The ISME Journal (2011), 1–10

www.nature.com/ismej

ORIGINAL ARTICLE

Microbial gene functions enriched in the Deepwater Horizon deep-sea oil plume

Zhenmei Lu[1,2], Ye Deng[2], Joy D Van Nostrand[2], Zhili He[2], James Voordeckers[2], Aifen Zhou[2], Yong-Jin Lee[2], Olivia U Mason[3], Eric A Dubinsky[3], Krystle L Chavarria[3], Lauren M Tom[3], Julian L Fortney[3], Regina Lamendella[3], Janet K Jansson[3], Patrik D'haeseleer[3], Terry C Hazen[3] and Jizhong Zhou[2,3,4]

[1]*College of Life Sciences, Zhejiang University, Hangzhou, China;* [2]*Department of Botany and Microbiology, Institute for Environmental Genomics, University of Oklahoma, Norman, OK, USA;* [3]*Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA and* [4]*Department of Environmental Science and Engineering, Tsinghua University, Beijing, China*

The Deepwater Horizon oil spill in the Gulf of Mexico is the deepest and largest offshore spill in the United State history and its impacts on marine ecosystems are largely unknown. Here, we showed that the microbial community functional composition and structure were dramatically altered in a deep-sea oil plume resulting from the spill. A variety of metabolic genes involved in both aerobic and anaerobic hydrocarbon degradation were highly enriched in the plume compared with outside the plume, indicating a great potential for *intrinsic* bioremediation or natural attenuation in the deep sea. Various other microbial functional genes that are relevant to carbon, nitrogen, phosphorus, sulfur and iron cycling, metal resistance and bacteriophage replication were also enriched in the plume. Together, these results suggest that the indigenous marine microbial communities could have a significant role in biodegradation of oil spills in deep-sea environments.

The ISME Journal advance online publication, 4 August 2011; doi:10.1038/ismej.2011.91
Subject Category: microbial ecology and functional diversity of natural habitats
Keywords: oil spill; deep-sea plume; microbial community; metagenomics; functional gene arrays; GeoChip

Introduction

On 20 April 2010, a massive oil leak occurred in the Gulf of Mexico's Mississippi Canyon area at a depth of 1544 m, releasing ~4.9 million barrels of crude oil into the deep ocean before the wellhead was finally capped on 15 July 2010 (The Federal Interagency Solutions Group, Oil Budget Calculator Science and Engineering Team, November 2010). Chemical dispersants, including COREXIT EC9500A and COREXIT EC9527A, were used on site at one of the highest rates in history to accelerate oil dispersal. A deep-water oil plume was initially detected at a depth of 1000–1200 m below the surface (Camilli *et al.*, 2010; Hazen *et al.*, 2010; Mascarelli, 2010a), but at last account (Mascarelli, 2010b) could no longer be detected, presumably as a result of dispersion and microbial degradation (OSAT, 2010). Significant environmental differences in the deep sea of Gulf of Mexico from other historic offshore oil spills present an urgent need to better understand the fate and impacts of the oil on this specific habitat (Kerr *et al.*, 2010a,b).

In marine ecosystems, microorganisms are known to have predominant roles in degradation of oil contaminants (Head *et al.*, 2003; Larter *et al.*, 2003). Therefore, it was expected that the indigenous microbial communities would have a significant role in degradation of the deep oil plume. This hypothesis was supported by two recent studies that explored the microbial and chemical properties of samples collected from the deep oil plume (Camilli *et al.*, 2010; Hazen *et al.*, 2010). Hazen *et al.* (2010) used a combination of molecular, chemical and physiological approaches to investigate the microbial and chemical composition in the deep-sea plume compared with uncontaminated water from the same depth outside the plume. They demonstrated that the oil depletion was due to a combination of mixing, dispersion and biodegradation by microbes residing in the deep sea (Hazen *et al.*, 2010).

In this study, samples from the deep-sea plume, oil-contaminated seawater (hereafter referred to as 'oil plume' in the following text) and non-plume controls (seawater samples at same depth that were not contaminated with oil) were analyzed with a functional gene microarray, the GeoChip 4.0

Correspondence: J Zhou, Department of Botany and Microbiology, Institute for Environmental Genomics, University of Oklahoma, Norman, OK 73019, USA.
E-mail: jzhou@ou.edu
Received 3 March 2011; revised 6 June 2011; accepted 6 June 2011

(Hazen *et al.*, 2010), to address the following questions: (i) How did the oil contamination affect the marine microbial community functional composition and structure? (ii) How did different microbial functional genes involved in key microbial processes shift in response to the oil spill? (iii) Were functional genes specific to hydrocarbon (HC) degradation processes enriched in the oil plume? Our results indicated that the oil spill dramatically altered microbial community functional structure, the marine microbial communities present were metabolically diverse, and that these communities were able to respond to the oil spill.

Materials and methods

The following is the summary of methods used in this study. More detailed information is provided in Supplementary Data A.

Sample description
Between 27 May and 2 June 2010, seawater samples were collected from the Gulf of Mexico during two monitoring cruises on the R/V Ocean Veritas and R/V Brooks McCall (Supplementary Table S1) as previously described (Hazen *et al.*, 2010). Briefly, two colored dissolved organic matter WETstar fluorometers (WET Labs, Philomath, OR, USA) were attached to a CTD sampling rosette (Sea-Bird Electronics Inc., Bellevue, WA, USA) and used to detect the presence of oil. The fluorometer results were subsequently confirmed by laboratory HC analysis. Niskin bottles attached to the CTD rosette were used to capture water samples at various depths with detected HCs. Eight samples (BM053, BM054, BM057, BM058, BM064, OV201, OV401 and OV501) from the MC252 dispersed oil plume, and five samples (OV003, OV004, OV009, OV013 and OV014) from non-plume at a depth of 1099–1219 m were analyzed in this study.

To better define the geochemical properties of the plume and non-plume samples, two sets of variables were measured: (i) seawater variables (dissolved oxygen, temperature, small particle counts, total ammonia nitrogen, nitrite (NO_2-N), total iron), orthophosphate (PO_4-P) and acridine orange direct count) and (ii) oil composition variables (fluorometer detection of oil, benzene, toluene, ethylbenzene, isopropylbenzene, *n*-propylbenzene, 1,3,5-trimethylbenzene, tert-butylbenzene, 1,2,4-trimethylbenzene, sec-butylbenzene, *p*-isopropyltoluene, *n*-butylbenzene, naphthalene and *o*-xylene, *m*- and *p*-xylenes) (Hazen *et al.*, 2010).

DNA amplification and labeling
Approximately 100 ng of DNA that was previously extracted from the samples (Hazen *et al.*, 2010) was amplified using a modification of the Templiphi kit (GE Healthcare, Piscataway, NJ, USA). The amplified DNA (2 μg) was then labeled with Cy3 using random primers and the Klenow fragment of DNA polymerase I (Wu *et al.*, 2006) and then purified and dried in a SpeedVac (45 °C, 45 min; ThermoSavant, Milford, MA, USA) before hybridization.

GeoChip 4.0 hybridization and data pre-processing
The GeoChip 4.0, containing 83 992 50-mer oligonucleotide probes targeting 152 414 genes in 410 gene categories for different microbial functional and biogeochemical processes, was synthesized by NimbleGen (Madison, WI, USA). All hybridizations were carried out at 42 °C with 40% formamide for 16 h on a MAUI hybridization station (BioMicro, Salt Lake City, UT, USA). After hybridization, the arrays were scanned (NimbleGen MS200, Madison, WI, USA) at a laser power of 100%. Signal intensities were measured based on scanned images, and spots with signal-to-noise ratios lower than 2 were removed before statistical analysis as described previously (He *et al.*, 2010).

Statistical analysis
Pre-processed GeoChip data were further analyzed with different statistical methods: (i) microbial diversity index, the two-tailed *t*-test and response ratio (Luo *et al.*, 2006); (ii) hierarchical clustering for microbial community structure and composition (de Hoon *et al.*, 2004); (iii) analysis of similarity, permutational multivariate analysis of variance using distance matrices and multiresponse permutation procedure analysis of differences of microbial communities (Anderson, 2001); (iv) canonical correspondence analysis (CCA) for linking microbial communities to environmental variables (Ramette and Tiedje, 2007; Zhou *et al.*, 2008); and (v) partial CCA for co-variation analysis of wellhead distance and environmental variables (variation partitioning analysis). Details for all methods are provided in the Supplementary Information.

Results

Functional gene changes in response to oil spill
To assess the dynamic changes of microbial communities in response to oil spill, microbial community functional composition and structure were analyzed using functional gene arrays (GeoChip 4.0). Significantly more functional genes ($P<0.01$) were detected in the oil plume samples than in non-plume control (Supplementary Table S2). The overall microbial functional diversity was also significantly ($P<0.01$) higher in the plume samples based on Shannon–Weiner (H') and Simpson's ($1/D$) indices. Consistent with geochemical ordination patterns, hierarchical clustering analysis showed that all plume samples were clustered together and well separated from non-plume samples (Figure 1

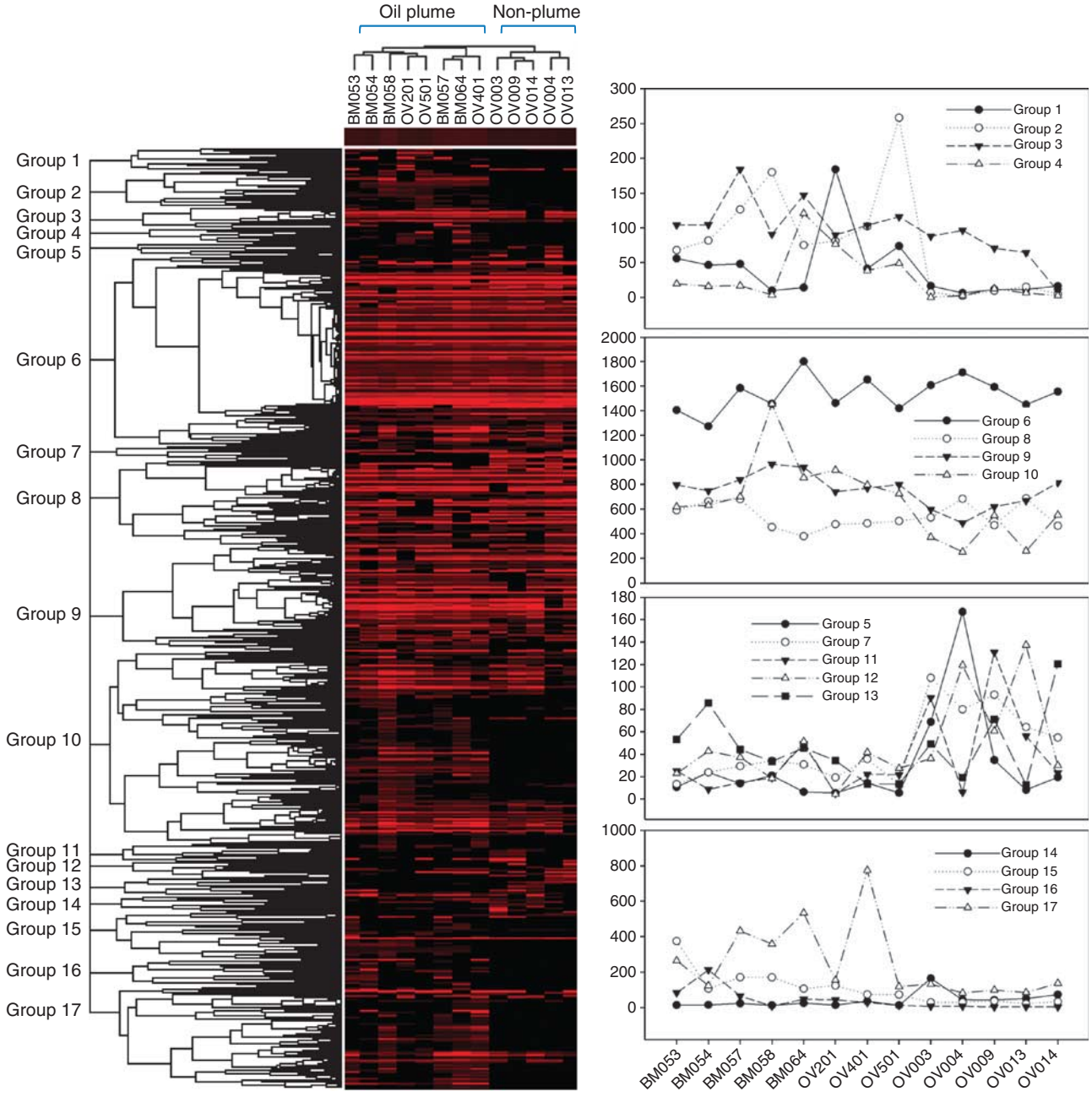

Figure 1 Hierarchical cluster analysis of all genes present in at least two out of the five samples. Results were generated in CLUSTER and visualized using TREEVIEW. Red indicates signal intensities above background, whereas black indicates signal intensities below background. Brighter red coloring indicates higher signal intensities. All oil plume samples clustered together and were well separated from non-plume samples.

and Supplementary Figure S1), as also shown for the microbial communities at a phylogenetic level (Hazen *et al.*, 2010). However, considerable variability in functional gene distribution was observed among different samples and some functional genes were common to all samples, although others were unique to oil plume samples (Figure 1). For example, Group 6, with 1439 or 20.14% of all genes detected, largely involved in organic remediation, carbon degradation, denitrification, sulfate reduction, metal resistance and stress response, was generally detected in all samples. Groups 1, 2, 10 and 17, with 2.2%, 3.9%, 20.5% and 10.8% of all genes detected, were mainly detected in the plume samples (Figure 1). In addition, the microbial community functional structure was significantly ($P<0.05$) different between the plume and non-plume samples as revealed by the three complementary non-parametric multivariate statistical tests (analysis of similarity, permutational

Table 1 Significance of the effects of the oil spill on the overall microbial community structure and geochemical pattern using three statistical analyses

Method	*Geochemical parameters*[a]		*Microbial community*	
	Statistic	P-*value*	*Statistic*	P-*value*
MRPP[b]	233.112	0.037	53.617	0.003
ANOSIM[c]	0.057	0.046	0.501	0.002
Adonis[d]	0.258	0.043	0.192	<0.001

Abbreviations: ANOSIM, analysis of similarity; MRPP, multi response permutation procedure.
[a]Geochemical parameters included temperature, DO concentration, fluorometer detection of oil, small particle concentrations, Fe, nitrate, phosphate, benzene, toluene, naphthalene, ethylbenzene, isopropylbenzene, *n*-propylbenzene, 1,3,5-trimethylbenzene, tert-butylbenzene, 1,2,4-trimethylbenzene, sec-butylbenzene, *p*-isopropyltoluene, *n*-butylbenzene, total xylenes, total volatile HC and total petroleum hydrocarbons—extractable (DRO).
[b]Multiple response permutation procedure, a nonparametric procedure that does not depend on assumptions such as normally distributed data or homogeneous variances, but rather depends on the internal variability of the data.
[c]Analysis of similarities.
[d]Non-parametric multivariate analysis of variance (MANOVA) with the adonis function.
All three tests are non-parametric multivariate analyses based on dissimilarities among samples.

multivariate analysis of variance using distance matrices and multiresponse permutation procedure) (Table 1).

Oil as a predominant factor shaping microbial community functional structure

CCA was performed to determine the most significant environmental variables shaping microbial community structure. On the basis of variance in inflation factors, seven variables were selected: dissolved oxygen, temperature, total volatile HC, total extractable petroleum HC, fluorometer detection of oil, phosphate and iron. The specified CCA model was significant ($P=0.026$). Of these, the total volatile HC, extractable petroleum HC, fluorometer detection of oil and dissolved oxygen were the most significantly correlated with plume samples (Figure 2). To separate the effects of seawater geochemical variables, geographic distance and oil composition on microbial community structure, a CCA-based variation partitioning analysis (Ramette and Tiedje, 2007; Zhou *et al.*, 2008) was performed. Seawater geochemical variables, oil composition and wellhead distance showed a significant correlation ($P=0.041$) with the functional gene structure of the community. Oil composition explained substantially more variations (48.34%, $P=0.03$) than seawater variables (21.76%, $P=0.017$), whereas distance independently explained 9.1% ($P=0.43$) of the observed variation (Figure 3). About 28% of the community functional variation based on GeoChip data remained unexplained by the above selected variables, which is significantly lower than those observed in other systems such as soils

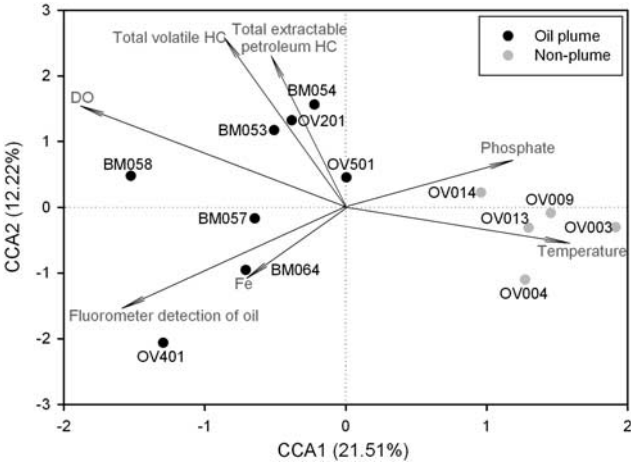

Figure 2 CCA compares the GeoChip hybridization signal intensities (symbols) and environmental variables (arrows). Environmental variables were chosen based on significance calculated from individual CCA results and variance inflation factors (VIFs) calculated during CCA. The percentage of variation explained by each axis is shown, and the relationship is significant ($P=0.026$).

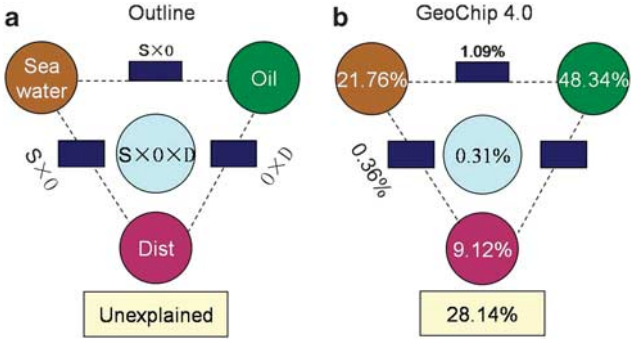

Figure 3 Variation partitioning based on CCA for all functional gene signal intensities. (**a**) General outline, (**b**) all functional genes. A CCA-based VIF was performed to identify common sets of oil composition and seawater variables important to the microbial community structure. Oil composition variables included fluorometer detection of oil, the concentration of total volatile HCs, xylenes and petroleum HCs—extractable (DRO). Seawater geochemical variables included temperature, dissolved oxygen (DO), Fe and phosphate.

(Ramette and Tiedje, 2007; Zhou *et al.*, 2008). These results indicate that oil contaminants could be a dominant factor shaping microbial community functional structure and potentially regulating associated microbial functional processes.

Oil spill stimulated increase in functional genes for HC degradation

A substantial number of genes involved in HC degradation were detected in the oil plume samples (Hazen *et al.*, 2010), especially those involved in degrading alkanes, alkynes and cycloalkanes, BTEX and related aromatics, chlorinated aromatics, heterocyclic aromatics, nitroaromatics, polycyclic aromatics and aromatic carboxylic acids. For

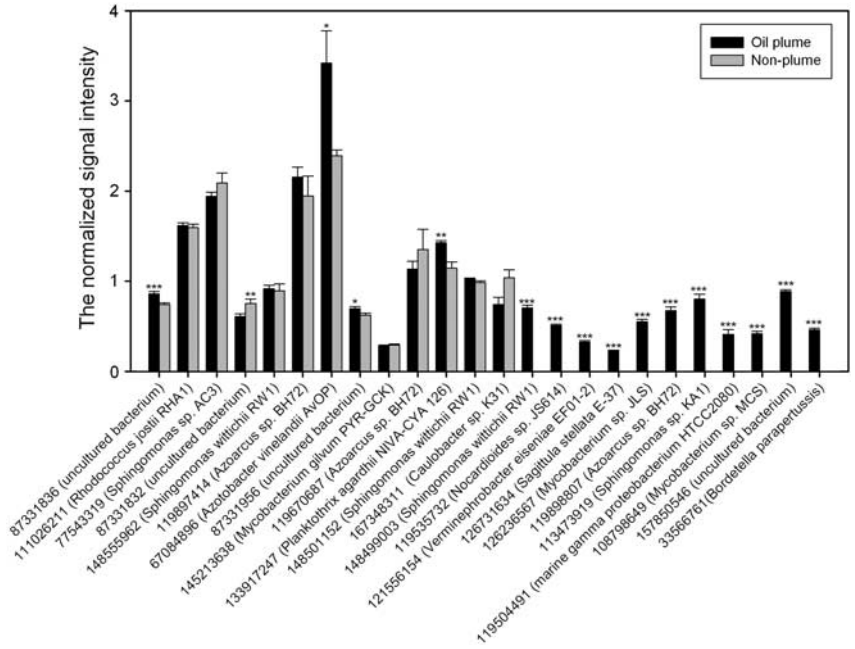

Figure 4 The normalized signal intensity of the *nahA* genes (naphthalene 1,2-dioxygenase) for the initial oxidation of naphthalene. The signal intensity for each sequence was the average of the total signal intensity from all the replicates. Gene number is the protein ID number for each gene as listed in the GenBank database. All data are presented as mean ± s.e. $***P<0.01$, $**P<0.05$, $*P<0.1$.

example, gene *alkB* encoding alkane 1-monooxygenase, a key enzyme responsible for the initial oxidation of inactivated alkanes, showed a significantly ($P<0.05$) higher abundance, with 19–26 genes detected in the oil contaminated samples and 11–15 detected in the non-oil contaminated samples. The *alkB* genes derived from *Rhodospirillum centenum* SW, *Bdellovibrio bacteriovorus* HD100, *Prauserella rugosa*, *Roseobacter* sp. CCS2, *Mycobacterium bovis* AF2122/97, *Bacillus* sp. BTRH40, *Gordonia* sp. Cg and *Rhodococcus* sp. RHA1 appeared to be dominant in all oil plume samples (Supplementary Figure S2).

GeoChip analysis also detected many aerobic PAH degradation genes from a variety of microorganisms (Figure 4 and Supplementary Figure S3). PAH degradation genes were more abundant in the plume samples, while some were unique to the plume samples. Although oxygen was still present in the plume samples (Camilli *et al.*, 2010; Hazen *et al.*, 2010), the gene *bbs* (beta-oxidation of benzylsuccinate) for anaerobic toluene degradation was also enriched in plume samples. These *bbs* genes were derived from putative E-phenylitaconyl-CoA hydratase of *Azoarcus* sp. EbN1 and *Thauera aromatic*, and benzylsuccinyl-CoA dehydrogenase of *Azoarcus* sp. EbN1 (Figure 5).

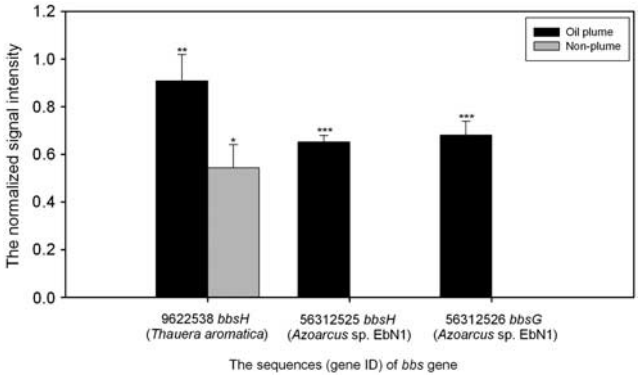

Figure 5 The normalized signal intensity of *bbs* (β-oxidation of benzylsuccinate) genes for anaerobic toluene degradation. The signal intensity for each sequence was the average of the total signal intensity from all the replicates. Gene number is the protein ID number for each gene as listed in the GenBank database. All data are presented as mean ± s.e. $***P<0.01$, $**P<0.05$, $*P<0.1$. In total, seven probes were designed for *bbs* genes in GeoChip 4.0 and three probes were detected in the samples.

Shifts of the genes involved in key biogeochemical cycling processes

Carbon. Among the carbon cycling genes detected, 798 genes involved in the degradation of complex carbon compounds, such as starch, hemicellulose, cellulose, chitin, lignin and aromatics, showed positive hybridization signals. Most of these genes (for example, *pulA*, *xylA*, *xynA*, *lip, limEH* and *vanA*) showed significantly ($P<0.05$) higher abundance in plume than in non-plume samples (Supplementary Figure S4). These types of genes could also be important in degradation of various oil components and their intermediates.

In this study, 9–14 *mcrA* genes encoding the α subunit of methyl coenzyme *M* reductase and 5–8

pmoA genes for methane monooxygenase were detected in the plume samples. Specifically, *mcrA* genes from *Methanococcus aeolicus* Nankai-3, *Methanoculleus marisnigri* JR1 and *Methanocorpusculum labreanum* Z were detected in all of the oil plume samples, but most of them were from uncultured microorganisms. Significantly ($P<0.05$) higher signal intensities were observed for *mcrA* in the plume than in the non-plume samples (Supplementary Figure S5). However, no significant differences were found for *pmoA* and *mmoX* (particulate methane monooxygenase) between plume and non-plume samples.

Nitrogen. Petroleum generally contains about 0.1–2% nitrogen, and given the large quantities of oil involved, it may act as an N pool in this ecosystem. Interestingly, *nasA* (nitrate reductase) and *nir* (nitrite reductase) for assimilatory N reduction, and *gdh* (glutamate dehydrogenase) for ammonia assimilation exhibited significantly ($P<0.05$ or 0.01) higher signal intensities in plume samples (Figure 6). The observed stimulation of N assimilation processes could be due to an increase of microbial biomass (Hazen *et al*., 2010). However, no significant differences were observed for other N-cycling genes, for example, nitrification, denitrification and N fixation (Figure 6).

Sulfur. Sulfite reduction genes were highly abundant in the deep-sea plume: 81–102 *dsrA/B* genes for dissimilatory sulfite reductase, and 8–12 *AprA* genes for dissimilatory adenosine-5′-phosphosulfate reductase were detected with significantly ($P<0.05$) higher abundance in the plume than in non-plume samples (Supplementary Figure S6). Microbial populations similar to *Alkalilimnicola ehrlichei* MLHE-1, *Chlorobium ferrooxidans* DSM 13031, *Clostridium leptum* DSM 753, *Desulfomicrobium thermophilum*, *Pyrobaculum calidifontis* JCM 11548, *Thermodesulforhabdus norvegica*, *Magnetococcus* sp. MC-1, *Pyrobaculum aerophilum* str. IM2, *Alkalilimnicola ehrlichei* MLHE-1, *Desulfohalobium retbaense* DSM 5692, sulfate-reducing bacterium QLNR1 and *Syntrophobacter fumaroxidans* MPOB were frequently detected in each sample, while most of the genes detected were from uncultured microorganisms (for example, sulfate-reducing bacteria) from various environments. The results suggest that sulfate reduction could be enhanced when coupled with HC degradation.

Phosphorus and iron reduction. As phosphorus is often a limiting factor for oil bioremediation, it is essential to understand phosphorus cycling in marine ecosystems. Genes encoding exopolyphosphatase (*ppx*) for inorganic polyphosphate degradation and phytase for phytate degradation were detected with significantly ($P<0.01$ and $P<0.05$, respectively) increased abundance in plume samples (Supplementary Figure S7). These results

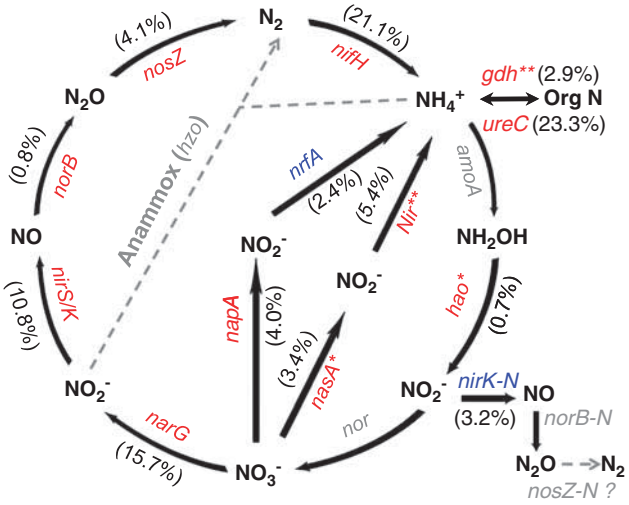

Figure 6 The relative changes of the detected genes involved in the N cycle in oil plume. The signal intensity for each gene detected was normalized by all detected gene sequences using the mean. The percentage of a functional gene in a bracket was the sum of signal intensity of all detected sequences of this gene divided by the grand sum of signal intensity of the detected N cycle genes, and weighted by the fold change of the signal intensity of this gene in plume to that in non-plume. For each functional gene, red indicates that this gene had a higher signal intensity in plume than in non-plume and their significance was indicated with two stars (**) at $P<0.01$, whereas blue indicates that this gene had a lower signal intensity in oil-plume than in non-plume. Grey-colored genes were not targeted by this GeoChip, or not detected in those samples. It remains unknown if *nosZ* homologs exist in nitrifiers. Description of the genes: (a) *gdh*, encoding glutamate dehydrogenase, *ureC*, encoding urease responsible for ammonification; (b) *nasA*, encoding nitrate reductase, *NiR*, encoding nitrite reductase, responsible for assimilatory N reduction; (c) *nifH*, encoding nitrogenase responsible for N_2 fixation; (d) *narG* encoding nitrate reductase, *nirS* and *nirK-D* (with denitrification activity), encoding nitrite reductase; *nosZ*, encoding nitrous oxide reductase, *norB*, encoding nitric oxide reducatse, responsible for denitrification (e) *napA*, encoding periplasmic nitrate reductase, *nrfA*, encoding c-type cytochrome nitrite reducatse, responsible for dissimilatory N reduction to ammonium; (f) *hao*, encoding hydroxylamine oxidoreductase, and *nirK-N* encoding nitrite reductase for nitrifiers (an indication of nitrification activity), responsible for nitrification.

suggested that organic phosphorus release could be stimulated by oil contamination. In addition, higher ($P<0.1$) signal intensities for 61 detected *cytochrome c* genes were observed in plume samples (Supplementary Figure S8), suggesting that HC degradation coupled with metal reduction could occur in the deep water.

Metal resistance. A substantial number (917) of the genes involved in resistance to various metals were detected, many of which showed significantly ($P<0.05$) increased abundance in plume samples (Supplementary Figure S9). Genes encoding reductases for As (*arsC*) and Hg (*mer*), efflux transporters for Cd (*cadA*), Cu, Co and Zn (*czcA* and *czcD*), Cr (*ChrA*), Cu (*copA*), Hg (*merT*), Ag (*silC*) and Zn (*zntA*), and the proteins involved in Te resistance

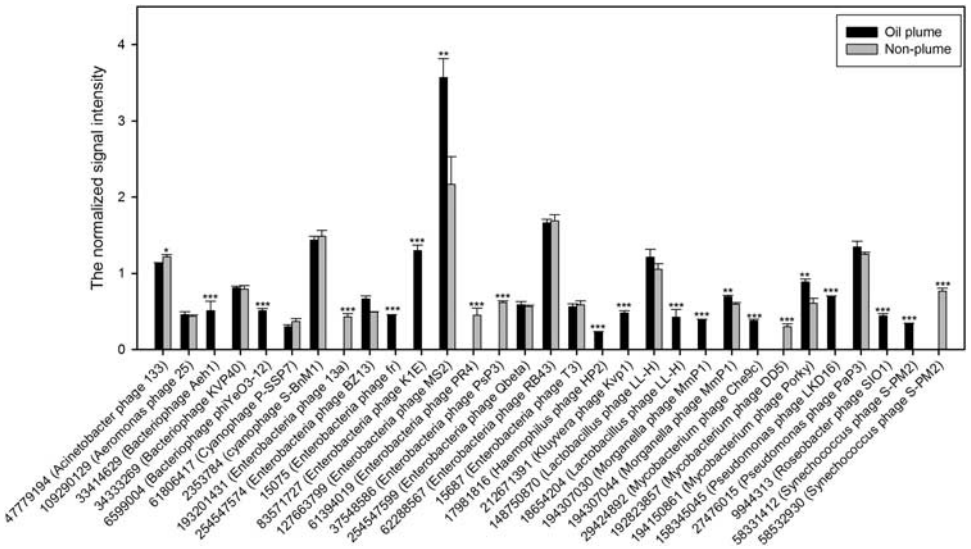

Figure 7 The normalized signal intensity of the replication genes for bacteriophage. The signal intensity for each sequence was the average of the total signal intensity from all the replicates. Gene number is the protein ID number for each gene as listed in the GenBank database. All data are presented as mean ± s.e. $***P<0.01$, $**P<0.05$, $*P<0.1$.

(*terC*, *terD* and *terZ*) were more ($P<0.05$ or 0.01) abundant in the plume samples.

Bacteriophages were also significantly stimulated
In total, 52 bacterial phage genes associated with host recognition, lysis, replication and structure were observed in all samples. The signal intensities for many of the genes involved in replication were significantly ($P<0.05$) higher in the plume than in the non-plume samples (Figure 7), supporting the suggestion by Head *et al*. (2006) that bacteriophages could be an important factor for intrinsic bioremediation of HCs.

Discussion

The Deepwater Horizon oil spill in the Gulf of Mexico was one of the worst environmental disasters in the United State history. The impact of an oil spill of such an unprecedented magnitude and depth on marine ecosystems is largely unknown. Using the GeoChip-based high-throughput microarray technology, we showed that diverse microbial functional groups (a group of genes involved in certain functional processes), including those important to HC degradation, carbon metabolism, methanogenesis, nitrogen assimilation, sulfate reduction, phosphorus release, metal resistance and bacteriophage replication, were more highly represented in the oil plume samples than in non-plume samples from the same depth. Also, the changes in community functional structure were highly correlated to the changes in geochemistry, with oil being the predominant factor shaping the functional composition and structure of the microbial communities. Our results support the phylogeny-based study by Hazen *et al*. (2010) that the deep-sea marine microbial communities underwent a dynamic change in response to the oil spill and associated geochemical changes. Our results are also consistent with previous studies of oil spill and petroleum contamination (Harayama *et al*., 2004; Head *et al*., 2006; Bordenave *et al*., 2007), which showed that microorganisms able to utilize HCs became dominant in oil-contaminated sites. Such functional gene information is useful for assessing the impacts of oil spills and should facilitate the design of appropriate strategies and approaches to deal with petroleum contamination.

The clean-up of the deep-sea oil plume will primarily depend on the indigenous microbes present in this environment, as current technology does not allow removing the dispersed oil and gas at such great depths. One of the critical environmental questions is whether microorganisms for degrading various HCs exist in the community and whether they respond to oil spill. Our GeoChip results indicated that many functional genes/populations involved in both aerobic and anaerobic degradation of various oil components are detected and/or enriched in the oil plume, indicating that the indigenous HC-degrading populations are capable of responding to the oil spill. For example, *alkB* for alkanes, *Xamo* for alkene, genes *bco*, *ohbAB*, *GCoADH* and *pimF* for benzoate, genes *mdlA, mdlB* and *mdlC* for mandelate, and genes *Apc* and *catB* for BTEX metabolic pathway exhibited a significantly ($P<0.05$) higher abundance in the oil plume than in the non-oil plume. The changes in relative abundance of these genes/populations were significantly correlated with the concentrations of various oil contaminants in the samples (Hazen *et al*., 2010).

Especially, several genes for PAH degradation were enriched in the oil plume samples, which could be important in determining the long-term effects of the oil spill on the marine ecosystems. Also, consistent with phylogenetic gene distribution obtained using a phylogenetic microarray 'PhyloChip' (Hazen *et al.*, 2010), functional genes representative of the order *Oceanospirillales* appeared to have significantly higher ($P<0.01$) abundance in the plume samples than in non-plume samples, although the dominance of the *Oceanospirillum* population consuming the oil in the plume was based on clone library and sequence analysis of 16S rRNA genes (Hazen *et al.*, 2010). GeoChip was not originally designed to link the detection of functional genes to the existence of related microbial population and it contains 567 functional genes derived from the order *Oceanospirillales*, with 25 genes detected in this study. In addition, a large number of metal resistance genes were enriched in plume samples, which are usually linked to organic degradation genes, for example, on plasmids (Parales and Haddock, 2004; Kunapuli *et al.*, 2007). Our GeoChip results demonstrated that there is a great potential for *intrinsic* bioremediation of oil contamination in the deep-sea environment.

Anaerobic HC degradation associated with sulfate reduction, denitrification and methanogenesis has long been considered the prevailing mechanism for petroleum biodegradation in the deep subsurface (Head *et al.*, 2003; Aitken *et al.*, 2004; Kniemeyer *et al.*, 2007; Jones *et al.*, 2008). Recent investigations have demonstrated that several classes of petroleum HCs, including alkanes (So *et al.*, 2003), mono- and polycyclic aromatic compounds (Meckenstock *et al.*, 2000; Widdel and Rabus, 2001), and short-chain HCs (Kniemeyer *et al.*, 2007), can be degraded anaerobically under nitrate-, iron- or sulfate-reducing conditions, or under methanogenic conditions (Harayama *et al.*, 2004; Jones *et al.*, 2008). Indeed, a substantial number of *dsrA/B* genes for sulfate reduction, *mcrA* genes for methanogenesis, *narG*, *nirS*, *nirK* and *nosZ* responsible for denitrification and populations for metal reduction were detected in this study. Also, *dsrA/B* and *mcrA* genes showed significantly ($P<0.05$ or 0.01) higher abundance in the plume than in the non-plume samples. In addition, *bbs* genes for the strict anaerobic toluene degradation were detected and enriched in the plume samples. It is possible that anaerobic HC degradation could have most likely occured through microaggregate formation as reported in Hazen *et al.* (2010).

Hydrocarbon degradation is generally limited by nutrient availability, which can be improved by nutrient recycling through phage-mediated biomass turnover (Jiang *et al.*, 1998; Head *et al.*, 2006; Paul, 2008). As significant biomass increase was observed (Hazen *et al.*, 2010) in the plume samples, bacteriophages could have critical roles in HC degradation. Approximately 43% of marine bacterial isolates have been found to contain prophages (Jiang *et al.*, 1998; Paul, 2008), which are induced by various environmental contaminants, such as fuel oil (Cochran *et al.*, 1998). The oil spill may stimulate the growth of pathogenic bacteria in marine environments and many pathogens are capable of efficiently degrading HCs (Rojo and Martínez, 2010). The research on phages has been heavily slanted to those that affect human-related activities, health/medical and industry. As no target genes for *Oceanospirillum* phages were designed on GeoChip 4.0, the *Oceanospirillum* phages were not detected. Genes for both iron uptake (*iro*) and adherence (*pap* and *pilin*) were significantly ($P<0.01$ or 0.05) enriched in the plume samples. The increase in the abundance of microorganisms capable of producing siderophores, highly specific iron-chelating compounds, may facilitate microbial acquisition of iron, a limiting nutrient in marine systems (Barbeau *et al.*, 2001a, b), thereby potentially increasing HC degradation.

A substantial quantity of methane gas was released together with the oil (The Federal Interagency Solutions Group, Oil Budget Calculator Science and Engineering Team, November 2010; Kessler *et al.*, 2011), which may result in more methane in the oil plume ecosystem and have the potential to greatly impact methane metabolism. GeoChip targets three key genes/enzymes involved in methane metabolism, with *mcrA* encoding methyl coenzyme M reductase for methanogenesis and two enzymes/genes (methane monooxygenase/*mmoX* and particulate methane monooxygenase/*pmoA*) for methanotrophy (He *et al.*, 2010). In this study, *pmoA* and *mmoX* genes for aerobic methane oxidation did not show a statistically significant change though their abundance was higher in plume samples than in non-plume samples. There are two possible explanations for this: one is that the aerobic methane oxidation was inhibited owing to the presence of easier to degrade alkanes in the deep sea, and the other is that the methane gas was moved up to the surface more directly and did not accumulate in the deep oil plume. Also, unlike propane, methane may form gas hydrates at the deep plume temperature and pressure, making it unavailable to microorganisms (Valentine *et al.*, 2010). However, significantly ($P<0.05$) higher signal intensities were observed for *mcrA* in the plume than non-plume samples, indicating that those enriched *mcrA* genes derived from methanogens likely link to HC degradation rather than plume methane release (Harayama *et al.*, 2004; Jones *et al.*, 2008). Enzymes or genes involved in anaerobic methane oxidation, however, remain unclear; thus, we could not detect this functional process.

In this study, many functional genes were detected in the uncontaminated samples that were not detected in the contaminated samples (Supplementary Table S3 and Supplementary Figure S10). These results suggest that oil spills can select against those populations containing these genes,

or that specific members of the community have a selective advantage if they are capable of HC degradation and these grow to represent a greater proportion of the functional gene repertoire.

In conclusion, our results indicate that a variety of HC-degrading functional genes were enriched in response to oil contamination and associated environmental changes. Our results also imply that there is a great potential for *in situ* bioremediation of oil contaminants in the deep-seawater ecosystem, and such oil-degrading populations and associated microbial communities may have a significant role in determining the ultimate fates and consequences of the spilled oil. However, to further understand and evaluate the potential impacts of this unprecedented oil spill on the marine ecosystem structure and function, it is essential to launch an integrated and comprehensive monitoring program to track the dynamics and adaptive responses of microbial communities together with other physical and chemical analysis of tracing oil contaminants and their products.

Acknowledgements

We thank the Captain, crew and science teams aboard the R/V Ocean Veritas and R/V Brooks McCall. This work was part of ENIGMA, a Scientific Focus Area Program supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomics: GTL Foundational Science through Contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the US Department of Energy. This study was also supported by University of Oklahoma Research Foundation, the National Key Science and Technology Project of China: Water Pollution Control and Treatment (No. 2008ZX07101-006), the Natural Science Foundation of Zhejiang Province (No. R5080124) and the National Key Technologies Research and Development Program of China during the 11th Five-Year Plan Period (No. 2006BAJ08B01).

References

Aitken CM, Jones DM, Larter SR. (2004). Anaerobic hydrocarbon biodegradation in deep subsurface oil reservoirs. *Nature* **431**: 291–294.

Anderson M. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecol* **26**: 32–46.

Barbeau K, Rue EL, Bruland KW, Butler A. (2001a). Photochemical cycling of iron in the surface ocean mediated by microbial iron(III)-binding ligands. *Nature* **413**: 409–413.

Barbeau K, Zhang G, Live DH, Butler A. (2001b). Petrobactin, a photoreactive siderophore produced by the oil-degrading marine bacterium *Marinobacter hydrocarbonoclasticus*. *J Am Chem Soc* **124**: 378–379.

Bordenave S, Goni-Urriza MS, Caumette P, Duran R. (2007). Effects of heavy fuel oil on the bacterial community structure of a pristine microbial mat. *Appl Environ Microbiol* **73**: 6089–6097.

Camilli R, Reddy CM, Yoerger DR, Van Mooy BAS, Jakuba MV, Kinsey JC *et al.* (2010). Tracking hydrocarbon plume transport and biodegradation at Deepwater Horizon. *Science* **330**: 201–204.

Cochran PK, Kellogg CA, Paul JH. (1998). Prophage induction of indigenous marine lysogenic bacteria by environmental pollutants. *Mar Ecol Prog Ser* **164**: 125–133.

de Hoon MJ, Imoto S, Nolan J, Miyano S. (2004). Open source clustering software. *Bioinformatics* **20**: 1453–1454.

Harayama S, Kasai Y, Hara A. (2004). Microbial communities in oil-contaminated seawater. *Curr Opin Biotechnol* **15**: 205–214.

Hazen TC, Dubinsky EA, DeSantis TZ, Andersen GL, Piceno YM, Singh N *et al.* (2010). Deep-sea oil plume enriches indigenous oil-degrading bacteria. *Science* **330**: 204–208.

He Z, Deng Y, Van Nostrand JD, Tu Q, Xu M, Hemme CL *et al.* (2010). GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. *ISME J* **4**: 1167–1179.

Head IM, Jones DM, Larter SR. (2003). Biological activity in the deep subsurface and the origin of heavy oil. *Nature* **426**: 344–352.

Head IM, Jones DM, Roling WFM. (2006). Marine microorganisms make a meal of oil. *Nat Rev Micro* **4**: 173–182.

Jiang SC, Kellogg CA, Paul JH. (1998). Characterization of marine temperate phage-host systems isolated from Mamala Bay, Oahu, Hawaii. *Appl Environ Microbiol* **64**: 535–542.

Jones DM, Head IM, Gray ND, Adams JJ, Rowan AK, Aitken CM *et al.* (2008). Crude-oil biodegradation via methanogenesis in subsurface petroleum reservoirs. *Nature* **451**: 176–180.

Kerr R, Kintisch E, Stokstad E. (2010a). Will deepwater horizon set a new standard for catastrophe? *Science* **328**: 674–675.

Kerr RA, Kintisch E, Schenkman L, Stokstad E. (2010b). Five questions on the spill. *Science* **328**: 962–963.

Kessler JD, Valentine DL, Redmond MC, Du MR, Chan EW, Mendes SD *et al.* (2011). A Persistent oxygen anomaly reveals the fate of spilled methane in the deep Gulf of Mexico. *Science* **331**: 312–315.

Kniemeyer O, Musat F, Sievert SM, Knittel K, Wilkes H, Blumenberg M *et al.* (2007). Anaerobic oxidation of short-chain hydrocarbons by marine sulphate-reducing bacteria. *Nature* **449**: 898–901.

Kunapuli U, Lueders T, Meckenstock RU. (2007). The use of stable isotope probing to identify key iron-reducing microorganisms involved in anaerobic benzene degradation. *ISME J* **1**: 643–653.

Larter S, Wilhelms A, Head I, Koopmans M, Aplin A, Di Primio R *et al.* (2003). The controls on the composition of biodegraded oils in the deep subsurface-part 1: biodegradation rates in petroleum reservoirs. *Org Geochem* **34**: 601–613.

Luo Y, Hui D, Zhang D. (2006). Elevated CO_2 stimulates net accumulations of carbon and nitrogen in land ecosystems: a meta-analysis. *Ecology* **87**: 53–63.

Mascarelli A. (2010a). Extent of lingering Gulf oil plume revealed. *Nature*; doi: 10.1038/news.2010.1420. http://www.nature.com/nature/archive/category.html?code=archive_news&year=2010&month=08&page=2.

Mascarelli A. (2010b). Deepwater horizon: after the oil. *Nature* **467**: 22–24.

Meckenstock RU, Annweiler E, Michaelis W, Richnow HH, Schink B. (2000). Anaerobic naphthalene degradation by a sulfate-reducing enrichment culture. *Appl Environ Microbiol* **66**: 2743–2747.

Operational Science Advisory Team (OSAT) (2010). Summary report for sub-sea and sub-surface oil and dispersant detection: sampling monitoring, 17 December 2010. http://www.restorethegulf.gov/release/2010/12/16/data-analysis-and-findings.

Parales RE, Haddock JD. (2004). Biocatalytic degradation of pollutants. *Curr Opin Biotechnol* **15**: 374–379.

Paul JH. (2008). Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? *ISME J* **2**: 579–589.

Ramette A, Tiedje J. (2007). Biogeography: an emerging cornerstone for understanding prokaryotic diversity, ecology, and evolution. *Microb Ecol* **53**: 197–207.

Rojo F, Martínez JL. (2010). Oil Degraders as Pathogens. In: Timmis KN (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin, Heidelberg, pp 3293–3303.

So CM, Phelps CD, Young LY. (2003). Anaerobic transformation of alkanes to fatty acids by a sulfate-reducing bacterium, strain Hxd3. *Appl Environ Microbiol* **69**: 3892–3900.

The Federal Interagency Solutions Group (2010). Oil Budget Calculator Science and Engineering Team. Oil budget calculator deepwater horizon technical documentation. November 2010. http://www.restorethegulf.gov/sites/default/files/documents/pdf/OilBudgetCalc_Full_HQ-Print_111110.pdf.

Valentine DL, Kessler JD, Redmond MC, Mendes SD, Heintz MB, Farwell C *et al.* (2010). Propane respiration jump-starts microbial response to a deep oil spill. *Science* **330**: 208–211.

Widdel F, Rabus R. (2001). Anaerobic biodegradation of saturated and aromatic hydrocarbons. *Curr Opin Biotechnol* **12**: 259–276.

Wu L, Liu X, Schadt CW, Zhou J. (2006). Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. *Appl Environ Microbiol* **72**: 4931–4941.

Zhou J, Kang S, Schadt CW, Garten CT. (2008). Spatial scaling of functional gene diversity across various microbial taxa. *Proc Natl Acad Sci USA* **105**: 7768–7773.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)

Supporting Information

A. MATERIALS AND METHOD

B. SUPPORTING TABLES

1. **Table S1**. Dispersed MC252 plume and control parameters at 1099-1219 m.
2. **Table S2**. The sampling site, the number of genes detected, and diversity indices of contaminated and control samples.
3. **Table S3**. All sequences present in non-plume samples but absent in plume samples.

C. SUPPORTING FIGURES

1. **Fig. S1** Ordination plot produced from principal-component analysis (PCA) of geochemical data for all the monitoring samples.
2. **Fig. S2** Canonical correspondence analysis (CCA) compares the GeoChip hybridization signal intensities and environmental variables.
3. **Fig. S3** Hierarchical cluster analysis of *alkB* gene, encoding alkane 1-monooxygenase.
4. **Fig. S4** The normalized signal intensity of the *arhA* (PAH dioxygenase) genes.
5. **Fig. S5** The normalized signal intensity of the detected key genes involved in carbon degradation.
6. **Fig. S6** The normalized signal intensity of the detected genes involved in methane metabolism.
7. **Fig. S7** The relative changes of the detected genes involved in the N cycle in plume.
8. **Fig. S8** The normalized signal intensity of the detected genes involved in sulfur cycling.
9. **Fig. S9** The normalized signal intensity of the detected genes involved in phosphorus cycling.
10. **Fig. S10** The normalized signal intensity of the cytochrome genes.
11. **Fig. S11** The normalized signal intensity of the detected key genes involved in metal resistance.
12. **Fig. S12** The normalized signal intensity of the replication genes for bacteriophage.
13. **Fig. S13** The sequences present in five replicates of control (non-plume) samples but absent in oil plume samples.

D. SUPPLEMENTAL REFERENCES

A. MATERIALS AND METHODS

1. Sample Collection

Water samples were collected from the Gulf of Mexico during two monitoring cruises from May 27-June 2 aboard the R/V Ocean Veritas and R/V Brooks McCall. The cruises were conducted as part of the monitoring effort to assess the effect of subsea dispersant use during the MC252 oil leak (http://www.epa.gov/bpspill/dispersants.html#directives). A colored dissolved organic matter (CDOM) WETstar fluorometer (WET Labs, Philomath, OR) was attached to a CTD sampling rosette (Sea-Bird Electronics Inc., Bellevue, WA) and used to detect the presence of oil. Fluorometer results were subsequently confirmed by laboratory hydrocarbon analysis. Eight samples (BM053, BM054, BM057, BM058, BM064, OV201, OV401 and OV501) from the MC252 dispersed oil plume, and five samples (OV003, OV004, OV009, OV013, OV014) from non-plume at depth of 1099-1219m were analyzed using GeoChip 4.0 (Hazen *et al*., 2010) (Table S2).

Niskin bottles attached to the CTD rosette were used to capture water samples at various depths where hydrocarbons were detected. From each sample 800-2000 mL of water were filtered through sterile filter units containing 47 mm diameter polyethylsulfone membranes with 0.22 μm pore size (MO BIO Laboratories, Inc., Carlsbad, CA) and then immediately frozen and stored at -20 ℃ for the remainder of the cruise. Filters were shipped on dry ice during transportation and stored at -80 ℃ until DNA extraction.

100 mL of water was syringe-filtered and injected into evacuated 25 mL serum bottles capped with thick butyl rubber stoppers to determine hydrocarbon concentrations and stable isotopes. 100 mL of water was frozen in 125 mL HDPE bottles for nutrient analyses. For AODC 36 mL water was preserved in 4% formaldehyde (final concentration).

2. Geochemical parameter analysis

The dispersed oil droplet size distribution was measured using a laser *in situ* scattering and transmissometry (LISST-100X, Sequoia Scientific, Seattle, WA) following the same procedure used for previous crude oil dispersion experiments (Li *et al.*, 2007)

Total ammonia nitrogen (TAN) was quantified using the TL-2800 ammonia analyzer made by Timberline Instruments (Boulder, CO) (Carlson *et al*., 1990). Nitrite (NO_2-N) was measured colormetrically using SM 4500-NO_2-N. Total Iron (Tot Fe) was measured using a reaction with phenanthroline according to SM 3500-Fe B. Ortho-phosphate (PO_4-P) was quantified on unfiltered samples by the ascorbic acid method adapted from SM 4500-P-E (APHA, 2005).

To determine hydrocarbon concentrations derived from the presence of oil in the samples, 200 µL of chloroform was added to the neutral lipid extract which was then vortexed followed by a 30 sec sonication. The extract was analyzed on an Agilent GC/FID and peaks were identified by GC/MS. Quantification was accomplished by comparison to a known hexadecane standard.

Volatile aromatic hydrocarbons were measured using USEPA methods 5030/8260b using an Agilent 6890 GC with 5973 mass spectrometer detector. Initial oven temperature 10℃, initial time 3.00 min, ramp 8 ℃/min to 188℃, then 16℃/min to 220℃, hold for 9.00 min. Split ratio 25:1. Restek Rtx-VMS capillary column, 60 m length by 250 µm diameter, 1.40 µm film. Scan 50 to 550 m/z.

Samples for direct counts were preserved with 4% formaldehyde and stored at 4℃. 1 to 10 ml sample were filtered through a 0.2 µm pore size black polycarbonate membrane (Whatman International Ltd., Piscataway, NJ) supported by a vacuum filtration sampling manifold (Millipore Corp., Billerica, MA). Filtered cells were stained with 25 mg/ml acridine orange for 2 min in the dark. Unbound acridine orange was filtered through the membrane with 10 ml filter sterilized 1X PBS (Sigma Aldrich Corp., St. Louis, MI) and the rinsed membrane was mounted on a slide for microscopy. Cells were imaged with a FITC filter on a Zeiss Axioskop (Carl Zeiss, Inc., Germany) (Francisco *et al*., 1973).

3. DNA Extraction

Filters were extracted using a modified Miller method (Miller *et al*., 1999). One quarter of each filter was cut into small pieces and placed in a Lysing Marix E tube (MP Biomedicals, Solon, OH). 300 µL of Miller phosphate buffer and 300 µL of Miller SDS lysis buffer were added and mixed. 600 µL phenol: chloroform: isoamyl alcohol (25:24:1) was then added, and the tubes were bead-beat at 5.5m/s for 45sec in a FastPrep instrument. The tubes were spun at 16,000 × g for 5 min at 4℃. 540 µL of

supernatant was transferred to a 2 mL tube and an equal volume of chloroform was added. Tubes were mixed and then spun at 10,000 ×g for 5 min, 400 µL aqueous phase was transferred to another tube and 2 volumes of Solution S3 (MoBio, Carlsbad, CA) was added and mixed by inversion. The rest of the clean-up procedures followed the instructions in the MoBio Soil DNA extraction kit. Samples were recovered in 60 µL Solution S5 and stored at -20 ℃.

4. GeoChip-based functional gene array hybridization

For assessing the impacts of oil plume on microbial community functional structure, DNA extracted from the oil plume and non-plume was used for functional gene array hybridization. Aliquots of DNA (4 µL) were amplified with the Templiphi kit (GE Healthcare; Piscataway, NJ) using WCAG (whole community genome amplification) (Wu *et al*., 2006) with modifications to increase DNA yield and minimize bias. All samples yielded between 2.8-3.3 µg amplified DNA. The amplified DNA (2 µg) was then labeled with Cy-3 using random primers and the Klenow fragment of DNA polymerase I (Wu *et al*., 2006). Labeled DNA was then dried in a SpeedVac (45 ℃, 45 min; ThermoSavant).

Dried DNA was rehydrated with 2.68 µL sample tracking control (NimbleGen, Madison, WI, USA) to confirm sample identity. The samples were incubated at 50 ℃ for 5 min, vortexed for 30 sec, and then centrifuged to collect all liquid at the bottom of the tube. Hybridization buffer (7.32 µL), containing 40% formamide, 25% SSC, 1% SDS, 2% Cy5-labeled common oligo reference standard (CORS) target, and 2.38% Cy3-labeled alignment oligo (NimbleGen) and 2.8% Cy5-labeled common oligonucleotide reference standard (CORS) target (Liang *et al*.,. 2009) for data normalization, was then added to the samples, vortexed to mix, spun down, incubated at 95 ℃ for 5 min, and then maintained at 42 ℃ until ready for hybridization. CORS probes were placed randomly throughout the array and are used for signal normalization (Liang *et al*., 2010).

GeoChip 4.0 is a new generation of functional gene array (He *et al*., 2010a, He *et al*., 2007), which contained 83,992 50-mer oligonucleotide probes targeting 152,414 genes in 410 gene categories for different microbial functional and biogeochemical processes including carbon, nitrogen, phosphorus, and sulfur cycling, energy processing, metal resistance and reduction, organic contaminant degradation, stress responses, antibiotic resistance, and bacteriophages. GeoChip 4.0 is synthesized by

NimbleGen in their 12-plex format (i.e., 12 arrays per slide). An HX12 mixer (NimbleGen) was placed onto the array using NimbleGen's Precision Mixer Alignment Tool (PMAT), and then the array is preheated to 42 ℃ on a Hybridization Station (MAUI, BioMicro Systems, Salt Lake City, UT, USA) for at least 5 min. Samples (6.8 μL) were then loaded onto the array surface and hybridized approximately 16 h with mixing.

After hybridization, the arrays were scanned with a laser power of 100% and 100% PMT (photomultiplier tube) (MS 200 Microarray Scanner, NimbleGen). Low quality spots were removed prior to statistical analysis as described previously (He *et al*., 2010b). Spots were scored as positive if the signal-to-noise ratio (SNR) was ≥2.0 and the CV of the background was <0.8. Genes that were detected in only one sample were removed.

5. Statistical analysis

All GeoChip 4.0 hybridization data are available at the Institute for Environmental Genomics, University of Oklahoma (http://ieg.ou.edu/). Pre-processed data were then used for further analysis. Hierarchical clustering was performed with CLUSTER 3.0 using uncentered correlations and the complete average linkage for both genes and samples, and trees were visualized in TREEVIEW (de Hoon *et al*., 2004). Functional gene diversity was calculated using Simpson's 1/D, Shannon-Weiner's H' and evenness. The effects of oil-plume on functional microbial communities were analyzed by two-tailed t-test or response ratio (RR) using the formula described by Luo *et al*., (2006). Based on the standard error, the 95% confident interval for each response variable was obtained and the statistical difference between the oil-plume and non-plume was estimated. For t-test and the response ratio analysis, the total abundance of each gene category or family was simply the sum of the normalized intensity for the gene category or family.

In this study, three different non-parametric analyses for multivariate data were used to examine whether oil plume has significant effects on deep sea microbial communities: analysis of similarity (anosim) analysis of similarities (ANOSIM) (Clarke, 1993), non-parametric multivariate analysis of variance (adonis) using distance matrices (Anderson, 2001), and multi-response permutation procedure (MRPP). All three methods are based on dissimilarities among samples and their rank

order in different ways to calculate test statistics, and the Monte Carlo permutation is used to test the significance of statistics.

Multivariate statistical analyses of GeoChip data including canonical correspondence analysis (CCA) for linking microbial communities to environmental variables (Zhou *et al*., 2008), partial CCA for co-variation analysis of wellhead distance and environmental variables (variation partitioning analysis, VPA) were performed. Selection for CCA modeling was conducted by an iterative procedure of eliminating redundant environmental variable based on variance inflation factor (VIF). All the analyses were performed by the vegan package in R 2.9.1 (R Development Core Team, 2006).

B. SUPPORTING TABLES

Table S1. Dispersed MC252 plume and control parameters at 1099-1219 m. Parameters with significant differences are highlighted (Student's T-test) (Hazen *et al*., 2010).

	Plume mean (s.d.)	**Non-plume** mean (s.d.)	**T-test** p-value
Physical-Chemical			
Fluorescence (mg/m^3)	24.2 (18.2)	5.9 (0.5)	**0.018**
Phosphate (µg/L)	39.8 (6.7)	40.7 (4.3)	0.781
Ammonia-N (µg/L)	23.6 (5.3)	20.8 (2.9)	0.347
Nitrate-N (µg/L)	277 (80)	359 (99)	**0.003**
d13C DIC	-0.57 (0.06)	-0.46 (0.14)	0.174
Total iron (µg/L)	47.9(2.2)	46.5(6.6)	0.702
Oil composition			
Fluorometer detection of oil (mg/m^3)	22.95(12.87)	5.98(0.21)	**0.018**
Benzene (µg/L)	47.12 (25.96)	0.38 (0.19)	**0.004**
Toluene (µg/L)	99 (55.63)	0.54 (0.25)	**0.004**
Isopropylbenzene (µg/L)	3.42 (1.39)	1.37 (0.31)	**0.012**
n-Propylbenzene (µg/L)	4.4 (2.96)	0.50 (0.31)	**0.019**
tert-Butylbenzene (µg/L)	1.52 (0.79)	0.42 (0.18)	**0.025**
1,2,4-Trimethylbenzene (µg/L)	29.56 (16.80)	0.72 (0.16)	**0.005**
n-Butylbenzene (µg/L)	1.32 (0.37)	0.71 (0.39)	**0.033**
Naphthalene (µg/L)	13.52 (8.12)	0.88 (0.82)	**0.008**
Total Xylenes (µg/L)	113.28 (64.05)	0.93 (0.77)	**0.004**
octadecane (ppb)	4.2 (2.4)	0.13 (0.18)	**<0.001**
n-docosane (ppb)	4.7 (2.7)	0.12 (0.17)	**<0.001**
Total volatile aromatic hydrocarbons[1]	139 (179)	0.5 (1.8)	**<0.001**
Total Petroleum Hydrocarbons - extractable (DRO)	6.4 (5.08)	0.45 (0.21)	**0.032**
Biological			
Bacteria density (Log(AODC))	4.59 (0.63)	4.01 (0.11)	**0.030**

[1]Benzene, toluene, ethylbenzene, isopropylbenzene, n-propylbenzene, 1,3,5-trimethylbenzene, tert-butylbenzene, 1,2,4-trimethylbenzene, sec-butylbenzene, p-isopropyltoluene, n-butylbenzene, naphthalene, o-xylene, m,p-xylenes.

Table S2. The sampling site, the number of genes detected, and diversity indices of contaminated and control samples.

	Sample name	Depth (m)	Latitude	Longitude	Distance[a] (km)	Gene No. (S)	Shannon (H')	Simpson (1/D)	SimpsonE
Oil plume	BM053	1219	28.735145	-88.381937	1.65	4460	8.21	2899.83	0.65
	BM054	1194	28.732133	-88.376850	1.32	4110	8.13	2740.43	0.67
	BM057	1174	28.705093	-88.401650	5.14	4834	8.30	3346.35	0.69
	BM058	1179	28.672323	-88.435935	10.08	5004	8.36	3624.74	0.72
	BM064	1099	28.683393	-88.448712	10.18	4973	8.33	3415.09	0.69
	OV201	1207	28.732011	-88.376789	1.33	4334	8.16	2696.70	0.62
	OV401	1181	28.732011	-88.376789	1.33	4789	8.27	3087.63	0.64
	OV501	1100	28.730275	-88.416872	5.09	4156	8.13	2676.41	0.64
	Average (SE)	**1169 (16)**			**4.52 (1.35)**	**4583 (128)**	**8.24 (0.03)**	**3060.90 (129.19)**	**0.67 (0.01)**
Control	OV003	1020	28.666022	-88.756806	39.01	3535	7.98	2369.19	0.67
	OV004	1100	28.676717	-88.362856	6.90	3428	7.98	2451.70	0.72
	OV009	1100	28.740994	-88.168814	19.18	3598	8.01	2486.86	0.69
	OV013	1100	28.801976	-88.391856	7.48	3323	7.95	2415.92	0.73
	OV014	1100	28.770928	-88.392046	4.41	3556	7.99	2399.98	0.67
	Average (SE)	**1084 (16)**			**18.14 (6.72)**	**3488 (50)**	**7.98 (0.01)**	**2424.73 (20.45)**	**0.70 (0.01)**
p value[b]					0.09	**<0.01**	**<0.01**	**<0.01**	0.16

[a]: Distance from the wellhead; [b]: *p* values from the Student's t-test between the oil plume and the non-plume (control) samples.

Table S3. All sequences present in non-plume samples but absent in plume samples.

Gene ID	Gene/enzyme	Gene category	Sub-category	Organism
50236433	*B_lactamase*	Antibiotic resistance	Beta-lactamases	*Ralstonia pickettii*
241463126	*Tet*	Antibiotic resistance	Other	*Allochromatium vinosum* DSM 180
170140550	*MFS_antibiotic*	Antibiotic resistance	Transporter	*Burkholderia graminis* C4D1M
148259577	*MFS_antibiotic*	Antibiotic resistance	Transporter	*Acidiphilium cryptum* JF-5
133738122	*MFS_antibiotic*	Antibiotic resistance	Transporter	*Herminiimonas arsenicoxydans*
219869375	*SMR_antibiotics*	Antibiotic resistance	Transporter	*Desulfovibrio desulfuricans* subsp. desulfuricans str. ATCC 27774
254449746	*AceA*	Carbon cycling	Carbon degradation (Others)	*Octadecabacter antarcticus* 238
145569035	*AceB*	Carbon cycling	Carbon degradation (Others)	*Pseudomonas stutzeri* A1501
262031871	*acetylglucosaminidase*	Carbon cycling	Carbon degradation (Chitin)	*Vibrio cholerae* CT 5369-93
170742400	*amyA*	Carbon cycling	Carbon degradation (Starch)	*Methylobacterium* sp. 4-46
145250957	*endoglucanase*	Carbon cycling	Carbon degradation (Cellulose)	*Aspergillus niger*
242222987	*glx*	Carbon cycling	Carbon degradation (Lignin)	*Postia placenta* Mad-698-R
119525483	*nplT*	Carbon cycling	Carbon degradation (Starch)	*Thermofilum pendens* Hrk 5
218564086	*phenol_oxidase*	Carbon cycling	Carbon degradation (Lignin)	Uncultured fungus
19171198	*phenol_oxidase*	Carbon cycling	Carbon degradation (Lignin)	*Gaeumannomyces graminis* var. tritici
10184710	*vdh*	Carbon cycling	Carbon degradation (Aromatics)	*Pseudomonas* sp. HR199
116098389	*xylanase*	Carbon cycling	Carbon degradation (Hemicellulose)	*Lactobacillus brevis* ATCC 367
89351430	*pcc*	Carbon cycling	Carbon fixation	*Xanthobacter autotrophicus* Py2
221082971	*pcc*	Carbon cycling	Carbon fixation	*Variovorax paradoxus* S110
196243766	*ArsC*	Metal Resistance	As	*Cyanothece* sp. PCC 8802
148654633	*CadA*	Metal Resistance	Cd	*Roseiflexus* sp. RS-1
187725787	*CadA*	Metal Resistance	Cd	*Ralstonia pickettii* 12J
90336192	*CadA*	Metal Resistance	Cd	*Aurantimonas* sp. SI85-9A1
134093787	*czcA*	Metal Resistance	Cd, Co, Zn	*Herminiimonas arsenicoxydans*
67985836	*czcD*	Metal Resistance	Cd, Co, Zn	*Kineococcus radiotolerans* SRS30216
14025262	*ChrA*	Metal Resistance	Ce	*Mesorhizobium loti* MAFF303099
188029622	*CopA*	Metal Resistance	Cu	*Erwinia tasmaniensis* Et1/99
77385595	*CopA*	Metal Resistance	Cu	*Pseudomonas fluorescens* PfO-1
193213306	*CopA*	Metal Resistance	Cu	*Chlorobaculum parvum* NCIB 8327
52080381	*TehB*	Metal Resistance	Te	*Bacillus licheniformis* ATCC 14580
254430241	*arsM*	Metal Resistance	As	*Cyanobium* sp. PCC 7001
194359405	*merF*	Metal Resistance	As	*Pseudomonas aeruginosa*
113941017	*ureC*	Nitrogen	Ammonification	*Herpetosiphon aurantiacus* ATCC 23779

76056949	*narG*	Nitrogen	Denitrification	Uncultured bacterium
119391563	*narG*	Nitrogen	Denitrification	Uncultured bacterium
73763054	*nirK*	Nitrogen	Denitrification	Uncultured bacterium
32895106	*nirS*	Nitrogen	Denitrification	Uncultured bacterium
74038362	*nirS*	Nitrogen	Denitrification	Uncultured bacterium
77378473	*nirS*	Nitrogen	Denitrification	Uncultured bacterium
24421301	*nirS*	Nitrogen	Denitrification	Uncultured bacterium
116698891	*nrfA*	Nitrogen	Dissimilatory N reduction	*Syntrophobacter fumaroxidans* MPOB
197334514	*nrfA*	Nitrogen	Dissimilatory N reduction	*Vibrio fischeri* MJ11
62149162	*nifH*	Nitrogen	Nitrogen fixation	Uncultured bacterium
62529086	*nifH*	Nitrogen	Nitrogen fixation	Gamma Proteobacterium BAL286
121716444	*BpH*	Organic Remediation	Aromatics (Aromatic carboxylic acid)	*Aspergillus clavatus* NRRL 1
1685013	*bphA*	Organic Remediation	Aromatics (Polycyclic aromatics)	
209535597	*Catechol*	Organic Remediation	Aromatics (Other aromatics)	*Rhizobium leguminosarum* bv. trifolii WSM2304
78221781	*Catechol*	Organic Remediation	Aromatics (Other aromatics)	*Geobacter metallireducens* GS-15
133948800	*Catechol*	Organic Remediation	Aromatics (Other aromatics)	Uncultured bacterium
170740754	*hmgA*	Organic Remediation	Aromatics (Aromatic carboxylic acid)	*Methylobacterium* sp. 4-46
126705740	*hmgC*	Organic Remediation	Aromatics (Aromatic carboxylic acid)	*Rhodobacterales bacterium* HTCC2150
148251653	*mdlA*	Organic Remediation	Aromatics (Aromatic carboxylic acid)	*Bradyrhizobium* sp. BTAi1
111611144	*mdlC*	Organic Remediation	Aromatics (Aromatic carboxylic acid)	*Verminephrobacter eiseniae* EF01-2
169242378	*nagG*	Organic Remediation	Aromatics (Aromatic carboxylic acid)	*Mycobacterium abscessus*
160899137	*nagI*	Organic Remediation	Aromatics (Aromatic carboxylic acid)	*Delftia acidovorans* SPH-1
170735912	*nitA*	Organic Remediation	Aromatics (Other aromatics)	*Burkholderia cenocepacia* MC0-3
182678157	*nmoA*	Organic Remediation	Aromatics (Nitroaromatics)	*Beijerinckia indica* subsp. indica ATCC 9039
126436248	*pimF*	Organic Remediation	Aromatics (Aromatic carboxylic acid)	*Mycobacterium* sp. JLS
163259205	*pimF*	Organic Remediation	Aromatics (Aromatic carboxylic acid)	*Bordetella petrii*
50122001	*pimF*	Organic Remediation	Aromatics (Aromatic carboxylic acid)	*Pectobacterium atrosepticum* SCRI1043
145216180	*tfdA*	Organic Remediation	Aromatics (Chlorinated aromatics)	*Mycobacterium gilvum* PYR-GCK
113524845	*tfdA*	Organic Remediation	Aromatics (Chlorinated aromatics)	*Ralstonia eutropha* H16
126236581	*xylJ*	Organic Remediation	Aromatics (BTEX and related aromatics)	*Mycobacterium* sp. JLS
148978109	*atzB*	Organic Remediation	Herbicides related compound	*Vibrionales bacterium* SWAT-3
28851229	*atzB*	Organic Remediation	Herbicides related compound	*Pseudomonas syringae* pv. tomato str. DC3000
186472614	*phn*	Organic Remediation	Herbicides related compound	*Burkholderia phymatum* STM815
170061235	*cpnA*	Organic Remediation	Other Hydrocarbons	*Culex quinquefasciatus*
71556174	*nitro*	Organic Remediation	Others	*Pseudomonas syringae* pv. phaseolicola 1448A

66046416	*linB*	Organic Remediation	Pesticides related compound	*Pseudomonas syringae* pv. syringae B728a
81251090	*gyrB*	other category	Phylogenetic marker	*Nitrosovibrio* sp. FJI82
196257859	*gyrB*	other category	Phylogenetic marker	*Cyanothece* sp. PCC 7822
163743742	*gyrB*	other category	Phylogenetic marker	*Phaeobacter gallaeciensis* 2.10
198251727	*gyrB*	other category	Phylogenetic marker	*Octadecabacter antarcticus* 307
28270519	*ppk*	Phosphorus	Phosphorus utilization	*Lactobacillus plantarum* WCFS1
198253442	*ppx*	Phosphorus	Phosphorus utilization	*Octadecabacter* antarcticus 307
223986236	*bglH*	Stress	Glucose limitation	*Holdemania filiformis* DSM 12042
254475115	*grpE*	Stress	Heat shock	*Ruegeria* sp. R11
162285320	*hrcA*	Stress	Heat shock	*Hoeflea phototrophica* DFL-43
89094999	*glnA*	Stress	Nitrogen limitation	*Oceanospirillum* sp. MED92
256357221	*glnA*	Stress	Nitrogen limitation	*Catenulispora acidiphila* DSM 44928
238794434	*proV*	Stress	Osmotic stress	*Yersinia intermedia* ATCC 29909
163858141	*narH*	Stress	Oxygen limitation	*Bordetella petrii* DSM 12804
259169062	*narH*	Stress	Oxygen limitation	*Lactobacillus antri* DSM 16041
124266896	*narI*	Stress	Oxygen limitation	*Methylibium petroleiphilum* PM1
126618476	*ahpC*	Stress	Oxygen stress	*Cyanothece* sp. CCY0110
88799477	*ahpF*	Stress	Oxygen stress	*Reinekea* sp. MED297
91796201	*fnr*	Stress	Oxygen stress	*Chromohalobacter salexigens* DSM 3043
120324818	*fnr*	Stress	Oxygen stress	*Marinobacter aquaeolei* VT8
169757354	*oxyR*	Stress	Oxygen stress	*Pseudomonas putida* W619
260665486	*phoB*	Stress	Phosphate limitation	*Lactobacillus jensenii* SJ-7A-US
225569383	*phoB*	Stress	Phosphate limitation	*Clostridium hylemonae* DSM 15053
239906660	*pstB*	Stress	Phosphate limitation	*Desulfovibrio magneticus* RS-1
184192106	*pstS*	Stress	Phosphate limitation	*Burkholderia phymatum* STM815
149926943	*sigma_24*	Stress	Sigma factors	*Limnobacter* sp. MED105
229474868	*sigma_24*	Stress	Sigma factors	*Planctomyces limnophilus* DSM 3776
239977836	*sigma_24*	Stress	Sigma factors	*Streptomyces albus* J1074
221737130	*sigma_32*	Stress	Sigma factors	*Agrobacterium vitis* S4
261855188	*sigma_32*	Stress	Sigma factors	*Halothiobacillus neapolitanus* c2
157804763	*AprA*	Sulfur	Other	*Thiothrix* sp. 12730
109452455	*dsrA*	Sulfur	Sulfite reductase	Uncultured sulfate-reducing bacterium
28974732	*dsrB*	Sulfur	Sulfite reductase	Uncultured bacterium
229512091	*iro*	virulence		*Vibrio cholerae* B33
86751345	*iro*	virulence		*Rhodopseudomonas palustris* HaA2
163802506	*pilin*	virulence		*Vibrio* sp. AND4
283102101	*srt*	virulence		*Bifidobacterium dentium* Bd1

1

C. SUPPORTING FIGURES

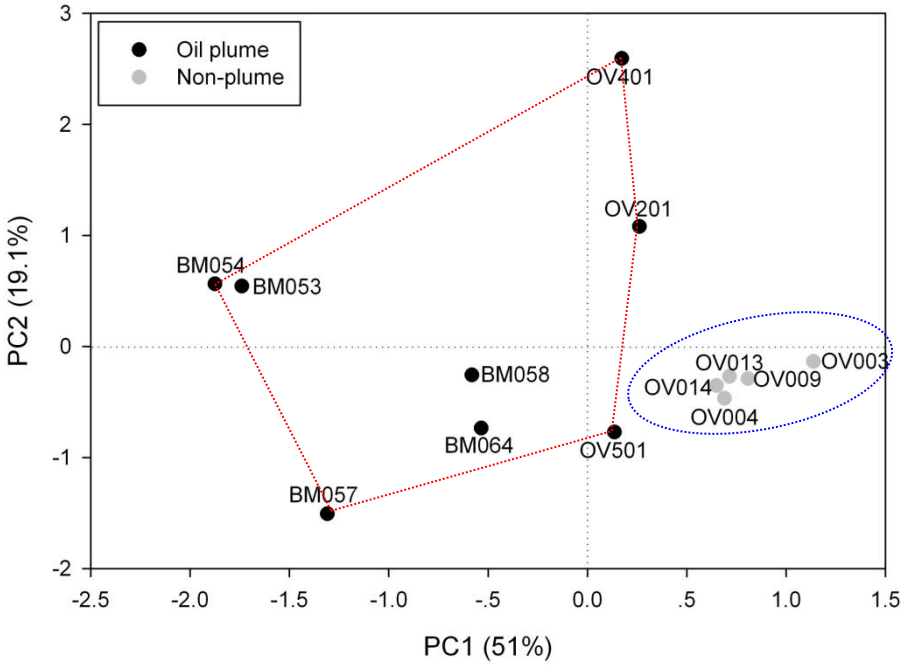

Fig S1. Ordination plot produced from principal-component analysis (PCA) of geochemical data for all the monitoring samples. The overall geochemical pattern was considerably different between the oil plume and non-plume samples. The geochemical parameters used for PCA include: temperature, DO concentration, fluorometer detection of oil, small particle concentrations, Fe, nitrate, phosphate, benzene, toluene, naphthalene, ethylbenzene, isopropylbenzene, n-propylbenzene, 1,3,5-trimethylbenzene, tert-butylbenzene, 1,2,4-trimethylbenzene, sec-butylbenzene, p-isopropyltoluene, n-butylbenzene, total xylenes, total volatile HC, and total petroleum hydrocarbons - extractable (DRO).

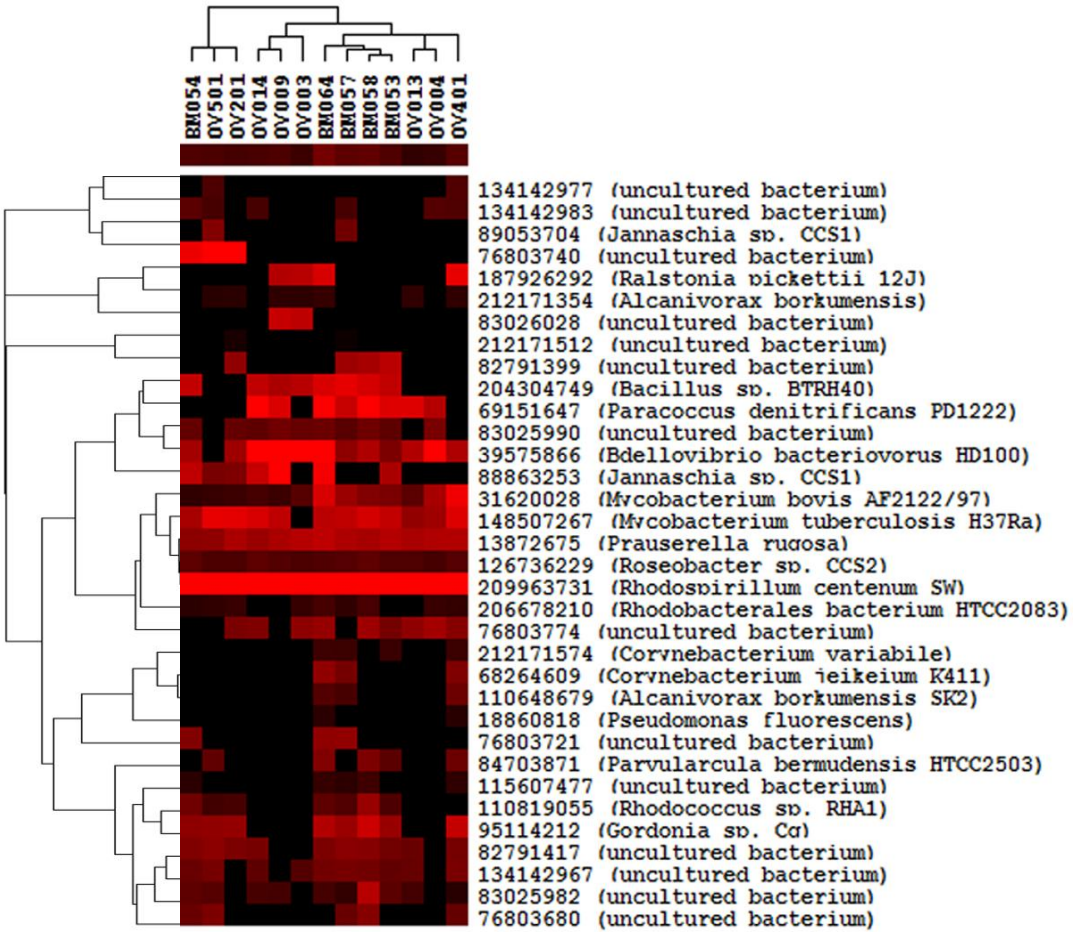

Fig. S2 Hierarchical cluster analysis of *alkB* gene, encoding alkane 1-monooxygenase. All genes were used for cluster analysis. Results were generated in CLUSTER and visualized using TREEVIEW. Red indicates signal intensities above background while black indicates signal intensities below background. Brighter red coloring indicates higher signal intensities.

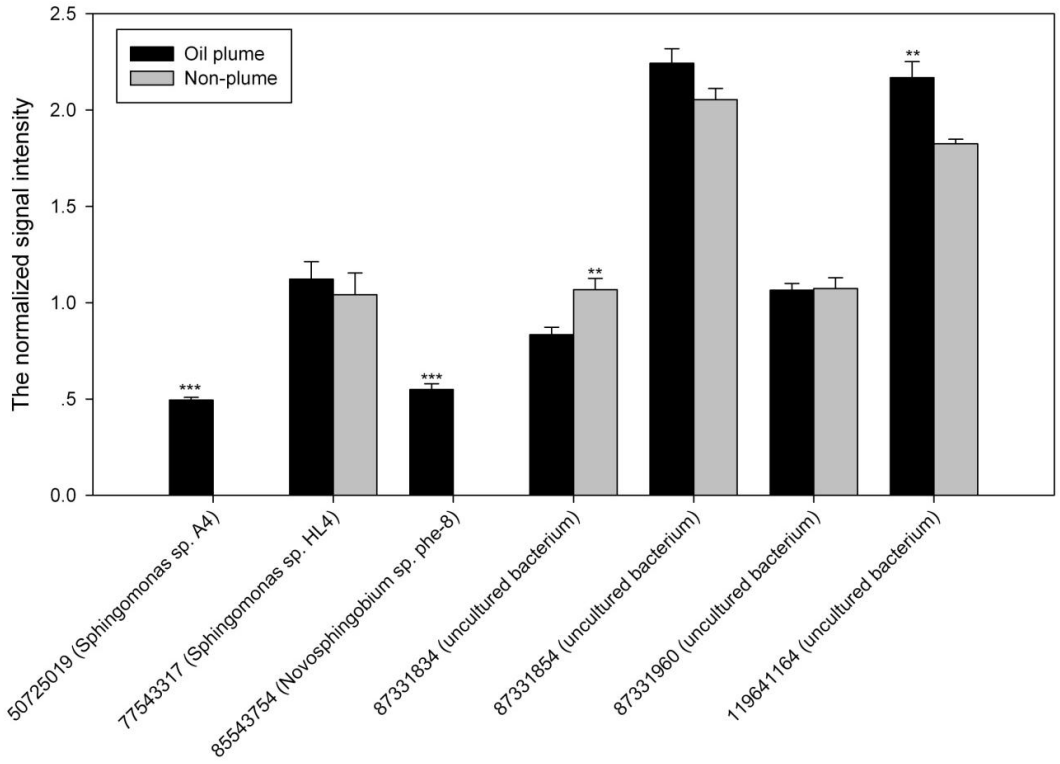

Fig. S3 The normalized signal intensity of the *arhA* (PAH dioxygenase) genes. The signal intensity for each sequence was the average of the total signal intensity from all the replicates. Gene number is the protein ID number for each gene as listed in the GenBank database. All data are presented as mean ± SE. ***p<0.01, **p<0.05, *p<0.1.

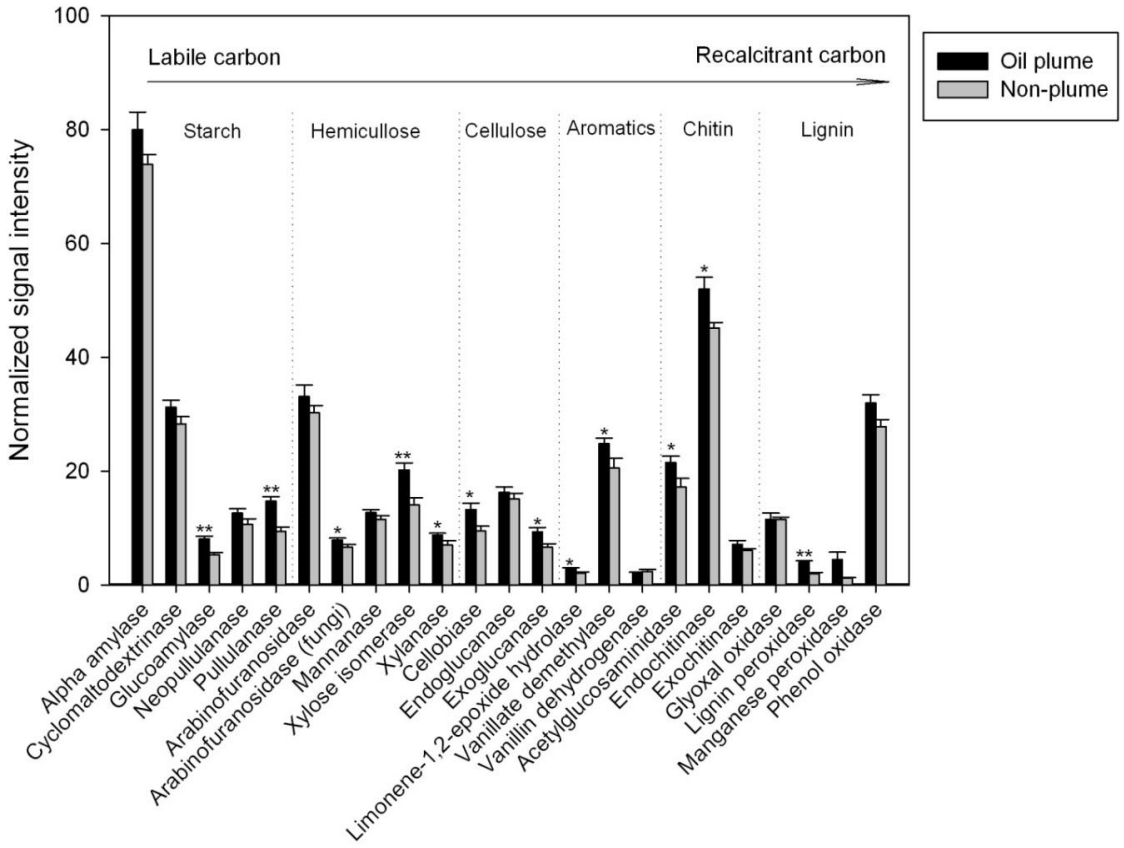

Fig. S4 The normalized signal intensity of the detected key genes involved in carbon degradation. The signal intensity for each function gene was the average of the total signal intensity from all the replicates. All data are presented as mean ± SE. ***p*<0.01, **p*<0.05. Many genes involved in carbon degradation (e.g., starch, cellulose, hemicelluloses, lignin, chitin, and aromatics) showed greater abundance in plume, including those encoding glucoamylase, and pullulanase (*pulA*) for starch, xylose isomerase (*xylA*), xylanase (*xynA*) and arabinofuranosidase (*ara_fungi*) for hemicellulose, cellobiase and exoglucanse for cellulose, endochitinase for chitin, lignin peroxidase and ligninase (*lip*) for lignin, limonene-1,2-epoxide hydrolase (*limEH*) and vanillate monooxygenase (*vanA*) for other C compounds (e.g., aromatics).

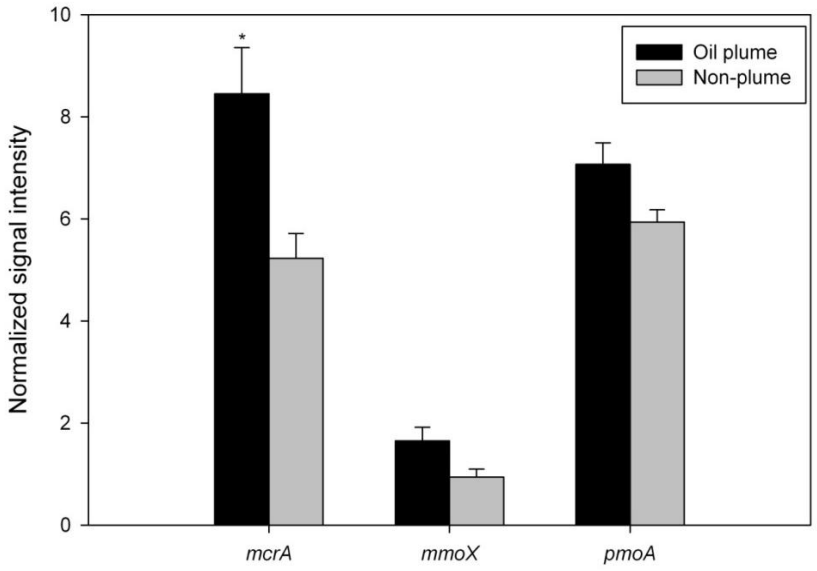

Fig. S5 The normalized signal intensity of the detected genes involved in methane metabolism. The signal intensity for each function gene was the average of the total signal intensity from all the replicates. All data are presented as mean ± SE. *p<0.05. *mcrA*, alpha subunit of methyl coenzyme M reductase; *mmoX*, particulate methane monooxygenase; *pmoA*, methane monooxygenase.

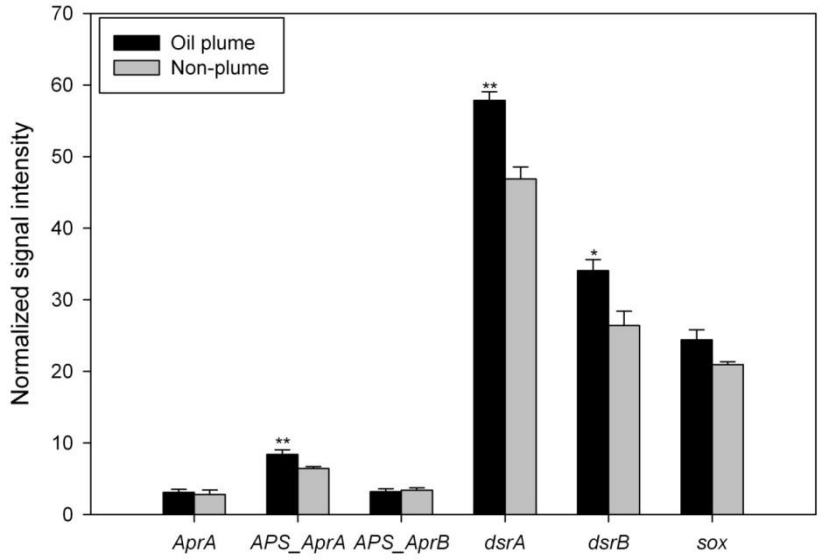

Fig. S6 The normalized signal intensity of the detected genes involved in sulfur cycling. The signal intensity for each function gene was the average of the total signal intensity from all the replicates. All data are presented as mean ± SE. $^{**}p<0.01$, $^{*}p<0.05$. *APS_AprA* encoding dissimilatory adenosine-5'-phosphosulfate (APS) reductase and *dsrA/B* encoding dissimilatory sulfite reductase responsible for sulfur reduction were significantly increased in oil plume.

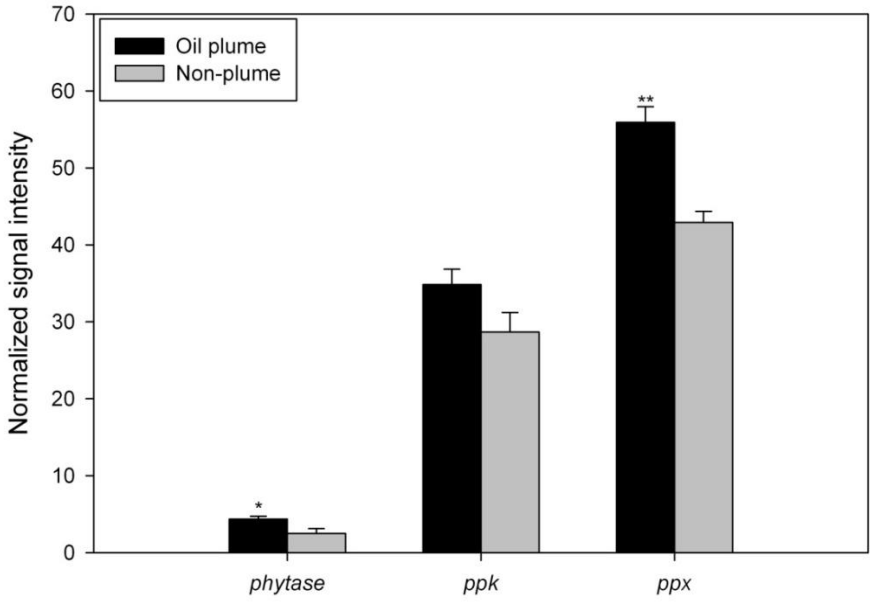

Fig. S7 The normalized signal intensity of the detected genes involved in phosphorus cycling. The signal intensity for each function gene was the average of the total signal intensity from all the replicates. All data are presented as mean ± SE. $^{**}p<0.01$, $^{*}p<0.05$. Phytase, responsible for phytate degradation; *ppk*, ATP-polyP phosphotransferase responsible for polyP biosynthesis; *ppx*, encoding exopolyphosphatase for inorganic polyphosphate degradation. The abundance of phytase and *ppx* genes was significantly increased suggesting an increased release of phosphorus from inorganic polyphosphate and phytate degradation may occur in the plume.

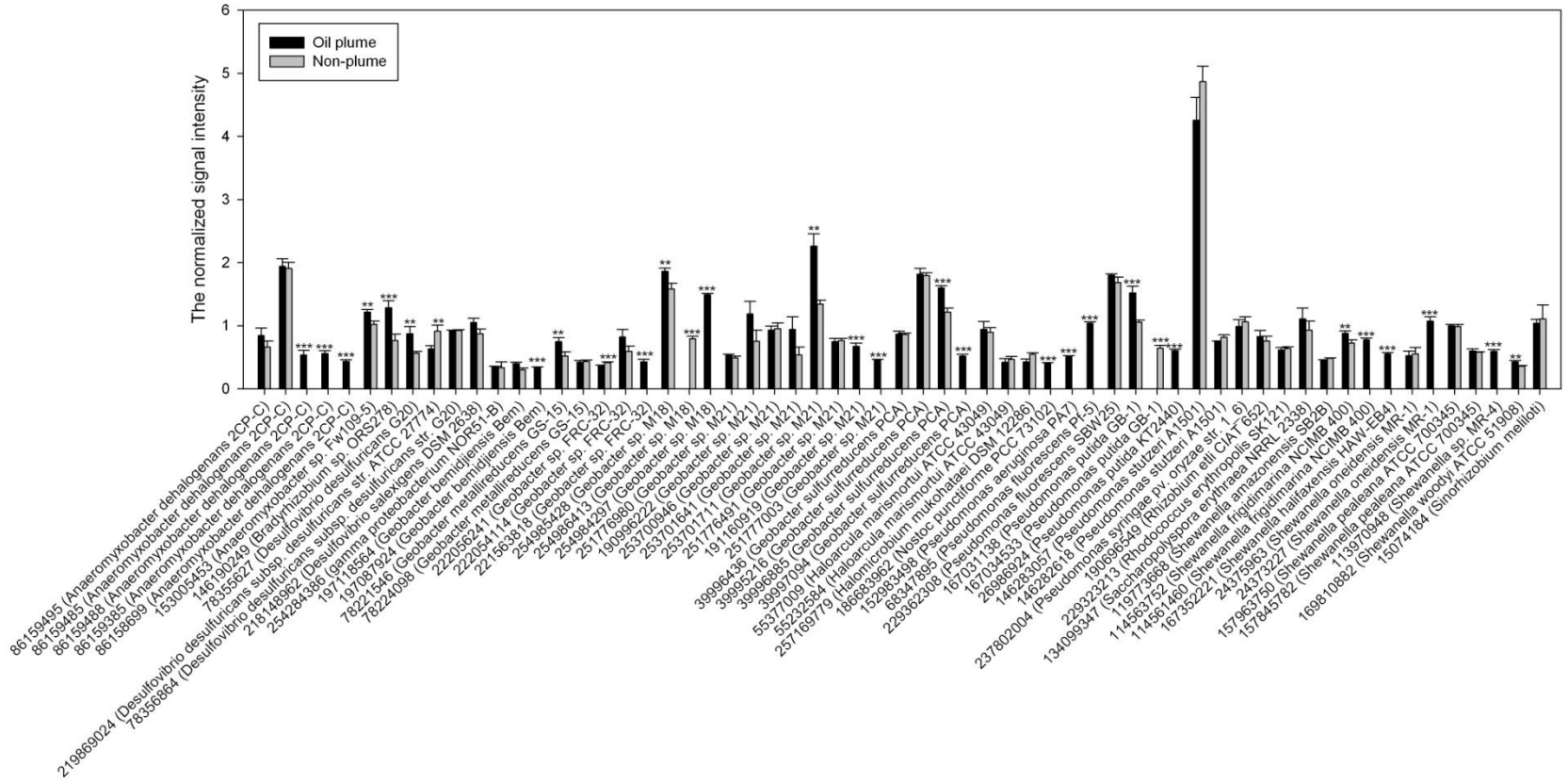

Fig. S8 The normalized signal intensity of the cytochrome genes. The signal intensity for each sequence was the average of the total signal intensity from all the replicates. Gene number is the protein ID number for each gene as listed in the GenBank database. All data are presented as mean ± SE. ***p<0.01, **p<0.05, *p<0.1.

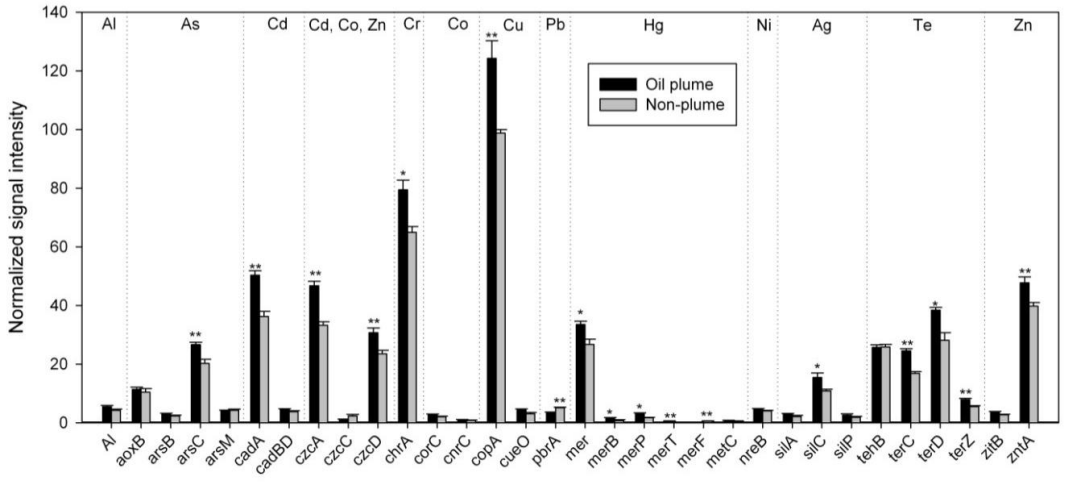

Fig. S9 The normalized signal intensity of the detected key genes involved in metal resistance. The signal intensity for each function gene was the average of the total signal intensity from all the replicates. All data are presented as mean ± SE. **$p<0.01$, *$p<0.05$.

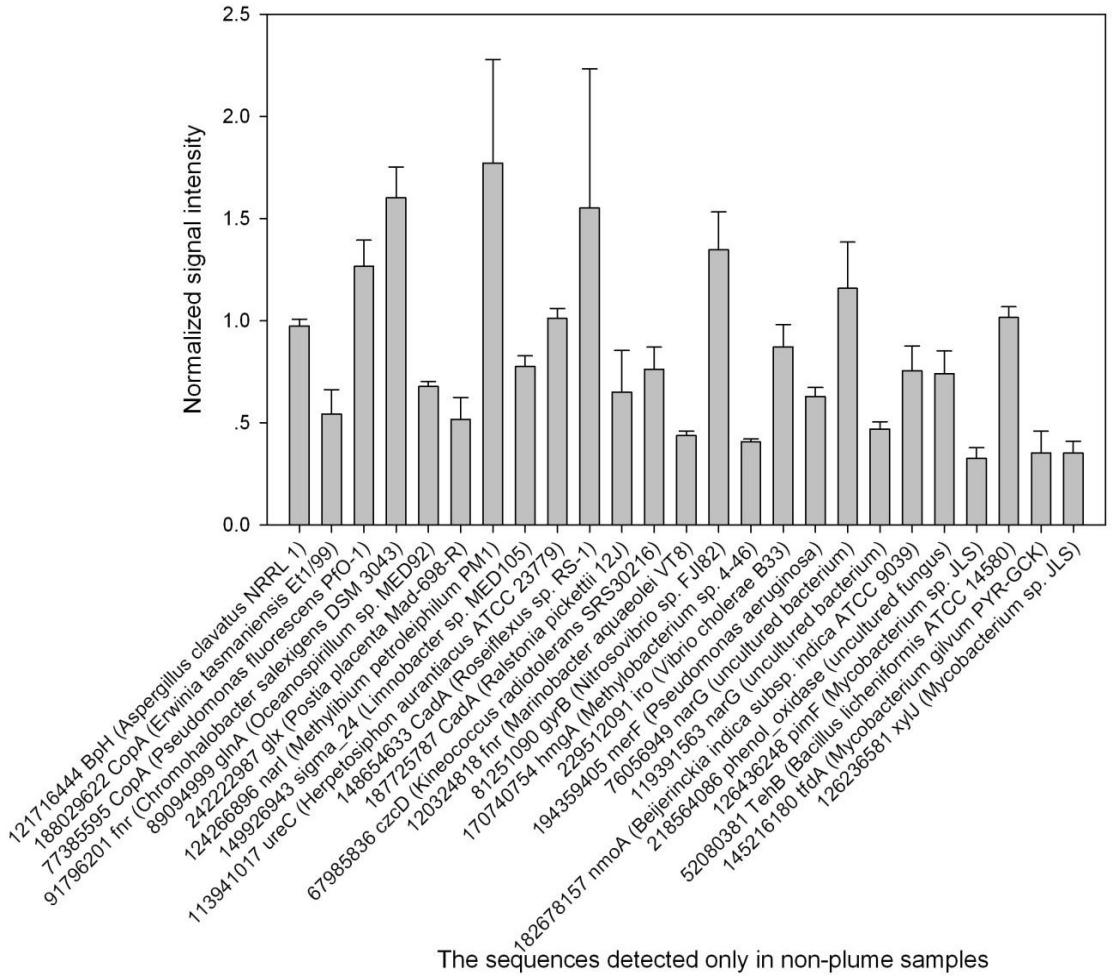

Fig. S10 The sequences present in five replicates of control (non-plume) samples but absent in oil plume samples. The signal intensity for each sequence was the average of the total signal intensity from all the replicates. All data are presented as mean ± SE. Gene number is the protein ID number for each gene as listed in the GenBank database.

D. SUPPLEMENTARY REFERENCES

Anderson M. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecol* **26:** 32-46.

APHA. (2005). Standard Methods of the Examination of Water and Wastewater. *American Public Health Association*.

Carlson RM, Cabrera RI, Paul JL, Quick J, Evans RY. (1990). Rapid direct determination of ammonium and nitrate in soil and plant tissue extracts. *Comm Soil Sci Plant Anal* **21:** 1519 - 1529.

Clarke KR. (1993). Non-parametric multivariate analyses of changes in community structure. *Austral Ecol* **18:** 117-143.

de Hoon MJ, Imoto S, Nolan J, Miyano S. (2004). Open source clustering software. *Bioinformatics* **20:** 1453-1454.

Francisco DE, Mah RA, Rabin AC. (1973). Acridine orange-epifluorescence technique for counting bacteria in natural waters. *Trans Am Microsc Soc* **92:** 416-421.

Hazen TC, Dubinsky EA, DeSantis TZ, Andersen GL, Piceno YM, Singh N *et al.* (2010). Deep-sea oil plume enriches indigenous oil-degrading bacteria. *Science* **330**: 204-208.

He Z, Deng Y, Van Nostrand JD, Tu Q, Xu M, Hemme CL *et al.* (2010a). GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. *ISME J* **4:** 1167-1179.

He Z, Xu M, Deng Y, Kang S, Kellogg L, Wu L *et al.* (2010b). Metagenomic analysis reveals a marked divergence in the structure of belowground microbial communities at elevated CO_2. *Ecol Lett* **13:** 564-575.

He ZL, Gentry TJ, Schadt CW, Wu LY, Liebich J, Chong SC *et al.* (2007). GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *ISME J* **1:** 67-77.

Li Z, Kepkay P, Lee K, King T, Boufadel MC, Venosa AD. (2007). Effects of chemical dispersants and mineral fines on crude oil dispersion in a wave tank under breaking waves. *Mar Pollut Bull* **54:** 983-993.

Liang Y, He Z, Wu L, Deng Y, Li G, Zhou J. (2010). Development of a common oligonucleotide reference standard for microarray data normalization and comparison across different microbial communities. *Appl Environ Microbiol* **76:** 1088-1094.

Luo Y, Hui D, Zhang D. (2006). Elevated CO_2 stimulates net accumulations of carbon and nitrogen in land ecosystems: a meta-analysis. *Ecology* **87:** 53-63.

Miller DN, Bryant JE, Madsen EL, Ghiorse WC. (1999). Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. *Appl Environ Microbiol* **65:** 4715-4724.

Wu L, Liu X, Schadt CW, Zhou J. (2006). Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. *Appl Environ Microbiol* **72:** 4931-4941.

Zhou J, Kang S, Schadt CW, Garten CT. (2008). Spatial scaling of functional gene diversity across various microbial taxa. *Proc Natl Acad Sci USA* **105:** 7768-7773.

R Development Core Team. (2006). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. *http://www.R-project.org*.