CULTIVATION-BASED DESCRIPTION OF MICROBIAL DIVERSITY IN SEDIMENT SAMPLES AFTER A MOLASSES SPILL IN HONOLULU HARBOR, O'AHU, HAWAI'I (USA)

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI'I AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

MARINE BIOLOGY

MAY 2016

By

Siobhán L. Burns, M.P.A.

Thesis Committee:

Stuart Donachie, Chairperson

Robert Richmond

Thomas Oliver

ACKNOWLEDGEMENTS

I would like to thank the members of my committee for their guidance and support during the course of my studies: Dr. Stuart Donachie, Dr. Robert Richmond, and Dr. Thomas Oliver. I would also like to thank the members of the Department of Microbiology who have been supportive of me during my studies: Dr. Susan Ayin, Dr. Sean Callahan, Dr. Tung Hoang, Mark Chaplin, Blake Ushijima, Christina Runyon, Silvia Beurmann, and Dr. Maqsudul Alam. Additionally, I would like to acknowledge Lindsay Root, Dr. Xin Wang, Dr. Durrell Kapan, Dr. Eric de Carlo, Dr. Mahdi Belcaid, Dr. Jimmy Saw, Solomon Champion, Ariell Colis, Tenny Pan, Kazukuni Hayashi, Shannon Morgan, Norman Wang, and Dr. Nathalie Coffineau for their contributions to this project. This thesis is dedicated with gratitude to my friends and family.

TABLE OF CONTENTS

TABLES

FIGURES

ABBREVIATIONS

Advanced Studies of Genomics, Proteomics, and Bioinformatics……ASGPB

Colony Forming Units……CFU

Department of Health……..……DoH

Department of Transportation……..DoT

Marine Agar……...MA

Operational Taxonomic Units……OTUs

Polymerase Chain Reaction……...PCR

Reasoner's 2A……..……R2A

Ribosomal Database Project……..RDP-II

Ribosomal Ribonucleic Acid……..…..rRNA

CHAPTER 1. INTRODUCTION

1.1 Event Summary

In September 2013, Matson Navigation Company, Inc. ("Matson") reported a major spill of molasses had occurred during a routine transfer. Over 233,000 gallons of molasses were released into Honolulu Harbor, killing approximately 26,000 fish (Matson, 2015). A molasses spill of this scale was unprecedented in Hawai'i. Honolulu Harbor is the largest commercial marine shipping facility in the State of Hawai'i, and among the largest container handling ports in the US, with over 8 million short tons of cargo handled annually (HI-DoT, 2012a). Honolulu Harbor is the primary port for the State, serving as the distribution point for cargo shipments to the neighbor islands. The harbor is well protected from wind and surge action, and offers ideal conditions for transferring freight; it has been dredged to accomodate large commercial cargo ships (HI-DoT, 2012a). Honolulu Harbor also serves passenger, research, fishing, and excursion vessels, and is located approximately 3 miles east of the Honolulu International Airport (HI-DoT, 2012a).

1.2 Timeline

At 6:30 p.m. on Saturday, September 7, 2013, a routine overnight transfer of molasses was initiated from shore-to-ship. This was completed before sunrise the following morning, apparently without incident (KHON2, 2013). According to a Hawai'i State Department of Health news release, the molasses was to be shipped to the west coast of the United States for processing. The ship left the harbor the following day before sunrise on Monday, September 9th, 2013 (HI-DoH, 2013a). A report of an unusual substance in the water was called in at 8 a.m. to the local Coast Guard station (Watkins,

2013). The Coast Guard originally responded to the call at Pier 41 & 42, directly across the Kāpalama Basin in Honolulu Harbor from the actual spill site, slowing response efforts (Figure 1). Reports of dead fish at Kāpalama Stream, added to the confusion about the source of the substance (Bidigare, 2015).

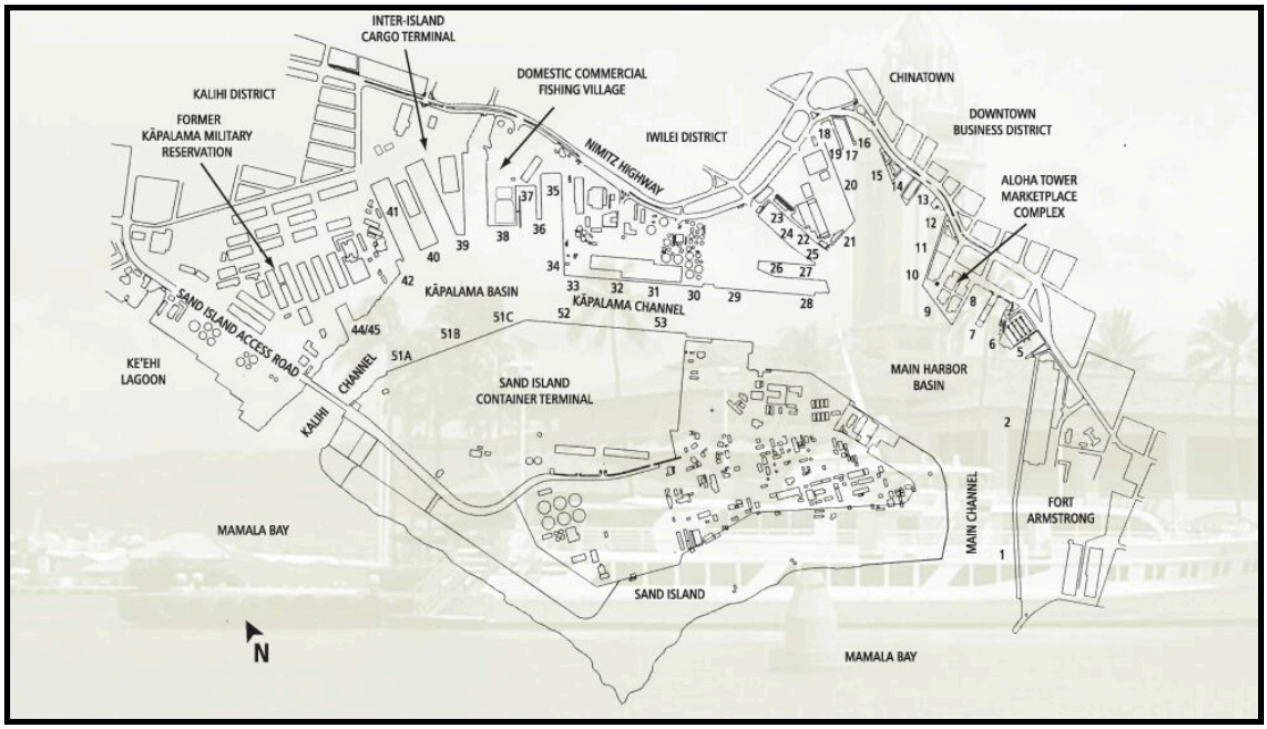


Figure 1. Honolulu Harbor, Hawai'i (HI-DoT[2], 2012).

The following day, the Hawai'i State Department of Health (DoH) issued a press release stating that a molasses spill at Pier 52 was the "apparent cause of death for many marine animals" (HI-DOH, 2013a). The DoH closed beaches in the Ke'ehi Lagoon area, and advised ocean users to stay out of the contaminated area of the spill, or any areas with a brown color in the water, in part due to concerns of harmful algal or bacterial blooms (HI-DoH, 2013a). The DoH also warned the public not to consume any fish that

had been killed by the spill, and to remain out of the area due to the concern of an increase in predator species, such as sharks or eels (HI-DoH, 2013a).

Figure 2. Department of Health Molasses Spill sampling sites – September 11-17, 2013 (HI-DoH, 2013b).

In response to the spill, a multi-agency response team was formed that included State and Federal partners:

State:

- Hawai'i Department of Health,
- Hawai'i Department of Land and Natural Resources,
- Hawai'i Department of Transportation,
- University of Hawai'i at Mānoa,

Federal:

- U.S. Coast Guard,
- U.S. Fish and Wildlife Service,
- U.S. Environmental Protection Agency, and
- National Oceanic and Atmospheric Administration.

The Hawai‘i Department of Health was the lead response agency, and oversaw aerial surveys, collection of water samples, removal and testing of the dead fish, posting of warning signs, assessment of damage to corals, monitoring molasses dispersion, and establishment of an Incident Command (HI-DoH, 2013b). Water samples were collected by the State of Hawai‘i Department of Health (DoH) and the Department of Land and Natural Resources (DLNR) at 15 locations (Figure 2) in order to monitor the movement of the molasses plume (Figure 3) and determine when the plume was effectively washed out to sea (HI-DoH, 2013a; HI-DoH, 2013b).

1.3 Impacts of a Major Molasses Spill in Honolulu Harbor

In the week following the spill, thousands of dead fish were removed from the harbor (Joaquin, 2013a), corals were bleached white, and dead crabs lined the shallows, with live crabs standing on rocks just above the water line; the air was pungent with the smell of rotting fish (pers. obs.). Over a nine-day period following the spill (September 9-19, 2013), a total of 26,492 marine organisms were collected, reportedly having been killed by the hypoxic conditions caused by the spill (Joaquin, 2013b; Joaquin, 2013c), with a peak of 11,604 organisms collected on September 12, 2013 (HI-DoH, 2013b). Necropsies performed on the fish were unable to distinguish mortality by hypoxia or osmotic shock due to ionic stress with the input of molasses (Richmond, 2015).

Roger Smith, a SCUBA diver hired by a local news channel (Joaquin, 2013a) reported, "…[T]he entire bottom is covered with dead fish. Small fish, crabs, mole crabs, eels… Every type of fish that you don't usually see, but now they're dead… Every single thing is dead." A member of the State's Environmental Health Division described the spill as "the worst environmental damage to sea life that I have come across…" (Memmott, 2013). At the time, no one knew what the long-term effects might be on the harbor and marine life after the spill, and how long it would take for the marine community to recover.

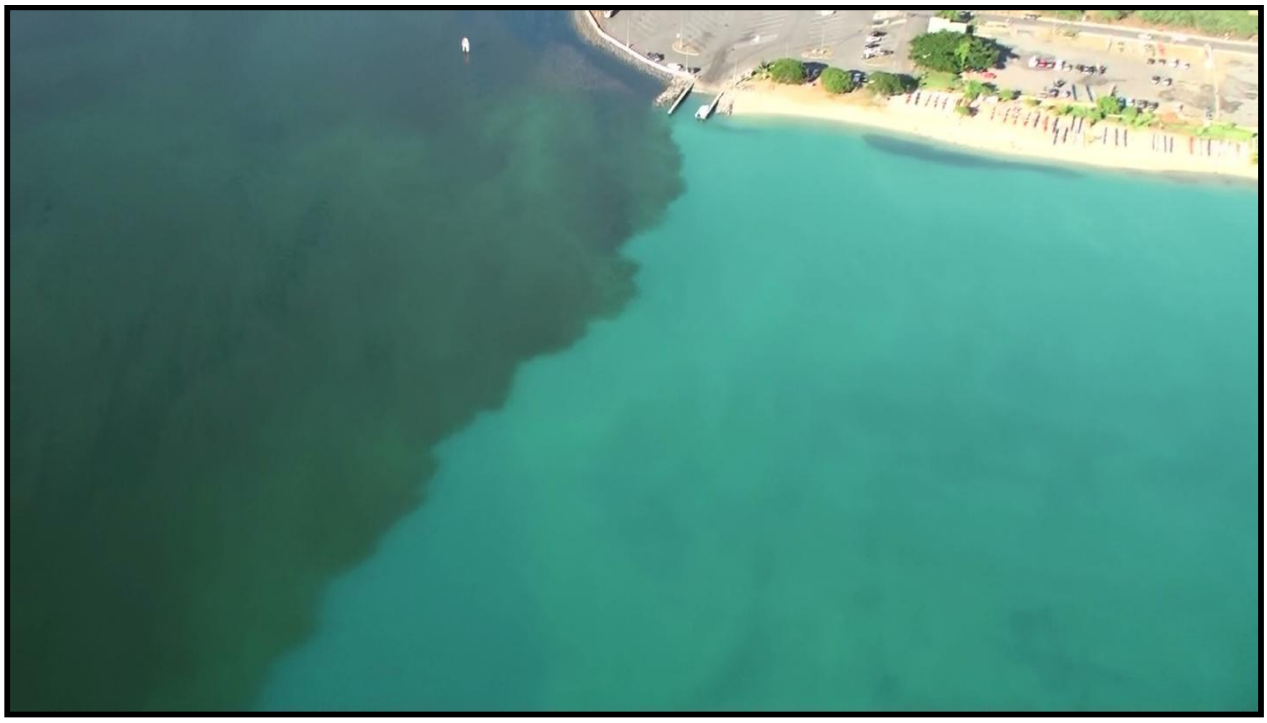


Figure 3. Sand Island Launch Ramp on September 11, 2013 at ~5:00 p.m. (HI-DoH, 2013b). (Additional photographs in Appendices 1-5).

By Tuesday, September 17, 2013, there was no visible discoloration in Honolulu Harbor, and there was no evidence of molasses within 200-feet of the spill's origin (HI-DoH, 2013b).

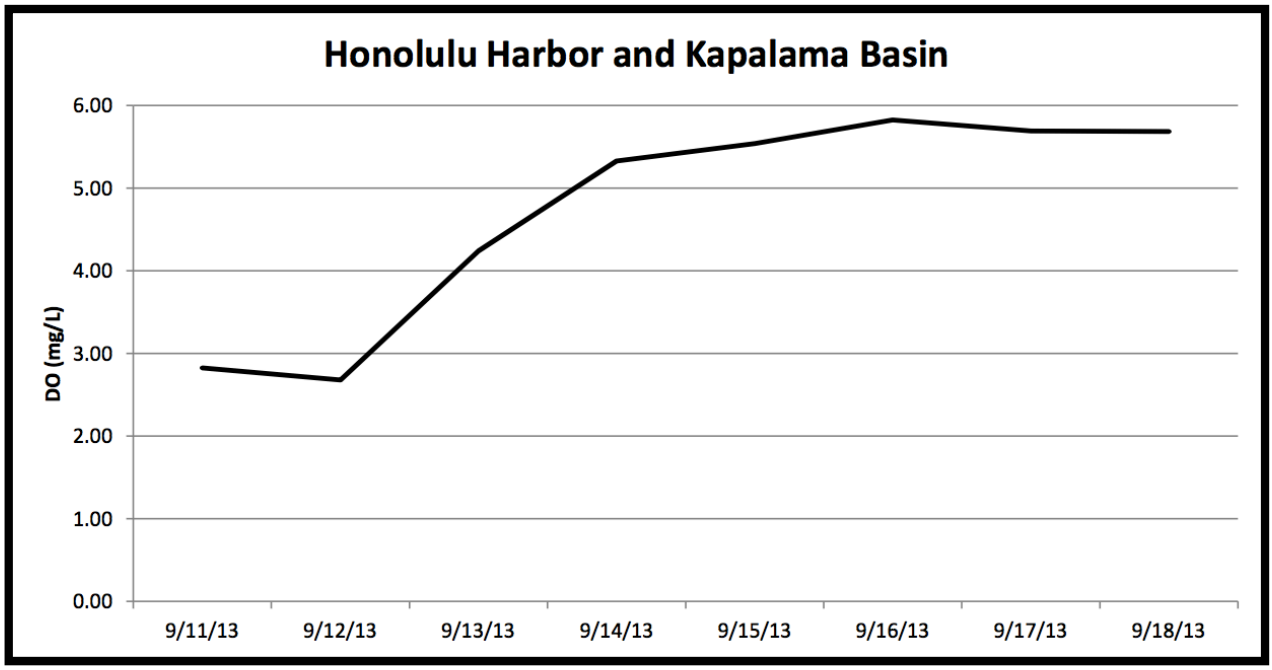


Figure 4. Dissolved oxygen concentration (mg/L) in Honolulu Harbor and Kāpalama Basin (09/11/2013 to 09/18/2013) (HI-DoH, 2013b).

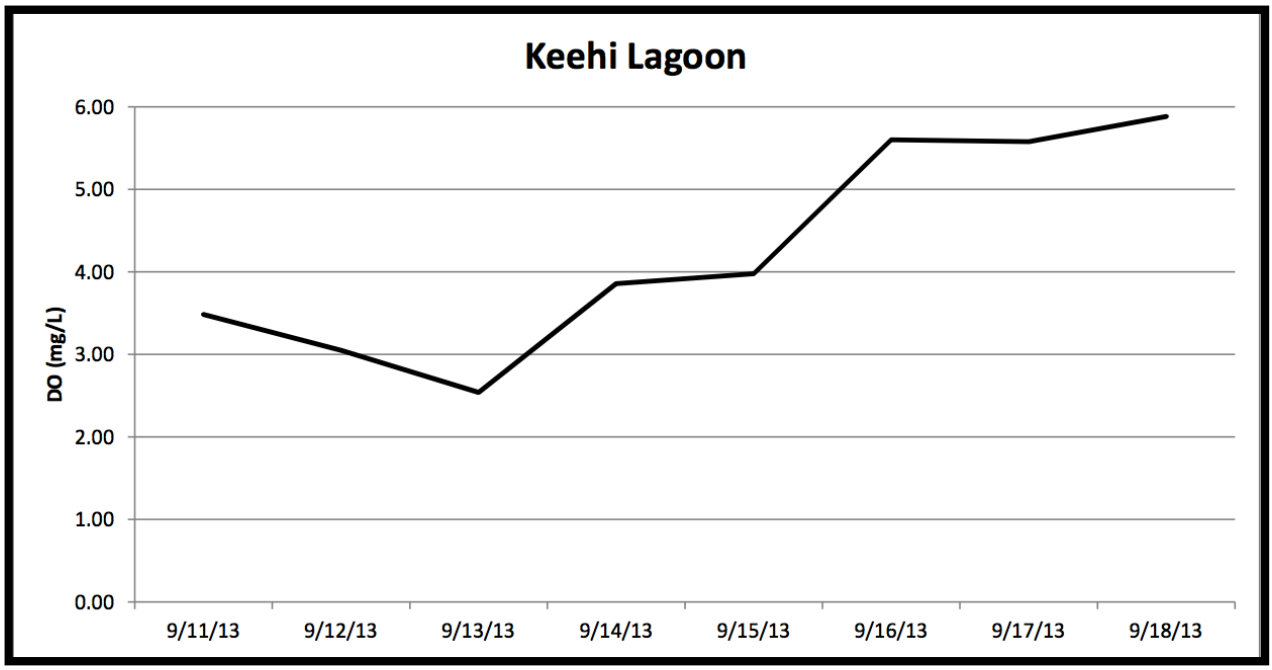


Figure 5. Dissolved oxygen concentration (mg/L) in Keʻehi Lagoon (09/11/2013 to 09/18/2013) (HI-DoH, 2013b).

On September 18th, 2013, dissolved oxygen concentrations were reported as having returned to historically normal levels for that area (Figure 4 & Figure 5; HI-DoH, 2013b), with slower recovery to such levels at Keʻehi Lagoon where there is less mixing by tidal flushing. This is evident in the patterns of molasses flow trends mapped by the Hawaiʻi State Department of Health that show the areas of restricted water flow (i.e., near the airport runway) or circulation (e.g., Keʻehi Lagoon) remained impacted longer than the site of the initial spill (Figure 6).

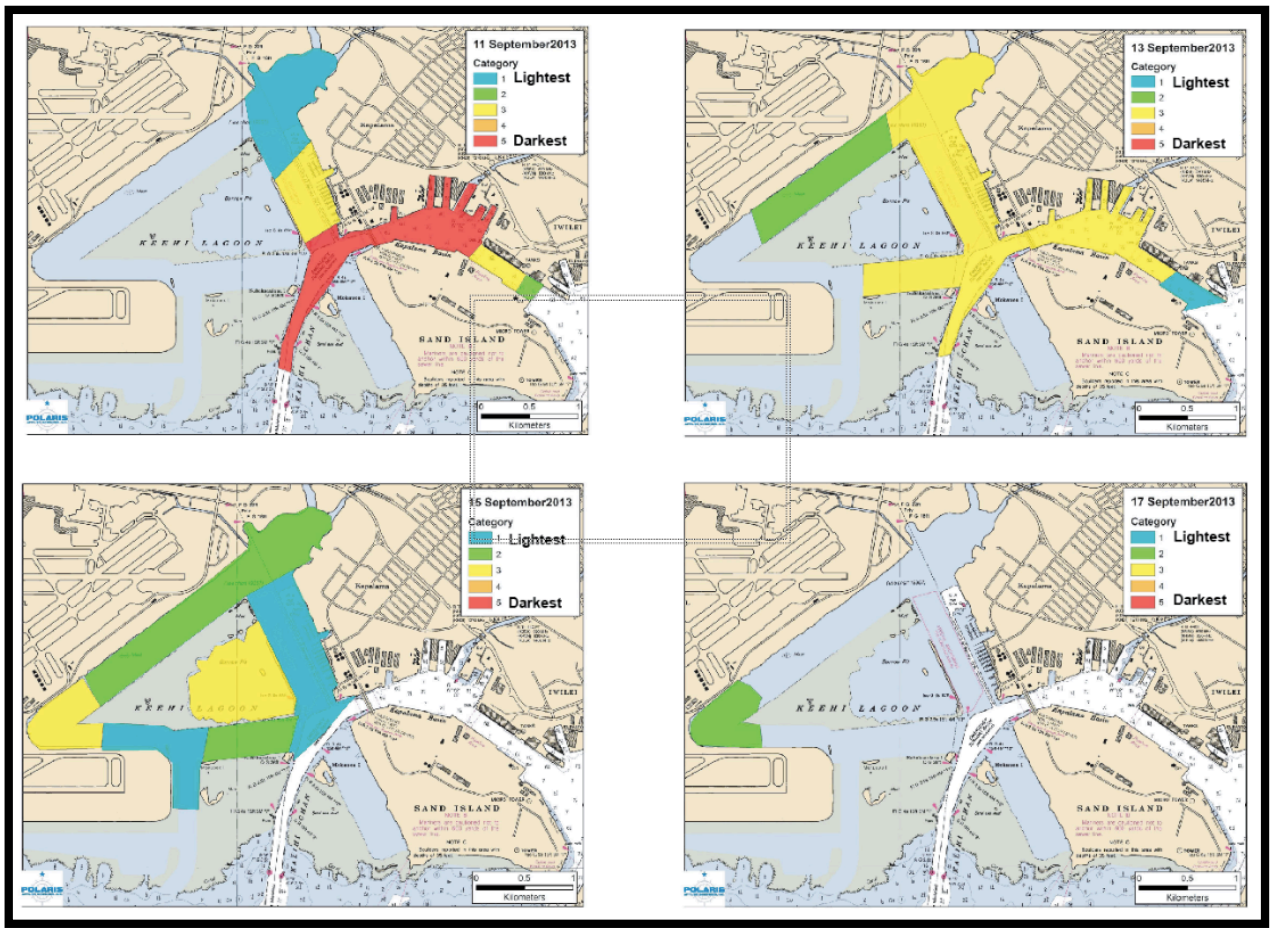


Figure 6. Molasses flow trends in Honolulu Harbor September 11-17, 2013 (HI-DoH, 2013b).

By Thursday, September 19, 2013, aerial surveys had been discontinued because boat surveys found only slight discoloration in Suicide Channel (near the end of the airport runway). Dive surveys to assess coral health on Thursday, September 19th, found *Porites* spp. in normal condition, but dead *Pocillopora damicornis* near Sand Island Bridge and Pier 42 (HI-DoH, 2013b). No endangered species were reportedly impacted by the spill (HI-DoH, 2013b).

1.4 Investigation into Point-source Pollution

A corroded elbow pipe below the dock at Pier 51B created a breach in the transfer line between shipping containers (Joaquin, 2013d; Figure 7. Piers 51 through 53 at Sand Island (HI-DoT[2], 2012).). In total, over 1400 tons (233,000 gallons) of a 1600-ton load of molasses was spilled into Honolulu Harbor during the transfer (HI-DoH, 2013).

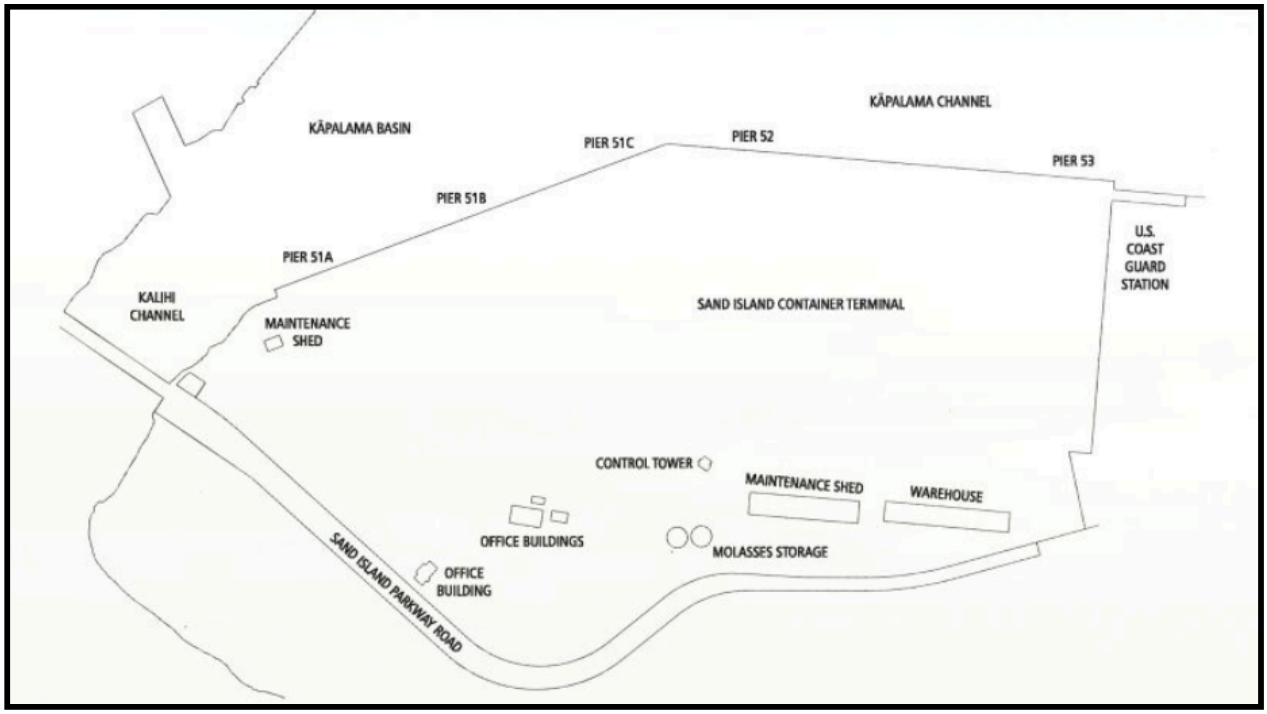


Figure 7. Piers 51 through 53 at Sand Island (HI-DoT[2], 2012).

1.5 Remediation Plan

Molasses is not classified as a hazardous substance. As noted in Matson's press release, molasses is a non-regulated product (Angoco, 2013), so no contingency plan for response and remediation existed. Molasses is denser than seawater, causing it to sink and settle on the sediment, unlike an oil spill which can be skimmed from surface water (Joaquin, 2013b). Local responders were not prepared for such an incident, and there was little that could be done to remove the molasses beyond natural tidal flushing and freshwater mixing from nearby streams.

1.6 Outcomes

Assessment of the spill focused on the fish kill and impact on coral health, and to a lesser extent on impact on benthic marine life. Approximately 26,000 fish deaths were recorded; 20,000 of these fish were under 12.7 cm (5 in) in length (Matson, 2015). An estimated 17,000 corals were classified as "recently dead", although the percentage of deaths related to the spill was not determined. Most of the impacted corals were less than 10 cm in size, and considered to have high natural mortality rates (Matson, 2015). The corals in Honolulu Harbor were considered to be in a "non-natural" and "sub-optimal" environment, and to have a low ecosystem-service value compared to pristine areas that were impacted by shipping in the state (Matson, 2015). The density of the spilled molasses caused it to sink to the sediment-seawater interface because of the density of molasses, where it impacted non-coral marine benthos, including polychaetes, tunicates, bivalves, bryozoans and sponges. A one-year, post-spill survey determined that the benthos was thriving and appeared to have been minimally impacted over that time scale by the spill (Matson, 2015). Despite having low-

ecosystem service value, corals that can withstand the stressful conditions of the harbor, including pollution, thermotolerance or changes in salinity, represent resilient genotypes that could serve to reseed reefs in Hawai'i as urbanization continues (Richmond, 2015) and the effects of climate change become more prevalent in sub-tropical areas (Dalton & Roff, 2013).

1.7 Industrial & Agricultural Uses of Molasses

Molasses became an important feed-stock additive beginning in the 1960s, with research out of Cuba showing that it could be used to reduce the need for forage while simultaneously fattening cattle (Preston, 1986). "Molasses" is considered any liquid feed ingredient with a total sugar content of greater than 43%, and can be sourced from various sugar sources, including cane, beet, citrus, wood (hemillose), and starch (Curtin, 1983). Cane molasses is a byproduct of sucrose production from sugar cane, and is created through the repeated evaporation, crystallization, and centrifugation of cane juice (Curtin, 1983). In the state of Hawai'i, the majority of sugar cane molasses is produced on the island of Maui, where The Hawaiian Commercial & Sugar Company generates approximately 1,500 tons weekly, with annual production of estimated at 50,000 tons of molasses (Grube, 2013). For commercial use, molasses is used in the production of bio-ethanol for both fuel and liquor production (Gopal & Kammen, 2009; Schworm, 2015).

1.8 Historical Molasses Spills

1.8.1 *The Great Molasses Flood of 1919 – Boston, Massachusetts (USA)*

On January 15, 1919, a 2.3-million-gallon (~8.7 x 10^6 L) tank of molasses exploded at North End Park in Boston, Massachusetts (Schworm, 2015). The explosion burst the

steel storage tank and created a 25-foot-high (7.6 m) wave of molasses, over 100 yards (91 m) wide, covering a 2-block area, injuring 150 people and killing 21, many of whom were in nearby buildings damaged in the wake of the flow (Schworm, 2015).

Investigation of the cause of the explosion suggested the tank had obvious signs of failed structural integrity, including frequent leaks (Schworm, 2015). Poor steel quality and flawed structural design were cited for the failure of the 58 feet (17.7 m) high by 98 feet (29.9 m) diameter storage tank (Schworm, 2015). It was also suggested that the molasses had been slowly fermenting over 6 months from when the tank was filled the previous July, causing a build-up of carbon dioxide, which in turn raised the internal pressure inside the cylinder (Jabr, 2013). The unseasonably warm winter day may have been enough to cause a spike in bacterial fermentation, leading to the tank's failure (Jabr, 2013). Molasses is a non-Newtonian fluid, so its viscosity is dependent upon applied forces, including stress or increasing the shear rate (Jabr, 2013). The explosive release of molasses created sufficient force to bend steel girders in its path, and move the firehouse from its foundation, before the molasses settled as a sticky syrup that trapped survivors after the initial wave subsided (Schworm, 2015). The spill was 10-times greater than that which affected Honolulu Harbor, and it took over 6 months to clean up seawater pumped in from Boston Harbor, discoloring the harbor waters for months (Basu, 2013).

1.8.2 *Kahului Harbor – Maui, Hawai'i (USA)*

In December 2003, a two-inch (5 cm) hole in a transmission line caused a spill of 50,000 gallons (1.89×10^5 L) of molasses into Kahului Harbor, Maui, Hawai'i (Kubota, 2003). Harbor workers immediately shut down the transmission line after seeing discoloration in the water. After the spill was reported, officials discovered a discrepancy between

the amount pumped from the storage tank and the amount received by the barge (Kubota, 2003). According to the State of Hawai'i Clean Water Branch, molasses was not considered a pollutant, and natural flushing was allowed to remediate the spill (Kubota, 2003). The two-inch hole in the state-owned pipeline caused the spill (Grube, 2013), and cost Kahului Trucking & Storage approximately $15,000 in losses (Kubota, 2003). In 1998, another molasses spill occurred on Maui, when a broken seam in the storage tank allowed 500,000 gallons (1.89×10^6 L) of molasses to be released, but the spill was reportedly contained before reaching the harbor (Grube, 2013).

1.9 Specific Aims

The study described here investigated the effect of a molasses spill on that component of the benthic bacterial community in Honolulu Harbor considered to be 'culturable'. Bacteria isolated on nutrient media were tentatively identified through 16S rRNA gene sequencing to determine if the molasses spill affected the phylogenetic diversity of this bacteria fraction over time since the spill.

The sequence data were used to determine species richness, or the number of Operational Taxonomic Units (OTUs) at the genus-level in a given sample (Schloss & Westcott, 2011), and evenness, or the distribution of individuals among OTUs (Hughes & Bohannan, 2004), between sample time points and sample sites. This study aimed to determine if there is a significant difference in the distribution or representation of OTUs between sample time points and to determine if any such temporal differences may be related to the molasses spill. Sample sites were similarly assessed in terms of community structure variability between sites.

CHAPTER 2. MATERIALS & METHODS

2.1 Sample Area

This was a case study to determine if the input of molasses to Honolulu Harbor altered the relative taxonomic representation in the "culturable" fraction of the benthic microbial community, and if any such alteration would reverse over time. Sediment samples from which to cultivate bacteria were collected from the area impacted by the molasses spill for up to one year, ending in September 2014. Four sample sites were in the harbor, and one was outside.

Samples were collected from five sites in the area impacted by the molasses spill in Honolulu Harbor: Ke'ehi Lagoon Beach Park (A: 21°19'44.87"N, 157°53'51.91"W), La Mariana Small Boat Harbor (B: 21°19'21.56"N, 157°53'36.91"W), Sand Island Launch Ramp (C: 21°18'43.42"N, 157°53'22.51"W), Sand Island Beach Park (D: 21°18'3.48"N, 157°52'41.77"W), and Pier 37 (E: 21°19'2.10"N, 157°52'40.29"W) between Nico's Fish Market and the Fish Auction House. All sites were in the harbor, except the Sand Island Beach Park site, which is directly exposed to the open ocean (Figure 8).

Figure 8. Source of molasses spill at Pier 51B (red triangle) and sediment collection sites (red pins) in Honolulu Harbor, Oʻahu, Hawaiʻi, USA.

2.2 Sample Collection

Samples were collected at each site beginning one week after the spill. Collections continued every four days for one month. Thereafter, samples were taken at two-week intervals for one month, then monthly for three months, with a final sample taken in September 2014, one year after the spill. Sediment push cores on any collection day were taken in triplicate within a 1 m radius of each other directly into metal-free 50 mL Falcon tubes (VWR International; Radnor, PA) from shallow (<1 m) depths. Tubes containing samples were stored on ice in a cooler for transport to the University of

Hawai'i at Mānoa, where they were stored in a refrigerator overnight, before subsampling and preparation for plate counts of culturable bacteria.

2.3 Enumeration of Culturable *Bacteria*

Subsamples of ~2-4 g wet weight of sediment were weighed, then homogenized in 3% w/v sterile saline, and serially diluted (to 1:1000) to determine the number of Colony Forming Units (CFU) per gram of wet sediment. Sediment subsamples were then aseptically diluted in sterile saline for plate counts [*i.e.*, Colony Forming Units per gram of sediment (CFU/g)] on ZoBell's Marine Agar 2216 (MA) and Reasoner's 2A (R2A) (HiMedia Laboratories: Mumbai, India; Difco™, BD Scientific: Franklin Lakes, NJ). A decimal dilution series to 10^{-3} was prepared in sterile 15 mL screw-capped polypropylene centrifuge tubes. For the first samples collected, each plate of nutrient medium was inoculated with 100 µL of the 10^{-3} dilution (effective dilution 1:10000) of the initial 1:1 sediment-in sterile saline (samples #125-129). The dilution series was adjusted thereafter in an effort to optimize counts to the range of 30-300 CFU/g colony counts for all samples; subsequent plates were inoculated with 100 µL of the 10^{-2} dilution (effective dilution of 1:1000). The remainder of each undiluted sediment sample was stored at -80°C. Control plates were inoculated with 100 µL from 9 mL of the 3% sterile saline used to prepare the dilution series of the sediment samples.

2.4 Isolation of *Bacteria*

CFU were counted on inoculated plates after two weeks of incubation at room temperature (~25°C). However, plates were checked weekly for the presence of agarolytic bacteria; CFU were counted and subcultured immediately if such activity was observed. Colonies considered representative of others on each plate on the basis of

color, luster, texture, margin, elevation, size (diameter in mm), and opacity were transferred from the original plate onto the same medium (MA or R2A). Such isolated colonies were subsequently described on the basis of the characteristics described above and presumed to be pure cultures. Pure cultures were prepared for long-term storage in cryovials (Table 1). Cells from a single colony of each pure culture were used in PCRs for amplification of a fragment of their 16S rRNA gene using the *Bacteria*-specific primer 27F and the universal primer 1492R (Table 2).

Table 1. Cryovial solution for long-term storage.

Cryovial Solution	
Marine Broth	2:1:1 Marine Broth:Glycerol:3% w/v NaCl
R2A Broth	2:1:1 R2A Broth:Glycerol:0.5% w/v NaCl

2.5 Identification of Pure *Bacteria* Cultures

Cells from a single colony of each pure *Bacteria* culture were collected on a sterile toothpick and aseptically suspended into 49 µL of PCR buffer in a PCR tube. Each PCR was performed in 50 µL reactions containing 23.5 µL 2X Promega (Madison, WI) GoTaq Green Master Mix (M712), 23.5 µL water, 1 µL (10 µM) 27F, and 1 µL (10 µM) 1492R (Table 2) (Klindworth et al., 2012). The PCR protocol comprised an initial "hot start" of 2 min at 95°C to denature the DNA, followed by 30 cycles of amplification (95°C for 30 sec), annealing (51°C for 30 sec), and elongation (72°C for 2 min), with an additional 2 min elongation step at 72°C prior to cooling at 4°C. PCRs were run in an Eppendorf Mastercycler gradient PCR machine (Eppendorf: Hamburg, Germany).

Table 2. Primers for PCR and Sanger sequencing (Weisberg et al., 1991).

Primer	Sequence	T_m	Length (nt)
27F	5′ – AGA GTT TGA TCM TGG CTC AG – 3′	53.2°C	20
1492R	5′ – TAC GGY TAC CTT GTT ACG ACT T – 3′	54.6°C	22

PCR products were visualized on a 1% agarose gel (110 V, 60 mA, 20 minutes). To reduce the number of poor quality reads, the samples were run both post-PCR and post-PCR clean-up to visually determine if the DNA concentration was satisfactory for sequencing.

Cultures for which no amplification product was visible were run through a second PCR using whole cells, as described above. If this second PCR also failed, genomic DNA was extracted from a cell pellet using the MoBio UltraClean® Microbial DNA Isolation Kit, or the MoBio Power Soil DNA Isolation Kit (MoBio Laboratories, Inc., Carlsbad, CA).

Of 1420 total sequences, 1121 sequenced from PCR product; 53 were derived from DNA extracted after initial PCR as described above failed. An additional 35 duplicate samples run for quality control between sequencing runs were excluded. A total of 157 cultures were not used for further analysis because they provided only low-quality reads that failed to amplify for sequencing or did not pass the screening parameters; among these, 89 were potentially mixed cultures. PCR products were cleaned in the MoBio UltraClean® PCR Clean-Up Kit (MoBio Laboratories, Inc.: Carlsbad, CA), and visualized on a 1% agarose gel.

2.6 DNA Sequencing & Quality Control

Pure PCR products were submitted for Sanger sequencing in the University of Hawai'i at Mānoa's core facility, Advanced Studies of Genomic, Proteomics, and Bioinformatics (ASGPB) using only the 27F *Bacteria*-specific primer.

Nucleotide sequence data were analyzed using the original base calls and Q-values (i.e., quality values) in the KB™ Base Caller v1.1.1 (ABI, 2004) and compared to Phred (Ewing et al., 1998) base calls using CodonCode Aligner (CodonCode Corp: Centerville, MA). Sequence chromatograms were visually inspected for accuracy of the base-calling software and edited manually using CodonCode Aligner. Q-values reported by KB™ Base Caller conform to the industry-standard established by Phred and are interchangeable for downstream pipelines. Sequence quality control was performed using CodonCode Aligner, which allows for trimming of sequences according to standardized or user-defined parameters, to generate FASTA and associated quality files. Sequences were trimmed using the following parameters, with user-defined conditions for optimized end clipping (CodonCode Aligner Corp., 2015):

- Trim from Start Until: There are fewer than (1) base with a quality of (20) in a (2) base window.
- Trim from End Until: There are fewer than (4) bases with a quality of (20) in a (20) base window.
- After End Clipping: Move all sequences shorter than (25) bases to the trash. Move all sequences with fewer than (50) Phred 20 bases to the trash.

Sequences were checked for chimeric sequences, or hybrid sequences created from mixed cultures or PCR artifacts (Haas et al., 2011), using DECIPHER (Wright, 2012), an on-line tool (decipher.cee.wisc.edu).

2.7 Identification and Taxonomic Assignment of 16S rRNA Amplicons

Curated sequences from 1086 cultures were aligned and grouped into Operational Taxonomic Units (OTUs) using the Ribosomal Database Project (RDP-II) Naïve Bayesian rRNA Classifier v2.6 (Cole et al., 2014) on-line web tool "Classifier" (rdp.cme.msu.edu/classifier/classifier.jsp) with a bootstrap value of 80% (i.e., confidence threshold) with a recommended minimum 250 bases, with an absolute minimum of 50 bases, and assigned to taxonomic groups (Wang et al., 2007). RDP-II uses a 7-mer query algorithm to match the highest number of 7-mers to the nearest neighbors in the sequence database compared to BLAST, based on alignment of 250 bases to generate E-values (Cole et al., 2014). To compare classification methods, sequences were also aligned and taxonomically classified according to the SILVA (Pruesse et al., 2007; Quast et al., 2013) and RDP-II (Maidak et al., 1997) databases with the SINA Alignment Service v.1.2.11 (http://www.arb-silva.de/aligner/) using the least common ancestor method with a minimum identity of 97% (Pruesse et al. 2012). Both SILVA and RDP-II databases are quality controlled. SILVA also provides the option of searching cultivated (including type) strains only (Pruesse et al., 2012).

FASTA and FASTAqual files generated from CodonCode Aligner were imported into mothur v.1.36.1 (University of Michigan) for alignment to the SILVA Release v.119, a database of 137,879 full-length bacterial sequences for alignment and taxonomic reference (Schloss et al., 2009). The sequences were aligned to the SILVA Release v.119 using kmer (*i.e.*, 8-mer) searching and the Needleman-Wunsch pairwise alignment method. The resultant file was filtered to remove gaps in order to create a distance matrix in Phylip format (Felsenstein, 1989), and cluster the OTUs at a distance threshold of 0.03 using the average-neighbor algorithm. The "clearcut" command was

used to generate five (5) phylogenetic trees using a stringent bootstrap value (*n* = 1000), with relaxed Neighbor-Joining method (Evans et al., 2006), a modified version of the Neighbor-Joining method (Saitou & Nei, 1987) that is less computationally intensive than Maximum Likelihood. The output file was imported into FigTree, a tree-building software program (A. Rambaut, University of Edinburgh), that allows visual representation of the data through color-coding of OTUs (Appendix 6).

The distance matrix and tree files were used in analyses through which the frequency of the distribution of sequences, or level of similarity in OTU "bins", were determined (Schloss et al., 2009). OTUs were grouped by sample time point and by sample site. To compare sample sites, an additional "site" was created to incorporate contaminants recovered and isolated from the control plates. Venn diagrams illustrating overlap between sites were generated based on the basis of results of the unweighted UniFrac analysis. A heat map illustrating overlap between OTUs was generated for both sample time points and sample sites. The "tree.shared" command generated trees showing the relatedness between sample time points and sample sites. Rarefaction curves were generated to determine if the level of sampling effort sufficiently captured each level of taxonomic association, specifically, phylum, genus, and species (Hughes & Bohannan, 2004).

mothur was used to run hypothesis-based analyses to determine if the difference between groups, i.e., sample time point or sample site, was significant, and thus to determine if phylogenetic diversity in the microbial community diversity associated with the molasses spill changed with time and at each sampling point. The "parsimony" command was run to determine the likelihood that the communities have the same structure by chance, through generation of multiple random-joining trees (*n* = 10,000

randomizations) and scoring each tree to generate a distribution (Slatkin & Maddison, 1990). The weighted UniFrac analysis was then run to quantitatively determine through pair-wise comparisons of randomly generated trees (n = 1000) if any of the groups structurally differ from others (Lozupone et al, 2007). The unweighted UniFrac analysis was run as a qualitative measurement of the evolutionary distance between a pair of communities as measured by unique evolution in each community compared on the basis of branch length (Lozupone and Knight, 2005). This is a useful statistic when testing spatial or temporal relatedness between communities.

CHAPTER 3. RESULTS

3.1 Cultivated *Bacteria* Abundance

Serially diluted sediment samples from 5 sites in Honolulu Harbor collected between September 15, 2013 and September 17, 2014, covering the time span between 8 and 375 days after the molasses spill, were plated to Marine Agar and R2A for enumeration of CFU. CFU on each medium were counted after a two-week incubation at room temperature. Representative colonies were streaked for purification to the same medium. Pure cultures provided material for amplification and sequencing of a fragment of their 16S rRNA gene.

Higher CFU/g were recorded on Marine Agar (MA) than on R2A at all time points except at 127-days post-spill (T12: 01/12/14) (Figure 9). At 127-days post-spill, the count on R2A was 14.2% greater than the mean CFU/g on MA at all sites for that sampling. The highest CFU/g for all sites were determined at the first sample time point (T1: 09/15/13), i.e., 8-days post-spill, on both Marine Agar and R2A (Figure 10 & Figure 11).

CFU counts in samples collected in the vicinity of the Sand Island Launch Ramp peaked at 1.55E x 10^5 CFU/g on R2A at 20-days post-spill (T4: 09/27/13), which was within the range of the confidence interval seen at 8-days post-spill (T1: 09/15/13) for which the count was 8.25E x 10^4 CFU/g. Sand Island Launch Ramp also had increased counts at 24-days post-spill (T5: 10/01/13) and 127-days post-spill (T12: 01/12/14).

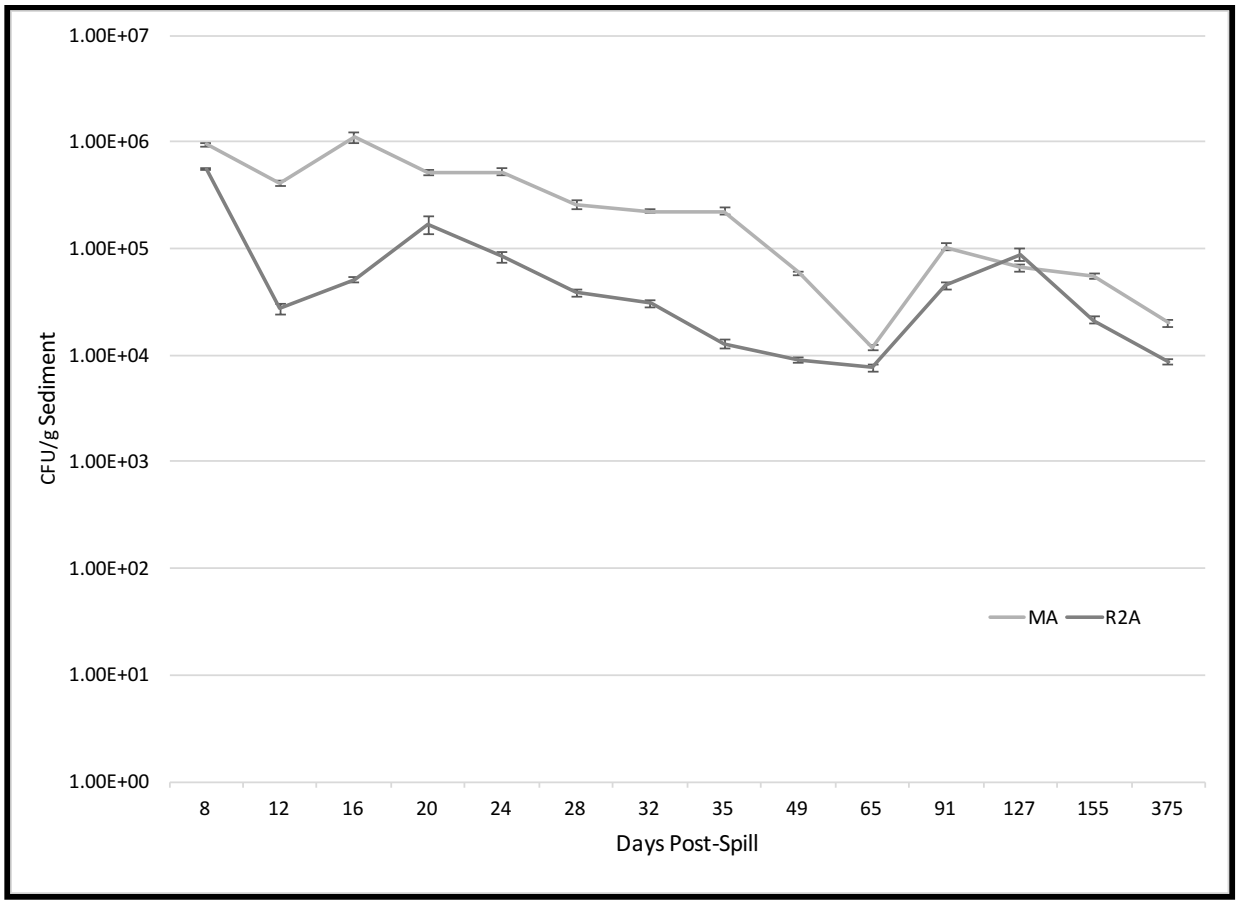


Figure 9. Mean CFU/g sediment counts for all 5 sites on MA and R2A at each time point. Bars represent standard error of the means.

CFU/g in samples collected at the third sampling point, i.e., 16-days post-spill (T03: 09/23/13), showed an increase at La Mariana and Pier 37 on Marine Agar (Figure 10), after a decrease in CFU/g on Marine Agar at 12-days post-spill (T02: 09/19/13). Counts of CFU on MA stabilized to an assumed *a posteriori* background level by 49-days post spill (T09: 10/26/13), with no site exceeding 100,000 CFU/g two months after the spill (Figure 10).

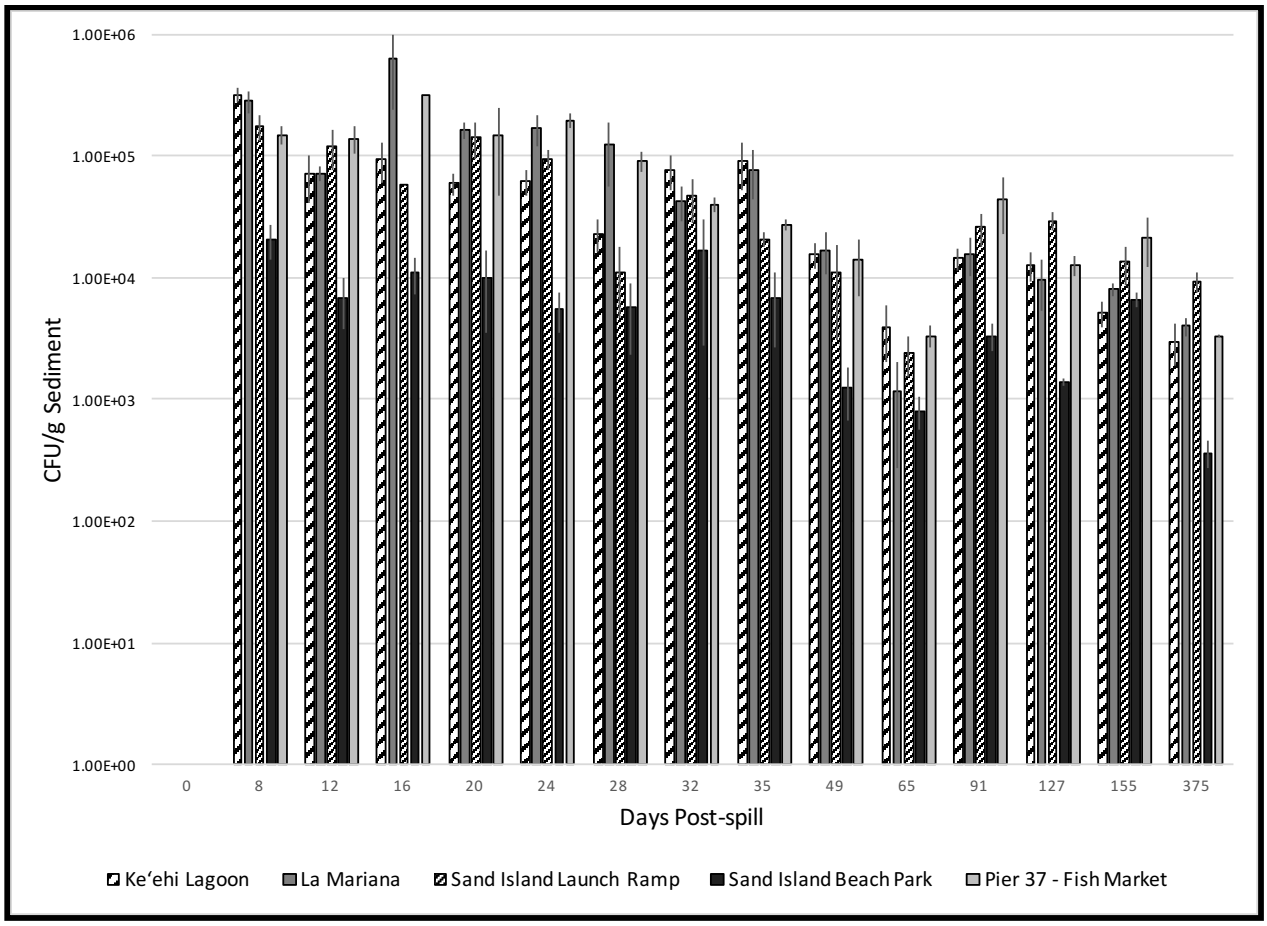


Figure 10. CFU/g sediment on MA. Bars represent standard error of the means. (See App. 6 for mean CFU/g sediment on MA.)

After the initial sampling 8-days post-spill (T1: 09/15/13), CFU from the Sand Island Beach Park site did not exceed 1000 CFU/g on R2A. There was a general trend for CFU/g on R2A at all sites to decrease after the initial sampling at T1, except for at the Sand Island Launch Ramp Site (Figure 11), where the CFU/g peaked at 1.55E+05 CFU/g at 20-days post-spill (T4: 09/27/13). Increased CFU/g on R2A were counted at 24-days post-spill (T5: 10/01/13; CFU/g=5.36E+04) and at 127-days post-spill (T12: 01/12/14; CFU/g=5.71E+04) at Sand Island Launch Ramp.

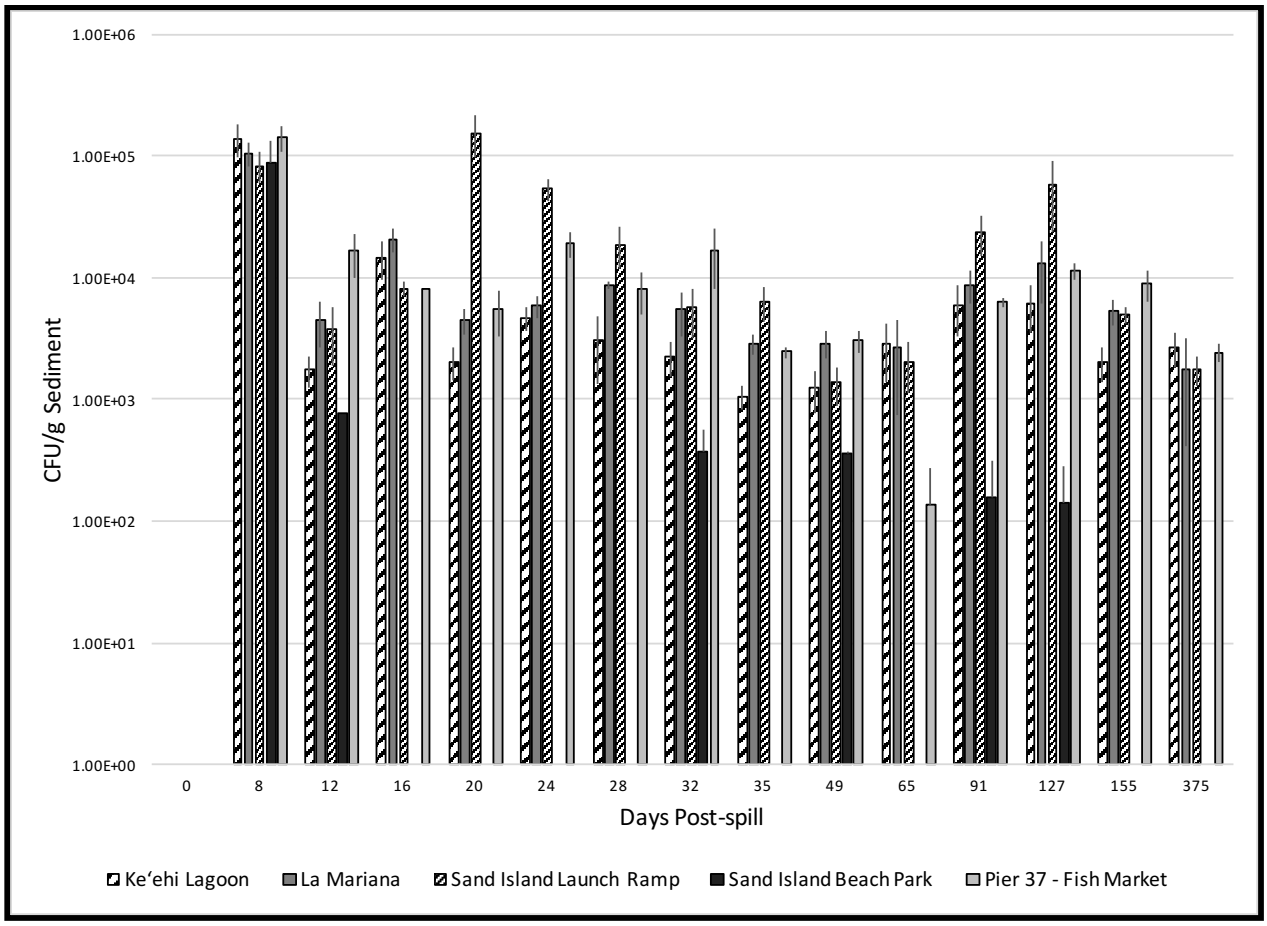


Figure 11. CFU/g sediment on R2A. Bars represent standard error of the means. (See App. 7 for mean CFU/g sediment on R2A.)

3.2 16S rRNA Gene Sequencing & Phylogenetic Analysis

The ABI 3730xl DNA Analyzer utilizes the KB™ Base Caller v1.1.1 (ABI, 2004), an incorporated software program that determines base identity. The KB™ Base Caller program was used here rather than the Phred base call by CodonCode Aligner because the former program is better at resolving base calls. A total of 1087 16S rRNA gene sequences from pure cultures passed quality control in the CodonCode Aligner software with the aforementioned parameters (Section 2.6).

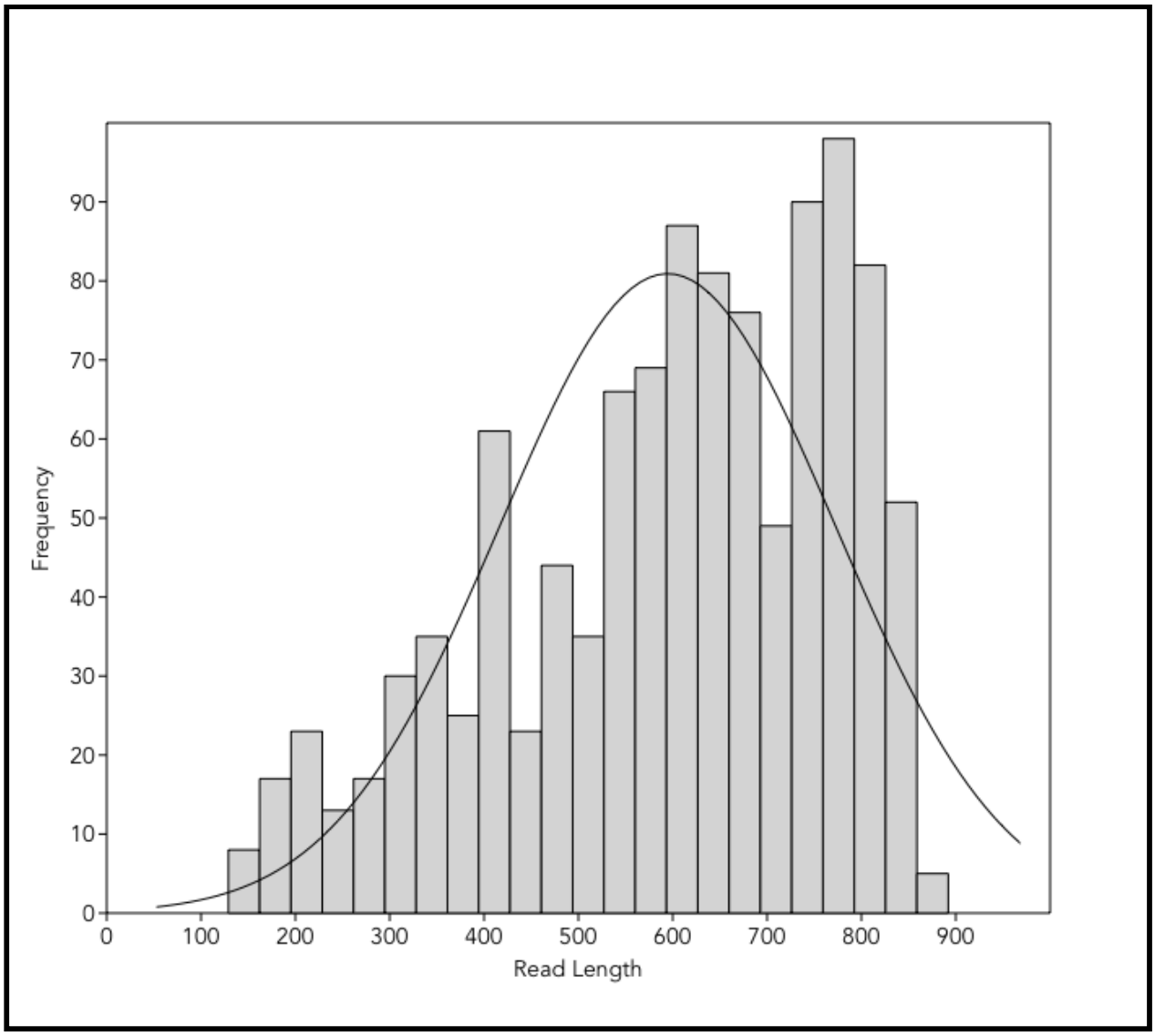


Figure 12. PCR amplicon read length with normalized frequency distribution curve.

No chimeras were detected using DECIPHER, although two sequences were considered "indecipherable" because they were classified as part of a group that lacked substantial reference sequences to be classified by the program: A31 (nearest neighbor: *Kiloniella* sp.) and B30 (nearest neighbor: *Cohaesibacter* sp.).

The 1086 sequences that passed quality control were analyzed using mothur for read length; read lengths varies from 129 to 892 nt (Figure 12). Median read length was 620

nt; mean read length was 595 nt (SD=178). Sequences smaller than the recommended 250 base pair for classification (Claesson et al., 2009) were retained for analysis if they passed the quality control parameters set in CodonCode Aligner and visual inspection. The shortest sequence assigned a taxonomic position was 129 nt (tentative *Bacillus* sp.); a total of 57 sequences <250 nt were assigned a taxonomic position.

Sequences were aligned and de-gapped in mothur to create a distance matrix in Phylip format for generation of phylogenetic trees. This program is optimized for short reads generated by next-generation sequencing, including the Illumina and 454-pyrosequencing platforms (Schloss et al., 2009), but can also be used with Sanger data with sample tutorials available on-line (www.mothur.org/wiki). The first of five trees was selected as a representative tree for visual representation of OTUs with color coding (App. 8) in FigTree. The distance matrix and tree files were used for additional OTU-based or hypothesis-based analysis in mothur. To incorporate sequences from isolates recovered from control plates, an additional "site" was created for the comparison of sample sites in the group file [*Bacillus* spp. (n = 3) and *Vibrio* spp. (n = 2)].

OTUs were grouped by sample date or location for unweighted UniFrac analysis. The following comparisons were made using Venn diagrams to illustrate which OTUs were common to two or more sample time points or sample sites: A-C-D-E (App. 9), B-C-D-E (App. 10), A-B-C-E (App. 11), T01-T03-T05-T10 (App. 12), T01-T03-T10-T12 (App. 13), and T01-T02-T03-T12 (App. 14). A heat map for sample time points and sample sites was generated to illustrate the 200 most commonly identified OTUs for each site and by sample time point (Apps. 15 & 16). A rarefaction curve was generated for OTUs based on distance thresholds of 0.03 (species), 0.05 (genus), and 0.10 (phylum) to determine if the sampling effort provided adequate coverage (Figure 13). The

rarefaction curve generated for all isolates shows sampling effort was adequate at the phylum level, as the curve is leveling off, but was insufficient at the genus and species level (Figure 13) (Hughes & Hellmann, 2005).

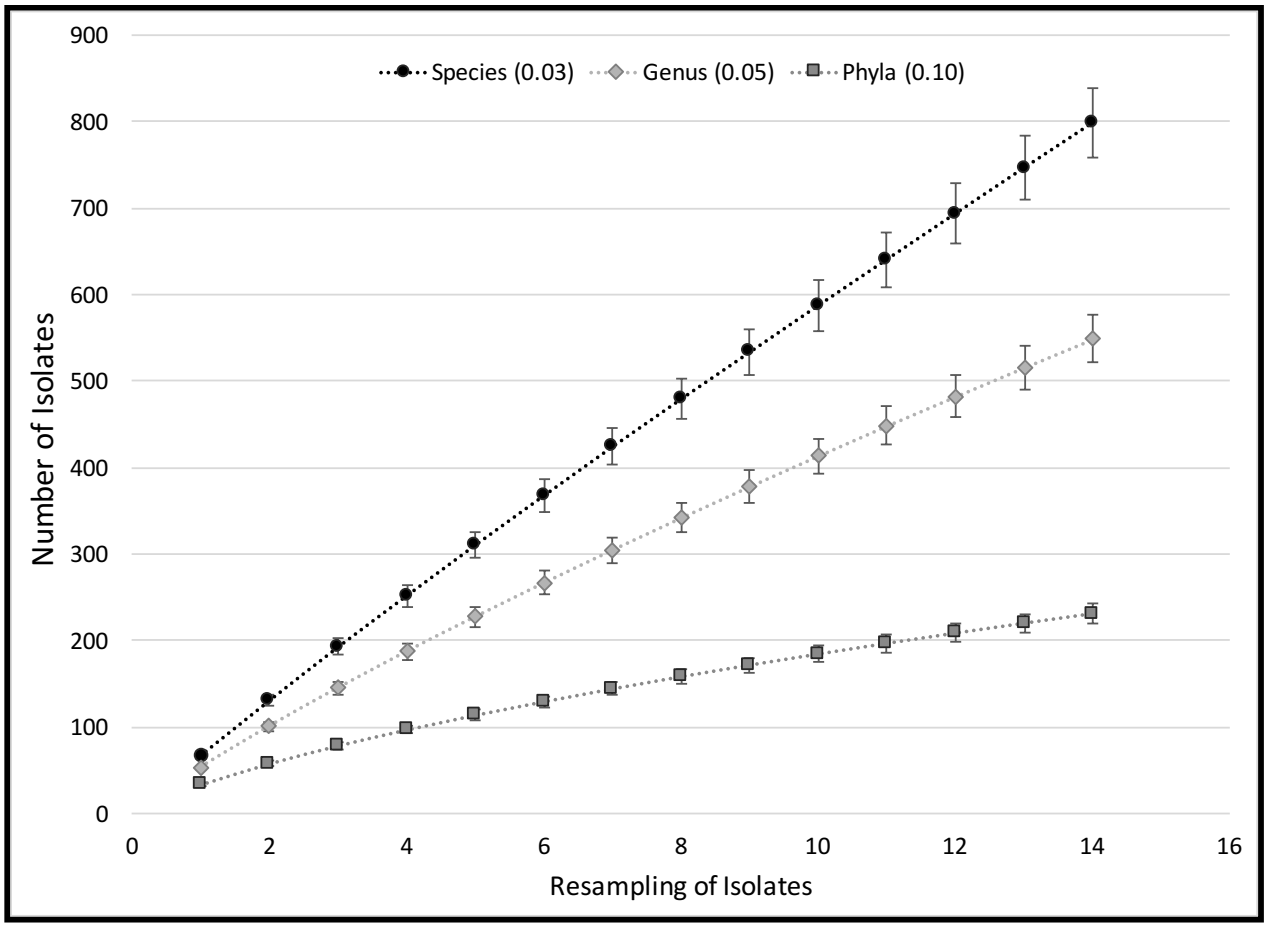


Figure 13. Rarefaction of all sequences at distance thresholds of 0.03 (species), 0.05 (genus), and 0.01 (phyla) with standard error bars (5%).

Various tests were performed to determine if microbial diversity in terms of phylogenetic affiliation of OTUs (from cultivated Bacteria) varied significantly between sample time points or sample sites after the molasses spill. The following reference pipelines were used to for analysis: Esophageal Community Analysis and MiSeq SOP (Kozich et al., 2013). Parsimony analysis requires correction for multiple comparisons to

integrate the number of comparisons relative to the significance threshold ($P \leq 0.05$). For parsimony analysis between sample time points ($n = 14$), the following correction was used to determine the significance threshold: 0.05/14 = 0.0036. This calculation gives the corrected P-value for significance ($P \leq 0.0036$) (Table 3). Parsimony analysis of sample time points showed that T01 and T02 differed significantly from all other time points in terms of community structure (App. 17).

Table 3. Parsimony analysis of sample sites ($P \leq 0.0083$) and determination of the effect of spatial isolation.

Sample Site	A	B	C	D	E	F
A		<0.0001	<0.0001	<0.0001	<0.0001	1
B	<0.0001		<0.0001	<0.0001	<0.0001	0.0278
C	<0.0001	<0.0001		<0.0001	<0.0001	1
D	<0.0001	<0.0001	<0.0001		<0.0001	0.0735
E	<0.0001	<0.0001	<0.0001	<0.0001		1
F	1	0.0278	1	0.0735	1	

Use of mothur to perform parsimony analysis offers insight into the statistical probability of spatial relatedness between sample sites. A correction factor (0.05/6 = 0.0083) was used to account for multiple comparisons, which provides the corrected P-value for significance ($P \leq 0.0083$) (Schloss et al., 2009). Parsimony analysis of sampled sites in terms of OTUs present at each site showed they differed significantly with respect to each other site, with the exception of the contaminants recovered from control samples, grouped as "Site F" for analysis (Table 3), which was not significantly

different from any site. This method recommends a minimum of 10 randomly selected individuals per location for comparison, with better resolving power provided by more locations (Slatkin & Maddison, 1990). Thus, comparing OTUs at specific sampling sites to the 'contaminant' OTUs (n = 5) lacks the sampling depth necessary for this statistic.

The weighted UniFrac analysis determined that all time points and all sites differed significantly from each other ($P \leq 0.001$) in terms of community (i.e., OTU) structure (Schloss et al., 2009). This qualitative analysis that accounts for differences in relative abundance that may skew the statistic based on sampling effort by assigning equal weight to all OTUs (Luzopone et al., 2011). The unweighted UniFrac analysis was used as a quantitative statistic to determine if the OTUs in samples are significantly different; this test measures the fractional branch length as a means of estimating evolutionary divergence between two communities (Luzopone et al., 2011). This statistic is sensitive to sampling depth based on small sample sizes from communities with diverse assemblages, but that does not negate the usefulness as a distance metric (Luzopone et al., 2011).

According to the unweighted UniFrac analysis, the paired time points were significantly different ($P \leq 0.001$) for all five trees generated in terms of their component OTUs: T01-T03, T01-T05, T01-T10, T01-T12, and T02-T12. With respect to sample sites, the all sites were considered significantly different from each other ($P \leq 0.001$) in terms of component OTUs, except Ke'ehi Lagoon (A) and La Mariana (B), which were similar to each other.

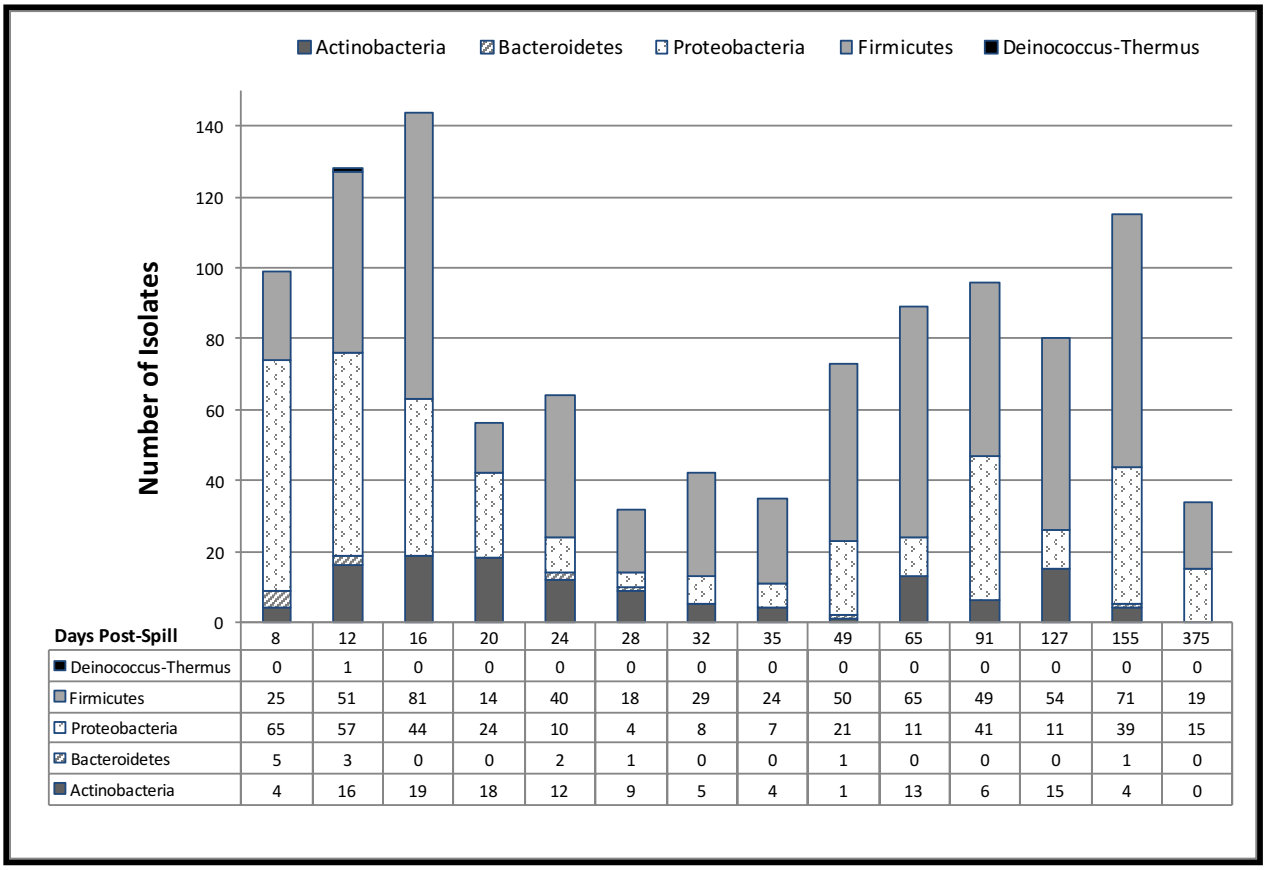
3.3 Phylogenetic Assignment of OTUs

RDP and SILVA databases use slightly different nomenclature in their taxonomic assignments, but the putative taxonomies assigned by both databases were consistent for all sequences to the level of genus. RDP and SILVA were used preferentially over the NCBI GenBank database because they are peer-reviewed and curated databases (Cole et al., 2014; Pruesse et al., 2007). The RDP assignments are shown later (Apps. 16 & 17). Five phyla were represented among the 1086 OTUs: *Firmicutes* (*n* = 590), *Proteobacteria* (*n* = 357), *Actinobacteria* (*n* = 126), *Bacteroidetes* (*n* = 12), and *Deinococcus-Thermus* (*n* = 1). The most common genera assigned were *Bacillus* (*n* = 452), followed by *Vibrio* (*n* = 150), and *Microbacterium* (*n* = 22) (Apps. 18 & 19). The putative OTU assignment of the isolates grouped into a total of 799 OTUs: *Firmicutes* (*n* = 413; 70% different), *Proteobacteria* (*n* = 269; 75.4% different), *Actinobacteria* (*n* = 105; 83.3% different), *Bacteriodetes* (*n* = 11; 91.7% different), and *Deinococcus-Thermus* (*n* = 1; 100% different).

The number of sequences affiliated with *Proteobacteria* and *Bacteriodetes* increased at 12-days post-spill, followed in subsequent days by increases in those affiliated with *Firmicutes* and *Actinobacteria*. *Bacteroidetes*–affiliated bacteria were not recovered from samples collected at Ke'ehi Lagoon or Sand Island Beach Park. A single *Deinococcus-Thermus* was cultivated during the second sampling at Sand Island Beach Park. Over the long-term, the majority of the cultures isolated on each sampling day were *Firmicutes*, with lower but varying numbers of *Proteobacteria*-affiliated cultures.

The third sampling, 16-days post-spill (T03: 09/23/13), showed an increase in CFU/g on MA at La Mariana and Pier 37 on Marine Agar was seen approximately 5-days after the

HI-DoH reported dissolved oxygen levels had returned to normal levels (Figure 4 & Figure 5), with counts on MA stabilized by 49-days post-spill (T9: 10/26/13).

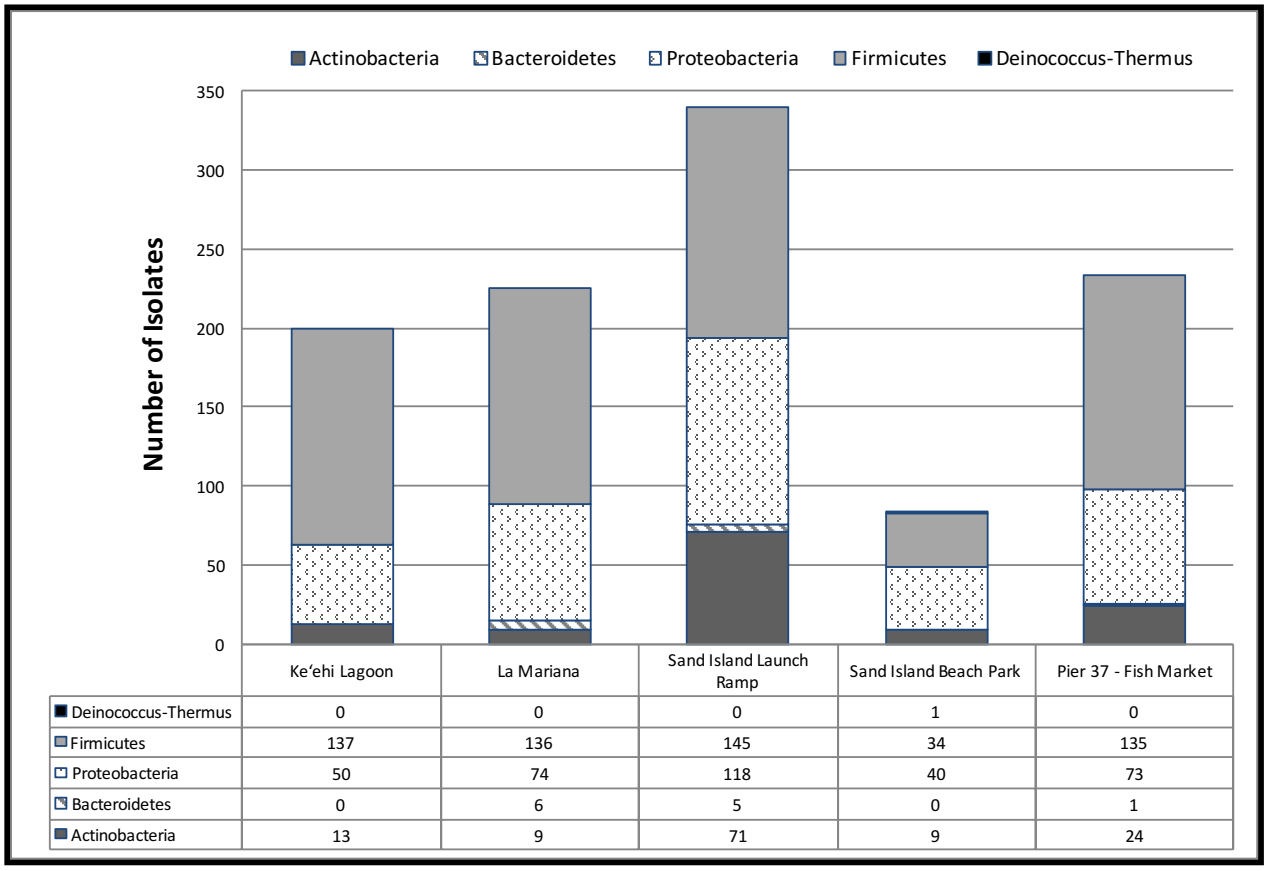


Days Post-Spill	8	12	16	20	24	28	32	35	49	65	91	127	155	375
Deinococcus-Thermus	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Firmicutes	25	51	81	14	40	18	29	24	50	65	49	54	71	19
Proteobacteria	65	57	44	24	10	4	8	7	21	11	41	11	39	15
Bacteroidetes	5	3	0	0	2	1	0	0	1	0	0	0	1	0
Actinobacteria	4	16	19	18	12	9	5	4	1	13	6	15	4	0

Figure 14. Phyletic distribution of isolates from all sites by number of days post-spill.

From 24-days post-spill through all other sample points, most cultivated bacteria belonged in the *Firmicutes*, with occasional increases in the degree of representation by the *Proteobacteria*. During the first two samplings (8- and 12-days post-spill) there was a decrease in *Bacteriodetes* (*n* = 5, 3, respectively), with occasional isolation of *Bacteroidetes* at later time points. The fraction of culturable *Actinobacteria* in the community appeared inversely related to that of *Proteobacteria*. Just one *Deinococcus-Thermus* was cultivated throughout the entire project (at 12-days post-spill) from Sand

Island Beach Park. At the final sample point one-year post-spill (375-days post-spill), only *Firmicutes* and *Proteobacteria* were cultivated (Figure 14).

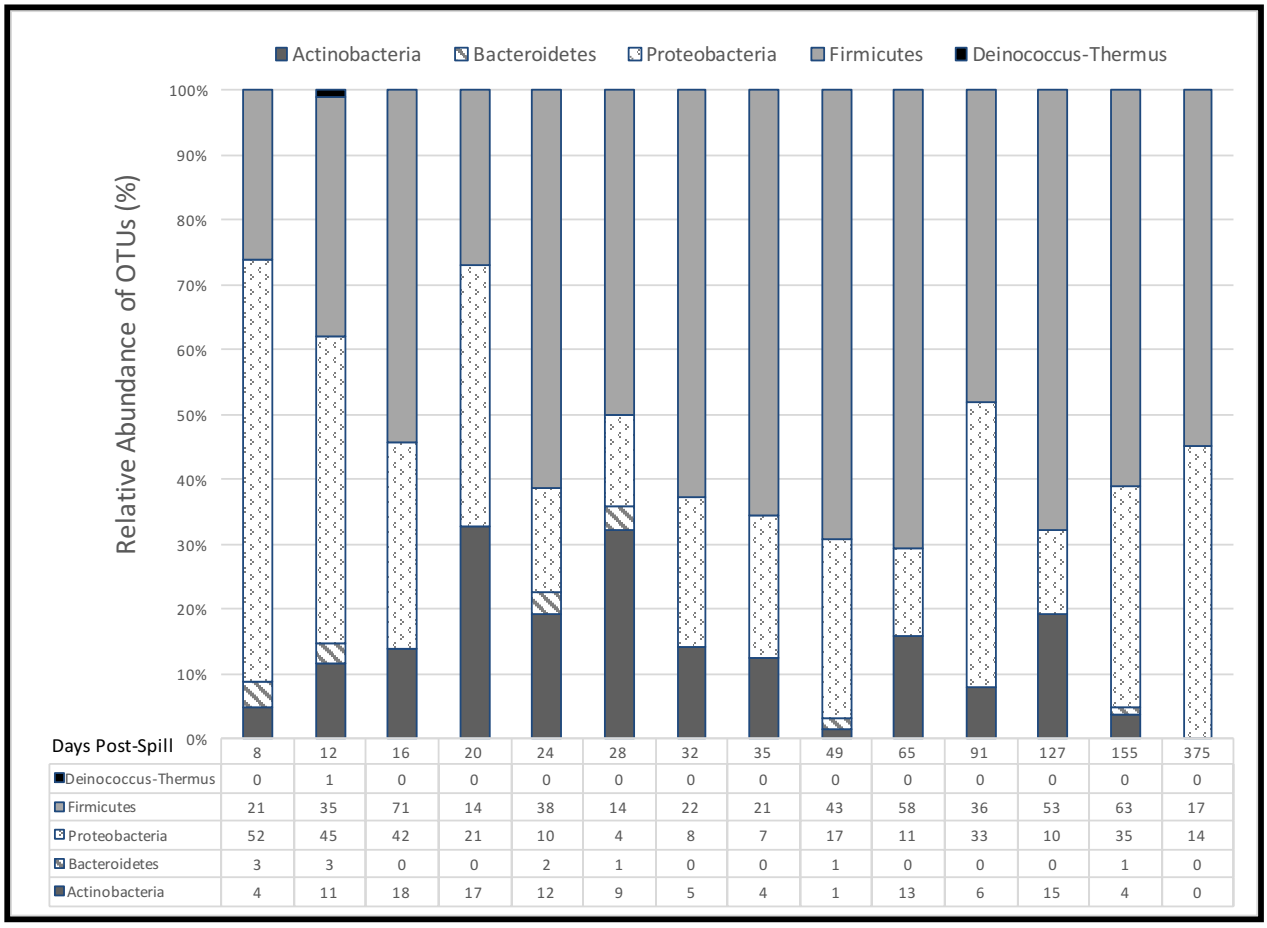


	Ke'ehi Lagoon	La Mariana	Sand Island Launch Ramp	Sand Island Beach Park	Pier 37 - Fish Market
Deinococcus-Thermus	0	0	0	1	0
Firmicutes	137	136	145	34	135
Proteobacteria	50	74	118	40	73
Bacteroidetes	0	6	5	0	1
Actinobacteria	13	9	71	9	24

Figure 15. Phyletic distribution of all isolates for each site over the duration of sampling.

The majority of different isolates defined on the basis of the nucleotide sequence of a fragment of their 16S rRNA gene originated from the Sand Island Launch Ramp (*n* = 339), followed by Pier 37 (*n* = 223), La Mariana (*n* = 225), and Ke'ehi Lagoon (*n* = 200) (Figure 15). Fewest originated from the Sand Island Beach Park (*n* = 84). *Firmicutes*, *Proteobacteria*, and *Actinobacteria* were recovered at all sites. Isolates of *Bacteroidetes* were only recovered from La Mariana, Sand Island Launch Ramp, and Pier 37. One *Deinoccus-Thermus* was recovered from Sand Island Beach Park. The identification of a

Deinococcus sp. from the Sand Island Beach Park from T02: 09/19/13 was likely a chance recovery of a rare species.

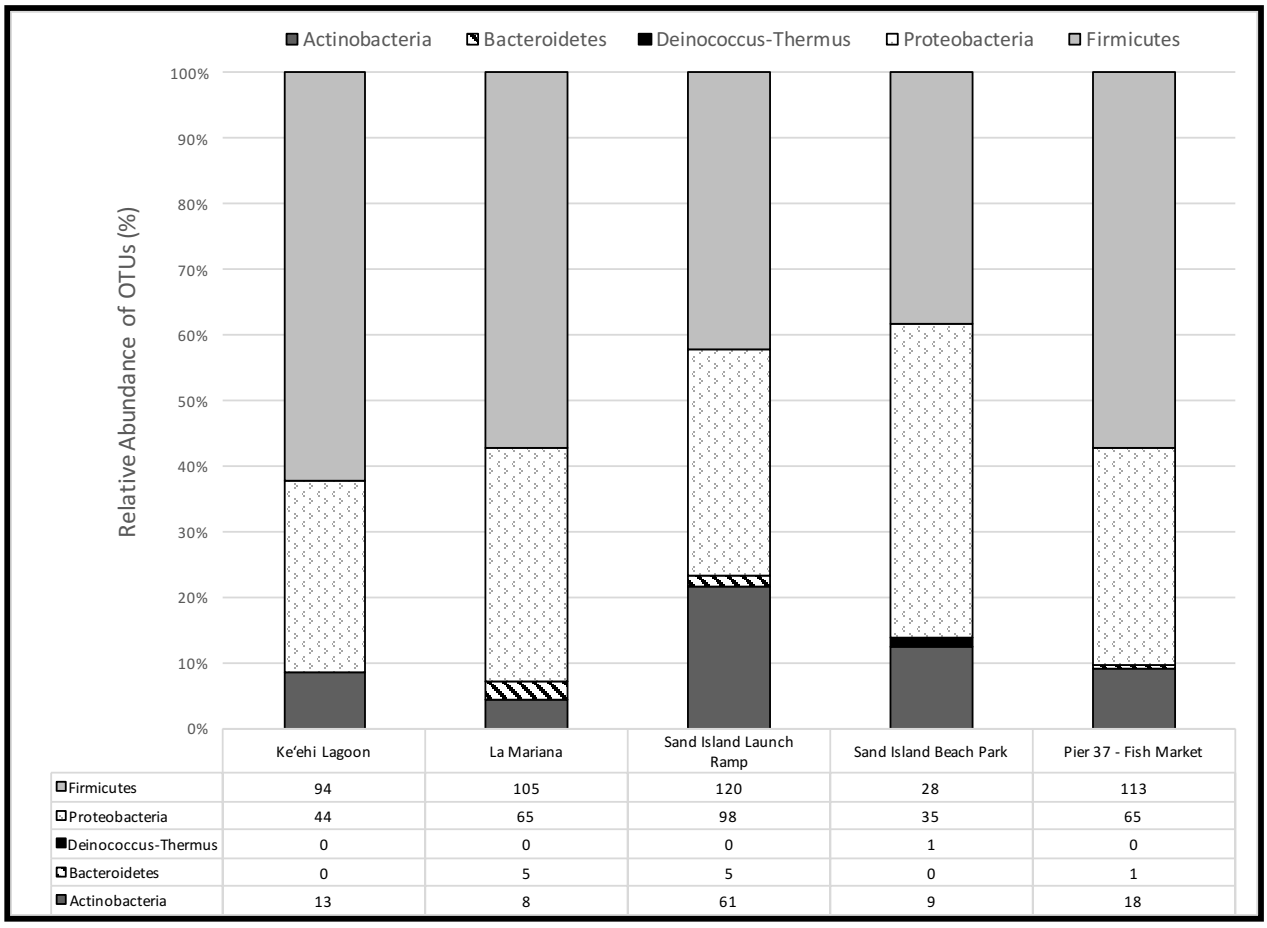


Days Post-Spill	8	12	16	20	24	28	32	35	49	65	91	127	155	375
Deinococcus-Thermus	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Firmicutes	21	35	71	14	38	14	22	21	43	58	36	53	63	17
Proteobacteria	52	45	42	21	10	4	8	7	17	11	33	10	35	14
Bacteroidetes	3	3	0	0	2	1	0	0	1	0	0	0	1	0
Actinobacteria	4	11	18	17	12	9	5	4	1	13	6	15	4	0

Figure 16. Relative abundance and phyletic assignments of OTUs (%) from all sites by days post-spill.

Of 1086 *Bacteria* isolates tentatively identified on the basis of a fragment of their 16S rRNA gene nucleotide sequence (OTUs), 799 unique OTUs (73.6%) were classified by mothur at a 0.03 distance threshold for 'species'. The greatest number of different OTUs in such terms was 131 isolates, or 91% of 144 sequenced, from samples collected 16-days post-spill (T03: 09/23/13) (Figure 16). The highest percentage of different OTUs among those recovered were from samples taken 24-days post-spill (T5:

10/01/13), with 62 of 64 isolates (96.9%) being so defined. The lowest fraction of different OTUs was determined in samples collected 12-days post-spill (T2: 09/19/13), with 95 of 128 unique isolates (74.2%). The fewest different OTUs *senso stricto* was recovered from samples taken 28-days post-spill (T06: 10/05/13), with 28 of 32 unique OTUs (87.5%). An average of 78 *Bacteria* isolates were cultivated at each time point, which represented 68 OTUs. The median number of *Bacteria* isolates per time point was 77, comprising 69 different OTUs.



	Ke'ehi Lagoon	La Mariana	Sand Island Launch Ramp	Sand Island Beach Park	Pier 37 - Fish Market
Firmicutes	94	105	120	28	113
Proteobacteria	44	65	98	35	65
Deinococcus-Thermus	0	0	0	1	0
Bacteroidetes	0	5	5	0	1
Actinobacteria	13	8	61	9	18

Figure 17. Relative phyletic distribution of OTUs (%) from all samples by site.

Firmicutes dominated the total number of OTUs represented by cultures from all sites, with the exception of Sand Island Beach Park, at which *Proteobacteria* comprised the

majority of OTUs represented (Figure 17). Most *Actinobacteria* OTUs were isolated from Sand Island Launch Ramp (n = 61).

CHAPTER 4. DISCUSSION

4.1 Effects of a Molasses Spill on Benthic Heterotrophic *Bacteria*

Surface sediment cores provided material for a time-series analysis of the effects of a molasses spill on culturable aerobic and non-obligately anaerobic benthic *Bacteria* in Honolulu Harbor over a period of 375 days. The two media used were selected to cultivate heterotrophic marine bacteria and non-obligately halophilic bacteria.

Plate counts and taxonomic assignments of the *Bacteria* cultivated suggest that the molasses spill caused a temporary shift in the composition of the culturable *Bacteria* community. For example, the fraction of OTUs which affiliated with *Proteobacteria* and *Bacteriodetes* increased in the 12-days following the molasses spill, and this abundance subsequently declined through the remainder of the investigation (Figure 16).

Honolulu Harbor receives freshwater inputs from Kāpalama Stream with two outfalls draining into Kāpalama Channel near Pier 39; this freshwater flows predominantly through Kalihi Channel into Māmala Bay near La Mariana and the Sand Island Launch Ramp. Nu'uanu Stream also provides freshwater input into the main harbor channel near Pier 15 with 19 outfalls draining into the area (HI-DoT, 2012c). Storm drain discharging at various points around the harbor from commercial, industrial, and residential areas complicate source-tracking of up-stream pollution (HI-DoH, 2013c). Indeed, harbor water regularly fails to meet Hawai'i State water quality standards for nitrogen, phosphorus, and turbidity, and that significant levels of copper, zinc,

chromium, nickel, lead, chlordane, and dieldrin have also been reported (Hawai'i DoH, 2013c).

Ke'ehi Lagoon is the largest lagoon in the State, and receives freshwater inputs from Kalihi and Moanalua Streams (HI-DoH, 2013c). Water quality in the lagoon is poor, in term of elevated phosphorus concentration and turbidity; currents may also transport pollutants from Honolulu Harbor into the lagoon (HI-DoH, 2013c). Drainage from Moanalua and Kalihi Streams includes suspended sediment, trash, and plant debris, and storm runoff from the industrial areas of Mapunapuna, Shafter Flats, Kāpalama, and Kalihi Kai, which bring additional pollution into the lagoon (HI-DoH, 2013c). Cultivated bacteria in sediment samples from Ke'ehi Lagoon were dominated by *Bacillus* sp., followed by *Gammaproteobacteria* (e.g., *Vibrio* sp.) and *Actinobacteria*. This is expected in an area that receives turbid freshwater inputs and undergoes saltwater mixing in a brackish environment (Kirchman et al., 2005). Members of the *Actinobacteria* are a diverse group of high G+C Gram-positive bacteria, and are found in both terrestrial and aquatic environments, including marine ecosystems, and are critical in recycling refractory biogenic material (Ventura et al., 2007).

Bacteroidetes were only cultivated from La Mariana, Sand Island Launch Ramp, and Pier 37. Members of this phylum are generally anaerobes or facultative anaerobes, and are considered indicative of fecal contamination in environmental samples (Santoro & Boehm, 2007; Dick & Field, 2004). Such inputs may exist in these areas given the presence of recreational swimmers (i.e., Sand Island Launch Ramp) and boaters who lack access to a pump house for sanitary removal of waste in the harbor (e.g., La Mariana and Pier 37). Both Ke'ehi Lagoon and Sand Island Beach Park have adequate public restroom facilities on site. La Mariana and neighboring marinas provide showers

and restroom facilities for customers, but not the general public. The Sand Island Beach Park site faces the ocean; this site was thus chosen as a "control" site for comparison to other areas in terms of the impact of the molasses spill due to its location outside the main harbor.

From approximately 8-days post-spill, the fraction of OTUs affiliates with *Proteobacteria* increased until 20-days post-spill (Figure 16); this was most evident for *Alphaproteobacteria* and *Gammaproteobacteria* sub-classes, which host facultative or obligate anaerobes with diverse metabolisms (Gupta, 2000). OTUs affiliating with the *Proteobacteria* were principally from the *Gammaproteobacteria*, followed by *Alphaproteobacteria* (App. 18 & 19), which coincidentally reflects their distribution in a meta-analysis of the taxonomic distribution of sequences submitted to GenBank (Hagström et al., 2002).

The final sample, taken approximately one-year post-spill, was represented by only *Firmicutes* and *Actinobacteria*. The *Firmicutes* are separated into classes: the anaerobic "Clostridia", the aerobic "Bacilli", "Erysipelotrichi", "Negativicutes" and the "Thermolithobacteria". However, the cultivation method used here did not select for strict anaerobes, such as the *Clostridia* (Ludwig et al. 2009). *Actinobacteria* contains some facultatively anaerobic genera, but most members of the phylum are aerobic, with diverse physiological and metabolic properties, including the production of extracellular enzymes and secondary metabolites (e.g., antibiotics) (Ventura et al., 2007).

The presence of contaminating bacteria on the control plates was likely from carry-over contamination of the pipetteman, which was used to dispense samples onto the media,

incomplete flame sterilization of the glass spreader used to distribute the inoculum on the plates, or from contaminated gloves. The dilution series was prepared in a biosafety cabinet, reducing the likelihood of airborne contamination. Bacteria recovered from the control plates affiliated with the two genera most often recovered from sediment samples throughout this project, namely *Bacillus* spp. (*n* = 3) and *Vibrio* spp. (*n* = 2). That they belonged to the most commonly cultivated genera suggests a methodological failure rather than airborne or human contamination. Different batches of 3% sterile saline were used to dilute sediment collected on different days, so there could not have been cross-contamination between samples.

4.2 Analysis of 16S rRNA Sequences

The KB™ Base Caller program was used instead of Phred base call by CodonCode Aligner because the former program has been optimized for increased read length from the 3' and 5' ends of the sequences, as well as increased accuracy in regions with anomalous signal artifacts or low signal-to-noise ratios (ABI, 2004). The KB™ Base Caller program better resolves base calls, but has less capacity to detect heterozygous alleles.

Sequences that failed quality control were either mixed, as evidenced by double-bands in the chromatogram, or of poor amplification quality. In hindsight, the cultures that either failed amplification or resulted in low product yield should have been checked by Gram stain to determine if they were pure *Bacteria* cultures, mixed, or possibly eukaryotes. This study focused on the bacterial-fraction of the microbial community, and *Eucarya*-specific primers were not used. It is possible that the primer set used (27F, 1492R) generated cross-Kingdom amplification of eukaryotes (Galkiewicz & Kellogg, 2008). Sequencing using the forward-, middle-, and reverse-primers would have

yielded a 1500 bp product under ideal conditions, which would have increased the confidence of taxonomic affiliation and potential carryover of eukaryotic amplicons. The primer 27F was previously reported as *Bacteria*-specific (Frank et al., 2008) but according to work by Galkiewicz & Kellogg (2008), the primer set 27F/1492R generated coral amplicons with a mean molecular weight of 1.5 kbp, which is within 200 bp of the length as the expected of a *Bacteria* amplicon.

To reduce cross-Kingdom amplification of non-target eukaryotic DNA, the use of 8F/1492R with a higher annealing temperature (62°C) or 63F/1592R with a lower annealing temperature is recommended (52°C) to reduce non-specific amplification (cf. Galkiewicz & Kellogg, 2008). Also, the cultures should be examined via wet-mount and Gram stained prior to PCR to to determine if they are eukaryotic or Gram-positive organisms that may require extractions for effective cell lysis. A small subset of (approximately 10) cultures for which no PCR product was obtained after DNA extraction were checked via wet-mount to determine if they were eukaryotes, but the samples did not appear nucleated, or significantly larger than would be expected if they were eukaryotic.

The taxonomic affiliations assigned by the RDP and SILVA databases were in agreement for each OTU at the phylum level, but not so at the family level (Apps. 16, & 17). The SILVA database was used for comparison in mothur because the database was more recently updated and includes more sequences than the RDP database. The OTUs were aligned to the SILVA Release v.119 and filtered to remove gaps in order to decrease calculation time in generating distance files. The distance files were used for tree-building by the "neighbor-joining" method (Saitou & Nei, 1987). Bootstrapping allows for multiple random resamplings of the input data to reduce the effect of non-

Normal distributions and small sample sizes, both of which were concerns due to the number of isolates recovered (Moore et al., 2009). mothur automatically roots the tree using the incorporated algorithm (Schloss et al., 2009), unlike other programs that require a reference sequence of a distantly related relative to the samples (*e.g.*, *Thermotoga* sp.) to root the phylogentic tree (Huelsenbeck et al., 2002),

Heat maps generated in mothur offer another means of visualizing the number of OTUs in only a single sample, as well as those common to two or more samples (Apps. 15 & 16). This feature is useful for rapid visualization of OTUs common to two or more samples, with red denoting OTUs common to two or more samples and black indicating those detected in just a single sample (Apps. 15 & 16).

The fact that the rarefaction curves for the genera and species did not level off indicated under-sampling at these levels. This limits the quantitative and qualitative value of the data in these respects.

Since the number of isolates cultured from each time point differed, a Similarity/Distance matrix, i.e., Bray-Curtis index (Hughes & Bohannan, 2004), compared similarity between both sample sites and sample time points (Apps. 20 & 21). Separate OTUs were defined according to an identity threshold of 3% (or ≤97% 16S rRNA nucleotide identity). However, Stackebrandt (2011) suggests a more stringent cutoff of ≥98.5% sequence identity may be more appropriate when defining OTUs. The less stringent cut-off of ≤97% was used here because the more stringent cut-off values have not been universally adopted for publication standards (Janda & Abbott, 2007) and software limitations do not allow half-percentage cut-off values. The minimum recommended read length for classification is 50 bases for RDP classification

(Wang et al., 2007), although Cole et al. (2014) recommends a 500 bases minimum length to allow sufficient context for taxonomic classification. These standards were defined for next generation sequencing of paired-end reads, not single-primer Sanger sequences, which were used for this study. Wang et al. (2007) reports an 88.7% confidence in assignment of genera at 400 base segments, with a diminishing 83.2% accuracy for samples of only 200 base segments. At 50 bases, the reliability of RDP to define members at a genus-level is diminished to 51.5% due to poor coverage of the variable regions, but the classification at the phylum-level remains 94.1% (Wang et al., 2007). This suggests that even with retention of short-reads (>500 nt), the taxonomic assignment at the phylum level is relatively reliable, but for the 57 sequences less than 200 bases, the assignment at the genera-level should be considered as unreliable taxonomic assignments. Sequencing of the the complementary strand would have improved the coverage per sequence, and thus the reliability of the taxonomic assignment at the genus-level. For both *Vibrio* and *Bacillus*, the two most common genera identified in this study, identification at the species- or strain-level would have required additional sequencing of another housekeeping gene, as 16S rRNA gene sequencing is unreliable due to high sequence homology (Thompson et al., 2005; Janda & Abbott, 2007).

The benefit of using RDP or SINA over GenBank searches using BLAST or EZ Taxon, is that the sequences can be submitted *in toto*. The naïve Bayesian classifier used by RDP is also more efficient at placing rare or undescribed query sequences with no nearest close relatives (Kim et al., 2013), which is possible in culturing environmental samples.

Parsimony analysis compares multiple random-joining trees to provide statistical information on the likelihood that the communities have the same structure by chance

(n = 10,000 randomizations), with each tree scored to generate a distribution (Slatkin & Maddison, 1990). With respect to the *Bacteria* cultivated at each time point, T1 and T2 differed significantly both from each other and from all other time points. In terms of the taxonomic affiliations of the *Bacteria* cultivated from each sample site, none differed significantly from the controls, meaning that each contained representatives of *Bacillus* spp. and *Vibrio* spp., the most commonly cultivated isolates, yet all sample sites were significantly different from each other (App. 17).

The weighted UniFrac algorithm is used to determine if the groups (sample time point or sample site) have significantly different structure (Lozupone et al. 2007). According to the values generated by the weighted UniFrac analysis, all comparisons among sample times and among sample sites showed they were significantly different ($P \leq 0.001$). This may be a reflection of undersampling, or uneven sampling between sites/time points, rather than significant changes in community structure. Species-level OTUs were also defined at 97% nucleotide identity, but the outcome may have been different were the OTUs were defined at the 98.5% level recommended by Stackebrandt (2011), which would have likely increased the number of reported OTUs and likely in turn decreased increased the statistical differences between groups. The unweighted UniFrac analysis was more useful in determining temporal and spatial relatedness between samples. The first sample point 8-day post-spill was the most distinctive site among all time points in terms of diversity and community composition. The results of the sample sites suggest a spatial pattern between sample location. For example, water circulation patterns and proximity to the coastline may be as critical in establishing and maintaining microbial communities.

It would have been prudent to determine the dissolved oxygen concentrations in the sediments in order to determine the depth of the oxic-anoxic boundary (Brune et al., 2000) and the depth of oxygen penetration. This would have offered another means of determining if the composition of the microbial community reflected that of the oxygen distribution. Samples should also have been collected as soon as possible after the molasses spill in order to capture the 'time-zero' or baseline data. Determination of the effect of intra-seasonal, seasonal, and inter-annual fluctuations in environmental conditions, and the resultant effect on beta-diversity (Hatosy et al., 2013), could not be clearly determined with the collection methods used for this study.

Rather than sampling all representative colonies found on nutrient media, a random subsample of colonies would offer the statistical strength to measure species richness while reducing sampling bias and preventing pseudoreplication (Hurlbert, 1984) based on sampling effort per plate or selection based on colony size, color, etc. The method applied also failed to cultivate representatives of anaerobic genera, such as *Clostridium*; adding anaerobic cultivation methods would have been useful for compiling a wider profile of the culturable community. A complementary method, such as DNA extraction from the same sediment samples, to determine which organisms are present through a 'molecular' approach such as high-throughput sequencing of community DNA using an Illumina platform, or a ribosomal tag pyrosequencing analysis (Brown et al., 2009), would provide information about the 'unculturable' fraction of the microbial community.

4.3 Conclusions

The molasses spill into Honolulu Harbor on September 7, 2013 provided an opportunity to determine how the taxonomic composition of the culturable *Bacteria*

community in the harbor's surface sediment changed in response. However, the results of this study offer only some insight in this respect because the depth of sampling proved to be inadequate for robust statistical outcomes. Recommendations for any subsequent spill assessments include: 1) A more rapid first response (i.e., initiate sampling within 24-48 hours post-spill); 2) A broader array of culture media and incubation conditions, especially with respect to providing different atmospheres to target both facultative and obligate anaerobes. Such groups can reasonably be expected to be abundant in benthic marine sediments, so conditions that reduce oxygen availability near the surface of such sediments or in the overlying water may well lead to enrichment of such *Bacteria*; 3) Purifying more colonies from each medium would provide a more representative overview of the culturable community; 4) Examining all pure cultures by microscopy before PCR would reveal which are prokaryotes (cf. *Bacteria* or *Archaea*) as opposed to *Eucarya*, either through Gram-staining or wet mounts with simple stain; 5) Sequencing full or almost full 16S rRNA genes in all cultivated *Bacteria* would provide more reliable and deeper insight into their taxonomic affiliations. The relatively short sequences used here simply could not resolve some sequences from others, so OTUs here are likely more abundant (i.e., comprise more sequences) and thus comprise larger fractions of the community than they would if OTUs were based on full sequences; 6) A complimentary analysis should be based on determining which organisms are present through a 'molecular' approach, such as high-throughput sequencing on an Illumina platform, or a ribosomal tag pyrosequencing analysis (Brown et al., 2009). This would complement the culture-based approach (sensu Donachie et al., 2007).

REFERENCES

Angoco, V. (2013) Statement from Matson Navigation on DOT letter. *Hawai'i News Now*. 20 Sept. 2013: <www.hawaiinewsnow.com/story/23490464/statement-from-matson-navigation-on-dot-letter>.

Applied Biosystems, Inc. (2004) The KB™ Basecaller v1.1.1 from Applied Biosystems Frequently Asked Questions. Accessed: 17 Sept. 2015: <http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_040412.pdf>.

Basu, T. (2013) Hawai'i Molasses Spill: Better or worse than oil? *Nat. Geo*. 17 Sept 2013: <http://news.nationalgeographic.com/news/2013/13/130917-molasses-oil-spill-hawaii-honolulu-cleanup/>.

Bidigare, Robert (2015) Biogeochemical responses to a massive molasses spill in Honolulu Harbor, Hawai'i. Hawai'i Institute of Marine Biology. Presentation: 09 Sept 2015.

Brown, M.V., G.K. Philip, J.A. Bunge, M.C. Smith, A. Bissett, F.M. Lauro, J.A. Fuhrman, S.P. Donachie (2009) Microbial community structure in the North Pacific Ocean. *The ISME Journal* **3**: 1374-1386.

Brune, A., P. Frenzel, & H. Cypionka (2000) Life at the oxic-anoxic interface: microbial activities and adaptations. *FEMS Microbiol Rev*. **24**(5): 691-710.

Claesson, M.J., O. O'Sullivan, Q. Wang, J. Nikkilä, J.R. Marchesi, H. Smidt, W.M. de Vos, R.P. Ross, & P.W. O'Toole (2009) Comparative analysis of pyrosequencing and phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS One* **4**(8): e6669.

CodonCode Aligner Corporation (2015) CodonCode Aligner User Manual. Accessed: 28 August 2015. <http://www.codoncode.com/aligner/AlignerHelp.pdf>.

Cole, J.R., Q. Wang, J.A. Fish, B. Chai, D.M. McGarrell, Y. Sun, C.T. Brown, A. Porras-Alfaro, C.R. Kuske, & J. M Tiedje (2014) Ribosomal Database Project: Data and tools for high-throughput rRNA analysis. *Nuc. Acid. Res*. **42**(Database Issue): D633-D642.

Curtin, L.V. (1983) Molasses – General Considerations. From: *Molasses in Animal Nutrition*. West Des Moines (IA): National Feed Ingredients Association. p.1-11.

Dalton, S.J. & G. Roff (2013) Spatial and temporal patterns of eastern Australia subtropical coral communities. *PLoS One* **8**(9): e75873.

Dick, L.K. & K.G. Field (2004) Rapid estimation of numbers of fecal *Bacteroidetes* by use of quantitative PCR assay for 16S rRNA genes. *Appl. Environ. Microbiol*. **70**(9): 5695-5697.

Donachie, S.P., J.S. Foster, M.V. Brown (2007) Culture Clash: Challenging the dogma of microbial diversity. *The ISME Journal* **1**: 97-99.

Evans, J., L. Sheneman, & J.A. Foster (2006) Relaxed Neighbor-Joining: A fast distance-based phylogenetic tree construction method. *J. Mol. Evol.* **62**: 785-792.

Ewing, B., L.D. Hillier, M.C. Wendl, & P. Green (1998) Base-calling of automatic sequencer traces using Phred. *Gen. Res.* **8**: 175-185.

Felsenstein, J. (1989) PHYLIP – PHYLogeny Inference Package (Version 3.2). *Cladistics* **5**: 164-166.

Frank, J.A., C.I. Reich, S. Sharma, J.S. Weisbaum, B.A. Wilson, & G.J. Olsen (2008) Critical evaluation of two primers commonly used for amplification of Bacterial 16S rRNA genes. *Appl. Environ. Microbiol.* **74**(8): 2461-2470.

Galkiewicz, J.P. & C.A. Kellogg (2008) Cross-Kingdom amplification using Bacteria-specific primers: Complications for studies of coral microbial ecology. *Appl. Envrion. Microbiol.* **74**(24): 7828-7831.

Gopal, A.R. & D.M. Kammen (2009) Molasses for ethanol: The economic and environmental impacts of a new pathway for the lifecycle greenhouse gas analysis of sugarcane ethanol. *Environ. Res. Lett.* **4**(4): 044005.

Grube, N. (2013) New Molasses Spill Response Plan concedes environmental danger. *Civil Beat*. 13 Nov 2013: <http://www.civilbeat.com/2013/11/20382-new-molasses-spill-response-plan-concedes-environmental-danger/>.

Gupta, R.S. (2000) The phylogeny of Proteobacteria: Relationships to other eubacterial phyla and eukaryotes. *FEMS Microbiol. Rev.* **24**: 367-402.

Haas, B.J., D. Gevers, A.M. Earl, M. Feldgarden, D.V. Ward, G. Giannoukos, D. Ciulla, D. Tabbaa, S.K. Highlander, E. Sodergren, B. Methé, T.Z. DeSantis, The Human Microbiome Consortium, J. F. Petrosino, R. Knight, & B.W. Birren (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Gen. Res.* **21**: 494-504.

Hagström, Å., T. Pommier, F. Rohwer, K. Simu, W. Stolte, D. Svensson, & U.L. Zweifel (2002) Use of 16S Ribosomal DNA for delineation of marine bacterioplankton species. *Appl. Environ. Microbiol.* **68**(7): 3628-3633.

Hatosy, S.M., J.B. Martiny, R. Sachdeva, J. Steele, J.A. Fuhrman, & A.C. Martiny (2013) Beta-diversity of marine bacteria depends on temporal scale. *Ecology* **94**(9): 1898-904.

Hawai'i State Department of Health (HI-DoH) (2013a) "News Release: Molasses spill in Honolulu Harbor causes fish kill – Department of Health closes beaches in Ke'ehi Lagoon area." 10 Sept 2013: <http://health.hawaii.gov/news/files/2013/05/13-050-Molasses-Spill-in-Honolulu-Harbor-Causes-Fish-Kill.pdf>.

Hawai'i State Department of Health (HI-DoH) (2013b) "Honolulu Harbor Molasses Spill 2013" Last updated: 20 September 2013: <http://eha-web.doh.hawaii.gov/eha-cma/Leaders/HEER/Honolulu-Harbor-Molasses-Spill-September-2013>.

Hawai'i State Department of Health (HI-DoH) (2013c) Hawai'i's Implementation Plan for Polluted Runoff Control: Appendix F – Detailed descriptions of Hawai'i's 18 Water Quality Limited Segments. Clean Water Branch. Accessed: 08 Nov. 2015 <http://health.hawaii.gov/cwb/files/2013/05/PRC_Implan_AppendixF.pdf>.

Hawai'i State Department of Transportation – Harbors Division (HI-DoT) (2012a) "A Guide to Port Hawai'i." Accessed 09 Nov 2014: <http://hidot.hawaii.gov/harbors/files/2012/10/A-Guide-To-Port-Hawaii.pdf>.

Hawai'i State Department of Transportation – Harbors Division (HI-DoT) (2012b) "Honolulu Harbor – O'ahu." Accessed 09 Sept 2015: <hitdot.hawaii.gov/harbors/files/2012/10/Honolulu-Harbor-O'ahu.pdf>.

Hawai'i State Department of Transportation – Harbors Division (HI-DoT) (2012c) Draft Stormwater Management Plan: Honolulu Harbor, Hawai'i. Weston Solutions, Inc.: Honolulu, Hawai'i.

Hurlbert, S.H. (1984) Pseudoreplication and the design of ecological field experiments. *Ecol. Monograph.* **54**(2): 187-211.

Huelsenbeck, J.P., J.P. Bollback, & A.M. Levine (2002) Inferring the root of a phylogenetic tree. *Syst. Biol*. **51**(1): 32-43.

Hughes, J.B. & B.J.M. Bohannan (2004) Application of ecological diversity statistics in microbial ecology. *Molec. Micro. Eco. Man.* 2 ed. **7.01**: 1321-1344.

Hughes, J. B. & J. J. Hellmann (2005) The application of rarefaction techniques to molecular inventories of microbial diversity. *Meth. Enzymol*. **397**: 292-308.

Jabr, F. (2013) The science of the Great Molasses Flood. *Sci. Am.* 17 Jul. 2013: <http://www.scientificamerican.com/article/molasses-flood-physics-science/>.

Janda, J.M. & S.L. Abbott (2007) 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls. *J Clin Microbiol.* **45**(9): 2761-2764.

Joaquin, T. & B. Gutierrez (2013a) Underwater video uncovers mass kill from Matson molasses spill. *Hawaii News Now*. 09 Sept 2013: <http://www.hawaiinewsnow.com/story/23409767/underwater-video-uncovers-mass-kill-from-matson-molasses-spill?autoStart=true&topVideoCatNo=default&clipId=9296935&autoStart=true&topVideoCatNo=default&clipId=9296935>.

Joaquin, T. (2013b) 233,000 gallons of molasses spills into Honolulu Harbor causing fish kill. *Hawaii News Now*. 10 Sept 2013: <http://www.hawaiinewsnow.com/story/23396971/faulty-pipe-caused-honolulu-harbor-molasses-spill>.

Joaquin, T. (2013c) 25,000 fish killed in Matson Molasses Spill. *Hawaii News Now*. 16 Sept 2013: <http://www.hawaiinewsnow.com/story/23448407/25-thousand-fish-killed-in-matson-molasses-spill>.

Joaquin, T. (2013d) State: Matson pipe was leaking year before molasses spill. *Hawaii News Now*. 16 Sept 2013: <http://www.hawaiinewsnow.com/story/23490408/state-inspector-saw-molasses-leak-from-same-matson-pipe-1-year>.

KHON2 (Web Staff) (2013) Matson lays out timeline of massive spill in Honolulu Harbor. 14 Sept 2013: <http://khon2.com/2013/09/14/matson-lays-out-timeline-of-massive-molasses-spill-in-honolulu-harbor/>.

Kim, M., K. Lee, S. Yoon, B. Kim, J. Chun, & H. Yi (2013) Analytical tools and databases for metagenomics in the Next-Generation sequencing era. *Genomics Inform.* **11**(3): 102-113.

Kirchman, D.L., A.I. Dittel, R.R. Malmstrom, & M.T. Cottrell (2005) Biogeography of major bacterial groups in the Delaware Estuary. *Limnol. Oceanogr.* **50**(5): 1697-1706.

Klindworth, A., E. Pruesse, T. Schweer, J. Peplies, C. Quast, M. Hom, & F.O. Glöckner (2012) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nuc. Acid. Res.* **41**(1): e1 doi:10.1093/nar/gks808.

Kozich, J.J., S.L. Westcott, N.T. Baxter, S.K. Highlander, & P.D. Schloss (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**(17): 5112-5120.

Kubota, G. (2003) 50,000 gallons of molasses spills into Maui harbor. *Star Bulletin*. 10 Dec. 2003: <http://archives.starbulletin.com/2003/12/10/news/story7.html>.

Kunin, V., A. Copeland, A. Lapidus, K. Mavromatis, & P. Hugenholtz (2008) A bioinformatician's guide to metagenomics." *Microbiol. Mol. Rev.* **72**(4): 557-578.

Lozupone, C., M. Hamady, S.T. Kelley, & R. Knight (2007) Quantitative and qualitative β-diversity measures lead to different insights into factors that structure microbial communities. *Appl. Environ. Microbiol.* **73**(5): 1576-1585.

Lozupone, C. & R. Knight (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71**(12): 8228-8235.

Ludwig, W., K-H. Schleifer, & W.B. Whitman (2009) Revised road map to the phylum *Firmicutes*. *Bergey's Manual of Systematic Bacteriology*. Vol. 3 Springer: New York. p. 1-13.

Maidak, B.L., G.J. Olsen, N. Larsen, R. Overbeek, M.J. McCaughey, & C.R. Woese (1997) The RDP (Ribosomal Database Project). *Nuc. Acid. Res.* **25**(1): 109-110.

Matson Navigation Company, Inc. ("Matson") (2015) Final Matson Report to Shareholders: "Molasses Report". 29 July 2015: <http://files.shareholder.com/downloads/AMDA-ZJRXJ/0x0x842089/BD8A7C73-0F36-4228-AF5C-A75F9DA705B9/Matson_Molasses_Report_7.29.15.pdf>.

Memmott, M. (2013) Massive molasses spill devastates Honolulu marine life. *NPR*. 12 Sept 2013: <http://www.npr.org/blogs/thetwo-way/2013/09/12/221709158/massive-molasses-spill-devastates-honolulu-marine-life>.

Moore, D.S., G.P. McCabe, & B.A. Craig (2009) Introduction to the practice of statistics. New York (NY): W.H. Freeman & Co. (6th Ed.) p. 368.

Preston, T.R. (1986) Molasses as animal feed: An overview. From: *Sugarcane as Feed. Proceedings from an FAO Expert Consultation*, Santo Domingo, D.R. 7-11 Jul. 1986: <http://www.fao.org/docrep/003/s8850e/S8850E00.htm#TOC>.

Pruesse, E., J. Peplies, & F.O. Glöckner (2012) SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinform*. **28**: 1823-1829.

Pruesse, E., C. Quast, K. Knittel, B. Fuchs, W. Ludwig, J. Peplies, & F.O. Glöckner (2007) SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nuc. Acid. Res*. **35**(21): 7188-7196.

Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Scheweer, P. Yarza, J. Peplies, & F.O. Glöckner (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nuc. Acid. Res*. **41**(D1): D590-D596.

Richmond, Robert (2015) Personal Interview. 04 December 2015.

Saitou, N. & M. Nei (1987) The Neighbor-Joining Method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol*. **4**(4): 406-425.

Santoro, A.E. & A.B. Boehm (2007) Frequent occurrence of the human-specific *Bacteroidetes* fecal marker at an open coast marine beach: Relationship to waves, tides, and traditional indicators. *Environ. Microbiol*. **9**(8): 2038-2049.

Schloss, P.D. & S.L. Westcott (2011) Assessing and improving methods used in Operational Taxonomic Unit-Based approaches for 16S rRNA gene sequence analysis. *Appl. Environ. Microbiol.* **77**(10): 3219-3226.

Schloss, P.D., S.L. Westcott, T. Ryabin, J.R. Hall, M. Hartmann, E.B. Hollister, R.A. Lesniewski, B.B. Oakley, D.H. Parks, C.J. Robinson, J.W. Sahl, B. Stres, G.G. Thallinger, D.J. Van Horn, & C.F. Weber (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Micro*. **75**(23): 7537-7541.

Schworm, P. (2015) Nearly a century later, structural flaw in molasses tank revealed. *The Boston Globe*. 14 Jan. 2015: <https://www.bostonglobe.com/metro/2015/01/14/nearly-century-later-new-insight-into-cause-great-molasses-flood >

Slatkin, M. & W.P. Maddison (1990) Detecting isolation by distance using phylogenies of genes. *Genetics* **126**(1): 249-260.

Stackebrandt, E. (2011) Molecular taxonomic parameters. *Microbiol. Aust.* 59-61.

Thompson, F.L., D. Gevers, C.C. Thompson, P. Dawyndt, S. Naser, B. Hoste, C.B. Munn, & J. Swings (2005) Phylogeny and molecular identification of Vibrios on the basis of multilocus sequence analysis. *Appl. Environ. Microbiol.* **71**(9): 5107-5115.

Ventura, M., C. Canchaya, A. Tauch, G. Chandra, G.F. Fitzgerald, K.F. Chater, & D. van Sinderen (2007) Genomics of *Actinobacteria*: Tracing the evolutionary history of an ancient phylum. *Microbiol. Mol. Biol. Rev*. **71**(3): 495-548.

Wang, Q., G.M. Garrity, J.M. Tiedje, & J.R. Cole (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol*. **73**(16): 5261-5267.

Watkins, T. (2013) Hawai'i molasses spill could lead to more predator species, agency says. *CNN*. 23 Sept 2013: <http://www.cnn.com/2013/09/12/us/hawaii-molasses-spill/>.

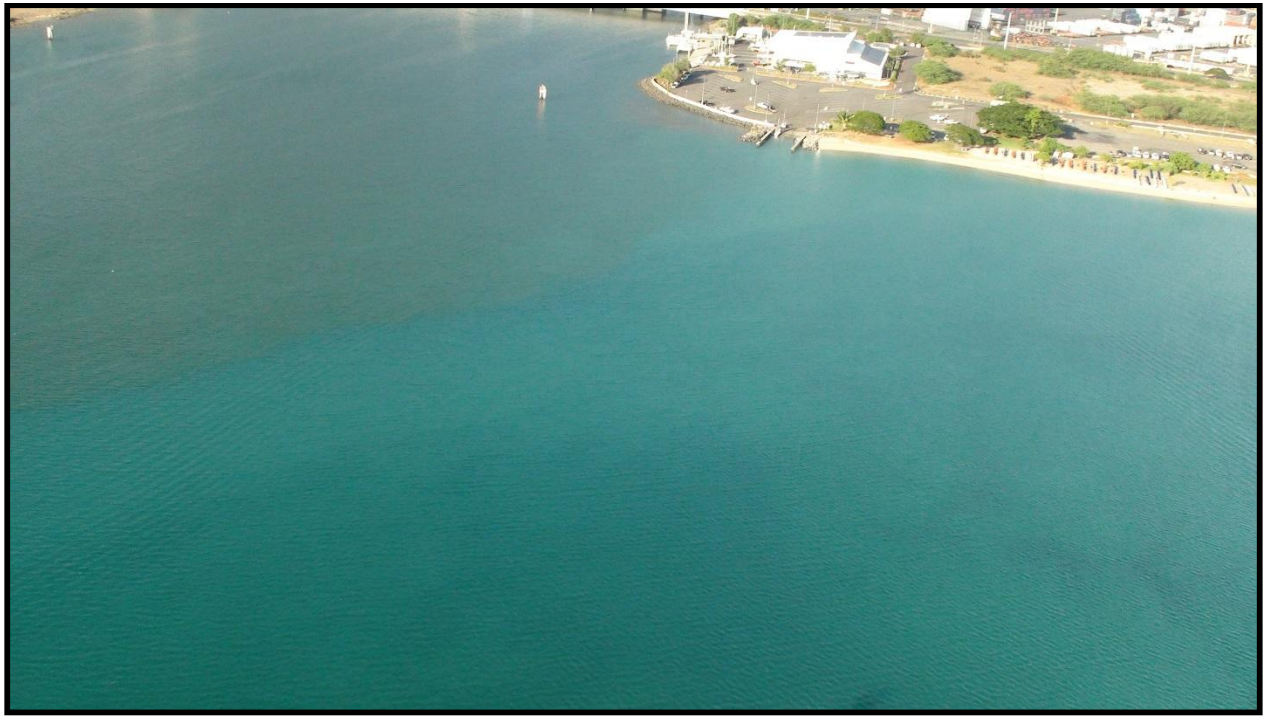
Weisberg, W.G., S.M. Barns, D.A. Pelletier, D.J. Lane (1991) 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol*. **173**(2): 697-703.

Wisch, J.A. (2015) Governor of the State of Hawai'i – Office of the Attorney General News Release: "Matson to pay more than $15 million for 2013 molasses spill in settlement agreement with State of Hawai'i." Posted: 29 July 2015. <http://governor.hawaii.gov/newsroom/ag-news-release-matson-to-pay-more-than-15-million-for-2013-molasses-spill-in-settlement-agreement-with-state-of-hawaii/>.

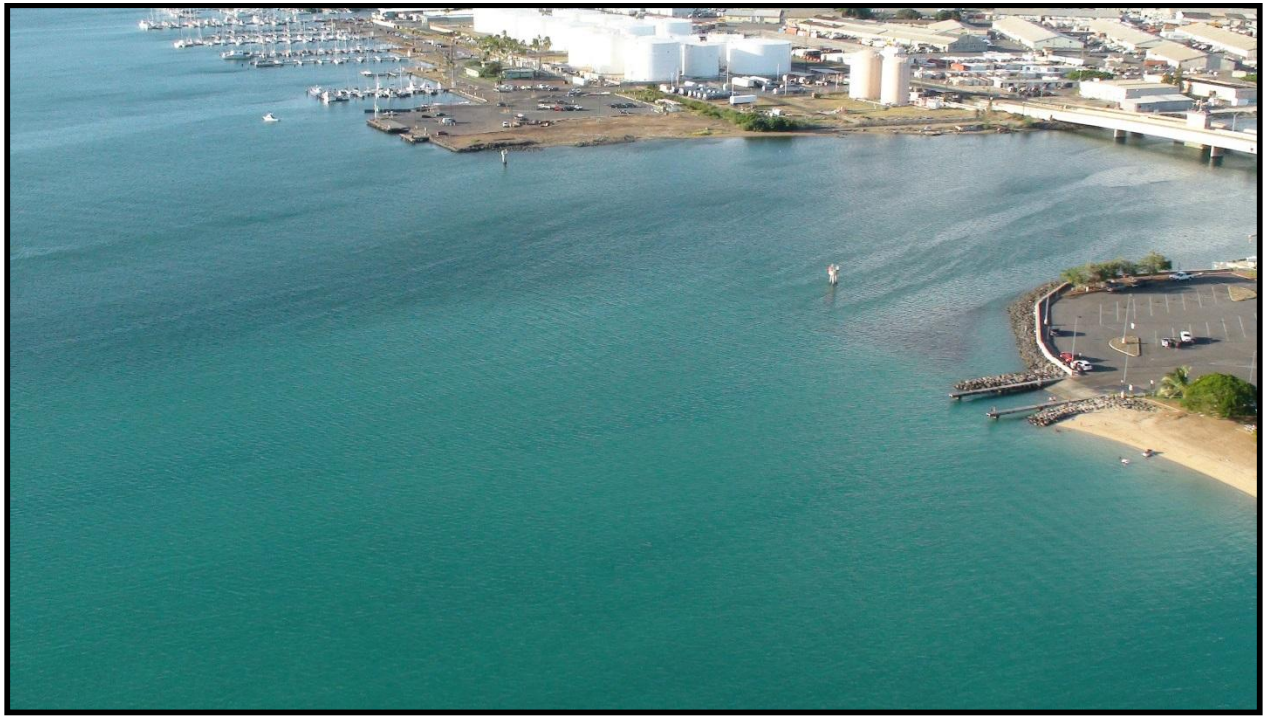
Wright, E.S. (2012) DECIPHER: A search-based approach to chimera identification for 16S rRNA sequences." *Appl. Environ. Microbiol*. **78**(3): 717-725.

APPENDICES

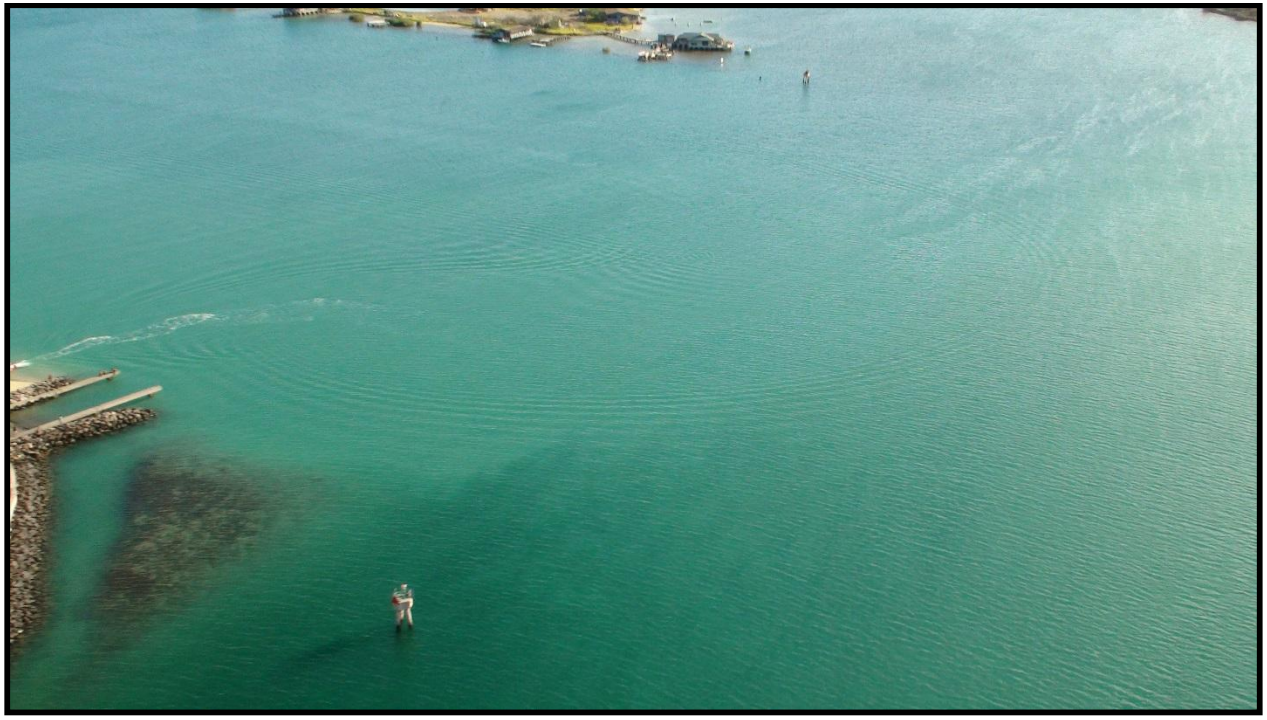
Appendix 1. Aerial photograph of Sand Island Launch Ramp on September 12, 2013 at approximately 5:00 p.m. (HI-DoH, 2013b).



Appendix 2. Aerial photograph of Sand Island Launch Ramp on September 13, 2013 at approximately 5:00 p.m. (HI-DoH, 2013b).



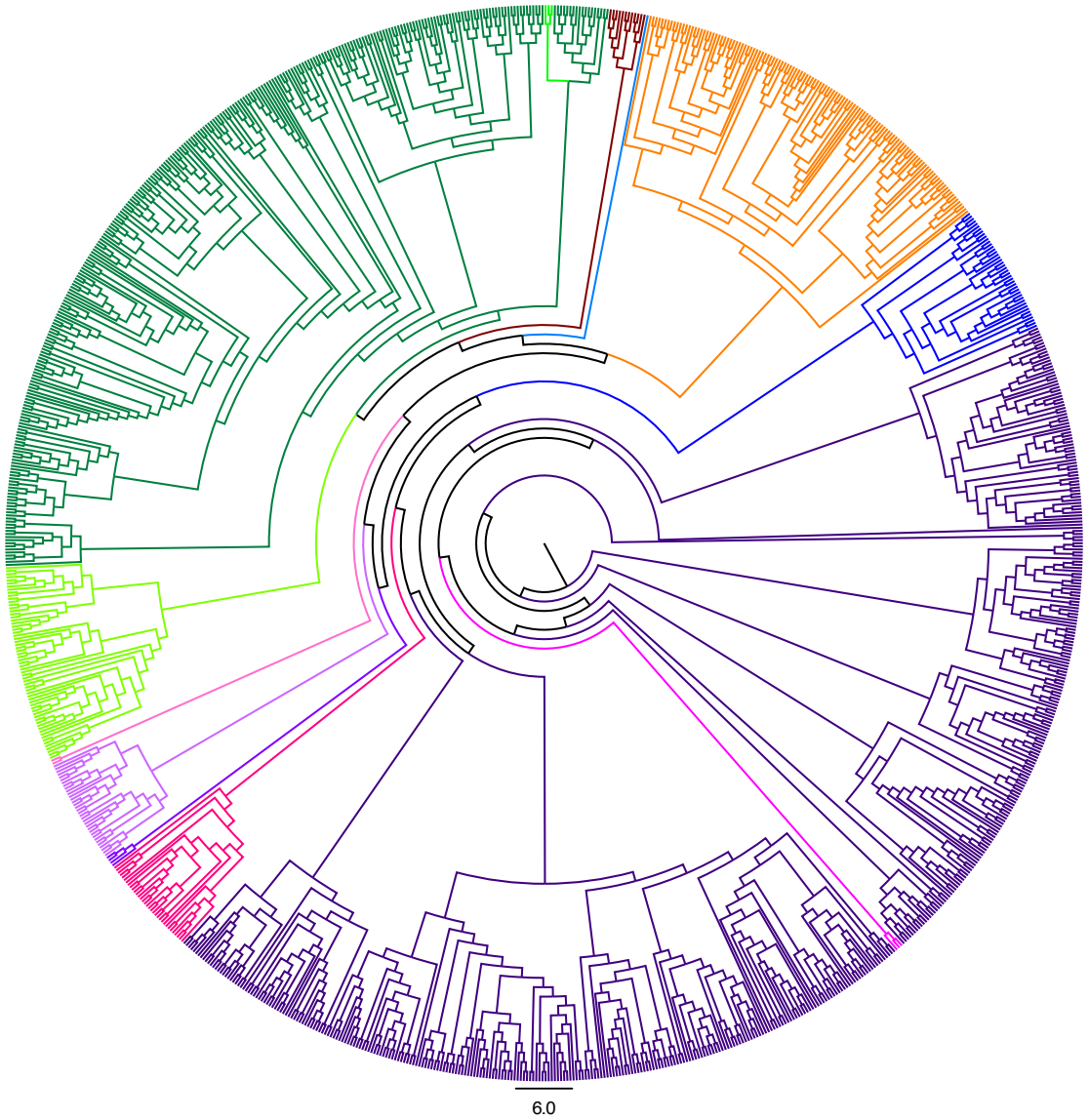
Appendix 3. Aerial photograph of Sand Island Launch Ramp on September 14, 2013 at approximately 5:00 p.m. (HI-DoH, 2013b).



Appendix 4. Aerial photograph of Sand Island Launch Ramp on September 15, 2013 at approximately 5:00 p.m. (HI-DoH, 2013b).

Appendix 5. Aerial photograph of Sand Island Launch Ramp on September 16, 2013 at approximately 5:00 p.m. (HI-DoH, 2013b).

Phyletic Association:	Color:
Firmicutes (F: Bacillaceae 2)	44
Firmicutes (G: Bacillus)	467
Firmicutes (G: Staphylococcus)	3
Firmicutes (G: Rumellibacillus/ Lysinibacillus)	35
Firmicutes (G: Exiguobacterium)	4
Firmicutes (F: Paenibacillaceae)	35
Firmicutes (G: Brevibacillus)	2
Proteobacteria (C: Alphaproteobacteria)	55
Proteobacteria (C: Gammaproteobacteria)	299
Proteobacteria (C: Betaproteobacteria)	3
Bacteriodetes	12
Deinococcus-Thermus	1
Actinobacteria	126



Appendix 6. Phylum affiliation of all pure *Bacteria* cultures [mothur (v1.36.1); FigTree v1.4.0)].

Appendix 7. Mean CFU/g sediment on MA.

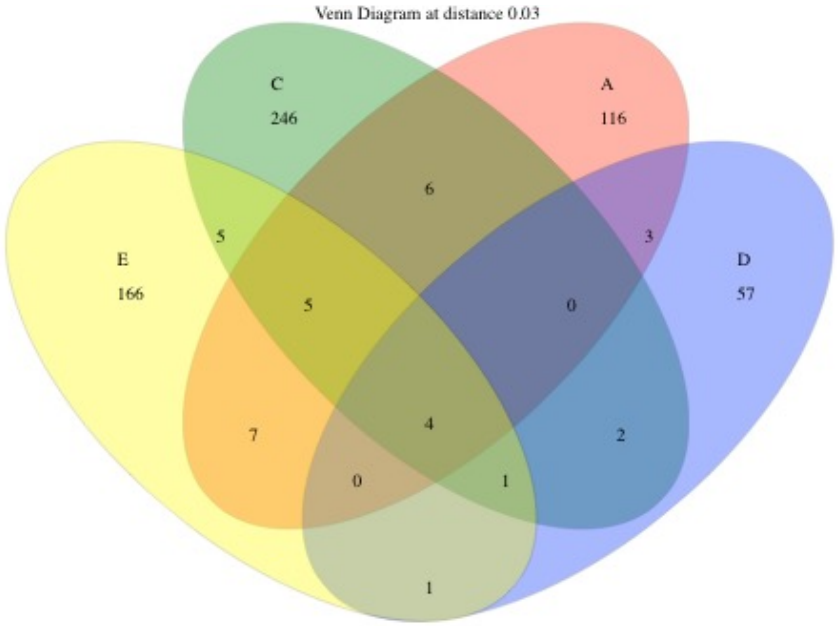
Time Point	Date Sampled	Days Post-Spill	Ke'ehi Lagoon	La Mariana	Sand Island Launch Ramp	Sand Island Beach Park	Pier 37 – Fish Market
T0	9/7/12	0	N/A	N/A	N/A	N/A	N/A
T1	9/15/13	8	3.13E+05	2.79E+05	1.73E+05	2.04E+04	1.49E+05
T2	9/19/13	12	7.08E+04	7.22E+04	1.18E+05	6.84E+03	1.39E+05
T3	9/23/13	16	9.43E+04	6.18E+05*	5.77E+04*	1.10E+04	3.14E+05*
T4	9/27/13	20	5.87E+04	1.63E+05	1.40E+05*	9.98E+03	1.48E+05
T5	10/1/13	24	6.14E+04	1.67E+05	9.34E+04	5.53E+03	1.95E+05
T6	10/5/13	28	2.31E+04	1.23E+05	1.11E+04*	5.61E+03	9.09E+04
T7	10/9/13	32	7.67E+04	4.21E+04	4.63E+04	1.65E+04	3.98E+04
T8	10/12/13	35	9.00E+04	7.68E+04	2.07E+04	6.74E+03	2.72E+04
T9	10/26/13	49	1.53E+04	1.68E+04	1.09E+04	1.23E+03	1.39E+04
T10	11/11/13	65	3.91E+03	1.15E+03	2.42E+03	7.98E+02	3.33E+03
T11	12/7/13	91	1.44E+04	1.57E+04	2.57E+04*	3.33E+03	4.42E+04
T12	1/12/14	127	1.28E+04	9.63E+03	2.94E+04	1.37E+03	1.28E+04
T13	2/9/14	155	5.05E+03	7.94E+03	1.33E+04	6.61E+03	2.14E+04
T14	9/17/14	375	3.00E+03	3.99E+03	9.18E+03	3.64E+02	3.32E+03

* Less than 3 plates used for calculating average CFU/g due to agarolytic activity.

Appendix 8. Mean CFU/g sediment on R2A.

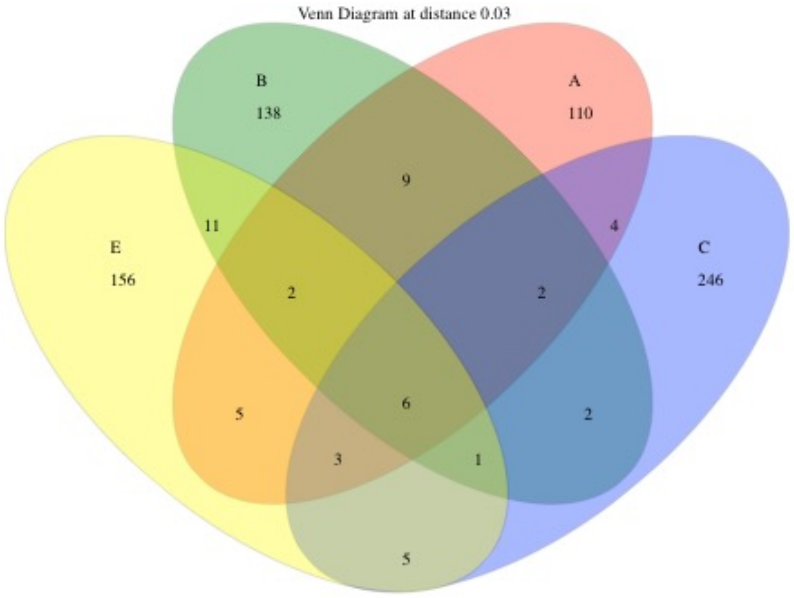
Time Point	Date Sampled	Days Post-Spill	Ke'ehi Lagoon	La Mariana	Sand Island Launch Ramp	Sand Island Beach Park	Pier 37 – Fish Market
T0	9/7/13	0	N/A	N/A	N/A	N/A	N/A
T1	9/15/13	8	1.37E+05	1.04E+05	8.25E+04	8.88E+04	1.41E+05
T2	9/19/13	12	1.75E+03	4.49E+03	3.78E+03	7.66E+02*	1.64E+04
T3	9/23/13	16	1.46E+04	2.04E+04	8.09E+03*	0.00E+00	8.07E+03*
T4	9/27/13	20	2.01E+03	4.48E+03	1.55E+05	0.00E+00	5.54E+03
T5	10/1/13	24	4.61E+03	5.83E+03	5.36E+04	0.00E+00	1.92E+04
T6	10/5/13	28	3.10E+03	8.62E+03	1.86E+04	0.00E+00	7.98E+03
T7	10/9/13	32	2.22E+03	5.44E+03	5.62E+03	3.68E+02	1.66E+04
T8	10/12/13	35	1.05E+03	2.85E+03	6.35E+03	0.00E+00	2.45E+03
T9	10/26/13	49	1.23E+03	2.90E+03	1.36E+03	3.62E+02	3.04E+03
T10	11/11/13	65	2.80E+03	2.64E+03	2.03E+03	0.00E+00	1.35E+02
T11	12/7/13	91	5.92E+03	8.72E+03	2.36E+04	1.56E+02	6.32E+03
T12	1/12/14	127	6.00E+03	1.31E+04	5.71E+04	1.43E+02	1.14E+04
T13	2/9/14	155	2.04E+03	5.28E+03	4.93E+03	0.00E+00	8.91E+03
T14	9/17/14	375	2.67E+03	1.76E+03	1.75E+03	0.00E+00	2.41E+03

* Less than 3 plates used for calculating average CFU/g due to agarolytic activity.



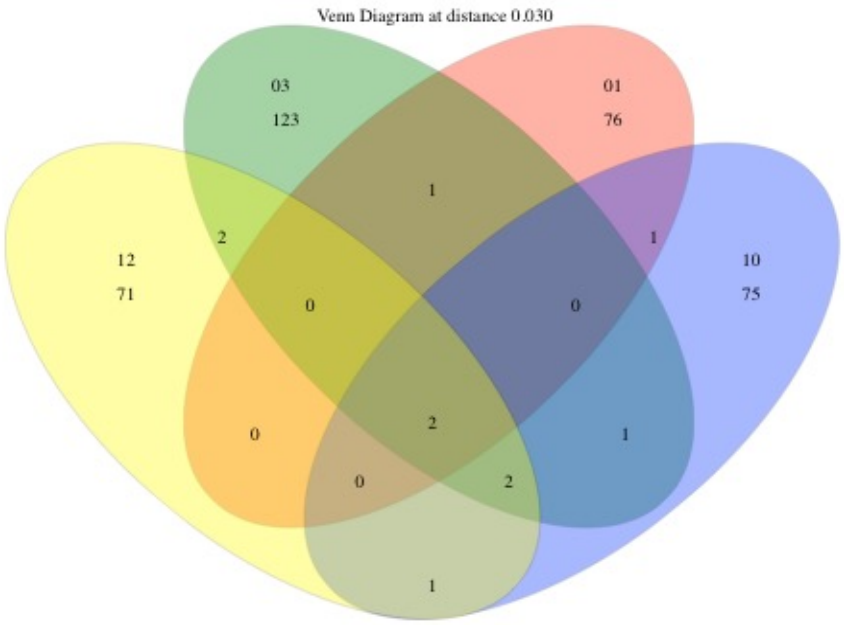
Appendix 9. Venn diagram of shared OTUs (97% 16S rRNA nucleotide identity) for isolates from Ke'ehi Lagoon (A), Sand Island Launch Ramp (C), Sand Island Beach Park (D), and Pier 37 (E).

Appendix 10. Venn diagram of shared OTUs (97% 16S rRNA nucleotide identity) for isolates from La Mariana (B), Sand Island Launch Ramp (C), Sand Island Beach Park (D), and Pier 37 (E).

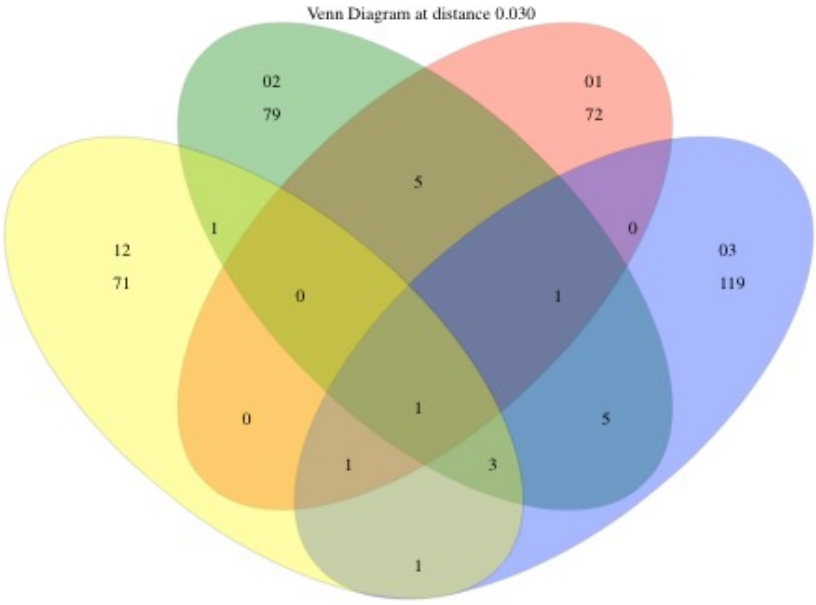


Appendix 11. Venn diagram of shared OTUs (97% 16S rRNA nucleotide identity) for isolates from Ke'ehi Lagoon (A), La Mariana (B), Sand Island Launch Ramp (C), and Pier 37 (E).

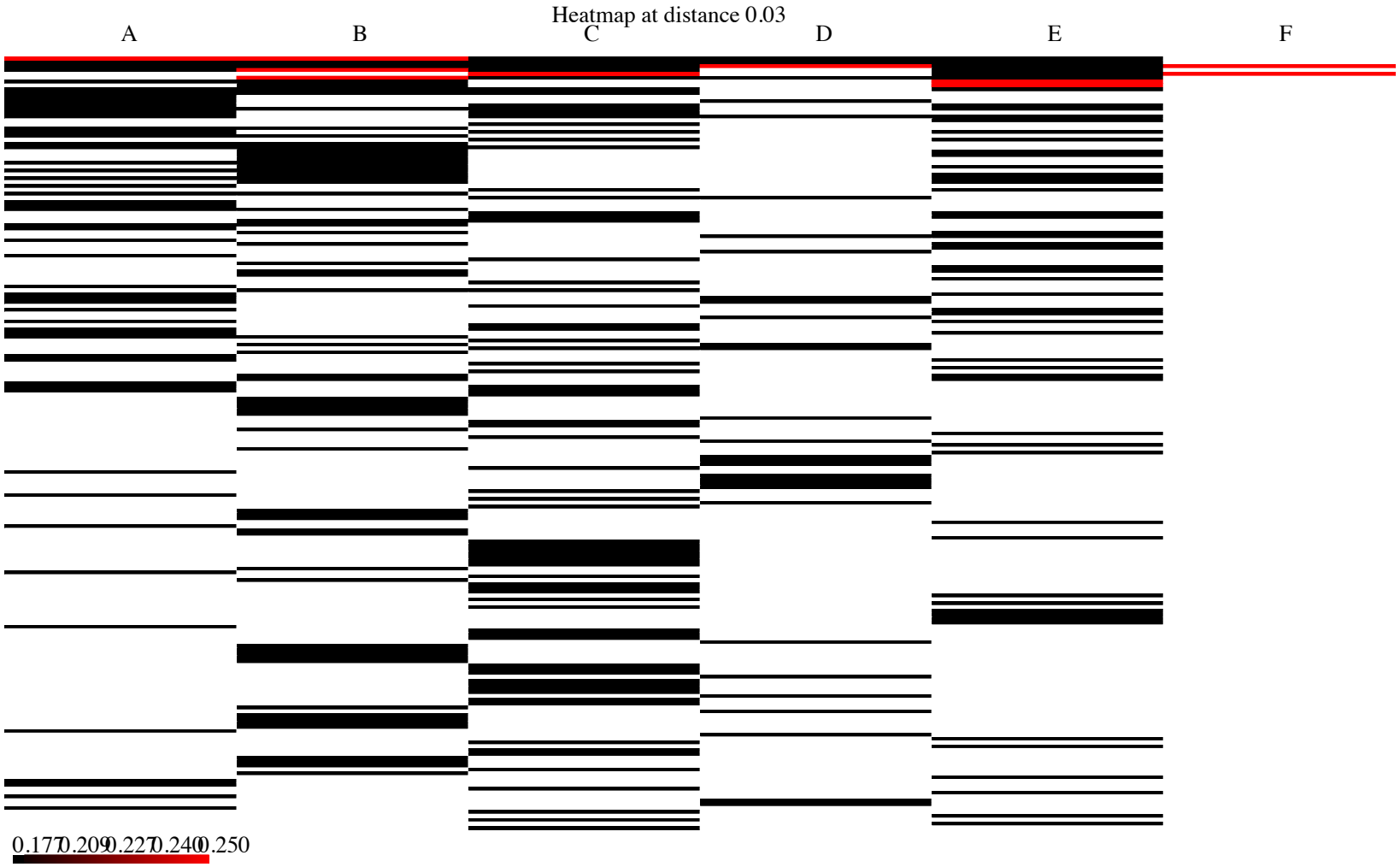
Appendix 12. Venn diagram of shared OTUs (97% 16S rRNA nucleotide identity) for isolates from sample time points T01 (09/15/13; 8-days post), T03 (09/23/13; 16-days post), T05 (10/01/13; 24-days post), and T10 (11/11/13; 65-days post).



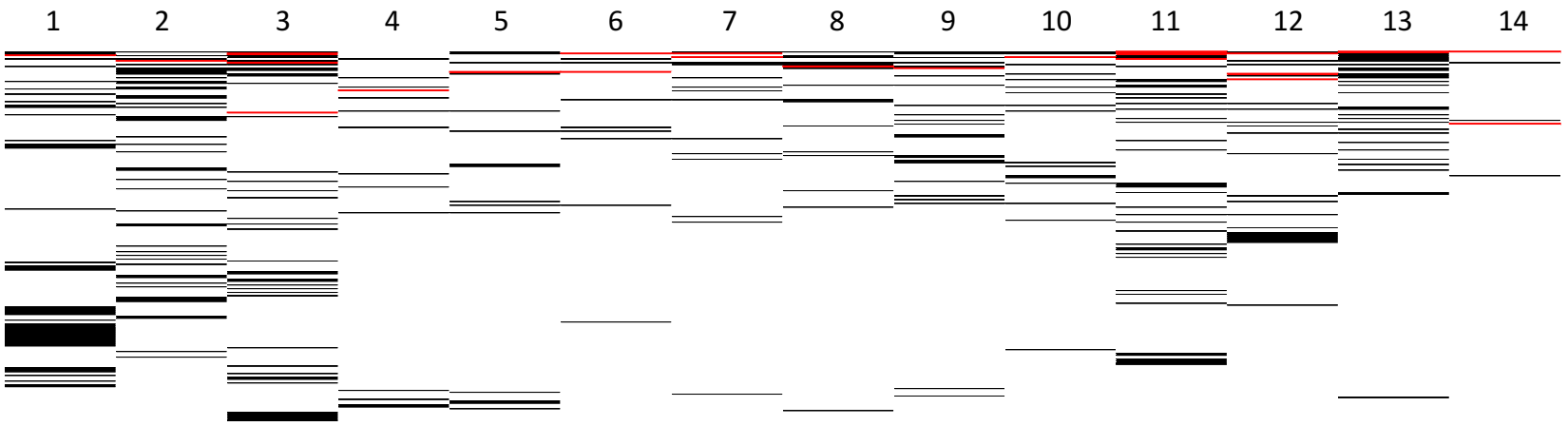
Appendix 13. Venn diagram of shared OTUs (97% 16S rRNA nucleotide identity) for isolates from sample time points T01 (09/15/13; 8-days post), T03 (09/23/13; 16-days post), T10 (11/11/13; 65-days post), and T12 (01/12/14; 127-days post).



Appendix 14. Venn diagram of shared OTUs (97% 16S rRNA nucleotide identity) for isolates from sample time points T01 (09/15/13; 8-days post), T02 (09/19/13; 12-days post), T03 (09/23/13; 16-days post), and T12 (01/12/14; 127-days post).



Appendix 15. Heat map showing OTU representation (97% 16S rRNA nucleotide identity) for sample sites: Ke'ehi Lagoon (A), La Mariana (B), Sand Island Launch Ramp (C), Sand Island Beach Park (D), Pier 37 – Fish Market (E), control contaminants (F). Red color indicates that the OTUs are shared between sample sites.



Appendix 16. Heat map showing OTU representation (97% 16S rRNA nucleotide identity) for sample time points. Red color indicates OTUs occurring at the same sample time points.

Appendix 17. Parsimony analysis of sample time points post-molasses spill.

Time Point	01	02	03	04	05	06	07	08	09	10	11	12	13	14
01		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
02	<0.0001		<0.0001	<0.0001	<0.0001	<0.0001	0.0003	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
03	<0.0001	<0.0001		<0.0001	<0.0001	0.0755	0.1939	0.0514	<0.0001	<0.0001	<0.0001	0.0010	<0.0001	0.0079
04	<0.0001	<0.0001	<0.0001		<0.0001	0.0164	0.0011	0.0033	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0033
05	<0.0001	<0.0001	<0.0001	<0.0001		0.1203	0.0145	0.0272	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
06	<0.0001	<0.0001	0.0755	0.0164	0.1203		0.1060	0.2890	0.0004	0.0072	0.0054	0.0345	<0.0001	0.0004
07	<0.0001	0.0003	0.1939	0.0011	0.0145	0.1060		0.3016	0.0004	0.0004	0.0007	0.0308	0.0024	0.0802
08	<0.0001	<0.0001	0.0514	0.0033	0.0272	0.2890	0.3016		0.1444	0.0216	0.0996	0.0873	0.0177	0.0622
09	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0004	0.0004	0.1444		<0.0001	<0.0001	<0.0001	<0.0001	0.0364
10	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0072	0.0004	0.0216	<0.0001		<0.0001	0.0002	<0.0001	<0.0001
11	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0054	0.0007	0.0996	<0.0001	<0.0001		<0.0001	0.0919	0.0071
12	<0.0001	<0.0001	0.0010	<0.0001	<0.0001	0.0345	0.0308	0.0873	<0.0001	0.0002	<0.0001		0.0016	<0.0001
13	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0024	0.0177	<0.0001	<0.0001	0.0919	0.0016		0.0002
14	<0.0001	<0.0001	0.0079	0.0033	<0.0001	0.0004	0.0802	0.0622	0.0364	<0.0001	0.0071	<0.0001	0.0002	

Appendix 18. Tentative taxonomic assignment of OTUs using the RDP Naïve Bayesian rRNA Classifier v2.6 (2013). (Taxonomic Level: P=Phylum, C=Class, O=Order, G=Genus.)

Bacteria OTU Assignment	Total	Actinobacteria	Actinobacteria	Knoellia	Janibacter	Ornithinimicrobium	Micrococcus	Cellulosimicrobium	Isoptericola	Piscicoccus	Georgenia	Brachybacterium	Cellulomonas	Agrococcus	Microbacterium	Brevibacterium	Rhodococcus	Gordonia	Mycobacterium	Dietzia	Nocardioides
Taxonomic Level	Total	P	C	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
Keehi Lagoon (A)	200	13	13	0	0	0	1	2	1	0	0	1	0	0	2	0	1	3	0	1	0
La Mariana (B)	225	9	9	0	1	0	0	3	0	0	0	0	0	0	1	0	0	1	2	0	0
Sand Island Launch (C)	339	71	71	0	0	1	0	9	1	1	1	0	4	2	21	1	7	4	2	5	0
Sand Island Beach Park (D)	84	9	9	3	1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Pier 37 - Fish Market (E)	233	24	24	0	0	0	0	2	0	0	0	0	0	0	3	1	4	2	10	0	1
Controls (F)	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Bacteria OTU Assignment	Aeromicrobium	Micromonospora	Streptomyces	Bacteroidetes	Unclassified "Bacteroidetes"	Flavobacteriia	Gelidibacter	Winogradskyella	Cytophagia	Pontibacter	Bacteroidia	Unclassified Marinilabiliaceae	Alkalitalea	Deinococcus-Thermus	Deinococci	Deinococcus
Taxonomic Level	G	G	G	P	P	C	G	G	C	G	C	F	G	P	C	G
Keehi Lagoon (A)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
La Mariana (B)	0	0	0	6	1	4	0	4	0	0	1	1	0	0	0	0
Sand Island Launch (C)	1	1	1	5	0	1	0	1	3	3	1	0	1	0	0	0
Sand Island Beach Park (D)	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
Pier 37 - Fish Market (E)	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0
Controls (F)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Appendix 18 (cont.): Tentative taxonomic assignment of OTUs using the RDP Naïve Bayesian rRNA Classifier v2.6 (2013). (Taxonomic Level: P=Phylum, C=Class, O=Order, G=Genus.)

Bacteria OTU Assignment	Proteobacteria	Betaproteobacteria	Achromobacter	Massilia	Variovorax	Alphaproteobacteria	Maricaulis	Unclassified Sphingomonadaceae	Sphingomonas	Erythrobacter	Afipia	Rhizobium	Cohaesibacter	Mesorhizobium	Pseudochrobactrum	Aurantimonas	Microvirga	Skermanella	Thalassospira	Kiloniella	Unclassified Rhodobacteraceae
Taxonomic Level	P	C	G	G	G	C	G	F	G	G	G	G	G	G	G	G	G	G	G	G	F
Keehi Lagoon (A)	50	0	0	0	0	5	0	0	0	0	0	0	0	0	3	0	0	0	0	0	1
La Mariana (B)	74	1	0	0	1	10	0	0	0	1	0	0	1	1	1	0	0	0	1	0	0
Sand Island Launch (C)	118	1	1	0	0	20	0	0	0	2	0	5	0	0	5	2	1	1	0	0	1
Sand Island Beach Park (D)	40	0	0	0	0	10	1	1	0	4	0	0	0	0	0	0	0	0	0	1	0
Pier 37 - Fish Market (E)	73	1	0	1	0	10	0	0	1	0	1	0	0	0	6	0	1	0	0	0	0
Controls (F)	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Bacteria OTU Assignment	Phaeobacter	Paracoccus	Marivita	Oceanicola	Loktanella	Thalassobius	Ruegeria	Pseudovibrio	Labrenzia	Gammaproteobacteria	Unclassified Pseudomonadales	Pseudomonas	Rheinheimera	Alcanivorax	Unclassified Oceanospirillaceae	Halomonas	Unclassified Enterobacteriaceae	Erwinia	Unclassified Xanthomonadaceae	Luteimonas	Lysobacter
Taxonomic Level	G	G	G	G	G	G	G	G	G	C	O	G	G	G	F	G	F	G	F	G	G
Keehi Lagoon (A)	0	0	0	0	0	1	0	0	0	45	0	0	0	0	0	2	1	0	0	0	1
La Mariana (B)	0	0	1	1	0	1	1	0	0	63	0	3	0	1	0	1	0	0	0	0	0
Sand Island Launch (C)	0	1	0	0	0	0	2	0	0	97	0	14	1	3	0	2	1	1	1	2	1
Sand Island Beach Park (D)	0	0	0	0	1	0	0	1	1	30	1	1	0	0	2	1	0	0	0	0	0
Pier 37 - Fish Market (E)	1	0	0	0	0	0	0	0	0	62	0	1	0	7	0	1	0	0	0	1	6
Controls (F)	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0

Appendix 18 (cont.): Tentative taxonomic assignment of OTUs using the RDP Naïve Bayesian rRNA Classifier v2.6 (2013). (Taxonomic Level: P=Phylum, C=Class, O=Order, G=Genus.)

Bacteria OTU Assignment	Xanthomonas	Stenotrophomonas	Pseudoxanthomonas	Unclassified Vibrionaceae	Vibrio	Photobacterium	Catenococcus	Unclassified Alteromonadales	Ferrimonas	Idiomarina	Unclassified Alteromonadaceae	Alteromonas	Marinobacter	Microbulbifer	Marinobacterium	Pseudoalteromonas	Psychrosphaera	Firmicutes	Unclassified Firmicutes	Bacilli	Unclassified Bacilli
Taxonomic Level	G	G	G	F	G	G	G	O	G	G	F	G	G	G	G	G	G	P	P	C	C
Keehi Lagoon (A)	0	0	0	4	27	1	0	1	4	0	0	0	3	1	0	0	0	137	0	137	1
La Mariana (B)	0	1	0	4	24	3	0	0	9	1	1	3	6	5	0	1	0	136	1	135	0
Sand Island Launch (C)	1	0	4	2	53	2	0	0	0	0	0	0	7	1	0	1	0	144	0	144	1
Sand Island Beach Park (D)	0	0	0	0	17	0	0	0	0	0	0	2	4	0	0	1	1	34	0	34	0
Pier 37 - Fish Market (E)	0	0	0	2	27	3	1	0	2	0	0	1	6	3	1	0	0	135	0	135	0
Controls (F)	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	3	0	3	0

Bacteria OTU Assignment	Unclassified Bacillales	Unclassified Bacillaceae 1	Bacillus	Halobacillus	Oceanobacillus	Terribacillus	Virgibacillus	Unclassified Paenibacillaceae 1	Cohnella	Brevibacillus	Paenibacillus	Exiguobacterium	Unclassified Planococcaceae	Rummeliibacillus	Sporosarcina	Lysinibacillus	Bhargavaea	Planomicrobium	Staphylococcus
Taxonomic Level	O	F	G	G	G	G	G	F	G	G	G	G	F	G	G	G	G	G	G
Keehi Lagoon (A)	0	1	113	9	0	0	0	0	0	0	3	0	0	1	2	7	0	0	0
La Mariana (B)	3	1	119	2	0	0	1	0	0	2	4	2	1	0	0	0	0	0	0
Sand Island Launch (C)	6	0	89	4	4	5	6	1	0	0	9	1	5	0	1	10	1	0	1
Sand Island Beach Park (D)	1	0	23	6	0	0	3	0	0	0	0	0	0	0	0	0	0	1	0
Pier 37 - Fish Market (E)	3	0	105	1	2	0	0	1	5	0	9	1	1	0	0	5	0	0	2
Controls (F)	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Appendix 19. Tentative taxonomic assignment of OTUs using the RDP Naïve Bayesian rRNA Classifier v2.6 (2013).

		Sampling Time Period (Days) Post-Spill													
		T1 (8d)	T2 (12d)	T3 (16d)	T4 (20d)	T5 (24d)	T6 (28d)	T7 (32d)	T8 (35d)	T9 (49d)	T10 (65d)	T11 (91d)	T12 (127d)	T13 (155d)	T14 (375d)
Tax. Level	Bacteria OTU Assignment	*n* = 98	*n* = 128	*n* = 144	*n* = 56	*n* = 64	*n* = 32	*n* = 42	*n* = 35	*n* = 73	*n* = 89	*n* = 96	*n* = 81	*n* = 115	*n* = 33
Phylum	Actinobacteria	4	16	19	18	12	9	5	4	1	13	6	15	4	0
Class	Actinobacteria	4	16	19	18	12	9	5	4	1	13	6	15	4	0
Genus	Knoellia	0	3	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Janibacter	0	0	0	1	1	0	0	0	0	0	0	0	0	0
Genus	Ornithinimicrobium	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Genus	Micrococcus	0	0	0	0	0	0	0	0	0	1	3	0	0	0
Genus	Cellulosimicrobium	2	1	1	3	1	1	1	0	1	1	1	1	2	0
Genus	Isoptericola	0	0	1	0	0	0	0	0	0	0	0	0	1	0
Genus	Piscicoccus	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Genus	Georgenia	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Genus	Brachybacterium	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Cellulomonas	0	0	0	0	0	1	0	1	0	0	0	2	0	0
Genus	Agrococcus	1	0	0	0	1	0	0	0	0	0	0	0	0	0
Genus	Microbacterium	1	0	2	3	6	1	1	1	0	6	0	6	0	0
Genus	Brevibacterium	0	0	1	0	0	0	0	0	0	1	0	0	0	0
Genus	Rhodococcus	0	0	4	0	2	4	1	0	0	0	0	1	0	0
Genus	Gordonia	0	0	5	1	0	0	1	0	0	3	0	0	0	0
Genus	Mycobacterium	0	8	2	2	0	0	0	0	0	0	0	1	1	0
Genus	Dietzia	0	1	0	3	0	2	0	0	0	0	0	0	0	0
Genus	Nocardioides	0	0	0	2	0	0	0	0	0	0	0	0	0	0
Genus	Aeromicrobium	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Genus	Micromonospora	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Genus	Streptomyces	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Phylum	Bacteroidetes	4	3	0	0	2	1	0	0	1	0	0	0	1	0
Phylum	Unclassified "Bacteroidetes"	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Class	Flavobacteriia	3	2	0	0	0	0	0	0	1	0	0	0	0	0
Genus	Gelidibacter	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Genus	Winogradskyella	3	2	0	0	0	0	0	0	0	0	0	0	0	0
Class	Cytophagia	0	0	0	0	2	1	0	0	0	0	0	0	0	0
Genus	Pontibacter	0	0	0	0	2	1	0	0	0	0	0	0	0	0
Class	Bacteroidia	0	1	0	0	0	0	0	0	0	0	0	0	1	0
Family	Unclassified Marinilabiliaceae	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Alkalitalea	0	0	0	0	0	0	0	0	0	0	0	0	1	0

Appendix 19 (cont.): Tentative taxonomic assignment of OTUs using the RDP Naïve Bayesian rRNA Classifier v2.6 (2013).

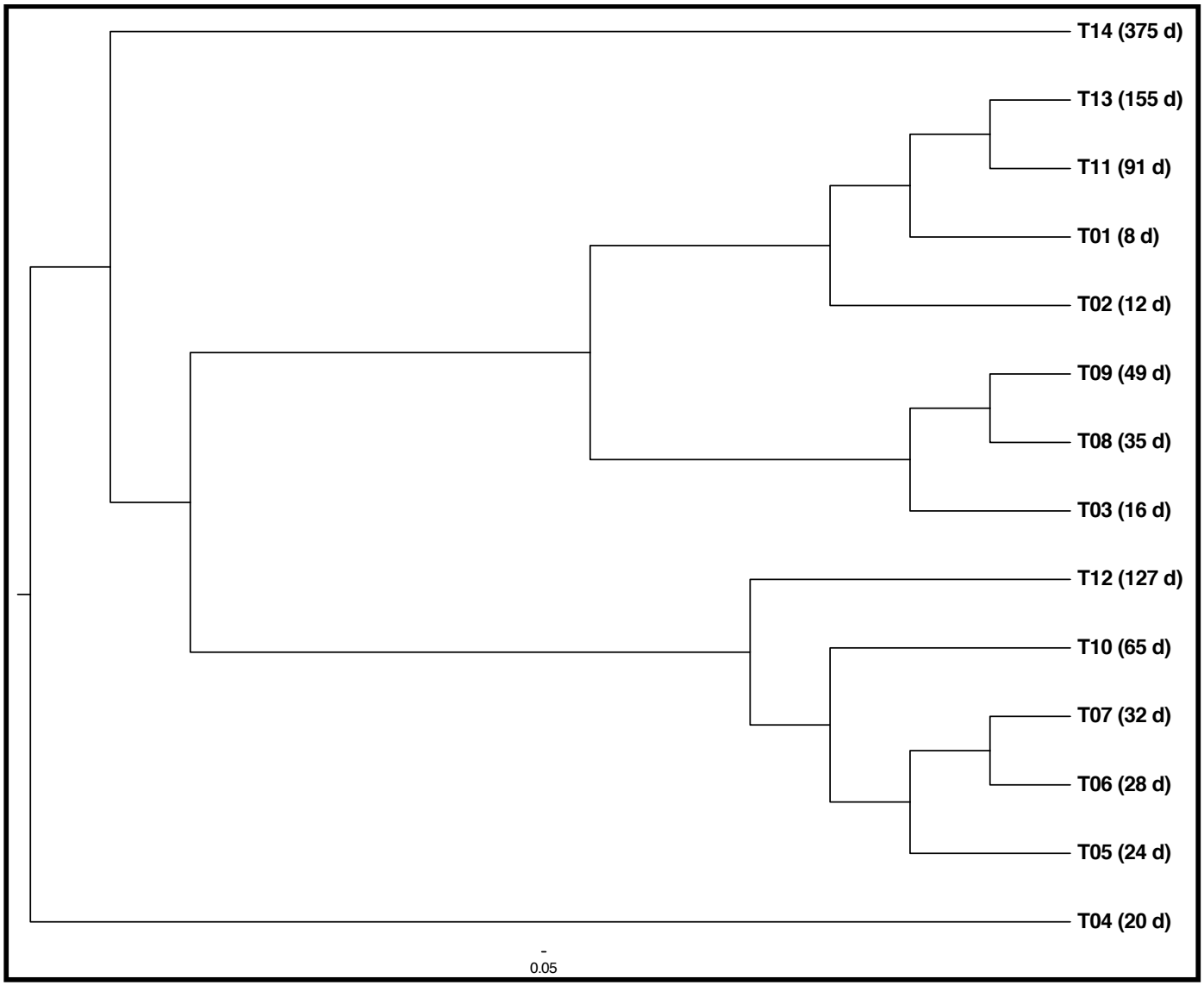
		Sampling Time Period (Days) Post-Spill													
Tax. Level	Bacteria OTU Assignment	T1 (8d)	T2 (12d)	T3 (16d)	T4 (20d)	T5 (24d)	T6 (28d)	T7 (32d)	T8 (35d)	T9 (49d)	T10 (65d)	T11 (91d)	T12 (127d)	T13 (155d)	T14 (375d)
Phylum	Proteobacteria	65	57	44	24	10	4	8	7	21	11	41	11	39	15
Class	Betaproteobacteria	0	0	0	2	0	0	0	0	0	0	0	0	1	0
Genus	Achromobacter	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Genus	Massilia	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Genus	Variovorax	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Class	Alphaproteobacteria	11	17	9	0	1	0	0	1	1	4	2	3	5	1
Genus	Maricaulis	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Family	Unclassified Sphingomonadaceae	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Sphingomonas	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Erythrobacter	1	4	2	0	0	0	0	0	0	0	0	0	0	0
Genus	Afipia	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Genus	Rhizobium	1	1	0	0	0	0	0	0	0	3	0	0	0	0
Genus	Cohaesibacter	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Mesorhizobium	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Genus	Pseudochrobactrum	0	1	5	0	1	0	0	0	0	1	2	2	3	0
Genus	Aurantimonas	0	2	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Microvirga	0	0	1	0	0	0	0	0	0	0	0	1	0	0
Genus	Skermanella	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Genus	Thalassospira	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Genus	Kiloniella	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Family	Unclassified Rhodobacteraceae	2	0	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Phaeobacter	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Paracoccus	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Marivita	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Oceanicola	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Loktanella	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Thalassobius	2	0	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Ruegeria	1	1	1	0	0	0	0	0	0	0	0	0	0	0
Genus	Pseudovibrio	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Genus	Labrenzia	0	1	0	0	0	0	0	0	0	0	0	0	0	0

Appendix 19 (cont.): Tentative taxonomic assignment of OTUs using the RDP Naïve Bayesian rRNA Classifier v2.6 (2013).

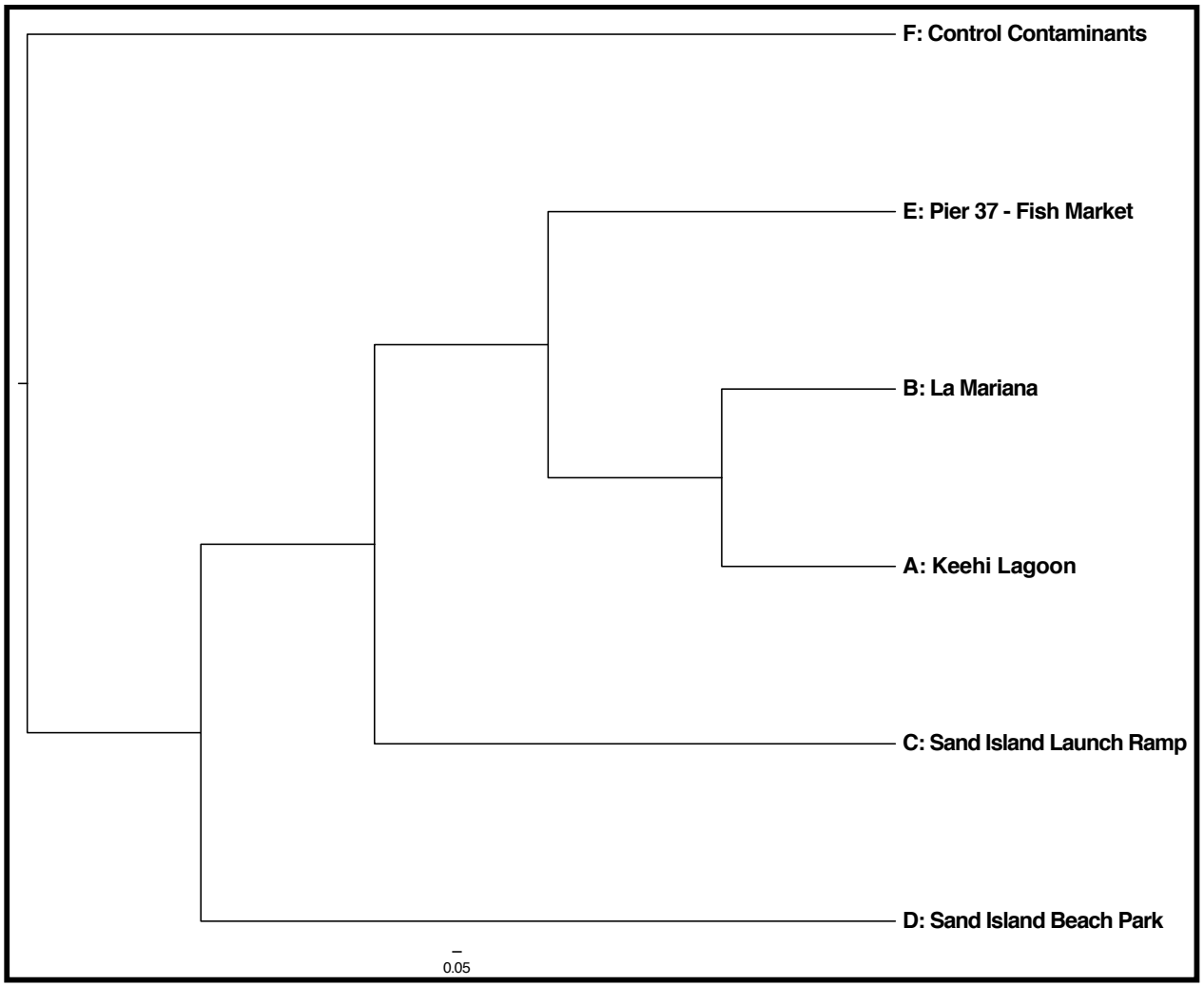
		Sampling Time Period (Days) Post-Spill													
Tax. Level	Bacteria OTU Assignment	T1 (8d)	T2 (12d)	T3 (16d)	T4 (20d)	T5 (24d)	T6 (28d)	T7 (32d)	T8 (35d)	T9 (49d)	T10 (65d)	T11 (91d)	T12 (127d)	T13 (155d)	T14 (375d)
Class	Gammaproteobacteria	54	40	35	22	9	4	8	6	20	7	39	8	33	14
Order	Unclassified Pseudomonadales	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Genus	Pseudomonas	0	0	0	5	3	1	1	0	0	0	1	3	5	0
Genus	Rheinheimera	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Alcanivorax	0	0	4	0	0	0	0	0	7	0	0	0	0	0
Family	Unclassified Oceanospirillaceae	0	1	0	0	0	0	0	0	0	0	0	0	1	0
Genus	Halomonas	0	0	1	1	2	0	0	1	1	0	0	0	1	0
Family	Unclassified Enterobacteriaceae	0	0	1	0	0	0	0	0	0	1	0	0	0	0
Genus	Erwinia	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Family	Unclassified Xanthomonadaceae	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Genus	Luteimonas	0	0	1	2	0	0	0	0	0	0	0	0	0	0
Genus	Lysobacter	1	5	0	1	1	0	0	0	0	0	0	0	0	0
Genus	Xanthomonas	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Genus	Stenotrophomonas	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Genus	Pseudoxanthomonas	0	1	0	2	1	0	0	0	0	0	0	0	0	0
Family	Unclassified Vibrionaceae	3	2	0	1	0	0	1	0	0	0	2	0	3	0
Genus	Vibrio	37	16	17	8	0	1	6	4	2	3	27	4	18	7
Genus	Photobacterium	0	1	1	0	0	0	0	0	0	0	5	0	1	1
Genus	Catenococcus	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Order	Unclassified Alteromonadales	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Ferrimonas	4	6	2	1	0	0	0	0	0	0	0	0	1	1
Genus	Idiomarina	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Family	Unclassified Alteromonadaceae	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Alteromonas	2	1	1	0	0	0	0	0	0	0	1	0	0	1
Genus	Marinobacter	0	1	5	1	0	1	0	1	9	1	1	0	2	4
Genus	Microbulbifer	4	3	1	0	0	1	0	0	0	1	0	0	0	0
Genus	Marinobacterium	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Genus	Pseudoalteromonas	1	1	0	0	0	0	0	0	1	0	0	0	0	0
Genus	Psychrosphaera	0	0	0	0	0	0	0	0	0	0	0	0	1	0

Appendix 19 (cont.): Tentative taxonomic assignment of OTUs using the RDP Naïve Bayesian rRNA Classifier v2.6 (2013).

		Sampling Time Period (Days) Post-Spill													
Tax. Level	Bacteria OTU Assignment	T1 (8d)	T2 (12d)	T3 (16d)	T4 (20d)	T5 (24d)	T6 (28d)	T7 (32d)	T8 (35d)	T9 (49d)	T10 (65d)	T11 (91d)	T12 (127d)	T13 (155d)	T14 (375d)
Phylum	Firmicutes	25	51	81	14	40	18	29	24	50	65	49	54	71	18
Phylum	Unclassified Firmicutes	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Class	Bacilli	25	51	81	14	40	18	29	24	50	65	49	54	71	17
Class	Unclassified Bacilli	0	0	0	0	0	1	1	0	0	0	0	0	0	0
Order	Unclassified Bacillales	0	1	1	0	0	2	0	1	2	1	0	3	1	1
Family	Unclassified Bacillaceae 1	0	0	0	0	0	0	0	0	0	2	0	0	0	0
Genus	Bacillus	14	45	66	10	27	11	23	22	41	51	43	36	53	10
Genus	Halobacillus	9	1	1	0	5	3	0	0	0	2	1	0	0	0
Genus	Oceanobacillus	0	0	0	0	0	0	0	0	4	0	0	0	0	2
Genus	Terribacillus	0	0	2	0	0	0	0	0	0	3	0	0	0	0
Genus	Virgibacillus	0	0	1	0	3	1	2	0	0	2	0	0	0	1
Family	Unclassified Paenibacillaceae 1	0	0	2	0	0	0	0	0	0	0	0	0	0	0
Genus	Cohnella	0	0	0	0	5	0	0	0	0	0	0	0	0	0
Genus	Brevibacillus	0	0	0	0	0	0	0	0	0	0	0	2	0	0
Genus	Paenibacillus	0	2	4	1	0	0	2	0	0	2	1	7	6	0
Genus	Exiguobacterium	0	0	1	0	0	0	0	0	0	1	0	0	2	0
Family	Unclassified Planococcaceae	0	0	0	0	0	0	1	0	0	0	0	1	5	0
Genus	Rummeliibacillus	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Genus	Sporosarcina	0	0	0	0	0	0	0	0	0	0	2	1	0	0
Genus	Lysinibacillus	2	1	2	1	0	0	0	1	3	1	1	4	4	2
Genus	Bhargavaea	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Genus	Planomicrobium	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Genus	Staphylococcus	0	1	0	2	0	0	0	0	0	0	0	0	0	0
Phylum	Deinococcus-Thermus	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Class	Deinococci	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Genus	Deinococcus	0	1	0	0	0	0	0	0	0	0	0	0	0	0



Appendix 20. Bray-Curtis distance matrix describing the similarity of community structure between sample time points (days post-spill) using the UPGMA algorithm.



Appendix 21. Bray-Curtis distance matrix describing the similarity of community structure between sample sites using the UPGMA algorithm.