Defining the seafloor microbiome of the Gulf of Mexico and its response to oil perturbation

Will A. Overholt[1], Patrick Schwing[2], David Hastings[3], David J. Hollander[2], Joel E. Kostka[1,4]

[1]School of Biology, Georgia Institute of Technology, Atlanta GA; [2]College of Marine Science, University of South Florida, St. Petersburg, FL; [3]Department of Marine Science, Eckerd College, St. Petersburg, FL; [4]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA

Keywords: biogeography, deep ocean sediment, microbial community, Deepwater Horizon, response and recovery, perturbations

Running title: Deepsea benthic microbial biogeography

Abstract

The microbial ecology of oligotrophic deep ocean sediments is understudied relative to their shallow counterparts, and this lack of understanding hampers our ability to predict responses to current and future perturbations. The Gulf of Mexico has experienced two of the largest accidental marine oil spills, i.e., the 1979 Ixtoc-1 blowout in the Southern Gulf of Mexico (GoM) and the 2010 Deepwater Horizon (DWH) discharge in the Northern GoM. Microbial communities were characterized via next generation sequencing of SSU rRNA gene sequences for 29 sites across multiple years in the Gulf of Mexico, represented by >700 samples. The distribution of seafloor microbial communities was elucidated and found to be surprisingly consistent across the entire region in terms of the OTUs detected and their relative abundances. The composition of the seafloor microbiome was well approximated by the overlying water depth and depth within the sediment column, which together explained 38% of the observed variation. In contrast, geographic distance had a limited role and explained only 6%. Biogeographical distributions were used to generate a depth-stratified machine-learning based predictive model for over 4000 dominant OTUs that relies on easy-to-obtain geospatial variables. Microbial community structure is linked to oxygen penetration depth and sediment geochemical regime, which are likely controlled through carbon delivery. Our results further demonstrate that sediments impacted by the DWH spill had returned to near baseline conditions after two years. The distributions of key microbial populations can now be calculated and constrained across the region while deviations from these predictions may be evaluated to pinpoint impacted sites, and more importantly, in future response efforts or long-term stability studies.

Introduction

The deep ocean floor at greater than 1000 m water depth covers approximately 50% of the Earth's surface. The seafloor of the deep ocean is comprised predominately of muddy sediments, and solute transport rates in these sediments are dominated by diffusion, while larger particles are transported through burial (Burdige, 2006). The biogeochemistry regimes of deep ocean sediments are largely governed by the quality and quantity of deposited organic matter, and nearly all organic matter introduced into these sediments originated in the photic zone via photosynthesis or via the transport from terrestrial sources (Arndt et al., 2013; Jørgensen and Boetius, 2007). Furthermore, in deep ocean environments, the majority of organic matter is decomposed while traveling through the water column prior to reaching the seafloor (Canfield et al. 2006). Therefore, oligotrophic sediments typically have very low organic carbon contents and <20% is recognizable as one of the main biomolecular classes (Canfield et al. 2006; Teske 2013).

In contrast to the complexity and patchiness of the sedimentary carbon content, the oligotropic deep ocean is characterized by consistent physiochemical gradients (temperature, pressure, salinity, pH) that are often noted as the primary variables in controlling bacterial diversities in other environments (Bienhold et al., 2012; Lozupone and Knight, 2007). In part, this temporal and spatial stability is thought to contribute to the higher diversity and lower variation in community structure observed in deep marine sediments compared to their coastal and pelagic counterparts (Bienhold et al., 2016; Jørgensen and Boetius, 2007; Zinger et al., 2011). Additionally, bacterial communities sampled from deep benthic marine sediments across the globe have thus far demonstrated few cosmopolitan groups with most populations geographically limited to one sample or one ocean region (Bienhold et al., 2016). However, apparent cosmopolitan populations detected from relatively sparse datasets generally consist of the most dominant groups detected, namely the JTB255 clade (Gammaproteobacteria, Xanthomonadales) and the OM1 clade (Actinobacteria, Acidimicrobiales) (Bienhold et al., 2016). The main environmental

variables identified driving community composition are the total organic carbon content, water depth, geographic distance, and overlying ecosystem productivity (Bienhold et al., 2016; Wei et al., 2010; Zinger et al., 2011).

While deep marine sediments are thought to be relatively stable, they are not impervious to ecological disturbances including phytoplankton and marine snow deposition (Billett et al., 1983), whale-falls (Smith and Baco, 2003), biogenic burial (Kukert and Smith, 1992), and sediment transport (Aller, 1989). In macrofaunal communities, these processes along with slow recovery rates contribute to a high level of diversity and patchy composition (Ebbe et al., 2010; Kukert and Smith, 1992). Much less is known about how benthic microbial communities respond to disturbances.

One glaring example of such a disturbance is the Deepwater Horizon oil spill (DWH), which represents the largest accidental marine oil spill in history, released an estimated 4.1 – 4.9 million barrels of crude oil and a third as much natural gas at approximately 1500 m water depth (Joye et al., 2011; Lubchenco et al., 2010). Although the amounts and chemical composition of DWH oil remaining in the environment remain uncertain, the largest reservoir is thought to be primarily in deep ocean sediments with a considerable, but unknown, amount deposited near coastal sediments or onshore (King et al., 2014). Approximately 12% of the 2 million barrels entrained in the deep ocean was transported to a 3200 km^2 region of the seafloor (Joye, 2015; Valentine et al., 2014). The fate of this oil will largely be determined by microbial metabolisms in these environments and by burial.

A substantial challenge in determining the impacts of this oil being deposited into deep ocean sediments is the nearly complete lack of studies on benthic microbial communities performed prior to DWH. Due to interest in the Gulf for petroleum exploration, previous studies focused almost entirely on deep subsurface sediments with the shallowest samples collected at ~4 meters below the seafloor, or on sediments impacted by natural hydrocarbon seeps (Biddle et al., 2011; Nunoura et al., 2009; Orcutt et al., 2010). Due to this limitation in the knowledge of surficial microbial communities, the majority of

studies on deep ocean sediments associated with the DWH event used samples that were collected outside of impacted areas as controls (Kimes et al., 2013; Liu and Liu, 2013; Mason et al., 2014; Yang et al., 2016). Sediments were primarily impacted as deep water oil intrusion layers impinged on the seafloor and by a large sedimentation event of marine oil-derived snow (Brooks et al., 2015; Hollander et al., 2013; Passow et al., 2012; Romero et al., 2015; Valentine et al., 2014). While the datasets on benthic microbes are limited to relatively few samples, results indicate that microbial communities responded quickly to oil perturbation and shifts in community composition as well as total metabolic potential were observed (Kimes et al., 2013; Mason et al., 2014; Yang et al., 2016). While there are conflicting reports and few samples (< 10), sedimentary microbial communities may have returned to baseline conditions within one year following the spill (Liu and Liu, 2013; Yang et al., 2016).

With some similarities to the Deepwater Horizon oil spill, the Ixtoc-I well blowout that began in 1979, released approximately 3.5 million barrels of crude oil, and represents the second largest accidental marine oil spill in history (Jernelöv and Lindén, 1981). Similar to the petroleum hydrocarbons released during the DWH disaster, Ixtoc-I crude oil was saturated with gaseous hydrocarbons and had a similar chemical composition. While the Ixtoc-I blowout occurred in shallow water (52 m water depth), subsurface plumes also appear to have formed in some regions owning to temperature and salinity density stratifications. Also similar to the DWH blowout, oil appears to have been transported to the seafloor through marine oil-derived snow and approximately 25% of the total oil released during Ixtoc is thought to have remained in seafloor sediments (Jernelöv and Lindén, 1981). Unfortunately, the long-term fate of this Ixtoc-I oil is not well understood (Schrope, 2010), and due to technical limitations at the time, little is known about the *in situ* microbial community response (Atlas, 1981).

Here, we present the largest benthic marine microbial community study conducted to date, and by far the most extensive in the Gulf of Mexico. We focus on two regions in the Gulf of Mexico that were both impacted by massive marine oil spills. Over 700 samples were analyzed representing 29 distinct

sites, across up to four years, and we generated >120 million SSU rRNA gene sequences, including both archaeal and bacterial members. Our objectives were to (1) characterize un-impacted sedimentary microbial communities to establish a baseline, (2) map the biogeographical patterns in microbial community structure across the Gulf of Mexico, (3) using this map, generate a Gulf of Mexico biogeography model of microbial community structure that can predict the abundance of dominant microbial populations, and (4) determine if impacted regions had returned to baseline conditions. For the first time on a region-wide scale, we determine the environmental parameters structuring microbial communities, including the stratification of microbial communities according to sediment depth and across oxygen gradients. In addition, we address the temporal stability of microbial communities across the region.

Material and Methods

Sediment Sample Collection

Samples were collected during five research cruises from 2012 to 2015 across the Northern and Southern Gulf of Mexico (see Table S1, Figure 1). Sediment samples were collected using a MC-800 multicorer capable of collecting up to eight, 10-cm diameter by 70 cm long cores per deployment. Cores were sectioned at either sub-centimeter (2 mm sections) or 1 cm resolution using a calibrated threaded rod attached to a core clamp stand with a tight fitting plunger (see Table S1 for the full sample list) (Schwing et al., 2016). Sectioned sediments were frozen immediately in a portable -80 °C freezer onboard the vessel and transported back to the laboratory for DNA extraction.

In the Northern Gulf of Mexico (GoM), we collected sediment samples across approximately 100,000 square kilometers, primarily focusing on the Desoto Canyon feature and in close proximity to the Deepwater Horizon well-head (Table S1, Figure 1). Sampling sites ranged from 56 m (SL1040) near the Mississippi River Delta to 2293 m (DSH09) southeast of the DWH wellhead, spanning three years (2012 – 2014) and four sampling time points (Table S1). In the Southern Gulf of Mexico, sampling sites spanned approximately 40,000 square kilometers ranging from 16-1470 m, primarily focusing on the Ixtoc-I well blowout site and the western transect (Figure 1) (Soto et al., 2014). Cores from the southern GoM were all collected from one cruise in 2015. For each region, sediment core profiles extend to at least 10 cm depth below the seafloor, with a subset including sections up to 20 cm of the sediment column. Our experimental design incorporated multiple levels of replication to address variation due to sampling effort, as well as spatial and temporal heterogeneity. These include: (1) technical replicates with duplicate DNA extractions performed and separate libraries generated on the same depth section, (2) multicore cast replicates where triplicate cores from the same cast were analyzed separately representing variation across approximately 1m^2 of the seafloor, (3) replicates at the site level where

triplicate casts were deployed at the same site and cores from cast was analyzed separately, and (4) temporal replicates where the same site was sampled up to 4 times in 3 years (Figure S1).

Oxygen Profiles

Oxygen concentration profiles were determined using a Presens Microx4 needle-microoptode (Presens, Regensburg, Germany) along with a CHPT micromanipulator (Composite High Pressure Technologies, Lewes, Delaware). Immediately after the cores were retrieved from the multicorer, a core was mounted on the extruder for stabilization and oxygen was profiled at mm-scales until the sensor measured 0 umol/L oxygen concentration. This was typically accomplished in < 15 minutes, although deeper sites required longer profiling times. Between sites, the optode was calibrated following the manufacturer's protocol using sodium sulfite for the 0% solution and bubbling with air for the 100% solution. The depth of oxygen penetration was imported into Ocean Data View v4.6.3 (Schlitzer, 2012) and predictive surfaces were generated using DIVA gridding with automatic scale lengths and with a signal to noise ratio of 50.

DNA Extraction, SSU rRNA Gene Sequencing and Processing

Sediment was sub-sectioned using a flame sterilized drillbit and total genomic DNA extractions performed on approximately 0.25 g using a MoBio Powersoil DNA extraction kit following the manufacturer's protocol. Samples collected in 2012 and 2013 were extracted using the MoBio Powersoil Highthroughput extraction kit, while those from 2014 and 2015 were extracted using the single-prep kit. The V4 region of the SSU rRNA was amplified using the commonly used 515F/806R primer set that targets both Archaeal and Bacterial domains and sequenced on an Illumina MiSeq with 2x250 v2 chemistry (Caporaso et al., 2012; Green et al., 2015, see Supplementary Information). All paired-end reads were merged using PEAR v0.9.6 with default parameters (Zhang et al., 2013) and assembled reads were quality filtered in QIIME using a minimum phred score of 30 (Q=30) (J. Caporaso et al., 2010). Mothur was used to remove sequences less than 250bp or greater than 255bp (Schloss et al., 2009). D*e*

novo based chimera detection followed by reference based chimera detection against the SILVA gold database was performed with vsearch v1.9.6 (Rognes et al., 2016). The remaining high quality sequences were dereplicated using vsearch v1.9.6 and operational taxonomic units (OTUs) were picked using Swarm v2.1.8 with default parameters (Mahé et al., 2014; Rognes et al., 2016). Overall patterns were similar as determined with uclust_reference and de novo based methods, but these would not complete on the full sample set (data not shown, a brief discussion is available here: http://waoverholt.github.io/testing_otu_picking/).

Singletons and OTUs present in less than 10 samples were removed, producing approximately 300,000 OTUs using 47 million reads (75% of initial raw reads). Samples with less than 5000 sequence counts were removed from further analysis. Taxonomic affiliations for each OTU were determined using the SILVA v123 database formatted for use with QIIME (J. G. Caporaso, Kuczynski, et al., 2010; Quast et al., 2012). Representative sequences were aligned using the PyNAST implementation in QIIME with the SILVA alignment and a phylogenetic tree for determining phylogenetic diversity metrics was generated with fasttree v2.1.3 (J. G. Caporaso, Bittinger, et al., 2010; J. Caporaso et al., 2010; Price et al., 2010). All raw sequences have been uploaded to NCBI under Bioproject PRJNA414249.

OTU Analysis

OTU counts were normalized using the metagenomeSeq package in R, scaling by 1000 and without a log-transform (Paulson et al., 2013). Trends were confirmed with rarefaction to 5000 sequences per sample (data not shown). Beta diversity was determined using Bray Curtis dissimilarity implemented in the R package 'vegan' (Oksanen et al., 2007; R Core Team, 2014), while weighted and unweighted Unifrac distances were calculated using the implementation found in the R package 'phyloseq' (McMurdie and Holmes, 2013). Patterns were consistent and Bray Curtis dissimilarities are shown. Nonmetric multidimensional scaling (NMDS) ordination was generated using the R package 'vegan' without transformation and patterns were indistinguishable to those observed using Principal

Coordinates Analysis (PCoA) also generated with 'vegan' (Oksanen et al., 2007; R Core Team, 2014). Patterns in Bray Curtis dissimilarities were tested using multiple linear regressions in R (R Core Team, 2014). The most significant regression used the formula (dissimilarity ~ sediment depth difference * water depth difference * geographic distance * Sampling Year). Linear regression models were compared using Analysis of Variance (ANOVA), and the percent each variable contributed to the patterns observed was determined with ANOVA (Table S2, Table S3). While all interaction terms were significant, only those that contributed >1% variance were retrained and the model was reconstructed after dropping those terms to reduce its complexity (dissimilarity ~ sediment depth difference + water depth difference + geographic distance + sampling year + sed:water). All plots generated use the ggplot2 package in R (R Core Team, 2014; Wickham, 2009).

Meta-Analysis: Incorporating Impacted Surficial Sediments

Sequences generated by Mason and colleagues targeted the same region of the 16S rRNA gene, although sequence lengths were substantially shorter (on average 100 bp after Q30 trimming) (Mason et al., 2014). These sequences were incorporated into this study by using our previously identified OTUs as reference sequences and using the QIIME script 'parallel_pick_otus_uclust_ref.py'. Without chimera checking (as the references were already cleared) we matched 75% of all Mason 16S sequences (20 million reads) to one of our OTUs at 3% sequence difference or less. Since the Mason et al., 2014 study used only surficial sediments (top 1 cm), we restricted our comparison samples to include only depths in the same interval.

Constructing a Gulf-Wide Model for Microbial Community Composition

A model for microbial community composition across the Gulf of Mexico was constructed using random forest regression implemented in the R package 'randomForest' at the OTU level (Liaw and Wiener, 2002). Random Forest regression is an ensemble machine learning method where multiple individual decision trees are constructed and ultimately averaged to minimize over-fitting and improve

prediction accuracy. Individual decision trees are generated from a subset of the response variable (OTU abundance) and each fork or node in the tree is guided by predictor variables (again a randomized subset) to maximize differences in the branches. Random Forest is robust to outliers, does not assume specific distributions (nonparametric), and tests/trains the model by subsampling the response variable at each step and testing the decision node with the remaining data (Wei et al., 2010).

Water depth, sediment depth, latitude, and longitude were chosen as predictor variables and were used as proxies for sediment carbon input, oxygen concentrations, geological regime shifts from E-W and N-S as well as riverine inputs. Latitude, longitude, and water depths were extracted from the NGDC Coastal Relief model (Divins and Metzger, 2008). The Random Forest regression model for each response variable was constructed with 500 trees, a minimum of 5 terminal nodes per tree, and randomly subsampling 1/3 of the response variable (OTU normalized counts) for each node of each tree. Only response variables of which >50% of the variance could be explained by the regression forest were retained. The effectiveness of the model was evaluated by randomizing the predictor variables 1000 times and comparing the total root mean squared error calculated using:

$$RMSE = \sqrt{\frac{\sum_1^n (\hat{x}-x)^2}{n}} \quad (1)$$

where $\hat{x}$ is the predicted OTU relative abundance, x is the observed OTU relative abundance, and n is the number of OTUs. In this calculation, the OTU abundance from every sample is used, giving one value for the trained model, and one value for each of the 1000 randomized models. The RMSE of the trained model was compared to the randomly constructed models using a Z-test. Additionally, individual sample deviations from the OTU predictions were determined using equation (1). Instead of evaluating the model, this was used to identify samples that were not representative oligotrophic Gulf of Mexico sediments (e.g. those that differed from baseline). All the code and raw data that was generated for this

study can be found at GRIIDC (https://data.gulfresearchinitiative.org) and can be accessed at (doi: 10.7266/N70G3HN3).

Results and Discussion

Environmental parameters structuring benthic microbial distributions across the Gulf of Mexico

In the Northern Gulf of Mexico (GoM), we collected sediment samples across approximately 100,000 square kilometers, and in the Southern Gulf of Mexico samples span approximately 40,000 square kilometers. Sampling sites ranged from 16 to 2300 m water depth, focusing on regions adjacent to the Deepwater Horizon (DWH) and Ixtoc I blowouts. As part of this study, we address variation in community structure at different spatial resolutions (Figure S1). Technical and cast level (1 m^2 bottom area) replicates showed similar Bray-Curtis dissimilarity values (mean = 0.23, sd= ± 0.05; 0.25 ± 0.05, respectively) and site level replicates (10-100 m^2) were less similar (0.34 ± 0.09). Sampling temporally (across 3 years), lower similarity (0.54 ± 0.12) was observed, although this value may be inflated due to a small change in DNA extraction technique between 2013 and 2014 sampling years. All sites sampled at greater than 500 m water depth for the entire study show a mean dissimilarity of 0.61 ± 0.15 in comparison to a mean dissimilarity of 0.51 ± 0.13 for sites sampled at less than 100 m. However, the shallow sites were primarily sampled from the SGoM during the same year (Figure 2 C). We note that our data suggest these communities to be more similar than other studies, likely since we are sampling the same relatively small ocean basin and our analyses employ a more restrictive OTU abundance filtering, although our ranges overlap with observations by Zinger et al. (2011) conducted over larger spatial scales in multiple oceans.

Overall, the distribution of microbial populations with sediment depth across these geographically distant regions of the Gulf were remarkably consistent, and were structured primarily based on overlying water depth along sediment core depth (Figure 2 A, B). Our sampling design focused primarily on shallow sites (<100 m) and deep sites (>1000 m), and thus the down core sediment progressions seen in Figure 2 A are structured into two main clusters primarily along the Y-axis. However, a continuous gradient with increasing water depth is expected, and indeed such a gradient is

observed at the sites >1000m water depth. Performing a multiple linear regression analysis and partitioning the variance resulted in water depth explaining 22.1% of the total variation ($F = 1.29e5$, $p < 2e-16$, Table S2).

The effect of sediment depth is primarily observed along the x-axis with surficial sediment samples clustering in the upper right of Figure 2 proceeding down core towards the bottom left, consistent with other studies (Figure 2 B) (Teske, 2013). Sediment depth explained 14.2% of the total microbial community variation ($F = 8.3e4$, $p < 2e-16$), and together with water depth, could explain 36% of our observed variance (74% of our total explained variance, Table S2). In contrast, geographic distance was of limited effectiveness in distinguishing microbial community compositions, explaining just 6% of the variance ($F = 3.6e4$, $p < 2e-16$). If we change this to a discrete variable (NGoM vs SGoM), it instead explains 4% of the variance ($F = 3.2e4$, $p < 2e-16$) and this is visualized in Figure 2 C. We note that we have fewer core profiles from the southern GoM, all collected during 2015. Similarly, we found the sampling time point explained 5% of the variance ($F = 2.9e4$, $p < 2e-16$) and differences were not visible in Figure 2 D. Due to a slight change in DNA extraction technique, the change in microbial communities with time may be partially or solely a technical issue. In total, water depth, sediment depth, geographic distance, sampling time point, and their interactions explained 52% of the observed variation. Simplifying this model by removing variables that contributed little to the percent variance explained (see methods), 49% of the observed variance can be explained (Table S2). Examining patterns in alpha diversity (Shannon, estimated OTUs, rarefaction curves) revealed lower metrics in samples collected at sites less than 100 m than those collected at greater than 400 m water depth, while changes in sediment depth were minor (Figure S2).

This study represents the largest marine benthic biogeography study to date based on sample size and sequencing effort by a significant margin. In addition, this sampling effort was focused on a relatively small ocean region in comparison to similar studies, which provides both greater spatial

resolution, while the focus on distinct regions in the Gulf of Mexico allowed for better generalization across the basin. One of the better studied regional ocean basins, with respect to biogeographic patterns in microbial community structure is the Arctic ocean (Bienhold et al., 2012; Jacob et al., 2013). Jacob et al. (2013) used spatial variables similar to this study consisting of latitude, longitude, geographic distance, and water depth at the Arctic Long-Term Ecological Research (LTER) site to constrain microbial community variation in 13 surficial sediment samples collected along two transects. Using redundancy analysis, Jacob et al. (2013) found that bacterial community structure significantly varied with water depth and longitude (7% and 3% of the variance, respectively), while geographic distance was not significant. Somewhat surprisingly, a proxy for phytodetrius input (CPE) explained less variation than water depth, although this was potentially due to the low sampling effort. Also in the Arctic, Bienhold et al. (2012) linked microbial community structure and diversity to energy availability in benthic sediments using pyrosequencing on 10 samples and fingerprinting methods on 42 samples. Contrasting with the results from this study and with Jacob et al. (2013), the authors linked changes in community composition primarily with geographic distance (10% of the variation), followed by sediment pigment concentrations (5%). Sediment depth and water depth had low explanatory power, although again, this may have been due in part to a much smaller sample set (Bienhold et al., 2012). In the southern Atlantic Ocean, a study by Schauer et al. (2010) examined biogeographic patterns in deep marine surficial sediments across 3 basins (Cape, Angola, and Guinea) generating three SSU rRNA clone libraries and analyzing another 51 samples using a fingerprinting method. Their results showed that microbial communities were significantly different between the basins examined and that within a basin there was a spatial effect. In another regional study of surficial benthic microbial communities, the authors sampled 24 surficial sediments in two geographically distinct regions in a 5500 km transect in the Southern Ocean off the coast of Western Antarctica (Learman et al., 2016). This study consisted of SSU rRNA gene amplicons generated from the same primers used in this study and found that microbial

community composition was linked to the bioavailability of carbon and nutrient availability. However, this finding was confounded by large physiochemical differences between the distinct regions, that had unique community compositions (Learman et al., 2016).

On a global scale, water depth appears to have less explanatory power than observed in this study, while geographic distance tends to have more power (Bienhold et al., 2016; Danovaro et al., 2016; Zinger et al., 2011). This phenomenon is well documented, poorly understood, and is unlikely to be related to a breakdown in the correlation between water depth and the quality and quantity of carbon inputs (Danovaro et al., 2016 and the refs within). In a global study, Zinger et al. (2011) constrained 22 % of the benthic microbial community variation and found that geographic distance best explained surficial microbial community variation at 6 %, similar in explanatory power to the results in this study. Zinger et al. (2011) also found that coastal benthic microbial communities varied more than their deep ocean counterparts, which is the opposite of what was observed in this study (Figure S1). This could be due to several reasons including our shallow samples were mostly retrieved from the same year and sampled more extensively in one geographic region. In a follow-up, global-scale study focusing exclusively on deep marine benthic microbial communities (>1000m depths), (Bienhold et al., 2016) observed that microbial communities were increasingly distinct with increasing geographic distance, and this distance-decay relationship was even greater when one considers water flow paths instead of straight lines between sites. Total organic carbon explained the most variation in microbial community structure (10%) followed by geographic distance (4%), although the sample set consisted of 1-6 samples per ocean region (Bienhold et al., 2016). Using non-sequence based techniques, Danovaro et al. (2016) found that the absolute abundances of bacteria, total archaea, Marine Group I Archaea, and Marine Group II Archaea did not significantly vary within the same ocean basin, however, there were differences between ocean basins. Danovaro et al. (2016) also demonstrated that archaeal and bacterial abundances were partially controlled by the same factors, namely biopolymeric carbon concentrations,

and bacterial abundances were primarily influenced by carbon inputs while archaeal abundances were primarily influenced by bottom water temperature.

The results presented herein greatly contribute to the growing body of knowledge surrounding the environmental parameters structuring the seafloor microbiome. In contrast to previous work, this is the first study to examine large scale biogeographic patterns within a sediment core-profile in the deep ocean. Future global studies are encouraged to explore constraining the variation in microbial communities using a basin-based nested approach to improve the statistical models. However, sufficient samples from the region need to be collected in order to explore basin-wide biogeographic patterns. Here we demonstrate within the Gulf of Mexico water depth explains a relatively high percent of total variation compared to similar studies even though we do not focus solely on surficial sediments. This result is likely due to close coupling with organic matter inputs in the region (discussed below).

Sedimentary oxygen concentrations throughout the Gulf of Mexico

Sedimentary oxygen profiles were determined for 42 sites across the northern and southern regions of the Gulf of Mexico using a micro-optode and micromanipulator at millimeter resolution. Oxygen profiles were obtained at the majority of sites where cores were sectioned for microbial community composition analysis, as well as over a substantially larger area (Figure 1, Table S1). Typical core profiles are found in Figure 4 (B, D), while the interpolated maximum oxygen penetration depths across the study area are shown in Figure 3 A. At shallow, neritic sites, oxygen was depleted within the first millimeter, while oxygen penetration depths extended to 16 cm at the deepest sites (~ 3700 m). As anticipated, oxygen penetration increased nearly linearly with increasing water depths (Figure 3 B, Wenzhöfer and Glud, 2002). This likely results from a greater amount of phytodetritus reaching the seafloor in shallower environments as well as increases in primary productivity closer to the coast (Glud et al., 2013; Honjo et al., 2008). On the shelf, elevated chlorophyll α and β concentrations and depleted major nutrients are indicative of higher primary production (Moss et al., 2016). The oxic-anoxic

interface was closer to the sediment surface in the northern Gulf in comparison to the southern Gulf at similar water depths, likely due to a lower sedimentary carbon deposition regulated by riverine inputs (Figure 3 C). The northern Gulf of Mexico is greatly influenced by the Mississippi River, the fifth largest river by volume in the world and the primary source (~90%) of both sediment and freshwater to the entire Gulf (Dunn, 1996; Goñi et al., 1997; Meybeck and Ragu, 2012). This region also has one of the largest and best studied anthropogenically induced "dead zones" in the world, directly related to the high nutrient loadings linked to large phytoplankton blooms and massive deposition events to the northern Gulf of Mexico (Bianchi et al., 2010; Turner et al., 2008). Furthermore, in late 2010 following the Deepwater Horizon oil spill, a massive offshore (>300m – 1500m sites) sedimentation event was observed where oily marine snow was deposited at 4-5 fold higher rates than typically observed (Brooks et al., 2015; Passow et al., 2012). Relative to the northern Gulf of Mexico, the southern Gulf of Mexico generally receives lower riverine inputs (~10% of the northern Gulf input) (Meybeck and Ragu, 2012). However, our southern Gulf samples were collected in close proximity to the second largest riverine system in the Gulf of Mexico, the Grijalva-Usumacinta system as well as the smaller Coatzacoalcos and Papaloapan systems (Day et al., 1982; Soto et al., 2014).

Inter-basin microbial biogeography

Significant differences in microbial community composition ($p < 2e-16$, var. explained = 4%) were observed between the northern and southern Gulf of Mexico, although these differences were small relative to the impacts of overlying water column depth and depth within the sediments (Figure 2). These modest differences can potentially be attributed to differences in oxygen penetration depths (Figure 3). When comparing the microbial community composition of sediments from similar water depths in each region, community composition in the northern Gulf of Mexico more closely resembled shallower sites in the southern Gulf of Mexico (Figure 4). This effect became more apparent when selecting sediments collected from 1000-1300 m water depths (Figure 4 C). Both southern Gulf microbial

community core profiles cluster below (in the direction of increasing water depth profiles) similar northern Gulf profiles. This is clearly reflected in the deeper depths of oxygen penetration in the southern Gulf of Mexico and the maintenance of higher oxygen concentrations through the core profile (Figure 4 D). The microbial community profile from the deepest site (DSH09, 2293 m) from the northern Gulf of Mexico, with the deepest oxygen penetration depths, was similar to microbial communities obtained from 1300 m in the southern Gulf of Mexico (Figure 4 A, B). Overall, these data are consistent with greater organic matter deposition in the northern Gulf of Mexico due to the greater carbon and nutrient contributions of the Mississippi River relative to the Grijalva- Usumacinta system (Dunn, 1996). It was remarkable how similar the sedimentary microbial communities in these geographically distinct regions with such different quantities of riverine inputs were, following a simple water depth adjustment.

The core seafloor microbiome across the Gulf of Mexico

Shifts in community composition with changes in water depth and sediment depth were apparent at the class level. In the deep Gulf of Mexico (>1000 m), surficial sediments (0-1 cm) were dominated by Gammaproteobacteria (25.6% ± 4.0, n = 126) followed by Marine Group I Thaumarchaeota (21.0% ± 0.08), Deltaproteobacteria (13.1% ± 3.4), and Alphaproteobacteria (10.2% ± 3.1) (Figure 5). The second most abundant class we detected, the MG-I Thaumarchaeota, were missed in previous studies due to recent advances in sequencing technology and primer design (Bienhold et al., 2016; Zinger et al., 2011). Commonly detected, but at low relative abundances in surficial sediments, were the classes Gemmatimonadetes (2.6% ± 0.63), Phycisphaerae (2.6% ± 0.84), Acidobacteria (2.0% ± 3.9), and Planctomycetacia (2.1% ± 1.5). Microbial communities were strongly stratified with sediment depth. Progressing downcore, class level shifts were apparent with decreasing Gammaproteobacteria (11.3% ± 5.3 at 8-10 cm), Thaumarchaeota (4.5% ± 3.7), and Alphaproteobacteria (6.6% ± 3.5) relative

abundances and corresponding increases in Deltaproteobacteria (20.1% ± 5.5), Planctomycetacia (7.2% ± 3.7), and Phycisphaerae (10.1% ± 6.1) groups (Figure 5).

In general, the same progression in class level distributions occurred in each core profile. However, the inflection points in these class distributions shifted down the sediment column with increasing water depth (Figures 2, 5). This was clearly observed by comparing the relative abundance of the aerobic Marine Group I (Thaumarchaeota) in sites DSH08 (1110 m), DSH10 (1500 m) and DSH09 (2293 m) (Figure 5). In addition, the transition between MG-I Thaumarchaeota and Phycisphaerae typically coincided with the oxygen gradient and the oxic-anoxic interface, which increased in sediment column with increasing water depth. This oxic-anoxic transition zone also coincided with highest relative abundances of Planctomycetacia.

In deep ocean oligotrophic sediment studies conducted across the globe, Gammaproteobacteria and Deltaproteobacteria also dominate surficial communities (Bienhold et al., 2012, 2016; Learman et al., 2016; Schauer et al., 2010; Yang et al., 2016; Zinger et al., 2011). Following these two classes, Alphaproteobacteria, Actinobacteria, Planctomycetacia, and Acidobacteria are often dominant and well represented. Less abundant groups include Flavobacteria, Betaproteobacteria, Gemmatimonadetes, and Chloroflexi. The majority of studies have focused on bacterial groups, however, when universal primers are used, Marine Group I Thaumarchaea are shown to be very abundant (Learman et al., 2016). Although previous work in the Gulf of Mexico did not incorporate Archaea, our results are corroborated by clone libraries of bacterial communities constructed from surficial sediments that were sampled prior to oil impacts which showed similar dominant classes (Yang et al., 2016). We observed much lower relative abundances of Bacteriodetes and Betaproteobacteria across all the surficial sediments, potentially due to the different primers used or related to the sequencing size differences between these studies.

Microbial population specific distributions across the Gulf of Mexico

These biogeographic patterns were evident at higher taxonomic resolution, and the most abundant and representative OTU for 6 dominant classes were selected for further study (Figures 6, S3, S4, S7-11). An overview for the distributions of dominant OTU across the Gulf of Mexico are presented here while the specifics are available in the Supplementary Information. The four most abundant Marine Group I OTU (Thaumarchaeota) were all phylogenetically related to *Nitrosopumilus maritimus* (3 OTU >97% sequence identity and 1 OTU at 93% identity). Specific OTU abundances were related to the overlying water depth, showing population level niche differences. Niche differences for these four abundant *Nitrosopumilus*-like populations were not driven by depth within the sediment column as was found by Durbin and Teske (2010).

Marine Group I Thaumarchaeota are well known members of oxic and suboxic sediments (Danovaro et al., 2016; Durbin and Teske, 2010, 2011, 2012; Gillan and Danis, 2007; Learman et al., 2016; Pester et al., 2011). All characterized isolates similar to these 4 OTU are aerobic autotrophs capable of oxidizing ammonia at very low concentrations. Increases in their relative abundances have been shown to be concurrent with decreases in the quality and quantity of organic carbon within sediments (Learman et al., 2016). It has also been shown that in deep, low energy pelagic environments a large portion of *Nitrosopumilus*-like populations were capable of using urea as a carbon and ammonia source to fuel autotrophic nitrification (Alonso-Saez et al., 2012). Conversely, in a global survey of bacterial and archaeal abundances in surficial oxygenated deep sea sediments, MG-I Thaumarchaeal absolute abundances were primarily related to bottom water temperatures and not to quality and quantity of carbon sources (Danovaro et al., 2016). In the Gulf of Mexico, however, there were high relative abundances of these groups at both shallow, warm water, sites as well as deep permanently cold sites and there was not clear evidence that this group as a whole was driven by water temperature instead of quality and quantity of carbon in the distribution of dominant OTUs.

Unlike the Thaumarchaea, the dominant Gammaproteobacteria detected in the deep GoM sediments were distinct from those in the shallower sediments. While there were still water depth and sediment depth dependent shifts in these top 4 Gammaproteobacterial OTU, there was a greater distinction between shallow and deep-water sites than was observed for the MG-I. Three of the dominant OTU belonged to the JTB255 marine benthic group, currently within the Xanthomonadales. The JTB225 marine benthic group was previously shown to be abundant and cosmopolitan at the seafloor, and was previously detected in abundance in Gulf of Mexico surficial sediments (Bienhold et al., 2016; Yang et al., 2016; Zinger et al., 2011). A recent genomic study suggested this group is affiliated with the *Woeseiaceae* family and members within this group are metabolically diverse, capable of aerobic or anaerobic respiration as well as chemolithoautotrophy using reduced sulfur-compounds or hydrogen (Du et al., 2016; Mußmann et al., 2017). These three populations show clear niche partitioning likely related to electron acceptor differences (Figure 6, 7).

All dominant OTU affiliated with the Planctomycetacia were closely related to *Candidatus Scalindua* spp (99-100%). These dominant OTU are likely anaerobic ammonium oxidizing bacteria (anammox), which are chemoautotrophic, ubiquitously found in anoxic systems, and couple the oxidation of ammonium with nitrite as the electron acceptor (Oshiki et al., 2016; Penton et al., 2006). Members from this genus tend to be psychrophilic, and are inhibited by low sulfide concentrations (4-100 µM), which potentially explains their depth limits in our study, although it may also be due to ammonium or nitrite availability (Oshiki et al., 2016). The Planctomycetes were nearly absent in the shallowest sites, and absent in surficial sediments, in agreement with previous studies that indicate ca. *Scalindua* is more abundant in oligotrophic systems (Canion et al., 2014; Learman et al., 2016). All dominant Planctomycetes OTU tended to reach a maximal relative abundance at mid-sediment depths below the oxic-anoxic interface with an overlapping distribution to the *Nitrosopumulis*-like OTUs (Figures 6, S4, S10). Ammonium oxidizing archaea may provide a nitrite source *Scalindua*, and the mostly

exclusive zonation patterns observed may limit direct competition for ammonium (Lipsewers et al., 2014).

While speculative, there was evidence of inorganic nitrogen species being important electron acceptors and donors in deep Gulf of Mexico sediments, particularly for fueling autotrophic metabolisms. In addition to the putative *Nitrosopumilus-like* aerobic ammonium oxidizers and *Scalindua-like* anaerobic ammonium oxidizers, we observed putative bacterial ammonium oxidizers (*Nitrosococcus-like*, and an unclassified Nitrosomonadaceae OTU) that were not present in shallow sediments and increase in relative abundance with increasing water depth (Figure S11). While the uncharacterized Nitrosomonadaceae OTU (denovo23) co-occured with *Nitrosopumilus*-like OTU, the *Nitrosococcus*-like OTU populations increased in relative abundance below the *Nitrosopumilus-like* populations. Abundant putative nitrite oxidizers affiliated with *Nitrospira* and *Nitropina* groups were detected that were also only present in deep water sediments below the aerobic-anaerobic transition zone, similar to the *Scalindua*-like population distributions.

Construction of a predictive microbial community composition model throughout the Gulf of Mexico

After visualizing the consistent biogeographic patterns portrayed in Figure 2, it was recognized that a regression-based model may be able to estimate the relative abundances of microbial groups across these regions. We employed the machine learning method, random forest regression, that has previously been successful in population distribution models (Bradter et al., 2013; Cutler et al., 2007; Hannigan et al., 2018; Lima-Mendez et al., 2015; Oppel and Huettmann, 2010; Smith et al., 2010; Wei et al., 2010). As stated in the methods section, the model was constructed using the variables water depth, sediment depth, latitude, and longitude as proxies for carbon input, redox state, and geographic

variations due to riverine inputs and differing geochronology across the Gulf of Mexico (Brooks et al., 2015).

This analysis was run for every abundantly detected OTU (9588 OTU). All OTU with greater than 50% variance explained by the full model were retained (4129 OTU). Overall, and as expected, the models fit the observed data well ($R^2 = 0.88$). The residuals were weakly correlated with OTU abundance as the model tends to underestimate the very abundant OTU. The model was further tested by calculating the root mean squared error (RMSE) compared to 1000 bootstrapped randomly permuted RMSE and was found to be highly significant ($p < 2e^{-16}$, Z-test).

To verify the utility of the model, model predictions were compared with observed counts for the dominant OTU (Figure S5, S6-S10). However, it is anticipated that the model can be used as a tool to assess future impacted sediments that may be poorly characterized. In order to test this, we hindcasted the microbial community composition at previously impacted sediments by determining how far the observed microbial community deviated from the model's prediction. This was accomplished by calculating the total root mean squared error (RMSE) for each sample using observed OTU abundance and predicted OTU abundances, including the Mason et al. (2014) surficial sample set (Figure 7 C). The most impacted sites had higher RMSE values than the below EPA limits sites, although both were skewed to the right of the model predictions for our data (Figure 7 C). Therefore, while controlling for experimental based variation, this model can still be effective at determining impacts.

Impacts of the DWH and internal testing of the predictive model

The earliest samples included in this study were collected in August 2012, over two years after the Deepwater Horizon well blowout. No evidence of a large oil spill signature was detected in sediments sampled from areas known to be impacted by sedimentary oil deposition related to the DWH blowout (Chanton et al., 2012; Romero et al., 2015, 2017; Valentine et al., 2014). At the broad

biogeographic scale visualized in Figure 2, the lack of an “oil-impacted cluster” indicated that any potential impact on community composition was small relative to the strong environmental gradients influenced primarily by water depth and sediment depth. To compare our data explicitly with sediments that were shown to be strongly impacted by oil, a meta-analysis based approach was employed. We observe that all surficial sediment samples from this study, including those collected close to the DWH wellhead, clustered adjacent to samples classified as below EPA limits and were distinct from those classified as most impacted by oiling (Mason et al., 2014) (Figure 7 B). Furthermore, the similarities between the below EPA limit samples and our dataset indicated a temporally stable and Gulf-wide conserved microbial community relative to a strong pulse disturbance. Finally, previously impacted sediments that were collected and analyzed during this study did not deviate from model predictions, unlike those collected in 2010 (Figure7 C).

These results are corroborated by a previous study that detailed a microbial community succession event in oil impacted surficial sediments immediately following the oiled marine snow sedimentation event, through the generation of full length SSU rRNA gene clone libraries (Yang et al., 2016). The authors suggested that by November 2010 the microbial community composition once again resembled pre-impacted surficial microbial communities (Yang et al., 2016). Our first samples were collected in August 2012 and much of our dataset comes from 2013 and 2014 for the NGoM, years following a potential return to baseline condition. We also do not observe a disturbance signature in cores collected at the Ixtoc-I wellhead or within the PEMEX (Mexican state-owned petroleum company) restricted area that has extensive oil extraction pressures. The lack of an Ixtoc-I signature in the microbial community is expected concerning the resilience observed in the sedimentary communities in the northern Gulf of Mexico, and the lack of a PEMEX oil signature indicates we did not sample sediments that were recently exposed to oil.

Conclusions

The primary objectives of this study were to define the core microbiome across the Gulf of Mexico seafloor, characterize biogeographic patterns in microbial populations in Gulf sediments, and use these results to constrain the impacts of petroleum hydrocarbons from the Deepwater Horizon oil spill to sedimentary microbial communities. Using the largest marine sediment microbial community composition dataset generated to date, we show that the distribution of microbial communities is broadly consistent across the entire region in terms of the populations (OTU) detected and their relative abundances. We further demonstrate that the distribution of seafloor microbial communities is well approximated by the overlying water depth and depth within the sediment column, which together explain 38% of our observed variation, regardless of geographic sampling region. The small differences between the regions are potentially related to the amount of organic carbon reaching the seafloor, with the southern Gulf showing greater depths of oxygen penetration as well as microbial communities that had greater similarity to deeper counterparts in the northern Gulf. Microbial community composition across the Gulf revealed similar taxonomic groups as observed in other marine sediment biogeography studies across the globe and is dominated primarily by groups that are poorly characterized with unknown ecological functions within the sediment.

The lack of a strongly perturbed community in our dataset, along with the incorporation of previous studies demonstrating an oil related perturbation signal, suggests that impacted sediment communities had returned or are returning to baseline conditions. Finally, the biogeographical distributions of seafloor microbial populations were used to generate a predictive model for over 4000 dominant populations that relies on easy-to-obtain geospatial variables. With this model, it is possible to predict community compositions across the Gulf of Mexico, and use deviations from such a prediction as a tool to assess future impacted sediments that may be poorly characterized along with natural ecosystems that are conspicuously different from oligotrophic ecosystems such as cold seeps.

Acknowledgements

This work was made possible in part by a grant from The Gulf of Mexico Research Initiative to the C-IMAGE II consortium and in part by a National Science Foundation Graduate Research Fellowship (WAO) under grant no. 2013172310. The authors would like to thank the ship crews on the R/V Weatherbird III and the R/V Justo Sierra that made the sample collection possible. Data are publicly available through the Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC) at https://data.gulfresearchinitiative.org (doi: 10.7266/N70G3HN3). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Aller, J. Y. 1989. Quantifying sediment disturbance by bottom currents and its effect on benthic communities in a deep-sea western boundary zone. *Deep Sea Research Part A, Oceanographic Research Papers*, *36*(6):, 901–934. doi: 10.1016/0198-0149(89)90035-6

Alonso-Saez, L., a. S. Waller, D. R. Mende, K. Bakker, H. Farnelid, P. L. Yager, … S. Bertilsson. 2012. Role for urea in nitrification by polar marine Archaea. *Proceedings of the National Academy of Sciences*, *109*(44):, 17989–17994. doi: 10.1073/pnas.1201914109

Arndt, S., B. B. Jørgensen, D. E. LaRowe, J. J. Middelburg, R. D. Pancost, and P. Regnier. 2013. Quantifying the degradation of organic matter in marine sediments: A review and synthesis. *Earth-Science Reviews*, *123*, 53–86. doi: 10.1016/j.earscirev.2013.02.008

Atlas, R. M. 1981. Fate of oil from two major oil spills: Role of microbial degradation in removing oil from the Amoco Cadiz and IXTOC I spills. *Environment International*, *5*(1):, 33–38.

Bianchi, T. S., S. F. DiMarco, J. H. Cowan, R. D. Hetland, P. Chapman, J. W. Day, and M. A. Allison. 2010. The science of hypoxia in the Northern Gulf of Mexico: a review. *Science of the Total Environment*, *408*(7):, 1471–1484.

Biddle, J. F., J. R. White, A. P. Teske, and C. H. House. 2011. Metagenomics of the subsurface Brazos-Trinity Basin (IODP site 1320): comparison with other sediment and pyrosequenced metagenomes. *The ISME journal*, *5*(6):, 1038–47. doi: 10.1038/ismej.2010.199

Bienhold, C., A. Boetius, and A. Ramette. 2012. The energy–diversity relationship of complex bacterial communities in Arctic deep-sea sediments. *The ISME Journal*, *6*(4):, 724–732. doi: 10.1038/ismej.2011.140

Bienhold, C., L. Zinger, A. Boetius, and A. Ramette. 2016. Diversity and biogeography of bathyal and abyssal seafloor bacteria. *PLoS ONE*, *11*(1):, 1–20. doi: 10.1371/journal.pone.0148016

Billett, D. S. M., R. S. Lampitt, A. L. Rice, and R. F. C. Mantoura. 1983. Seasonal sedimentation of phytoplankton to the deep-sea benthos. *Nature*. doi: 10.1038/302520a0

Bradter, U., W. E. Kunin, J. D. Altringham, T. J. Thom, and T. G. Benton. 2013. Identifying appropriate spatial scales of predictors in species distribution models with the random forest algorithm. *Methods in Ecology and Evolution*, *4*(2):, 167–174. doi: 10.1111/j.2041-210x.2012.00253.x

Brooks, G. R., R. A. Larson, P. T. Schwing, I. Romero, C. Moore, G. J. Reichart, … D. Hollander. 2015. Sedimentation pulse in the NE Gulf of Mexico following the 2010 DWH blowout. *PLoS ONE*, *10*(7):, 1–24. doi: 10.1371/journal.pone.0132341

Burdige, D. J. 2006. *Geochemistry of marine sediments*.

Canion, A., W. A. Overholt, J. E. Kostka, M. Huettel, G. Lavik, and M. M. M. Kuypers. 2014. Temperature response of denitrification and anaerobic ammonium oxidation rates and microbial community structure in Arctic fjord sediments. *Environmental Microbiology*, *16*(10):, 3331–3344. doi: 10.1111/1462-2920.12593

Caporaso, J. G., K. Bittinger, F. D. Bushman, T. Z. DeSantis, G. L. Andersen, and R. Knight. 2010. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics (Oxford, England)*, *26*(2):, 266–7. doi: 10.1093/bioinformatics/btp636

Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, … R. Knight. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, *7*(5):, 335–336. doi: 10.1038/nmeth.f.303

Caporaso, J. G., C. L. Lauber, W. a Walters, D. Berg-Lyons, J. Huntley, N. Fierer, … R. Knight. 2012. Ultra-

high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME journal*, *6*(8):, 1621–4. doi: 10.1038/ismej.2012.8

Caporaso, J., J. Kuczynski, and J. Stombaugh. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, *7*(5):, 335–336. doi: 10.1038/nmeth0510-335

Carrigg, C., O. Rice, S. Kavanagh, G. Collins, and V. O'Flaherty. 2007. DNA extraction method affects microbial community profiles from soils and sediment. *Applied Microbiology and Biotechnology*, *77*(4):, 955–964. doi: 10.1007/s00253-007-1219-y

Chanton, J. P., J. Cherrier, R. M. Wilson, J. Sarkodee-Adoo, S. Bosman, A. Mickle, and W. M. Graham. 2012. Radiocarbon evidence that carbon from the Deepwater Horizon spill entered the planktonic food web of the Gulf of Mexico. *Environmental Research Letters*, *7*(4):, 045303. doi: 10.1088/1748-9326/7/4/045303

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random forests for classification in ecology. *Ecology*, *88*(11):, 2783–2792.

Danovaro, R., M. Molari, C. Corinaldesi, and A. Dell'Anno. 2016. Macroecological drivers of archaea and bacteria in benthic deep-sea ecosystems. *Science Advances*, *2*(4):, 1–12. doi: 10.1126/sciadv.1500961

Day, J. W., R. H. Day, M. T. Barreiro, F. Ley-Lou, and C. J. Madden. 1982. Primary production in the Laguna de Terminos, a tropical estuary in the southern Gulf of Mexico. *Oceanologica Acta, Special issue*.

Divins, D. L., and D. Metzger. 2008. NGDC coastal relief model. *National Geophysical Data Center, National Oceanic and Atmospheric Administration, US Department of Commerce, www. ngdc. noaa. gov/mgg/(last accessed December 2007)*.

Du, Z. J., Z. J. Wang, J. X. Zhao, and G. J. Chen. 2016. Woeseia oceani gen. Nov., sp. nov., a chemoheterotrophic member of the order Chromatiales, and proposal of Woeseiaceae fam. nov. *International Journal of Systematic and Evolutionary Microbiology*, *66*(1):, 107–112. doi: 10.1099/ijsem.0.000683

Dunn, D. E. 1996. *Trends in nutrient inflows to the Gulf of Mexico from streams draining the conterminous United States, 1972-93*.

Durbin, A. M., and A. Teske. 2010. Sediment-associated microdiversity within the Marine Group I Crenarchaeota. *Environmental microbiology reports*, *2*(5):, 693–703. doi: 10.1111/j.1758-2229.2010.00163.x

Durbin, A. M., and A. Teske. 2011. Microbial diversity and stratification of South Pacific abyssal marine sediments. *Environmental microbiology*, *13*(12):, 3219–34. doi: 10.1111/j.1462-2920.2011.02544.x

Durbin, A. M., and A. Teske. 2012. Archaea in organic-lean and organic-rich marine subsurface sediments: an environmental gradient reflected in distinct phylogenetic lineages. *Frontiers in microbiology*, *3*(May):, 168. doi: 10.3389/fmicb.2012.00168

Ebbe, B., D. S. M. Billett, A. Brandt, K. Ellingsen, A. Glover, S. Keller, … others. 2010. Diversity of abyssal marine life. *Life in the World's Oceans: Diversity, Distribution, and Abundance, edited by: McIntyre, A*, 139–160.

Gillan, D. C., and B. Danis. 2007. The archaebacterial communities in Antarctic bathypelagic sediments. *Deep Sea Research Part II: Topical Studies in Oceanography*, *54*(16):, 1682–1690.

Glud, R. N., F. Wenzhöfer, M. Middelboe, K. Oguri, R. Turnewitsch, D. E. Canfield, and H. Kitazato. 2013. High rates of microbial carbon turnover in sediments in the deepest oceanic trench on Earth. *Nature Geoscience*, *6*(4):, 284–288. doi: 10.1038/ngeo1773

Goñi, M. A., K. C. Ruttenberg, and T. I. Eglinton. 1997. Sources and contribution of terrigenous organic carbon to surface sediments in the Gulf of mexico. *Letters To Nature*, *389*(SEPTEMBER):, 275–279.

Green, S. J., R. Venkatramanan, and A. Naqib. 2015. Deconstructing the polymerase chain reaction: Understanding and correcting bias associated with primer degeneracies and primer-template mismatches. *PLoS ONE*, *10*(5):, 1–21. doi: 10.1371/journal.pone.0128122

Hannigan, G. D., M. B. Duhaime, D. Koutra, and P. D. Schloss. 2018. Biogeography and environmental conditions shape bacteriophage-bacteria networks across the human microbiome. *PLOS Computational Biology*, *14*(4):, e1006099. Retrieved from https://doi.org/10.1371/journal.pcbi.1006099

Hollander, D., G. Brooks, R. Larson, I. Romero, and P. Schwing. 2013. Testing the mechanisms of sedimentary oil deposition in the deep-sea. *Proceedings of the Gulf of Mexico Oil Spill & Ecosystem Sciences Conference*.

Honjo, S., S. J. Manganini, R. A. Krishfield, and R. Francois. 2008. Particulate organic carbon fluxes to the ocean interior and factors controlling the biological pump: A synthesis of global sediment trap programs since 1983. *Progress in Oceanography*, *76*(3):, 217–285.

Jacob, M., T. Soltwedel, A. Boetius, and A. Ramette. 2013. Biogeography of Deep-Sea Benthic Bacteria at Regional Scale (LTER HAUSGARTEN, Fram Strait, Arctic). *PLoS ONE*, *8*(9):, e72779. doi: 10.1371/journal.pone.0072779

Jernelöv, A., and O. Lindén. 1981. Ixtoc I: a case study of the world's largest oil spill. *Ambio*, 299–306.

Jørgensen, B. B., and A. Boetius. 2007. Feast and famine — microbial life in the deep-sea bed. *Nature Reviews Microbiology*, *5*(10):, 770–781. doi: 10.1038/nrmicro1745

Joye, S. B. 2015. Deepwater Horizon, 5 years on. *Science*, *349*(6248):, 592–593.

Joye, S. B., I. R. MacDonald, I. Leifer, and V. Asper. 2011. Magnitude and oxidation potential of hydrocarbon gases released from the BP oil well blowout. *Nature Geoscience*, *4*(3):, 160–164. doi: 10.1038/ngeo1067

Kimes, N. E., A. V Callaghan, D. F. Aktas, W. L. Smith, J. Sunner, B. Golding, … P. J. Morris. 2013. Metagenomic analysis and metabolite profiling of deep-sea sediments from the Gulf of Mexico following the Deepwater Horizon oil spill. *Frontiers in microbiology*, *4*(March):, 50. doi: 10.3389/fmicb.2013.00050

King, G. M. M., J. E. Kostka, T. C. C. Hazen, and P. A. A. Sobecky. 2014. Microbial Responses to the Deepwater Horizon oil spill: From Coastal Wetlands to the Deep Sea. *Annual review of marine science*.

Kukert, H., and C. R. Smith. 1992. Disturbance, colonization and succession in a deep-sea sediment community: artificial-mound experiments. *Deep Sea Research Part A, Oceanographic Research Papers*, *39*(7–8):, 1349–1371. doi: 10.1016/0198-0149(92)90073-3

Learman, D. R., M. W. Henson, J. C. Thrash, B. Temperton, P. M. Brannock, S. R. Santos, … K. M. Halanych. 2016. Biogeochemical and microbial variation across 5500 km of Antarctic surface sediment implicates organic matter as a driver of benthic community structure. *Frontiers in Microbiology*, *7*(MAR):, 1–11. doi: 10.3389/fmicb.2016.00284

Liaw, A., and M. Wiener. 2002. Classification and Regression by randomForest. *R News*, *2*(3):, 18–22. Retrieved from http://cran.r-project.org/doc/Rnews/

Lima-Mendez, G., K. Faust, N. Henry, J. Decelle, S. Colin, F. Carcillo, … L. Bittner. 2015. Determinants of community structure in the grobal plankton interactome. *Science*, *348*(6237):, 1262073. doi: 10.1126/science.1262073

Lipsewers, Y. A., N. J. Bale, E. C. Hopmans, S. Schouten, J. S. Sinninghe Damst??, and L. Villanueva. 2014. Seasonality and depth distribution of the abundance and activity of ammonia oxidizing microorganisms in marine coastal sediments (North Sea). *Frontiers in Microbiology*, *5*(SEP):, 1–12. doi: 10.3389/fmicb.2014.00472

Liu, Z., and J. Liu. 2013. Evaluating bacterial community structures in oil collected from the sea surface and sediment in the northern Gulf of Mexico after the Deepwater Horizon oil spill. *MicrobiologyOpen*. doi: 10.1002/mbo3.89

Lozupone, C. a, and R. Knight. 2007. Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(27):, 11436–11440. doi: 10.1073/pnas.0611525104

Lubchenco, J., M. McNutt, B. Lehr, M. Sogge, M. Miller, S. Hammond, and W. Conner. 2010. *Deepwater Horizon/BP Oil Budget: What happened to the oil? National Oceanic and Atmospheric Administration Report*.

Mahé, F., T. Rognes, C. Quince, C. de Vargas, and M. Dunthorn. 2014. Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ*, *2*, e593. doi: 10.7717/peerj.593

Mason, O. U., N. M. Scott, A. Gonzalez, A. Robbins-Pianka, J. Bælum, J. Kimbrel, … J. K. Jansson. 2014. Metagenomics reveals sediment microbial community response to Deepwater Horizon oil spill. *The ISME Journal*, (October 2013):, 1–12. doi: 10.1038/ismej.2013.254

McMurdie, P. J., and S. Holmes. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*, *8*(4):, e61217. doi: 10.1371/journal.pone.0061217

Meybeck, M., and A. Ragu. 2012. GEMS-GLORI world river discharge database. doi: 10.1594/PANGAEA.804574

Moss, J. A., C. McCurry, P. Schwing, W. H. Jeffrey, I. C. Romero, D. J. Hollander, and R. A. Snyder. 2016. Molecular characterization of benthic foraminifera communities from the Northeastern Gulf of Mexico shelf and slope following the Deepwater Horizon event. *Deep-Sea Research Part I: Oceanographic Research Papers*, *115*, 1–9. doi: 10.1016/j.dsr.2016.04.010

Mußmann, M., P. Pjevac, K. Krüger, and S. Dyksma. 2017. Genomic repertoire of the Woeseiaceae/JTB255, cosmopolitan and abundant core members of microbial communities in marine sediments. *The ISME Journal*, 1–6. doi: 10.1038/ismej.2016.185

Nunoura, T., B. Soffientino, A. Blazejak, J. Kakuta, H. Oida, A. Schippers, and K. Takai. 2009. Subseafloor microbial communities associated with rapid turbidite deposition in the Gulf of Mexico continental slope (IODP Expedition 308). *FEMS microbiology ecology*, *69*(3):, 410–24. doi: 10.1111/j.1574-6941.2009.00718.x

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, … H. Wagner. 2007. The Vegan Package. *Community Ecology Package*, *10*, 631–637.

Oppel, S., and F. Huettmann. 2010. Using a Random Forest Model and Public Data to Predict the Distribution of Prey for Marine Wildlife Management. In S. A. Cushman and F. Huettmann (Eds.), *Spatial Complexity, Informatics, and Wildlife Conservation* (pp. 151–163). doi: 10.1007/978-4-431-87771-4_8

Orcutt, B. N., S. B. Joye, S. Kleindienst, K. Knittel, A. Ramette, A. Reitz, … A. Boetius. 2010. Impact of natural oil and higher hydrocarbons on microbial diversity, distribution, and activity in Gulf of Mexico cold-seep sediments. *Deep Sea Research Part II: Topical Studies in Oceanography*, *57*(21):, 2008–2021.

Oshiki, M., H. Satoh, and S. Okabe. 2016. Ecology and physiology of anaerobic ammonium oxidizing

bacteria. *Environmental Microbiology*, *18*(9):, 2784–2796. doi: 10.1111/1462-2920.13134

Passow, U., K. Ziervogel, V. Asper, and a Diercks. 2012. Marine snow formation in the aftermath of the Deepwater Horizon oil spill in the Gulf of Mexico. *Environmental Research Letters*, *7*(3):, 035301. doi: 10.1088/1748-9326/7/3/035301

Paulson, J. N., O. C. Stine, H. C. Bravo, and M. Pop. 2013. Differential abundance analysis for microbial marker-gene surveys. *Nature methods*, *10*(12):, 1200–2. doi: 10.1038/nmeth.2658

Penton, C. R., A. H. Devol, and J. M. Tiedje. 2006. Molecular evidence for the broad distribution of anaerobic ammonium-oxidizing bacteria in freshwater and marine sediments. *Applied and Environmental Microbiology*, *72*(10):, 6829–6832. doi: 10.1128/AEM.01254-06

Pester, M., C. Schleper, and M. Wagner. 2011. The Thaumarchaeota: An emerging view of their phylogeny and ecophysiology. *Current Opinion in Microbiology*, *14*(3):, 300–306. doi: 10.1016/j.mib.2011.04.007

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

Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, … F. O. Glockner. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1):, D590–D596. doi: 10.1093/nar/gks1219

R Core Team. 2014. R: A Language and Environment for Statistical Computing. Retrieved from http://www.r-project.org/

Rognes, T., T. Flouri, B. Nichols, C. Quince, and F. Mahe. 2016. VSEARCH : a versatile open source tool for metagenomics (# 13057). *PeerJ Preprints*, 1–30.

Romero, I. C., P. T. Schwing, G. R. Brooks, R. A. Larson, D. W. Hastings, G. Ellis, … D. J. Hollander. 2015. Hydrocarbons in deep-sea sediments following the 2010 Deepwater Horizon blowout in the northeast Gulf of Mexico. *PLoS ONE*, *10*(5):, 1–23. doi: 10.1371/journal.pone.0128371

Romero, I. C., G. Toro-Farmer, A. R. Diercks, P. Schwing, F. Muller-Karger, S. Murawski, and D. J. Hollander. 2017. Large-scale deposition of weathered oil in the Gulf of Mexico following a deep-water oil spill. *Environmental Pollution*, *228*, 179–189. doi: 10.1016/j.envpol.2017.05.019

Schauer, R., C. Bienhold, A. Ramette, and J. Harder. 2010. Bacterial diversity and biogeography in deep-sea surface sediments of the South Atlantic Ocean. *The ISME Journal*, *4*(2):, 159–170. doi: 10.1038/ismej.2009.106

Schlitzer, R. 2012. Ocean data view. Retrieved from http://odv.awi.de/

Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, … C. F. Weber. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, *75*(23):, 7537–7541. doi: 10.1128/AEM.01541-09

Schrope, M. 2010. The lost legacy of the last great oil spill.

Schwing, P. T., I. C. Romero, R. A. Larson, B. J. O'Malley, E. E. Fridrik, E. A. Goddard, … J. Mulhollan. 2016. Sediment Core Extrusion Method at Millimeter Resolution Using a Calibrated, Threaded-rod. *Journal of Visualized Experiments : JoVE*. doi: 10.3791/54363

Smith, A., B. Sterba-Boatwright, and J. Mott. 2010. Novel application of a statistical technique, Random Forests, in a bacterial source tracking study. *Water Research*, *44*(14):, 4067–4076. doi: 10.1016/j.watres.2010.05.019

Smith, C. R., and A. R. Baco. 2003. Ecology of whale falls at the deep-sea floor. *Oceanography and*

marine biology, *41*, 311–354.

Soto, L. A., A. V. Botello, S. Licea-DurÃ¡n, M. L. LizÃ¡rraga-Partida, and A. YÃ¡Ã±ez-Arancibia. 2014. The environmental legacy of the Ixtoc-I oil spill in Campeche Sound, southwestern Gulf of Mexico. *Frontiers in Marine Science*, *1*(November):, 1–9. doi: 10.3389/fmars.2014.00057

Teske, A. 2013. Marine deep sediment microbial communities. In E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Eds.), *The Prokaryotes* (pp. 123–133). doi: 10.1007/978-3-642-30123-0

Turner, R. E., N. N. Rabalais, and D. Justic. 2008. Gulf of Mexico hypoxia: aleternate states and a legacy. *Environmental Science and Technology*, *42*(7):, 2323–2327.

Valentine, D. L., G. B. Fisher, S. C. Bagby, R. K. Nelson, C. M. Reddy, S. P. Sylva, and M. a. Woo. 2014. Fallout plume of submerged oil from Deepwater Horizon. *Proceedings of the National Academy of Sciences*, 1–6. doi: 10.1073/pnas.1414873111

Wei, C.-L., G. T. Rowe, E. Escobar-Briones, A. Boetius, T. Soltwedel, M. J. Caley, … B. E. Narayanaswamy. 2010. Global Patterns and Predictions of Seafloor Biomass Using Random Forests. *PLoS ONE*, *5*(12):, e15323. doi: 10.1371/journal.pone.0015323

Wenzhöfer, F., and R. N. Glud. 2002. Benthic carbon mineralization in the Atlantic : a synthesis based on in situ data from the last decade. *Deep-Sea Research I*, *49*, 1255–1279.

Wickham, H. 2009. *ggplot2: elegant graphics for data analysis*. Retrieved from http://had.co.nz/ggplot2/book

Yang, T., K. Speare, L. McKay, B. J. MacGregor, S. B. Joye, and A. Teske. 2016. Distinct Bacterial Communities in Surficial Seafloor Sediments Following the 2010 Deepwater Horizon Blowout. *Frontiers in Microbiology*, *7*(September):, 1–18. doi: 10.3389/fmicb.2016.01384

Zhang, J., K. Kobert, T. Flouri, and A. Stamatakis. 2013. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, 1–7.

Zinger, L., L. a Amaral-Zettler, J. a Fuhrman, M. C. Horner-Devine, S. M. Huse, D. B. M. Welch, … A. Ramette. 2011. Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. *PloS one*, *6*(9):, e24570. doi: 10.1371/journal.pone.0024570

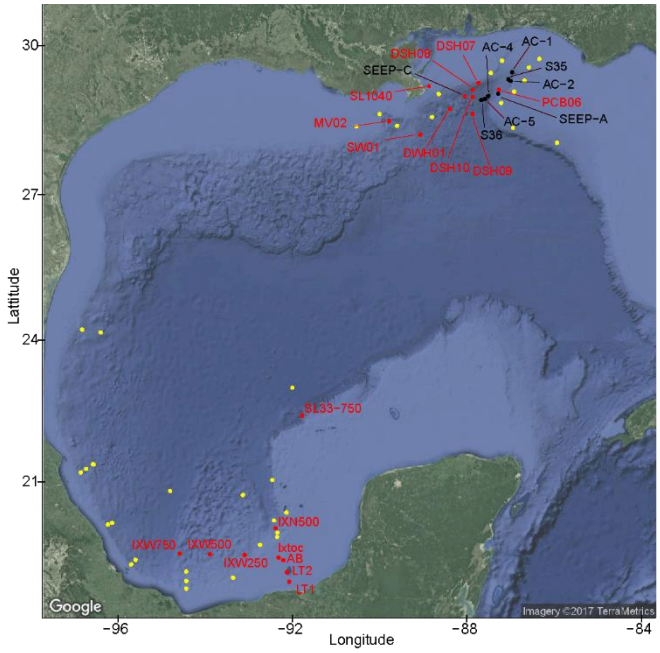

Figure 1 Map of the Gulf of Mexico. All yellow points are sites that were profiled for oxygen concentrations, while sites with black points are those that only have microbial community data. Sites in red are represented by both oxygen and microbial community data. All sites sampled for microbial data are labeled with the site name, which can be seen in Table S1.

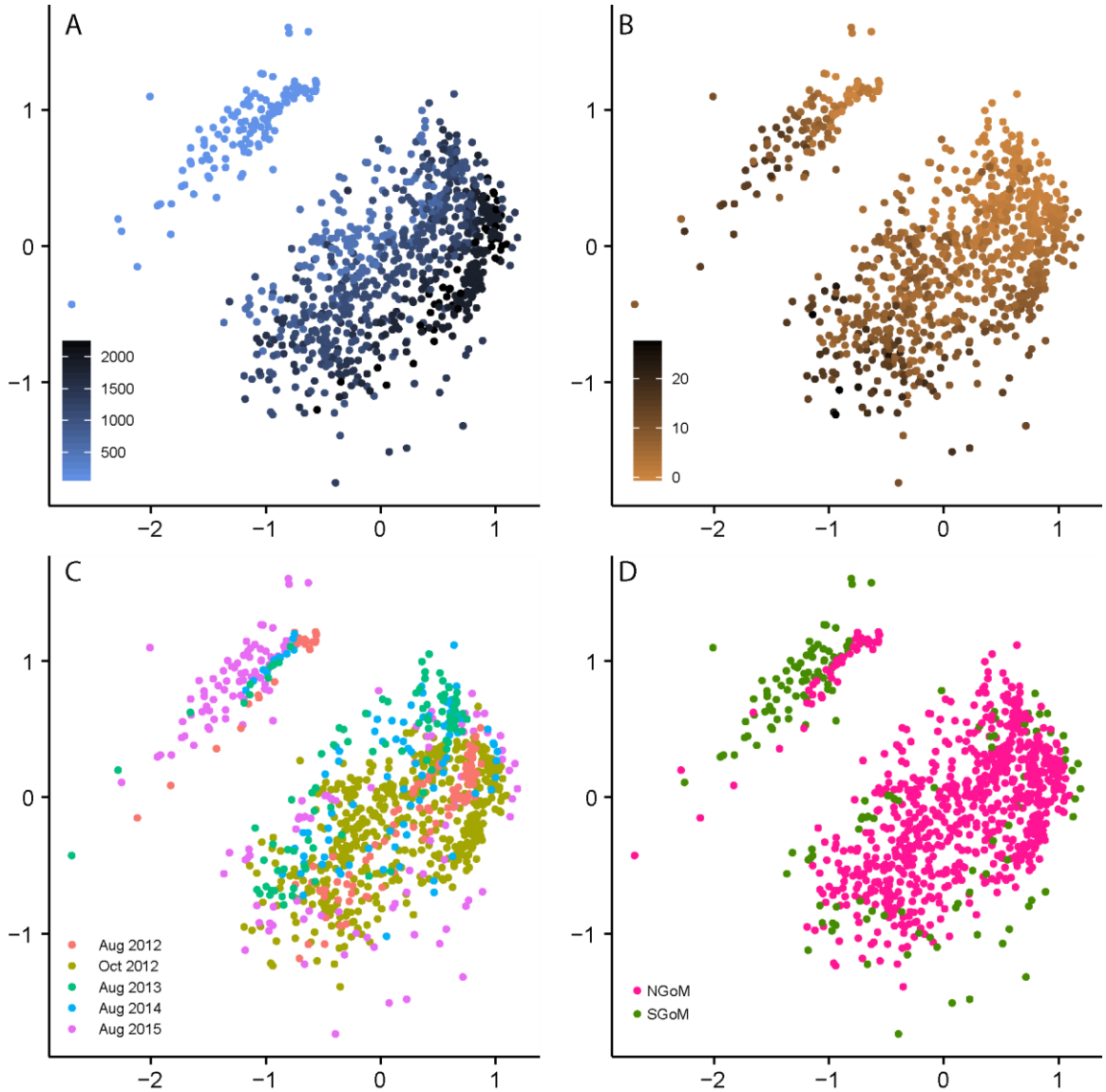

Figure 2 Large-scale patterns in microbial biogeography. All plots are generated from the same NMDS ordination of Bray-Curtis dissimilarity. The stress for each plot is 0.152. (A) Samples are colored based on increasing water depth. The shallow water samples (< 100m) clearly separate from the deeper water samples (> 400m). Note the intermediate water depths were not sampled. (B) Samples are colored darker with increasing sediment depths. Although there are a few outliers, there is a clear and consistent downcore progression from the "top-right" of the NMDS plot towards the bottom left in every core sampled. (C) Samples are colored based on collection date. (D) Cores were collected from the Southern Gulf of Mexico (colored in green) and the Northern Gulf of Mexico (colored in pink)

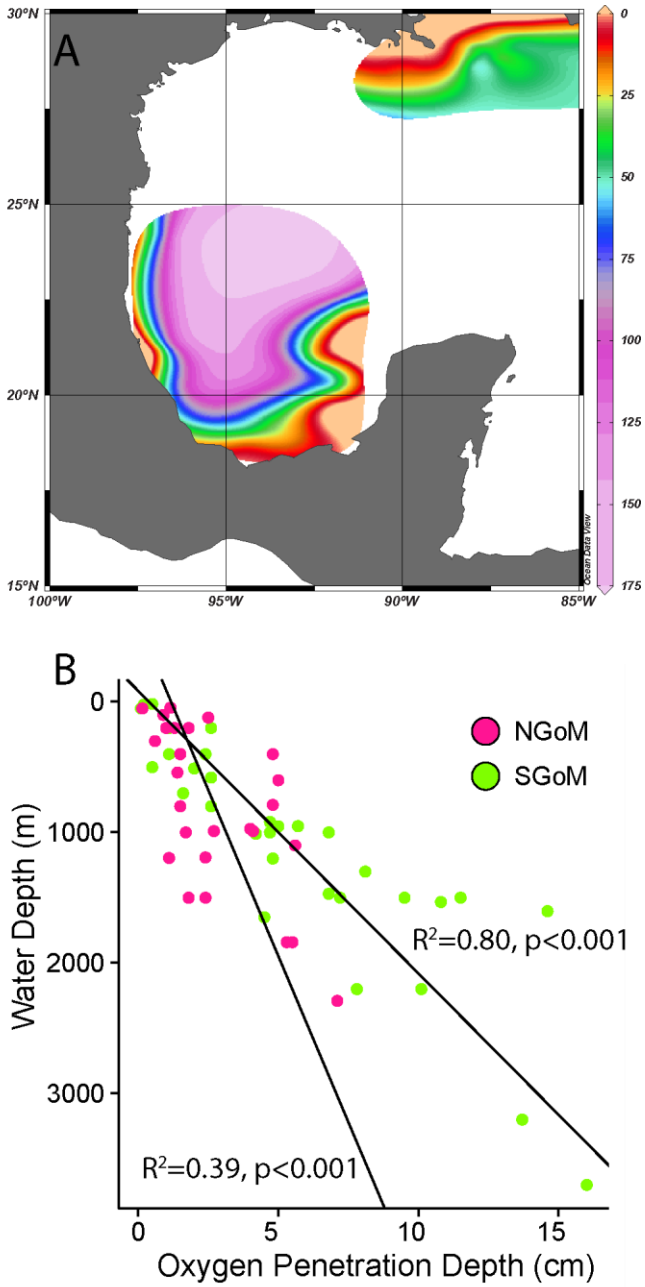

Figure 3 Oyxgen profiles collected throughout the Gulf of Mexico. (A) Interpolated plot showing predicted oxygen penetration depths across our study site. (B) Oxygen penetration plotted against water column depth.

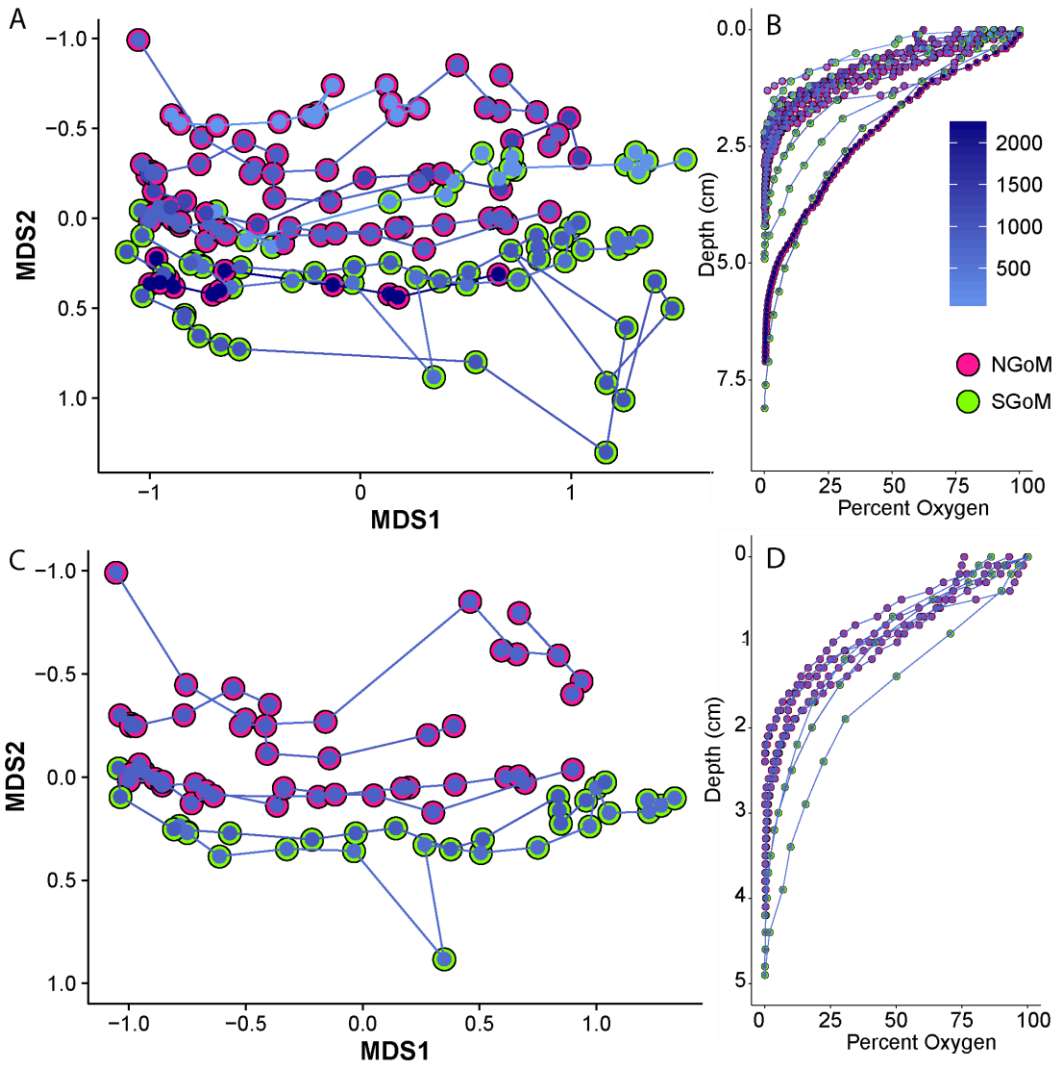

Figure 4 Subtle differences in community progression between the Northern and Southern Gulf of Mexico. The SGoM sites sampled have deeper depths of oxygen penetration and correspond to a community composition more similar to shallower sites in the NGoM. (A) NMDS of Bray-Curtis dissimilarity values for cores collected in 2014 (NGoM, pink) or 2015 (SGoM, green) from 500m – 2300m water depths. Down core progression is indicated by the connecting lines from surficial sediments on the left towards deeper samples on the right. (B) Corresponding oxygen profiles to all cores presented in A. (C, D) Same as plots A & B except only displaying sites collected between 1000-1300m.

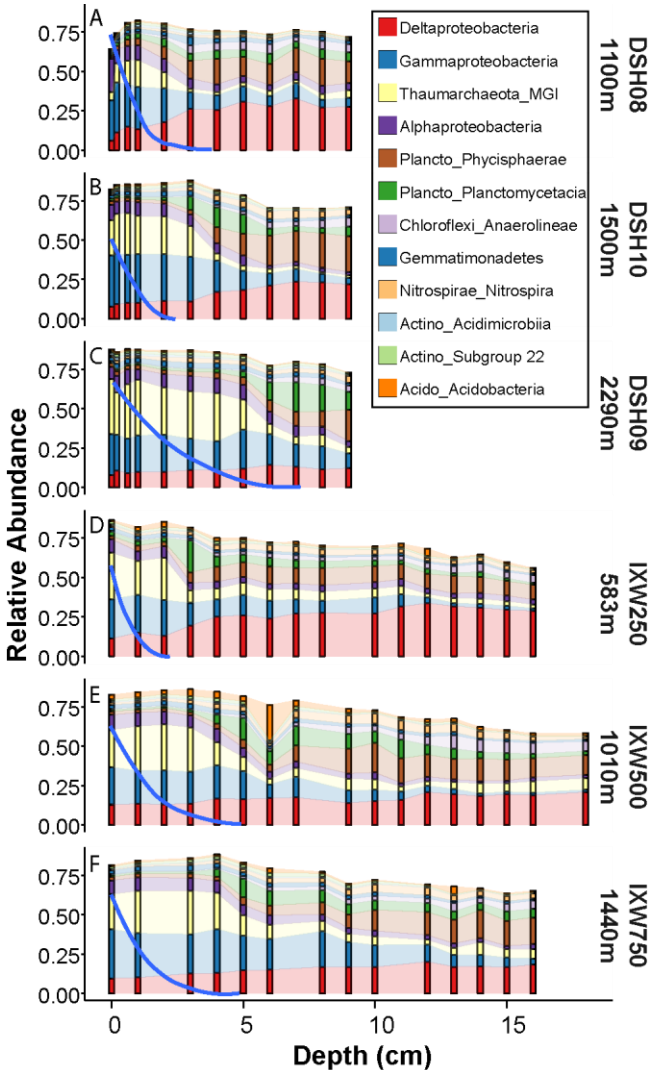

Figure 5 Core profiles for the 12 most abundant classes detected from deep water transects. (A-C) Representative profiles from the Northern Gulf of Mexico from the DSH core transect going offshore towards increasing water depth. (D-F) Representative profiles from the Southern Gulf of Mexico along a westward transect from the Ixtoc wellhead, also organized from shallowest to deepest site in the transect. In all plots, blue lines represent the percent oxygen concentration relative to the initial overlying water oxygen concentration.

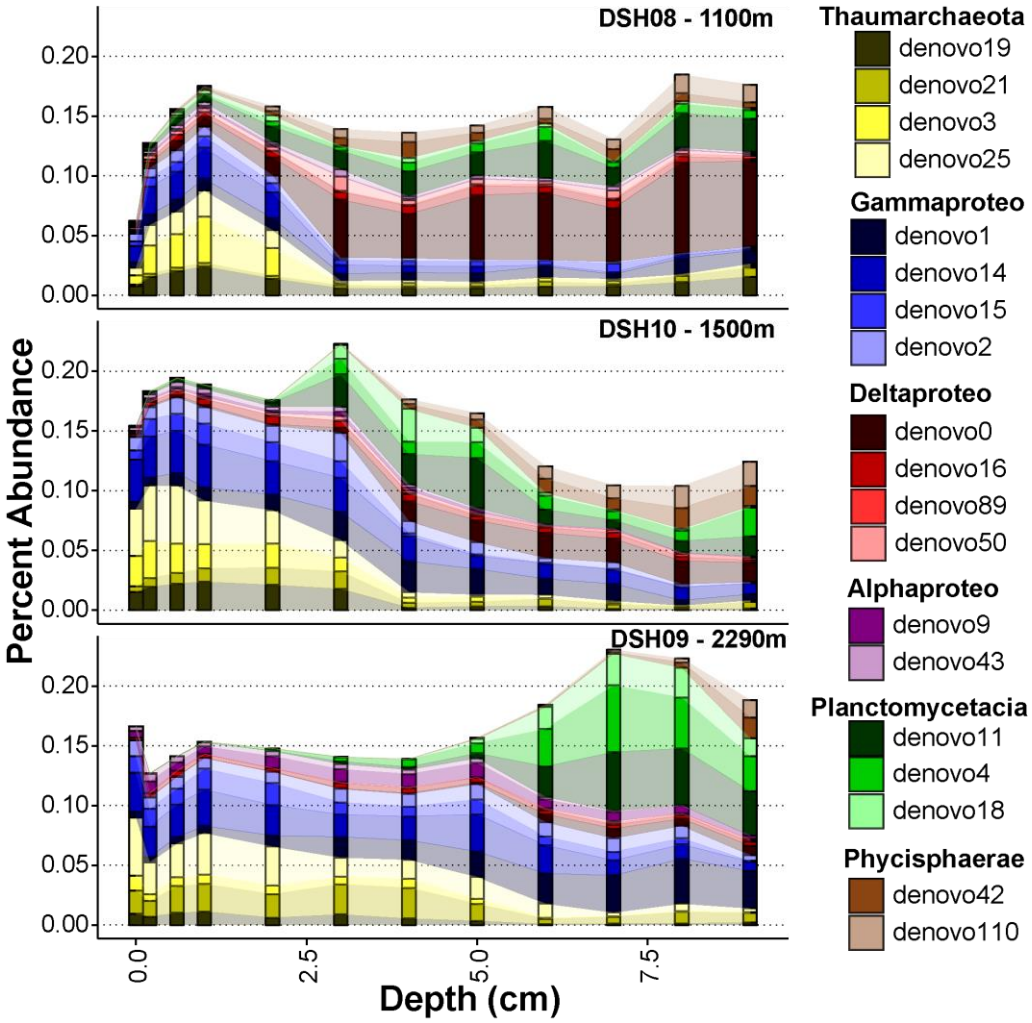

Figure 6 Core profiles representing deep water sites in the Northern Gulf of Mexico showing population (OTU) level distributions with increasing sediment and water depths. Core profiles are taken from the DSH transect and are arranged in order of increasing water depth. See Figure S4 for similar trends within the Southern Gulf of Mexico.

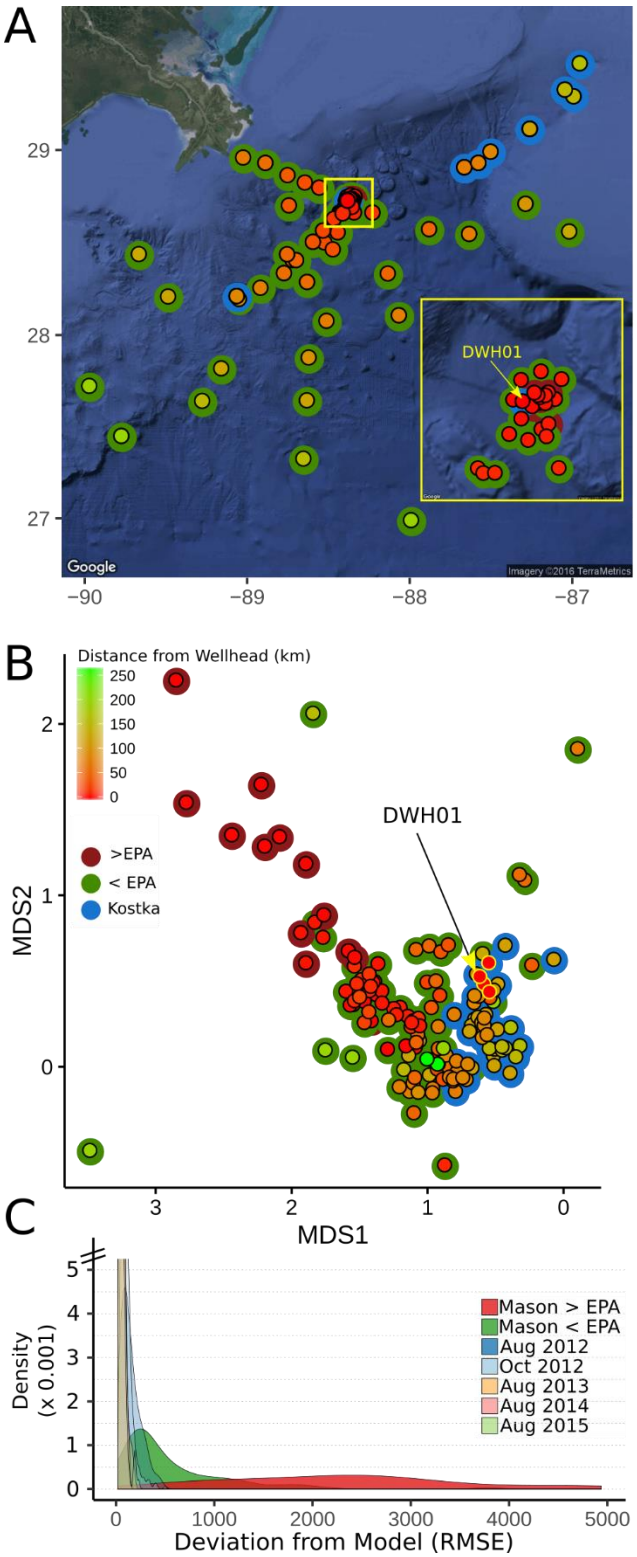

Figure 7 Meta-study revealing localized impact and subsequent recovery. Data generated by Mason et al., 2014 from surficial samples (red, green) was re-analyzed with surficial samples from this study (blue). (A) Geographic map plotting each sampling location. Samples outlined in red

from Mason et al., exceeded EPA benchmarks while samples outlined in green were below EPA benchmarks. The fill color for each point represents distance from well head, going from red to green with increasing distance. (B) NMDS of Bray-Curtis dissimilarity values, plotted using the same color scheme as in A. The wellhead samples (DWH01) are indicated by yellow outlines and were collected in 2012 and 2013. (C) Deviations of every sample collected, including the full sediment core for samples from this study, from the OTU distribution models. Samples are colored by the sampling date or by oil contamination state. High RMSE (root mean squared error) values represent samples that strongly differ from OTU relative abundance predictions.