**Michigan
Technological
University**

Michigan Technological University
Digital Commons @ Michigan Tech

Dissertations, Master's Theses and Master's Reports

2018

MICROBIAL COMMUNITY RESPONSE TO LIGHT AND HEAVY CRUDE OIL IN FRESHWATER SYSTEMS

Timothy M. Butler

Michigan Technological University, tbutler@mtu.edu

Copyright 2018 Timothy M. Butler

Recommended Citation

Butler, Timothy M., "MICROBIAL COMMUNITY RESPONSE TO LIGHT AND HEAVY CRUDE OIL IN FRESHWATER SYSTEMS", Open Access Master's Thesis, Michigan Technological University, 2018.
<https://doi.org/10.37099/mtu.dc.etr/715>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etr>

Part of the [Environmental Microbiology and Microbial Ecology Commons](#), [Integrative Biology Commons](#), and the [Terrestrial and Aquatic Ecology Commons](#)

**MICROBIAL COMMUNITY RESPONSE TO LIGHT AND HEAVY
CRUDE OIL IN FRESHWATER SYSTEMS**

By
Timothy M. Butler

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Biological Sciences

MICHIGAN TECHNOLOGICAL UNIVERSITY

2018

© 2018 Timothy M. Butler

This thesis has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Biological Sciences.

Department of Biological Sciences

Thesis Advisor: *Stephen Techtmann*

Committee Member: *Amy Marcarelli*

Committee Member: *Rupali Datta*

Committee Member: *Gordon Paterson*

Department Chair: *Chandrashekhar Joshi*

Table of Contents

Acknowledgments	v
Abstract.....	vii
1 Introduction.....	8
1.1 Oil Spill/Oil Transport	2
1.2 Comparison of BP, Valdez. and Kalamazoo	4
1.3 Oil Types.....	5
1.4 Physical and Biological Remediation	7
1.5 Constraints on Bioremediation.....	9
1.6 Bacteria Involved with Bioremediation	9
1.7 Freshwater Oil.....	10
1.8 Study Goals.....	12
2 Methods.....	13
2.1 Sampling Location and Water Collection	13
2.2 Microcosm Setup.....	16
2.3 Water Filtration/DNA Extraction	17
2.4 16S PCR and Sequencing.....	18
2.5 Sequencing Analysis Using QIIME1	19
2.6 Analysis in R.....	20
2.7 Alpha Diversity Analysis	20
2.8 Analysis of Changes in Microbial Community Composition.....	21
2.9 Differential Abundance Analysis to Identify Responsive OTUs.	21
2.10 Hydrocarbon Extraction	22
2.11 Gas Chromatography Analysis	23
3 Results	24
3.1 Differences in Microbial Communities Between Sites.....	24
3.2 Community Response to Oil Exposure	29
3.3 Bakken vs DilBit.....	44
4 Discussion.....	52
4.1 Differences Between Sites.....	52

4.2 Microbial Community Response to Oil	54
4.3 Comparison of the Microbial Response to Heavy and Light Crude Oil	60
5 Conclusion	64
6 References	65
7 Additional community data	70
7.1 Taxa Bar Plot	70

Acknowledgments

The amount of people that have helped me along the way in finishing this research is unmeasurable. My appreciation and thanks for everyone is without limit.

I would like to begin by first thanking my advisor Stephen Techtmann who has been the most patient and thorough teacher I have ever had. Throughout the years you have become a very good friend and words cannot describe how much I appreciate the help and support that you have given me. I can never repay you for the endless hours spent in your office where you were answering questions and guiding me through how to do different lab techniques, bioinformatics, and for letting me pick your brain for random inquiries of mine.

I would also like to thank my committee members; Dr. Gordon Paterson, Dr. Rupali Datta, and Dr. Amy Marcarelli. Your knowledge, advice, and lending of materials has been a large help towards completing this thesis.

Next, I would like to thank my fiancé Paige Webb who has been there throughout my whole Master's program and has dedicated countless hours supporting me and helping me through the process of research and thesis writing. I would also like to thank her for all her work as an undergraduate worker helping with DNA extractions, and other lab-based experiments.

Emma Byrne also deserves special thanks even though she wasn't part of the project. Without her help running multiple samples on the GC we wouldn't have figured out the GC data was so off as easily. Additionally, I would like to

thank all the undergraduates that have been a part of my research (Matthew Breneman, Joshua Bocker, Caitayln Nielson).

I also owe a large thanks to Jorge Santo Domingo and Michael Elk from the EPA office of Research and Development for sequencing the 16S rRNA genes from the Straits Microcosms, Edith Holder from Pegasus, Inc., Cincinnati, OH, for supplying the DilBit used in my research, Tim Schultz from Intertek for supplying the Bakken crude oil used in my research and the Captain and Crew of the NOAA 5501 for assistance in sample collection for this study and making the voyage through the Great Lakes a memorable experience. I would also like to thank Karen Butler my mother for all the support and help that she has put in.

I would finally like to thank the Department of Biological Sciences for the ability to use their facilities and for employing me as a Graduate Teaching Assistant so I could conduct my research and afford to live.

Abstract

With increased demand for oil, there is an increased risk for oil spills in many environments. A number of pipelines transport oil near or across freshwater systems including the Great Lakes. Microbes are capable of breaking down oil and have thus been proposed as tools for oil spill response through bioremediation. There is a need to understand the microbial response to diverse oil types in freshwater environments due to the lack of research into this topic. This study's main objectives are to understand how the freshwater microbial communities respond to oil, and how the bacterial communities may respond to different oil types. The bacterial community response to oil was examined at seven different geographical locations in the Great Lakes. Additionally, the microbial community response to two very different oil types. A heavy oil, Cold Lake Diluted Bitumen (DilBit), and a light oil (Bakken) were examined. Our results demonstrated a distinct community composition at different sites throughout the Great Lakes. Furthermore, there was a distinct response to oil depending on the location. Additionally, our results showed a distinct community response to the two oil types tested Bakken and DilBit crudes. The primary organisms that responded to oil in our microcosms in the Great Lakes were bacteria from the families; Sphingomonadaceae, Rhodocyclaceae, Burkholderiales, and Comamonadaceae. Our results also indicated that the extent of response to oil varied greatly between offshore, the Straits, and inland systems. These findings suggest that in the case of an oil spill in the Great Lakes, the location of the spill and type of oil should be taken into account in planning

bioremediation efforts. Our results demonstrate that in most locations in the Great Lakes, a common group of bacteria can be expected to respond to the oil exposure indicating the potential for oil biodegradation throughout the Great Lakes.

1 Introduction

Prospecting for and transporting of oil has expanded with increasing demand for oil throughout the world. There is great concern regarding the potential impact of an accidental release of oil from pipelines transporting oil across freshwater systems. However to date there is very little research that characterizes the microbial community response to oil in freshwater systems. This study aims to answer these questions and provide insight into the identity of oil degrading taxa and the microbial community's role in oil response in freshwater systems. Pipelines such as the Dakota Access Pipeline and the Enbridge Line 5 pipeline transport oil across freshwater bodies. The Enbridge Line 5 oil pipeline currently crosses the Great Lakes at the Straits of Mackinac. These pipelines can carry a variety of products from light crude oil to heavy diluted bitumen (DilBit). The value and sensitivity of freshwater systems including the Great Lakes would make an oil spill devastating to the environment and to human communities that rely on these systems¹⁻⁵. In order to better protect freshwater environments, there is a need to better understand the full ecological effects of oil spills. Research from the oceans has shown that the wildlife and fisheries can be negatively affected through oil

exposure ¹. In previous spills, microbial communities have been relied upon to assist in cleanup through the breakdown of the released oil. Therefore, proper oil spill response planning and management in freshwater systems must consider the impact of oil spills on freshwater microbial communities. Currently there is a dearth of information regarding the identity of the dominant oil degrading microbial taxa in surface water, with even less known about the microbial community response to oil in the Great Lakes. To address these data gaps, this study undertook a microcosm-based study to examine the impact of crude oil on microbial community structure. The primary hypothesis of this study is that the type of oil is a major influence on the environmental impacts of oil spills. Additionally, we are interested in identifying the bioremediation potential of microbial communities in the Great Lakes and their potential utility to aid in oil spill cleanup.

1.1 Oil Spill/Oil Transport

Today oil is extracted from onshore and offshore reservoirs, then transported all over the world by pipelines, rail, and oil tankers. During the transportation processes the risk of accidental spills is high ⁷. Across the United States alone there are approximately 79,000 miles of pipelines that transport crude oil ⁸. Many of these pipelines cross sensitive freshwater bodies. An oil spill in these water bodies could have devastating impacts on the environment. The proposed Keystone XL Pipeline is a 2,634 mile long pipeline that transports oil from Canada to refineries in Texas, and would cross 14 bodies of freshwater,

making them all at a higher risk of oil contamination ⁹. The Enbridge Company Line 5 oil pipeline sits along the bottom of the Straits of Mackinac transporting oil from Alberta Canada across the Upper Peninsula of Michigan, then through the lower peninsula to Sarnia Ontario (FIGURE 1.1) ³. In addition to the approximately five mile long Straits of Mackinac crossing, Line 5 has an additional 74 water crossings in the state of Michigan ³. In the event that a spill happens from any pipeline that is close to a freshwater body, there must be swift action to stop the spill and then to begin remediation efforts.

FIGURE 1.1: Map showing location of Enbridge Company Line 5 and highlighting location where it crosses the Straits of Mackinac ².

1.2 Comparison of BP, Valdez. and Kalamazoo

There have been a number of catastrophic oil spills in the U.S in both marine and freshwater settings. One of the most notable marine spills was the BP *Deepwater Horizon (DWH)*. During the *Deepwater Horizon* spill, 4.9 million barrels of Macondo light crude oil were spilled into the Gulf of Mexico, making the *Deepwater Horizon* the worst marine spill on record ¹⁰. Previous to *DWH*, the worst marine spill in the U.S. was the *Exxon Valdez* spill in Prince William Sound, Alaska. The *Exxon Valdez* released 260,000 barrels of Alaska North Slope crude oil from an oil tanker ⁶. Response to the *Exxon Valdez* spill was delayed due to weather. This delay allowed more of the oil to end up reaching the beach which resulted in oil coating the rocks and becoming trapped on the beaches. Oil that is beached is often more difficult to clean due the tendency of oil to coat and remain on substrates such as rock and sand as well as the potential to limit biodegradability ⁸⁻¹¹. A major goal of oil spill response teams is to avoid allowing the oil to reach beaches. Limiting beach exposure will aid in clean up and limit the ecological damage to the coastal ecosystems ^{6,13,14}. During the *DWH* oil spill, approximately 1,312 miles of beaches were oiled, which was similar in the amount of shoreline impacted during the *Exxon Valdez* spill ^{6,10,15}. Despite the similar amounts of shoreline impacted, the types of oil released in the *DWH* and the *Exxon Valdez* spill varied greatly in their composition and physical properties, which had an impact on the cleanup efforts.

Another major oil spill occurred in Marshall, MI. The Enbridge Line 6B spill took place in July of 2010 when this pipeline ruptured, and is one of the

largest inland freshwater oil spills in U.S. history. The amount of spilled oil was estimated to be between 20,880 and 23,809 barrels of Diluted Bitumen (DilBit) crude, spilled into Talmadge Creek which is a tributary feeding into the Kalamazoo River ¹⁶⁻¹⁸. This spill oiled approximately 38 miles of the river ^{16,18}. The spill cleanup required 4 years of dredging the river and cost over 1.3 billion dollars to clean up ¹⁸.

While the amount of oil, extent of shoreline impacted, and type of oil spilled was different between the *Deepwater Horizon*, *Exxon Valdez*, and the Kalamazoo River Line 6B spills, all of these spills were catastrophes and required multiple response strategies to remediate the spills as quickly as possible.

1.3 Oil Types

Oil is a complex mixture composed of many distinct chemicals ¹⁹. No two oils are the same and can vary from region to region and depth of the formation ²⁰. A substantial component of crude oil are hydrocarbons, which are any compound that contains only hydrogen and carbon bonds. In addition to hydrocarbons such as alkanes and aromatics, crude oil is also composed of asphaltenes and resins²¹. The ratio of these classes and size of the compounds in these classes can give different oils distinct properties. One way of classifying crude oils is based on the American Petroleum Institute's (API) measurement known as API gravity. API gravity is a measure of the density of oil relative to water²².

Light crude oil has an API gravity that is greater than 31.1° ²². It is a liquid that can flow freely at room temperature. Light crude oils are often what is

refined into diesel and gasoline because they contain higher amounts of compounds needed for gasoline and diesel fuels. In this study, Bakken crude was used as the light crude oil with an API gravity of 40.6°. Bakken crude oil is extracted from the Bakken shale formation in North Dakota²³. The composition of Bakken Crude used in this study is outlined in TABLE 1.1. Heavy crude oils have an API gravity below 22.3°²². Cold Lake Diluted Bitumen crude is extracted from oil sand deposits in Alberta, Canada. Bitumen is a heavily biodegraded form of oil. For transport of bitumen it is often blended with natural gas condensates or light hydrocarbons and diluents giving it the name of DilBit^{17,24}. In this study Cold Lake Diluted Bitumen (DilBit) was used as the heavy crude oil with an API gravity of 21.7°. The makeup of Cold Lake DilBit used in this study is outlined in TABLE 1.1.

TABLE 1.1: Bakken Crude and Cold Lake Bitumen (DilBit) crude, chemical characteristics^{17,23}.

<u>North Dakota Bakken</u>			<u>Cold Lake Diluted Bitumen</u>	
API Gravity (° API)	40.6		API Gravity	21.7
Absolute Density (% weight)	0.04		Absolute Density (kg/m3)	922.8
Total Sulphur (% weight)	0.08		Total Sulphur (mass %)	3.68
Total Alkanes (µg/g of bakken)	5,879		Total Alkanes (µg/g of dilbit)	3,100
Total PAHs (µg/g of bakken)	12,487.7		Total PAHs (µg/g of dilbit)	3,540
BTEX (mg/g)	16.9		BTEX (mg/g)	7.2

1.4 Physical and Biological Remediation

Response strategies to oil spills can take the form of either physical, chemical, or biological approaches. The most immediate actions in spill response is to contain as much of the spilled oil as possible. In the case of the *DWH* in the Gulf of Mexico, responders used skimmers, booms, and *in situ* burning to remove the oil from the system¹⁰. Booms are used to corral the oil slick into a defined location. Skimmers can then be used to recover the oil within the boomed area. Alternatively, the corralled oil can be removed through *in situ* burning. In the case of the Line 6B oil spill in the Kalamazoo River, booms were set up to contain the oil and prevent it from moving further down river, and from reaching shore¹³. These physical methods are effective at addressing the immediate problem of preventing oil from reaching shorelines, but there are limitations to these physical remediation efforts. The effectiveness of many of the physical remediation strategies is limited by weather such as wind speed, and wave height. In addition to weather limitations, recovery of oil using physical means is highly dependent on the thickness of the oil slick^{10,13,16,25}. So as the slick becomes thinner and moves, the efficiency of recovery decreases²⁶. There are some estimates that in the *Deepwater Horizon*, only 25% of the oil was recovered using physical means¹⁰.

Another remediation tactic is the use of chemical dispersants. Chemical dispersants are used to break the oil spill up into smaller droplets¹⁰. In the *DWH* spill 1.8 million U.S. gallons of the chemical dispersant Corexit 9500, was added

to the Gulf of Mexico at the well head and at sites of surface oil ¹⁰. This was done in an attempt to break up the oil into smaller droplets, so it could disperse throughout the water column. Dispersants of this type were designed to be used in aquatic spills and we know that they work in the marine environment, however they are not approved for use in the Great Lakes.

Biological breakdown of oil is an appealing option for its cost and efficacy. Natural microbial communities are capable of using oil as a carbon and energy source ^{14,27-30}. Biological removal of contaminants, such as oil, is called bioremediation. During the *Deepwater Horizon* spill, monitored natural attenuation was put into action as a response strategy for the deep-water oil plume ^{10,27,30}. Monitored natural attenuation is allowing the bacteria to naturally break down the oil without human intervention. Researchers have observed that the microbial community in samples of oil contaminated water, were able to degrade the oil into carbon dioxide and non-toxic daughter products through natural attenuation ^{27,30,31}. Other studies have shown that the addition of nutrients to the system can stimulate bacterial growth and encourage the breakdown of the oil in systems where nutrients are limiting ^{10,11,14,25,29,31,32}. This method is called biostimulation and was used during the *Exxon Valdez* spill to stimulate the natural microbial communities on the beaches for a more rapid oil biodegradation. ¹¹. The third type of bioremediation is bioaugmentation which involves addition of oil degrading bacteria to the environment that has been affected, to accelerate oil biodegradation ¹⁴.

1.5 Constraints on Bioremediation

In marine environments, oil biodegradation occurs from the surface waters to deep ocean waters^{28,30,33–35}. Aerobic oil-degrading metabolisms dominate throughout the water column of the oceans due to the elevated oxygen levels throughout the oxic water column²⁸. Anaerobic oil biodegradation can happen in sediments or some anoxic parts of the water column^{33,36}. Due to the fact that most oils are less dense than water, it is assumed that most of the spilled oil is going to accumulate at the surface of the water, thus allowing for aerobic biodegradation to dominate in oil spills in oxygenated surface waters^{11,14,28,37}. However, at the surface of the water column there can be a lower availability of nutrients^{32,38–40}. These low nutrient levels can limit bioremediation processes, resulting in slowed growth as the bacteria compete for scarce nutrients³².

Several factors can limit biodegradation. The input of oil into a system results in excess carbon. This excess carbon can drive the systems to be limited in Nitrogen and Phosphorus. This limitation of Nitrogen and Phosphorus can constrain microbial growth and oil biodegradation³². In addition to nutrient availability, other factors such as temperature and oxygen availability can impact the composition of the oil-degrading microbial community and the rates of oil biodegradation^{29,32,34,35,41}.

1.6 Bacteria Involved with Bioremediation

A lot of researchers, and environmental engineers have turned to the bioremediation potential of bacteria as the answer to clean up many contaminated

sites ^{10,11,14,18}. An informed use of bioremediation should take into account the identity and capabilities of the microbial community. Research has found 175 prokaryote genera that can use the hydrocarbons in crude oil as a main source of carbon ²⁹. The primary oil-degrading bacteria in the oceans include species such as *Alcanivorax spp.*, *Marinobacter spp.*, *Thalassolituus spp.*, *Cycloclasticus spp.*, and *Oleispira spp* ²⁹. Many of the dominant oil-degrading bacteria in the marine environment are within the class Gammaproteobacteria. More importantly, in many places across the globe, when there is an oil spill in a marine environment, there is an increase in the abundance of Gammaproteobacteria ^{10,30,42}. During the *Deepwater Horizon spill* the microbial community in the deep ocean was dominated by a single microbial group, the Oceanospirillales, which were present in some samples up to 57% of the community ^{30,43}. This response is common to what was observed with *Alcanivorax spp.* in surface water spills, where it was present at very low levels in uncontaminated seawater and then bloomed in response to oil ⁴⁴. Additionally, similar oil degrading taxa, as seen in the *DWH*, are shown to bloom in response to oil in many places outside of the Gulf of Mexico suggesting that the microbes that grow in response to oil appear to be globally distributed ^{6,11,27,29,34,44,45}.

1.7 Freshwater Oil

In the present study, we sought to understand how two types of oil can affect community composition in freshwater. Another study that complements this research looked at the biodegradability of DilBit crude using the Kalamazoo

River as a water collection site¹⁷. Their study was aimed at determining how temperature affected the native oil-degrading bacterial communities' ability to degrade two types of DilBit. Their work indicated that known oil degrading bacteria were enriched in response to oil at both temperature conditions (25 °C and 5 °C). These oil degraders included members of the *Pseudomonas*, *Rhodococcus*, *Hydrogenophaga*, *Parvibaculum*, *Arthrobacter*, *Acidovorax*., with members of the Gammaproteobacteria dominating the enrichment at 25 °C and Betaproteobacteria being the dominant Proteobacterial group at 5 °C. This study also found that there was no difference in the microbial community composition between the two DilBit types, and that temperature was a stronger factor affecting microbial community composition. The two types of oil used in Deshpande *et al.* were both heavy dilbit oil types. Deshpande *et al.* concluded that DilBit can be degraded by autochthonous microbes if the site had recent exposure to DilBit¹⁷.

While, Deshpande *et al.* showed that DilBit could be degraded by enriched consortia, it is unclear how efficient natural microbial communities would be at degrading oil. Another study examined the fate of polycyclic aromatic hydrocarbon PAHs in an estuarine system⁴⁶. PAHs are harmful compounds that can be found in most crudes and have been known to cause health issues for humans. Short term effects of PAH exposure include asthmatics, thrombotic effects, and irritations to the skin or eyes. Long term effects can be more severe such as increased risk of skin, lung, bladder, and gastrointestinal cancers⁴⁶⁻⁴⁹. This study found that PAH mineralization took two to four times longer in the ecosystem never exposed to oil than in an ecosystem that has had chronic

exposure to hydrocarbon contamination. This finding suggests that systems without recent exposure to oil may down hydrocarbons at a slower rate. Since there has not been significant oil exposure in the Great Lakes, these findings may suggest a decreased ability of the natural microbial community to respond to oil, which may have implications for the decision-making process when considering bioremediation in response to a spill in the Great Lakes.

1.8 Study Goals

The overall goal of this study is to characterize the impact of light and heavy crude oil on microbial diversity and microbial community composition of the Great Lakes. A microcosm-approach was used where two oil types (Bakken crude oil and Cold Lake Diluted Bitumen) were amended to water samples. This study hypothesized that due to the distinct chemical composition of these oils, biodegradation of these compounds requires a distinct microbial community for each individual oil type. Previous microcosm based studies in the oceans and in freshwater indicate that upon oil addition the microbial community rapidly changed and became dominated by oil degraders, mostly related to Gammaproteobacteria. Therefore, we set out to test hypothesis 1:

- 1) Upon oil addition, the microbial community in the Great Lakes will change dramatically with a common set of oil degrading taxa being present throughout the Great Lakes.

Previous studies have indicated a preference of certain bacteria for distinct components of oil, with some microbes specialized for alkane degradation

compared to aromatic degradation. Furthermore, it is hypothesized that microbes preferentially degrade the lighter components of oil. Light and heavy crude oils differ greatly in their chemical composition. While previous studies with samples from the Kalamazoo River showed little difference in the microbial community between two types of dilbit. However, these dilbits were quite similar in their composition. Based off these studies done in the in the Kalamazoo river, we sought to test hypothesis two.

- 2) The type of oil is a primary factor shaping the microbial community composition.

The data gained from this research will give insights into the oil biodegradation potential within the Great Lakes and may provide insights that could extend to other major freshwater settings at risk for oil spills. Furthermore, this research will aid in understanding the potential risk associated with transport of these oil types through and near freshwater systems.

2 Methods

2.1 Sampling Location and Water Collection

To test the impact of light and heavy crude oil on freshwater bacterial community composition and the fate of oil in freshwater settings, samples were collected from seven sites in the waters of the Great Lakes surrounding the State of Michigan. Sampling sites included; one site in Lake Michigan, three sites in the Straits of Mackinac, two sites in Lake Superior, and a site in the Keweenaw

Waterway close to Michigan Technological University. Each location was chosen to represent a range of environments including multiple lakes, near-shore, offshore, and inland environments, in order to gain a broad understanding of the fate of oil in the Great Lakes. Three sites were chosen for the Straits of Mackinac due to the risk of oil exposure from the Enbridge Line 5 pipeline, which runs across the lake bed in the Straits of Mackinac.

During mid-August of 2016 water samples were collected aboard the NOAA R/V 5501. Sample MI141 was the first sample collected, followed by Straits 1 (STR1), Straits 2 (STR2), and Straits 3 (STR3). Then Superior 1 (SUP1), Superior 4 (SUP4), and GLRC was collected last. Sample collection occurred over a five-day period. Surface water was collected from each site using a bucket, it was rinsed three times with lake water before sample collection. The water was then put into a 20 L carboy that was sealed and placed inside an insulated ice-chest at approximately the ambient temperature of the water. Upon returning to the lab, the water was stored at 4 °C until microcosm setup. Microcosms were setup within two weeks of sample collection.

FIGURE 2.1. Seven sampling locations in the Great Lakes: MI141, Straits 1, Straits 2, Straits 3, Superior 1, Superior 4, and GLRC. Sampling locations are shown in red.

TABLE 2.1. Individual sampling location sites showing: longitude, latitude, date of sample collection and environmental factors (Temperature, Conductivity, PAR, and Oxygen Saturation) (ND=No Data).

SITE	LAT	LONG	Date	Temp	Conductivity	PAR	OxySat
MI141M	44° 44.180' N	86° 43.308' W	8/9/16	23.28	278.07	3.19E+02	8.506
Straits_1	45°50.46' N	85° 6.40' W	8/10/16	22.59	269.99	7.31E+02	8.619
Straits_2	45°48.4' N	84°44.52'W	8/10/16	23.72	274.03	5.21E+02	8.431
Straits_3	45°42.31' N	84°31.26'W	8/10/16	22.73	254.38	8.77E+02	8.159
Superior_1	46°40.21'N	84°49.45'W	8/11/16	20.94	94.088	6.84E+02	8.904
Superior_4	46°54.57'N	86°35.85'W	8/13/16	20.94	94.088	6.84E+02	8.904
GLRC	47°7' 14.0556" N 88°	88° 32' 43.53" W	8/13/16	ND	ND	ND	ND

2.2 Microcosm Setup

In-lab microcosms were used to simulate oil exposure. Microcosms were set up with the following conditions for each location: Control (no oil), Bakken crude oil (light oil), and Cold Lake Diluted Bitumen (heavy crude oil) (FIGURE 2.2). Each condition was set up in triplicate and incubated at room temperature in the dark. Sacrificial sampling of microcosm bottles occurred one time per week for five weeks. To all microcosms 100 ml of collected water was used. In the oil-amended bottles, oil was added to a final concentration of 25 ppm of oil. Triplicate bottles of each condition were set up for microbial community analysis and an additional set of triplicate bottles were set up for the oil-amended conditions for hydrocarbon analysis. Killed controls were set up for each location

and oil type. Killed control bottles contained autoclaved water from that location amended with oil to a concentration of 25 ppm of oil. The killed controls were treated exactly as the experimental conditions and harvested at the last time point for hydrocarbon analysis. These killed controls were used to determine the amount of abiotic loss of hydrocarbons during the incubation.

FIGURE 2.2. Microcosm bottle set up for each sample location. There were six time points separated by a week. The first T0 was an initial sample collected at the time when the microcosms were set up, then T1, T2, T3, T4, T5 were collected each week. Each time point had triplicate control microcosms (shown in blue), Bakken amendment microcosm set up in triplicate (shown in green), and triplicate DilBit amended microcosms (shown in red).

2.3 Water Filtration/DNA Extraction

Once per week microcosms for microbial community analysis were sacrificially sampled by filtering the 100 ml microcosm through a 0.2 μ m pore size 47 mm diameter PES filter on a glass vacuum filtration apparatus. The filter was then stored in a 2 ml Eppendorf tube and frozen at -80 °C until sample processing. Filters were cut in half using flame sterilized scissors. One half of the filter was used for DNA extractions and the other half was stored at -80 °C as an archived sample. DNA extraction of filter halves were done using a Modified Miller Method ⁵⁰. While it is possible that some bacterial species may be present

in the oil used in these microcosms (ref), we did not measure the microbial community present in the oil itself. We expect that the impact of microbes present in the oil would be minimal due to the higher abundances of microbes present in the surface water.

2.4 16S PCR and Sequencing

The Straits 1, 2 and 3 samples were sequenced at the EPA Office of Research and Development. The samples sequenced at the EPA were sequenced using the V4 region of the 16S rRNA using primers 515F and 806R as described in Kapoor et al ⁵⁰. The other samples (MI141, Superior 1 and 4, and GLRC) were sequenced at Michigan Technological University. The samples sequenced at Michigan Tech amplified the V4-V5 region of the bacterial 16S rRNA gene using 1 µl of each of the 20 µM PCR primers (515F-Y and 926R)⁵¹, 1 µl of DNA, 25 µl of ThermoFisher Phusion Master Mix, and 22 µl of autoclaved ultra-pure water. The thermocycler protocol was 95°C for 3 minutes, 25 cycles of (95°C for 30 seconds, 55°C for 30 seconds, 72°C for 30 seconds) then 72°C for 5 minutes and ending with a hold at 4°C ⁵¹. The amplified PCR products were cleaned using AxyPrep PCR Clean up magnetic beads (Axigen, Tewksbury, MA). Each sample was indexed using an indexing PCR with primers that contained the Illumina sequencing adapters and a 12 base pair index sequence specific for each sample. The indexing PCR contained 5 µl of the purified 16S rRNA product from the previous PCR and was amplified using ThermoFisher Phusion Master Mix. The PCR was performed with the following conditions for 8 cycle PCR: 95°C for 30

seconds, 55°C for 30 seconds, 72°C for 30 seconds ⁵¹. Sequencing was done using an Illumina MiSeq using a v3 2x300 cycle kit following the Illumina protocol for amplicon sequencing.

2.5 Sequencing Analysis Using QIIME1

Sequencing reads from EPA and in-house data sets were analyzed using the QIIME1 pipeline ⁵². For both sequencing runs, paired end forward and reverse reads were assembled using fastq-join ⁵³. Sequencing reads were demultiplexed and quality filtered using the split_libraries_fastq.py command from Qiime1. Quality filtering was performed to remove reads with quality scores below 20. A mapping file was then made using the Qiime_Map.txt. Chimeric sequences were identified using VSEARCH ⁵⁴. The chimeric sequences were then filtered. Quality filtered data from the two sequencing runs were then combined and trimmed to ensure similar length of reads. Operational Taxonomic Units (OTUs) were picked using uclust ⁵⁴ as implemented in the pick_open_reference_otus.py command in QIIME using a percent similarity cutoff of 97%. Reference-based OTU picking was done against SILVA release 128 ⁵⁵. Taxonomy assignments were performed using the rdp classifier ⁵⁶ trained against the SILVA release 128 reference database.

2.6 Analysis in R

In R ⁵⁷ the package phyloseq was used for the diversity and microbial community analysis ⁵⁸. The OTU table from Qiime was imported into R and used for the data analysis. Additionally, a metadata file was also included that contained important information about each sample including the; treatment, timepoint, and other import data for those samples. In the dataset the initial filter samples taken at time of water collection were given the designation of T0. Samples with less than 1000 reads were removed from the analysis and the data was then rarefied using the phyloseq command `rarefy_even_depth` with the seed of 12 to rarefy the OTU table to the minimum number of reads found in the remaining samples (1140 reads).

2.7 Alpha Diversity Analysis

Shannon diversity was calculated from non-rarified OTU tables using phyloseq. To test if there was a significant difference of the variance between the Shannon diversity of the microbial communities between sites, a one-way analysis of variance (ANOVA) was performed. A Tukey Honest Significant Difference post hoc test was used to identify between which sites there was a significant difference in Shannon diversity. Similarly, ANOVA and Tukey HSD tests were performed to determine if there was a significant difference in the richness of oil-amended samples compared to controls, and if there were significant differences in the Shannon diversity between Bakken and diluted bitumen-amended conditions.

2.8 Analysis of Changes in Microbial Community Composition

To show there were differences in the microbial community composition between sites and treatments, non-metric multidimensional scaling plots (NMDS) were constructed. The NMDS plots were made using a Bray Curtis dissimilarity matrix constructed from the rarefied OTU table. To test the significance of the response in the microbial community composition, permutational multivariate analysis of variance (PERMANOVA) statistics were performed as implemented in the vegan package in R ^{59,60}. PERMANOVA analysis was performed comparing the microbial community composition between sites to test the hypothesis that there is a different community composition at each site. Testing the hypothesis that there is distinct community composition in oil-amended conditions relative to the controls, was also performed using PERMANOVA analysis between control and oil-amended conditions. To test the hypothesis that there was a distinct community composition between the two oil-types, PERMANOVA analysis was used comparing the microbial community composition between Bakken-amended samples compared to DilBit-amended samples.

2.9 Differential Abundance Analysis to Identify Responsive OTUs.

To identify which OTUs were responsive to oil, differential abundance analysis was performed using DESeq ⁶¹. DESeq was performed using the non-rarefied OTU table. DESeq estimates the variance-mean dependence in count data from high-throughput sequencing and tests for differential abundance based

on a model using the negative binomial distribution and Wald Test ⁶¹. P-values for determination of significance were corrected using FDR correction for multiple comparisons. OTUs were considered to be differentially abundant if they had an adjusted p-value of less than 0.01 and a log2 fold change of greater than 2. The results from DESeq were visualized using volcano plots constructed using ggplot2 ⁶².

2.10 Hydrocarbon Extraction

Bottles set up for hydrocarbon extractions were stored in the freezer at -20 °C at the time of sampling for later hydrocarbon extraction. During freezing of bottles many became cracked or broken. All glassware used in hydrocarbon extractions was combusted in a muffle oven at 450 °C for four hours. Bottles were thawed in sets of six. Bottles were thawed in 400 ml beakers allowing the microcosms to thaw without loss of liquid. Once thawed 1,3,5-Trichlorobenzene (TCB) was used as an internal standard. 30 µL of TCB was added to each bottle to allow for correction of variation in recoveries during the extraction process. 10 ml of Dichloromethane (DCM) was added to the bottles and they were shaken by hand. The water/DCM mixture was then poured using a 250 ml separatory funnel. An additional 20 ml of DCM was added to the bottles and they were shaken and then poured into the separatory funnel. Following this, three separate additions of 30 ml DCM were added to the bottle, shaken and poured into separatory funnel after each addition. The 250 ml separatory funnel was shaken vigorously for ~20 seconds. The aqueous and organic phases were then allowed to separate for

approximately one hour. The lower organic phase was separated from the funnel into two 60ml VOA vials. The extracts were then concentrated under Nitrogen gas flow while partially submerged in a hot water bath that was kept at ~35 °C. Once the volume of the extract was concentrated to approximately 5 ml, the organic phase was pipetted using glass pipettes into 15 ml glass collection tube. The VOA vials were rinsed inside with 3 ml of DCM and the rinse was added to the collection tube. The collection tube was concentrated down to near evaporation using the Nitrogen gas/warm water bath. The nearly evaporated mixture from the collection tube was transferred by glass syringe to 1.5 ml Gas Chromatograph GC vials. The collection tube was rinsed inside using 300 μ l DCM which was then added to the GC vial. The final volume of the hydrocarbon extraction solution was adjusted to 1ml using DCM.

2.11 Gas Chromatography Analysis

Hydrocarbon extracted samples were attempted to be analyzed using a Gas Chromatograph. The total petroleum values that were determined were highly variable, making the data inconclusive. This was in part due to the wide variance in the extraction efficiencies as determined by TCB recover, which ranged from 2% to 141%. Due to the high variance and inconsistent recoveries, the data was not included.

3 Results

3.1 Differences in Microbial Communities Between Sites

To characterize how the diversity of microbial communities varied among geographical locations, we calculated the Shannon diversity of the OTUs at each site and then plotted the values as a box plot separated by site (FIGURE 3.1).

FIGURE 3.1. Shannon Index of microbial communities Alpha Diversity displayed as box plot. The colors represent the individual sites in the data set. Each point is colored as is the box plot for the sites. Outliers and dots along each line make up sample points and box region represent the upper and lower quartile of that sites samples.

This study observed that there were large differences in the Shannon diversity between the sites within the Great Lakes. The means of the Shannon diversity were calculated for each site and found that the site with the highest mean Shannon diversity was the GLRC at 4.715 ± 0.6996 . The Shannon diversity

of the GLRC difference between its position on the plot and the other site's Shannon Diversity. MI141 and the two Superior samples had a similar Shannon diversity, but they were also distinct from the other sites. SUP1 had a mean Shannon diversity of 4.323 ± 0.3082 . MI14 had a mean Shannon diversity of 4.262 ± 0.2469 . SUP4 had a mean Shannon diversity of 4.212 ± 0.1972 . Samples from the Straits sites had similar Shannon Diversity to other samples from Straits that was distinct from and lower than the rest of the sites. STR1 had a Shannon diversity mean of 3.736 ± 0.2789 . STR2 had a Shannon diversity mean of 3.829 ± 0.1865 . The site with the lowest Shannon diversity mean was STR3 with a Shannon diversity mean of 3.65 ± 0.3656 . There was a significant difference in the Shannon diversity overall as determined by ANOVA (p-value $< 2e^{-16}$, F-stat = 83.2, degrees of freedom = 6). When broken down by location there was significance found between most sites, determined by Tukey Post Hoc analysis (TABLE 3.1). There was no significant difference between MI141 and either of the Lake Superior sites, or between SUP1 and SUP4. Additionally, there was no significant difference between STR1 and STR2, and STR1 and STR3. (TABLE 3.1)

TABLE 3.1. Tukey HSD post hoc test results displaying adjusted p-value. Bolded adjusted p-values indicate a significant difference.

	GLRC	MI141	STR1	STR2	STR3	SUP1	SUP4
GLRC		<0.001	<0.001	<0.001	<0.001	<0.003	<0.001
MI141			<0.001	<0.001	<0.001	0.948	0.979
STR1				0.502	0.582	<0.001	<0.001
STR2					0.004	<0.001	<0.001
STR3						<0.001	<0.001
SUP1							0.441
SUP4							

An NMDS was constructed to visualize the microbial community composition in the controls for each site (FIGURE 3.2). The lowest stress value for this plot was 0.1282 after 20 iterations and no convergence was reached (FIGURE 3.2). The NMDS plot showing only the control microcosms, shows that samples from the same site clustered closely together and distinctly from other sites (FIGURE 3.2). This suggests that there is a distinct microbial community composition between sites.

The GLRC samples are clustered far away from other sites indicating a distinct microbial community composition in this inland system. The clustering of the other sites suggests that community composition of individual communities can be broken down by lake. Lake Superior includes samples, Superior1, and Superior 4. All these samples are clustering as distinct from the Straits samples, and MI141. Superior 1 and 4 community samples clustered together meaning that their community compositions were similar to each other but distinct from the GLRC's community composition. There is little overlap between MI141 in open water Lake Michigan and the Straits samples suggesting a difference in their community structures.

Permutational multivariate analysis of variance (PERMANOVA) was used to test the hypothesis that there was a significantly different community composition between sites (TABLE 3.2). PERMANOVA analysis indicated that

there was a significant difference in the community composition between each site.

TABLE 3.2. Table matrix of r-squared and statistical significance by p-value, comparing microbial community samples by site against each other. The top right triangle shows the p value adjusted using FDR comparisons showing significance as bolded values, while the lower left triangle shows the R^2 value.

	Straits 1	Straits 2	Straits 3	Superior 1	Superior 4	MI141	GLRC
Straits 1		0.002	0.002	0.002	0.002	0.002	0.002
Straits 2	0.078		0.002	0.002	0.002	0.002	0.002
Straits 3	0.066	0.071		0.002	0.002	0.002	0.002
Superior 1	0.346	0.331	0.380		0.002	0.002	0.002
Superior 4	0.329	0.332	0.375	0.153		0.002	0.002
MI141	0.277	0.265	0.329	0.210	0.250		0.002
GLRC	0.344	0.325	0.372	0.243	0.312	0.308	

NMDS SITE MICROBIAL COMMUNITIES: CONTROLS

FIGURE 3.2. NMDS showing microbial communities from the control samples only. Sites are distinguished by color. Clustering is seen by each independent site, by lakes, by offshore and Straits samples.

3.2 Community Response to Oil Exposure

To test the hypothesis that oil addition would influence the diversity of the microbial community or the number of distinct microbes present in these samples,

we examined the Shannon diversity of the oil amended samples in each site compared to the the control samples and tested for a significant difference (FIGURE 3.3)(TABLE 3.3). FIGURE 3.3 shows the Shannon diversity of the microbial communities for each location comparing each treatment. Diversity remained fairly constant in the control versus oil treatment, with the only significant difference in Shannon diversity between control and oil conditions occurring in the GLRC, the STR2, and STR3 samples. In the GLRC there was a significant decrease in Shannon diversity in the Bakken-amended samples (mean 4.353) compared to the controls (mean 4.943) (ANOVA p-value = 0.00837, F-stat = 6.217, degrees of freedom (dof) = 2, and Tukey HSD adjusted p-value =0.008178), and between the Bakken (mean 4.353) and DilBit (mean 4.815) amended samples (ANOVA p-value = 0.00837, F-stat = 6.217, degrees of freedom (dof) = 2, and Tukey HSD adjusted p-value =0.046772). However, there was no difference in the Shannon diversity when comparing the DilBit-amended samples with the controls. In the STR1 there was a significant decrease in Shannon diversity in the Bakken-amended samples (mean 3.942) against the DilBit-amended samples (mean 3.724) (ANOVA p-value = 0.0118, F-stat = 4.85, degrees of freedom (dof) = 2 ,and Tukey HSD adjusted p-value =0.008533). However, there was no difference in the Shannon diversity of the DilBit-amended samples against the controls, or the Bakken-amended samples against the controls. In the STR3 there was a significant decrease in Shannon diversity in the Bakken-amended samples (mean 3.789) against the DilBit-amended samples (mean 3.568) (ANOVA p-value = 0.0343 and Tukey HSD adjusted p-value

=0.041196). However, there was no difference in the Shannon diversity of the DilBit-amended samples against the controls, or the Bakken-amended samples against the controls.

FIGURE 3.3. Alpha diversity of the microbial community as a function of oil treatment samples and controls. Each facet and color are an individual site which shows the Shannon diversity as a box plot separated by treatment. Color of points represent the sites from which they were derived, and the shape of the points represent the time of sample collection.

TABLE 3.3. ANOVA results of Shannon Diversity comparison of treatments and controls. Bolded values represent significant difference.

	dof	F value	P value
SUP1	2	1.102	0.344
SUP4	2	2.594	0.0883
GLRC	2	6.217	0.00837
MI141	2	1.816	0.181
STR1	2	0.154	0.858
STR2	2	4.85	0.0118
STR3	2	3.61	0.0343

TABLE 3.4. Tukey HSD analysis for sites indicated as having significant difference based on ANOVA analysis. Bolded values show significance based off adjusted p-value.

	padj
GLRC	
CONTROL-BAKKEN	0.008178
CONTROL-BITUMEN	0.745924
BAKKEN-BITUMEN	0.046772
STR2	
CONTROL-BAKKEN	0.199437
CONTROL-BITUMEN	0.367221
BAKKEN-BITUMEN	0.008533
STR3	
CONTROL-BAKKEN	0.093213
CONTROL-BITUMEN	0.927695
BAKKEN-BITUMEN	0.041196

The NMDS plot in FIGURE 3.4 (lowest stress value 0.1282 after 20 iterations and no convergence) shows the microbial communities of the controls and the oil-amended samples for all of the seven sites. In the plot, samples from each region clustered separately from other regions, in a similar fashion to the control samples shown in FIGURE 3.2. For the most part, samples from each site

clustered separately from other sites. However, all three samples from Straits appear to have a similar community composition and were intermingled in a Straits cluster. When adding the oil samples on the NMDS plot, it was seen that oil-amended samples still cluster with the controls from their site. This indicates that while there is a response in the microbial community composition in the Great Lakes, that this response is not so strong that all of the oil-amended samples converge on a similar microbial community composition. PERMANOVA analysis indicated that there was a significant difference in the microbial community composition between control and oil-amended samples in all of the locations (TABLE 3.5). This suggests that while there were distinct community members found in the oil-amended conditions, the clustering on the NMDS suggests that the oil amended samples maintained some similarity to the control conditions.

NMDS SITE MICROBIAL COMMUNITIES: All SAMPLES

FIGURE 3.4. NMDS of microbial community composition in response to oil exposure. Sites are displayed by color and oil-amendment is displayed by shape. Clustering is representation of similar microbial community composition.

TABLE 3.5. Table showing the Pairwise PERMANOVA significance comparison of each oil type in comparison with the control. Bolded values represent a significant difference based on adjusted FDR p-value.

		R2	pvalFDR
GLRC			
BAKKEN	CONTROL	0.189953361	0.003
BITUMEN	CONTROL	0.166718728	0.006
MI141			
BAKKEN	CONTROL	0.098863847	0.003
BITUMEN	CONTROL	0.087039273	0.018
STRAITS			
BAKKEN	CONTROL	0.108998786	0.002
BITUMEN	CONTROL	0.097068693	0.002
Superior			
BAKKEN	CONTROL	0.058491943	0.006
BITUMEN	CONTROL	0.040600097	0.015

Due to the significant differences in the overall community composition between sites, the shifts in microbial community composition in response to oil was more closely examined in each individual site. We used DESeq to identify the OTUs that were differentially abundant in the oil-amended samples. We considered an OTU to be enriched if the adjusted p-value was less than 0.01 and the log2 fold change was greater than 2.

TABLE 3.6. Table of top 5 enriched OTUs in oil-amended conditions for each different site location. Samples from similar regions were combined.

Top 5 Enriched Taxa in Oil							
	log2FoldChange	padj	Domain	Phylum	Class	Order	Family
All							
GDIP01031912.2.1304	-4.825034181	1.15E-09	Bacteria	Proteobacteria	Betaproteobacteria	Rhodocyclales	Rhodocyclaceae
HM773508.1.1259	-4.813390899	1.01E-80	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae
New.ReferenceOTU193	-4.76857179	2.89E-05	Bacteria	Proteobacteria	Alphaproteobacteria	Alphaproteobacteria Incertae Sedis	Unknown Family
HQ216358.1.1245	-4.437041976	0.000114849	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae
EU803640.1.1257	-4.351315443	6.79E-12	Bacteria	Proteobacteria	Alphaproteobacteria	SAR11 clade	LD12 freshwater group
Straits Samples							
New.ReferenceOTU193	-6.119622319	6.25E-10	Bacteria	Proteobacteria	Alphaproteobacteria	Alphaproteobacteria Incertae Sedis	Unknown Family
HM773508.1.1259	-4.872466705	1.36E-70	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae
EU803640.1.1257	-4.73201764	1.71E-18	Bacteria	Proteobacteria	Alphaproteobacteria	SAR11 clade	LD12 freshwater group
HJ352592.1.1456	-4.614116279	5.93E-16	Bacteria	Proteobacteria	Betaproteobacteria	Methylophilales	Methylophilaceae
GU208284.1.1442	-4.603554241	3.09E-26	Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae
Superior							
JF729004.1.1504	-7.933076096	5.05E-07	Bacteria	Proteobacteria	Betaproteobacteria	Rhodocyclales	Rhodocyclaceae
EF520602.1.1486	-6.458186995	0.002190011	Bacteria	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	uncultured
New.ReferenceOTU354	-4.88550351	0.000867456	Bacteria	NA	NA	NA	NA
HQ216358.1.1245	-4.707141154	0.000401003	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae
LN560506.1.1381	-4.556961251	0.001451539	Bacteria	NA	NA	NA	NA
MI141							
FM200864.1.1352	-5.138801692	6.82E-05	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae
HM773508.1.1259	-4.818102468	1.96E-10	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae
KJ572477.1.1355	-4.69988192	6.82E-05	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	NA
KC437364.1.1517	-4.299821669	0.003045147	Bacteria	Proteobacteria	Deltaproteobacteria	Myxococcales	mle1-27
HM129128.1.1449	-4.011868582	1.05E-06	Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae
GLRC							
AM935018.1.1375	-25.6489331	9.70E-34	Bacteria	Proteobacteria	Gammaproteobacteria	Cellvibrionales	Porticoccaceae
New.CleanUp.ReferenceOTU16027	-22.45919799	6.21E-12	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Nevskiaceae
KP636071.1.1500	-21.91179197	1.98E-11	Bacteria	Proteobacteria	Betaproteobacteria	Rhodocyclales	Rhodocyclaceae
KP398531.1.1350	-10.19755851	3.49E-14	Bacteria	Proteobacteria	Gammaproteobacteria	PYR10d3	NA
GU291353.1.1403	-9.526184676	2.93E-18	Bacteria	Proteobacteria	Betaproteobacteria	Rhodocyclales	Rhodocyclaceae

TABLE 3.7. Table of top 5 enriched OTUs in the control treatments for each different site location. Samples from the same region were combined for this comparison.

Top 5 Enriched Taxa in the Controls							
	log2FoldChange	padj	Domain	Phylum	Class	Order	Family
All							
GQ249367.1.1507	4.916518199	9.13E-14	Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae
AF418976.1.1456	4.409444746	3.70E-05	Bacteria	Cyanobacteria	Chloroplast	uncultured diatom	uncultured diatom
EU803264.1.1258	4.407054997	2.84E-28	Bacteria	Actinobacteria	Actinobacteria	Frankiales	Sporichthyaceae
New.CleanUp.ReferenceOTU42329	3.949906563	1.16E-16	Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae
New.CleanUp.ReferenceOTU1670	3.908584356	5.68E-10	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae
Straits Samples							
GQ249367.1.1507	5.528800283	1.37E-14	Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae
JQ941774.1.1392	5.295182327	0.000627321	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Oceanospirillaceae
AB435574.1.1459	4.698372533	4.28E-08	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae
New.CleanUp.ReferenceOTU53641	4.571374674	1.27E-07	Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae
EU803264.1.1258	4.451886727	1.57E-20	Bacteria	Actinobacteria	Actinobacteria	Frankiales	Sporichthyaceae
Superior							
AF418976.1.1456	5.403814959	0.000183422	Bacteria	Cyanobacteria	Chloroplast	uncultured diatom	uncultured diatom
New.ReferenceOTU134	4.22489184	0.000183422	Bacteria	Proteobacteria	NA	NA	NA
New.CleanUp.ReferenceOTU20005	4.084700534	0.000401003	Bacteria	Bacteroidetes	Cytophagia	Cytophagales	NA
KC253350.1.1487	3.6767058	0.000670166	Bacteria	NA	NA	NA	NA
New.CleanUp.ReferenceOTU23970	3.528761152	0.000326019	Bacteria	Cyanobacteria	Chloroplast	NA	NA
MI141							
none							
GLRC							
FJ612352.1.1449	23.99376669	1.13E-18	Bacteria	Cyanobacteria	Chloroplast	NA	NA
JQ195581.1.1311	23.36844955	6.06E-13	Bacteria	Cyanobacteria	Chloroplast	Guillardia theta	Guillardia theta
FR720644.1.1449	23.13379658	8.91E-13	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Nevskiaceae
New.ReferenceOTU352	22.18316241	8.92E-12	Bacteria	Cyanobacteria	Chloroplast	NA	NA
FN668181.1.1494	10.64649583	1.42E-16	Bacteria	Bacteroidetes	Flavobacteriia	Flavobacteriales	Cryomorphaceae

When identifying differentially abundant OTUs in oil-amended samples compared the controls across all sites, it was found that 51 OTUs were enriched in the presence of oil. The top 5 most enriched OTUs in oil-amended samples across all sites, and by each individual site are outlined in TABLE 3.6 (Complete list of OTUs enriched in the oil-amended conditions is available as Supplemental Data).

In the controls 96 OTUs were found to be enriched. The top 5 OTUs that were at a higher enrichment in the control samples across all sites, and by each individual site are outlined in Table 3.7. (Complete list of OTUs enriched in the control treatments relative to the controls is available as Supplemental Data).

Differentially abundant OTUs were plotted as a volcano plot comparing significance of enrichment and fold change for each OTU (FIGURE 3.5). Many more OTUs were enriched in the control treatments. However, in the presence of oil many of the enriched OTUs were more highly enriched ($\log_2$ fold change of >2) compared to the controls. These differentially abundant OTUs represent the set of OTUs that respond to oil across all of the Great Lakes samples. The OTUs that are more highly abundant in the oil-amended treatments are hypothesized to be the oil-degrading organisms that are able to grow and increase in abundance in response to oil. Conversely, the OTUs that are more abundant in the controls relative to the oil-amended treatments would represent the oil-sensitive OTUs that potentially could be inhibited by oil. These differentially abundant OTUs represent the set of organisms that are impacted by oil across all of the Great Lakes samples in this study. What we ended up finding was that the majority of OTUs enriched in oil amended samples were from the Alphaproteobacteria and

Betaproteobacteria (TABLE 3.6). Across all the sites the most enriched families were Sphingomonadaceae, LD12 freshwater group, and Rhodocyclaceae. The OTUs found to be highly enriched in the controls were classified as chloroplast sequences from an uncharacterized family of diatoms, Actinobacteria, and a Gammaproteobacteria from the Pseudomonadales order (TABLE 3.7).

FIGURE 3.5. Volcano plot of enriched OTUs in oil vs controls using all the sites. The enriched OTUs in the presence of oils are represented as the points on the negative side of the plot while the control OTUs are represented on the positive side of the plot. Enriched OTUs are shown in orange.

To test if there were a different set of OTUs that were impacted at each region, we determined the differentially abundant OTUs between oil and control conditions at each region. For samples from Lake Superior it was found that 21 OTUs were enriched in the presence of oil, while there were 17 OTUs enriched in the controls relative to oil. These differences are plotted as a volcano plot

(FIGURE 3.6). This plot shows that fewer OTUs were enriched in the controls, and many of the enriched OTUs in the presence of oil were more highly enriched (log2fold change of >5 enrichment) as compared to the control samples. The top 5 enriched in oil were from the families Rhodocyclales, and Sphingomonadaceae (TABLE 3.6). The top 5 enriched in the controls were from the class chloroplast, and cytophagia (TABLE 3.7).

FIGURE 3.6. Volcano plot of enriched OTUs in oil vs controls in the Superior sites. The enriched OTUs in the presence of oils are represented as the points on the negative side of the plot while the control OTUs are represented on the positive side of the plot. Enriched OTUs are shown in orange.

For samples from the Straits region, it was found that 58 OTUs were enriched in the presence of oil, while there were 126 OTUs enriched in the controls relative to oil. These differences are plotted as a volcano plot (FIGURE 3.7). This plot shows that many more OTUs were enriched in the controls than

enriched in the oil amendments. There were few OTUs that were highly enriched (log2fold change of >5 enrichment) in both the controls and oil samples. The top 5 enriched in oil were from the families Sphingomonadaceae, LD12 freshwater group, Methylophilaceae, and Comamonadaceae (TABLE 3.6). The top 5 enriched in the controls were from the families Pseudohongiella, Perlucidibaca, Comamonas, and hgcI clade (TABLE 3.7).

FIGURE 3.7. Volcano plot of enriched OTUs in oil vs controls in the Straits sites. The enriched OTUs in the presence of oils are represented as the points on the negative side of the plot while the control OTUs are represented on the positive side of the plot. Enriched OTUs are shown in orange.

For the MI141 samples, it was found that 13 OTUs were enriched in the presence of oil, while there were no OTUs enriched in the controls relative to oil. These differences are plotted as a volcano plot (FIGURE 3.8). This plot shows that the differences in the microbial community composition observed was due

exclusively to OTUs that were enriched in the oil amendments. The OTUs enriched in the oil-amended conditions had only a few highly enriched (log2fold change of >5 enrichment) in the oil samples to similar levels to oil-enriched OTUs from other sites. The top 5 enriched in oil were from the families Sphingomonadaceae, Comamonadaceae, and mle1-27 (TABLE 3.6). There were none enriched in the controls.

FIGURE 3.8. Volcano plot of enriched OTUs in oil vs controls in the MI141 site. The enriched OTUs in the presence of oils are represented as the points on the negative side of the plot while the control OTUs are represented on the positive side of the plot. Enriched OTUs are shown in orange.

When determining differentially abundant OTUs from the GLRC, it was found that 68 OTUs were enriched in the presence of oil, while there were 177 OTUs enriched in the controls relative to oil. These differences are plotted as a volcano plot (FIGURE 3.9). Similar to many of the regions, there were more

OTUs enriched in the controls than enriched in the oil amendments. There were OTUs that were highly enriched (log2fold change of >10 enrichment) in both the controls and oil samples. Of all the sites examined, the GLRC had the most OTUs that were differentially enriched, and those OTUs had the highest fold change. The top 5 enriched in oil were from the families Porticoccaceae, Nevskiaceae, and Rhodocyclaceae (TABLE 3.6). The top 5 enriched in the controls were from the families Guillardia theta, Alkanibacter, Fluvicola (TABLE 3.7).

FIGURE 3.9. Volcano plot of enriched OTUs in oil vs controls in the GLRC site. The enriched OTUs in the presence of oils are represented as the points on the negative side of the plot while the control OTUs are represented on the positive side of the plot. Enriched OTUs are shown in orange.

3.3 Bakken vs DilBit

To test the hypothesis that the type of oil is a primary factor shaping the microbial community composition, Pairwise PERMANOVA were used to measure if there was a significant difference between the microbial communities in the Bakken and DilBit samples at each site. Significance was determined by adjusted p value using the FDR method of adjustment (TABLE 3.8).

TABLE 3.8. Bakken vs DilBit significance table. This table shows the comparison of Bakken and DilBit. Control shows significant difference based on FDR adjusted p value. Bolded values are not statistically different.

		R2	pvalFDR
GLRC			
BAKKEN	BITUMEN	0.156523669	0.009
MI141			
BAKKEN	BITUMEN	0.073032072	0.043
STRAITS			
BAKKEN	BITUMEN	0.027570479	0.002
Superior			
BAKKEN	BITUMEN	0.02868799	0.102

The results of the PERMANOVA calculations showed that the Superior sites were the only region that did not have a significantly distinct microbial community composition between Bakken-amended and DilBit-amended conditions. In all of the other regions, there was a significant difference in microbial community composition between Bakken-amended samples and DilBit-amended samples. To show the difference in microbial community composition between oil types, an NMDS plot was made for each region to display the

difference in community compositions for oil-amendments (Bakken and Diluted Bitumen) and their controls (FIGURE 3.10).

NMDS BAKKEN DILBIT COMMUNITY COMPOSITION

FIGURE 3.10. NMDS plots separated by region. Plots show the NMDS clustering of the microbial communities by oil treatment (Bakken, DilBit, Control). Each point represents an individual sample. Treatments are colored with Bakken being gray, Diluted Bitumen being blue and controls being cyan. Stress values after 20 runs for each plot were as follows: A) Straits NMDS=0.2042 no convergence, B) Superior NMDS= 0.1804 no convergence, C) MI141 NMDS=0.1877 convergence reached, D) GLRC NMDS=0.1161 no convergence

Differential abundance analysis was also performed to identify OTUs that were differentially abundant between Bakken and DilBit samples. The analysis was performed across the whole data set to see if there was a difference in the number and identity of OTUs that were enriched in the Bakken-amendments or the DilBit-amendments. It was found that there were more OTUs enriched in the Bakken-amended samples than in the DilBit-amended samples (13 OTUs enriched in Bakken and 2 in DilBit), and that the Bakken enriched OTUs were more enriched than the OTUs enriched in the DilBit (FIGURE 3.11).

FIGURE 3.11. Volcano plot of enriched OTUs in Bakken-amendments vs DilBit-amendments using all sites. The enriched OTUs in the presence of Bakken are represented as the points on the right or positive side of the plot while the DilBit enriched OTUs are represented on the left or negative side of the plot.

This study also looked specifically at the OTU enrichments between the Bakken and DilBit within each region. The top 5 OTUs that were enriched in Bakken amended samples across all sites, and by each individual region are outlined in TABLE 3.12. Within the top 5 OTUs enriched in Bakken, this study found bacteria from the alpha- and beta-proteobacteria class, and mainly the Comamonadaceae, Burkholderiaceae, and LD12 freshwater group families being the most enriched.

The top 5 OTUs that were enriched in the DilBit samples across all sites, and by each individual region are outlined in TABLE 3.13. . Within the top 5 OTUs enriched in DilBit, this study found bacteria from the alpha- and beta-proteobacteria class. The top 5 families that were found were Comamonadaceae, Sphingomonadaceae, LD12 freshwater group, and Oxalobacteraceae.

TABLE 3.12. This plot shows the top 5 OTUs that were most highly enriched in the Bakken amended samples determined by DESeq2 analysis.

Top 5 OTUs enriched in Bakken amendments							
SampleID	log2FoldChange	padj	Phylum	Class	Order	Family	Genus
All Sites							
JN038730.1.1487	-7.267517002	4.83E-12	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Aquabacterium
JX458424.1.1460	-6.885120619	9.00E-12	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Aquabacterium
FJ375396.1.1642	-5.246257406	4.22E-12	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	NA
EU801674.1.1424	-3.833837931	0.004553027	Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Polynucleobacter
EU800187.1.1330	-3.758914796	2.83E-07	Proteobacteria	Alphaproteobacteria	SAR11 clade	LD12 freshwater group	uncultured bacterium
MI141							
Superior							
GU305726.1.1441	-3.335148594	0.006806145	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Zymomonas
AY752098.1.1391	-2.90466031	0.002882075	Actinobacteria	Actinobacteria	Frankiales	Sporichthyaceae	hgcI clade
Straits							
JX458424.1.1460	-7.999918418	1.59E-10	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Aquabacterium
JN038730.1.1487	-7.453487716	1.14E-14	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Aquabacterium
EU802044.1.1501	-5.859737821	0.002830436	Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Polynucleobacter
FJ375396.1.1642	-5.583072359	7.15E-13	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	NA
FM209324.1.1502	-4.961528364	1.20E-05	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	NA
GLRC							
HQ856458.1.1457	-25.26121416	3.13E-14	Acidobacteria	Solibacteres	Solibacterales	Solibacteraceae (Subgroup 3)	Paludibaculum
KC252876.1.1359	-24.62955468	1.23E-13	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Sphingobacteriaceae	Pedobacter
New.CleanUp. Reference OTU53505	-24.1779438	3.21E-13	Proteobacteria	Gammaproteobacteria	Legionellales	Legionellaceae	Legionella
New.ReferenceOTU61	-23.32984615	1.57E-12	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	NA
AJ518799.1.1286	-23.25017524	1.72E-12	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bdellovibrionaceae	OM27 clade

TABLE 3.13. Top 5 OTUs enriched in Bakken oil. This plot shows the top 5 OTUs that were most highly enriched in the Bakken samples determined by DESeq2 analysis

Top 5 OTUs enriched in Bakken amendments							
SampleID	log2FoldChange	padj	Phylum	Class	Order	Family	Genus
All Sites							
GQ480068.1.1499	2.771554143	0.00492125	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Brachymonas
KF010658.1.1492	2.189047269	0.002388574	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	NA
GQ379601.1.1243	1.886806751	0.00492125	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium
HQ178852.1.1434	1.87143764	0.004692176	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	NA
EU803786.1.1256	1.509995256	0.001452114	Proteobacteria	Alphaproteobacteria	SAR11 clade	LD12 freshwater group	uncultured bacterium
MI141							
Superior							
GU305795.1.1496	3.281463816	0.000485003	Proteobacteria	Betaproteobacteria	Burkholderiales	Alcaligenaceae	GKS98 freshwater group
Straits							
GU127217.1.1226	3.259472783	1.39E-05	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	NA
EU802021.1.1480	3.093166168	0.003194578	Bacteroidetes	Cytophagia	Cytophagales	Cytophagaceae	Pseudarcicella
KC633510.1.1495	2.444494544	5.94E-07	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	NA
GU305807.1.1464	2.260515525	9.30E-06	Planctomycetes	Phycisphaerae	Phycisphaerales	Phycisphaeraceae	CL500-3
FM955622.1.1508	2.213527138	0.003194578	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Hydrogenophaga
GLRC							
AB930605.1.1444	27.76384573	2.33E-19	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	env.OPS 17	uncultured bacterium
New.ReferenceOT U133	24.03550465	3.75E-13	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	Coxiella
LN561283.1.1365	23.36737773	1.57E-12	Actinobacteria	Actinobacteria	Micrococcales	Microbacteriaceae	NA
New.ReferenceOT U286	23.09031033	2.32E-12	Proteobacteria	Gammaproteobacteria	Legionellales	Legionellaceae	Legionella
KC619577.1.1503	22.86017142	1.57E-12	Proteobacteria	Gammaproteobacteria	Legionellales	Legionellaceae	Legionella

When comparing the OTUs enriched in the Bakken conditions compared to the DilBit-amended conditions in the Straits, it was found that 23 OTUs were enriched in the presence of Bakken-amended, while there were 7 OTUs enriched in the presence of DilBit. When comparing the Bakken and DilBit-amended OTUs enriched in Lake Superior, we found 2 OTUs were enriched in the presence of Bakken, while there was 1 OTU enriched in the presence of DilBit. When comparing the Bakken-amended and DilBit-amended OTUs enriched in the GLRC site, it was found that 23 OTUs were enriched in the presence of Bakken, while there were 99 OTUs enriched in the presence of DilBit. This finding is different than the rest of the comparisons where there was consistently more OTUs enriched in the Bakken compared to the DilBit. When comparing the Bakken-amended and DilBit-amended OTUs enriched in the MI141 site, we found that neither Bakken or DilBit had any OTUs that were enriched. This is contradictory to the PERMANOVA analysis which showed that MI141 had a significant difference between its Bakken and DilBit communities. This difference is attributed to how the data was processed. PERMANOVA analysis is done using a rarefied OTU table, while DESeq uses a normalization analysis on an unprocessed OTU table. Regardless, MI141 has too few OTUs present to determine significance and the PERMANOVA analysis p-value is close to 0.05. The results of the OTUs that were enriched between Bakken and Dilbit are shown in the volcano plots below FIGURE 3.12 (A-D). Based off these results DESeq analysis showed that there was a difference in the identity of OTUs present in

each oil amendment and that different OTUs grew in the presence of the different crude oil types.

Volcano Plots of Sites Bakken vs DilBit

FIGURE 3.12: Volcano Plots of OTUs enriched in Bakken-amendment or DilBit-amendment for each site. Yellow dots represent enriched OTUs, while blue dots are not considered enriched OTUs. A) Shows OTUs enriched in oil treatments in Lake Superior, B) Shows OTUs enriched in oil treatments in the Straits, C) Shows OTUs enriched in oil treatments in MI141, D) Shows OTUs enriched in oil treatments in the GLRC.

4 Discussion

The goal of this study was to characterize the impact of light and heavy crude oils on microbial diversity and composition of the Great Lakes. This study hypothesized that after the addition of oil, the microbial community will change dramatically with a common set of oil degrading taxa present throughout the Great Lakes. It was also hypothesized that the type of oil was a primary factor that shaped the microbial community composition.

4.1 Differences Between Sites

The initial expectation was that upon oil addition, the microbial community composition across all sample sites would converge to a common community composition dominated by the same oil-degrading taxa. This expectation was based off studies done in oceans that saw a common microbial community response to oil across environments and in diverse ocean contexts^{27,30,34,63}. However, we observed a community response that was individualized by lake and furthermore by location. Our data indicates similarity in the microbial community composition for samples from the same location, which is distinct from the microbial community composition in other regions. In all cases, the oil-amended microcosms are more similar to the control conditions from the sites of origin than they are to other oil-amended samples from other sites. There were also similarities in the community composition between samples from the same geographical region. We believe a few things could explain this region and site-

specific community composition. The difference in the community composition of samples from Lake Superior compared to those in Lake Michigan and the Straits could be due to little exchange of water between the lakes. Separating the two bodies of water (Lake Superior and Lake Michigan) is the Saint Mary's River and a set of locks at Sault Ste. Marie. Because of this geographical separation we believe that the bacterial communities are not able to mix in high amounts causing the different community composition seen in our data.

We speculate that the differences in the microbial community in the controls could be due to differences in the environmental conditions in the Straits compared to the offshore environments. These environmental differences could be differences in nutrient content as seen in other research in the Great Lakes when comparing nearshore and offshore locations. Previous research has shown that excessive nutrient loads are introduced to the nearshore environments in the Great Lakes through non-point sources in run-off from agriculture and sewage ³⁹. This could cause elevated nutrients in nearshore waters of the Straits of Mackinaw. While in comparison, in the offshore waters in the Great Lakes (Lake Michigan and Superior), nutrients are more limited ^{38,39,64}. Therefore, the microbial community in offshore samples could be more adapted for growth under lower nutrient conditions, which may select for a distinct community composition, compared to the Straits' environment^{38,39,65}. The Straits are a highly dynamic system with rapidly currents and high ship traffic. These factors also may select for distinct community composition in the Straits compared to the other sites. The microbial community in the Keweenaw Waterway may have

adapted to higher nutrients levels also, that are supplied by runoff since the Keweenaw Water is an inland canal^{38–40,65}. The Keweenaw waterway is also a highly dynamic system with water inputs from Lake Superior as well as Portage Lake, which may select for a distinct starting microbial community.

4.2 Microbial Community Response to Oil

The overarching goal of this work was to characterize the complement of oil degrading bacteria that respond to oil across the Great Lakes. Based on previous studies in the oceans, we expected that select oil degrading members of the overall community would increase in relative abundance to dominate the microbial community. This had been previously seen in a number of ocean environments, where the oil degraders bloomed to dominate the microbial community^{29,30,45}. In the Kalamazoo River study, a similar phenomenon was observed where there were a select set of OTUs that dramatically grew in the presence of DilBit, similar to the blooms observed in the oceans¹⁷. We expected that this bloom of select oil degraders would result in a decrease in Shannon diversity as the large majority of the bacteria would drop in relative abundance with these oil degraders increasing in abundance. Our results indicate that there were no universal trends regarding the impact of oil on the Shannon diversity. Instead, our results indicate that the effects on richness were site dependent.

In the GLRC, we saw a significant decrease in richness in the Bakken-amended samples. The decrease in richness for the Bakken oil amended samples could be due in part to the Bakken having more readily degradable components in

the oil compared to DilBit. Therefore, we believe that there is a set of bacteria in the community that are more efficient at breaking down some of the components in Bakken oil and that the conditions were such that a bloom of oil degraders occurred.

Work done as part of studying the Kalamazoo River spill showed as time progressed the microbial community composition changed as more of the DilBit was biodegraded and its chemistry changed ¹⁷. This change in microbial community composition during the Kalamazoo experiment was believed to be due to microbes preferentially degrading the alkanes first, then the community shifting to degrade more recalcitrant compounds. Similar observations of microbial dynamics being caused by preference for particular components of the oil were seen in other studies ⁶⁶. We believe that this difference in richness in Bakken versus DilBit conditions could be due to the distinct chemical composition of the oil selecting for particular organisms ^{24,34,41,45,67}. The distinct community composition in response to these oil types is supported by the DESeq enrichments.

In the marine environment, we see that when an oil spill occurs, the richness of the microbial community decreases and only a small subset of OTUs become the dominant OTUs in the community ^{17,27,29,30,34,63}. This is thought to be due to the enhanced ability of marine microbes to use the oil as a carbon source and their fast growth causes the decrease in the richness. The effect of a decrease in community richness as seen in the oceans is not as prevalent in our data, where

we did see changes in the community richness between sites but no overall decrease of richness in every site.

Our results demonstrate that there was a significant difference in the community composition between the controls and the oil amended samples in all seven of our sites. This finding is in line with other studies in the ocean which demonstrate that the microbial community composition in oil-amended conditions was distinct from no-oil controls^{27,29,30,34,67,68}. These significant changes could mean that the microbial community is shifting towards a subset of the original community that can live in the presence of oil or potentially degrade oil. Our results also match what was seen in the Kalamazoo River oil spill, where researchers found that when an enriched microbial community was exposed to DilBit, the bacterial community composition shifted towards one dominated by OTUs related to known oil-degrading microbes¹⁷. While we see a significant change in the community composition across all sites, the extent of change was different between locations, with the strongest changes observed in the GLRC and the Straits. Other sites such as Lake Superior and MI141 show a marginally significant different response. These differences in the extent of the response is potentially due to differences in starting microbial communities or nutrient levels found in open water vs. near shore differences³⁹. Based off previous research pointing to higher nutrient loads in nearshore environments such as the straits and inland environments such as the GLRC, nutrient loads may be a primary explanatory factor for the extent of microbial response to oil in the Great Lakes. Both the Straits and the GLRC have distinct environmental conditions which may

enable a more robust change in the microbial community, compared to the open water sites of Lake Michigan and Superior^{38,39,65}. Additionally, the response seen in the GLRC samples could be due to past oil exposure history as there have been small oil spills in the Keweenaw Waterway.

To clarify the difference in microbial community composition, we used DESeq to identify which OTUs were differentially abundant between the oil-amended conditions and the controls. We expected to see a low number of specific OTUs become highly enriched in our oil treatments. Our results indicate that there were a few specific OTUs that were enriched in the oil conditions relative to the controls across all of the sites. We believe that the enrichment of just a few OTUs is due to the microbial community shifting towards an oil degrading community. These OTUs that are enriched in the oil-amended conditions are most likely the common set of oil-degrading taxa found across the Great Lakes. We also observed a substantial number of OTUs that were enriched in the controls relative to the oil. These OTUs appear to be inhibited by oil and thus may be the members of the community that are most sensitive to oil. This type of enrichment response is similar to what others have seen. In the Kalamazoo River oil spill where the microbial communities shifted towards a community that was dominated with oil-degrading bacteria that could begin to break down oil¹⁷.

Based on previous studies that have been done in the oceans and in the Kalamazoo River, we expected the primary responding microbes to oil would be members of the Gammaproteobacteria. In the oceans, members of the Gammaproteobacteria such as *Marinobacter*, *Alcanivorax*, and *Oceanospirillales*

are often the dominant first responders to oil in the ocean, sometimes making up 57% of the responding community^{29,42}. In the Kalamazoo River the dominant oil-degraders that responded in the 25 °C incubations were *Pseudomonas* which are members of the Gammaproteobacteria¹⁷. In the enrichment from the Kalamazoo River, *Pseudomonas* sequences were 41±6% of the 25 °C oil-amended microbial community¹⁷. However, Gammaproteobacteria are often not particularly abundant in lake microbial communities⁶⁹. Despite the fact that Gammaproteobacteria are only a fraction of the ambient microbial community in lakes, it is possible that the oil-degrading communities in lakes may still be composed of Gammaproteobacteria as oil degraders. Oil degraders are normally minor constituents of the community and become dominant after oil addition^{17,29,30}. Our results however, demonstrated that Alpha and Betaproteobacteria were the primary oil-responding organisms across all the sites in the oil amended samples instead of Gammaproteobacteria. However, Gammaproteobacteria were mainly enriched in our controls.

Across all the sites, the most enriched OTUs in the oil-amended conditions were classified as Sphingomonadaceae, and Rhodocyclaceae families. The Sphingomonadaceae family was enriched in all the sites amended with oil. Sphingomonadaceae is a family within the Alphaproteobacteria that have been shown to be able to breakdown aromatic compounds⁷⁰. Similarly, members of the Rhodocyclaceae family have been shown to degrade hydrocarbons⁷¹. Most of the oil-amended enriched OTUs were not classifiable after the family level. However, one of the Sphingomonadaceae OTUs was classified down to the genus level as

Novosphingobium. Some representatives of the *Novosphingobium* have been shown to be responsible for the breakdown of hydrocarbon compounds such as; phenol, aniline, nitrobenzene, and phenanthrene⁷⁰. These findings indicate that known oil-degrading microbes are present throughout the Great Lakes and are able to respond to released oil under natural conditions.

As stated earlier in all the samples Alpha and Betaproteobacteria were the dominant class of bacteria responding to oil. The site-specific responses seen in our data set is believed to be derived from other site-specific OTUs persisting in the samples at the same time as these common oil-degraders are growing. When considering differences in the microbial community response to oil across the sites, there were some overlap in the individual OTUs that are enriched in the sites. For example, OTUs classified as Sphingomonadaceae and Rhodocyclaceae were present in multiple sites suggesting that they may be universal responders to the presence of oil in the Great Lakes. This finding indicates that while the overall community composition of the oil-amended treatments remains distinct between locations, the core set of oil-degraders are present throughout the Great Lakes.

We saw a number of OTUs that were enriched in the controls compared to oil. These OTUs are potentially sensitive to oil. The most enriched OTUs in the controls across all sites were classified as; chloroplast sequence from an uncharacterized family of diatoms, Actinobacteria from the order Sporichthyaceae of the genus hgcl clade, a Gammaproteobacteria from the Oceanospirillaceae order, and a member of the *Psedohongiella* genus. Even though our main responding OTUs to oil exposure were Alpha and Betaproteobacteria we still saw

one Betaproteobacteria from the order Burkholderiales and the family Comamonadaceae was enriched in the controls across all the sites as well as in the Straits. We believe that this may be because this OTU is an important player in the natural communities in the Great Lakes and may be unaffected by oil exposure. This would attribute to why we saw the OTU show up in the top 5 OTUs enriched in oil also. The enrichment of the chloroplast sequences in the controls could indicate a preference of diatoms for non-oil impacted samples. This finding suggests that oil may be toxic to algae. Similar results were shown in previous studies where they found that oil had a severe toxic effect on benthic species in freshwater ⁷². Overall we observe that nearshore sites had more OTUs that had enrichment in the controls.

4.3 Comparison of the Microbial Response to Heavy and Light Crude Oil

Our second hypothesis for this study hypothesized that the type of oil influenced the microbial community composition. In the study examining the microbial community response in enrichments from the Kalamazoo River, they found that there was no difference in the microbial community composition between two different types of DilBit. They did state however that as the hydrocarbons in the DilBit changed, the community composition and abundance changed ¹⁷. Additionally, the two oil types used in the Kalamazoo river study were both heavy crude oils with API gravities of 21.7 and 22.1. In other research, looking at how the chemical composition of the crude oil flowing from the *DWH*

well changed from alkanes to aromatic compounds, this change in oil composition caused a change in the dominant responding community of bacteria ⁷³. Both of these studies show that the chemical composition of the oil plays a large role in the microbial community dynamics. Since we were testing two chemically different crude oils, we expected to see a different community composition between the oil amendments. Our data indicates that there was a distinct community composition in Bakken-amended samples compared to DilBit-amended samples. The reason that we see this difference in diversity between the two microbial communities could be explained by the chemical differences in the oil types.

According to the PERMANOVA analysis there was a statistically significant difference between the Bakken and DilBit treatments in all the sites except in Lake superior sites. These differences were represented in a number of differentially abundant OTUs enriched in one of the oil types relative to the other. There were significantly more OTUs enriched in Bakken oil treatments relative to DilBit in most sites. Comamonadaceae and Burkholderiaceae were the most enriched bacteria families in the Bakken amended samples when comparing across all sites.

Comamonadaceae was found as the second most abundant responder in the presence of oil degrading communities in other research looking at oil contamination in ground water ⁷⁴. Not many OTUs were enriched in DilBit in most sites. However, in the GLRC site, more OTUs were highly enriched in DilBit. The OTUs that were consistently enriched the most across all the sites in

the DilBit amendments were from the families Comamonadaceae, Sphingomonadaceae, and Oxalobacteraceae. Of these families enriched in DilBit, only Comamonadaceae has been reported to be present in other oil biodegradation studies. This may mean that Comamonadaceae could be a primary DilBit degrading bacteria in the Great Lakes.

It is possible that significant differences in abundance of OTUs between these two oil could be due to differences the ability of these taxa metabolize different chemicals in these oils as they have distinct chemical makeups. Bakken oil is a light crude oil that has a higher proportion of alkanes which may be more favorable for degradation to some bacteria ⁴¹. DilBit oil is a heavier crude oil that is composed of more aromatic rings and has a higher content of asphaltenes and resins in comparison to Bakken crude, so it may be more favorable to aromatic degrading bacteria ^{17,41}. However as seen in the enrichment tables members of the Comamonadaceae were prevalent in both the Bakken and Dilbit enrichments. This could be because members of the Comamonadaceae were capable of consuming both alkanes and aromatic compounds. While Comamonadaceae were present in both oil-types, there was a significant difference in the microbial community composition in the different oil treatments. This difference could be explained by the chemical difference between Bakken and DilBit .

We saw a different set of OTUs enriched in Bakken compared to DilBit. Based on the different OTUs that were enriched in the Bakken vs. DilBit table we found that in the Great Lakes a universal microbial community would respond to oil with accessory microbes specifically responding to particular oil types

depending on its composition. We found that the community that will respond to oil is dependent on the type of oil that is present. We believe that this is also due to the different chemical makeup of the two oils tested and the microbes that are naturally present at each location.

5 Conclusion

There is very little knowledge currently that can attest to what would happen to the microbial community in the case of an oil exposure event in the Great Lakes. This study was designed to understand how oil impacted microbial communities in fresh water, and specifically the Great Lakes. Seven sample locations were chosen throughout two of the Great Lakes. The collected water was used in a five-week microcosm study where the microbial communities were amended with either Bakken or Diluted Bitumen crude oil. The resulting microbial communities were sequenced and analyzed yielding the results of the study. This study found that in the Great Lakes there is a site dependent response to oil exposure by the microbial community, and an oil-type dependent response within the microbial community. This study also found what is possibly a variation in response of Straits and offshore microbial communities. This study did find though that a systemic response in the Great Lakes to oil is possible from bacteria from the families; Sphingomonadaceae, Rhodocyclaceae, Burkholderiales, and Comamonadaceae, which have all previously been shown to contain species able to break down hydrocarbons. Therefore, in the event of an oil spill in the Great Lakes a response from the microbial community is possible in most locations. However, in the case of a spill in the Great Lakes the location of the spill and the type of oil should be assessed before bioremediation is a possibility.

6 References

1. Colwell, R. R. Toxic Effects of Pollutants on Microorganisms. 20
2. Goldman, S. S., Maya. New report highlights risks of Enbridge Line 5 pipeline. Available at: <http://www.michiganradio.org/post/new-report-highlights-risks-enbridge-line-5-pipeline>. (Accessed: 13th November 2018)
3. Risk Analysis for the Straits Pipelines | MI Petroleum Pipelines. Available at: <https://mipetroleumpipelines.com/document/risk-analysis-straits-pipelines>. (Accessed: 13th November 2018)
4. McCauley, R. N. The Biological Effects of Oil Pollution in a River1. *Limnol. Oceanogr.* **11**, 475–486 (1966).
5. ATSDR. Toxicity of Polycyclic Aromatic Hydrocarbons (PAHs): Health Effects Associated With PAH Exposure | ATSDR - Environmental Medicine & Environmental Health Education - CSEM. Available at: <https://www.atsdr.cdc.gov/csem/csem.asp?csem=13&po=11>. (Accessed: 24th April 2017)
6. Peterson, C. H. *et al.* Long-Term Ecosystem Response to the Exxon Valdez Oil Spill. *Science* **302**, 2082–2086 (2003).
7. Skogdalen, J. E. & Vinnem, J. E. Quantitative risk analysis of oil and gas drilling, using Deepwater Horizon as case study. *Reliability Engineering & System Safety* **100**, 58–66 (2012).
8. Oracle BI Interactive Dashboards - Public Reports. Available at: <https://hip.phmsa.dot.gov/analyticsSOAP/saw.dll?Portalpages>. (Accessed: 6th November 2018)
9. The Rivers Crossed by the Keystone XL Pipeline. Available at: <https://www.arcgis.com/home/item.html?id=da30f3d659e04d03bc1b80cf56b7d2d0>. (Accessed: 11th November 2018)
10. DeepwaterHorizonReport -31Aug2011 -CD_2.pdf.
11. Atlas, R. M. & Hazen, T. C. Oil Biodegradation and Bioremediation: A Tale of the Two Worst Spills in U.S. History. *Environ Sci Technol* **45**, 6709–6715 (2011).
12. Lee, R. F. & Page, D. S. Petroleum hydrocarbons and their effects in subtidal regions after major oil spills. *Marine Pollution Bulletin* **34**, 928–940 (1997).
13. US EPA, R. 05. EPA Response to Enbridge Spill in Michigan. Available at: <https://www.epa.gov/enbridge-spill-michigan>. (Accessed: 21st February 2017)
14. Usepa, A. D. V. NRT FACT SHEET: BIOREMEDIATION IN OIL SPILL RESPONSE. 4
15. Geochemical changes in crude oil spilled from the Exxon Valdez supertanker into Prince William Sound, Alaska - ScienceDirect. Available at: <https://www.sciencedirect.com/science/article/pii/0146638094900515>. (Accessed: 11th November 2018)
16. Dollhopf, R. & Durno, M. Kalamazoo River\Enbridge Pipeline Spill 2010. *International Oil Spill Conference Proceedings* **2011**, abs422 (2011).

17. Ruta S.DeshpandeaDeviSundaravadeluaStephenTechtmannbRobyn N.ConmycJorge W.Santo DomingocPabloCampod. Microbial degradation of Cold Lake Blend and Western Canadian select dilbits by freshwater enrichments - ScienceDirect. Available at: <https://www.sciencedirect.com/science/article/pii/S0304389418301821>. (Accessed: 11th November 2018)
18. enbridge-fosc-report-20160407-241pp.pdf.
19. Crude oil | petroleum product | Britannica.com. Available at: <https://www.britannica.com/science/crude-oil>. (Accessed: 11th November 2018)
20. Learn more about ExxonMobil Crude Trading sales through this listing by geographic location. *ExxonMobil* Available at: <http://corporate.exxonmobil.comhttps://corporate.exxonmobil.com/en/company/worldwide-operations/crude-oils/crude-by-region>. (Accessed: 11th November 2018)
21. Chinenyeze, M. A. J. & Ekene, U. R. Physical and Chemical Properties of Crude Oils and Their Geologic Significances. **6**, 8 (2015).
22. Petroleum - Classification - API. Available at: <http://www.petroleum.co.uk/api>. (Accessed: 11th November 2018)
23. Characterization of chemical fingerprints of unconventional Bakken crude oil - PDF Free Download. *kundoc.com* Available at: <https://kundoc.com/pdf-characterization-of-chemical-fingerprints-of-unconventional-bakken-crude-oil-.html>. (Accessed: 15th November 2018)
24. Cold Lake Analysis - CrudeMonitor.ca. Available at: <https://crudemonitor.ca/crudes/index.php?acr=CL>. (Accessed: 11th November 2018)
25. Baca, B. J. & Getter, C. D. Freshwater oil spill considerations: protection and cleanup. *International Oil Spill Conference Proceedings* **1985**, 385–390 (1985).
26. McKinney, K., Caplis, J., DeVitis, D. & Van Dyke, K. Evaluation of oleophilic skimmer performance in diminishing oil slick thicknesses. *International Oil Spill Conference Proceedings* **2017**, 1366–1381 (2017).
27. Head, I. M., Jones, D. M. & Röling, W. F. M. Marine microorganisms make a meal of oil. *Nature Reviews Microbiology* **4**, 173–182 (2006).
28. Olajire AA, E. J. Aerobic Degradation of Petroleum Components by Microbial Consortia. *Journal of Petroleum & Environmental Biotechnology* **05**, (2014).
29. Hazen, T. C., Prince, R. C. & Mahmoudi, N. Marine Oil Biodegradation. *Environ. Sci. Technol.* **50**, 2121–2129 (2016).
30. Hazen, T. C. *et al.* Deep-Sea Oil Plume Enriches Indigenous Oil-Degrading Bacteria. *Science* **330**, 204–208 (2010).
31. FOGHT, J. M. & WESTLAKE, D. W. S. BIODEGRADATION OF HYDROCARBONS IN FRESHWATER. in *Oil in Freshwater: Chemistry, Biology, Countermeasure Technology* (eds. VANDERMEULEN, J. H. &

- HRUDEY, S. E.) 217–230 (Pergamon, 1987). doi:10.1016/B978-0-08-031862-2.50022-8
32. Howarth, R. W. Nutrient Limitation of Net Primary Production in Marine Ecosystems. *Annual Review of Ecology and Systematics* **19**, 89–110 (1988).
 33. Sherry, A. *et al.* Anaerobic biodegradation of crude oil under sulphate-reducing conditions leads to only modest enrichment of recognized sulphate-reducing taxa. *International Biodeterioration & Biodegradation* **81**, 105–113 (2013).
 34. MacNaughton, S. J. *et al.* Microbial Population Changes during Bioremediation of an Experimental Oil Spill. *Appl. Environ. Microbiol.* **65**, 3566–3574 (1999).
 35. Edwards, B. R. *et al.* Rapid microbial respiration of oil from the Deepwater Horizon spill in offshore surface waters of the Gulf of Mexico. *Environ. Res. Lett.* **6**, 035301 (2011).
 36. Salminen, J. M., Tuomi, P. M., Suortti, A.-M. & Jørgensen, K. S. Potential for aerobic and anaerobic biodegradation of petroleum hydrocarbons in boreal subsurface. *Biodegradation* **15**, 29–39 (2004).
 37. Pu, F., Philip, R. P., Zhenxi, L. & Guangguo, Y. Geochemical characteristics of aromatic hydrocarbons of crude oils and source rocks from different sedimentary environments. *Organic Geochemistry* **16**, 427–435 (1990).
 38. Nutrients in Michigan - Overview. 3 - DEQ
 39. (PDF) Feast and Famine in the Great Lakes: How Nutrients and Invasive Species Interact to Overwhelm the Coasts and Starve Offshore Waters. *ResearchGate* doi:<http://dx.doi.org/10.13140/RG.2.1.3710.1609>
 40. Impacts of suspended sediment on the ecosystem in Lake Michigan: A comparison between the 1998 and 1999 plume events - Chen - 2004 - Journal of Geophysical Research: Oceans - Wiley Online Library. Available at: <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2002JC001687>. (Accessed: 15th November 2018)
 41. Atlas, R. M. Effects of Temperature and Crude Oil Composition on Petroleum Biodegradation. *Appl. Environ. Microbiol.* **30**, 396–403 (1975).
 42. Liu, Z. & Liu, J. Evaluating bacterial community structures in oil collected from the sea surface and sediment in the northern Gulf of Mexico after the Deepwater Horizon oil spill. *MicrobiologyOpen* **2**, 492–504 (2013).
 43. Microbial gene functions enriched in the Deepwater Horizon deep-sea oil plume | The ISME Journal. Available at: <https://www.nature.com/articles/ismej201191#ref10>. (Accessed: 11th November 2018)
 44. Røling, W. F. M. *et al.* Bacterial Community Dynamics and Hydrocarbon Degradation during a Field-Scale Evaluation of Bioremediation on a Mudflat Beach Contaminated with Buried Oil. *Appl. Environ. Microbiol.* **70**, 2603–2613 (2004).
 45. Albokari, M., Mashhour, I., Alshehri, M., Boothman, C. & Al-Enezi, M. Characterization of microbial communities in heavy crude oil from Saudi Arabia. *Annals of Microbiology* **65**, 95–104 (2015).

46. Heitkamp, M. A. & Cerniglia, C. E. Effects of chemical structure and exposure on the microbial degradation of polycyclic aromatic hydrocarbons in freshwater and estuarine ecosystems. *Environmental Toxicology and Chemistry* **6**, 535–546 (1987).
47. Djomo, J. E. *et al.* Toxic effects of some major polyaromatic hydrocarbons found in crude oil and aquatic sediments on *Scenedesmus subspicatus*. *Water Research* **38**, 1817–1821 (2004).
48. Kim, K.-H., Jahan, S. A., Kabir, E. & Brown, R. J. C. A review of airborne polycyclic aromatic hydrocarbons (PAHs) and their human health effects. *Environment International* **60**, 71–80 (2013).
49. ATSDR - Public Health Statement: Polycyclic Aromatic Hydrocarbons (PAHs). Available at: <https://www.atsdr.cdc.gov/phs/phs.asp?id=120&tid=25>. (Accessed: 14th November 2018)
50. The Unique Chemistry of Eastern Mediterranean Water Masses Selects for Distinct Microbial Communities by Depth. Available at: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0120605>. (Accessed: 13th November 2018)
51. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology* **18**, 1403–1414 (2016).
52. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**, 335–336 (2010).
53. Aronesty, E. Comparison of Sequencing Utility Programs. *The Open Bioinformatics Journal* **7**, 1–8 (2013).
54. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
55. Release 128. Available at: <https://www.arb-silva.de/documentation/release-128/>. (Accessed: 15th November 2018)
56. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy | Applied and Environmental Microbiology. Available at: <https://aem.asm.org/content/73/16/5261.short>. (Accessed: 14th November 2018)
57. R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
58. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**, e61217 (2013).
59. *vegan.pdf*.
60. Anderson, M. J. Permutational Multivariate Analysis of Variance (PERMANOVA). in *Wiley StatsRef: Statistics Reference Online* 1–15 (American Cancer Society, 2017). doi:10.1002/9781118445112.stat07841

61. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
62. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer, 2016).
63. Yang, S. *et al.* Hydrocarbon degraders establish at the costs of microbial richness, abundance and keystone taxa after crude oil contamination in permafrost environments. *Scientific Reports* **6**, 37473 (2016).
64. Green, S. A. & Eadie, B. J. Introduction to special section: Transport and transformation of biogeochemically important materials in coastal waters. *Journal of Geophysical Research: Oceans* **109**, (2004).
65. Edsall, T. A., Arbor, A. & Charlton, M. N. NEARSHORE WATERS OF THE GREAT LAKES. 180
66. Dubinsky, E. A. *et al.* Succession of Hydrocarbon-Degrading Bacteria in the Aftermath of the Deepwater Horizon Oil Spill in the Gulf of Mexico. *Environ. Sci. Technol.* **47**, 10860–10867 (2013).
67. Influence of crude oil on changes of bacterial communities in Arctic sea-ice | FEMS Microbiology Ecology | Oxford Academic. Available at: <https://academic.oup.com/femsec/article/53/1/129/515893>. (Accessed: 13th November 2018)
68. Rapid Response of Eastern Mediterranean Deep Sea Microbial Communities to Oil | Scientific Reports. Available at: <https://www.nature.com/articles/s41598-017-05958-x>. (Accessed: 11th November 2018)
69. A guide to the natural history of freshwater lake bacteria. - PubMed - NCBI. Available at: <https://www.ncbi.nlm.nih.gov/pubmed/21372319>. (Accessed: 15th November 2018)
70. *The Prokaryotes: Vol. 7: Proteobacteria: Delta and Epsilon Subclasses. Deeply Rooting Bacteria*. (Springer-Verlag, 2006).
71. Singleton, D. R., Dickey, A. N., Scholl, E. H., Wright, F. A. & Aitken, M. D. Complete Genome Sequence of a Novel Bacterium within the Family Rhodocyclaceae That Degrades Polycyclic Aromatic Hydrocarbons. *Genome Announc* **3**, (2015).
72. Bhattacharyya, S., Klerks, P. L. & Nyman, J. A. Toxicity to freshwater organisms from oils and oil spill chemical treatments in laboratory microcosms. *Environmental Pollution* **122**, 205–215 (2003).
73. Ribicic, D. *et al.* Microbial community and metagenome dynamics during biodegradation of dispersed oil reveals potential key-players in cold Norwegian seawater. *Marine Pollution Bulletin* **129**, 370–378 (2018).
74. Sun, Y., Lu, S., Zhao, X., Ding, A. & Wang, L. Long-Term Oil Pollution and In Situ Microbial Response of Groundwater in Northwest China. *Arch. Environ. Contam. Toxicol.* **72**, 519–529 (2017).

7 Additional community data

7.1 Taxa Bar Plot

FIGURE 3.4.1 Taxa plot showing individual taxa enriched in oil treatment separated by site. Colors represent different taxa at the class level of classification.