Annals of Microbiology (2019) 69:495–513
https://doi.org/10.1007/s13213-019-1438-3

ORIGINAL ARTICLE

Comparative metagenomics and functional profiling of crude oil-polluted soils in Bodo West Community, Ogoni, with other sites of varying pollution history

Chioma Blaise Chikere[1] · Ijeoma Jessie Mordi[1] · Blaise Ositadinma Chikere[2] · Ramganesh Selvarajan[3] · Tom Omotayo Ashafa[3] · Chinedu Christopher Obieze[4]

Received: 6 October 2018 / Accepted: 7 January 2019 / Published online: 13 February 2019

Abstract

The impact of long-term crude oil pollution on soil microbial community structure in Bodo West Community, Ogoniland, Nigeria, was investigated to determine the amenability of the soil to microbial mediated remediation. Crude oil-polluted and pristine soil samples were collected approximately from 0 to 30 cm depth for both chemical and microbiological analyses. Total petroleum hydrocarbons (TPH) and polycyclic aromatic hydrocarbons (PAH) were determined using gas chromatograph–mass spectrophotometer (GC-MS). The soil microbiome was determined using the Illumina MiSeq platform. Results from this study were then compared with publicly available data from other oil-polluted sites. Taxonomic biomarkers and pathways associated with oil-polluted soils were detected using bioinformatics pipelines. TPH in the polluted and pristine soils were 7591 mg/kg and 199.70 mg/kg respectively, while the values of PAHs were significantly higher ($p < 0.05$) in the oil-polluted soil. Predictive functional and biomarker analysis demonstrated that microbes detected in the oil-polluted environment were involved in different metabolic pathways for degradation of a broad set of xenobiotic aromatic compounds. Established hydrocarbon degraders belonging to the families *Alcanivoracaceae* and *Oceanospirillaceae* were mostly detected in the oil-polluted soils. *Sneathiella*, *Parvibaculum*, *Sphingobium*, and *Oceanicaulis* were among biomarker taxa. The bacterial families *Acidithiobacillaceae* and *Desulfobacteraceae* were differentially more abundant in Bodo West spill site than any other site used for comparison. Furthermore, differentially represented species in our study site and other oil-polluted sites ranged from 21 to 42 bacterial families. The findings from this study revealed the bacterial community had a strong dependence on hydrocarbons and that acid-tolerant bacterial families can as well contribute significantly to biodegradation in the site and other polluted sites in Ogoniland usually known to have an acidic pH. Further research on Bodo West spill site will reveal the novel enzymes and pathways for enhanced microbial mediated eco-restoration.

Keywords Crude oil pollution · Metagenomics · Biodegradation · Soil microbiome · Niger Delta

Introduction

The economic benefits of crude oil cannot be over highlighted; however, the accidental release of oil and related products leads to an extensive pollution of soil and aquifers, stimulating the need for an upgrade in bioremediation processes (Joshi et al. 2014). Soil microorganisms play dynamic roles in the ecosystem and they are responsible for most biological transformations and drive the nutrient cycles to facilitate the subsequent establishment of plant communities (Schulz et al. 2013). Furthermore, the diversity of the microbial community in soil is closely related to the function and structure of its surrounding ecosystem (Ataikiru et al. 2017). For instance, increased pollution of hydrocarbon in the soil will inevitably

✉ Chioma Blaise Chikere
chioma.chikere@uniport.edu.ng

[1] Department of Microbiology, University of Port Harcourt, East-West Road, Choba. P.M.B. 5323, Port Harcourt, Nigeria

[2] Environmental Studies Unit, Shell Petroleum Development Company (SPDC) Port Harcourt, P. O. Box 263, Rumuobiakani, Port Harcourt, Nigeria

[3] Department of Plant Sciences, University of the Free State, Qwaqwa Campus, Phuthaditjhaba, South Africa

[4] Africa Centre of Excellence in Oilfield Chemicals Research, University of Port Harcourt, East-West Road, Choba. P.M.B. 2, Port Harcourt, Nigeria

alter the microbial communities and their nutrient fluxes (Reid et al. 2018). Such microorganisms play an important role in determining the rate of hydrocarbon degradation and the prospects for the recovery of novel enzymes for industrial applications.

Numerous researches have been carried out to advance the application of soil microorganisms in bioremediation as a number of novel enzymes and pathways have been developed, likewise the establishment of some bacterial species as key to hydrocarbon degradation (Gibson et al. 2002; Hassanshahian et al. 2014; Nwinyi et al. 2016; Parthipan et al. 2017). However, the complexity of crude oil does not allow for a successful hydrocarbon degradation using single organism, pathway, or enzyme. In recent years, it has been well documented that crude oil contamination is better degraded by a consortia of microorganisms, interacting among themselves and all contributing in one way or the other to their overall survival under real-life and highly variable in situ conditions (Wang et al. 2016; Chikere et al. 2017; Kumari et al. 2018). Understanding the complex relationships among microorganisms in oil-polluted soils, their interaction with pollutants, response to chemical stress and the geophysical properties of the polluted site is of utmost importance if progress is to be made in advancing engineered bioremediation.

Microbial species that are able to degrade hydrocarbons are not easy to isolate using the normal nutrient media and laboratory conditions (Sibanda et al. 2017). Although culture-dependent methods generally recover approximately 0.3% of the microbial population from soil environments, they are still a critical component of bioremediation development and research (Stefani et al. 2015), whereas culture-independent analysis has become the most widely used method to determine the taxonomic fingerprints of microbial populations in different environments (Paul et al. 2016; Selvarajan et al. 2018a, b). Conventional molecular approaches including cloning of 16S rRNA genes, denaturing gradient gel electrophoresis (DGGE), terminal restriction fragment length polymorphism (T-RFLP), and fluorescence in situ hybridization (FISH) lack sufficient depth to cover the comprehensive information on various microbial communities (Lu et al. 2012). However, the recent next-generation sequencing (NGS) technologies are promising techniques for exploring soil microbiomes and their metabolic potentials in situ. A great deal of research has been performed using this technique and their findings have consistently shown that soil microbiomes and their functional potential are a direct reflection of the prevailing geophysical and chemical properties of the sites studied (Caporaso et al. 2011a; Bao et al. 2017; Mukherjee et al. 2017; Feng et al. 2018; Abia et al. 2018).

Within the last two decades, the use of molecular techniques has led to a significant improvement in our knowledge of microbial diversity and functional profiling in different complex environments (Selvarajan et al. 2014). Gałązka et al. (2018) used quantitative molecular approaches to determine the diversity of bacterial microbiome and the functional profile of a long-term hydrocarbon-impacted soil. While Bao et al. (2017) and Mukherjee et al. (2017) determined microbial responses to hydrocarbon pollution with respect to diversity and function using predictive metagenomic methods. Ogoniland in the Niger Delta region has a history of over five decades of colossal oil pollution in both coastal and terrestrial settings resulting in massive destruction of vulnerable ecosystems like mangroves and wetlands (Lindén and Pålsson 2013). Hence, it becomes necessary to characterize microbial communities and its function in polluted environments, especially when the pollutant is as complex as crude oil. The aim of this research was, therefore, to investigate the taxonomic and functional profiles of the soil microbiomes of long-term oil-polluted and pristine soils in Bodo West Community, Ogoniland, Nigeria. In addition, we compared our samples with publicly available datasets with varying history of oil pollution to determine the differential nature of soil microbial community structures and their functional responses to the pollutant and to compare differences in biomarker taxa and metabolic pathways in crude oil-polluted and pristine environments. To the best of our knowledge, this is the first time quantitative molecular approaches and predictive metagenomic study will be applied to determine the microbial and functional profiles of Bodo West, Ogoniland oil pollution.

Materials and methods

Site description and sample collection

Crude oil-polluted and pristine soil samples were collected from Bodo West community in Gokana Local Government Area of Rivers State, Nigeria (Fig. 1). Bodo Community is host to Shell Petroleum Development Company's oil operations and is installed with 24 and 28-in. Trans-Niger pipelines. It is estimated that the 20-km^2 network of creeks and inlets in Bodo have been devastated by crude oil spills particularly as a result of sabotage, leaking pipelines (The Guardian 2018), and more recently artisinal refining. Composite samples from the crude oil-polluted and pristine soils were collected from 0 to 30 cm depths using a soil auger (GPS coordinates: latitude 4.6090150 longitude 7.2242150 E; latitude 4.6163330 longitude 7.2254576). Samples were immediately kept at 4 °C in a cooler box and transported to the laboratory at the University of Port Harcourt, Nigeria, for further analyses.

One set of soil samples was used for the analysis of physicochemical parameters such as pH, electrical conductivity, moisture content, phosphate (PO_4), potassium (K), nitrate (NO_3) total organic carbon (TOC), and heavy metal analysis such as zinc (Zn), nickel (Ni), and lead (Pb). All the

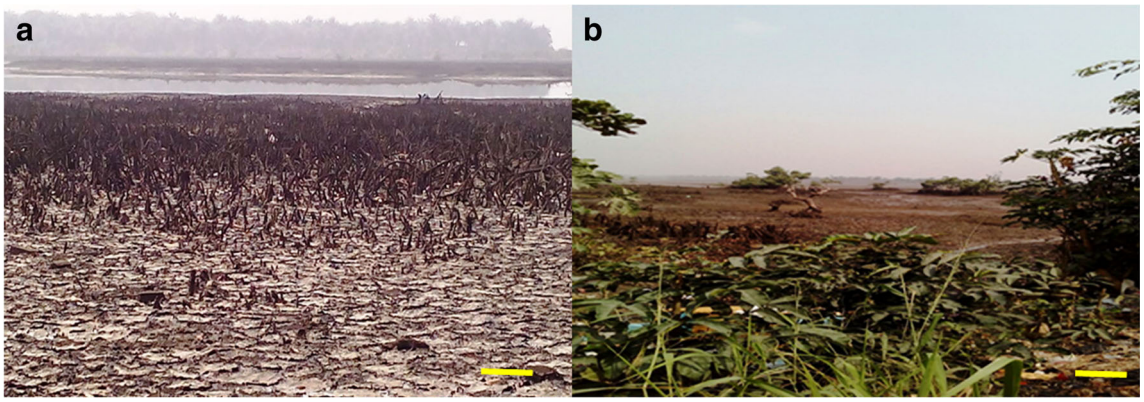


Fig. 1 Polluted soil (**a**) and pristine soil (**b**) samples collected from Bodo West community in Gokana Local Government Area of Rivers State, Nigeria. Scale bar = 0.1 m

physicochemical variables were determined according to methods of APHA 4500 and ASTM D1691 (APHA 2012), respectively, apart from heavy metals which were determined using inductively coupled plasma optical emission spectrometry (ICP-OES) (PerkinElmer Optima 5300 DV). The second set of samples was used to profile the soil microbiome of both the polluted and pristine soil.

Oil analysis of Bodo West Community, Ogoniland spill

The residual TPH and PAHs were extracted from the soil samples using the USEPA EPA 418.1 method (USEPA 1978) and quantified using a gas chromatograph–mass spectrophotometer (GC-MS) (an Agilent 6890GC, Agilent Technologies, Wilmington, USA GC equipped with 5975B MSD chemstation version D. 03.00), according to the methods of USEPA 8270. The carrier gas was helium set at a constant flow rate of 1 mL/min and at a pressure of 75 kpa. Residual TPH extracts were introduced into the GC-MS machine equipped with a narrow-bore fused silica capillary column. The column separated GC analytes were then detected with a mass spectrometer. The mass spectra of the eluted analytes were compared to authentic standards for hydrocarbons. For the identification of PAHs, the polycyclic aromatic hydrocarbons mixture with 16 common PAHs (ULTRA SCIENTIFIC PM-610) such as acenaphthene, acenaphthylene, anthracene, benz[a]anthracene, benzo[b]fluoranthene, benzo[k]fluoranthene, benzo[ghi]perylene, benzo[a]pyrene, chrysene, dibenz[a,h]anthracene, fluoranthene, fluorene, indeno[1,2,3-cd]pyrene, naphthalene, phenanthrene, and pyrene were used as our calibration standards.

Total DNA extraction and Illumina sequencing

Genomic DNA was directly extracted from the crude oil-polluted and pristine soils obtained from Bodo West community, Ogoniland, using the ZR Soil Microbe DNA Microprep (Zymo Research, CA, USA) according to the manufacturer's instructions. The eluted DNA was assessed for purity on 1.0% agarose gel and then quantified using a NanoDrop spectrophotometer (Nanodrop 2000, Thermo Scientific, Japan). Polymerase chain reaction (PCR) was performed on the extracted DNA samples using the universal bacterial primers 27F and 518R (Weisburg et al. 1991; Muyzer et al. 1993) targeting the variable region V1-V3 of the 16S ribosomal DNA. PCR reactions were prepared using 25 μL of one *Taq* 2X Master Mix, 22 μL of nuclease-free water, 1.5 μL of both forward and reverse primers at a concentration of 0.2 μM and 2 μL of extracted DNA (50–100 ng μL^{-1}). The following thermal cycler program was used for the 16 s rRNA gene amplification; initial denaturation step at 95 °C for 10 min, followed by 32 cycles of denaturation at 95 °C for 30 s; annealing at 55 °C for 30 s; extension at 72 °C for 1 min; final extension at 72 °C for 10 min. PCR amplicons were purified using a DNA Clean & Concentrator Kit (Zymo Research Corporation, USA) according to the manufacturer's instructions. The purified PCR products were then sequenced along with multiplex sample identifiers on the Illumina MiSeq platform by German Sequencing Centre, Hamburg, Germany.

Sequence data analysis

The obtained raw sequences (fastq files) were initially analyzed for PCR artifacts and low-quality reads using *ngsShoRT* (next-generation sequencing Short Reads) trimmer as described by Chen et al. (2014). Following the initial screening process, all the sequence data sets were processed using QIIME (v.1.9.0) pipeline as described by Caporaso et al. (2011b). Sequence reads containing less than 50 nucleotides, reads with more than 2% of ambiguities or 7% of homopolymers were excluded during the course of analysis. Sequences were clustered into operational taxonomic units (OTUs) at a 97% identity threshold using the UCLUST algorithm (Edgar et al. 2011). The SILVA reference database (version 132. April 10, 2018 release) was used for both open reference OTU picking and taxonomic assignment. Alpha

and beta diversity analysis was done using the Phyloseq package (McMurdie and Holmes 2013) in R statistical software. Prior to diversity analysis, the sequences were first normalized to even sampling depth. The alpha diversity indices such as Chao1, Shannon, and Simpson were measured. Nonmetric multidimensional scaling (NMDS) applying Bray–Curtis measure for estimation of beta diversity was used to compare the similarities between and within the samples. Permutation-based multivariate analysis of variance (PERMANOVA) was used to determine significant differences in samples and homogeneity of taxonomic spread among the samples using PAST v.3.11 (Hammer et al. 2001). Sequence reads for Bodo West, Ogoniland, samples were deposited in GenBank (Sequence Reads Archive) under the SRA accession number SRP133543. To compare the metagenomes of the oil-polluted soil drawn from Bodo West community, Nigeria, 12 16S rRNA meta-sequences were obtained from publicly available sequence database for both oil-polluted and pristine soils and used for this study. The sequences were retrieved from a study by Huettel et al. (2018), who analyzed the samples from oil-polluted and pristine soil at Pensacola beach following the Deep Water Horizon (DWH) oil spill in the year 2010. The other datasets used in this study were 16S rRNA sequences from Noonmati oil refinery and Barhola oilfields, both in Assam, India, by Mukherjee et al. (2017). Prior to our analysis, the samples were grouped according to history of oil pollution. A summary of the datasets used in this study is shown in Table 4 in the Appendix..

Functional profiling using 16S rRNA datasets

Prior to metagenome prediction using PICRUSt (phylogenetic investigation of communities by reconstruction of unobserved states) (Langille et al. 2013), the detected OTUs were reclassified using The GREENGENES reference database (May, 2013 release). PICRUSt and Kyoto Encyclopaedia of Genes and Genomes (KEGG) were used to obtain the relative abundance of gene families within the crude oil-polluted and pristine soils based on a constructed phylogenetic workflow of 16S rRNA marker gene sequences. The input data was first normalized by copy number by dividing each OTU by the known 16S copy number abundance prior to metagenome predictions and subsequent collapse into functional pathways. The Nearest Sequenced Taxon Index (NSTI) value was used to validate the reliability of predicted metagenomes and functional pathways.

Metabolic reconstruction of the predicted metagenomes and biomarkers detection

The metagenomes predicted using PICRUSt were reconstructed in HUMAnN2 (Abubucker et al. 2012) using KEGG pathways. The abundance and coverage of the predicted KEGG orthology (KO) was inferred by MinPath (Ye and Doak 2009) implemented in HUMAnN2. The generated output of gene abundance was then used to detect biomarker KEGG pathways. To achieve this, analysis was carried out on the data to determine differentially abundant metabolic pathways in the unique environments under study. Linear Discriminant Analysis Effect Size (LEfSe) was used to determine the presence of biomarkers by applying the Kruskal–Wallis alpha significance threshold of ≤ 0.05 and an LDA (linear discriminant analysis) score of 2.0. The GraPhlan software was then used to visualize graphically the detected metabolic biomarkers. The same method was also applied on the collapsed functional pathways to detect differentially abundant functional pathways and biomarker taxa in the crude oil-polluted soils. The dataset from other oil-polluted sites was also used for comparison with the dataset obtained from this study using STAMP (Parks et al. 2014). The comparison was carried out on a per sample basis to determine differentially abundant species by calculating the odds ratio. Only differentially abundant bacterial families with a minimum abundance of 20 were considered. Benjamini–Hochberg adjusted p value (Benjamini and Hochberg 1995) was calculated to control the false discovery rate (FDR). Bacterial families with odds ratio ≥ 1 and FDR corrected p value ≥ 0.05 were considered significantly enriched, while the significantly over-represented bacterial families satisfied an odds ratio of ≥ 2 and FDR corrected p value ≥ 0.05.

Results

Physicochemical analysis of Bodo West, Ogoniland, samples

The physicochemical variables of the polluted and pristine soil samples are presented in Table 1. Polluted soil sample had slight alkaline pH (7.8) while pristine soil had slightly acidic (6.5) pH. The concentration of total organic carbon was high in polluted soil (0.8 mg/kg) as compared to the pristine soil (0.6 mg/kg). Among the nutrients, the concentrations of nitrate and phosphate were significantly higher ($p < 0.05$) in pristine soil (10.8; 11.6 mg/kg) compared to the polluted soil, while there was no significant difference in the concentration of potassium between the sampling sites. Heavy metals like lead (Pb), zinc (Zn), and nickel (Ni) were detected in considerably higher concentrations in polluted soil than in pristine soil.

The amount of total petroleum hydrocarbon (TPH) in the polluted and pristine soil was 7591 mg/kg and 199.70 mg/kg respectively (Table 1). Further quantification revealed that the pristine soil had minimal hydrocarbon peaks while longer carbon chain hydrocarbons from C_{12} to C_{34} were detected in the oil-polluted soil. The carbon atom distribution of TPH in

Table 1 Comparison of physicochemical characteristics of the collected soil samples

Parameter	Concentration for polluted soil (IES)	Concentration of pristine soil (ICES)
pH	7.8	6.5
Electrical conductivity (μS/cm)	2020	105
Total organic carbon (%)	0.80	0.60
Total petroleum hydrocarbon (mg/kg)	7591	199.70
Polycyclic aromatic hydrocarbon (mg/kg)	26.12	2.76
Moisture content (%)	30.00	26.00
Nitrate (mg/kg)	1.30	10.80
Phosphate (mg/kg)	1.01	11.60
Zinc (mg/kg)	5.50	0.66
Potassium (mg/kg)	26.70	26.70
Lead (mg/kg)	48.70	0.09
Nickel (mg/kg)	0.07	< 0.05

polluted and pristine soil samples and the ratios of occurrence are represented in Fig. 2.

The PAH content in the polluted soil was significantly higher ($p < 0.05$) than that of the pristine soil, which confirmed the presence of a range of aromatic hydrocarbons (Table 2). The amount of PAHs having two or more single or fused aromatic rings was higher in oil-contaminated soil sample, whereas the pristine soil contained only three-member aromatic hydrocarbons such as ANTHRACENE and PHENANTHRENE. Table 2 shows all the PAHs detected in both the polluted and pristine soils.

Diversity profile of the distinct environments

A total of 50 phyla, 107 classes, 144 orders, 512 families, and 1285 bacterial genera were detected for all the samples. *Proteobacteria* was the dominant phylum in all the samples studied (Fig. 7 in the Appendix) followed by the phyla

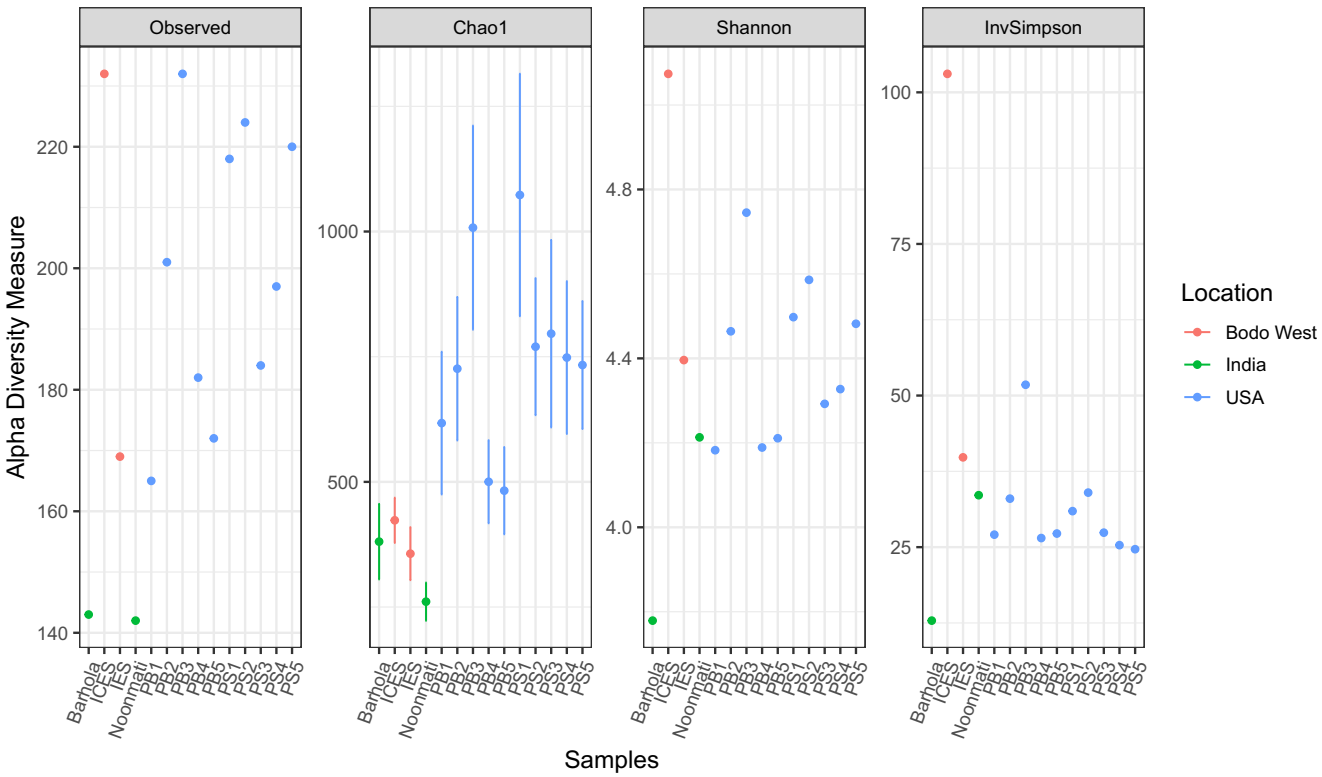


Fig. 2 Alpha diversity analysis of the samples colored according to pollution location, using diversity indices Chao1, Shannon, observed and inverse Simpson after normalizing reads from each sample.*PB1-PB5 and PS1-5 (Pensacola beach polluted and pristine soils). Noonmati (Noonmati oil refinery India). Barhola (Barhola oil refinery India). IES and ICES (Bodo West, Ogoniland, Nigeria)

Table 2 Complete profile of PAHs distribution in polluted and pristine soil samples

No. of rings	Polluted soil sample (IES)	Pristine soil sample (ICES)
2	Naphthalene	–
2	Acenaphthylene	–
2	Acenaphthene	–
3	Fluorene	–
3	Anthracene	Anthracene
3	Phenanthrene	Phenanthrene
4	Fluoranthene	–
4	Pyrene	–
4	Chrysene	–
4 or 5	Benz(a)anthracene	–
5	Benzo(b)fluoranthene	–
5	Benzo(k)fluoranthene	–
5	Benzo(a)pyrene	–
5	Dibenz(g,h,i)anthracene	–
6	Indeno(1,2,3-c,d)perylene	–
6	Benzo(ghi)perylene	–

Acidobacteria, *Actinobacteria*, *Bacteroidetes*, *Chlorobi*, *Chloroflexi*, *Cyanobacteria*, *Firmicutes*, *Planctomycetes*, and *Verrucomicrobia*. Whereas the pristine soil had an increased abundance of *Planctomycetes*, *Bacteroidetes*, *Actinobacteria*, and *Acidobacteria*, while comparing with the datasets, the crude oil-polluted soil obtained from Noonmati oil refinery had a significant high abundance of *Acidobacteria*. Statistical analysis of the taxonomic abundance for the distinct environments showed significant differences between the pristine and oil-polluted soils (Table 5 in the Appendix). Further, the pair-wise analysis of the samples confirmed that the polluted environment taxonomic profile was significantly different ($p < 0.05$) from the pristine environment (Table 5 in the Appendix). Alpha diversity analysis (Fig. 2.) of the samples that make up the distinct environments showed that the oil-polluted samples from Noonmati oil refinery and Barhola oil fields were the least diverse. NMDS plots revealed a clustering of the samples according to oil pollution exposure; the pristine soil samples clustered together on the plot, while the oil-polluted soil samples clustered separately, indicating that the microbiome of the pristine environment were different from the hydrocarbon stressed environment (Fig. 3). Furthermore, the samples obtained from Bodo West, Ogoniland, and the oil-polluted samples from India clustered according to geographical location indicating that the factors that determine microbial selection were beyond just the presence of hydrocarbons (Table 6).

A closer comparison of the bacterial structure of our study site to other oil-polluted sites revealed significant differences in bacterial families ranging from 21 to 42 bacterial families. The bacterial dataset sequenced from Barhola oil refinery in India had 42 bacterial families (Fig. 6a) significantly differentially abundant in comparison with the bacterial data from this study. *Burkholderaceae* and *Anaerolineaceae* were among the bacterial families with the highest odds ratio. Noonmati oil refinery dataset revealed that 33 bacterial families (Fig. 6b) were differentially abundant. The dataset obtained from Pensacola Beach, USA, with sample codes PB1, PB2, PB3, PB4, and PB5 had significant differences in microbial structure when compared with our study site ranging from 21 to 33 bacterial families (Fig. 6c–g). The overall analysis revealed the bacterial families *Acidithiobacillaceae* and *Desulfobacteraceae* were significantly over-represented in our study site compared to other oil-polluted sites used in this study.

Detection of biomarker taxa

Differentially abundant bacterial and achaeal taxa referred to as biomarkers were detected using LEfSe algorithm. For this analysis, the alpha parameter significance for the Kruskal–Wallis (KW) and the logarithmic LDA score cut-off were set to 0.05 and 2.0 respectively. Both analyses were to detect features with significant differential abundance with respect to the class of interest and to estimate the effect size of each differentially abundant feature. The biomarker taxa detected for the oil-polluted and pristine soils are shown in Fig. 4. The genera *Sneathiella*, *Parvibaculum*, *Oceanicaulis*, *Thalassospira*, C1 B045 belonging to the family *Porticoccaceae* and the genera *Solimonas*, *Fontimonas*, *Thioalkalispira*, *Sphingobium*, KCM 112 belonging to the family *Acidithiobacillaceae* were the predominant biomarker taxa detected in the oil-polluted samples.

Predicted functional profile of the distinct environments

In order to determine the effect of the long-term crude oil spill on soil microbiome functional potential, the obtained taxonomic profile data were subjected to PICRUSt analysis for functional predictions. Predicted proteins were further classified by KEGG orthologs (KOs) and this resulted in the identification of 6909 KOs which were collapsed to 302 metabolic pathways. The collapsed pathways were analyzed using LEfSe for differentially abundant pathways. Twenty-six pathways were found to be differentially abundant (Fig. 5). Fifteen of the biomarker pathways were associated with the crude oil-polluted environment, while the remaining 11 differentially represented pathways were biomarkers for the pristine environment Fig. 6. The biomarker pathways for the distinct environments and their p values are shown in Table 3. The core biomarkers detected in the crude oil-polluted environments were mostly pathways for degradation and metabolism processes which included naphthalene degradation (KO00626),

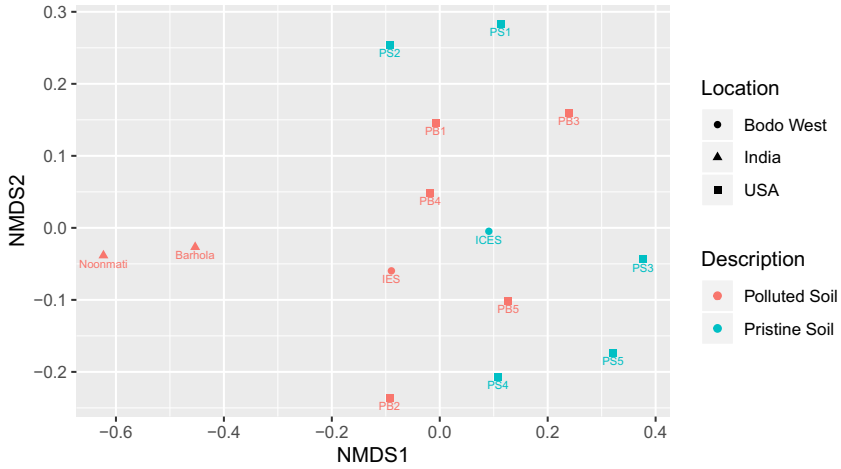


Fig. 3 NMDS plot for visualization of sample similarity or differences based on bacterial structure

fatty acid metabolism (ko00071), lysine degradation (ko00310), metabolism of xenobiotics by cytochrome P450 (ko00980), C5 branched dibasic acid metabolism (ko00660), beta alanine metabolism (ko00410), tryptophan metabolism (ko00380), propanoate metabolism (ko00640), limonene and pinene degradation (ko00903), geraniol degradation (ko00281), caprolactam degradation (ko00930), toluene degradation (ko00623) and drug metabolism enzymes (ko00983), whereas the core biomarkers for the pristine environment included biosynthesis of vancomycin group antibiotics

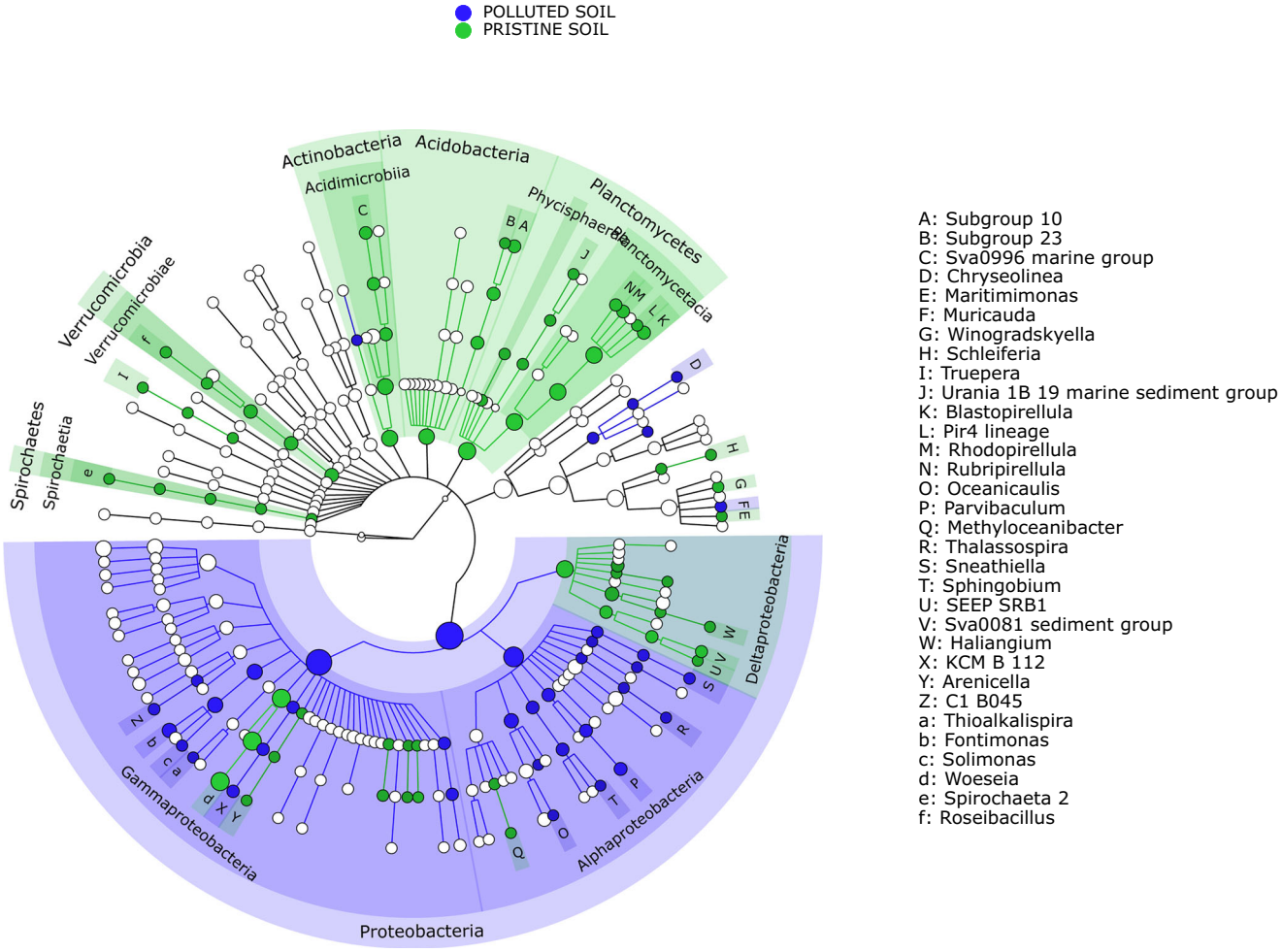


Fig. 4 Phyletic representation of taxonomic biomarkers detected for the pristine and crude oil-polluted samples. The external legend represent the detected biomarker taxa, while the innermost ring represent levels 3 to 6 taxonomic hierarchy of the detected biomarker taxa. Differentially abundant taxa are colored according to oil pollution history of the distinct environments

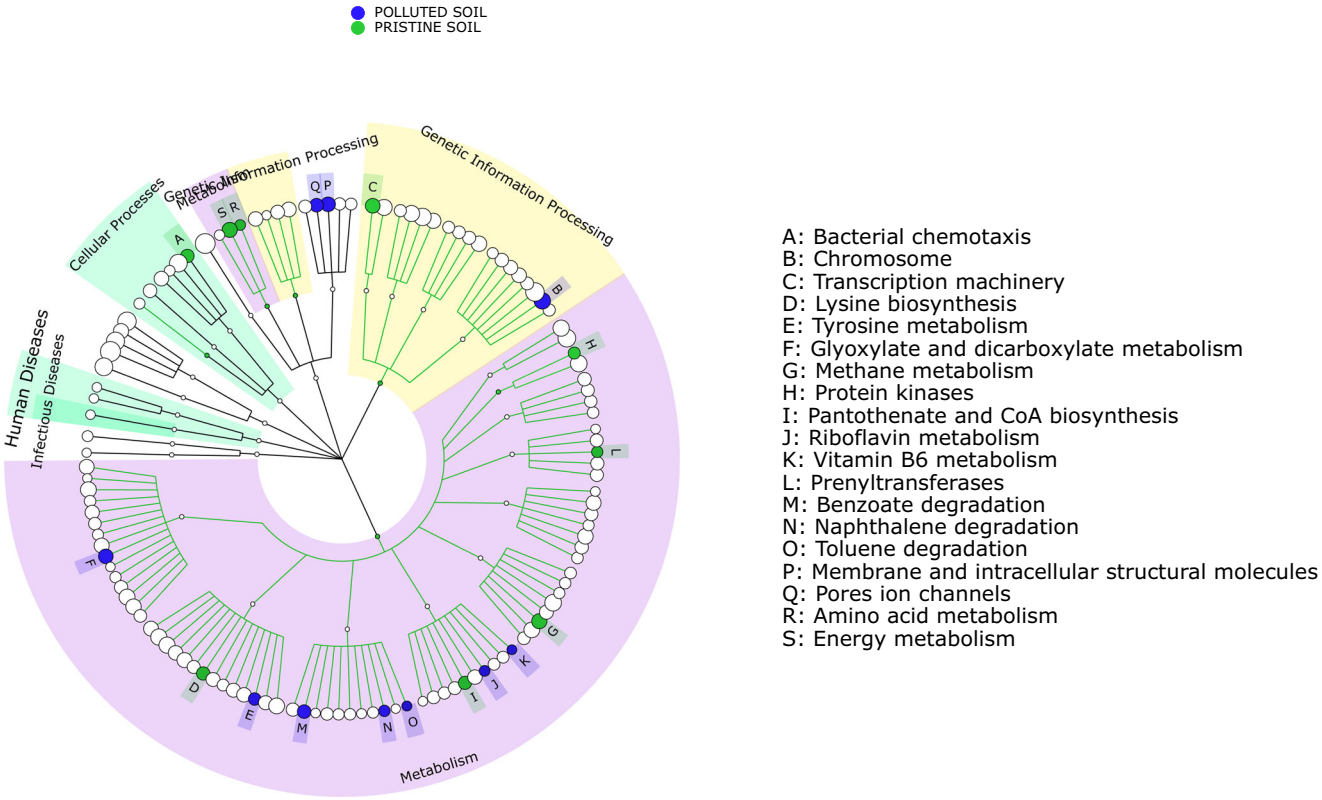


Fig. 5 Collapsed metagenomes prediction at the functional level. The external legend represents level III KEGG functional pathways. Differentially abundant KEGG pathways are colored corresponding to study environment (crude oil-polluted and pristine soils). KEGG functional pathways not differentially represented in any of the study sites are colorless. The size of each circle corresponds with the abundance of that particular biomarker

(ko01055), seleno-compound metabolism (ko00450), pantothenate and CoA biosynthesis (ko00770), bacterial chemotaxis (ko02030), and valine leucine and isoleucine biosynthesis (ko00290). Detailed biomarkers obtained for all environments are shown in Table 3.

Permutation-based multivariate analysis was carried out on the gene abundance using Bray–Curtis similarity scores to determine significant differences in KEGG pathways for the distinct environments. The p value ($p = 0.004$) obtained indicated significant differences exist in gene abundance between the pristine and oil-polluted environments. The collapsed pathway biomarkers (Fig. 5) showed that pathways for toluene degradation, glyoxylate and dicarboxylate metabolism, naphthalene degradation and benzoate degradation were major pathways for the crude oil-polluted environment.

Discussion

Physicochemical analyses

The physicochemical characteristics of the soil samples obtained from Bodo West community, Ogoniland spill site, showed a high concentration of residual total petroleum hydrocarbon and exceeded the permissible (50 mg/kg) and intervention (5000 mg/kg) limits as set out in the Environmental Guidelines and Standards for the Petroleum Industry in Nigeria (DPR 2002). The electrical conductivity of the polluted soil was found to be higher than that of the pristine soil (Table 1). A previous study at the Cason City site, USA, revealed an increase in the proportion of hydrocarbon-degrading microbes with an increase in electrical conductivity (Allen et al. 2007). Also, pH has been found to contribute significantly to explaining the observed variation in community composition (Sutton et al. 2012). The pH values of the crude oil-polluted soil was 7.8 while that of the pristine soil was slightly acidic (6.5), a finding that is in line with the report of Lindén and Pålsson (2013) that carried out an extensive chemical analyses of over 40 sites in Ogoniland. They reported the average pH to be within 6.1. The nitrate and phosphate contents in the pristine soil sample were significantly higher when compared to the polluted soil sample; this could be as a result of an abundance of organic matter in the pristine soil environments. Additionally, microbial community composition is highly correlated to physicochemical parameters, as described by previous investigations (Kostka et al. 2011; Kadali et al. 2012; Keshri et al. 2015).

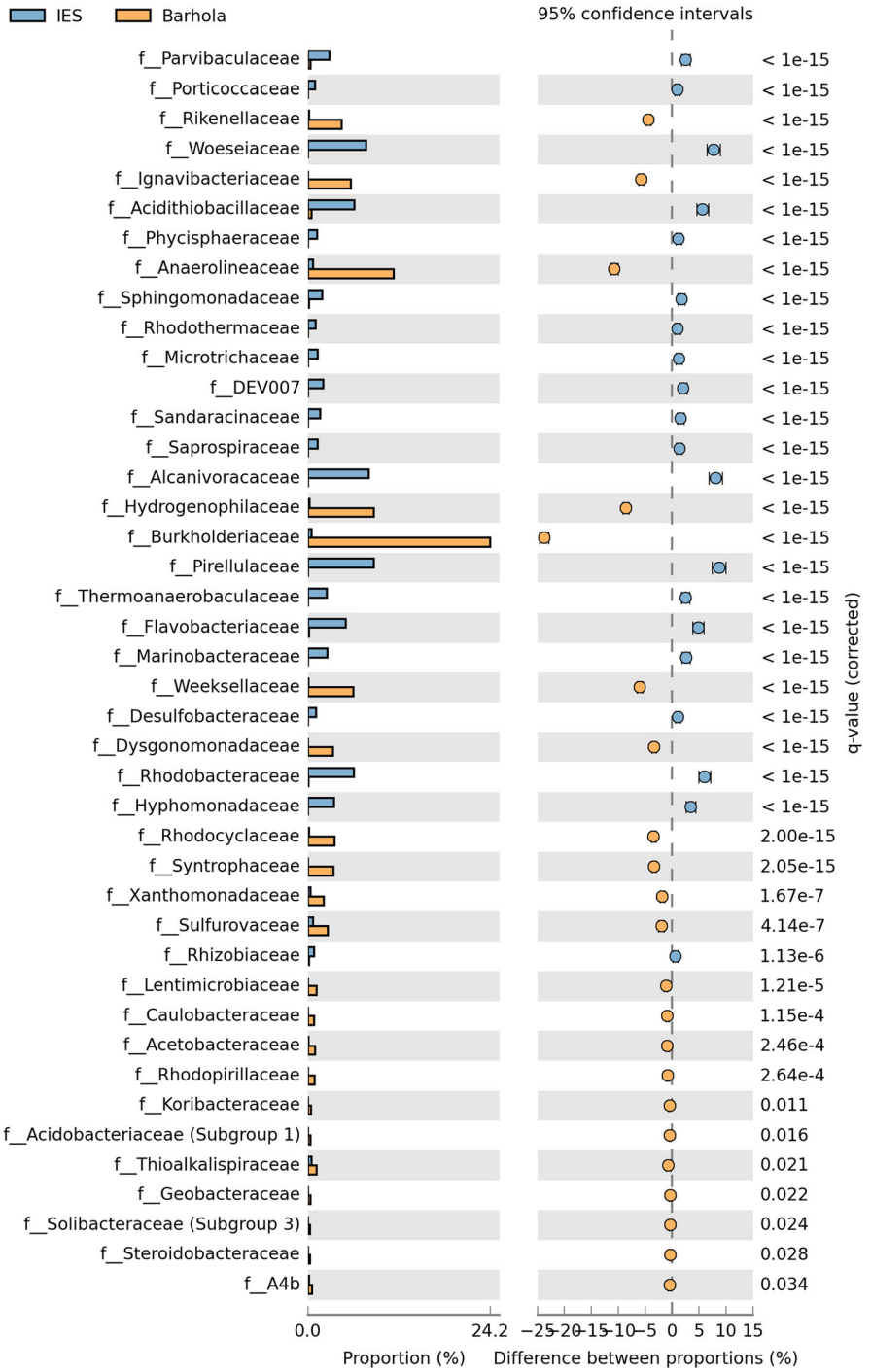


Fig. 6 **a** Differentially abundant bacterial families and the odds ratio and significant *p* values for Bodo West spill site (IES) and Barhola refinery (Barhola), India. Extended bar plots are arranged according to significant *p* values. **b** Differentially abundant bacterial families and the odds ratio and significant *p* values for Bodo West spill site (IES) and Noonmati refinery (Noonmati), India. Extended bar plots are arranged according to significant *p* values. **c** Differentially abundant bacterial families and the odds ratio and significant *p* values for Bodo West spill site (IES) and Pensacola beach sample (PB1), USA. Extended bar plots are arranged according to significant *p* values. **d** Differentially abundant bacterial families and the odds ratio and significant *p* values for Bodo West spill site (IES) and Pensacola beach sample (PB2), USA. Extended bar plots are arranged according to significant *p* values. **e** Differentially abundant bacterial families and the odds ratio and significant *p* values for Bodo West spill site (IES) and Pensacola beach sample (PB3), USA. Extended bar plots are arranged according to significant *p* values. **f** Differentially abundant bacterial families and the odds ratio and significant *p* values for Bodo West spill site (IES) and Pensacola beach sample (PB4), USA. Extended bar plots are arranged according to significant *p* values. **g** Differentially abundant bacterial families and the odds ratio and significant *p* values for Bodo West spill site (IES) and Pensacola beach sample (PB5), USA. Extended bar plots are arranged according to significant *p* values

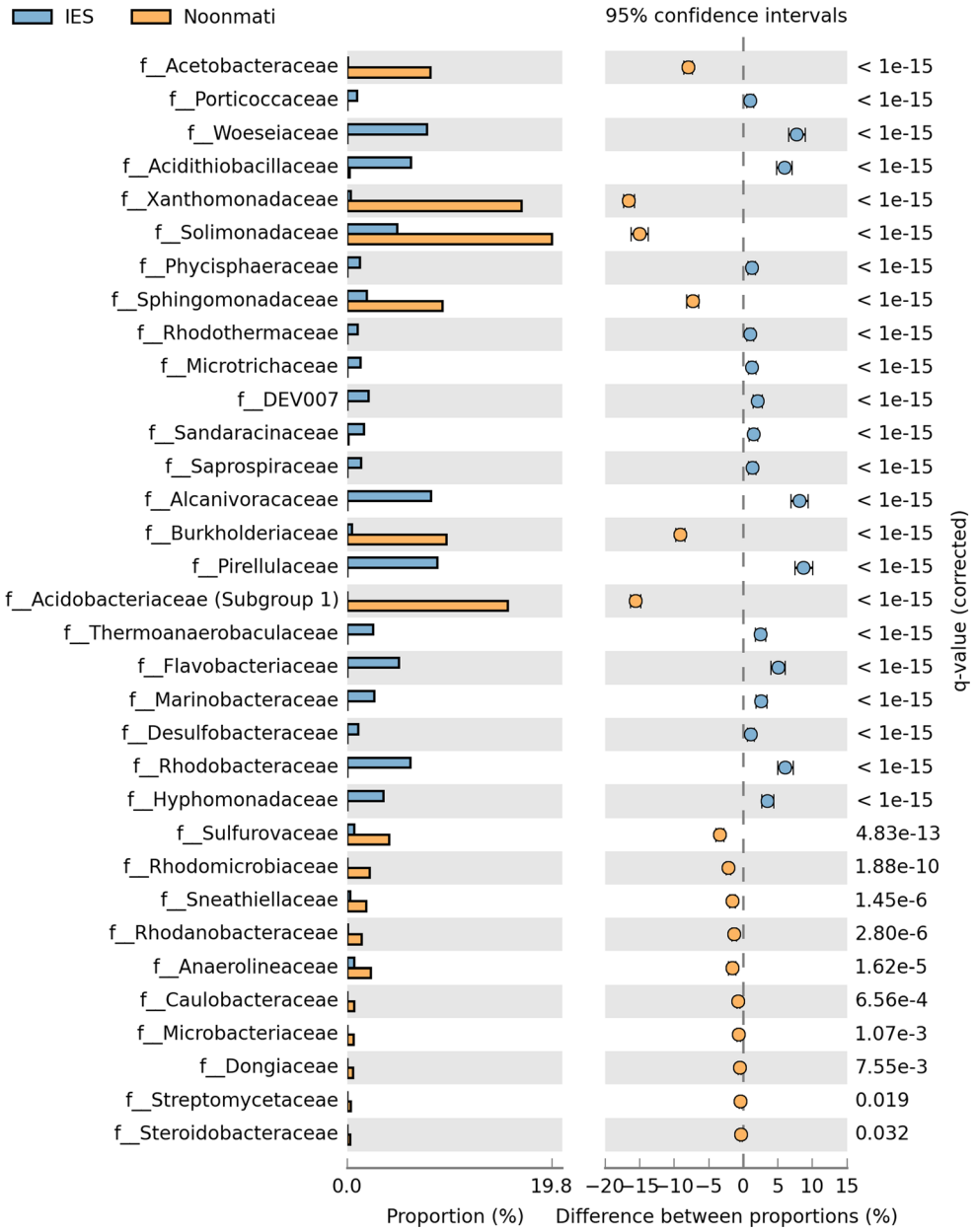
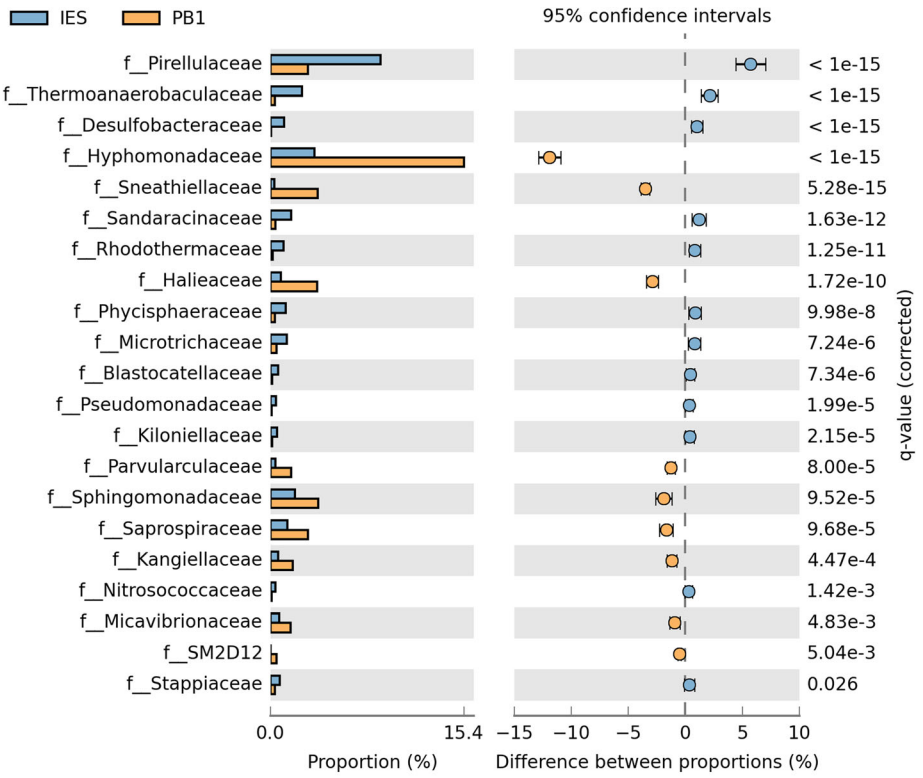


Fig. 6 (continued)

Microbial community structure and biomarker taxa

Comparison of microbial community structure among the studied environments revealed distinct taxa were differentially abundant depending on environmental stress exerted on the microbial community. Proteobacteria was predominant in all the samples analyzed and has been identified in many studies as the predominant phylum in soil samples (Roesch et al. 2007; Militon et al. 2010; dos Santos et al. 2011; Nacke et al. 2011; Fahrenfeld et al. 2014). Within this phylum, Gammaproteobacteria and to a lesser extent the Epsilonproteobacteria, Alphaproteobacteria, and Deltaproteobacteria predominated the bacterial communities of the oil-polluted environment. Several researchers have made similar findings in oil-polluted soils (Head et al. 2006; Yakimov et al. 2007; Berthe-Corti and Nachtkamp 2010; Greer 2010; Sutton et al. 2012)

The dominant Gammaproteobacteria in the oil-impacted Bodo West community were *Alcanivorax* and *Marinobacter* (Fig. 8 in the Appendix). These groups of bacteria are well-known hydrocarbon degraders that usually increase in response to oil contamination in marine and brackish water environments (Yakimov et al. 2007; dos Santos et al. 2010; Kostka et al. 2011; Huettel et al. 2018). They are also known to be excellent degraders of alkane and their presence usually suggests alkane degradation occurring naturally.



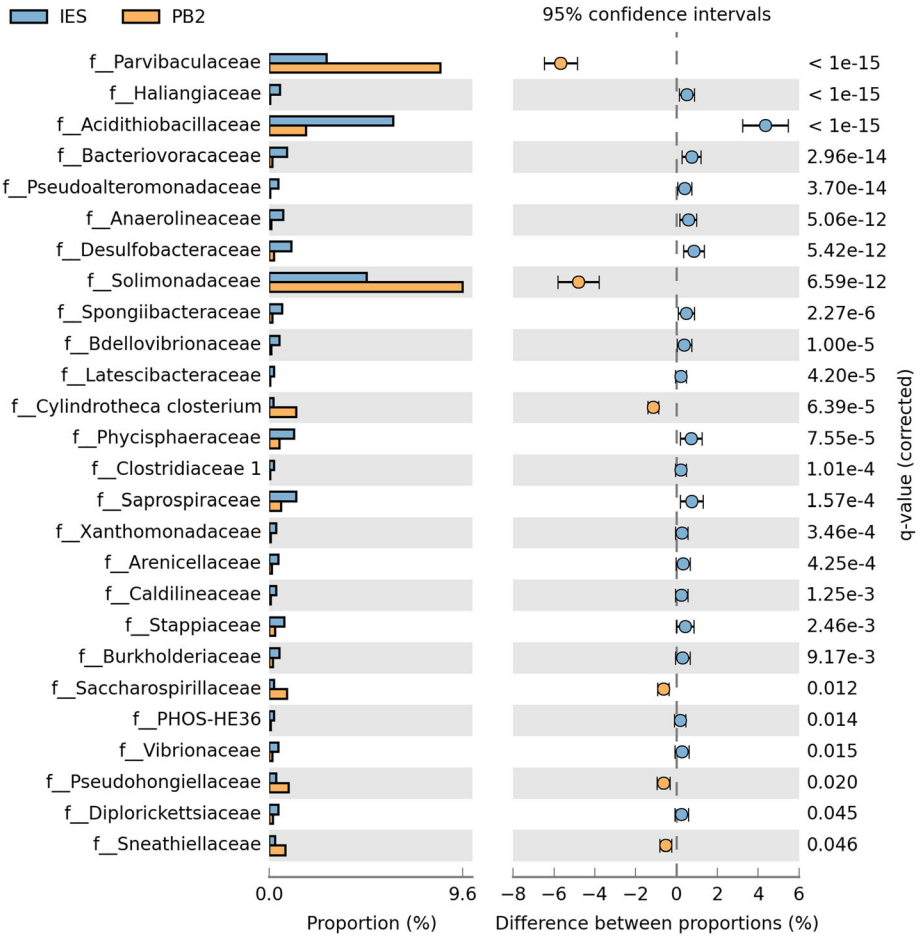


Fig. 6 (continued)

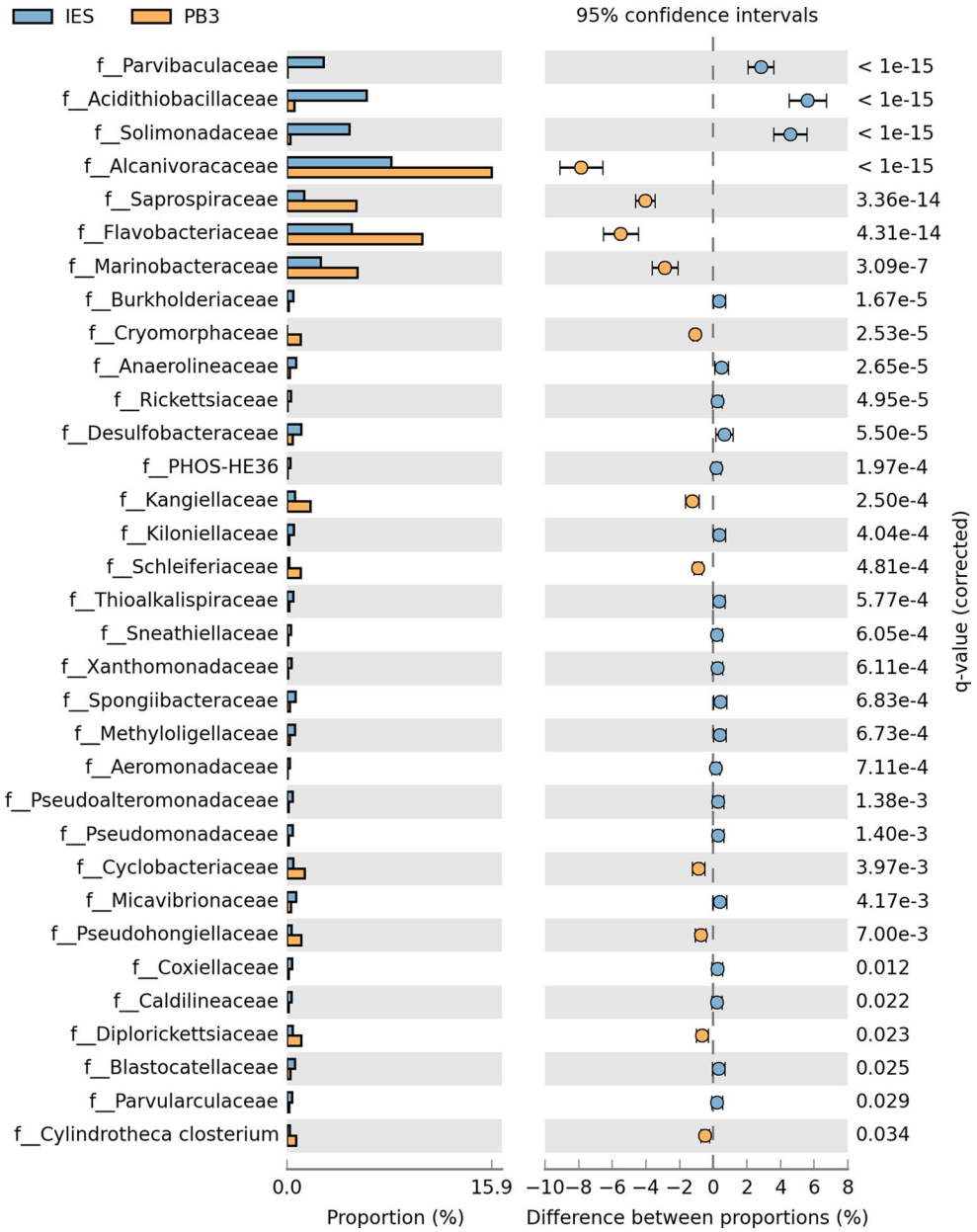


Fig. 6 (continued)

As expected, all of the pristine soil samples had higher bacterial diversity compared to the oil-impacted soils where the presence of hydrocarbon pollutant had triggered a shift in bacteria composition. The phyla *Planctomycetes*, *Actinobacteria*, and *Acidobacteria* were most predominant in the pristine soil samples and they have been reported in earlier works that examined uncontaminated soil samples (Roesch et al. 2007; Nacke et al. 2011; Sutton et al. 2012). Alpha diversity analysis (Fig. 2) showed the oil-polluted environment had the least diversity of bacterial and archaeal species. This observation is in line with the results of several researchers (Lamendella et al. 2014; Chikere and Obieze 2018; Huettel et al. 2018) that studied the effect of hydrocarbon stress on microbial diversity.

All the taxonomic biomarkers detected in the oil-polluted environment are established degraders of various fractions of hydrocarbon (Wang et al. 2010; Kimes et al. 2013; Newton et al. 2013; Kappell et al. 2014). Kappell et al. (2014) reported that the genus *Sneathiella* played a key role in the degradation of naphthalene and fluorene following the DWH oil spill of 2010. In addition, *Thalassospira* is also an established degrader of naphthalene and fluorene as it was reported to be associated with oil on the water surface during the DWH oil spill (Liu and Liu 2013). In another study of the DWH spill by Looper et al. (2013), *Parvibaculum* was detected as the dominant oil degrader and they are known to be very important degraders of both alkane and aromatics in oil-

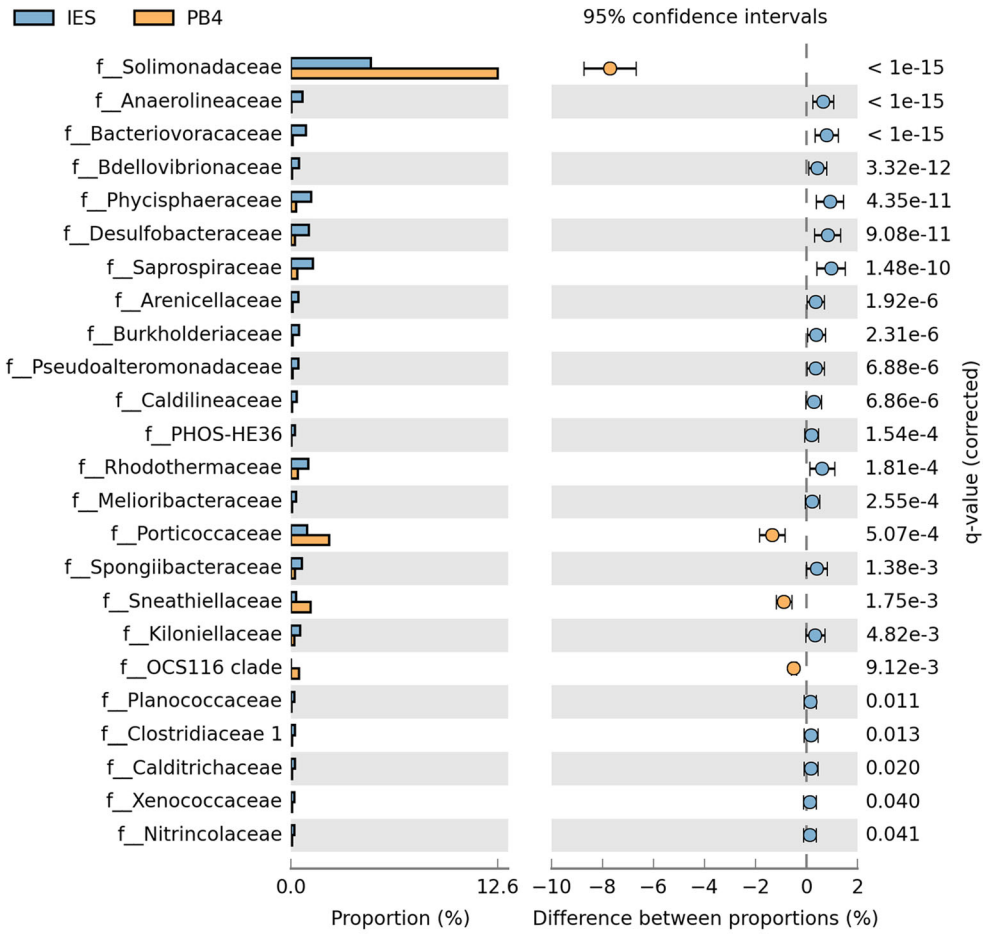


Fig. 6 (continued)

Table 3 Differentially abundant KEGG pathways detected in the studied environments

KEGG pathway	Sample history	*p* value
ko01055: Biosynthesis of vancomycin group antibiotics	Pristine soil	0.04378
ko00941: Flavonoid biosynthesis	Pristine soil	0.014298
ko00300: Lysine biosynthesis	Pristine soil	0.010951
ko03020: RNA polymerase	Pristine soil	0.044453
ko00450: Seleno-compound metabolism	Pristine soil	0.005517
ko00633: Nitrotoluene degradation	Pristine soil	0.025295
ko00770: Pantothenate and CoA biosynthesis	Pristine soil	0.010704
ko02030: Bacterial chemotaxis	Pristine soil	0.024373
ko00270: Cysteine and methioninemetabolism	Pristine soil	0.011109
ko03420: Nucleotide excision repair	Pristine soil	0.011109
ko00290: Valine leucine and isoleucine biosynthesis	Pristine soil	0.023262
ko00983: Drug metabolism other enzymes	Polluted soil	0.004732
ko04210: Apoptosis	Polluted soil	0.002693
ko00623: Toluene degradation	Polluted soil	0.022332
ko00830: Retinol metabolism	Polluted soil	0.00923
ko00281: Geraniol degradation	Polluted soil	0.015082
ko00903: Limonene and pinene degradation	Polluted soil	0.031318
ko00640: Propanoate metabolism	Polluted soil	0.016014
ko00312: Beta lactam resistance	Polluted soil	0.003855
ko04310: Wnt signaling pathway	Polluted soil	0.006108
ko00410: Beta alanine metabolism	Polluted soil	0.016864
ko00660: C5 branched dibasic acid metabolism	Polluted soil	0.047972
ko00980: Metabolism of xenobiotics by cytochrome P450	Polluted soil	0.007518
ko00071: Fatty acid metabolism	Polluted soil	0.007773
ko00626: Naphthalene degradation	Polluted soil	0.010674
ko00910: Nitrogen metabolism	Polluted soil	0.046532

polluted environments (Joye et al. 2016). The TPH analysis of the oil-polluted soil obtained from Bodo West Community showed the presence of mostly C_{17}–C_{20} hydrocarbons and the presence of a high number of polycyclic aromatic hydrocarbons. Other datasets used in this study also had a high hydrocarbon concentration of both saturated and aromatic hydrocarbons which resulted in the abundance of these key hydrocarbon degraders. Having knowledge of these important biomarker taxa in oil-polluted environments will impact heavily on engineered bioremediation of oil spills.

The presence of *Acidithiobacillaceae* in significantly differential abundance in the site of study could be related to the relatively low pH predominant in most oil-polluted soils in the Niger Delta of Nigeria and particularly in Ogoniland. Lindén and Pålsson (2013) in an extensive study of over 40 sites as part of the United Nations Environmental Programme (UNEP) assessment of crude oil-polluted soils, sediments and water in Ogoniland, reported a pH range of between 4.7 and 7.3 with an average of 6.1 for all the sites investigated. The bacterial family *Acidithiobacillaceae* is usually present in soils with high sulfate concentration and low pH. Hamamura et al. (2005) detected the presence of iron- and sulfur-oxidizing chemolithotroph *Acidithiobacillus* spp. in an environment associated with natural hydrocarbon seeps. Valdés et al. (2008) reported the genome of *Acidithiobacillus ferrooxidans* to contain toluene tolerance protein TtgD (AFE1830) as well as xylene and related aromatic hydrocarbon compounds.

Another differentially abundant bacterial family in our study site is *Desulfobacteraceae*. Members of *Desulfobacteraceae* have been shown to play important roles in the degradation of aromatic compounds, particularly under sulfate-reducing conditions. Kümmel et al. (2015) investigated the role of the bacterial family *Desulfobacteraceae* in the degradation of naphthalene. They demonstrated the mineralization of naphthalene by monitoring sulfide formation which is usually concomitant with naphthalene depletion. Further investigation will be required to determine the exact role of these two important differentially abundant bacterial families in the study site and in Ogoniland in general since the soil chemistry around this area will support their proliferation.

Comparison of the functional profiles of the pristine and oil-impacted soils

Predicted metabolic profiles for the oil-polluted and pristine environments showed distinct functional capabilities. Results of this study showed that the differentially abundant metagenomes in the oil-polluted environment were majorly the pathways for metabolism and biodegradation, while the differentially represented pathways for the pristine soil were mostly pathways for sugar and nucleotide biosynthesis apart from the pathway for nitrotoluene degradation. The detection of nitrotoluene degradation pathway in the pristine soil is not surprising, as pathways for the degradation of alkanes, poly-aromatics, and xenobiotics are known to be ubiquitous with the only difference being that environments under hydrocarbons stress will have a higher abundance of the degradative genes responsible for hydrocarbons breakdown (Looper et al. 2013).

The predicted metabolic profile of the oil-polluted environment suggests heavy reliance of the microbial community on hydrocarbon as the sole source of carbon. In total contrast to the pristine environment, several metabolic pathways for the degradation of alkanes and polycyclic aromatic hydrocarbons were shown to be differentially abundant in the oil-polluted samples. Pathways for naphthalene degradation, lysine degradation, metabolism of xenobiotics by cytochrome P_{450}, and toluene degradation were among the core KEGG pathways discovered as biomarkers for the oil-polluted environment. This observation suggests that this increase is as a result of the environmental stress triggered by the presence of petroleum hydrocarbons in the environment. The study of Kappell et al. (2014) showed that the majority of differentially detected functional genes responsible for both saturated and aromatic hydrocarbon degradation in their study were associated with oil-polluted soils. Similar observations were made by Bao et al. (2017) and Mukherjee et al. (2017) who successfully used predictive metagenomics to determine bacterial response to hydrocarbon contamination. Their findings revealed that petroleum hydrocarbon contaminated sites had a significant differential functional profile and particularly higher functional KEGG pathways for hydrocarbons degradation.

The limitation of macro-nutrients such as nitrogen is not uncommon in oil-contaminated environments (Leahy and Colwell 1990). A study by Hazen et al. (2010) showed that an increased abundance of genes associated with nitrogen assimilation led to an increase in biomass in the water column during the DWH spill; hence, the differential abundance of KEGG pathway for nitrogen metabolism in the oil-polluted environment could suggest that the microbial community is responding to nitrate limitation in the oil-polluted environment. Similarly, Kappell et al. (2014) demonstrated that microorganisms respond to nitrate limitation by either fixing nitrogen or through some yet to be detected mechanisms. Our findings indicate that several factors including availability of different organic and inorganic compounds and environmental stress can heavily influence the evolution of the microbiome. The hydrocarbon stress led to an increase in bacteria harboring genes responsible for the degradation of various hydrocarbon fractions.

Conclusions

In conclusion, high-throughput sequencing and predictive metagenomics revealed a significant difference in biomarker pathways and taxa for the studied environments. The presence of petroleum hydrocarbons influenced microbial diversity and its functions in the oil-polluted environments. Furthermore, the predictive metagenomic analysis revealed that members of the microbial community in the hydrocarbon polluted environment relied mostly on hydrocarbons as their source of carbon due to the high abundance of pathways for hydrocarbons degradation as well as a corresponding increase of established oil-degrading bacterial taxa detected as biomarkers in this environment. However, further investigations are required to gather information on possible distinctive roles played by these bacterial species in these habitats. The taxonomic and functional profile of Bodo West Community oil spill site indicated the presence of versatile and well-established hydrocarbon-degrading bacterial species and a significantly higher abundance of acid-tolerant bacteria families that can be exploited for the recovery of oil-polluted soils in Ogoniland and other hydrocarbon-impacted soils with acidic pH.

Acknowledgements The authors would like to thank the Centre for High Performance Computing (CHPC) facility, South Africa, for providing computational support for sequence data analysis.

Data accessibility statement Sequence reads for Bodo West, Ogoniland, samples were deposited in GenBank (Sequence Reads Archive) under the SRA accession number SRP133543.

Funding This study was self-funded by the authors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Research involving human participants and/or animals N/A

Informed consent N/A

Appendix

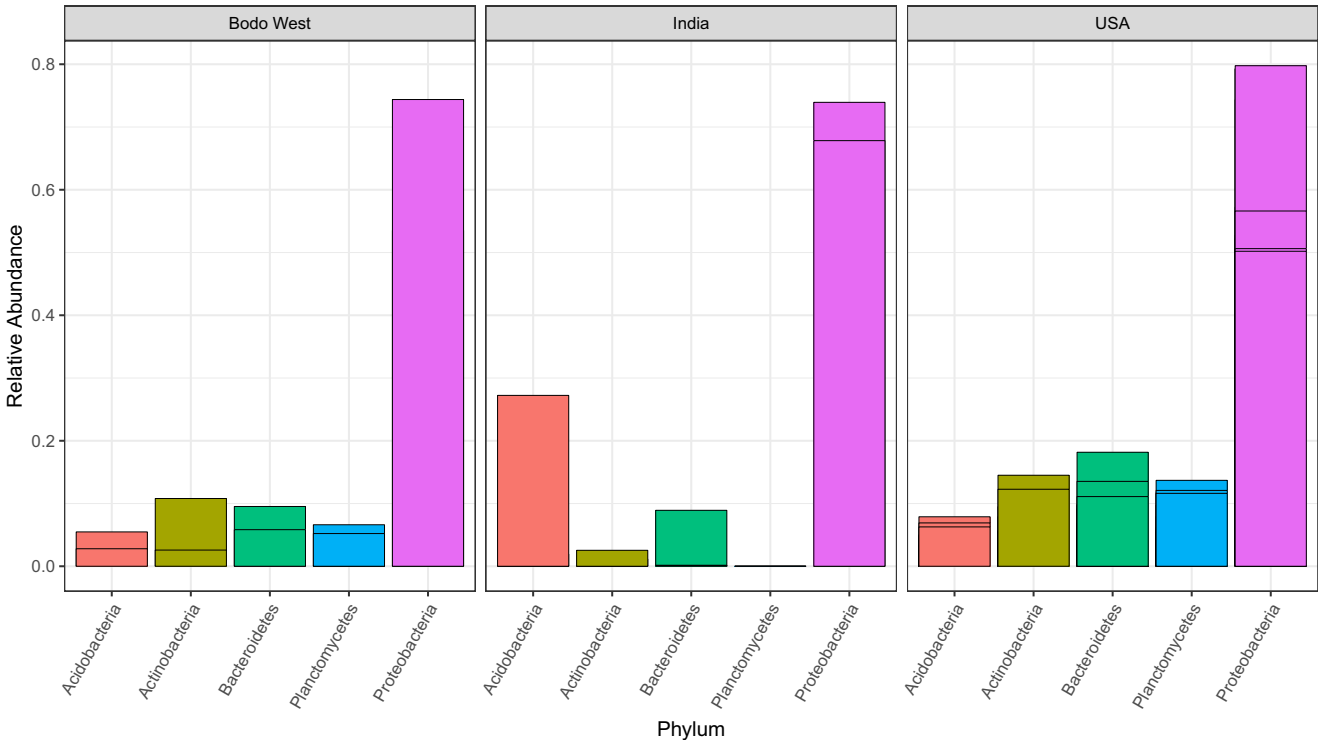


Fig. 7 Abundance and distribution of the five most abundant phyla across the samples. *USA (Pensacola beach oil-polluted and pristine soils). India (oil-polluted soils from Noonmati and Barhola oil refineries India). Bodo West (oil-polluted soils, Bodo West, Ogoniland, Nigeria)

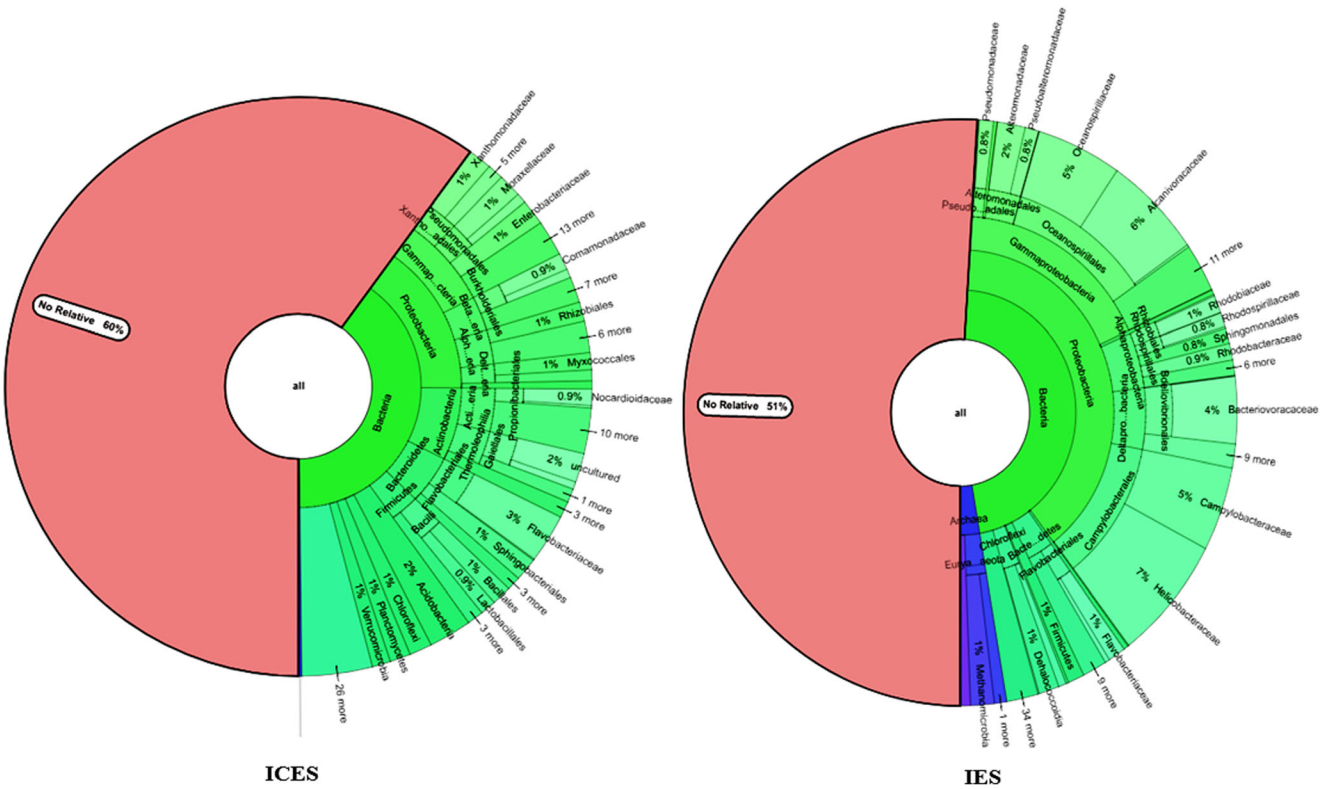


Fig. 8 Comparison of microbial community structure and its abundance in the crude oil-polluted and the pristine soil samples obtained from Bodo West, Ogoniland, Nigeria. *IES (polluted soil). *ICES (pristine soil)

Table 4 Summary information on datasets used for this study

Sample code	Pollution history	Location	Sample type	SRA Accession number	Reference
ICES	Crude oil-polluted soil	Bodo West, Ogoniland, Nigeria	Soil	SRP133543	This study
IES	Pristine soil	Bodo West, Ogoniland, Nigeria	Soil	SRP133543	This study
PS1	Pristine soil	Pensacola beach, FL, USA	Soil	SRX1434718	Huettel et al. (2018)
PS2	Pristine soil	Pensacola beach, Florida, USA	Soil	SRX1434719	Huettel et al. (2018)
PS3	Pristine soil	Pensacola beach, FL, USA	Soil	SRX1434720	Huettel et al. (2018)
PS4	Pristine soil	Pensacola beach, FL, USA	Soil	SRX1434721	Huettel et al. (2018)
PS5	Pristine soil	Pensacola beach, FL, USA	Soil	SRX1434713	Huettel et al. (2018)
Noonmati	Crude oil-polluted soil	Guwahati, Assam, India	Soil	SRX1584453	Mukherjee et al. (2017)
Barhola	Crude oil-polluted soil	Barhola, Assam, India	Soil	SRX1584455	Mukherjee et al. (2017)
PB1	Crude oil-polluted soil	Pensacola beach, FL, USA	Soil	SRX1434694	Huettel et al. (2018)
PB2	Crude oil-polluted soil	Pensacola beach, FL, USA	Soil	SRX1434696	Huettel et al. (2018)
PB3	Crude oil-polluted soil	Pensacola beach, FL, USA	Soil	SRX1434704	Huettel et al. (2018)
PB4	Crude oil-polluted soil	Pensacola beach, FL, USA	Soil	SRX1434712	Huettel et al. (2018)
PB5	Crude oil-polluted soil	Pensacola beach, FL, USA	Soil	SRX1434713	Huettel et al. (2018)

Table 5 PERMANOVA analysis for bacterial diversity using Bray–Curtis

Summary		
Permutation *N*	9999	
Total sum of squares	1.749	
Within-group sum of squares	1.339	
F	3.673	
p (same)	0.0052	
Pair-wise		
Samples	Polluted soil	Pristine soil
Polluted soil		0.0053
Pristine soil	0.0053	

Table 6 PERMANOVA analysis for gene abundance using Bray–Curtis

Summary		
Permutation *N*	9999	
Total sum of squares	0.4436	
Within-group sum of squares	0.2885	
F	2.958	
p (same)	0.0041	
Pair-wise		
Samples	Polluted soil	Pristine soil
Polluted soil		0.018

References

Abia ALK, Alisoltani A, Keshri J, Ubomba-Jaswa E, (2018) Metagenomic analysis of the bacterial communities and their functional profiles in water and sediments of the Apies River, South Africa, as a function of land use. Sci Total Environ 616–617:326–334. https://doi.org/10.1016/j.scitotenv.2017.10.322

Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, … Huttenhower C (2012) Metabolic reconstruction for metagenomic data and its application to the human microbiome, 8(6). https://doi.org/10.1371/journal.pcbi.1002358

Allen JP, Atekwana EA, Duris JW, Werkema DD, Rossbach S (2007) The microbial community structure in petroleum-contaminated sediments corresponds to geophysical signatures. Appl Environ Microbiol 73(9):2860–2870

APHA (2012) Standard methods for the examination of water and wastewater. Standard Methods, 541. ISBN 9780875532356

Ataikiru TL, Okorhi-damisa BF, Akpaiboh JI (2017) Microbial community structure of an oil polluted site in. International Research Journal of Public and Environmental Health, 4(3), 41–47

Bao YJ, Xu Z, Li Y, Yao Z, Sun J, Song H (2017) High-throughput metagenomic analysis of petroleum-contaminated soil microbiome reveals the versatility in xenobiotic aromatics metabolism. J Environ Sci (China) 56:25–35. https://doi.org/10.1016/j.jes.2016.08.022

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol. WileyRoyal Statistical Society. https://doi.org/10.2307/2346101

Berthe-Corti L, Nachtkamp M (2010) Bacterial communities in hydrocarbon-contaminated marine coastal environments. In: Timmis KN (ed) Handbook of hydrocarbon and lipid microbiology. Springer, Berlin

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ et al (2011a) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci 108(Supplement_1):4516–4522. https://doi.org/10.1073/pnas.1000080107

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, … Walters WA (2011b). QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7(5), 335–336. https://doi.org/10.1038/nmeth.f.303

Chen C, Khaleel SS, Huang H, Wu CH (2014) Software for pre-processing Illumina next-generation sequencing short read sequences. Source Code Biol Med 9(8):1–11. https://doi.org/10.1038/nbt1486

Chikere CB, Azubuike CC, Fubara EM (2017). Shift in microbial group during remediation by enhanced natural attenuation (RENA) of a crude oil-impacted soil: A case study of Ikarama Community, Bayelsa, Nigeria. 3Biotech 7:152. https://doi.org/10.1007/s13205-017-0782-x

Chikere CB, Obieze CC (2018) Effect of hydrocarbon pollution on microbial diversity and implication for bioremediation. In Proceedings of the 2018 International Women in Science Without Borders (WISWB) – Indaba. (pp. 10–11). Johannesburg. Retrieved from https://researchspace.csir.co.za/dspace/handle/10204/10251

dos Santos M, Sabirova V, Timmis J, Yakimov MM (2010) Alcanivorax borkumensis. In: Timmis KN (ed) Handbook of hydrocarbon and lipid microbiology. Springer-Verlag, Berlin, pp 1266–1285

dos Santos HF, Cury JC, do Carmo FL, dos Santos AL, Tiedje J, van Elsas JD, Rosado AS, Peixoto RS, Heimesaat M (2011) Mangrove bacterial diversity and the impact of oil contamination revealed by pyrosequencing: bacterial proxies for oil pollution. PLoS ONE 6(3): e16943

DPR (2002) Environmental Guidelines and Standards for the Petroleum Industry in Nigeria (Egaspin). Lagos. Retrieved from http://www.ngfcp.gov.ng/media/1066/dprs-egaspin-2002-revised-edition.pdf

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27(16):2194–2200. https://doi.org/10.1093/bioinformatics/btr381

Fahrenfeld N, Cozzarelli IM, Bailey Z, Pruden A (2014) Insights into biodegradation through depth-resolved microbial community functional and structural profiling of a crude-oil contaminant plume. Microb Ecol 68(3):453–462

Feng G, Xie T, Wang X, Bai J, Tang L, Zhao H et al (2018) Metagenomics analysis of microbial community and function involved in Cd-contaminated soil. BMC Microbiol 18(11):1–13. https://doi.org/10.1186/s12866-018-1152-5

Gałązka A, Grządziel J, Gałązka R, Ukalska-Jaruga A, Strzelecka J, Smreczak B (2018) Genetic and functional diversity of bacterial microbiome in soils with long term impacts of petroleum hydrocarbons. Front Microbiol 9:1923. https://doi.org/10.3389/fmicb.2018.01923

Gibson J, Harwood CS, Harwood SC (2002) Metabolic diversity in aromatic compound utilization by anaerobic microbes. Annu Rev Microbiol. 56(1):345–369. https://doi.org/10.1146/annurev.micro.56.012302.160749

Greer CW (2010) Bacterial diversity in hydrocarbon-polluted rivers, estuaries and sediments. In: Timmis KN (ed) Handbook of hydrocarbon and lipid microbiology. Springer, Berlin

Hamamura N, Olson SH, Ward DM, Inskeep WP (2005) Diversity and functional analysis of bacterial communities associated with natural hydrocarbon seeps in acidic soils at Rainbow Springs, Yellowstone National Park. Appl Environ Microbiol 71(10):5943–5950

Hammer O, Harper DAT, Ryan PD (2001) PAST: Palentological Statistics Software Package for Education and Data Analysis, 4(1), 1–9

Hassanshahian M, Zeynalipour MS, Musa FH, Hassanshahian M, Zeynalipour MS, Musa FH (2014) Isolation and characterization of crude oil degrading bacteria from the Persian Gulf (Khorramshahr provenance). Mar Pollut Bull 82(1–2):39–44. https://doi.org/10.1016/j.marpolbul.2014.03.027

Hazen TC, Dubinsky EA, DeSantis TZ, Andersen GL, Piceno YM et al (2010) Deep-sea oil plume enriches indigenous oil-degrading bacteria. Science 330(6001):204–208

Head IM, Jones DM, Röling WFM (2006) Marine microorganisms make a meal of oil. Nat Rev Microbiol 4(3):173–182

Huettel M, Overholt WA, Kostka JE, Hagan C, Kaba J, Brian W, Dudley S (2018) Degradation of Deepwater Horizon oil buried in a Florida beach influenced by tidal pumping Gulf of Mexico, 126(December 2017), 488–500

Joshi MN, Dhebar SV, Bhargava P, Pandit AS, Patel RP, Saxena AK, Bagatharia SB (2014) Metagenomic approach for understanding microbial population from petroleum muck. Genome Announc. 2(3):e00533–e00514. https://doi.org/10.1128/genomeA.00533-14. Copyright

Joye S, Kleindienst S, Gilbert JA, Handley KM, Weisenhorn P, Overholt WA, Kostka JE (2016) Responses of microbial communities to hydrocarbon exposures. Oceanography 29(3):1–15

Kadali KK, Simons KL, Skuza PP, Moore RB, Ball AS (2012) A complementary approach to identifying and assessing the remediation potential of hydrocarbonoclastic bacteria. J Microbiol Methods 88(3):348–355

Kappell AD, Wei Y, Newton RJ, Van Nostrand JD, Zhou J, McLellan SL, Yakimov MM (2014) The polycyclic aromatic hydrocarbon degradation potential of Gulf of Mexico native coastal microbial communities after the deepwater horizon oil spill. Front Microbiol 5:1–13

Keshri J, Mankazana BJ, Momba MN (2015) Profile of bacterial communities in South African mine-water samples using Illumina next-generation sequencing platform. Appl Microbiol Biotechnol 99(7): 3233–3242

Kimes NE, Callaghan AV, Aktas DF, Smith WL, Sunner J, Golding B, Morris PJ (2013) Metagenomic analysis and metabolite profiling of deep–sea sediments from the Gulf of Mexico following the Deepwater Horizon oil spill. Front Microbiol 4:1–17

Kostka JE, Prakash O, Overholt WA, Green SJ, Freyer G, Canion A, Delgardio J, Norton N, Hazen TC, Huettel M (2011) Hydrocarbon-degrading bacteria and the bacterial community response in Gulf of Mexico beach sands impacted by the deepwater horizon oil spill. Appl Environ Microbiol 77(22):7962–7974

Kumari S, Kumar R, Manickam N (2018) Improved polycyclic aromatic hydrocarbon degradation in a crude oil by individual and a consortium of bacteria. Bioresour Technol 254(November 2017):1–26. https://doi.org/10.1016/j.biortech.2018.01.075

Kümmel S, Herbst F, Bahr A, Duarte M, Pieper DH, Jehmlich N, Seifert J, von Bergen M, Bombach P, Richnow HH, Vogt C (2015) Anaerobic naphthalene degradation by sulfate-reducing Desulfobacteraceae from various anoxic aquifers. FEMS Microbiol Ecol 91(3)

Lamendella R, Strutt S, Borglin S, Chakraborty R, Tas N, Mason OU, Hultman J, Prestat E, Hazen TC, Jansson JK (2014) Assessment of the deepwater horizon oil spill impact on Gulf coast microbial communities. Front Microbiol 5

Langille MGI, Zaneveld J, Caporaso JG, Mcdonald D, Knights D, Reyes JA et al (2013) Analysis predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 31(9):814–821. https://doi.org/10.1038/nbt.2676

Leahy JG, Colwell RR (1990) Microbial degradation of hydrocarbons in the environment. Microbiol Rev 54(3):305–315

Lindén O, Pålsson J (2013) Oil contamination in ogoniland, Niger delta. Ambio 42(6):685–701. https://doi.org/10.1007/s13280-013-0412-8

Looper JK, Cotto A, Kim B, Lee M, Liles MR, Chadhain SM, Son A (2013) Microbial community analysis of Deepwater Horizon oil-spill impacted sites along the Gulf coast using functional and phylogenetic markers. Environ Sci Process Impacts 15(11):2068

Liu Z, Liu J (2013) Evaluating bacterial community structures in oil collected from the sea surface and sediment in the northern Gulf of Mexico after the oil spill. MicrobiologyOpen 2(3):492–504

Lu L, Xing D, Ren N (2012) Pyrosequencing reveals highly diverse microbial communities in microbial electrolysis cells involved in enhanced H 2 production from waste activated sludge. Water Res 46(7):2425–2434. https://doi.org/10.1016/j.watres.2012.02.005

McMurdie PJ, Holmes S (2013) Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8:e61217

Militon C, Boucher D, Vachelard C, Perchet G, Barra V, Troquet J, Peyretaillade E, Peyret P (2010) Bacterial community changes during bioremediation of aliphatic hydrocarbon-contaminated soil. FEMS Microbiol Ecol 74(3):669–681

Mukherjee A, Chettri B, Langpoklakpam JS, Basak P, Prasad A, Mukherjee AK, Bhattacharyya M, Singh AK, Chattopadhyay D (2017) Bioinformatic approaches including predictive metagenomic profiling reveal characteristics of bacterial response to petroleum hydrocarbon contamination in diverse environments. Sci Rep 7(1): 1108. https://doi.org/10.1038/s41598-017-01126-3

Muyzer G, de Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol 59(3):695–700 Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7683183

Nacke H, Thürmer A, Wollherr A, Will C, Hodac L, Herold N, Schöning I, Schrumpf M, Daniel R, Gilbert J (2011) Pyrosequencing-based assessment of bacterial community structure along different management types in German forest and grassland soils. PLoS ONE 6(2):e17000

Newton RJ, Huse SM, Morrison HG, Peake CS, Sogin ML, McLellan SL (2013) Shifts in the microbial community composition of gulf coast beaches following beach oiling. PLoS ONE 8:1–13

Nwinyi OC, Ajayi OO, Amund OO (2016) Degradation of polynuclear aromatic hydrocarbons by two strains of Pseudomonas. Braz J Microbiol 47(3):551–562. https://doi.org/10.1016/j.bjm.2016.04.026

Parks DH, Tyson GW, Hugenholtz P, Beiko RG, Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics 30(21):3123–3124. https://doi.org/10.1093/bioinformatics/btu494

Parthipan P, Preetham E, Machuca LL, Rahman PKSM, Murugan K, Rajasekar A (2017) Biosurfactant and degradative enzymes mediated crude oil degradation by bacterium Bacillus subtilis A1. Front Microbiol 8(FEB):1–14. https://doi.org/10.3389/fmicb.2017.00193

Paul S, Cortez Y, Vera N, Villena GK, Gutiérrez-Correa M (2016) Metagenomic analysis of microbial community of an Amazonian geothermal spring in Peru. Genom Data 9:63–66. https://doi.org/10.1016/j.gdata.2016.06.013

Selvarajan R, Maredza TA, Tekere M (2014). Microbial exploration in extreme conditions: metagenomic analysis and future perspectives. In: BENEDETTI C (ed) Metagenomics—methods, applications and perspectives. Nova Science Publishers, Inc., New York, pp 157–181

Selvarajan R, Sibanda T, Tekere M (2018a). Thermophilic bacterial communities inhabiting the microbial mats of "indifferent" and chalybeate (iron-rich) thermal springs: Diversity and biotechnological analysis. MicrobiologyOpen 7, 1–12.

Selvarajan R, Sibanda T, Venkatachalam S, Kamika I, Nel W (2018b) Industrial wastewaters harbor a unique diversity of bacterial communities revealed by high-throughput amplicon analysis. Ann Microbiol, 1–14. https://doi.org/10.1007/s13213-018-1349-8

Sutton NB, Maphosa F, Morillo JA, Al-Soud WA, Langenhoff AM, Grotenhuis T, Rijnaarts HM, Smidt H (2012) Impact of long-term diesel contamination on soil microbial community structure. Appl Environ Microbiol 79(2):619–630

Reid T, Chaganti S, Droppo IG, Weisener CG (2018) Novel insights into freshwater hydrocarbon-rich sediments using metatranscriptomics: opening the black box. Water Res, 1–27. https://doi.org/10.1016/j.watres.2018.02.039

Roesch LFW, Fulthorpe RR, Riva A, Casella G, Hadwin AKM, Kent AD, Daroub SH, Camargo FAO, Farmerie WG, Triplett EW (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. ISME J 1(4):283–290

Schulz S, Brankatschk R, Dümig A, Kögel-Knabner I, Schloter M, Zeyer J (2013) The role of microorganisms at different stages of ecosystem development for soil formation. Biogeosciences 10(6):3983–3996. https://doi.org/10.5194/bg-10-3983-2013

Sibanda T, Selvarajan R, Tekere M (2017) Synthetic extreme environments: overlooked sources of potential biotechnologically relevant microorganisms. Microb Biotechnol. https://doi.org/10.1002/mrd.22357

Stefani FOP, Bell TH, Marchand C, De La Providencia IE, El Yassimi A, St-Arnaud M, Hijri M (2015) Culture-dependent and -independent methods capture different microbial community fractions in hydrocarbon-contaminated soils. PLoS One 10(6):1–16. https://doi.org/10.1371/journal.pone.0128272

The Guardian. (2018) Fishing in oily waters: Bodo's long, painful wait for pollution clean-up. Retrieved June 26, 2018, From https://guardian.ng/features/fishing-in-oily-waters-bodos-long-painful-wait-for-pollution-clean-up/

USEPA (1978) Test method for evaluating total recoverable petroleum hydrocarbon, method 418.1 (spectrophotometric, infrared). Government Printing Office, Washington, D.C.

Valdés J, Pedroso I, Quatrini R, Dodson RJ, Tettelin H, Blake R, Eisen JA, Holmes DS (2008) Acidithiobacillus ferrooxidans metabolism: from genome sequence to industrial applications. BMC Genomics 9(1):597

Wang L, Wang W, Lai Q, Shao Z (2010) Gene diversity of CYP153A and AlkB alkane hydroxylases in oil-degrading bacteria isolated from the Atlantic Ocean. Environ Microbiol 12:1230–1242

Wang F, Li C, Wang H, Chen W, Huang Q, Wang F, Li C, Wang H, Chen W, Huang Q (2016) Characterization of a phenanthrene-degrading microbial consortium enriched from petrochemical contaminated environment. Int Biodeterior Biodegrad 98:4209–4221. https://doi.org/10.1016/j.ibiod.2016.08.028

Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173(2):697–703 Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1987160

Yakimov MM, Timmis KN, Golyshin PN (2007) Obligate oil-degrading marine bacteria. Curr Opin Biotechnol 18(3):257–266

Ye Y, Doak TG (2009) A parsimony approach to biological pathway reconstruction/inference for genomes and metagenomes. PLoS Comput Biol 5(8):e1000465. https://doi.org/10.1371/journal.pcbi.1000465