Bacterial communities associated with hydraulic fracturing fluids in thermogenic natural gas wells in North Central Texas, USA. *FEMS Microbiology Ecology*, *81*(1), 13–25. https://doi.org/10.1111/j.1574-6941.2011.01196.x

Taylor, N. T., Davis, K. M., Abad, H., McClung, M. R., & Moran, M. D. (2017). Ecosystem services of the Big Bend region of the Chihuahuan Desert. *Ecosystem Services*, *27*, 48–57. https://doi.org/10.1016/j.ecoser.2017.07.017

Torres-Cortés, G., Millán, V., Fernández-González, A. J., Aguirre-Garrido, J. F., Ramírez-Saad, H. C., Fernández-López, M., Toro, N., & Martínez-Abarca, F.

(2012). Bacterial community in the rhizosphere of the cactus species Mammillaria carnea during dry and rainy seasons assessed by deep sequencing. *Plant and Soil*, *357*(1), 275–288. https://doi.org/10.1007/s11104-012-1152-4

Tourna, M., Stieglmeier, M., Spang, A., Könneke, M., Schintlmeister, A., Urich, T., Engel, M., Schloter, M., Wagner, M., Richter, A., & Schleper, C. (2011). Nitrososphaera viennensis, an ammonia oxidizing archaeon from soil. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(20), 8420–8425. https://doi.org/10.1073/pnas.1013488108

U.S. Energy Information Administration. (2016). *Annual Energy Outlook 2016 with projections to 2040*. 256.

U.S. Energy Information Administration. (2019, April 5). *EIA - Annual Energy Outlook 2019*. https://www.eia.gov/outlooks/aeo/

US EPA, ORD. (2013, February 26). *EPA's Study of Hydraulic Fracturing for Oil and Gas and Its Potential Impact on Drinking Water Resources* [Overviews and Factsheets]. US EPA. https://www.epa.gov/hfstudy

US EPA, OW. (2015, March 27). *Class II Oil and Gas Related Injection Wells* [Overviews and Factsheets]. US EPA. https://www.epa.gov/uic/class-ii-oil-and-gas-related-injection-wells

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (2nd ed.). Springer International Publishing. https://doi.org/10.1007/978-3-319-24277-4

Wiesman, Z. (2009). Chapter 3—Key characteristics of the desert environment. In Z. Wiesman (Ed.), *Desert Olive Oil Cultivation* (pp. 31–53). Academic Press. https://doi.org/10.1016/B978-0-12-374257-5.00003-8

Xia, J., Xie, Z.-H., Dunlap, C. A., Rooney, A. P., & Du, Z.-J. (2017). Rhodohalobacter halophilus gen. Nov., sp. Nov., a moderately halophilic member of the family Balneolaceae. *International Journal of Systematic and Evolutionary Microbiology*, *67*(5), 1281–1287. https://doi.org/10.1099/ijsem.0.001806

Yakimov, M. M., Golyshin, P. N., Crisafi, F., Denaro, R., & Giuliano, L. (2019). Marine, Aerobic Hydrocarbon-Degrading Gammaproteobacteria: The Family Alcanivoracaceae. In T. J. McGenity (Ed.), *Taxonomy, Genomics and Ecophysiology of Hydrocarbon-Degrading Microbes* (pp. 167–179). Springer International Publishing. https://doi.org/10.1007/978-3-030-14796-9_24

Zaneveld, J. R., McMinds, R., & Vega Thurber, R. (2017). Stress and stability: Applying the Anna Karenina principle to animal microbiomes. *Nature Microbiology*, *2*, 17121. https://doi.org/10.1038/nmicrobiol.2017.121

Zeng, Y., Feng, F., Medová, H., Dean, J., & Koblížek, M. (2014). Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. *Proceedings of the National Academy of Sciences*, *111*(21), 7795–7800. https://doi.org/10.1073/pnas.1400295111

Zhang, K., Shi, Y., Cui, X., Yue, P., Li, K., Liu, X., Tripathi, B. M., & Chu, H. (2019). Salinity Is a Key Determinant for Soil Microbial Communities in a Desert Ecosystem. *MSystems*, *4*(1). https://doi.org/10.1128/mSystems.00225-18

Zheng, Y., Saitou, A., Wang, C.-M., Toyoda, A., Minakuchi, Y., Sekiguchi, Y., Ueda, K., Takano, H., Sakai, Y., Abe, K., Yokota, A., & Yabe, S. (2019). Genome Features and Secondary Metabolites Biosynthetic Potential of the Class Ktedonobacteria. *Frontiers in Microbiology, 10*. https://doi.org/10.3389/fmicb.2019.00893

Zvyagintsev, D., Zenova, G., Doroshenko, E., Gryadunova, A., Gracheva, T., & Sudnitsyn, I. (2007). Actinomycete growth in conditions of low moisture. *Biology Bulletin, 34*(3), 242–247. https://doi.org/10.1134/S1062359007030053

l

## BIOGRAPHY

Mitra Kashani graduated from Freedom High School, Chantilly, Virginia, in 2014. She received her Bachelor of Science in Biology from George Mason University in 2018. She was employed as a Graduate Teaching Assistant at George Mason University and conducted her Master's thesis research at the U.S. Geological Survey, Reston, Virginia for two years. Mitra will receive her Master of Science in Biology from George Mason University in Summer 2020.